entropy

Complexity in Economic and Social Systems

Edited by
Stanisław Drożdż, Jarosław Kwapień and Paweł Oświęcimka

Printed Edition of the Special Issue Published in *Entropy*

www.mdpi.com/journal/entropy

MDPI

Complexity in Economic and Social Systems

Complexity in Economic and Social Systems

Editors

Stanisław Drożdż
Jarosław Kwapień
Paweł Oświęcimka

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editors

Stanisław Drożdż
Polish Academy of Sciences
Poland

Jarosław Kwapień
Polish Academy of Sciences
Poland

Paweł Oświęcimka
Polish Academy of Sciences
Poland

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Entropy* (ISSN 1099-4300) (available at: https://www.mdpi.com/journal/entropy/special_issues/complexity_economic_social).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-0365-0794-1 (Hbk)
ISBN 978-3-0365-0795-8 (PDF)

Contents

About the Editors . . . . . **ix**

Stanisław Drożdż, Jarosław Kwapień, and Paweł Oświęcimka
Complexity in Economic and Social Systems
Reprinted from: *Entropy* **2021**, *23*, 133, doi:10.3390/e23020133 . . . . . **1**

Jarosław Klamut, Ryszard Kutner, and Zbigniew R. Struzik
Towards a Universal Measure of Complexity
Reprinted from: *Entropy* **2020**, *22*, 866, doi:10.3390/e22080866 . . . . . **5**

Aleksander Jakimowicz
The Role of Entropy in the Development of Economics
Reprinted from: *Entropy* **2020**, *22*, 452, doi:10.3390/e22040452 . . . . . **35**

Roy Cerqueti, Giulia Rotundo and Marcel Ausloos
Tsallis Entropy for Cross-Shareholding Network Configurations
Reprinted from: *Entropy* **2020**, *22*, 676, doi:10.3390/e22060676 . . . . . **61**

Pietro Murialdo, Linda Ponta and Anna Carbone
Long-Range Dependence in Financial Markets:
A Moving Average Cluster Entropy Approach
Reprinted from: *Entropy* **2020**, *22*, 634, doi:10.3390/e22060634 . . . . . **77**

Andrés García-Medinaand José B. Hernández C.
Network Analysis of Multivariate Transfer Entropy of Cryptocurrencies in Times of Turbulence
Reprinted from: *Entropy* **2020**, *22*, 760, doi:10.3390/e22070760 . . . . . **97**

Stanisław Drożdż, Jarosław Kwapień, Paweł Oświęcimka, Tomasz Stanisz and Marcin Wątorek
Complexity in Economic and Social Systems: Cryptocurrency Market at around COVID-19
Reprinted from: *Entropy* **2020**, *22*, 1043, doi:10.3390/e22091043 . . . . . **115**

Peng Yue, Yaodong Fan, Jonathan A. Batten and Wei-Xing Zhou
Information Transfer between Stock Market Sectors: A Comparison between the USA and China
Reprinted from: *Entropy* **2020**, *22*, 194, doi:10.3390/e22020194 . . . . . **141**

Gang Chu, Xiao Li, Dehua Shen and Yongjie Zhang
Unexpected Information Demand and Volatility Clustering of Chinese Stock Returns: Evidence from Baidu Index
Reprinted from: *Entropy* **2020**, *22*, 44, doi:10.3390/e22010044 . . . . . **155**

Zhiqiang Hu, Yuan Hu, Yushan Jiang and Zhen Peng
Pricing Constraint and the Complexity of IPO Timing in the Stock Market: A Dynamic Game Analysis
Reprinted from: *Entropy* **2020**, *22*, 546, doi:10.3390/e22050546 . . . . . **165**

Lucia Inglada-Perez
A Comprehensive Framework for Uncovering Non-Linearity and Chaos in Financial Markets: Empirical Evidence for Four Major Stock Market Indices
Reprinted from: *Entropy* **2020**, *22*, 1435, doi:10.3390/e22121435 . . . . . **183**

Barbara Będowska-Sójka and Krzysztof Echaust
Do Liquidity Proxies Based on Daily Prices and Quotes Really Measure Liquidity?
Reprinted from: *Entropy* **2020**, *22*, 783, doi:10.3390/e22070783 . . . . . **209**

Yong Shi, Yuanchun Zheng, Kun Guo, Zhenni Jin and Zili Huang
The Evolution Characteristics of Systemic Risk in China's Stock Market Based on a Dynamic Complex Network
Reprinted from: *Entropy* **2020**, *22*, 614, doi:10.3390/e22060614 . . . . . **231**

Zhen Peng and Changsheng Hu
The Threshold Effect of Leveraged Trading on the Stock Price Crash Risk: Evidence from China
Reprinted from: *Entropy* **2020**, *22*, 268, doi:10.3390/e22030268 . . . . . **247**

Katarzyna Bień–Barkowska
Looking at Extremes without Going to Extremes:
A New Self-Exciting Probability Model
for Extreme Losses in Financial Markets
Reprinted from: *Entropy* **2020**, *22*, 789, doi:10.3390/e22070789 . . . . . **261**

Xin Yang, Xian Zhao, Xu Gong, Xiaoguang Yang and Chuangxia Huang
Systemic Importance of China's Financial Institutions: A Jump Volatility Spillover Network Review
Reprinted from: *Entropy* **2020**, *22*, 588, doi:10.3390/e22050588 . . . . . **287**

Michał Cieśla and Małgorzata Snarska
A Simple Mechanism Causing Wealth Concentration
Reprinted from: *Entropy* **2020**, *22*, 1148, doi:10.3390/e22101148 . . . . . **303**

Jagoda Kaszowska-Mojsa and Mateusz Pipień
Macroprudential Policy in a Heterogeneous Environment—An Application of Agent-Based Approach in Systemic Risk Modelling
Reprinted from: *Entropy* **2020**, *22*, 129, doi:10.3390/e22020129 . . . . . **319**

Aleksander Jakimowicz and Daniel Rzeczkowski
Innovativeness of Industrial Processing Enterprises and Conjunctural Movement
Reprinted from: *Entropy* **2020**, *22*, 1177, doi:10.3390/e22101177 . . . . . **393**

Eduardo Viegas, Hayato Goto, Yuh Kobayashi, Misako Takayasu, Hideki Takayasu and Henrik Jeldtoft Jensen
Allometric Scaling of Mutual Information in Complex Networks: A Conceptual Framework and Empirical Approach
Reprinted from: *Entropy* **2020**, *22*, 206, doi:10.3390/e22020206 . . . . . **427**

Aleksander Jakimowicz and Daniel Rzeczkowski
New Measure of Economic Development Based on the Four-Colour Theorem
Reprinted from: *Entropy* **2021**, *23*, 61, doi:10.3390/e23010061 . . . . . **441**

Bedane S. Gemeda, Birhanu G. Abebe, Andrzej Paczoski, Yi Xie and Giuseppe T. Cirella
What Motivates Speculators to Speculate?
Reprinted from: *Entropy* **2020**, *22*, 59, doi:10.3390/e22010059 . . . . . **473**

José Martins and Alberto Pinto
The Value of Information Searching against Fake News
Reprinted from: *Entropy* **2020**, *22*, 1368, doi:10.3390/e22121368 . . . . . **491**

Dorian Wild, Margareta Jurcic and Boris Podobnik
The Gender Productivity Gap in Croatian Science: Women Are Catching up with Malesand Becoming Even Better
Reprinted from: *Entropy* **2020**, *22*, 1217, doi:10.3390/e22111217 . . . . . **509**

About the Editors

Stanisław Drożdż, Professor, Present professor of physics and head of the Complex Systems Theory Department at Institute of Nuclear Physics (INP), Polish Academy of Sciences and professor of computer science at the Cracow University of Technology. He received an M.S. (1978) and Ph.D. (1982) in physics, both from Jagiellonian University in Kraków, and a DSc (1988) in theoretical physics from INP. In 1994, he received the Polish state title of professor. His long-term scientific studies abroad include Forschungszentrum Juelich, Germany from 1983 to 1986 as a post-doc and from 1989 to 1991 as a senior scientist, as well as the University of Illinois at Urbana-Champaign, USA from 1993 to 1994 and Bonn University, Germany from 2001 to 2002 as a visiting professor. His research interests and activities include the quantum many-body problem, nonlinear dynamics, general aspects of complexity, brain-research-related issues, dynamics of financial markets, and quantitative linguistics.

Jarosław Kwapień, Assoc. Prof., Graduated from Jagiellonian University in Kraków in 1995 with a Master's Degree in physics. He started his doctoral programme at the Institute of Nuclear Physics (INP), Polish Academy of Sciences and earned a Ph.D. degree in 2001 based on his work on topics at the interface between statistical physics and neuroscience. As a post-doctoral fellow at Forschungszentrum Juelich, Germany, he started his research on topics in the physics of financial markets (econophysics) and data science. Later, he entered a tenure track at INP and completed his habilitation in 2010, after which he obtained the position of Associate Professor. Recently, he has continued to publish works in econophysics, but also began an interest in complex networks, quantitative linguistics, and other interdisciplinary topics.

Paweł Oświęcimka, Assoc. Prof., Paweł Oświecimka received an M.S. degree in physics from the AGH University of Science and Technology, Kraków, Poland, in 2001, and Ph.D. and D.Sc. degrees in physics from Institute of Nuclear Physics (INP), Polish Academy of Sciences, Kraków, in 2005 and 2015, respectively. Since 2016, he has been Associate Professor at the Complex Systems Theory Department INP. He also leads one of the research groups in the project on bio-inspired artificial neural networks at the Jagiellonian University and he is a lecturer at the Cracow University of Technology. His current research interests include interdisciplinary research in natural and social complex systems by means of advanced methods of time series analysis, with particular emphasis on multifractal analysis, complex network analysis, and modelling of multi-scaling processes. He is a Fellow of the Commission of Complex Systems at the Polish Academy of Arts and Sciences (PAU).

entropy

MDPI

Editorial

Complexity in Economic and Social Systems

Stanisław Drożdż [1,2,*], Jarosław Kwapień [1] and Paweł Oświęcimka [1,3]

[1] Complex Systems Theory Department, Institute of Nuclear Physics, Polish Academy of Sciences, ul. Radzikowskiego 152, 31-342 Kraków, Poland; jaroslaw.kwapien@ifj.edu.pl (J.K.); pawel.oswiecimka@ifj.edu.pl (P.O.)

[2] Faculty of Computer Science and Telecommunication, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland

[3] Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, ul. Prof. Stanisława Łojasiewicza 11, 30-348 Kraków, Poland

* Correspondence: stanislaw.drozdz@ifj.edu.pl

Received: 5 January 2021; Accepted: 18 January 2021; Published: 21 January 2021

During recent years we have witnessed a systematic progress in the understanding of complex systems, both in the case of particular systems that are classified into this group and, in general, as regards the phenomenon of complexity [1]. This is possible owing to an outburst of research interest in the science of complexity and a joint effort made by the researchers representing different disciplines and backgrounds which resulted in the enormous number of interdisciplinary studies carried out. This progress has been achieved on both the theoretical, model, and experimental levels. However, in order to comprehend the complexity in full detail, much is still to be done. This is particularly true in the case of the systems involving human society and behaviour.

This Special Issue of *Entropy* was intended to attract researchers specializing in interdisciplinary studies of complex systems, with the economic and social systems in particular, and to collect in one place their contributions that otherwise could be scattered among many journals and issues. We believe that the papers spanning this issue can be considered as valuable input to their specific fields, but also to complexity science in general. Some of them because they relate to general concepts and thus their conclusions can be exploited in various situations across many fields and others because of the methods that were used there, the knowledge of which can be disseminated more broadly. We are glad that our idea was met with a positive response and now we can present as many as 23 genuine research papers on a wide spectrum of topics. The largest set of papers is related to the economical systems, while smaller sets to the social systems and to general complexity, with such a proportion reflecting the total amount of the current scientific output in these fields.

Complexity still lacks a commonly accepted strict definition. In many practical cases it suffices to understand this notion intuitively as a nontrivial, irreducible order (i.e., other than simple regularity or a straightforward effect of a lower level of organization) that spontaneously emerges from an overall chaos [1], but there is also a strong need to provide a strict definition that, for instance, can be applied to categorize various systems based on their structure and dynamics or to construct a measure of complexity. There were a plenty of attempts in that direction but they largely failed. An interesting step towards resolving this issue is presented in a paper [2] where its authors propose a measure of complexity based on a nonlinear transformation of time-dependent entropy that attributes the highest complexity to the optimally mixed states between maximum regularity and maximum disorder.

Various tools based on entropy are frequently used in the context of complex systems and it is not surprising that they are applied in a few other contributed papers. One of the principal directions of research is looking for precursors of the oncoming structural phase transitions. For example, transfer entropy quantifying dependence asymmetry between two systems is used to construct a network of information transfers among cryptocurrencies. The resulting network topology reveals significant alteration during turmoils and forecasts a systemic risk increase [3]. The Tsallis nonextensive entropy has already proved useful in studying complex systems [4]. It is applied to analyze the

cross-shareholding networks of companies. In this context it offers a measure of market polarisation and a tool for analyzing market self-organization in response to external shocks [5]. Finally, the moving average cluster entropy is proposed to study the long-range dependence in time series and proves useful as a measure capturing endogenous sources of risk over different temporal horizons [6].

Risk, which quantifies market or asset stability and vulnerability to external shocks, has always been one of the key topics in economics, but it is also an important issue from the complexity perspective [7]. A few more papers from this Special Issue consider systemic risk as one of their central points. If such a risk is quantified in terms of some framework, it is possible to observe its evolution on a given market. For example, the Chinese stock market network topology analysis leads to a conclusion that the systemic risk can be decomposed into a clear trend and periodic fluctuations with the former reflecting the gradual improvement of the management and operation of the market and the latter reflecting the events of excessive strength [8]. Among the most important sources of risk is leverage trading but this relation can rather be non-trivial with either stabilizing or destabilizing impact depending on the leverage trading share in total market activity [9]. In order to manage risk, one needs to construct realistic models that are able to predict the probability of financial losses [10] and to use reliable measures able to provide one with sufficiently early alerts [3]. On the other hand, risk can also be managed by identifying the key companies or sectors that are its sources of centers [11].

Financial markets are among the most interesting complex systems from a perspective of the empirical data analysis, because they provide incomparably clean data. This is why much effort dedicated to studying these markets can be fruitful far beyond the field of economics [1]. Self-organization of the stock markets and their hierarchical structure can be approached from the angle of information transfer between different sectors in various time intervals [12]. This also refers to the cross-shareholding market structure which self-organizes under the influence of external shocks [5]. Both the external shocks and the internal market events can produce excessive demand for information, which, if properly quantified, may offer a way to monitor oncoming market events that are difficult to predict by using other methods. A new tool is proposed based on an internet search engine like the Chinese Baidu [13].

Another signature of market complexity is pricing and timing of the stocks during their initial public offerings. In [14], a few results on the IPO timing properties are considered and discussed. Studying the temporal properties of financial dynamics, which is a property related to their complexity, offers verification for one of the key paradigms of financial markets, namely the efficient market hypothesis [6], and testing for nonlinearity and chaos [15]. Temporal properties of financial dynamics can in turn be one of the consequences of stock liquidity and it is thus important to have a reliable method to quantify it [16].

Same as the most financial markets offer high quality data, the cryptocurrency market offers also a unique possibility of observation of the whole process of new complex system development from scratch [17]. The cryptocurrency market is also interesting as the possible future fate of money. So their evolution and reacting to external shocks like the COVID-19 pandemic is particularly interesting and instructive [3], especially as it seems that this market gradually reaches maturity [18].

Modelling of the economical systems can go beyond the financial markets and also be applied to more general problems like the wealth condensation in society (simple but effective an agent-based model in [19]), the innovation-related performance on a market (the innovation pressure effects of private-owned enterprises and public companies [20]), and the impact of the macroprudential policy on economy and the financial system (with the results on the stabilizing effects of such policy during the turmoils and crises [21]). From this general system level one may look downwards into the system component parts, which reveal complexity on their own: the geographical regions or administrative divisions. It is possible to quantify their economical development by a newly proposed method of the public administration website quality assessment [22] and to analyze differences in an inter-regional business ecosystem structure or economic activity efficiency level by means of a network approach [23].

Complex phenomena occurring on the interface between economical activity and spatial structure are the subject of a study of land speculation on the outskirts of a sample city in Ethiopia [24]. This study investigates motivations the speculators are driven by and concludes on a possible direction local governments should proceed in order to diminish negative impact of such practices on city development. An even more important social phenomenon with a negative impact on the society is fake news. A model of rumour spreading with evolutionary information search dynamics allows one for analyzing optimal search strategies that maximize pay-off for the society and potentially provides the policy makers with the recommendations how to minimize the harmful impact of fake news [25]. The most sociologically-oriented study of this Special Issue considers the research output of the male and female scientists quantified in terms of their publication citations from the perspective of the gender productivity gap [26]. It occurs that a larger gender inequality can be found in the STEM disciplines (i.e., science, technology, engineering, and mathematics) as compared with the non-STEM ones.

Finally, it is worth to mention a more history-oriented essay on the impact of physics (with thermodynamics in particular) on the development of ideas in the contemporary economics [27]. It is an interesting example of the innovation-generating potential of the interdisciplinary cooperation in science, which is the exceptionally welcome in the science of complex systems.

After this introduction, readers are warmly invited to read the papers collected in this Special Issue.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kwapień, J.; Drożdż, S. Physical approach to complex systems. *Phys. Rep.* **2012**, *515*, 115. [CrossRef]
2. Klamut, J.; Kutner, R.; Struzik, Z.R. Towards a Universal Measure of Complexity. *Entropy* **2020**, *22*, 866. [CrossRef] [PubMed]
3. García-Medina, A.; Hernandéz, C.J.B. Network Analysis of Multivariate Transfer Entropy of Cryptocurrencies in Times of Turbulence. *Entropy* **2020**, *22*, 760. [CrossRef] [PubMed]
4. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*; Springer: New York, NY, USA, 2009.
5. Cerqueti, R.; Rotundo, G.; Ausloos, M. Tsallis Entropy for Cross-Shareholding Network Configurations. *Entropy* **2020**, *22*, 676. [CrossRef] [PubMed]
6. Murialdo, P.; Ponta, L.; Carbone, A. Long-Range Dependence in Financial Markets: A Moving Average Cluster Entropy Approach. *Entropy* **2020**, *22*, 634. [CrossRef] [PubMed]
7. Sornette, D. *Why Stock Market Crash: Critical Events in Complex Financial Systems*; Princeton University Press: Princeton, NJ, USA, 2003.
8. Shi, Y.; Zheng, Y.; Guo, K.; Jin, Z.; Huang, Z. The Evolution Characteristics of Systemic Risk in China's Stock Market Based on a Dynamic Complex Network. *Entropy* **2020**, *22*, 614. [CrossRef]
9. Peng, Z.; Hu, C. The Threshold Effect of Leveraged Trading on the Stock Price Crash Risk: Evidence from China. *Entropy* **2020**, *22*, 268. [CrossRef]
10. Bień-Barkowska, K. Looking at Extremes without Going to Extremes: A New Self-Exciting Probability Model for Extreme Losses in Financial Markets. *Entropy* **2020**, *22*, 789. [CrossRef]
11. Yang, X.; Zhao, X.; Gong, X.; Yang, X.; Huang, C. Systemic Importance of China's Financial Institutions: A Jump Volatility Spillover Network Review. *Entropy* **2020**, *22*, 588. [CrossRef]
12. Yue, P.; Fan, Y.; Batten, J.A.; Zhou, W.-X. Information Transfer between Stock Market Sectors: A Comparison between the USA and China. *Entropy* **2020**, *22*, 194. [CrossRef]
13. Chu, G.; Li, X.; Shen, D.; Zhan, Y. Unexpected Information Demand and Volatility Clustering of Chinese Stock Returns: Evidence from Baidu Index. *Entropy* **2020**, *22*, 44. [CrossRef] [PubMed]
14. Hu, Z.; Hu, Y.; Jiang, Y.; Peng, Z. Pricing Constraint and the Complexity of IPO Timing in the Stock Market: A Dynamic Game Analysis. *Entropy* **2020**, *22*, 546. [CrossRef] [PubMed]

15. Inglada-Perez, L. A Comprehensive Framework for Uncovering Non-Linearity and Chaos in Financial Markets: Empirical Evidence for Four Major Stock Market Indices. *Entropy* **2020**, *22*, 1435. [CrossRef] [PubMed]
16. Będowska-Sójka, B.; Echaust, K. Do Liquidity Proxies Based on Daily Prices and Quotes Really Measure Liquidity? *Entropy* **2020**, *22*, 783. [CrossRef] [PubMed]
17. Wątorek, M.; Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* **2020**. [CrossRef]
18. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Stanisz, T.; Wątorek, M. Complexity in Economic and Social Systems: Cryptocurrency Market at around COVID-19. *Entropy* **2020**, *22*, 1043. [CrossRef]
19. Cieśla, M.; Snarska, M. A Simple Mechanism Causing Wealth Concentration. *Entropy* **2020**, *22*, 1148. [CrossRef]
20. Jakimowicz, A.; Rzeczkowski, D. Innovativeness of Industrial Processing Enterprises and Conjunctural Movement. *Entropy* **2020**, *22*, 1177. [CrossRef]
21. Kaszowska-Mojsa, J.; Pipień, M. Macroprudential Policy in a Heterogeneous Environment—An Application of Agent-Based Approach in Systemic Risk Modelling. *Entropy* **2020**, *22*, 129. [CrossRef]
22. Jakimowicz, A.; Rzeczkowski, D. New Measure of Economic Development Based on the Four-Colour Theorem. *Entropy* **2020**, *22*, 61. [CrossRef]
23. Viegas, E.; Goto, H.; Kobayashi, Y.; Takayasu, M.; Takayasu, H.; Jensen, H.J. Allometric Scaling of Mutual Information in Complex Networks: A Conceptual Framework and Empirical Approach. *Entropy* **2020**, *22*, 206. [CrossRef] [PubMed]
24. Gemeda, B.S.; Abebe, B.G.; Paczoski, A.; Xie, Y.; Cirella, G.T. What Motivates Speculators to Speculate? *Entropy* **2020**, 22, 59. [CrossRef] [PubMed]
25. Martins, J.; Pinto, A. The Value of Information Searching against Fake News. *Entropy* **2020**, *22*, 1368. [CrossRef]
26. Wild, D.; Jurcic, M.; Podobnik, B. The Gender Productivity Gap in Croatian Science: Women Are Catching up with Males and Becoming Even Better. *Entropy* **2020**, *22*, 1217. [CrossRef] [PubMed]
27. Jakimowicz, A. The Role of Entropy in the Development of Economics. *Entropy* **2020**, *22*, 452. [CrossRef]

entropy

MDPI

Article

Towards a Universal Measure of Complexity

Jarosław Klamut [1], Ryszard Kutner [1,*] and Zbigniew R. Struzik [1,2,3]

[1] Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland; jaroslaw.klamut@fuw.edu.pl (J.K.); z.r.struzik@p.u-tokyo.ac.jp (Z.R.S.)
[2] Graduate School of Education, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[3] Advanced Center for Computing and Communication, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
* Correspondence: ryszard.kutner@fuw.edu.pl

Received: 27 May 2020; Accepted: 3 August 2020; Published: 6 August 2020

Abstract: Recently, it has been argued that entropy can be a direct measure of complexity, where the smaller value of entropy indicates lower system complexity, while its larger value indicates higher system complexity. We dispute this view and propose a universal measure of complexity that is based on Gell-Mann's view of complexity. Our universal measure of complexity is based on a non-linear transformation of time-dependent entropy, where the system state with the highest complexity is the most distant from all the states of the system of lesser or no complexity. We have shown that the most complex is the optimally mixed state consisting of pure states, i.e., of the most regular and most disordered which the space of states of a given system allows. A parsimonious paradigmatic example of the simplest system with a small and a large number of degrees of freedom is shown to support this methodology. Several important features of this universal measure are pointed out, especially its flexibility (i.e., its openness to extensions), suitability to the analysis of system critical behaviour, and suitability to study the dynamic complexity.

Keywords: dynamical complexity; universal complexity measure; irreversible processes; entropies; entropic susceptibilities

1. Introduction

Analysis of the concept of complexity is a non-trivial task due to its diversity, arbitrariness, uncertainty, and contextual nature [1–10]. There are many different levels/scales, faces, and types of complexity, researched with very different technologies/techniques and tools [11–13] (and refs. therein). In the context of dynamical systems, Grassberger suggested [14] that a slow convergence of the entropy to its extensive asymptotic limit is a signature of complexity. This idea was materialized [15,16] further by information and statistical mechanics techniques. It generalizes many previous approaches to complexity, unifying physical ideas with ideas from learning and coding theory [17]. There also exists a connection of this approach to algorithmic or Kolmogorov complexity. The hidden pattern can be the essence of complexity [18–21]. Techniques adapted from the theories of information and computation have led physical science (in particular, the region extended between classical determinism and deterministic chaos) to discover hidden patterns and quantify their dynamic structural complexity [22]. The above approaches are not universal—they only capture small fragments of the concept of complexity.

We must remember that complexity also depends on the conditions imposed (e.g., boundary or initial conditions), as well as the restrictions adopted. This creates a challenge for every complexity study. It concerns the complexity that can appear in the movement of a single entity and collection of entities braided together. These entities can be irreducible or straightforward, simple systems, but they can also be complex systems.

When we talk about complexity, we mean irreducible complexity, which can no longer be divided into smaller sub-complexities. We refer to this as a primary complexity. Considering the primary complexity here, we mean one that can be expressed at least in an algorithmic way—it is an effective complexity if it also contains a logical depth [23–27]. We should take into account that our models (analytical and numerical) and theories describing reality are not fully deterministic. The evolution of a complex system is potentially multi-branched and the selection of an alternative trajectory (or branch selection) is based on decisions taken randomly.

One of the essential questions concerning a complex system is the problem of its stability/robustness and the question of the stationarity of its evolution [28]. Moreover, the relationship between complexity and disorder on the one hand, and complexity and pattern on the other is an important question—especially in the context of irreversible processes, where non-linear processes, running away from the equilibrium, play a central role. Financial markets can be a spectacular example of these processes [29–39].

The central question of whether entropy is a direct measure of complexity is one we answer in the negative. In our opinion, based on the Gell–Mann concept of complexity, the measure of complexity is appropriately, non-linearly transformed entropy. This work is devoted to finding this transformation and examining the resulting consequences.

2. Definition of a Universal Measure of Complexity and Its Properties

In this Section, we translate the Gell–Mann general qualitative concept of complexity into the language of mathematics, and we present the consequences of this.

2.1. The Gell–Mann Concept of Complexity

The problem of defining a universal measure of complexity is urgent. For this work, the Gell–Mann concept [23,40] of complexity is the inspirational starting point. We apply this concept to irreversible processes, by assuming that both fully ordered and fully disordered systems cannot be the complex. The fully ordered system essentially has no complexity because of maximal possible symmetry of the system, but the fully disordered system contains no information as it entirely dissipates. Hence, the maximum of complexity should be sought somewhere in between these pure extreme states. This point of view allows for the introduction of a formal quantitative phenomenological complexity measure based on entropy as a parameter of order [29,41]. This measure reflects the dynamics of the system through the dependence of entropy on time. The vast majority of works analyzing the general aspects of complexity, including its basis, are based on information theory and computational analysis. Such an approach requires supplementing with a provision allowing a return from a bit representation to physical representation—only this will allow physical interpretations, including understanding of the causes of complexity.

We define the phenomenological partial measure of complexity as a non-linear function of entropy S of the order of (m, n),

$$\begin{aligned} CX(S;m,n) &\stackrel{\text{def.}}{=} (S^{max}-S)^m (S-S^{min})^n \\ &= CX(S;m-1,n-1)\left[\left(\frac{Z}{2}\right)^2-(S-S^{arit})^2\right],\ m,n\geq 1, \end{aligned} \tag{1}$$

where S^{min} and S^{max} are minimal and maximal values of entropy S, respectively, $S^{arit}=\frac{S^{min}+S^{max}}{2}$, and the entropic span $Z \stackrel{\text{def.}}{=} S^{max}-S^{min}$, whereas m and n are natural numbers (an extension to real positive numbers is possible but this is not the subject of this work). They define the order (m, n) of the partial measure of complexity CX. Let us add that this formula is also applicable at a mesoscopic scale. In other words, complexity appears in all systems for which we can build entropy. Notably, S^{max} does not have to concern the state of thermodynamic equilibrium of the system. It may refer to the

state for which entropy reaches its maximum value in the observed time interval. However, in this work, we are only limited to systems having a state of thermodynamic equilibrium. Below, we discuss the Equation (1), indicating that it satisfies all properties of the measure of complexity. Of course, when $m = 0$ and $n = 1$ then CX simply becomes $S - S^{min}$, i.e., the entropy of the system (the constant is not important here). However, when $m = 1,\ n = 0$, we obtain the information contained in the system (constant does not play a role here). Equation (1) gives us a lot more—showing this is the purpose of this work (helpful features of CX are shown in Appendix A).

The partial measure of complexity given by Equation (1) is determined with the accuracy of the additive constant of S, i.e., this constant does not contribute to the measure of the complexity of the system.

Using Equation (1), we can also enter the partial measure of specific complexity, as follows,

$$cx(s; m, n) \stackrel{\text{def.}}{=} \frac{1}{N^{m+n}} CX(Ns; m, n), \tag{2}$$

where N is the number of entities that make up the system and specific entropy $s = S/N$. As one can see, the partial measure of specific complexity cx is independent of N for an extensive system. Specific entropy and specific complexity are particularly convenient when comparing different extensive systems and when we do not examine the complexity dependence on N.

However, the extraction of an additional multiplicative constant (e.g., particle number) to have s independent of N often presents a technical difficulty, or may even be impossible, especially for non-extensive systems. Then it is more convenient to use the entropy of the system instead of the specific entropy. It is also important to realize that determining extreme entropy values (or extreme specific entropy values) of actual systems can be complicated and it requires additional dedicated tools/technologies, algorithms, and models.

The partial measures of complexity are enslaved by entropy in every order (m, n) of complexity. However, the kind of entropy we use in Equation (1) depends on the specific situation of the system and what we want to know about the system, because our definition of complexity does not specify this. From our point of view, relative entropies formulated in the spirit of Kullback–Leibler seem to be the most appropriate (this is referred to in Appendix B). Using the Kullback–Leibler type of entropy, one can express both ordinary entropies and conditional entropies, in particular one can describe the entropy rate increasingly used in the context of complexity analysis.

The entropy here can be both additive (the Boltzmann–Gibbs thermodynamic one [42], Shanon information [17], Rényi [43]), and non-additive entropy (Tsallis [44]). The measure $CX(S)$ is a concave (or convex up) function of entropy S, which disappears on the edges at points $S = S^{min}$ and $S = S^{max}$.

It has a maximum

$$CX^{max} = CX(S = S^{max}_{CX}; m, n) = m^m\, n^n \left(\frac{Z}{m+n}\right)^{m+n} \tag{3}$$

at point

$$S = S^{max}_{CX} = \overline{S} = \frac{mS^{min} + nS^{max}}{m+n} = \frac{\frac{1}{m}S^{max} + \frac{1}{n}S^{min}}{\frac{1}{m} + \frac{1}{n}} \tag{4}$$

as at this point $\frac{dCX(S)}{dS}\,|_{S=\overline{S}} = 0$ and $\frac{d^2CX(S)}{dS^2}\,|_{S=\overline{S}} < 0$. The quantity S^{max}_{CX} is a characteristic also because it is a weighted average. The quantity CX^{max} is well suited to global universal measurements of complexity, because (at a given order (m, n)), it only depends on the entropy span Z. The quantity $cx^{max} \stackrel{\text{def.}}{=} CX^{max}/N^{m+n}$ might also be a good candidate for measuring the logic depth of complexity.

2.2. The Most Complex Structure

The question now arises about the structure of the system corresponding to entropy S_{CX}^{max} given by Equation (4). The answer is given by the following constitutive equation,

$$S\left(Y = Y^{CX^{max}}\right) = S_{CX}^{max}, \tag{5}$$

where Y is the set of variables and parameters (e.g., thermodynamic), on which the state of the system depends. However, $Y = Y^{CX^{max}}$ is a set of such values of these variables and parameters that are the solution of Equation (5). This solution gives the entropy value $S = S_{CX}^{max}$ that maximizes the partial measure of complexity, that is $CX = CX^{max}$. Hence, with the value of $Y^{CX^{max}}$, we can finally answer the key question: what structure/pattern is behind CX^{max} or what the structure of maximum complexity looks like.

There are a few comments to be made regarding the constitutive Equation (5) itself. It is a (non-linear) transcendental equation in the untangled form relative to the Y. This equation should be numerically solved, because we do not expect it to have an analytical solution for maximally complex systems. An instructive example of a specific form of this equation and its solution for a specific physical problem is presented in Section 3. However, this will help us to understand how our machinery works.

Equation (4) legitimizes the measure of complexity we have introduced. Namely, its maximum value falls on the weighted average entropy value, which describes the optimal mixture of completely ordered and completely disordered phases. To the left of $\overline{S}$, we have a phase with dominance of order and to the right a phase with dominance of disorder. The transition between both phases at $\overline{S}$ is continuous. Thus, we can say that the partial measure of complexity that we have introduced also defines a certain type of phase diagram in S and CX variables (phase diagram plain). Section 2.5 provides more detailed information.

2.3. Evolution of the Partial Measure of Complexity

Differentiating Equation (1) over time t, we obtain the following non-linear dynamics equation,

$$\frac{dCX(S(t);m,n)}{dt} = \chi_{CX}(S;m,n)\,\frac{dS(t)}{dt} = (m+n)\,[S_{CX}^{max} - S(t)]\,CX(S(t);m-1,n-1)\frac{dS(t)}{dt}, \tag{6}$$

where the entropic S-dependent (non-linear), susceptibility is defined by

$$\chi_{CX}(S;m,n) \stackrel{\text{def.}}{=} \frac{\partial CX(S;m,n)}{\partial S} = (m+n)\,[S_{CX}^{max} - S(t)]\,CX(S(t);m-1,n-1) \tag{7}$$

and $\frac{dS(t)}{dt}$ can be expressed, for example, using the right-hand side of the master Markov equation (see Ref. [45] for details). However, we must realize that the dependence of entropy on time can, in general, be non-monotonic, because real systems are not isolated (cf. the schematic plot in Figure 2). One can see how the dynamics of complexity is controlled in a non-linear way by the evolution of the entropy of the system.

In concluding this Section, we state that Equations (1)–(6) together provide a technology for studying the multi-scale aspects of complexity, including the dynamic complexity. However, it is still a simplified approach, as we show in Section 4.

2.4. Significant Partial Measure of Complexity

We consider the partial measure of complexity to be significant when the entropy of the system is located between two inflection points of the $CX(S;m,n)$ curve, i.e., in the range $S_{ip}^{-} \leq S \leq S_{ip}^{+}$. This case occurs for $n, m \geq 2$. We then obtain

$$S^{min} < S_{ip}^{\mp} = S^{min} + \frac{n(n-1)}{\sqrt{n(n+m-1)}} \frac{S^{max} - S^{min}}{\sqrt{n(n+m-1)} \pm \sqrt{m}} < S^{max}, \tag{8}$$

see Figure 1d for details.

There are two different cases where a single inflection point is present. Namely,

$$S^{min} < S_{ip}^{-} = \frac{2S^{max} + m(m-1)S^{min}}{2 + m(m-1)} < \overline{S}, \quad \text{for } m \geq 2, n = 1, \tag{9}$$

and

$$\overline{S} < S_{ip}^{+} = \frac{2S^{min} + n(n-1)S^{max}}{2 + n(n-1)} < S^{max}, \quad \text{for } m = 1,\ n \geq 2. \tag{10}$$

In Figure 1b, we present the case defined by Equation (9), while that defined by Equation (10) is shown in Figure 1c.

For $n = m = 1$, the curve $CX(S; m, n)$ vs. S has no inflection points and it looks like a horseshoe (cf. Figure 1a).

Notably, we can equivalently write

$$S^{min} < S_{ip}^{\mp} = S^{max} - \frac{m(m-1)}{\sqrt{m(n+m-1)}} \frac{S^{max} - S^{min}}{\sqrt{m(n+m-1)} \mp \sqrt{n}} < S^{max}, \quad \text{for } n, m \geq 2. \tag{11}$$

Let us consider the span $Z_{ip} = S_{ip}^{+} - S_{ip}^{-}$ of the two-phase area. From Equation (8), or equivalently from Equation (11), we obtain

$$Z_{ip} = \frac{2\sqrt{nm}}{(n+m)\sqrt{n+m-1}} Z. \tag{12}$$

As one can see, the span Z_{ip} depends linearly on the span Z and in a non-trivial way on the exponents n and m. Thus, with the Z set, only Z_{ip}'s non-trivial dependence on the order (m, n) of measure of complexity CX occurs, which is different from CX^{max} dependence. In other words, Z_{ip} is less sensitive to complexity than CX^{max}.

The significant partial measure of complexity ranges between the two inflection points only for the case $n, m \geq 2$ (cf. Figure 1d). Indeed, a mixture of phases is observed in this area. For areas where $S^{min} \leq S < S_{ip}^{-}$ and $S_{ip}^{+} < S \leq S^{max}$, we have (practically speaking) only single phases, ordered and disordered, respectively (see Section 2.5 for details). The case defined by Equation (8), and equivalently by Equation (11), is the most general, while taking into account the fullness of complexity behaviour as a function of entropy. Other cases impoverish the description of complexity. Therefore, we will continue to consider the situation, where $n, m \geq 2$.

The choice of any of the $CX(S; m, n)$ forms (i.e., exponents n and m) is a somewhat arbitrary function of the state of the system as it depends on the function of the state, that is on the entropy. In our opinion, the shape of the $CX(S; m, n)$ measure vs. S we present in Figure 1d is the most appropriate, because only then the significant complexity is ranging between non-vanishing inflection points S_{ip}^{-} and S_{ip}^{+}.

In generic case we should, however, use the series of partial measures defined by Equation (1). Then, we define the order of the partial complexity using the pair of exponents (n, m). The introduction of the order of the partial complexity is in line with our perception of the existence of multiple levels of (full) complexity.

We are able to discover the nature of the CX measure, i.e., its dynamics and, in particular, its dynamical structures, when we analyze the entropy dynamics $S(t)$ (see Figure 2 for details).

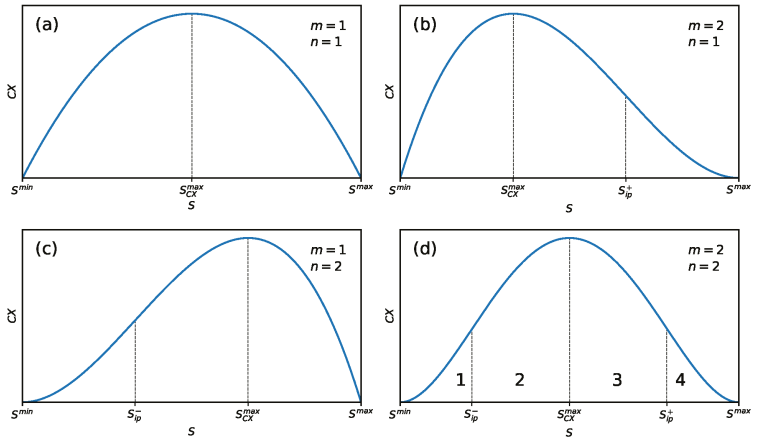


Figure 1. Plots of the partial measure of complexity $CX(S;m,n)$ vs. S given by Equation (1) for four characteristic cases: (**a**) Case $n = m = 1$ where no inflection points, $S_{ip}^{\mp}$ are present. (**b**) Case $m = 2$ and $n = 1$ where a single inflection point S_{ip}^{+} is present. (**c**) Case $m = 1$ and $n = 2$ where a single inflection point S_{ip}^{-} is present. (**d**) Case $m = 2$ and $n = 2$ where both inflection points are present. The shape of the curve, containing two inflection points, is typical for partial measures of complexity, characterized by exponents $m, n \geq 2$. Numbers 1–4 mark individual phases differing in the degree of order.

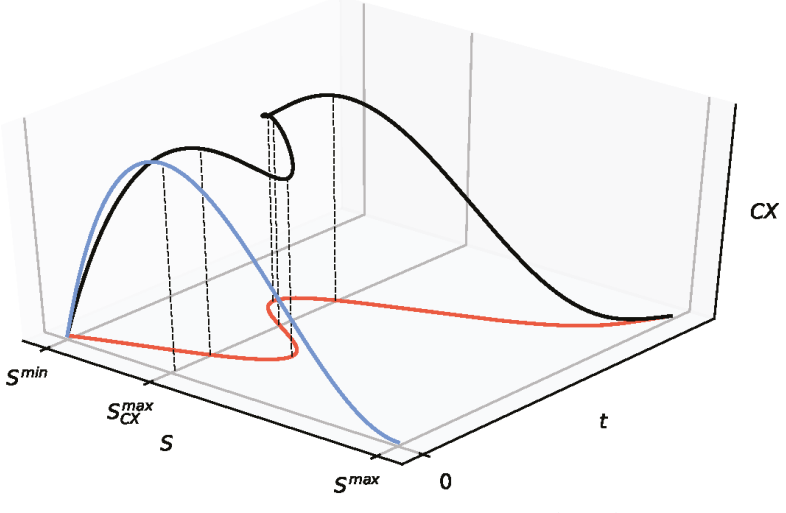


Figure 2. Schematic plot of the partial measure of complexity $CX(S;m,n)$ vs. S and t given by Equation (1). The red curve shows the dependence of entropy S on time t. The black curve represents $CX(S(t);m,n)$ in three dimensions. The blue curve represents projection of the black curve on the (S,CX) plane. We show different variants of this blue curve presented in Figure 1. The non-monotonic dependence of the entropy on time visible here indicates the open nature of the system. However, this non-monotonicity is not visible through the blue curve. For instance, the three local maxima of the black curve collapse to one of the blue curve.

The measurability of the partial measure of complexity is necessary for characterizing it quantitatively and to be able to compare different complexities. Following Gell–Mann [40], we must identify the

scales at which we perform the analysis and thus determine coarse-graining to define the entropy. Its dependence on complexity cannot be ruled out.

However, the question of direct measurement of the partial measure of complexity in an experiment (real or numerical) remains a challenge.

2.5. Remarks on the Partial Entropic Susceptibility

An essential tool for studying phase transitions is the system susceptibility—in our case, the partial entropic susceptibility of the partial measure of complexity. Here, it (additionally) plays the role of the partial order parameter.

The plot of susceptibility $\chi_{CX}(S; m, n)$ vs. S is presented in Figure 3. Four phases, already marked in Figure 1, are visible (also numbered 1 to 4).

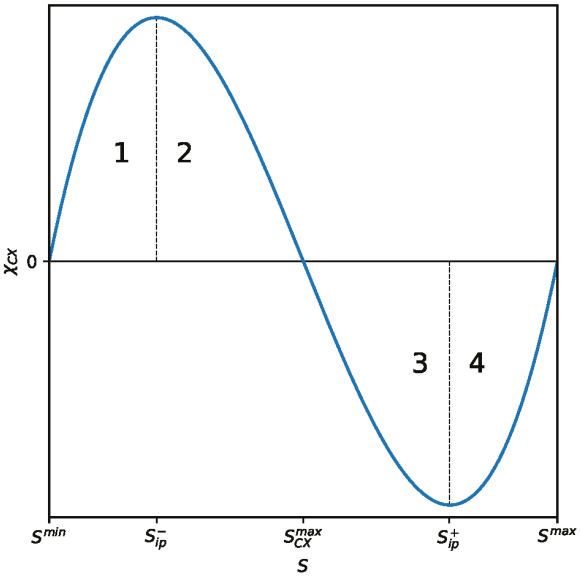


Figure 3. Plot of the partial entropic (non-equilibrium) susceptibility $\chi_{CX}(S; m, n)$ of the partial measure of complexity vs. S given by Equation (7) at fixed order $(m = 2, n = 2)$. The finite susceptibility value at the S_{ip}^{-} and S_{ip}^{+} phase transition points (cf. Figure 1) may be considered to correspond to finite susceptibility value in absorbing the non-equilibrium phase transition in the model of direct percolation at a critical point in the presence of an external field [21]. However, the situation presented here is richer, because susceptibility changes its sign, smoothly passing through zero at $S = S_{CX}^{max}$. At this point, the system is exceptionally robust and, therefore, is poorly affected by data artefacts, because its susceptibility vanishes there.

Phase number 1 is almost entirely ordered—the disordered phase input is residual. At point S_{ip}^{-}, there is a phase transition to the mixed-phase marked with number 2, still with the predominance of the ordered phase. At the S_{ip}^{-} inflection point, the entropic susceptibility reaches a local maximum. By further increasing the entropy of the system, it enters phase 3 as a result of phase transition at the very specific S_{CX}^{max} transition point. At this point, the entropic susceptibility of the partial measure of complexity disappears. This mixed phase (number 3) is already characterized by the advantage of the disordered phase over the ordered one. Finally, the last transition, which occurs at S_{ip}^{+}, leads the system to the dominating phase of the disordered phase—the input of the ordered phase is residual here. At this transition point, the susceptibility reaches a local minimum. Intriguingly, entropic

susceptibility can have both positive and negative value passing smoothly through zero at $S = S^{max}_{CX}$, where the system is exceptionally robust. The presence of phases with positive and negative entropic susceptibility is an exceptionally intriguing phenomenon. The phases discussed above, together with the above-mentioned inflection points, are also marked in Figure 1d. Let us add that the location of the phases mentioned above, i.e., the location of the inflection points, depends on the order (m, n) of the partial measure of complexity. This is clearly seen in Figures 4 and 5.

The values of local extremes of the entropic susceptibility of the partial measure of complexity are finite here and not divergent, as in the case of (equilibrium and non-equilibrium) phase transitions in the absence of an external field. We use this definition to describe the critical behaviour of a system that we demonstrate in Section 2.7, where it requires an explicit dependence on N.

2.6. Universal Full Measure of Complexity

The full universal measure of complexity X is a weighted sum of the partial measures of complexity $CX(S; m, n)$ for individual scales. That is,

$$X(S; m_0, n_0) = \sum_{m \geq m_0, n \geq n_0} w(m, n) CX(S; m, n), \; m_0, n_0 \geq 0, \tag{13}$$

where $w(m, n)$ is a normalized weight, which must be given in an explicit form. This form is to some extent imposed by the power-law form of partial complexity. Namely, we can assume

$$w(m, n) = \left(1 - \frac{1}{M}\right)^2 \frac{1}{M^{m-m_0+n-n_0}}, \; M > 1, \tag{14}$$

which seems to be particularly simple because

$$\frac{w(m+1, n)}{w(m, n)} = \frac{w(m, n+1)}{w(m, n)} = \frac{1}{M}, \tag{15}$$

independently of m and n.

As one can see, Equation (13), supported by Equation (15), is the product of the sums of two geometric series,

$$\begin{aligned} X(S; m_0, n_0) &= (S^{max} - S)^{m_0} \left(1 - \frac{1}{M}\right) \sum_{m \geq m_0} \frac{(S^{max} - S)^{m-m_0}}{M^{m-m_0}} \\ &\times (S^{max} - S)^{n_0} \left(1 - \frac{1}{M}\right) \sum_{n \geq n_0} \frac{(S^{max} - S)^{n-n_0}}{M^{n-n_0}}. \end{aligned} \tag{16}$$

If both series converge for any $S^{min} \leq S \leq S^{max}$, which is the case if and only if the condition $Z(= S^{max} - S^{min}) < M$ is met, then we directly obtain

$$X(S; m_0, n_0) = \left(1 - \frac{1}{M}\right)^2 \frac{(S^{max} - S)^{m_0}}{1 - \frac{S^{max} - S}{M}} \frac{(S - S^{min})^{n_0}}{1 - \frac{S - S^{min}}{M}}. \tag{17}$$

In other words, the M parameter can always be chosen, so that the sums of both series in Equation (21) diverge for all S values. Thus, $m_0, n_0 \geq 1$ is the natural lower limit of m_0, n_0, satisfying the condition of $X(S; m_0, n_0)$ disappearing for $S = S^{min}, S^{max}$. We still assume more strongly that $m_0, n_0 \geq 2$, which has already been explained above.

For extensive systems, Equation (17) can be presented in a form that clearly shows the dependence of the X complexity on the number of entities N, simply replacing S entropy by Ns, where s is already N-independent specific entropy. Subsequently,

$$X(Ns; m_0, n_0) = \left(1 - \frac{1}{M}\right)^2 N^{m_0+n_0} \frac{(s^{max} - s)^{m_0}}{1 - \frac{N}{M}(s^{max} - s)} \frac{(s - s^{min})^{n_0}}{1 - \frac{N}{M}(s - s^{min})}. \quad (18)$$

We emphasize that X does not scale with N, as opposed to partial measures of complexity.

In Figures 6 and 7, we show the dependence of X on N (on the plane) and on N and s (in three dimensions), respectively. We obtained the singularities of full complexity, $N_j^{cr}(s)$, $j = 1, 2$, as a result of the zeroing of denominators in the Equation (17) at nonzero numerators.

Note that, for $M \gg Z$, both measures of complexity have approximate values $X(S;\ m0,\ n0) \approx CX(S;\ m0,\ n0)$. Important differences between these two measures only appear for Z/M close to 1, because only then does the denominator in Equation (17) play an important role. Of course, M is a free parameter, and possibly its specific value could be obtained from some additional (e.g., external) constraint.

In Figure 4, we compare the behaviour of the partial (black curve) and full (orange curve) measures of complexity, where we used the entropy instead of the specific entropy. Whether CX lies below or above X depends both on M parameter (determining the weight at which individual measures of partial complexity enter the full measure of complexity), and on the Z/M ratio.

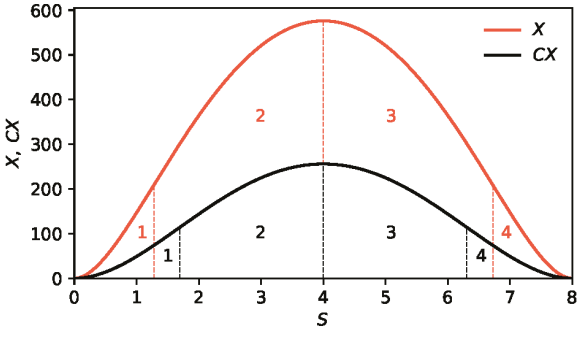


Figure 4. Comparison of the partial measure of complexity $CX(S; m = 2, n = 2)$ given by Equation (1) and full measure of complexity $X(S; m_0 = 2, n_0 = 2)$ given by Equation (17), for instance, for the symmetric case of $m = n = m_0 = n_0$. In addition, we assume that $S^{min} = 0, S^{max} = 8$ and $M = 10$. Vertical dashed lines indicate inflection points: black for the CX curve, orange for the X curve, while $S_{CX}^{max} = S_X^{max} = 4$. Notably, S_X^{max} maximizes X (here at a given ratio $Z/M = 0.8$). Vertical dashed lines mark the locations of inflection points on both curves.

We continue to determine the full entropic susceptibility of the full measure of complexity,

$$\begin{aligned} \chi_X(S; m_0, n_0) &= \frac{dX(S; m_0, n_0)}{dS} = (m_0 + n_0)(S_{CX}^{max} - S)X(S; m_0 - 1, n_0 - 1) \\ &+ \frac{2}{M^2} X(S; m_0, n_0) \frac{S - S^{arit}}{\left(1 - \frac{S^{max} - S}{M}\right)\left(1 - \frac{S - S^{min}}{M}\right)},\ m_0, n_0 \geq 1, \end{aligned} \quad (19)$$

where S_{CX}^{max} is given here by Equation (4) but for $m = m_0$ and $n = n_0$. Notably, for the symmetric cases $m = n$ and/or $m_0 = n_0$, we have $S_{CX}^{max} = S_X^{max} = S^{arit}$, which are independent of m, m_0.

Similarly to the partial entropic susceptibility of a partial measure of complexity, we obtain the full entropic susceptibility of a full measure of complexity,

$$\begin{aligned} \chi_X(Ns; m_0, n_0) &= \frac{dX(S; m_0, n_0)}{dS} = (m_0 + n_0)\ N\ (s_{CX}^{max} - s)X(Ns; m_0 - 1, n_0 - 1) \\ &+ \frac{2}{M^2} X(Ns; m_0, n_0) N \frac{s - s^{arit}}{\left(1 - \frac{N}{M}(s^{max} - s)\right)\left(1 - \frac{N}{M}(s - s^{min})\right)},\ m_0, n_0 \geq 1, \end{aligned} \quad (20)$$

where $s_{CX}^{max} = S_{CX}^{max}/N$, $s^{arit} = S^{arit}/N$, $s^{min} = S^{min}/N$, and $s^{max} = S^{max}/N$. The progression of susceptibility $\chi_X(S; m_0, n_0)$, depending on S, for selected parameter values is shown in Figure 5. This progression course is similar to the analogous one that is presented in Figure 3.

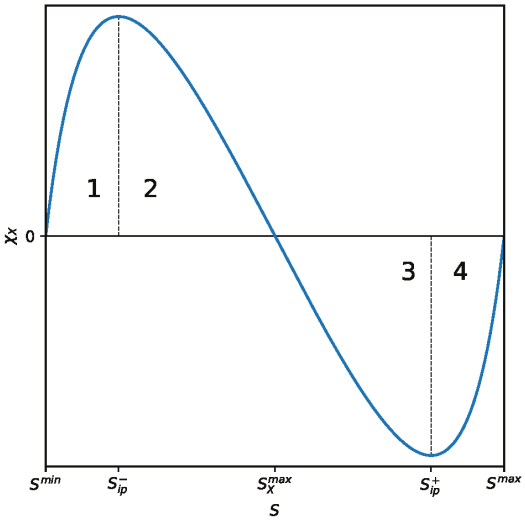


Figure 5. Plot of the full entropic susceptibility $\chi_X(S; m_0, n_0)$ of the full measure of complexity vs. S given by Equation (19), at arbitrary fixed order $(m_0 = 2, n_0 = 2)$. As expected from the comparison with Figure 3, the turning points of CX (cf. Figure 4) lie within the S interval bounded by inflection points of X.

Thus, the evolution of X is governed by an equation that is analogous to Equation (6), except that χ_{CX} present in that equation should be replaced by χ_X given by Equation (19). Therefore, we have

$$\frac{dX(S(t), m_0, n_0)}{dt} = \chi_X(S(t); m_0, n_0)\frac{dS(t)}{dt}. \tag{21}$$

The relationship between measures of complexity and time is implicit in our work—complexity indirectly depends on time through the dependence of entropy on time. It should be emphasized that the dependence of entropy on time is external in our approach—it can be taken into account based on additional modelling that is dedicated to specific real situations. We have already signalled this when discussing Equation (6).

2.7. Criticality in Extensive Systems

By using Equation (17), we show when the universal full measure of complexity diverges and, thus, the system enters a critical state. We assume that we are dealing with an extensive system, i.e., that Equation (17) can be represented as

$$X(Ns; m_0, n_0) = \left(1 - \frac{1}{M}\right)^2 N^{m_0+n_0} \frac{(s^{max} - s)^{m_0}}{1 - \frac{N}{M}(s^{max} - s)} \times \frac{(s - s^{min})^{n_0}}{1 - \frac{N}{M}(s - s^{min})}, \quad \frac{Nz}{M} < 1, \tag{22}$$

where entropy densities $s(= S/N), s^{min}(= S^{min}/N), s^{max}(= S^{max}/N)$ are (at most) slowly varying functions of the number N of elements making up the system and special entropy span $z = s^{max} - s^{min}$. As one can see, the measure X is divergent in two critical points $N_{cr}^{max}(s) = \frac{M}{s^{max} - s}$ and $N_{cr}^{min}(s) = \frac{M}{s^{min} - s}$,

where $s^{min} < s < s^{max}$. Moreover, the susceptibilities given by Equations (19) and (20) diverge at the same points where measures of complexity given by Equations (17) and (18) diverge, which underlines the self-consistency of our approach.

Equation (22) can now be written in a form that explicitly includes both critical points (both physical and non-physical):

$$X(Ns; m_0, n_0) = \left(1 - \frac{1}{M}\right)^2 N^{m_0+n_0} \frac{(s^{max} - s)^{m_0}}{\left(1 - \frac{N}{N_{cr}^{max}(s)}\right)^{\beta^{max}}} \times \frac{(s - s^{min})^{n_0}}{\left(1 - \frac{N}{N_{cr}^{min}(s)}\right)^{\beta^{min}}}, \tag{23}$$

where critical exponents assume the mean-field values $\beta^{max} = \beta^{min} = 1$. In this case, we could speak of two-criticality were it not for the fact that one of these criticalities is unphysical.

Figure 6 shows dependence $X(Ns; m_0, n_0)$ vs. N at fixed $s = 0.8$. The values of parameters are shown there, while the specific entropy s is chosen so that the condition $s - s^{min} < s^{max} - s$ is satisfied (this is equivalent to a condition $s < s^{arit}$). This means that s is closer to s^{min} than s^{max}. The existence of these divergences is a signature of criticality. However, the situation for borderline cases $s = s^{min}$ or $s = s^{max}$ changes rapidly—it is a different consideration.

Critical numbers of entities in the system $N_{cr}^{max}(s)$ and $N_{cr}^{min}(s)$ are determined by the ratio of the M parameter characterizing the hierarchy/cascade of scales in the system and the distance between entropy density s and its extreme values s^{min} and s^{max}. The construction of these critical numbers resembles the canonical critical temperature structure for the Ising model in the mean-field approximation, where $\beta_c J z = 1$ (here $\beta_c = 1/k_B T_c$ and k_B is the Boltzmann constant). In our case, the role of the inverse temperature β_c is played by N_{cr}^{max} and N_{cr}^{min}, the role of the coupling constant J is $1/M$, while the role of the mean coordination number z is played by $s^{max} - s$ and $s - s^{min}$, respectively.

The hierarchy is the source of criticality here. Criticality is an immanent feature of our full description of complexity. Nevertheless, in this work, we do not specify the sources of this hierarchy—it could be self-organized criticality or due to some other sources.

For the sake of completeness, note that the dependence on N of the partial measure of complexity is given by Equation (2). This means that for extensive systems this measure increases powerfully depending on N. Therefore, only the weighted infinite sum of these measures generates the existence of singularity.

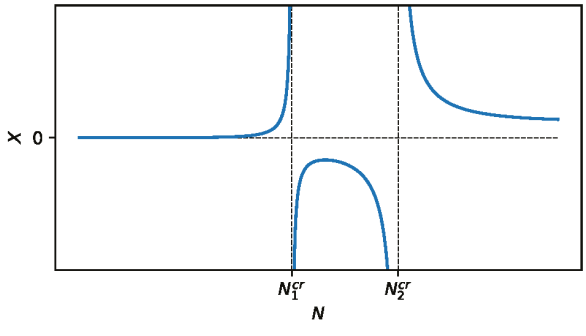


Figure 6. Dependence of the universal full measure of complexity X vs. number of entities N given by Equation (23). It should be emphasized that the full measure of complexity and its susceptibility have singularities in the same points. As one can see, we are dealing here with complexity barriers separating the phases/states of the system and the small and large number of objects forming them. The parameters we adopted here are as follows: $M = 30$, $s^{min} = 0$, $s^{max} = 2$, $s = 0.8$, $m_0 = n_0 = 2$, hence, point $N_{cr}^{max}(s = 0.8) = 25$ and point $N_{cr}^{min}(s = 0.8) = 37.5$.

Let us now consider in more detail the behaviour of $X(Ns; m_0, n_0)$ depending on N and s. A three-dimensional plot of Figure 7 will be helpful here. One can see how the mutual location of the singularities of $N_{cr}^{max}(s)$ and $N_{cr}^{min}(s)$ changes with the increase of s. From the situation of $s < s^{arit}$, in which $N_{cr}^{max}(s) > N_{cr}^{min}(s)$, through the situation when $s = s^{arit}$ in which $N_{cr}^{max}(s) = N_{cr}^{min}(s)$, up to the situation in which $N_{cr}^{max}(s) > N_{cr}^{min}(s)$ for $s > s^{arit}$.

It must be clearly stated that the area physically accessible is the one in front of the first singularity, which is further emphasized in Figure 7 by blue curves. Let us emphasize that the N range in which criticality occurs is sufficient to cover the corresponding values of N discussed in the literature to date, especially the Dunbar numbers [46–49] (e.g., $N = 5, 15, 50$, and $N = 150$). However, it should be noted that our view of complexity is complementary to that presented in the literature.

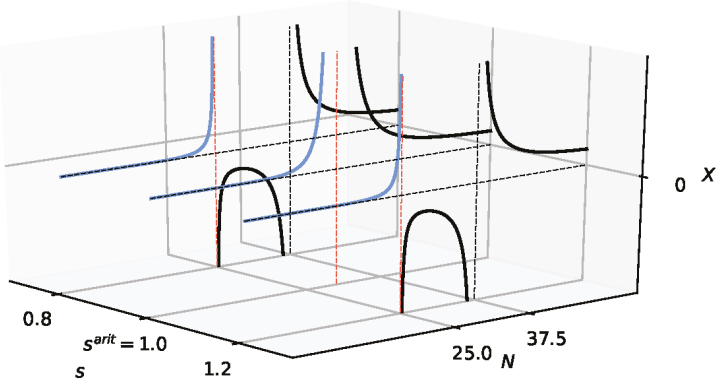


Figure 7. Dependence of the universal full measure of complexity X vs. number of entities N and specific entropy s given by Equation (23), for $m_0, n_0 \geq 1$. Notably, the full measure of complexity and its susceptibility have singularities at the same points $N_{cr}^{max}(s)$ and $N_{cr}^{min}(s)$. We are dealing here with complexity barriers separating the phases/states of the system and the small and large number of entities that form them. The parameters we adopted here are, as follows: $M = 30$, $s^{min} = 0$, $s^{max} = 2$, $s = 0.8$, $m_0 = n_0 = 2$. These are the same parameters that we used to construct the plain plot in Figure 6.

3. Finger Print of Complexity in Simplicity

Let us consider a perfect gas at a fixed temperature, which is initially closed in the left half of an isolated container. The partition is next removed, and the gas undergoes a spontaneous expansion. Here we are dealing (practically speaking) with an irreversible process, even for a small number of particles (at least the order of 10^2).

Let us recall the definition of 'perfect gas'. It is a gas of particles that cannot 'see' each other, i.e., there are no interactions between them. Thus, from a physical point of view, it is a dilute gas at high temperature. We further assume that all of the particles have the same kinetic energy. A legitimate question is whether such a gas will expand after the partition is removed. We notice that the thermodynamic force is at work here, being roughly proportional to the difference in the number of particles in the right and left parts of the container. This force causes the expansion process. Thus, we are dealing with the simplest paradigmatic irreversible process [50]. The particles remain stuck in the final state and will not leave it (with accuracy subject to slight fluctuations in the number of particles in the right half of the container). Such a final state of the whole system is referred to as the equilibrium state. The simple coarse-grain description of the system allows us to introduce here the concept of configuration entropy.

Note that the macroscopic state of the system (generally, the non-equilibrium and non-stationary/relaxing one) can be described by the instantaneous number of particles in the left

($N_L(t)$) and right ($N_R(t)$) parts of the container, with $N = N_L(t) + N_R(t)$, where N is the fixed total number of particles in the container (isolated system). It allows for one to define the weight of the macroscopic state $\Gamma(N_L(t))$, also called thermodynamic probability. This is the number of ways to arrange the $N_L(t)$ particles in the left part of the container and $N_R(t) = N - N_L(t)$ in the right. Hence,

$$\Gamma(N_L(t)) = \frac{N!}{N_L(t)!(N - N_L(t))!}. \tag{24}$$

Here we do not distinguish permutations of particles inside each part of the container separately. We only take into account permutations of particles located in different halves of the container. This is because our resolution here is too small to observe the location of particles inside each container separately. Such a coarse-graining creates an information barrier: more information can mask the complexity of the system. We will not be able to see the complexity, because we will not be able to construct entropy. This creates a paradoxical situation: the surplus of information makes the task difficult and does not facilitate obtaining the insight into the system. Here we have an analogy with chaotic dynamics, where chaos is only visible in the Poincaré surface cross-section of the phase space and not in the entire phase space.

The configuration entropy at a given time t we define, as follows,

$$S(N_L(t)) = \ln \Gamma(N_L(t)), \tag{25}$$

where $\Gamma(N_L(t))$ is given by Equation (24). The above expression can be used both for the equilibrium and non-equilibrium states.

It can be demonstrated using the Stirling formula that for large N, entropy S is reduced to the BGS form,

$$\ln \Gamma(N_L(t)) = -N\left[p_L(t) \ln p_L(t) + p_R(t) \ln p_R(t)\right] = Ns(t), \tag{26}$$

where $p_J(t) \stackrel{\text{def.}}{=} \frac{N_J(t)}{N}$, $J = L, R$, and $s(t)$ is a specific entropy. The law of entropy increase Equation (A8) is also fulfilled here, as expected.

We now prepare the equation for determining $N_L^{CX^{max}}$, i.e., the number of particles in the left part of the container that maximizes the partial complexity measure CX. To this end, we assume, for instance, the symmetric partial measure of complexity of the order of $(m = 2, n = 2)$. Next, we substitute $N_L = N_L^{CX^{max}}$ into the both sides of Equation (25) and according to constitutive Equation (5), we equate Equation (25) to S_{CX}^{max}. Hence, we obtain a constitutive equation for the relaxing perfect gas,

$$S(N_L(t) = N_L^{CX^{max}}) = S_{CX}^{max}, \tag{27}$$

where $N_L^{CX^{max}}$ is our sought quantity.

Now, we need to independently determine S_{CX}^{max}. Recall that the number of N_L particles that maximize entropy is the number of N_L^{eq} particles in the statistical/thermodynamic equilibrium state of the system. This number is equal to half of all particles in the container, i.e., $N_L^{eq} = N/2$. It can still be assumed (without reducing the general considerations) that $S^{min} = 0$. Therefore,

$$S^{max} = S(N/2). \tag{28}$$

However, from Equation (A5), we know that $S_{CX}^{max} = S^{max}/2$. By using it, we transform Equation (27) into the form,

$$S\left(N_L^{CX^{max}}\right) = \frac{1}{2} S(N/2). \tag{29}$$

Equation (27) is an example of the general constitutive Equation (5), where $N_L^{CX^{max}}$ plays the role of $Y^{CX^{max}}$. This equation has the following explicit form,

$$\left[\Pi_{j=1}^{N-N_L^{CX^{max}}}\left(1+\frac{N_L^{CX^{max}}}{j}\right)\right]^2=\Pi_{j=1}^{N/2}\left(1+\frac{N}{2j}\right),\ \text{for } n=m=2. \tag{30}$$

Just deriving Equation (30) (see Appendix C for details) is the primary purpose of this example. This is a transcendental equation of which the exact analytical solution is unknown. When deriving Equation (30), we used the initial condition for the entropy that is, $S(t=0)=S^{min}=\ln\Gamma(N_L=N)=0$, which follows from Equations (24) and (25). Even for such a simple toy model, determining the partial measure of complexity is a non-trivial task, also because N_L is different from $N/2$ (as we show below).

The numerical solutions of Equation (30), i.e., the relationship of $N_L^{CX^{max}}$ to N, are shown in Figure 8 (for simplicity, L defining the vertical axis on the plot means $N_L^{CX^{max}}$). Both of the solutions (small circles above and below the solid straight line) show that $N_L^{CX^{max}}$ is significantly different from $N/2$. Thus, the most complex state is significantly different from the equilibrium state.

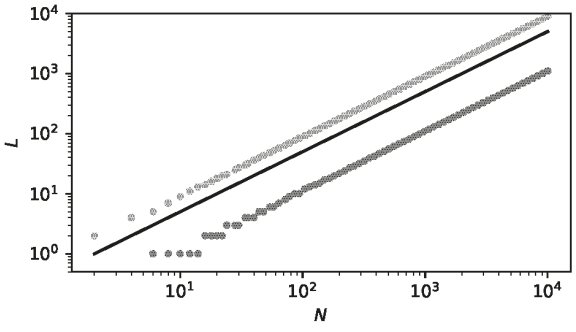


Figure 8. Dependence of $L(=N_L^{CX^{max}})$ vs. N. There are two solutions of Equation (30): one marked with blue circles and the other with orange ones. Above $N\approx 10^2$, both dependencies are linear, which is particularly clearly confirmed in Figure 9. That is, in a log-log scale, their slopes equal 1. However, in linear scale, the directional coefficients of these straight lines equal 0.11 and 0.89, respectively. This is clearly shown in Figure 9. Only the solution with orange circles is realistic, because the chance that 89% of particles will pass in a finite time to the second part of the container (as indicated by the solution marked with blue circles) is negligibly small. The black solid tangent straight line indicates a reference case $N_L^{CX^{max}}=N/2$.

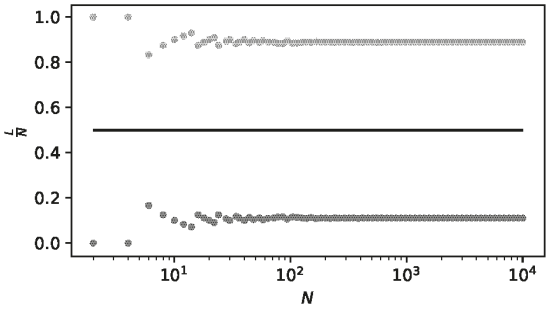


Figure 9. Directional coefficient of linear dependencies L vs. N as a function of N. For N greater than 10^2, no N-dependence of this coefficient is observed. Both of the solutions (having $L/N=0.11$ and $L/N=0.89$) are mutually symmetric about the straight horizontal line $L/N=1/2$, but we only consider the solution $L/N=0.89$ to be realistic. The black horizontal straight solid line indicates a reference case $N_L^{CX^{max}}=N/2$.

Having the $N_L^{CX^{max}}$ dependence on N, we can obtain the dependence of the partial measure of specific complexity $cx^{max} \stackrel{\text{def.}}{=} CX^{max}/N^{m+n}$ on N order $(m = 2, n = 2)$. We can write

$$cx^{max} = \left(\frac{s(N/2)}{2}\right)^4 = \left[\frac{1}{2N}\ln\left(\Pi_{j=1}^{N/2}\left(1+\frac{N}{2j}\right)\right)\right]^4, \tag{31}$$

as in our case $s^{max} = s(N/2)$ equals the logarithm of the right-hand side of Equation (30) divided by N. Notably, Equation (31) is based on Equation (A12).

In Figure 10, we present the dependence of cx^{max} on N. Quantity cx^{max} is a non-extensive function—it reaches the plateau for $N \gg 1$. For $N \approx 10^4$ the plateau is achieved with a good approximation. This is important for researching complexity. Namely, systems can attain complexity already on a mesoscopic scale. Although the absolute value of the complexity measure is relatively small, it is evident and possesses a structure that is related to the current inflection point there (near $N = 10$).

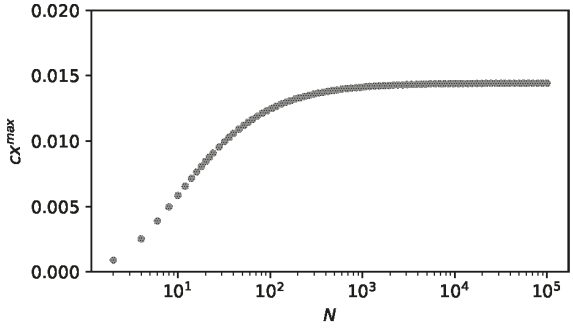


Figure 10. Dependence of cx^{max} on N given by Equation (31). As one can see, cx^{max} is a non-extensive function—it reaches the plateau for $N \gg 1$. For $N \approx 10^4$ the plateau is achieved with a good approximation. This is an important issue for researching complexity. Namely, systems can attain complexity already on a mesoscopic scale. It can be said that the curve's inflection point (located near $N = 10$) marks the beginning of the complexity stabilization region.

This example shows that even such a simple arrangement of non-interacting objects may have non-equilibrium non-stationary complexity. A necessary (but not sufficient) condition is the possibility of constructing entropy and the presence of a time arrow.

4. Concluding Remarks

In many recent publications [5,8,9,51] it is argued that entropy can be a direct measure of complexity. Namely, a smaller value of entropy indicates more regularity or lower system complexity, while its larger value indicates more disorder, randomness and higher system complexity. However, according to Gell–Mann, more disorder means less, and not more, system complexity. These two viewpoints are contradictory—this is a serious problem, which we have addressed.

Our motivation in solving the above problem was based on Gell–Mann's view of complexity. This is because we fail to agree that the loss of information by the system as it approaches equilibrium increases its complexity; notably, $\Delta I(p^{eq}, p^{eq})$ (see Appendix B for detail) takes its minimum value then, and complexity must decrease.

In addition, the differences between entropies in Equation (1) eliminate the useless dependence of complexity on the additive constant that may appear in the definition of entropy. It can be said that the

system state with the highest complexity is the state most distant from all of the states of the system of lesser or no complexity.

Thus, in the sense of Gell–Mann, the measure of complexity should supply complementary information to the entropy or its monotonic mapping.

Therefore, in this work, we have presented a methodology which allows building a universal measure of complexity as a function of a system state based on non-linearly transformed entropy. This is a non-extensive measure. This measure should meet a number of conditions/axioms, which we have indicated in this work. A parsimonious example, of the simplest system with a small and a large number of degrees of freedom, is presented in order to support our methodology. As a result of this approach, we have shown that (generally speaking) the most complex are optimally mixed states consisting of pure states, i.e., of the most regular and most disordered, which the space of states of a given system allows. This also applies to the distinctive examples outlined in Appendixes D and E (although this requires a redefinition of some variables and parameters).

We should pay attention to an essential issue regarding the definition of the phenomenological partial measure of complexity that is given by the Equation (1). This definition is open in the sense that if the description of complexity requires, for example, one additional quantity, then the Equation (1) takes on an extended form,

$$CX(S,E;m_1,n_1,m_2,n_2) \stackrel{\text{def.}}{=} (S^{max}-S)^{m_1}(S-S^{min})^{n_1}(E^{max}-E)^{m_2}(E-E^{min})^{n_2} \geq 0, \tag{32}$$

whereby $E^{min} \leq E \leq E^{max}$ this new quantity is marked. This definition has still an open character. Specifically, this definition also allows (if the situation requires) the replacement of one quantity with another, e.g., entropy with free energy, or considering some derivatives (e.g., of the type $\frac{\partial S}{\partial E}$). Openness and substitutability should be the key features of the measure of complexity. Moreover, exponents $m_j,\ n_j,\ j=1,2$, determine the order of complexity, i.e., its level or scale. We emphasize that the measure of complexity introduced can describe isolated and closed systems (although in contact with the reservoir), as well as open systems that can change their elements.

From Equations (13) and (32), we get the phenomenological universal full measure of complexity in the form, which extends Equation (17),

$$\begin{aligned} X(S,E;m_1^0,n_1^0,m_2^0,n_2^0) &= \left(1-\frac{1}{M_S}\right)^2 \frac{(S^{max}-S)^{m_1^0}}{1-\frac{S^{max}-S}{M_1}} \frac{(S^{min}-S)^{n_1^0}}{1-\frac{S^{min}-S}{M_1}} \\ &\times \left(1-\frac{1}{M_2}\right)^2 \frac{(E^{max}-E)^{m_2^0}}{1-\frac{E^{max}-E}{M_2}} \frac{(E^{min}-E)^{n_2^0}}{1-\frac{E^{min}-E}{M_2}} \geq 0. \end{aligned} \tag{33}$$

The full measure of complexity is a weighted sum of partial measures of complexity across all complexity scales. As one can see, this full measure may contain singularities. They are the necessary signatures of criticality existing in the system. This meets the expectations presented in the literature.

Definitions of measures of complexity Equations (1) and (17) and their possible extensions are universal and useful. It is due to entropy that is associated not only with thermodynamics (Carnot, Clausius, Kelvin) and statistical physics (Boltzmann, Gibbs, Planck, Rényi, Tsallis), but also with the information approach (Shannon, Kolmogorov, Lapunov, Takens, Grassberger, Hantschel, Procaccia), and with the approach from the side of cellular automata (von Neumann, Ulam, Turing, Conway, Wolfram, et al.), i.e., with any representation of the real world using a binary string. Today, we already have several very effective methods for counting entropy of such strings, as well as other macroscopic characteristics sensitive to organization and self-organizing systems, as well as to their synchronization (synergy, coherence), competition, reproduction, adaptation—all of them sometimes having local and sometimes global characters.

Our definition of complexity also extends to meet research into the complexity of the biologically active matter. In this, especially research on the consciousness of the human brain can derive a

fresh impulse. The point is that most researchers believe that the main feature of conscious action is a maximum complexity or even a critical complexity [52]. In our approach, it would be CX^{max} and $X(N_1^{cr}s)$.

We hope that our approach will enable: (i) the universal classification of complexity, (ii) the analysis of a system critical behaviour and its applications, and (iii) the study of dynamic complexity. All of these constitute the background to the science of complexity.

Appendix A. Properties of the Partial Measure of Complexity

It is worth paying attention to Equations (1)–(4). For a fixed span of Z and the order (m, n), there may still be systems of different complexities. The complexity description only using CX^{max} is insufficient, because there can be many systems with the same span and order. However, we assume systems to be equivalent, i.e., belonging to the same complexity class (Z, m, n), if they have the same span at a given order. We can still distinguish them as, in general, they differ in the location of S_{CX}^{max}. We can say that a given class has a greater partial measure of complexity if it has a larger CX^{max}. In a given class, the system has a larger complexity if the system stays closer to CX^{max}, i.e., its current entropy S is closer to S_{CX}^{max}. For a given CX with Equation (1), we obtain (for each order (m, n)) two S solutions: one on the left and the other on the right of CX^{max} (except when $S = S_{CX}^{max}$). Division into classes allows us to introduce an additional classification of complexity.

A distinction should be made between two cases of measuring complexity: (i) $Z < m + n$ and (ii) $Z > m + n$. This is particularly evident when we consider the ratio of both types of complexity measures for $m + n > 1$,

$$\frac{CX^{max}(Z_i)}{CX^{max}(Z_{ii})} = \left(\frac{Z_i}{Z_{ii}}\right)^{m+n} < 1, \tag{A1}$$

where Z_i belongs to case (i), while Z_{ii} to case (ii). Thus, the greater the exponent $m + n$, the greater the difference between $CX^{max}(Z_{ii})$ and $CX^{max}(Z_i)$.

The alternate form of Equation (1),

$$CX(\Delta) = \left(\frac{n}{n+m}Z + \Delta\right)^n \left(\frac{m}{n+m}Z - \Delta\right)^m, \tag{A2}$$

where deviation $\Delta = \Delta(t) = S(t) - S_{CX}^{max}$ makes the operating of the CX coefficient easier in the vicinity of S_{CX}^{max}, where the parabolic expansion is valid. We then have:

$$CX(\Delta) \approx CX^{max}\left[1 - \frac{1}{2mn}\left((n+m)\frac{\Delta}{Z}\right)^2\right] \approx CX^{max}\exp\left(-\frac{1}{2mn}\left((n+m)\frac{\Delta}{Z}\right)^2\right), \tag{A3}$$

that is a Gaussian form, which has variance $\sigma^2 = \frac{nm}{(n+m)^2}Z^2$. Only in the narrow range of S around S_{CX}^{max} is the measure of complexity CX symmetrical regardless of the order (m, n).

In fact, only the location of the maximum of $CX(S; m, n)$ is determined (for a given range of S) by the ratio of m to n. However, to have dependence of coefficient CX on entropy in the entire entropy range $S^{min} \leq S \leq S^{max}$, it is necessary to determine two extreme values of entropy (S^{min} and S^{max}) and two exponents (n and m). In general, finding these parameters and exponents is still far from trivial because they have a contextual (and not a universal) character.

However, in a particular situation, when the maximum complexity is symmetrical, i.e., when $m = n$, we obtain

$$S_{CX}^{max} = \overline{S} = \frac{S^{min} + S^{max}}{2} \tag{A4}$$

and

$$CX^{max} = \left(\frac{Z}{2}\right)^{2n}. \tag{A5}$$

Equation (1) of the partial measure of complexity applies both to single- and multi-particle problems because entropy can also be built even for a very long single-particle trajectory. Moreover, Equation (1) emphasizes our point of view that any evolving system for which one can introduce the concept of entropy and which has a state of thermodynamic equilibrium (for which entropy reaches a global maximum) contains at least a signature of complexity. For systems of negligible complexity, i.e., for which $S \approx S^{min}$ or $S \approx S^{max}$, the measure $CX(S; m, n)$ is close to zero. This does not mean, however, that we cannot locate S^{max}_{CX} near S^{min} or S^{max}. It is then sufficient to have strongly asymmetric situations when $n \ll m$ or $n \gg m$, respectively.

Appendix B. Non-Stationary Entropies

In this Appendix, we sketch a non-stationary situation, which is what we are dealing with throughout this work. We are dealing with systems evolving to a state of statistical equilibrium, even from states far from statistical equilibrium.

Non-stationary entropy is understood to be entropy based on coarse-grained time-dependent probability distributions—this type of entropy is most often used [43–45,51]. A very characteristic example is the entropy class built on time-dependent probability distributions, $\{p_j(t)\}$, satisfying the master (Markovian) or M-equation, in the presence of detailed balance conditions. Thus, we are only considering systems evolving to statistical equilibrium.

Here we give two very characteristic (non-equivalent) examples of non-stationary entropies (more preciselly: one should talk about specific entropies). In addition, entropies given by Equations (A6) and (A7) belong to the category of relative entropies. Namely,

$$S(t) = S_0 \left[1 - \sum_j p_j^{eq} f\left(\frac{p_j(t)}{p_j^{eq}}\right)\right] \tag{A6}$$

and

$$S(t) = -S_0 \ln \sum_j p_j^{eq} f\left(\frac{p_j(t)}{p_j^{eq}}\right), \tag{A7}$$

where $p_j(t)$ is a probability of finding a system in state j at time t, while p_j^{eq} is a corresponding equilibrium probability. We are considering only discrete states here. The function $S_0 f(x) \geq 0$, where domain $0 \leq x \leq \infty$, is a non-negative convex function obeying $S_0 \frac{d^2 f}{dx^2} \geq 0$. It can be shown [45] that entropies defined in this way meet the law of entropy increase, i.e., its derivative

$$\frac{dS(t)}{dt} \geq 0; \tag{A8}$$

therefore $S(t) \to S^{max}$ from below when $p_j(t) \to p_j^{eq}$, for any j. Equation (A8) is the key property of entropy. Let us add that at the limit $p_j(t) = p_j^{eq}$, for any j, entropy defined by Equations (A6) and (A7) disappears. In other words, these entropies are negative and grow to zero as the system tends to equilibrium.

It is worth paying attention to the possibility of defining generalized information gain, whereby this information gain is calculated here relative to the equilibrium distribution. We can write,

$$\Delta I(p(t), p^{eq}) = -S(t), \tag{A9}$$

where $p(t) = \{p_j(t)\}$ and $p^{eq} = \{p_j^{eq}\}$. Furthermore, entropy $S(t)$ is closely related to partition function. Therefore, in this approach, the entropy is a base function.

Most often the function $f(x)$ is selected in the form [45,56],

$$f(x) = x^{\alpha},\ \alpha > 1, \tag{A10}$$

coupled with a constant $S_0 = \frac{1}{\alpha - 1}$, where α can converge to 1. With these choices the entropy given by Equation (A6) is called Tsallis (relative) entropy and the entropy given by Equation (A7) Rényi (relative) entropy. Usually, the entropic index α is denoted by q in the case of Tsallis entropy.

Entropies given by Equations (A6) and (A7) converge, with the help of Equation (A10), to Kullback-Leibler entropy [9] when entropy index $\alpha \to 1$. However, the Rényi and Tsallis entropies are essentially different for $\alpha \neq 1$. The Rényi entropy is an additive function describing extensive systems, while the Tsallis is not. It is a non-additive function describing non-extensive systems.

Using relative entropy in the definition of complexity measures is productive. It is because other types of entropy can be derived from it, such as ordinary entropy (or Boltzmann-Gibbs-Shannon one) and conditional entropy.

Appendix C. Derivation of the Constitutive Equation for Perfect Gas

The derivation of constitutive Equation (30) comes down to presenting both sides of Equation (29) in explicit form, shortening of common factors, proper organization and presentation. Accordingly, the left side of Equation (29) takes the form,

$$\begin{aligned} S_{CX}^{max} &= S\left(N_L^{CX^{max}}\right) = \ln\Gamma\left(N_L^{CX^{max}}\right) = \ln\frac{N!}{N_L^{CX^{max}}!\left(N - N_L^{CX^{max}}\right)!} \\ &= \ln\frac{\left(N_L^{CX^{max}}+1\right)\left(N_L^{CX^{max}}+2\right)\cdot\ldots\cdot(N-1)N}{1\cdot 2\cdot 3\cdot\ldots\cdot\left(N - N_L^{CX^{max}}\right)} \\ &= \ln\Pi_{j=1}^{N-N_L^{CX^{max}}}\left(1 + \frac{N_L^{CX^{max}}}{j}\right), \end{aligned} \tag{A11}$$

where the number of factors in the numerator and denominator is the same and equals $N - N_L^{CX^{max}}$.

As for the right side of Equation (29), we present it in an explicit form,

$$\begin{aligned} \frac{1}{2}S(N/2) &= \frac{1}{2}\ln\Gamma(N/2) = \frac{1}{2}\ln\frac{N!}{\left(\frac{N}{2}\right)!\left(\frac{N}{2}\right)!} \\ &= \frac{1}{2}\ln\frac{\left(\frac{N}{2}+1\right)\left(\frac{N}{2}+2\right)\left(\frac{N}{2}+3\right)\cdot\ldots\cdot\left(\frac{N}{2}+\frac{N}{2}\right)}{1\cdot 2\cdot 3\cdot\ldots\cdot\frac{N}{2}} = \frac{1}{2}\ln\Pi_{j=1}^{N/2}\left(1+\frac{N}{2j}\right). \end{aligned} \tag{A12}$$

Comparison of Equations (A11) and (A12) just leads to Equation (30).

Appendix D. Entropies of Time Series—A Sketch

The entropy study of various time series is a crucial issue in system dynamics. The point is that the activity of the systems is perceived precisely through time series. The study of nonlinear time series is particularly important. Below we outline two essential methodologies for constructing entropy (including a multi-scale one). Then we show how to connect our complexity measure with these methodologies.

Appendix D.1. Entropy of Embedded Time Series

Various time series from stock exchanges or Forex quotations are the central sources of empirical data available from financial markets. The main question is about the entropy of the time series and hence about the measure of time series complexity. Following Li et al. [51], we present a method of constructing entropy for a finite time series.

We consider the time series $\{x_j\}_{j=1}^{N}$ consisting of N elements x_j, $j = 1, 2, \ldots$. From this we select $1 \leq i \leq N - l + 1$ sub-series indexed by i. Each sub-series consists of l components defined as follows, $y_i^l(k) : 0 \leq k \leq l - 1$, where (for given indexes i and l) component $y_i^l(k) \stackrel{\text{def.}}{=} x(i + k)$. As one can see, m here means the embedding dimension.

Moreover, two subsequent sub-series characterized by the same l have $l - 2$ elements in common. In the collection of sub-series, which create m-dimensional vector space, one can enter a topology defined by the metric d_{ij} defining the distance between arbitrary sub-series i and j. Then one can build a distribution (histogram), $p(d_{ij})$, of distances between vectors. Of course, the question of how to choose the embedding dimension l is fully justified. For example, one could assume that this dimension is equal to the correlation dimension [7]. However, we treat l as a free parameter, and we do not impose any additional restrictions on it. With the probability distribution, $p(d_{ij})$, one can build the entropy (for example, $S(l) = -\langle \ln p(d_{ij}) \rangle$) and hence the complexity of the time series (see Section 2.1 and Appendix B for details).

Similarly to the example with expanding gas in Section 3 and based on Equation (5), we can formulate the constitutive equation in the form

$$S(l^{CX^{max}}) = \frac{\frac{1}{m} S(l^{max}) + \frac{1}{n} S(l^{min})}{\frac{1}{m} + \frac{1}{n}}, \tag{A13}$$

where $S(l^{max}) = S^{max}$ and $S(l^{min}) = S^{min}$.

The transcendental Equation (A13) should be solved numerically due to the value of the unknown $l^{CX^{max}}$ sought. However, first, the values of l^{min} and l^{max} must be found (also numerically), which leads to the determination of S^{min} and S^{max}, respectively. It allows us to determine CX^{max}. In the final step, one can (similar to that presented in Appendix B) find informative relationships of the above quantities $l^{CX^{max}}$, $l^{CX^{max}}/N$ and CX^{max} from N.

This is all possible if one has sufficiently useful statistics, i.e., when $N - l + 1 \gg l \equiv \frac{N-1}{2} \gg l$.

Appendix D.2. Multi-Scale Entropy

We can now proceed to define multi-scale complexity, but first we need to define multi-scale entropy or the hierarchy of entropies. For this purpose, we prepare the coarse-grained scheme. The primary time series consists of N elements. We divide it into $n = Int[N/\tau]$ non-overlapping intervals, where τ is the time horizon/scale, i.e., the number of time steps that we use to separate the elements of the first time series. Now we can build a new time series, of which the non-overlapping elements are defined as arithmetic means in subsequent intervals of τ,

$$y_j^\tau = \frac{1}{\tau} \sum_{i=j\tau+1}^{(j+1)\tau} x_i \, , \; 0 \leq j \leq n - 1. \tag{A14}$$

More can be said about the choice of τ using the (bilinear) autocorrelation function

$$AC(t) \approx \frac{1}{n - t} \sum_{j=0}^{n-t} y_j^\tau y_{j+t}^\tau \, , \; n \gg t, \tag{A15}$$

if we are dealing with a stationary and a long time series. Then time $\tau = \tau_c$ can be considered, e.g., as the half-life of this function, i.e., $AC(\tau_c) \approx \frac{1}{2} AC(0)$. Other choices for τ can also

be considered [43]. Note that the time series $\{y_j^\tau\}$ can consist of $\tau > \tau_c$ with statistically independent elements.

With time series dependent on the time scale τ, we can build scale-dependent entropy $S^\tau(t)$ and the corresponding complexity $CX(S^\tau)$ by the corresponding methods presented in Appendix D.1. Thanks to this, for each scale τ separately (i.e., for each n separately), we can find quantities such as $n^{CX^{max}}$, $n^{CX^{max}}/N$ and CX^{max}.

Appendix D.3. Elements of Deterministic Chaos: A Cyclically Kicked Damped Rotor

Deterministic chaos can be an example of the complexity of single-particle motion (i.e., the complexity of its phase space). It is caused by instability due to initial conditions. A typical example of this is a cyclically kicked damped rotor [20] or a damped pendulum with cyclic driving force [19]. Here we sketch this example.

The starting point is Newton's equation for a rotor motion in the presence of viscous drag, which takes the form

$$\frac{d^2\phi(t)}{dt^2} = -\gamma \frac{d\phi(t)}{dt}, \tag{A16}$$

where $\phi(t)$ is a time-dependent rotation angle of the rotor, γ is the viscous drag coefficient, and the moment of inertia is equal to the unit. The exact solution of Equation (A16) is based on the time-dependent exponential function.

Next, we enrich this equation with a non-linear impulse forcing force,

$$F = \kappa f(\phi) \sum_{n=0}^{\infty} \delta(t - nT), \tag{A17}$$

where $f(\phi)$ is a non-linear function of ϕ and κ is its amplitude, while T is the period of this force. Hence, in the stroboscopic variables (or in the Poincaré representation) we obtain the Poincaré map,

$$\omega_{n+1} = \exp(-\gamma T)[\omega_n + \kappa f(\phi_n)], \tag{A18}$$

$$\phi_{n+1} = \phi_n + \frac{1 - \exp(-\gamma T)}{\gamma}[\omega_n + \kappa f(\phi_n)], \tag{A19}$$

where $\omega = \frac{d\phi}{dt}$. The above set of recursive equations allows us to examine on the Poincaré surface both dissipative deterministic chaos (e.g., logistic or Henon mapping) and conservative (i.e., Chirikov mapping). This depends on the values of γ and κ parameters and the form of the function f.

Belief in the complexity of the phase space $\{\omega, \phi\}$ of the system presented above is common. The complexity requires a knowledge of entropy. The total entropy of the system is the entropy of its long phase trajectory. However, we consider the complexity of the phase space structure mapped to the Poincaré surface, i.e., based on narrow entropy describing only this mapped structure. Obviously, this entropy depends on initial conditions and parameters assumed. Constructing entropy first requires defining a long time series. We have this time series as one consisting of N two-component elements $\{\omega_n, \phi_n\}_{n=0}^{N-1}$. Indeed, the approach presented in Appendix D.1 can be used to calculate entropy.

The goal here is to have the entropy value S_{CX}^{max} at which the complexity measure reaches the maximum value CX^{max}. It comes down to finding the dimension of the embedded subspace l_{CX}^{max} for which the entropy $S = S_{CX}^{max}$. It is l that is the optimized parameter (in Section 3 it was the dynamic variable N_L). It can be expected that l_{CX}^{max} is neither too small nor too large compared with the dimension N of the base space.

Appendix D.4. Elements of Anomalous Diffusion

Let us consider as a typical example the deterministic dissipative motion of an extremely massive single molecule in a viscous fluid. This simple motion is described by the Newton's dynamics, i.e., by the linear ordinary differential equation,

$$\frac{du(t)}{dt} = -\gamma\, u(t), \tag{A20}$$

where $u(t)$ is a time-dependent molecule velocity and γ is a constant friction coefficient. The exact solution of Equation (A20) is given by the time-dependent exponential function.

We are now moving from a deterministic level to a stochastic level by a relatively small extension of Equation (A20). That is, we extend Equation (A20) so as to obtain the retarded general Langevin equation (GLE) [50],

$$\frac{du(t)}{dt} = -\int_0^t \gamma(t-t')u(t')dt' + \frac{R(t)}{m}, \tag{A21}$$

where $\gamma(t)$ is the retarded friction coefficient, $R(t)$ is a random force of a thermal origin, i.e., caused by the random action of fluid particles, and m is the mass of the molecule. Notably, the retardation is present only when the mass of the molecule suspended in the fluid is not too heavy.

Equation (A21), which is the essence of the Ornstein-Uhlenbeck (OU) generalized theory of Brownian motion, takes into account both the feedback effect associated with the reverse fluid flow pushed by the molecule and the erratic nature of the molecule's motion. Although it is still a linear equation relative to u, the velocity the autocorrelation function, $C(t)$, is no longer expressed by a simple exponential, but exhibits a slower, power-law decay,

$$C(t) \propto \frac{1}{t^{2-\alpha}},\ \alpha = \frac{1}{2}, \tag{A22}$$

for a long time, which is indeed more realistic. Equation (A22) is a central result of OU generalized theory. As one can see, a relatively small extension of Equation (A20)—small because it still leaves this equation in the domain of linear equations relative to u—led to decay according the power law. Such a law is an essential attribute of a complex system. This algebraic fat tail was noticed for the first time by Adler and Wainwright in molecular-dynamic simulation of hard spheres' fluid [53], at least at intermediate fluid densities. Equation (A22) is the result of a cooperative phenomenon in the form of a positive feedback arising in the system between the molecule suspended in the fluid and the particles of the fluid. The non-linear nature of this coupling in time (that is the nonlinear dependence of γ on time) is contained in the integral kernel of Equation (A21). There is a wide class of physical problems that can be modelled using this equation [28,54] (for almost arbitrary α). Equation (A21) is the first level of complexity here.

Note that the direct consequence of Equation (A22) is the sub-diffusive behaviour of the suspended molecule, i.e., for long times the variance of the stochastic process $X(t)$ (without a drift) takes the form,

$$\langle X(t)^2 \rangle \propto t^{\alpha}. \tag{A23}$$

Now one can ask a question about what the distribution family with the variance given by the above formula looks like. The answer is almost instant namely,

$$P(X,t) = \frac{1}{t^{\alpha/2}} f\left(\frac{|\,X\,|}{t^{\alpha/2}}\right), \tag{A24}$$

which is a positive and normalized time-dependent probability distribution, where scaling function f is almost an arbitrary one.

Results Euqation (A23) and (A24) became the inspiration for extensive research on anomalous diffusion [28]. For example, by taking restrictions on the shape exponent α, i.e., α can be generally different from 1/2. This type of extension permits realistic considerations [28].

Characteristic Examples

Here we show how comprehensive the probability distribution given by Equation (A24) is.

Example A1. *Brownian and Gaussian random walk.*

Let us assume for $\alpha = 1$,

$$f\left(\frac{|X|}{t^{\alpha/2}}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{X^2}{2t}\right). \tag{A25}$$

As one can see, the variance of the distribution given by Equation (A24) with the substitution given by Equation (A25) satisfies Equation (A23)—it is a linear function of time, as it should be. This example is our reference case.

Example A2. *Brownian and non-Gaussian random walk.*

Let us now assume that still $\alpha = 1$ but

$$f\left(\frac{|X|}{t^{\alpha/2}}\right) = \frac{1}{2} \exp\left(-\frac{|X|}{t^{1/2}}\right), \tag{A26}$$

that is, probability distribution P is a Laplace distribution. Despite this, the variance of this distribution remains a linear function of time. Here we are dealing here with a Brownian and Laplace random walk.

Example A3. *Non-Brownian and Gaussian random walk.*

Let us here assume that

$$f\left(\frac{|X|}{t^{\alpha/2}}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{X^2}{2t^{\alpha}}\right), \tag{A27}$$

where α is an arbitrary shape exponent. As one can see, the variance is here (in general) a non-linear function of time: it can be both sub- and super-linear. Thus, we are dealing here with the fractional Brownian motion.

Example A4. *Non-Brownian and Non-Gaussian random walk.*

Let us consider the Weibull distribution—this gives f function in the form,

$$f\left(\frac{X}{t^{\alpha/2}}\right) = \kappa \left(\frac{X}{t^{\alpha/2}}\right)^{\kappa-1} \exp\left(-\left(\frac{X}{t^{\alpha/2}}\right)^{\kappa}\right), \; X \geq 0, \tag{A28}$$

where $\kappa > 1$. The variance for Weibull distribution is given by Equation (A23). This means that we can model both sub-diffusion (when $\alpha < 1$) and super-diffusion (when $\alpha > 1$) using the Weibull distribution. Of course, with this distribution one can also model a Brownian (when $\alpha = 1$) but non-Gaussian random walk.

Appendix D.5. Nonequilibrium Configuration Entropy

We are now constructing configuration entropy (i.e., based on configuration rather than phase space). We first discretize the space, i.e., the X axis. Let the size of the discretization step be ΔX. The time-dependent probability of finding a stochastic process at time t in the nth cell with size ΔX is,

$$p_n(\Delta X, t) = \int_{n\Delta X}^{(n+1)\Delta X} P(X,t)dX, \tag{A29}$$

where free parameter ΔX here defines the level of coarse-graining. Indeed, we use this probability in Equations (A6) or (A7). The equilibrium probability needed there $p_n^{eq}(\Delta X) = p_n(\Delta X, t \to \infty)$ constructs in the usual way. That is, we confine the system to a very large but finite size. Then time t is going to infinity. Next, both probabilities (the nonequilibrium for finite time t and equilibrium probability) substitute to Equations (A6) or Equation (A7). Finally, the system goes to the thermodynamic limit.

The goal is to find ΔX size equal to ΔX_{CX}^{max} that maximizes CX, that is CX^{max}. This can only be done by numerical means.

We emphasize that Equation (A21) is the dynamic basis of the (stochastic) Fokker–Planck equation [45] as well as the Langevin [55] fractal equation and hence the Fokker-Planck [28] fractal equation. The same Equation (A21) is a springboard to move to higher levels of complexity.

Appendix E. Dynamic Self-Organisation on a Complex Network—A Sketch

It is hard to imagine not using complex network technologies to analyze collective processes in the socio-economic world. One of the fascinating sources of complexity is its capacity for the self-organization—spontaneous (of endogenic) character or stimulated (of exogenic) character. We consider the dynamic phase transition here, in which the network of the stock market companies evolves towards a characteristic big-star structure—it condensates [6,32].

Appendix E.1. Minimal Spanning Tree of the Frankfurt Stock Exchange

Here we consider the canonical Minimal Spanning Tree (MST) of the Frankfurt Stock Exchange (FSE) companies belonging to the widely-exploited class of correlation complex networks (the content of this section was created with the participation of Mateusz Wiliński). They describe well the dynamics of relationships between companies, that is the evolution of the MST. The MST is a very simple type of a complex network that, although it resets the clustering rate, provides a lot of important information about the structure and dynamics of real networks.

For example, in Figure A1, we present the self-organized structure of the Frankfurt Stock Exchange as of 29 January 2007. As one can see, the prominent star centred at SALZGITTER (SZG) AG-Stahl und Technologie company (the network node marked by the red circle located at the centre of the network) dominates the stock market structure. The middle plot on the top and the plot on the right there correctly reflect this type of behaviour. The former plot clearly shows the local minimum determining the size of the average coordination zone of the SZG node. Its average radius (i.e., the MOL equals about 2.5 separation steps) confirms that the SZG company is the centre of the star. The last plot shows a red circle with a multiplicity of 1 and a degree close to 100, which represents the SZG node. This node is an outlier—an excellent example of the appearance of a super-extreme event, i.e., dragon-king event [29–31].

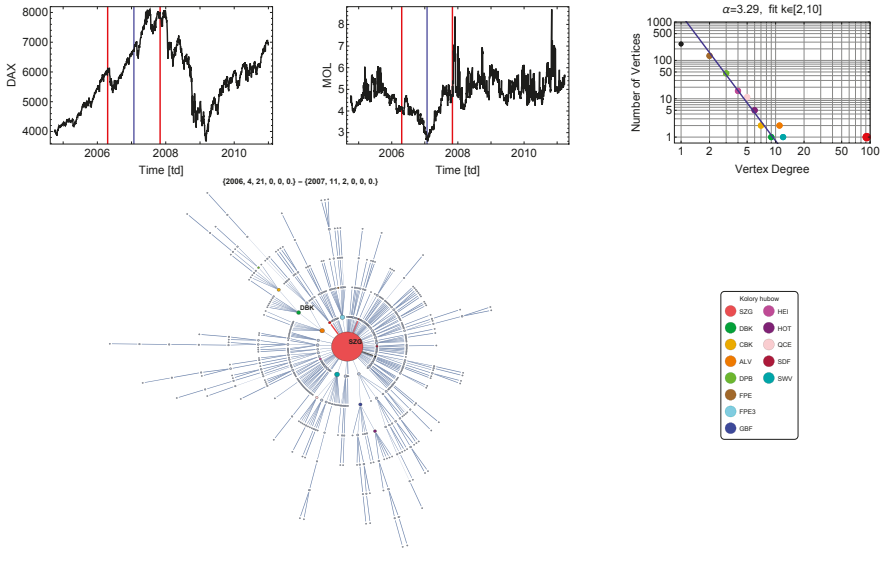


Figure A1. A snapshot picture of Monday 29 January 2007. The plot on the top left shows the DAX index in the years 2005–2010. This index is the emanation of the entire stock exchange. Of course, a computer analyzes quotations of all companies on this stock exchange that are in the range mentioned above. Two vertical red straight lines mark the scanning window, and its centre is marked by a blue line. The current window location, i.e., the beginning and end, are given dates just below the DAX and MOL plots. The same scanning window applied to the MOL indicator. The middle plot at the top shows the course of the widely used mean-occupation layer (MOL) [32]. The upper plot on the right shows the distribution of degrees of the vertices of the network shown below. The red circle on this plot, which stands out significantly from the power-law distribution, concerns the centre of the star (the most prominent red circle). The coloured circles in the legend and the abbreviations of the names mean the companies presented in the graph. The figure was taken from our publication [32].

The structure dynamics of the dragon-king is presented in Figure A2. That is, we presented there the dependence of the degree of this node on time, yielding a very meaningful λ peak characterizing here a structural condensate (for details see [32]).

In Figure A3 we compare empirical degree entropies $S(t) = -\sum_k P(k) \ln P(k)$, where $P(k)$ defines the empirical degree distribution, versus time in the presence and absence of SZG company on the network. In Figure A1, the upper plot on the right shows an example of the (non-normalized) distribution, $P(k)$, where k means the degree of a node. Various nodes are marked here with coloured circles defined in the legend. As one can see, the presence of the SZG node significantly changes the structure of the MST network. It is worth determining the complexity of such a network.

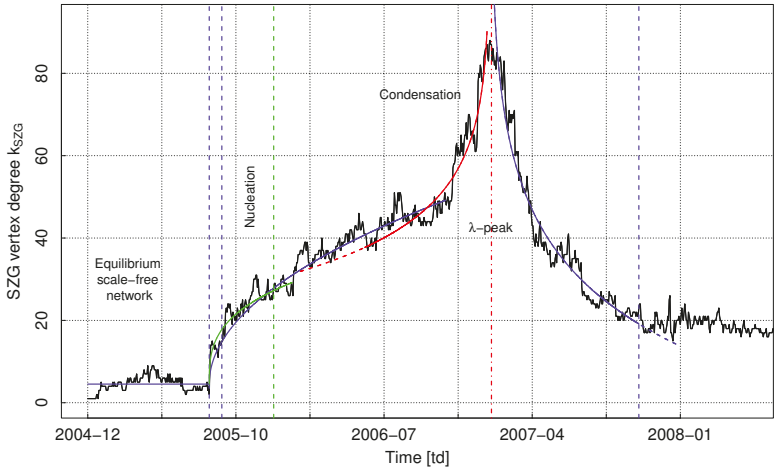


Figure A2. The temporal SZG vertex degree, k_{SZG}, vs. time. It forms the so-called λ-peak marked with a red vertical dashed straight line. It shows temporary edge condensation on the SZG node. The span of this peak is marked by the last blue vertical dashed straight lines. These lines are also plotted in Figure A3 below. The equilibrium scale-free networks are placed outside this area. The centre of this peak has been extrapolated on Thursday 25 January 2007—a deviation of two trading days from the result shown in Figure A1 is irrelevant here. The plot was taken from the publication [32] with the consent of the editors.

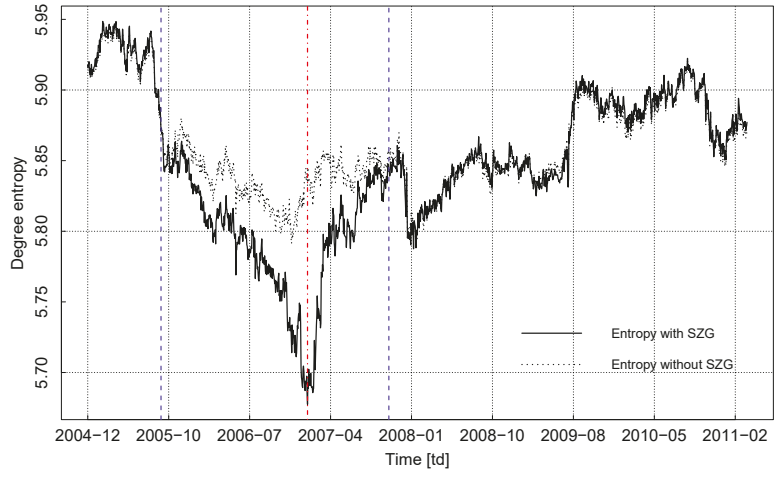


Figure A3. Degree entropy (defined in [32] by degree distributions) vs. time. The solid line marked entropy in the presence of the SZG vertex but dotted after removing this node. One can see the crucial role of this node in preparing the temporary network structure. The well-defined local absolute minimum of degree entropy is placed on Thursday, 25 January 2007. The plot was taken from the publication [32] with the consent of the editors.

Appendix E.2. Analysis for Case $Z < 1$

We note that the extreme entropy values, for the data shown in Figure A3, are as follows: $S^{min}(t = \text{2007-01-25}) = 5.690$ and $S^{max}(t = \text{2005-03-15}) = 5.948$, which gives a span of $Z = 0.258 < 1$.

Substituting these numbers into Equations (A4) and (A5) we get the entropy of the order of $(m = 2, n = 2)$,

$$S_{CX}^{max} = 5.819, \tag{A30}$$

and complexity measure of the same order

$$CX^{max} = 2.769\ 10^{-4}. \tag{A31}$$

From the plot data: 'Entropy with SZG 'in Figure A3, dates corresponding to the entropy given by Equation (A30) can be read. There are many of them—we chose three particular ones: 6 January 2006, 21 December 2007, 15 April 2008. All of them concern the most complex networks.

Figure A4 shows the network corresponding to the second date. One can see how much it differs from the least complex shown in Figure A1. This Figure shows a mixture of few mini stars with a degree not greater than about 20 (but no central one with degree almost 100, as shown in Figure A1). It is indicated by the degree distribution on the right-hand side of the plot, besides several skeletons and developed cascades arranged there in a somewhat disordered way. It is a richer and more disordered structure—more complex than that shown in Figure A1.

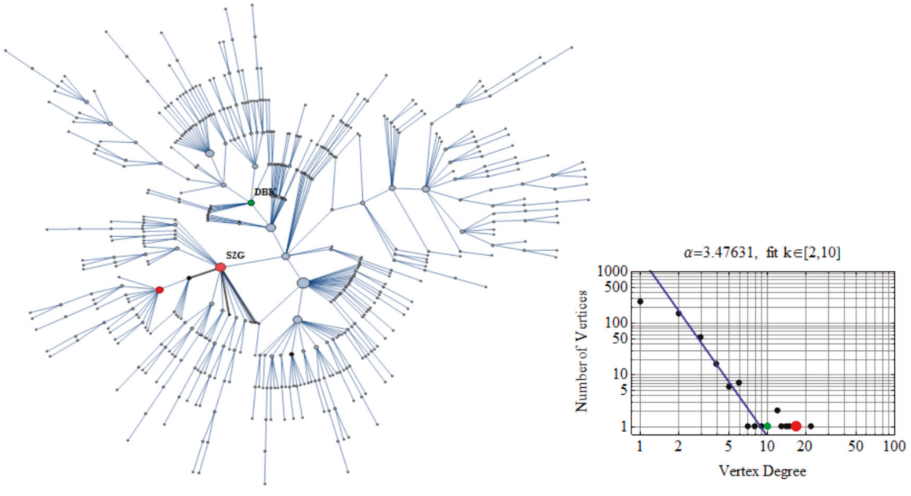


Figure A4. The snapshot picture of the Frankfurt Stock Exchange complex network on 21 December 2007. This network represents the most sophisticated state of the stock market, that is, CX^{max} given by Equation (A31). One can see how much this network differs from that shown in Figure A1—e.g., the dominant central star has ceased to exist. The network structure is now skeletal.

Author Contributions: R.K. and Z.R.S. conceptualised the work; R.K. wrote the draft and conducted the formal analysis and prepared the draft of figures; J.K. provided numerical calculations and provided the final figures; Z.R.S. finalised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: One of the authors of the work (Z.R.S.) partially benefited from the financial support of the ZIP Program. This program of integrated development activities of the University of Warsaw is implemented under the operational program Knowledge Education Development, priority axis III. Higher education for economy and development, action: 3.5 Comprehensive university programs, from 2 April 2018 to 31 March 2022, based on the contract signed between the University of Warsaw and the National Center for Research and Development. The program is co-financed by the European Union from the European Social Fund; http://zip.uw.edu.pl/node/192. The author (Z.R.S.) is also partially supported by the Center for Study of Systemic Risk, Faculty of "Artes Liberales" University of Warsaw; http://al.uw.edu.pl/. Apart from that, there was no other financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nicolis, G.; Nicolis, C. *Foundations of Complex Systems. Emergence, Information and Prediction*, 2nd ed.; World Science Publication: Singapore, 2012; ISBN 981-4366-60-9.
2. Kwapień, J.; Drożdż, S. Physical approach to complex systems. *Phys. Rep.* **2012**, *515*, 115–226. [CrossRef]
3. Dorogovtsev, S.N.; Goltsev, A.V. Critical phenomena in complex networks. *Rev. Mod. Phys.* **2008**, *80*, 1275–1335. [CrossRef]
4. Albert, R.; Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **2002**, *74*, 47–97. [CrossRef]
5. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [CrossRef] [PubMed]
6. Dorogovtsev, S.N. *Lectures on Complex Networks*; Clarendon Press: Oxford, UK, 2010.
7. Grassberger, P.; Procaccia, I. On the characterization of strange attractors. *Phys. Rev. Lett.* **1983**, *50*, 346–349. [CrossRef]
8. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, 278, 2039–2049. [CrossRef]
9. Prehl, J.; Boldt, F.; Essex, C.H.; Hoffmann, K.H. Time evolution of relative entropies for anomalous diffusion. *Entropy* **2013**, *15*, 2989–3006. [CrossRef]
10. Thurner, S.; Hanel, R.; Klimek, P. *Introduction to the Theory of Complex Systems*; Oxford Univ. Press: Oxford, UK, 2018; ISBN 9780198821939.
11. Popiel, N.J.M.; Khajehabdollahi, S.; Abeyasinghe, P.M.; Riganello, F.; Nichols, E.; Owen, A.M.; Soddu, A. The Emergence of Integrated Information, Complexity, and 'Consciousness' at Criticality. *Entropy* **2020**, *22*, 339. [CrossRef]
12. Available online: https://en.wikipedia.org/wiki/Complexity (accessed on 5 August 2020).
13. Thurner, S.; Hanel, R.; Gell-Mann, M. How multiplicity of random processes determines entropy and the derivation of the maximum entropy principle for complex systems. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6905–6910.
14. Grassbereger, P. Toward a quantitative theory of self-generated complexity. *Int. J. Theor. Phys.* **1986**, *25*, 907–938. [CrossRef]
15. Bialek, W.; Nemenmana, I.; Tishby, N. Complexity through nonextensivity. *Phys. A* **2001**, *302*, 89–99. [CrossRef]
16. Prokopenko, M.; Boschetti, F.; Ryan, A.J. An information-theoretic primer on complexity, self-organization, and emergence. *Complexity* **2008**, *15*, 11–28. [CrossRef]
17. Borda, M. *Fundamentals in Information Theory and Coding*; Springer: Berlin/Heidelberg, Germany, 2011.
18. Crutchfield, J.P. Between order and chaos. *Nat. Phys.* **2012**, *8*, 17–24. [CrossRef]
19. Baker, L.G.; Gollub, J.P. *Chaotic Dynamics: An Introduction*, 2nd ed.; Cambridge University Press: Cambridge, UK, 1996; Section 3.
20. Schuster, H.G. *Deterministic Chaos. An Introduction*, 2nd ed.; VCH: Verlagsgesellschaft, Germany, 1988.
21. Henkel, M.; Hinrichsen, H.; Lübeck, S. *Non-Equilibrium Phase Transitions. Volume I: Absorbing Phase Transitions*; Springer: Berlin/Heidelberg, Germany, 2008; Section 4.1.7, pp. 112–116.
22. Wolfram, S. *A New Kind of Science*; Wolfram Media Inc.: Champaign, IL, USA, 2002.
23. Gell-Mann, M. Plectics: The study of simplicity and complexity. *Europhysicsnews* **2002**, *1*, 17–20. [CrossRef]
24. Gell-Mann, M.; Lloyd, S. Information measures, effective complexity and total information. *Complexity* **1996**, *2*, 44–52. [CrossRef]
25. Gell-Mann, M.; Lloyd, S. Effective complexity. In *Nonextensive Entropy: Interdisciplinary Applications*; Gell-Mann, M., Tsallis, C., Eds.; Oxford University Press: New York, NY, USA, 2004; pp. 387–398, ISBN 0-19-515976-4.
26. Ay, N.; Müller, M.; Szkoła, A. Effective complexity and its relation to logical depth. *IEEE Trans. Inf. Theory* **2010**, *56*, 4593–4607. [CrossRef]
27. Gell-Mann, M. *The Quark and the Jaguar: Adventures in the Simple and the Complex*, 8th ed.; Freeman, W.H., Ed.; W.H. Freeman: New York, NY, USA, 1994; ISBN 0-7167-2725-0.
28. Kutner, R.; Masoliver, J. The continuous time random walk still trendy: fifty-year history, state of art, and outlook. *Eur. Phys. J.* **2017**, *90*, 50; doi:10.1140/epjb/e2016-70578-3. [CrossRef]

29. Sornette, D. Dragon-kings, black swans and the prediction of crises. *Int. J. Terraspace Sci. Eng.* **2009**, *1*, 1–17. [CrossRef]
30. Sornette, D.; Ouillon, G. Dragon-kings: Mechanisms, statistical methods and empirical evidence. *Eur. Phys. J. Spec. Top.* **2012**, *205*, 1–26. [CrossRef]
31. Wiliński, M.; Sienkiewicz, A.; Gubiec, T.; Kutner, R.; Struzik, Z.R. Structural and topological phase transition on the german stock exchange. *Phys. A* **2013**, *392*, 5963–5973. [CrossRef]
32. Wiliński, M.; Szewczak, B.; Gubiec, T.; Kutner, R.; Struzik, Z.R. Temporal condensation and dynamic λ-transition within the complex network: an application to real-life market evolution. *Eur. Phys. J. B* **2015**, *34*, 1–15.
33. Kozłowska, M.; Denys, M.; Wiliński, M.; Link, G.; Gubiec, T.; Werner, T.R.; Kutner, R.; Struzik, Z.R. Dynamic bifurcations on financial markets. *Chaos Solitons Fractals* **2016**, *88*, 126–142. [CrossRef]
34. Jakimowicz, A. The role of entropy in the development of economics. *Entropy* **2020**, *22*, 452. [CrossRef]
35. Jakimowicz, A. Fundamental sources of economic complexity. *Int. J. Nonlinear Sci. Numer.* **2016**, *17*, 1–13. [CrossRef]
36. Rossler, J.B., Jr. Econophysics and economic complexity. *Adv. Complex Syst.* **2008**, *11*, 745–760. [CrossRef]
37. Rossler, J.B., Jr. Entropy and econophysics. *Eur. Phys. J. Spec. Top.* **2016**, *225*, 3091–3104. [CrossRef]
38. Zambelli, S.; George, D.A.R. *Nonlinearity, Complexity, and Randomness in Economics: Towards Algorithmic Foundations for Economic*; Wiley–Blackwell: Chichester, UK, 2012; ISBN 978-1-4443-5031-9.
39. Kutner, R.; Ausloos, M.; Grech, D.; Di Matteo, T.; Schinckus, C.H.; Stanley, H.E. Econophysics and sociophysics: Their milestones & challenges. *Physica A* **2019**, *516*, 240–253.
40. Gell-Mann, M. What is Complexity? *Complexity* **1995**, *1*, 16–19. [CrossRef]
41. Bertin, E. *A Concise Introduction to the Statistical Physics of Complex Systems*; Springer Briefs in Complexity; Springer: Berlin/Heidelberg, Germany, 2012; pp. 33–38, ISBN 978-3-642-23922-9.
42. Jaynes, E.T. Gibbs vs boltzmann entropies. *Am. J. Phys.* **1965**, *33*, 391–398. [CrossRef]
43. Beck, C.H.; Schlögl, F. *Thermodynamics of Chaotic Systems. An Introduction*; Cambridge Nonlinear Science Series 4; Cambridge University Press: Cambridge, UK, 1995; pp. 50–55.
44. Tsallis, C. Possible generalization of boltzmann-gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
45. Van Kampen, G. *Stochastic Processes in Physics and Chemistry*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2007; pp. 111–114.
46. Bahcall, S. *Loonshots: How to Nurture the Crazy Ideas that Win Wars, Cure Diseases, and Transform Industries*; St. Martin's Press: New York, NY, USA, 2019; ISBN 978-1-250-18596-9.
47. West, G. *Scale: The Universal Laws of Life, Growth, and Death in Organisms, Cities, and Companies*; Penguin Books: London, UK, 2017; ISBN 978-1594205583
48. Dunbar, R.M. The social brain hypothesis. *Evol. Antropol.* **1998**, *6*, 178–190. [CrossRef]
49. Acharjee, S.; Bora, B.; Dunbar, R.I.M. On M-polynomials of dunbar graphs in social networks. *Symmetry* **2020**, *12*, 932. [CrossRef]
50. Kubo, R.; Toda, M.; Hashitsume, N. *Statistical Physics II. Nonequilibrium Statistical Mechanics*; Springer: Tokyo, Japan, 1985.
51. Li, P.; Liu, C.; Li, K.; Zheng, D.; Liu, C.; Hou, Y. Assessing the complexity of short-term heartbeat interval series by distribution entropy. *Med. Biol. Eng. Comput.* **2015**, *53*, 77–87. [CrossRef] [PubMed]
52. Sornette, D. *Critical Phenomena in Natural Sciences. Chaos, Fractals, Selforganization and Disorder: Concepts and Tools*; Springer: Berlin/Heidelberg, Germany, 2000; ISBN 3-540-67462-4.
53. Alder, B.J.; Wainwright, T.E. Decay of the velocity autocorrelation function. *Phys. Rev. A* **1970**, *1*, 18–21. [CrossRef]
54. Bunde, A.; Havlin, S. (Eds.) *Fractals and Disordered Systems*, Second Revised and Enlarged Edition; Springer: Berlin/Heidelberg, Germany, 1996.
55. Camargo, R.F.; Chiacchio, A.O.; Charnet, R.; de Oliveira, C.E. Solution of the fractional Langevin equation and the Mittag–Leffler functions. *J. Math. Phys.* **2009**, *50*, 063507. [CrossRef]
56. Wehrl, A. General properties of entropy. *Rev. Mod. Phys.* **1978**, *50*, 221–260. [CrossRef]

entropy

MDPI

Article

The Role of Entropy in the Development of Economics

Aleksander Jakimowicz

Department of World Economy, Institute of Economics, Polish Academy of Sciences, Palace of Culture and Science, 1 Defilad Sq., 00-901 Warsaw, Poland; ajakimowicz@inepan.waw.pl

Received: 25 February 2020; Accepted: 13 April 2020; Published: 16 April 2020

Abstract: The aim of this paper is to examine the role of thermodynamics, and in particular, entropy, for the development of economics within the last 150 years. The use of entropy has not only led to a significant increase in economic knowledge, but also to the emergence of such scientific disciplines as econophysics, complexity economics and quantum economics. Nowadays, an interesting phenomenon can be observed; namely, that rapid progress in economics is being made outside the mainstream. The first significant achievement was the emergence of entropy economics in the early 1970s, which introduced the second law of thermodynamics to considerations regarding production processes. In this way, not only was ecological economics born but also an entropy-based econometric approach developed. This paper shows that non-extensive cross-entropy econometrics is a valuable complement to traditional econometrics as it explains phenomena based on power-law probability distribution and enables econometric model estimation for non-ergodic ill-behaved (troublesome) inverse problems. Furthermore, the entropy economics has accelerated the emergence of modern econophysics and complexity economics. These new directions of research have led to many interesting discoveries that usually contradict the claims of conventional economics. Econophysics has questioned the efficient market hypothesis, while complexity economics has shown that markets and economies function best near the edge of chaos. Quantum economics has already appeared on the horizon, which recognizes money as a fundamental measurement device in the economy. The development of these sciences may indicate the need to reformulate all mainstream economics from its foundations.

Keywords: entropy economics; non-extensive cross-entropy econometrics; non-ergodic ill-behaved inverse problems; general system theory; econophysics; non-linear dynamics; complex adaptive systems; *homo oeconomicus*; edge of chaos; complexity economics

1. Introduction

Since its emergence, economics has been strongly methodologically linked to physics. These links made neoclassical economics possible. The aim of this paper is to review these relationships over the last 150 years and to point out a number of key discoveries that have resulted from the application of methods of physics in economics and that still remain outside mainstream economics. The probable reason that they are constantly ignored is that they undermine traditional economic knowledge.

The emergence of neoclassical economics can be dated to the first half of the nineteenth century. Its birth was influenced by economic issues related to various technical projects undertaken by French engineers who employed mathematical methods in order to solve these problems. The group of engineers included Jules Dupuit (1804–1866) and Charles Minard (1781–1870), but the achievements of the former are most significant [1,2]. The economic knowledge gained in this way was supplemented in the 1870s by Leon Walras, William Stanley Jevons and Carl Menger, but it was Alfred Marshall who gave it a coherent form [3]. In an effort to raise the scientific status of economics, neo-classicists decided to transfer to it the ideas and the mathematical apparatus from the leading science of that time—energy physics—which became the nucleus of later thermodynamics. The basic concepts of physics of the

mid-19th century were translated into economic language by Irving Fisher in 1892: material points (particles) became economic entities (individuals), force was replaced by marginal utility, and energy became equivalent to utility [4]. After this, the law of equilibrium could be transferred from physics to economics. In physics, an equilibrium point is determined by the maximum of the function of net energy, while the position of equilibrium in economics is determined by the maximum of the function of gain. As a result, neoclassical economics, which is still taught today, was created, but its methodological basis has been forgotten. This basis referred to thinking about markets and economies as closed systems striving to achieve a state of equilibrium.

There is no doubt that thermodynamics significantly contributed to the emergence of neoclassical economics and that in the early twentieth century it was a huge advance in science. Since then, however, economics and physics have gradually started to move away from each other. The emergence of the global financial crisis has made it clear that today we again need an economics based on physical methods, but this can no longer be 19th-century physics—it has to be replaced by 21st-century physics. This is proven by the results obtained by economics and econometrics of entropy, econophysics, complexity economics and, more recently, by quantum economics. Upgrading economic knowledge is not expensive; it is enough to abandon old, untrue dogmas.

The main aim of the paper is to examine the broadly understood impact of thermodynamics and entropy on the development of economics, which indicates the need to include in the research various types of entropy. The starting point for the analysis is entropy as a physical phenomenon, described by Rudolf Clausius, reflected in the second law of thermodynamics, which initiated the emergence of ecological economics. Next, various modifications of the entropy concept are taken into consideration. In recent decades, the science of entropy has been developing very fast. This has brought benefits to economics, in which the use of such forms of entropy as Shannon informational entropy or non-extensive Tsallis entropy has become a factor of progress. The analogies and metaphors combining entropy investigated in natural sciences with similar phenomena occurring in economic systems are also significant for the development of economic sciences [5]. However, despite the great importance of such analogies and metaphors, since the paper focuses mainly on similarities following the isomorphism principle, attempts have been made to limit the transfer of entropy formulae from physics to economics to logical homologies.

2. A Brief History of the Emergence and Development of the Entropy Concept

The emergence of entropy in physics was caused by an observation that in steam engines a large part of energy was lost due to friction and dissipation and, therefore, could not be converted into useful work. The research on this missing energy was conducted by Rudolf Clausius, who used the term entropy to describe it. In 1854, he presented the first mathematical definition of entropy in the following form [6] (p. 126):

$$S = \frac{1}{T}Q \tag{1}$$

or

$$\Delta S = \left(\frac{1}{T_2} - \frac{1}{T_1}\right)Q, \tag{2}$$

where ΔS stands for changes in entropy, while Q is the quantity of heat passing from the body with temperature T_1 to another body with temperature T_2. Clausius is also known for his concise and beautiful presentation of the first and the second law of thermodynamics [6] (p. 365):

1. *The energy of the universe is constant.*
2. *The entropy of the universe tends to a maximum.*

According to his other statement, the second law of thermodynamics takes the following form: *heat can never pass from a colder to a warmer body without some other change, connected therewith, occurring at the same time* [6] (p. 117).

A slightly different definition of entropy, being a measure of the molecular disorder of the system, was formulated by Boltzmann. It has the following form [7] (p. 137):

$$S = k_B \ln W, \tag{3}$$

where k_B is the Boltzmann constant, while W is the total number of microscopic states, corresponding to the macroscopic state of the system. Boltzmann entropy provides the basis for statistical mechanics, but the concept of probability, which is crucial in statistical theory, does not result directly from it. Thermodynamic probability W in formula (3) is not an ordinary probability, but an integer. However, it is possible to modify Boltzmann entropy to enable the introduction of the notion of probability distribution in Boltzmann statistics [8]. If an isolated system, consisting of N molecules belonging to n energy states is examined, then – assuming a fixed number of molecules and fixed values of the total energy – the total number of microscopic states of the system is given by the formula:

$$W = \frac{N!}{\prod_{i=1}^{n} N_i!}\,. \tag{4}$$

Substituting (4) to (3), Boltzmann entropy is obtained in the following form:

$$S = k_B \ln \frac{N!}{\prod_{i=1}^{n} N_i!} \approx -k_B N \sum_{i=1}^{n} p_i \ln p_i\,, \tag{5}$$

where $p_i = \frac{N_i}{N}$ and for a large N means the probability that the molecule is in the i-th energy state. Therefore, we obtain the entropy equation for the system consisting of N molecules distributed with probability distribution $p = (p_1, p_2, \ldots, p_n)$ among n energy states.

In the mid-20th century, the concept of entropy found its application in information theory. As it turned out, thermodynamic entropy is very similar to information entropy. While the former concerns energy loss, the latter concerns data loss in information transmission systems. In 1948, Claude E. Shannon published his groundbreaking work, *A Mathematical Theory of Communication*, in which he addressed the issues of measures of information, choice, and uncertainty. In this work, he formulated the following function [9]:

$$H = -K \sum_{i=1}^{n} p_i \log p_i\,, \tag{6}$$

where H stands for the Shannon entropy, whereas K is a positive constant, the role of which comes down only to a choice of a unit of measure. As can be easily observed, the H function represents entropy known in statistical mechanics, where p_i denotes the probability of a system being in cell i of its phase space. Equation (5) therefore presents Shannon entropy, which measures uncertainty related to a probability distribution $(p_1, p_2, \ldots, p_n)$. Thus, for large classical systems, Boltzmann entropy is proportional to Shannon entropy. At the same time, it should be emphasized that various types of entropy represented by Equations (1), (3), (5) and (6) are logical homologies since they present the same formal (mathematical) structure at various levels of reality. In particular, Shannon informational entropy can be treated as a more general concept compared to statistical thermodynamic entropy. In other words, this latter concept proves to be a special case of the former. In this way, the prediction of equilibrium thermodynamic properties can provide a form of statistical inference based on Shannon entropy as an information measure, while probabilities are interpreted in a subjective manner. Therefore, a reinterpretation of statistical mechanics in which statistical mechanics is based on information theory is possible. The entropy formula has a much deeper meaning than was initially believed, as it is completely independent of thermodynamics. Consequently, entropy can become a starting point for reflections, and the probability distribution maximising entropy can be used for statistical inference [10,11].

The idea of introducing the concept of entropy to information theory was put forward by John von Neumann. Shannon wondered for a long time how to name the measure of missing information he formulated (6). During a discussion, von Neumann gave him the following advice: *You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage* [12] (p. 180). According to another version of this anecdote, von Neumann said: *Why don't you call it entropy? In the first place, a mathematical development very much like yours already exists in Boltzmann's statistical mechanics, and in the second place, no one understands entropy very well, so in any discussion you will be in a position of advantage!* [13] (p. 81).

3. The Birth of Entropy Economics

Entropy first appeared in the economic sciences in 1971, when the American economist of Romanian descent, Nicholas Georgescu-Roegen, published his opus magnum entitled *The Entropy Law and Economic Process*. This created the foundations of a novel approach to the theory of production, which consisted of applying the second law of thermodynamics to economic considerations. This gave rise to ecological economics, also called bioeconomics. Georgescu-Roegen believes that—contrary to a widely held belief—thermodynamics did not originate from striving to explain physical phenomena based on heat transfer, but from efforts aimed at understanding phenomena based on pure economics [14]. For this reason, he called thermodynamics the physics of economic values. He regards the law of entropy expressed by the second law of thermodynamics as the most economic of all the laws of physics. Economic processes (production) turn the low entropy of the original goods and services into the high entropy of the final goods and services. It is a very convincing explanation of the fact that low entropy is responsible for the utility of a given good. Therefore, only thermodynamics can explain why goods have economic value.

The constantly dwindling resources of low entropy in the environment of man are the main reason for the scarcity of goods. Production processes are characterised by the fact that they reduce the resources of low entropy, so the principal feature of economic phenomena is their irreversibility. This leads one to the conclusion that—contrary to what conventional economics holds—economic flows do not create a circular flow of income, but rather they are one-directional. High entropy can be emitted to the environment both by natural physical processes and by economic processes. The latter are characterised by the fact that they result from human purposeful actions and they are ultimately proven right by joy and satisfaction with one's life. The economic value originates from the value that life presents to everyone. In this way, one can explain not only why people engage themselves in production processes, but also the basic goal of their economic activity, which is the preservation of mankind.

Paul A. Samuelson, a Nobel laureate in economics, said that entropy economics has changed the perception of economic processes and the creator of its concept deserves to have his achievements recognised and propagated in scientific circles. He also called him "a scholar's scholar, an economist's economist" [15] (p. 125). The concepts developed by Georgescu-Roegen coincided with a well-known work entitled *The Limits to Growth: A report for the Club of Rome's Project on the Predicament of Mankind*, published in 1972 [16], i.e., at nearly the same time as the basic work on entropy economics. This coincidence of opinions bore some fruit in the form of cooperation between Georgescu-Roegen and the team of authors of *The Limits to Growth* [17]. The Club of Rome reports refer to the entropy paradigm as a major trend in economic research. Examples include *Money and Sustainability: The Missing Link. A Report from the Club of Rome–EU Chapter to Finance Watch and the World Business Academy*, published in 2012, which suggests that taking entropy into account in economic processes is a necessary condition for the success of future monetary reforms [18].

Entropy economics contributed considerably to the development of economics by emphasising the necessity of including ecological issues in the theory of economic growth. Another, rather unnoticed and unappreciated achievement of entropy economics is the fact that it created favourable conditions for

the development of such novel disciplines as econophysics and complexity economics [19]. However, its creator, Nicholas Georgescu-Roegen, was never awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, despite the efforts of his followers [20]. As it turned out later, this was caused mainly by the fact that he operated outside the neoclassical paradigm, which is linked strongly with monetary reductionism, and it assumes that accelerated consumption of natural resources does not pose a considerable threat to economic growth [21]. Monetary reductionism is a frequently criticised feature of neoclassical economics. Its main idea is to reduce multidimensional relations between the natural environment and production processes to one-dimensional monetary issues. It consists in reducing non-monetary phenomena, such as health, socio-cultural, physical or ecological processes or impacts, to their monetary equivalents [22–24]. In this way, such features of economic phenomena as irreversibility and complexity are disregarded [25]. However, recent years have seen great progress within entropy economics concerning the development of thermodynamic techniques of modelling economic phenomena, which complement the standard econometric methods.

4. Thermodynamic Entropy as a Metaphor in Organization and Management Sciences

Thermodynamic entropy can be a useful metaphor in social science as a general measure of the disorder of a system. It is usually associated with a key feature of the system, which, for certain reasons, is to some extent irretrievably lost and cannot be used for its development. The management and organizational sciences use the notion of corporate entropy, which should be understood as an irretrievable loss of productive energy that cannot be transformed into useful work for the corporation. The task of enterprise management is to coordinate and integrate human work, which saves time in the entire enterprise, so that it is not lost on economic worries, complaints, beefing or partial and thus ineffective reforms. Redirecting wasted energy to productive uses is a basic condition for converting corporate entropy into useful work [26]. On the other hand, DeMarco and Lister talk about a new law, called by them the second thermodynamic law of management, according to which entropy in an organisation always grows, just like thermodynamic entropy in the universe. In their opinion, this law should be understood as follows: *most elderly institutions are tighter and a lot less fun than sprightly young companies* [27] (p. 97). They define corporate entropy as levelness or sameness. The uniformity of attitudes, appearances, and thought processes in a corporation is perceived as entropy, since it suppresses productive energy during work. This is somewhat reminiscent of Clausius' reasoning, initially using the term of equivalence-value (Equations 1 and 2) to describe entropy. An increase in corporate entropy implies a decrease in the potential to generate energy or to perform work. The way to fight corporate entropy is as follows: *The most successful manager is the one who shakes up the local entropy to bring in the right people and let them be themselves, even though they may deviate from the corporate norm* [27] (p. 97). Social sciences also refer to similar forms of entropy, such as cultural entropy, education and school entropy, organizational entropy, as well as leadership entropy [28,29]. This last one, for instance, means less efficient and effective work and a decrease in productivity. All of them, just like corporate entropy, are based on metaphors.

5. Levels of Measurement as a Prerequisite for the Application of Thermodynamic Entropy in Economics

Generally speaking, the level of measurement of a variable describes a type of information contained in numbers assigned to the examined objects or subjects within a given variable. It specifies to what extent data characteristics can be mathematically modelled [30] (p. 851). Economic applications of various forms of entropy adopt Steven's classification system, consisting of four levels of measurement: nominal, ordinal, interval, and ratio [31–33]. Each of them contains various properties of numbers or symbols, such as relations and operations, which specify measurements as well as the set of permissible transformations. The type of scale is defined by the group of transformations whose performance does not change the scale form. Permissible transformations include only those that do not infringe the empirical information presented by the scale. Below is a brief description of each of the scales.

1. The nominal scale is used only to classify or categorize the value of variables. It permits any one-to-one substitution of the assigned numbers. Therefore, for nominal scales the only empirical operation is the determination of equality.
2. The ordinal scale is used to order or rank the value of variables. It can be transformed using any increasing monotonic function.
3. Using the interval scale, variables can be presented in a quantitative form and used to examine usual statistical measures. The zero point, in this case, is a matter of convention or convenience since the form of the scale does not change after adding a constant. This scale can be subjected to linear transformation. A numerical value on one scale can be transformed to the value on the other scale, by using the following equation: $y = ax + b$.
4. A prerequisite for the existence of the ratio scale is an indication of operations enabling the determination of the following four relations: equality, rank-order, equality of intervals, and equality of ratios. After establishing such a scale, its numerical values can be transformed only by multiplying each of the values by a constant. The existence of an absolute zero is always assumed. All types of statistical measures can be applied to ratio scales. Additionally, only those scales enable logarithmic transformations.

Since the ratio scales are most commonly found in physics, their importance in research on various types of entropy is high. On the other hand, in economics and in social sciences, all four Stevens' scale types are used, which means that entropy can be examined based on each of them. Consequently, it seems that the isomorphism principle operates here, to some extent, regardless of the assumed level of measurement of variables. Even the use of metaphors can be valuable. In establishing logical homologies, the choice of scale type is certainly important because it clarifies the considerations. As a consequence, the development of economics induced by applications of entropy has resulted in shifting the emphasis to more frequent applications of the ratio scales in this science. This explains the emergence of such scientific disciplines as econophysics and complexity economics. Thus, the separation of these transdisciplinary research fields from mainstream economics should be treated as a sign of certain cognitive conservatism. On the other hand, it must be admitted that it is not yet fully clear whether the entire economics must be transformed into econophysics or complexity economics.

Stevens used the doctrine of operationalism to develop a system for classifying the scales of measurement. Individual classes of measurement scales are determined by empirical operations required in the measurement processes and by formal (mathematical) properties of the scales. Statistical measures, which can be in a justified way applied to empirical data, depend on the scale type used for data arrangement. In the broadest sense, the measurement consists in assigning numerals to objects or events following specific rules. The concept of operationalism is attributed to Percy W. Bridgman, an American physicist and the winner of the Nobel Prize in Physics in 1946. Bridgman believed that it was possible to redefine and at the same time to specify some unobservable entities only when we could indicate physical and mental operations enabling their measurement. As an example, he provides the concept of length. In order to determine the length of any object, the performance of certain physical operations is necessary. To define the concept of length, it is necessary to indicate operations permitting length to be measured. In other words, the concept of length is explained by providing the set of operations, based on which the length is determined. Bridgman's conclusion is as follows [34] (p. 5)*: we mean by any concept nothing more than a set of operations; the concept is synonymous with the corresponding set of operations*. The physical concept is determined by actual physical operations, and the mental concept, such as mathematical continuity, is described by mental operations. From an economic point of view, his general comments on the operational point of view are extremely interesting [34] (pp. 31–32). Adoption of operational definitions implies a far-reaching change in all our thinking habits, in that we will not be able to use concepts that are not properly explained by operations as tools in our reasoning. Initially, operational thinking may become an unsocial virtue. Later on, it has a chance to reform both the social art of conversation, and all our social relations. This is particularly true – as Bridgman claims

– of popular contemporary discussions on religious or moral topics. Thus operational thinking also includes the use of different definitions of entropy in economics. Such research is expected to result in moving the issues under consideration from mainstream economics towards econophysics and complexity economics.

The application of Stevens' classification system in economics, and in particular the growing use of the ratio scales, leads to more frequent occurrences of operational definitions of entropy in this science. As a result of such activities, the centre of thought is shifted from mainstream economics to econophysics and complexity economics. At the same time, in natural sciences, the very concept of entropy as a measure of disorder and uncertainty is constantly changing and developing, finding its expression in the theory of deterministic chaos, the theory of stochastic processes and quantum mechanics. This often involves replacing the notion of entropy with other, similar terms, better adjusted to the context of the research. These research trends are gradually penetrating economics, which results in the need to move on to econophysics and complexity economics in the further part of the paper. This does not imply the tendency to omit various definitions of entropy in these approaches, but it results from the partial transformation of this notion into related concepts such as the edge of chaos or complexity.

6. Thermoeconomics and Operational Definitions of Entropy in Economics

As an effect of the operationalisation of the entropy concept in economics, resulting from the introduction of the ratio scales, various definitions of entropy can be the subject of reification and be applied to specific economic phenomena, such as monetary flow in the economy. In such a case, the starting point for the discussion is the circular flow of income, in which flows of goods and services as well as factors of production are compensated for with equivalent, but monetary flows in the opposite direction. However, this process cannot be continued indefinitely, because the economy would then be a perpetual motion machine, which violates the second law of thermodynamics. Wastes, which have to be disposed of outside the economic system, are inevitably generated. This must be compensated for by an influx of new resources from the environment. For this reason, the circular flow must be supplemented with the linear throughput of matter-energy, which has to supply the constant movement of money, goods and services, as well as factors of production. Linear throughput means the inflow to the economy of low-entropy natural capital, such as solar energy, mines, wells, fisheries, croplands and the outflow of high-entropy wastes, which are no longer economically valuable. An entropic flow is a part of production processes and causes that the matter and the energy participating in such processes to become less useful [35]. Therefore, economic entropy should be linked to the utility of goods and services and factors of production. This leads to the conclusion that entropy is an inherent feature of production processes, participates in the circular flow, and as such has a direct effect on real money supply (including the consequences of inflation). It must, therefore, be to some extent a monetary phenomenon itself.

The development of operational definitions of entropy was possible within a science referred to as thermoeconomics, which deals with the application of laws of thermodynamics to study economic phenomena. The processes of entropy generation in economic systems have been thoroughly analysed by John Bryant in his book *Entropy Man* [36]. The production of entropy in the economy is formulated in a way resembling Boltzmann entropy as follows:

$$S = \ln\left(\frac{V}{X}\right), \tag{7}$$

where V stands for the volume of economic activity, and X represents the level of the constraints for that activity. Entropy changes in the economy per unit of time can be presented using a more detailed formula [36] (Ch. 4):

$$dS = (\omega - \omega n + 1)\frac{dV}{V} = \left(1 + \frac{1-n}{r}\right)\frac{dV}{V}, \tag{8}$$

where $\frac{dV}{V}$ stands for the growth rate of volume flow, ω is the lifetime coefficient, n represents the elastic index, while r is the natural rate of return. Factor $(\omega - \omega n + 1)$ is referred to as the marginal entropic index. After the integration of Equation (8), entropy generation per unit of time takes the following form:

$$S = (\omega - \omega n + 1) \ln V + C = \left(1 + \frac{1-n}{r}\right) \ln V + C \,, \tag{9}$$

where C is a constant of integration. Economic entropy can be measured in units referred to as nats, similar to entropy in information theory, where logarithms to the base of the mathematical constant $e \approx 2.71828$ are also used.

The above entropy formulas can be used to measure money entropy. In such a case, the rate of return is approximately the long-term average or natural level of the velocity of money circulation. When defining money entropy in the economy, one of the key potential constraints that should be also considered is the level of interest rates. The demand for money is represented by a function of liquidity preference formulated by John M. Keynes in his famous book *The General Theory of Employment, Interest and Money* [37]. According to this function, the interest rate is the price of money, i.e., the price that has to be paid to get people to part with liquidity for a specific period of time. Interest can be seen as a form of value flow constraint. Thus, the change in money entropy in the economy takes the following form [36] (Ch. 6):

$$dS = \frac{dG}{G} - \frac{dI}{I} \,, \tag{10}$$

where $\frac{dG}{G}$ is the rate of change in output value flow G, and $\frac{dI}{I}$ represents the rate of change in the Index of Money Interest. In other words, the change in money entropy is the difference between the growth rate of output value and the prevailing interest rate.

Changes in the interest rate are not the only factor influencing the output value flow and entropy change, as employment factors are also of great importance. In this case, the entropy formula takes the form [36] (Ch. 7):

$$dS = (\omega - \omega n + 1)\frac{dv}{v} = \left(1 + \frac{1-n}{r}\right)\frac{dw}{w} \,, \tag{11}$$

where $\frac{dv}{v}$ stands for the rate of change in output volume flow per head, and $\frac{dw}{w}$ represents the rate of change in the wage rate. Taking unemployment into consideration, the analysis can also use the Phillips curve, which is well-known in economics [38,39]. In line with this relationship, a high unemployment rate is usually accompanied by low rates of wage growth and inflation, and when the unemployment rate is low, rates of wage growth and inflation tend to increase. Taking this into account, the following formula for generating entropy in the economy can be presented [36] (Ch. 7) as:

$$dS = \frac{dG}{G} - \frac{dJ}{J} \,, \tag{12}$$

where $\frac{dG}{G}$ represents the rate of change in output value attributed to the labour sector, while $\frac{dJ}{J}$ stands for the rate of change in lost value from unemployment. In other words, a change in entropy generated in the economy is equal to the difference between the output value attributed to labour that an economy can support and the potential output value flow loss taken out through unemployment, which the economy cannot support.

7. Non-Extensive Cross-Entropy Econometrics

Non-extensive cross-entropy econometrics (NCEE) is one of the most interesting trends in entropy economics as it is based on power-law probability distribution and it is a method of econometric model estimation, which is particularly useful for non-ergodic ill-behaved inverse problems [40,41]. It is based on non-extensive Tsallis entropy, which is a generalisation of the widely known Boltzmann–Gibbs

entropy. Tsallis entropy, S_q, is particularly useful in studying systems with strong correlations between various microstates [42–45]. It is noted as the following formula:

$$S_q = \frac{k}{q-1}\left(1 - \sum_{i=1}^{W} p_i^q\right), \quad q \in R, \tag{13}$$

where: k is a conventional positive constant, $W \in N$ is the total number of possible (microscopic) configurations, $\{p_i\}$ describes a discrete set of probabilities with the condition $\sum_{i=1}^{W} p_i = 1$, and q is any real number. The parameter q is often called the Tsallis index. It is a measure of the strength of the correlation between different microstates of a system and it can take values smaller or greater than 1. If this parameter approaches 1, then the Tsallis entropy reduces to the usual Boltzmann–Gibbs entropy (S_{BG}). We can write [44,45] it as:

$$p_i^q = p_i p_i^{q-1} = p_i e^{(q-1)\ln p_i} \sim p_i[1 + (q-1)\ln p_i], \tag{14}$$

hence,

$$S_1 \equiv \lim_{q\to 1} S_q = k\lim_{q\to 1} \frac{1 - \sum_{i=1}^{W} p_i \exp[(q-1)\ln p_i]}{q-1} = -k\sum_{i=1}^{W} p_i \ln p_i = S_{BG}, \tag{15}$$

where i denotes the number of different microstates each of which has its own probability p_i of occurring.

There is some terminological confusion in the physics literature concerning such properties of physical quantities as additivity and extensivity. This is of particular importance in research on the entropy measure proposed by Tsallis, where those two notions are often treated as synonyms. However, in this case, non-extensivity cannot be identified with non-additivity [46–49]. In fact, the Tsallis entropy is pseudo-additive entropy of degree-q. If we have two independent systems A and B, such as the probability distribution of the composite system $A + B$ is expressed by the formula $P(A + B) = P(A)P(B)$, the Tsallis entropy of the $A + B$ system is as follows:

$$S_q(A + B) = S_q(A) + S_q(B) + {}^{(1-q)S_q(A)S_q(B)}/_{k}\,. \tag{16}$$

This property is referred to as pseudo-additivity. If $q = 1$, we are dealing with an additive system.

Integration of the classic econometric model with one of Tsallis non-extensive entropy, which involves a proper choice of constraints, allows for generating new data, which reduce the entropy of a system under study, thereby allowing for more precise estimation of the model parameters. This approach allows for analysing complex systems described with heavy-tailed distributions, with long memory and characterised by scale invariance. Since rare events occurring in non-ergodic systems are associated closely with heavy tails, the use of normal distribution is limited. In this manner, fractal and multifractal objects are included in considerations, which exhibit self-similarity in different temporal and spatial scales [50,51]. Therefore, power-law probability distributions become the focus of research attention; they are useful in describing systems both with low-frequency and high-frequency events.

Generalisations of the normal distribution using the non-extensive Tsallis entropy can describe ergodic and non-ergodic socio-economic systems. In the case of isolated macroscopic systems, ergodicity means that, with time, a phase trajectory runs through all the allowed microstates. On the other hand, non-ergodic systems are characterised by multi-level correlations between system microstates, with the consequent occurrence of non-extensive entropy in them, which is not a linear function of the number of possible states. There is quite convincing evidence that the amplitude and frequency of socio-economic phenomena do not substantially diverge from many extreme events occurring in nature if they are transformed by power law distribution into the same time or space scales. A distribution arising from the aggregation of statistical data depends on the nature of the phenomena

under study and it results from an effect of two attractors: the power-law attractor associated with extreme events and the Gaussian attractor, which occurs in the central part of the data series. The use of pseudo-additive Tsallis statistics in modelling allows for preserving long-range correlation and the observed time-scale invariance structure of high-frequency series in the process of transformation (aggregation) of high-frequency data to a lower frequency. The differences between an approach typical of NCEE and other econometric methods become visible when aggregated data approach the normal distribution, i.e., economic systems of relatively low complexity levels are considered. In such cases, entropy econometrics referring to Gibbs–Shannon's classic concepts gives a result similar to NCEE, because the Tsallis complexity index (q) goes to one. Similar results can be obtained with classic econometric techniques. However, NCEE methods are more useful in the case of complex systems described by Lévy processes [52]. Therefore, the non-ergodic approach is at least non-inferior to the parameter estimation methods now used in econometrics.

Entropy econometrics is particularly useful in forecasting input-output table parameters, which are presented as an inverse problem. We have an inverse problem when we use a set of observations to establish the causes that generated the set. In other words, the effects are used to establish the causes. It is the reverse of a forward problem, whose solution starts with causes, which are then used to determine the effects. A reverse problem can be defined as follows:

$$d_{obs} = F(p), \tag{17}$$

where d_{obs} is a set of noted observations, $F(p)$ denotes the forward map, and p stands for the set of model parameters. It denotes the determination of the parameters of the model p, which generated the set of observations d_{obs}. When it comes to input-output tables, they are often unbalanced, due to which a researcher faces an ill-behaved (troublesome) inverse problem. In such cases, the pseudo-additive Tsallis entropy gives a numerically more stable result than the Shannon entropy. Nevertheless, Shannon's entropy is useful for solving many other economic problems, such as developing a new interpretation for uncertainty and risk related to economic disparity [53] and analysing the quality of institutions in the European Union countries from the perspective of their influence on the speed of resource reallocation [54] or comparing the development level of the digital economy in various countries [55].

The NCEE method also enables obtaining good results for the estimation of production function parameters with constant elasticity of substitution (CES) between production factors, as it allows for demonstrating that a non-linear function CES can have an analytically closed-form solution. So far, solitons have been the best-known examples of such non-linear functions [56]. This result was achieved at the model variance minima and varying q-Tsallis complexity index when the estimated CES function was described with the power law. However, it required the previous demonstration of a relationship between the power-law distribution and the macroeconomic aggregate structure of national accounts. Moreover, calculations have demonstrated the superiority of the NCEE method over such traditional estimation techniques as Shannon entropy, the non-linear least squares, the generalized methods of moments, and the maximum likelihood approaches [57]. Non-extensive Tsallis entropy also finds a number of applications on financial markets, which includes studies concerning the distribution of return fluctuations for the Polish stock market index WIG20 [58], the origin of multifractality in the time series [59], the exchange rate return fluctuations [60], relationship between the stock market returns and corresponding trading volumes [61], the memory effect involved in returns of companies from WIG 30 index on the Warsaw Stock Exchange [62] and the asymmetry of price returns on stock and money markets [63]. Other types of entropy have also been applied to financial markets [64].

8. Program for the Development of Econophysics formulated by Zygmunt Rawita-Gawroński in 1958

The foundations of modern econophysics were formulated by the Polish economist Zygmunt Rawita-Gawroński. He pointed to the need to supplement the methodology of economics with some

ideas of physics, mainly with the theory of stochastic processes and the non-linear dynamics he anticipated, which he called an "aleatory momentum". Rawita-Gawroński's article is a collection of ideas aimed at improving economics by the large-scale introduction of physics methods. This work is still relevant and it seems that it will remain so for a very long time. Although written over 60 years ago, it reads as if it was written yesterday. Rawita-Gawroński's program of econophysical research consists of the following elements [65]:

8.1. General Formulation of the Econophysics Scientific Problem

The rapid development of physics and the related progress in mastering natural forces should be used to develop economic methodologies. First of all, the need to apply the concept of entropy and selected elements of thermodynamics, statistical mechanics and quantum theory is stressed.

8.2. A Criterion for Applying Physics Methods in Economics

The greatest benefits may be gained by economic theories with the same approach to the phenomena they study as in natural sciences. First of all, the microcosmic approach is mentioned, which has to be cleared of any metaphysical tarnish in order for the object of research to be quantitatively acceptable. This would allow for the application of mathematics. The importance of general principles applicable to all sciences is also emphasized, which indicates thinking in terms of general system theory.

8.3. The Principium Rationis Sufficientis as a Mental Aspect of Causality

This principle, which stems from the fundamental criterion of *praedicatus inest subjecto*, is based on divine wisdom. It is in line with Poincaré's views, for whom the physical phenomenon is the result of a combination of many causes, but their influences cannot be distinguished and evaluated. Causality is also the basis of statistical laws, manifested in mass phenomena.

8.4. Development of the Classic Concept of Equilibrium Based on the Works of Walras and Pareto

It is impossible to formulate a definition of equilibrium in economics without referring to the second law of thermodynamics. All isolated systems aim at a state of equilibrium corresponding to the minimum of free energy in relation to the total energy of the system, which is equivalent to the maximum of the entropy function. Movements leading to equilibrium can be reversible if they originate from the causal laws, or are irreversible if they are a consequence of an increase in entropy and thus the result of statistical laws.

8.5. Types of Equilibrium in Economics

Equilibrium should be classified according to how it is reached. This can be done by the cancellation of opposing moves, by fusion or by accumulation. The first type is the mechanical equilibrium achieved by cancellation, which occurs when all movements are homogeneous. Thermodynamic equilibrium, on the other hand, is achieved by a fusion of unobservable micro-movements, which leads to a macroscopic state associated with the maximum entropy. The third type is psychological equilibrium, whose importance in economics cannot be overestimated, based on the accumulation of past states. This is a fundamental issue in psychoanalysis and somewhat similar to inertia. This equilibrium is a kind of attitude of man towards the world he wants to know.

8.6. Anticipation of the Definition of Complex Adaptive Systems given by Gell-Mann Several Decades Later

In the statistical equilibrium, the principle of cognitive level difference between the phenomenon and its explanation scheme applies. The schemes are used to explain the forms of macroscopic phenomena. They are temporary because they are improved during cognitive processes.

8.7. Anticipating the Existence of Dissipative Structures in Economics and Physics

Statistical laws are a link between stochastic processes and life phenomena. The concept of personality has no clear bottom limit and it emerges slowly as the complexity of material structures increases. However, this explanation must not overlook the extremely important moment when particle coordination occurs. The usefulness of the ideas of evolution derived from entropy gives an important clue for further research in the sense that where organization and causality begins, statistical probabilism and determinism ends.

8.8. The Need to Take Open Systems into Account in Economics

Physical and economic systems are not completely isolated from the environment, so there is a possibility of fluctuations that counteract the trend towards the most probable state. This weakens the operation of statistical laws, and the significance and role of fluctuations bring the researcher closer to the problems of an individual and society and the organisation of the economy.

8.9. Anticipation of the Emergence of Ecological Economics (Bioeconomics) as Understood by Georgescu-Roegen and its Critical Judgment

In open systems, the phenomenon of entropy does not have to express a constant, irreversible degradation of energy in nature and reaching a state of complete homogeneity of the universe, where time and space lose all sense. A decrease in system energy quality due to entropy, which means a reduction in usable energy resources, is constantly hampered by technical progress in the use of energy sources. According to the first law of thermodynamics, with a constant amount of energy in the universe, entropy is equivalent to full energy deterioration only if we assume an unchanged level of technology.

8.10. Similar Limitations of Cognition in Physics and in Economics

In both sciences, there is a need to abstract from a great many causes, which affect the phenomena under study to a relatively small extent and thus can be omitted (*ceteris paribus* principle). In addition, the boundaries of scientific thought are marked by certain imperfections, antinomies or contradictions, such as Heisenberg's uncertainty principle, the unattainability of full statistical equilibrium and the impossibility of separating adjacent systems that are expressed by the same wave function or what might be called an anti-accident.

8.11. Collective Facts as the Basis of Complexity Economics and Quantum Economics

Methodological constructions in economics, based on deterministic physics of the nineteenth century, do not take into account uncertainty in human behaviour. On the other hand, the physics contemporary to Rawita-Gawroński took into account stochastic probabilism, which he called the aleatory momentum. The most important research problem he posed in this paper is whether methods of physics in combination with the mathematics of aleatoric phenomena will allow capturing the whole of human activities in economics. The criterion for assessing the scientific value of econophysics is its ability to describe collective facts, which are the result of many individual actions.

8.12. The Role of Physics in Economics Development

Postulated by physics, stochastics is connected with the problem of economic assumptions revision. It should be modified by replacing some causal or functional relations with stochastic relations.

8.13. The Basic Difference between Social and Physico-Chemical Sciences

The more accurate the statistical cognition and prediction in physics, the more related the condition of the object under study is to stochastic chaos. Difficulties occur only when molecule coordination occurs and causal relationships are revealed. The main role in the description of an object in social

sciences is played by an organizational element and the associated causality. The aleatory momentum, associated with stochastic chaos, is much less important, but it is the main connecting factor of these sciences.

8.14. Sources of Econophysics Advantage over Mainstream Economics

Different world views, traditionalisms, political orientations, environmental influences, one's own temperament, etc., place a much greater mental burden on a theorist of economics than on a physics theorist.

The well-known Italian physicist Ettore Majorana, much earlier than Rawita-Gawroński, stressed the importance of statistical laws, especially those in thermodynamics and quantum mechanics, for the development of economics and other social sciences. However, unlike the latter, he did not formulate a more specific program of economics development supported by physical methods, i.e., econophysics [66,67]. Another difference between Majorana and Rawita-Gawroński is that the former limited himself to an analogy between physics and social sciences, while the latter thought in terms of the isomorphism principle formulated by von Bertalanffy. Nevertheless, Majorana was perfectly aware of the importance of entropy for the development of social sciences. He treated it as an additive quantity. He stressed that the development of physics emphasizes the importance of statistical laws in the whole science. He also shared Sorel's view that determinism applies only to the phenomena classified as artificial nature, which do not occur in the presence of significant degradation of energy as provided for by the second law of thermodynamics. He pointed out that this principle must be understood as a statistical law, since entropy only describes macroscopic states, while the number of internal possibilities is so large that it is impossible to isolate and explore them all. Therefore, the number of these internal possibilities is a measure of the degree of hidden indeterminacy of the system. One can only assume that they are equally probable. Analogies with social systems are rather obvious.

Majorana perceived the role of quantum mechanics in the development of social sciences somewhat differently. It is based on statistical laws, which, in his opinion, differ significantly from the classical statistical laws, where uncertainty is the result of voluntary resignation for practical reasons from establishing the initial conditions of systems in every detail. Quantum mechanics lacks objectivity in the description of phenomena because the measurement itself changes the examined system. Moreover, this theory changes the rules of the determination for internal configurations in systems, which affects such phenomena as entropy. In classical theory, the number of internal possibilities could be infinite, whereas in quantum theory there is an essential discontinuity in the description of phenomena associated with the Planck constant. This means that the number of microstates is finite, although still huge. Therefore, in quantum physics, we have probabilistic laws, which are hidden at lower levels of reality than customary statistical laws. Rawita-Gawroński also sees this problem and claims that the introduction of quantum theory into thermodynamics removes the inability to accurately determine entropy and thus changes the definition of statistical equilibrium. It seems that both Majorana and Rawita-Gawroński present here the nucleus of the idea developed later by Gell-Mann, that deterministic chaos is a mechanism that reinforces indeterminacy inherent in quantum mechanics to macroscopic levels—occurring both in physics and in economics.

9. Modern Econophysics

Econophysics is a transdisciplinary science based on the observation that physical and economic objects can have a common theory. Since there are some logical homologies at its base it is an exemplification of a known isomorphism principle developed by Ludwig von Bertalanffy. The emergence of transdisciplinary areas of knowledge is consistent with the general system theory paradigm which he formulated [68]. According to the isomorphism principle, there are structural similarities between objects described by different science branches. Isomorphism is not associated with analogies, which are only superficial similarities of phenomena and processes, but with logical

homologies. They occur when factors affecting certain phenomena or processes are different, whereas the formal laws that govern the dynamics of apparently different objects are identical. Discovering homologies facilitates scientific work as it reduces the time needed to explain phenomena. An issue, which is still a challenge in one branch of science, may be formally similar to another in a different branch, which has long been explained. Imparting a relatively broad meaning to the concept of a system enables transferring models from one area to another. Philip Mirowski carried out a comprehensive analysis of the relationship between natural sciences, information theory and economics [69–73].

Mantegna and Stanley understand econophysics as activities of physicists who strive to solve economic problems by applying methods already tested in different branches of physics [74]. However, such a definition seems to be too one-sided because it assumes cognitive actions only by physicists. One must admit that studies of this type were initiated after 1995 only by physicists [75], but now an increasing number of economists can also apply modern methods and techniques originating in physics to describe markets and economies. It seems that the dynamic development of econophysics must be based on permanent cooperation between physicists and economists in solving economic problems, which may prove beneficial for the development of both economics and physics [76–78].

Development of a branch of science is usually measured by its ability to formulate new knowledge about reality. Progress in research can be identified both when the use of traditional methods has led to the discovery of new facts and when new scientific laws have been discovered using innovative methods. Econophysics is an attempt to develop economic science, which is based mainly on the transfer of research methods and techniques from physics to economics. Thus, in this case we are dealing with a second possibility of increasing the knowledge base. The methods of physics most frequently used in economics include non-linear dynamics, the theory of stochastic processes, cellular automata and quantum mechanics.

The very definition of econophysics shows a certain methodological primacy of physics over economics, but numerous, although less known, examples can be given to the contrary: that economics initiated the development of physics. The first power law was discovered in the late nineteenth century by the Italian economist Vilfredo Pareto, who studied income distribution in different societies. He observed a stable relationship, independent of time and space, which differed significantly from the Gaussian curve [79,80]. Income distributions were similar in different countries and periods. If $N(y)$ denotes the number of people with the income not lower than y, then the power law will describe changes in individuals' wealth under stable economic conditions:

$$N(y) = C\, y^{-\alpha}, \tag{18}$$

where C is a certain constant and $\alpha > 0$ is called the critical exponent that Pareto estimated to be 1.5. Such distributions were applied in physics much later, mainly owing to the work of Lévy and Mandelbrot.

It was similar to the random walk concept, which Bachelier developed and applied in the pricing of options in speculative markets as early as 1900 [81]. This idea appeared in physics only five years later in one of Einstein's works [82]. In this context, the story of the discovery of the butterfly effect, which means the sensitive dependence on initial conditions, occurring in non-linear dynamical systems, is also very interesting. This was discovered by Poincaré in 1890, when he dealt with the restricted three-body problem [83]. However, this discovery was forgotten for a long time. This phenomenon was again encountered in the 1930s by the Swedish economist Palander when he studied certain models of duopoly and oligopoly [84–86]. They were also observed in the early 1950s in the course of numerical exploration of Goodwin's non-linear business cycle model [87–89]. In physics, the butterfly effect was rediscovered only in 1963, when Edward Lorenz revealed the presence of the chaotic attractor, known today as the Lorenz attractor, in the non-linear model for atmospheric convection [90]. It can be seen that science is holistic, and the lack of cooperation between economists and physicists has caused considerable delays in its development. It is noteworthy that a certain cognitive balance between physics and economics can be seen in the formulation of entropy economics given by Georgescu-Roegen,

which speaks in favour of this approach. Therefore, it should be accepted that there is an inherent two-way relationship between these sciences.

The most interesting achievements of econophysics apply to financial markets [91–95]. This is because these markets generate sufficiently large and precise datasets, which is a prerequisite for the effective application of physical methods. The flagship achievement here is the questioning of the efficient market hypothesis, as a result of the finding of endogenous self-organisation processes in major world stock exchanges [96]. Notable successes have also been achieved in modelling phase transitions, catastrophic and critical phenomena [97–99]. Furthermore, the sphere of interests of econophysicists embraces issues related to business cycles, factors of economic growth, income and wealth distributions, economic equilibrium, property markets, mechanisms of hyperinflation and the development of enterprises. Empirical research carried out within the framework of econophysics either complements or contradicts traditional economic knowledge. If the latter occurs, it is necessary to verify and update this knowledge constantly. The broad front of empirical research also forces changes in economic theory. Progress, in this case, is two-directional. On the one hand, the logical consistency of conventional economic models is being tested, while on the other, the possibility of developing new theories of markets and economies is being created.

Although econophysics and mainstream economics have the same goal of solving economic problems, due to significant methodological differences they are treated as two separate sciences [100]. Economists usually build their models on empirically untested premises that are treated as religious dogmas [101,102]. Thus, mainstream economics is largely a deductive science that uses axiomatic mathematical modelling. The application of economics to government policies regarding science and technology is particularly unreliable. Its usefulness is especially limited as far as the rational management of research and the conduct of thoughtful science policy are concerned [103]. Econophysics, on the other hand, is an inductive science which emphasizes empirical research and discovers relationships contained in data using mathematical tools and logic. Its essence is not to match observations to a priori models, but to discover the mechanisms of real economic systems. The success of econophysics as a science and the proper use of its achievements will only be possible if it has a noticeable impact on the course of economic phenomena. For this to happen, econophysical models must allow for accurate predictions and be accepted by economists as useful for economic policy.

10. Entropy, Ignorance and Complexity

As shown above, entropy and information are closely linked. It follows that entropy can be understood as a measure of ignorance [104]. Information and ignorance are opposites, but what one of them measures is also measured by the other. The entropy of a certain macrostate of the system is a measure of our ignorance concerning the actual microstate and it is equal to the number of bits of additional information necessary to describe that microstate, as long as all microstates in the macrostate are equally probable. This reasoning also applies when the system is not in a single macrostate, but can take many various macrostates, each with a different probability of occurrence. The entropy of the macrostates should then be averaged using appropriate weights in the form of their probabilities. Furthermore, entropy includes the number of bits of information necessary to determine a macrostate. To sum up, entropy means the sum of the average ignorance referring to a specific microstate within the macrostate and the ignorance of the macrostate itself. Thus, it can be claimed that—according to the second law of thermodynamics—specification of the state is equivalent to order, while ignorance corresponds to disorder.

The definition of entropy as a measure of ignorance may also take into account the operation of Maxwell's demon, which controls a vessel filled with a gas at uniform temperature [105] (pp. 308–309). While individual gas molecules move at different speeds, the mean velocity of a large number of molecules is the same. After dividing the vessel into two parts, A and B, the demon sorts the molecules in such a way that faster ones are allowed to pass from A to B, and slower ones from B to A. As a result of the demon's action, the temperature of part B increases and the temperature of chamber A decreases,

without performing work, which is contradictory to the second law of thermodynamics. Heat flows from the cold gas to the hot one. The operation of the demon should be explained as follows. It can force the flow of heat from cold to hot bodies, but in addition to being able to measure molecular velocities, it must also collect information about these velocities. The entropy of the system consisting of cold and hot bodies can be reduced at the expense of increasing the number of records in the demon's memory. However, this memory is finite and after it is filled in, it will be necessary to erase old records to make room for new ones. The erasing of the last copy of the information record causes at least such an increase in entropy that restores the operation of the second law of thermodynamics. Consequently, Gell-Mann proposes a new definition of entropy of the entire system, which does not violate the second law of thermodynamics, even when records exist. Algorithmic information content of records with relevant information should be added to entropy being a measure of ignorance. The algorithmic information content of a bit string is the length of the shortest computer program containing specific information. This means the exchange of ignorance for records. A decrease in entropy results in the information record to emerge, and then ignorance is reduced, but at the same time algorithmic information content of the record increases. However, erasing the record reduces the information it contains, but causes at least the same increase in ignorance related to the entire closed system.

The relationship between information and ignorance can be expressed by a mathematical formula. The ignorance measure I can be defined based on Shannon informational entropy, in which the multiplicative constant is omitted [106]:

$$I = -\sum_r P_r \log P_r \,, \tag{19}$$

where P_r stands for coarse-grained probability for the individual member r of the ensemble, while log means logarithm to the base 2 and, therefore, uncertainty is measured in bits. Based on this, a generalized measure of ignorance I_q can be defined, taking the following form:

$$I_q = -\left[\sum_r (P_r)^q - 1\right] \Big/ (q-1) \,, \tag{20}$$

which is reduced to Equation (19) in the limit where q approaches 1. This measure can be applied to describe complex systems operating at the edge of chaos, which—as will be demonstrated—are crucial to understanding the functioning of markets and economies.

In order to solve problems related to the functioning of complex systems, such quantities as the effective complexity and total information are used [107]. The effective complexity should be understood as the length of a compact description of the regularities identified in the examined system, while total information is the sum of the effective complexity and an entropy term, which measures information necessary to describe the random aspects of the entity. Thus the total information describes both rule-based features and the probabilistic features of the perceived entity. Comparison of different sets of identified regularities of an entity permits us to show that the best set minimizes total information, which leads to minimizing the effective complexity. The effective complexity specified in such a way seems to be independent of the observer. Furthermore, there exists a relation between effective complexity and Bennett's logical depth [108]. If the effective complexity of the binary string exceeds a certain clearly defined threshold, then this string has astronomically large depth. Otherwise, the depth is arbitrarily small.

To conclude this part, it is worth mentioning a certain paradox associated with the growth of knowledge. In recent years, our knowledge has been increasing exponentially, but ignorance has been growing at an even faster rate [109]. Therefore, an increase in knowledge is accompanied by an increase in entropy. Thus, total information about the examined system increases not only due to an increase in effective complexity, but also as a result of increased ignorance. However, Maxwell's demon has limited memory, so it cannot indefinitely convert ignorance into records.

11. Complexity Economics

Although econophysics has a lot in common with complexity economics, it is not possible to equate these two sciences. Econophysics is a scientific discipline which consists of applying physical models and methods to study economic phenomena [110–114]. Complexity economics, in turn, is a broader discipline, since—apart from the fact that it contains almost all econophysics—it examines economic problems within complexity science. Since it is the opposite of equilibrium economics, it does not focus on order, determinacy, deduction or stasis, but as non-equilibrium economics, it emphasizes the importance of contingency, indeterminacy, sense-making and openness to change [115–119]. Equilibrium (neoclassical) economics is a special case of complexity economics. Additionally, complexity economics has derived many ideas and concepts from biology because it treats the economy as an evolutionary system [120,121].

Complexity economics is a response to certain shortcomings of conventional economics, such as the *homo oeconomicus* model, treating markets and economies as closed systems or relying on the assumption of linearity. The *homo oeconomicus* model defines a set of rules concerning human behaviour in economic processes, which include unbounded rationality, unbounded willpower and unbounded selfishness [122]. A closed system aiming at a state of thermodynamic equilibrium has become the main metaphor determining economic thinking throughout almost the entire 20th century. This assumption became the cornerstone of many elegant mathematical models of general equilibrium and certainly inspired the creators of the rational expectation hypothesis. Linearity, on the other hand, consists in reducing relations between phenomena to straight lines. If, in a certain system, a change of one variable at a given time by a certain multiple or fraction will change the same or another variable at a later time by the same multiple or fraction, then the system under consideration is linear. After plotting later values of a given variable as a function of previous values of the same or another variable, it turns out that the points lie on a straight line. These assumptions caused the cognitive gap between conventional economics and reality to grow steadily as a result of humanity's economic development. In such conditions, the phenomenon of innovation or economic growth could only be explained by referring to the random exogenous shock.

In recent years, it has become apparent that dissipative structures can be identified in economic systems, i.e., those with steady states existing far from equilibrium. In such cases, we are dealing with non-extensive entropy, without the features of additivity, i.e., which is not a sum of values calculated for all subsystems. Entropy in an economic system is the sum of the production of entropy in the system itself and its exchange with the environment. The inflow of negative entropy from the environment can, therefore, compensate for the production of entropy in the system itself. After the introduction of certain disorders into the dissipative systems, it appears that they do not have to go to a state of equilibrium corresponding to the maximum entropy. In states far from equilibrium, orderly structures based on long-range correlations are created, and the occurring phenomena are characterized by long-term memory. The sensitivity of these systems to interactions of internal elements and environmental influences is related to the dissipation of energy [123,124]. Identification of dissipative structures in real systems was a significant advance in the development of the science of complexity and laid the foundations for another scientific breakthrough in identifying complex adaptive systems in nature.

Complexity economics treats markets and economies as complex adaptive systems that are by their very nature open, non-linear, consisting of interacting agents and exhibiting emergence. The sources of interaction are agents, i.e., mainly people, although they may also be other biological organisms or even computer programs. Complexity should be understood as the dynamic properties of a system that cause its behaviour to be non-periodic or structural changes to occur in it due to endogenous forces. A very important feature of these systems is emergence, which consists in their ability to create ordered collective phenomena, which can only be described at a level higher than that used to describe the components (basic rules). Complex adaptive systems collect information about their environment and their relationships with it, discover regularities in them, and use this information to build cognitive schemes that they use to operate in the real world. The effects of these activities generate a new, return

flow of data used to evaluate the available set of cognitive patterns. As a result of this process, some of these schemes will survive and be improved, while others will be rejected [104].

Within complexity economics, markets and economies are understood as extremely dynamic, constantly evolving systems that resemble the brain, the Internet or the ecosystem in terms of their functionality. The greatest differences between complexity economics and conventional economics concern 10 issues: dynamics, agents, networks, emergence, evolution, technology, preferences, sources, elements and the horizon of predictability [125,126]. New ideas developed by complexity economics indicate that markets and economies tend to operate far from being balanced and that different agents should be modelled individually. They usually have incomplete information, make mistakes, but also have the ability to learn and adapt. Multilateral interactions between agents form networks that change over time. There is no difference between microeconomics and macroeconomics, as the latter is the emergent result of the interaction and behaviour of agents at the microeconomic level. The development of markets and economies is possible as a result of innovation, which is an essential factor for economic growth. Technological progress is endogenous to the economic system. The formation of individual preferences is a central issue in complexity economics, and agents do not necessarily have to be selfish. Complexity economics has its sources in biology, which is based on morphogenesis, structures, shapes, forms, capabilities, self-organization and life cycles. The source of conventional economics, on the other hand, is 19th-century physics, referring to balance and stability, which is why modern economics textbooks are mainly based on prices and quantities. In complexity economics, forecasting the dynamics of markets and economies is limited to the so-called Lapunov time, which is the inverse of the Lapunov exponent known in non-linear dynamics. The latter is a measure of the butterfly effect, so it determines whether the system is sensitive to initial conditions. The butterfly effect (deterministic chaos) occurs when this exponent is positive.

According to some critics, complexity science has difficulty in defining the very concept of complexity. Indeed, Horgan points to the existence of at least 45 different definitions of complexity, most of which refer to concepts as imprecise as entropy, randomness or information [127]. However, this is not evidence of the weakness of the concept of complexity, as this is the kind of differentiation that should be expected. The concept of complexity is only partially objective, as it depends on the level of knowledge of reality and the state of consciousness of a given agent. Therefore, the term is irreducibly subjective. Undoubtedly, however, all agents acting in certain conditions can be assigned shared initial knowledge, because what unites them is the search for order and regularity in a turbulent world. A certain amount of subjectivity in the definition of complexity must result from coarse-grained averaging, i.e., the accuracy with which a given agent (complex adaptive system) perceives reality. Therefore, it is best to adopt the definition of the effective complexity of an entity proposed by Gell-Mann. It should be understood as the length of a concise schema showing the regularity of this entity, which has been drawn up by the observer – the chosen complex adaptive system. The concept under consideration is, therefore, inherently and inextricably linked with the cognitive agent, and disregarding this relationship will always be a source of misunderstanding.

In the greatest discovery of complexity economics, it was proved that some markets and economies have the ability to tune in to the edge of chaos. A necessary, but not sufficient, condition is that they are complex adaptive systems. The edge of chaos is the distinguished state of a system situated between periodic and chaotic behaviour, where it has the maximum computing power. Around the edge of chaos, the complexity of the system is close to optimal and its adaptability is the greatest [128,129]. The edge of chaos may have one of the following forms: chaotic attractors and repellers, catastrophes of complexity, coexistence of attractors, sensitive dependence on parameters, final state sensitivity, effects of fractal basin boundaries, chaotic saddles [130].

12. Conclusions

There is no doubt that the emergence of entropy in economics has led to many surprising and important discoveries in the functioning of markets and national economies. However, not all of them

have been fully accepted by mainstream economics. There are many reasons for this and they are quite well known. Furthermore, entropy has contributed to the emergence of econophysics, complexity economics, and quantum economics. This makes the situation in modern economics quite strange. If one compares it to a tree, the branches representing mainstream economics are developing more and more slowly, while the branches based on entropy are developing rapidly, creating new economic knowledge quite quickly. This disequilibrium may end in two ways: either mainstream economics will accept the achievements of econophysics, complexity economics and quantum economics, which will require—as Rawita-Gawroński noted more than 60 years ago—a revision of the content of economic assumptions, or economics will become the next physical science [131].

The concepts discussed here have contributed to the remarkable development of economics, but some weaknesses, particularly in their practical application, have not been overcome. The most important of these is the inability of econometrics, both traditional and non-extensive entropy-based, to determine the value of parameters accurately. This boundary of cognition results directly from non-linear dynamics, i.e., from the existence of deterministic chaos in the real world. Therefore, it is a result of the very nature of reality. Econometricians assume the correctness of a predetermined model with a number of unknown parameters and then try to forcefully match it to a non-stationary time series by the best choice of parameters. However, non-linear dynamics indicate that by matching any infinite precision model (stochastic or deterministic) to inherently finite precision data, non-uniqueness cannot be avoided [132]. The source of such problems is the sensitive dependence on parameters in non-linear dynamical systems, in which the parameter values associated with stable periodic orbits may be close to chaotic parameter values. Moreover, being unable to determine the initial conditions of the systems under study with infinite accuracy, we are exposed to the butterfly effect, whether we want it or not. Thus, econometricians mislead themselves and others by thinking that their models are helpful in understanding economic processes.

Some hopes for solving certain methodological problems are placed on a relatively new research trend called quantum economics [133,134]. It also appears that many economic issues can be treated as quantum phenomena. Logical homology, in this case, is a bridge connecting the dynamics of economic objects with the laws of motion of particles, which are of interest to quantum mechanics. The basic source of isomorphism here is Heisenberg's uncertainty principle [135]. The more precisely we try to determine the position of some particle, the less precisely we are able to determine its momentum and vice versa. An example of a quantum phenomenon in economics can be the process of setting the price of a good or service. In quantum economics the value of a good is rather undefined, so the price reflecting it cannot be determined. Only making a transaction—by analogy to the uncertainty principle—is a measurement of value and allows for determining the price precisely. During this process, the exchange of money takes place and, therefore, money acts as a measuring device in the markets. Therefore, money is a fundamental element of economic analysis in quantum economics [136]. This concept is very promising [137–142], but time will tell if it will broaden our understanding of the regularities governing the markets and national economies. A number of arguments exist for including quantum mechanics in social scientific debates, since consciousness is probably a macroscopic quantum phenomenon [143].

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Ekelund, R.B., Jr.; Hébert, R.F. *Secret Origins of Modern Microeconomics: Dupuit and the Engineers*; The University of Chicago Press: Chicago, IL, USA, 1999; ISBN 0-226-19999-1.
2. Ekelund, R.B., Jr.; Hébert, R.F. Retrospectives: The origins of neoclassical microeconomics. *J. Econ. Perspect.* **2002**, *16*, 197–215. [CrossRef]

3. Marshall, A. *Principles of Economics. An Introductory Volume*; Macmillan: London, UK, 1947; ISBN 978-0879910518.
4. Fisher, I. *Mathematical Investigations in the Theory of Value and Prices*; Yale University Press: New Haven, CT, USA, 1925; ISBN 978-1-61427-305-9.
5. Amin, T.G.; Jeppsson, F.; Haglund, J.; Strömdahl, H. Arrow of time: Metaphorical construals of entropy and the second law of thermodynamics. *Sci. Educ.* **2012**, *96*, 818–848. [CrossRef]
6. Clausius, R. *The Mechanical Theory of Heat: With Its Applications to the Steam-Engine and to the Physical Properties of Bodies*; Hirst, T.A., Ed.; John van Voorst: London, UK, 1867; ISBN 9789353740962.
7. Planck, M. *Vorlesungen über die Theorie der Wärmestrahlung*; Verlag von Johann Ambrosius Barth: Leipzig, Germany, 1906; ISBN 9781334013331.
8. Chakrabarti, C.G.; Chakrabarty, I. Boltzmann-shannon entropy: Generalization and application. *Mod. Phys. Lett. B* **2006**, *20*, 1471–1479. [CrossRef]
9. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*, 16th ed.; University of Illinois Press: Urbana, IL, USA, 1971; ISBN 978-0252725487.
10. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [CrossRef]
11. Jaynes, E.T. Information theory and statistical mechanics. II. *Phys. Rev.* **1957**, *108*, 171–190. [CrossRef]
12. Tribus, M.; McIrvine, E.C. Energy and information. *Sci. Am.* **1971**, *225*, 179–188. [CrossRef]
13. Avery, J. *Information Theory and Evolution*; World Scientific Publishing: Singapore, 2003; ISBN 9812383999.
14. Georgescu-Roegen, N. *The Entropy Law and the Economic Process*; Harvard University Press: Cambridge, MA, USA, 1971; ISBN 9780674281646.
15. Daly, H.E. *Ecological Economics and Sustainable Development, Selected Essays of Herman Daly*; Edward Elgar: Cheltenham, UK, 2008; ISBN 978-1-84720-988-7.
16. Meadows, D.H.; Meadows, D.L.; Randers, J.; Behrens, W.W., III. *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*; Universe Books: New York, NY, USA, 1972; ISBN 0-87663-165-0. [CrossRef]
17. Levallois, C. Can de-growth be considered a policy option? A historical note on Nicholas Georgescu-Roegen and the Club of Rome. *Ecol. Econ.* **2010**, *69*, 2271–2278. [CrossRef]
18. Lietaer, B.; Arnsperger, C.; Goerner, S.; Brunnhuber, S. *Money and Sustainability: The Missing Link. A Report from the Club of Rome—EU Chapter to Finance Watch and the World Business Academy*; Triarchy Press: Axminster, Devon, UK, 2012; ISBN 978-1-908009-75-3.
19. Raine, A.; Foster, J.; Potts, J. The new entropy law and the economic process. *Ecol. Complex.* **2006**, *3*, 354–360. [CrossRef]
20. Røpke, I. Trends in the development of ecological economics from the late 1980s to the early 2000s. *Ecol. Econ.* **2005**, *55*, 262–290. [CrossRef]
21. Syll, L.P. *Nicholas Georgescu-Roegen and the Nobel Prize in Economics*. Available online: https://larspsyll.wordpress.com/2012/05/16/nicholas-georgescu-roegen-and-the-nobel-prize-in-economics/ (accessed on 11 January 2020).
22. Söderbaum, P. Economics, evaluation and environment. In *Economics of Ecosystem Management*; Hall, D.O., Myers, N., Margaris, N.S., Eds.; Dr. W. Junk Publishers: Dordrecht, The Netherlands, 1985; pp. 5–17, ISBN 978-94-010-8928-9.
23. Söderbaum, P. *Ecological Economics: A Political Economics Approach to Environment and Development*; Earthscan: New York, NY, USA, 2000; ISBN 978-1-85383-685-5.
24. Brown, J.; Söderbaum, P.; Dereniowska, M. *Positional Analysis for Sustainable Development: Reconsidering Policy, Economics and Accounting*; Routledge: New York, NY, USA, 2017; ISBN 978-1-138-63450-3.
25. Włodarczyk, J. Comparative analysis of the course of business cycles and thermodynamic cycles. *Equilibrium* **2011**, *6*, 127–139. [CrossRef]
26. Ackoff, R.L. *Creating the Corporate Future: Plan or be Planned For*; John Wiley and Sons: New York, NY, USA, 1981; ISBN 978-0-471-09009-0.
27. DeMarco, T.; Lister, T. *Peopleware: Productive Projects and Teams*, 3rd ed.; Addison-Wesley: Upper Saddle River, NJ, USA, 2013; ISBN 978-0-321-93411-6.
28. Erçetin, Ş.Ş.; Açıkalın, Ş.N. Lead-entropy: redefining leadership from the perspective of organizational entropy. In *Handbook of Research on Chaos and Complexity Theory in the Social Sciences*; Erçetin, Ş.Ş., Bağcı, H., Eds.; Information Science Reference: Hershey, PA, USA, 2016; pp. 183–194, ISBN 9781522501480. [CrossRef]

29. Coldwell, D. Entropic citizenship behavior and sustainability in urban organizations: Towards a theoretical model. *Entropy* **2016**, *18*, 453. [CrossRef]
30. Level of measurement. In *Encyclopedia of Public Health*; Kirch, W. (Ed.) Springer Science + Business Media: New York, NY, USA, 2008; pp. 851–852, ISBN 978-1-4020-5614-7.
31. Stevens, S.S. On the theory of scales of measurement. *Science* **1946**, *103*, 677–680. [CrossRef]
32. Stevens, S.S. Measurement, statistics and the schemapiric view. *Science* **1968**, *161*, 849–856. [CrossRef]
33. Stevens, S.S. *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects*; Transaction Publishers: New Brunswick, NJ, USA, 2008; ISBN 978-0-88738-643-5.
34. Bridgman, P.W. *The Logic of Modern Physics*; The Macmillan Company: New York, NY, USA, 1927; ISBN 978-9997497581.
35. Daly, H.E.; Farley, J. *Ecological Economics: Principles and Applications*; Island Press: Washington, DC, USA, 2004; ISBN 1-55963-312-3.
36. Bryant, J. *Entropy Man*; VOCAT International Ltd.: Harpenden, UK, 2015; ISBN 978-0-9562975-4-9.
37. Keynes, J.M. The general theory of employment, interest and money. In *The Collected Writings of John Maynard Keynes*; Robinson, A., Moggridge, D., Eds.; Cambridge University Press: Cambridge, UK, 2013; Volume 7, ISBN 978-1-107-67373-1.
38. Phillips, A.W. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. *Econ.–New Ser.* **1958**, *25*, 283–299. [CrossRef]
39. Samuelson, P.A.; Solow, R.M. Analytical aspects of anti-inflation policy. *Am. Econ. Rev.* **1960**, *50*, 177–194. Available online: https://www.jstor.org/stable/1815021 (accessed on 20 March 2020).
40. Bwanakare, S. Non-extensive entropy econometrics and CES production models: Country case study. *Stat. J. IAOS* **2016**, *32*, 709–713. [CrossRef]
41. Bwanakare, S. *Non-Extensive Entropy Econometrics for Low Frequency Series: National Accounts-Based Inverse Problems*; De Gruyter Open Ltd.: Warsaw, Poland; Berlin, Germany, 2017; ISBN 978-3-11-055044-3. [CrossRef]
42. Cartwright, J. Roll over, Boltzmann. *Phys. World* **2014**, *27*, 31–35. [CrossRef]
43. Gell-Mann, M.; Tsallis, C. *Nonextensive Entropy: Interdisciplinary Applications*; Oxford University Press: New York, NY, USA, 2004; ISBN 0-19-515976-4.
44. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
45. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*; Springer Science + Business Media: New York, NY, USA, 2009; ISBN 978-0-387-85358-1. [CrossRef]
46. Touchette, H. When is a quantity additive, and when is it extensive? *Physica A* **2002**, *305*, 84–88. [CrossRef]
47. Lavenda, B.H.; Dunning-Davies, J. Pseudo-additive entropies of degree-q and the Tsallis entropy. *J. Appl. Sci.* **2005**, *5*, 315–322. [CrossRef]
48. Jiulin, D. Property of Tsallis entropy and principle of entropy increase. *Bull. Astr. Soc. India* **2007**, *35*, 691–696.
49. Jizba, P.; Korbel, J. On the uniqueness theorem for pseudo-additive entropies. *Entropy* **2017**, *19*, 605. [CrossRef]
50. Mandelbrot, B.B. *The Fractal Geometry of Nature*; W.H. Freeman and Company: New York, NY, USA, 1983; ISBN 0-7167-1186-9.
51. Mandelbrot, B.B. A multifractal walk down Wall Street. *Sci. Am.* **1999**, *280*, 70–73. [CrossRef]
52. Lévy, P. *Calcul des Probabilités*; Gauthier-Villars: Paris, France, 1925; ISBN 978-2-87647-231-0.
53. Mishra, S.; Ayyub, B.M. Shannon entropy for quantifying uncertainty and risk in economic disparity. *Risk Anal.* **2019**, *39*, 2160–2181. [CrossRef]
54. Balcerzak, A.P. Quality of institutions in the European Union countries. Application of TOPSIS based on entropy measure for objective weighting. *Acta Polytech. Hung.* **2020**, *17*, 101–122. [CrossRef]
55. Balcerzak, A.P. Digital economy in Czech Republic, Slovakia and Hungary. Measurement with TOPSIS based on entropy measure for objective weighting. In Proceedings of the 11th International Days of Statistics and Economics. Conference Proceedings, Prague, Czech Republic, 14–16 September 2017; Löster, T., Pavelka, T., Eds.; Libuše Macáková, Melandrium: Slaný, Czech Republic, 2017; pp. 49–57, ISBN 978-80-87990-12-4.
56. Seifritz, W. What is sustainable development? An attempt to interpret it as a soliton-like phenomenon. *Chaos Solitons Fractals* **1996**, *7*, 2007–2018. [CrossRef]
57. Bwanakare, S. Non-extensive entropy econometrics: New statistical features of constant elasticity of substitution-related models. *Entropy* **2014**, *16*, 2713–2728. [CrossRef]
58. Rak, R.; Drożdż, S.; Kwapień, J. Nonextensive statistical features of the Polish stock market fluctuations. *Physica A* **2007**, *374*, 315–324. [CrossRef]

59. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Rak, R. Quantitative features of multifractal subtleties in time series. *Europhys. Lett.* **2009**, *88*, 60003. [CrossRef]
60. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Rak, R. The foreign exchange market: Return distributions, multifractality, anomalous multifractality and the Epps effect. *New J. Phys.* **2010**, *12*, 105003. [CrossRef]
61. Rak, R.; Drożdż, S.; Kwapień, J.; Oświęcimka, P. Stock returns versus trading volume: Is the correspondence more general? *Acta Phys. Pol. B* **2013**, *44*, 2035–2050. [CrossRef]
62. Bil, Ł.; Grech, D.; Podhajska, E. Methods of non-extensive statistical physics in analysis of price returns on Polish stock market. *Acta Phys. Pol. A* **2016**, *129*, 986–992. [CrossRef]
63. Bil, Ł.; Grech, D.; Zienowicz, M. Asymmetry of price returns—Analysis and perspectives from a non-extensive statistical physics point of view. *PLoS ONE* **2017**, *12*, e0188541. [CrossRef] [PubMed]
64. Vințe, C.; Smeureanu, I.; Furtună, T.-F.; Ausloos, M. An intrinsic entropy model for exchange-traded securities. *Entropy* **2019**, *21*, 1173. [CrossRef]
65. Rawita-Gawroński, Z. Pewne punkty styczne między metodologią najnowszych nauk fizycznych a metodologią teorii ekonomii. *Ekonomista* **1958**, *6*, 1691–1708.
66. Mantegna, R.N. Presentation of the English translation of Ettore Majorana's paper: The value of statistical laws in physics and social sciences. *Quant. Financ.* **2005**, *5*, 133–140. [CrossRef]
67. Mantegna, R.N. The tenth article of Ettore Majorana. *Europhys. News* **2006**, *37*, 15–17. [CrossRef]
68. Bertalanffy von, L. *General System Theory: Foundations, Development, Applications*; George Braziller: New York, NY, USA, 1968; ISBN 978-0-8076-0453-3.
69. Mirowski, P. *More Heat than Light: Economics as Social Physics, Physics as Nature's Economics*; Cambridge University Press: Cambridge, UK, 1991; ISBN 0-521-42689-8.
70. Mirowski, P. *Natural Images in Economic Thought: "Markets Read in Tooth and Claw"*; Cambridge University Press: Cambridge, UK, 2008; ISBN 978-0-521-47884-7.
71. Mirowski, P. *Machine Dreams: Economics Becomes a Cyborg Science*; Cambridge University Press: New York, NY, USA, 2002; ISBN 0-521-77283-4.
72. Mirowski, P. *Never Let a Serious Crisis Go to Waste: How Neoliberalism Survived the Financial Meltdown*; Verso: London, UK, 2014; ISBN 978-1-78168-302-6.
73. Mirowski, P.; Nik-Khah, E. *The Knowledge We Have Lost in Information: The History of Information in Modern Economics*; Oxford University Press: New York, NY, USA, 2017; ISBN 978-0190270056.
74. Mantegna, R.N.; Stanley, H.E. *An Introduction to Econophysics: Correlations and Complexity in Finance*; Cambridge University Press: Cambridge, UK, 2000; ISBN 0-521-62008-2.
75. Jakimowicz, A. Econophysics as a new school of economic thought: Twenty years of research. *Acta Phys. Pol. A* **2016**, *129*, 897–907. [CrossRef]
76. Jovanovic, F.; Schinckus, C. Breaking down the barriers between econophysics and financial economics. *Int. Rev. Financ. Anal.* **2016**, *47*, 256–266. [CrossRef]
77. Jovanovic, F.; Schinckus, C. *Econophysics and Financial Economics: An Emerging Dialogue*; Oxford University Press: New York, NY, USA, 2017; ISBN 978-0-19-020503-4.
78. Jovanovic, F.; Mantegna, R.N.; Schinckus, C. When financial economics influences physics: The role of econophysics. *Int. Rev. Financ. Anal.* **2019**, *65*, 101378. [CrossRef]
79. Pareto, V. Cours d'Économie Politique, Volumes 1–2. In *Oeuvres complètes publiées sous la direction de Giovanni Busino*; Bousquet, G.-H., Busino, G., Eds.; Librairie Droz: Genève, Switzerland, 1964; Volume 1, ISBN 978-2-600-04014-3.
80. Pareto, V. *Manual of Political Economy*; The Macmillan Press: London, UK, 1971; ISBN 978-0333135457.
81. Bachelier, L. Théorie de la spéculation. *Ann. Sci. Ecole Norm. Supérieure* **1900**, *17*, 21–86. [CrossRef]
82. Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys. (Leipzig)* **1905**, *17*, 549–560. [CrossRef]
83. Poincaré, H. Sur le problème des trois corps et les équations de la dynamique. *Acta Math.* **1890**, *13*, A3–A270.
84. Palander, T.F. *Instability in Competition between Two Sellers. Abstracts of Papers Presented at the Research Conference on Economics and Statistics Held by the Cowles Commission at Colorado College*; General Series No. 208, Studies Series No. 21; Colorado College Publications: Colorado Springs, CO, USA, 1936; pp. 53–59.
85. Palander, T.F. Konkurrens och marknadsjämvikt vid duopol och oligopol. I. Fullkomlig marknad och »autonomt» handlande. I. Duopol-oligopol-problemet. *Ekon. Tidskr.* **1939**, *41*, 123–145. [CrossRef]

86. Palander, T.F. Konkurrens och marknadsjämvikt vid duopol och oligopol. I. Fullkomlig marknad och »autonomt» handlande. II. Cournots duopol: Mängdanpassning. *Ekon. Tidskr.* **1939**, *41*, 222–250. [CrossRef]
87. Goodwin, R.M. The nonlinear accelerator and the persistence of business cycles. *Econometrica* **1951**, *19*, 1–17. [CrossRef]
88. Strotz, R.H.; McAnulty, J.C.; Naines, J.B., Jr. Goodwin's nonlinear theory of the business cycle: An electro-analog solution. *Econometrica* **1953**, *21*, 390–411. [CrossRef]
89. Rosser, J.B., Jr. The development of complex oligopoly dynamics theory. In *Oligopoly Dynamics: Models and Tools*; Puu, T., Sushko, I., Eds.; Springer: Berlin/Heidelberg, Germany, 2002; pp. 15–29, ISBN 978-3-642-07742-5. [CrossRef]
90. Lorenz, E.N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **1963**, *20*, 130–141. [CrossRef]
91. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L.A.N.; Stanley, H.E. Econophysics: Financial time series from a statistical physics point of view. *Physica A* **2000**, *279*, 443–456. [CrossRef]
92. Gębarowski, R.; Oświęcimka, P.; Wątorek, M.; Drożdż, S. Detecting correlations and triangular arbitrage opportunities in the Forex by means of multifractal detrended cross-correlations analysis. *Nonlinear Dyn.* **2019**, *98*, 2349–2364. [CrossRef]
93. Wątorek, M.; Drożdż, S.; Oświęcimka, P.; Stanuszek, M. Multifractal cross-correlations between the world oil and other financial markets in 2012–2017. *Energy Econ.* **2019**, *81*, 874–885. [CrossRef]
94. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Signatures of the crypto-currency market decoupling from the Forex. *Future Internet* **2019**, *11*, 154. [CrossRef]
95. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Competition of noise and collectivity in global cryptocurrency trading: Route to a self-contained market. *Chaos* **2020**, *30*, 023122. [CrossRef] [PubMed]
96. Andersen, J.V.; Sornette, D. A mechanism for pockets of predictability in complex adaptive systems. *Europhys. Lett.* **2005**, *70*, 697–703. [CrossRef]
97. Kiyono, K.; Struzik, Z.R.; Yamamoto, Y. Criticality and phase transition in stock-price fluctuations. *Phys. Rev. Lett.* **2006**, *96*, 068701. [CrossRef]
98. Ozun, A.; Contoyiannis, Y.F.; Diakonos, F.K.; Hanias, M.; Magafas, L. Intermittency in stock market dynamics. *J. Trading* **2014**, *9*, 34–41. [CrossRef]
99. Kutner, R.; Ausloos, M.; Grech, D.; Di Matteo, T.; Schinckus, C.; Stanley, H.E. Econophysics and sociophysics: Their milestones & challenges. *Physica A* **2019**, *516*, 240–253. [CrossRef]
100. Jakimowicz, A. Econophysics as a cause of a scientific revolution in mainstream economics. *Acta Phys. Pol. A* **2018**, *133*, 1339–1346. [CrossRef]
101. Nelson, R.H. *Economics as Religion: From Samuelson to Chicago and Beyond with a New Epilogue*; The Pennsylvania State University Press: University Park, PA, USA, 2014; ISBN 978-0271022840.
102. Rapley, J. *Twilight of the Money Gods: Economics as a Religion and How it all Went Wrong*; Simon & Schuster: London, UK, 2017; ISBN 978-1471152740.
103. Spriggs, T.W.; Pritchard, M.W. Scientists' guide to economics. *N. Z. J. Technol.* **1987**, *3*, 173–194.
104. Gell-Mann, M. *The Quark and the Jaguar: Adventures in the Simple and the Complex*, 8nd ed.; W.H. Freeman and Company: New York, NY, USA, 2002; ISBN 0-7167-2725-0.
105. Maxwell, J.C. *Theory of Heat*, 3nd ed.; Longmans, Green, and Co.: London, UK, 1872.
106. Gell-Mann, M.; Lloyd, S. Effective complexity. In *Nonextensive Entropy: Interdisciplinary Applications*; Gell-Mann, M., Tsallis, C., Eds.; Oxford University Press: New York, NY, USA, 2004; pp. 387–398, ISBN 0-19-515976-4.
107. Gell-Mann, M.; Lloyd, S. Information measures, effective complexity, and total information. *Complexity* **1996**, *2*, 44–52. [CrossRef]
108. Ay, N.; Müller, M.; Szkoła, A. Effective complexity and its relation to logical depth. *IEEE Trans. Inf. Theory* **2010**, *56*, 4593–4607. [CrossRef]
109. Ravetz, J.R. Usable knowledge, usable ignorance: Incomplete science with policy implications. *Knowledge* **1987**, *9*, 87–116. [CrossRef]
110. Săvoiu, G.; Simăn, I.I. History and role of econophysics in scientific research. In *Econophysics: Background and Applications in Economics, Finance, and Sociophysics*; Săvoiu, G., Ed.; Academic Press: Oxford, UK, 2013; pp. 3–16. [CrossRef]

111. Rosser, J.B., Jr. The nature and future of econophysics. In *Econophysics of Stock and other Markets. Proceedings of the Econophys-Kolkata II*; Chatterjee, A., Chakrabarti, B.K., Eds.; Springer: Milan, Italy, 2006; pp. 225–234, ISBN 978-88-470-0501-3. [CrossRef]
112. Rosser, J.B., Jr. Debating the role of econophysics. *Nonlinear Dyn. Psychol.* **2008**, *12*, 311–323.
113. Rosser, J.B., Jr. Econophysics and economic complexity. *Adv. Complex Syst.* **2008**, *11*, 745–760. [CrossRef]
114. Rosser, J.B., Jr. Entropy and econophysics. *Eur. Phys. J. Spec. Top.* **2016**, *225*, 3091–3104. [CrossRef]
115. Arthur, W.B. *Complexity and the Economy*; Oxford University Press: New York, NY, USA, 2015; ISBN 978-0-19-933429-2.
116. Holt, R.P.F.; Rosser, J.B., Jr.; Colander, D.C. The complexity era in economics. *Rev. Political Econ.* **2011**, *23*, 357–369. [CrossRef]
117. Schinckus, C. The econophysical turn and economic complexity. *Front. Phys.* **2014**, *2*, 41. [CrossRef]
118. Colander, D.C.; Kupers, R. *Complexity and the Art of Public Policy: Solving Society's Problems from the Bottom Up*; Princeton University Press: Princeton, NJ, USA, 2014; ISBN 978-0-691-15209-7.
119. Abu el Ata, N.; Schmandt, R. *The Tyranny of Uncertainty: A New Framework to Predict, Remediate and Monitor Risk*; Springer: Berlin, Germany, 2016; ISBN 978-3-662-49103-4. [CrossRef]
120. Beinhocker, E.D. *The Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*; Harvard Business School Press: Boston, MA, USA, 2006; ISBN 1-57851-777-X.
121. *Complexity and Evolution: Toward a New Synthesis for Economics*; Wilson, D.S.; Kirman, A. (Eds.) MIT Press: Cambridge, MA, USA, 2016; ISBN 9780262035385.
122. Mullainathan, S.; Thaler, R.H. Behavioral economics. *NBER Work. Pap.* **2000**, *No. 7948*, 1–13. [CrossRef]
123. Prigogine, I.; Stengers, I. *Order out of Chaos: Man's New Dialogue with Nature*; Bantam Books: New York, NY, USA, 1984; ISBN 9780553340822.
124. Prigogine, I.; Stengers, I. *The End of Certainty: Time, Chaos and the New Laws of Nature*; The Free Press: New York, NY, USA, 1997; ISBN 9780684837055.
125. Zambelli, S.; George, D.A.R. *Nonlinearity, Complexity and Randomness in Economics: Towards Algorithmic Foundations for Economics*; Wiley-Blackwell: Chichester, UK, 2012; ISBN 978-1-4443-5031-9.
126. Colander, D.C. *The Complexity Vision and the Teaching of Economics*; Edward Elgar Publishing: Cheltenham, UK, 2000; ISBN 978-1-84064-252-0.
127. Horgan, J. *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*; Helix Books/Addison-Wesley: Reading, MA, USA, 1996; ISBN 978-0201626797.
128. Jakimowicz, A. *Podstawy Interwencjonizmu Państwowego. Historiozofia Ekonomii*; Seria: Współczesna Ekonomia; Wydawnictwo Naukowe PWN: Warsaw, Poland, 2012; ISBN 9788301169831.
129. Jakimowicz, A. Nonlinear dynamical systems theory and economic complexity. *Chaotic Model. Sim.* **2013**, *4*, 657–667. Available online: http://www.cmsim.eu/October_2013_issue.html (accessed on 15 December 2019).
130. Jakimowicz, A. Fundamental sources of economic complexity. *Int. J. Nonlinear Sci. Numer.* **2016**, *17*, 1–13. [CrossRef]
131. Farmer, J.D.; Shubik, M.; Smith, E. Is economics the next physical science? *Phys. Today* **2005**, *58*, 37–42. [CrossRef]
132. McCauley, J.L. Nonintegrability, chaos, and complexity. *Physica A* **1997**, *237*, 387–404. [CrossRef]
133. Orrell, D. *Quantum Economics: The New Science of Money*; Icon Books: Duxford, UK, 2019; ISBN 9781785785085.
134. Goswami, A. *Quantum Economics: Unleashing the Power of an Economics of Consciousness*; Rainbow Ridge Books: Faber, VA, USA, 2015; ISBN 978-1-937907-34-1.
135. Heisenberg, W. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Phys.* **1927**, *43*, 172–198. [CrossRef]
136. Werner, R.A. A lost century in economics: Three theories of banking and the conclusive evidence. *Int. Rev. Financ. Anal.* **2016**, *46*, 361–379. [CrossRef]
137. Piotrowski, E.W.; Sładkowski, J. Trading by quantum rules: Quantum anthropic principle. *Int. J. Theor. Phys.* **2003**, *42*, 1101–1106. [CrossRef]
138. Piotrowski, E.W.; Sładkowski, J. Quantum games in finance. *Quant. Finance* **2004**, *4*, 61–67. [CrossRef]
139. Piotrowski, E.W.; Sładkowski, J. Quantum auctions: Facts and myths. *Physica A* **2008**, *387*, 3949–3953. [CrossRef]

140. Piotrowski, E.W.; Sładkowski, J. Quantum game theoretical frameworks in economics. In *The Palgrave Handbook of Quantum Models in Social Science*; Haven, E., Khrennikov, A., Eds.; Palgrave Macmillan: London, UK, 2017; pp. 39–57, ISBN 978-1-137-49275-3. [CrossRef]
141. Baaquie, B.E. *Quantum Finance: Path Integrals and Hamiltonians for Options and Interest Rates*; Cambridge University Press: Cambridge, UK, 2007; ISBN 978-0-521-71478-5. [CrossRef]
142. Baaquie, B.E. *Interest Rates and Coupon Bonds in Quantum Finance*; Cambridge University Press: Cambridge, UK, 2011; ISBN 9780511808715. [CrossRef]
143. Wendt, A. *Quantum Mind and Social Science: Unifying Physical and Social Ontology*; Cambridge University Press: Cambridge, UK, 2015; ISBN 978-1-107-08254-0.

entropy

MDPI

Article

Tsallis Entropy for Cross-Shareholding Network Configurations

Roy Cerqueti [1,2], Giulia Rotundo [3,*] and Marcel Ausloos [4,5,6]

1 Department of Social and Economic Sciences, Sapienza University of Rome, p.le A. Moro 5, 00185 Roma, Italy; roy.cerqueti@uniroma1.it
2 School of Business, London South Bank University, London SE1 0AA, UK
3 Department of Statistical Sciences, Sapienza University of Rome, p.le A. Moro 5, 00185 Roma, Italy
4 School of Business, College of Social Sciences, Arts, and Humanities, Brookfield, University of Leicester, Leicester LE2 1RQ, UK; marcel.ausloo@ulg.ac.be
5 Group of Researchers for Applications of Physics in Economy and Sociology (GRAPES), Rue de la belle jardinière, 483, Sart Tilman, B-4031 Angleur, Liege, Belgium
6 Department of Statistics and Econometrics, Bucharest University of Economic Studies, Calea Dorobantilor 15-17, 010552 Sector 1 Bucharest, Romania
* Correspondence: giulia.rotundo@uniroma1.it; Tel.: +39-06-4925-5323

Received: 27 May 2020; Accepted: 13 June 2020; Published: 17 June 2020

Abstract: In this work, we develop the Tsallis entropy approach for examining the cross-shareholding network of companies traded on the Italian stock market. In such a network, the nodes represent the companies, and the links represent the ownership. Within this context, we introduce the out-degree of the nodes—which represents the diversification—and the in-degree of them—capturing the integration. Diversification and integration allow a clear description of the industrial structure that were formed by the considered companies. The stochastic dependence of diversification and integration is modeled through copulas. We argue that copulas are well suited for modelling the joint distribution. The analysis of the stochastic dependence between integration and diversification by means of the Tsallis entropy gives a crucial information on the reaction of the market structure to the external shocks—on the basis of some relevant cases of dependence between the considered variables. In this respect, the considered entropy framework provides insights on the relationship between in-degree and out-degree dependence structure and market polarisation or fairness. Moreover, the interpretation of the results in the light of the Tsallis entropy parameter gives relevant suggestions for policymakers who aim at shaping the industrial context for having high polarisation or fair joint distribution of diversification and integration. Furthermore, a discussion of possible parametrisations of the in-degree and out-degree marginal distribution—by means of power laws or exponential functions— is also carried out. An empirical experiment on a large dataset of Italian companies validates the theoretical framework.

Keywords: Tsallis entropy; copula functions; cross-shareholding network; finance

1. Introduction

The presence of interconnections among companies is the ground for the propagation of shocks over the entire industrial structure of a country; see e.g., [1,2]. This evidence has led to a growing number of studies exploring such structure through networks theories; see e.g., [3,4].

In this respect, a single company can be intuitively seen as a network node. The ownership relationship can be represented through a network: there is a (directed) link from a company i to a company j if i holds shares of j. For what concerns the mutual connections among companies, several contexts can be explored on the basis of the topic under investigation. Here, we mention connections

that are driven by technological transfer [5], the presence of personal relationships [6–8], the interlock of directorates [9–12], and capabilities at the organisational level [7,13]. For a survey on this field, see e.g., [14].

We propose a specific focus on the cross-shareholding matrix, which is associated to the directed links, thereby capturing the so-called in-degree and out-degree of each node.

Specifically, the in-degree of a company—say, k_{in}—is the number of companies holding some ownership of the considered node. Such a concept has a clear interpretation on the integration of any given company in its reference industrial and productive environment. Similarly, the out-degree of a company—namely, k_{out}—counts the companies included in the portfolio of the considered node. Thus, k_{out} is associated to diversification, which, in turn, might point to information on the possible reaction of a considered company to markets fluctuations. For the concepts of integration and diversification, we refer the interested reader to [15].

Notice that the so-followed approach is grounded on the existence of a connection—in terms of ownership relations—between two companies. In so doing, we explore diversification and integration—along with market concentration, which is a synthesis of them—as a matter of pure shareholding strategies and through the singular attitude of companies to collect shares of other companies, and at the same time to have shares own by other companies—"other companies", which can be the same being owner and owned (Renault SA, which is part-owned by the French state, owns 43% of Nissan Motor Co, while the Japanese firm has 15% of the French carmaker—but with no voting rights in this case). Within such thinking, the amount of inter company flows leads to a discussion on the size of the connections between companies. In this setting, in-degrees and out-degrees should be reasonably written as sums of percentages of in-flows and out-flows. Thus, the in-degree can be high in both cases, i.e., when there is a large number of existing in-connections with small flows or small values of in-connections with large entities of flows; the same consideration applies also for the out-degree, of course. The numerical dimension of the connections is then lost—even if a new information on the size of the flows is available. Yet, the analysis of such flows is clearly beyond the scopes of the present paper.

While out-degrees are widely explored, for their natural connections with the resilience of an industrial system, see e.g., [16–21], scarce attention has been paid to in-degrees. Let us point to a noticeable contribution on the trade-off between diversification and integration in the analysis of economic crises in [22].

Here, we are concerned by the market concentration—which represents a synthesis of diversification and integration, by means of the entropy of the in-degree and out-degree distributions. The entropy concept allows for understanding the position of the considered industrial structure between the extreme cases of uniform diversification and integration and *a contrario* strong polarisation, with only one company playing the role of the leader.

Furthermore, we also include a deep analysis of the particular features of the distributions through a generalised concept [23,24] of Boltzmann–Gibbs (or equivalently Shannon information [24–26]) entropy. To this end, we move from [27] and deal with the Tsallis entropy for discussing the in- and out-degrees distributions of the companies.

Tsallis entropy—introduced in [23]—has been applied in a number of contexts related to economics and finance; see e.g., the excellent review in [28] and references therein. Most of the time, the studies concern risk or portfolio management [29–34]. Our present report seems to be the first contribution dealing with this powerful instrument in the context of the cross-shareholding matrix for its related network of companies.

Tsallis entropy depends on a (usually real, see a complex case in [35]) parameter, whose interpretation provides relevant information on the shape of the distributions. Indeed, when the parameter is negative (positive), then Tsallis entropy attains its maximum in the highest polarisation case (in the uniform distribution case). Moreover, a negative value of the parameter is associated to a strong relevance of fat tails and rare events; see e.g., [24].

To explore in depth the relationship between integration and diversification, we propose the analysis of the joint distributions between such terms in the relevant cases of independence—i.e., when the stochastic dependence is described by a product copula—and in the maximum level case of positive (negative) dependence—i.e., when the dependence is given by the upper (lower) Frechet bounds copulas. These represent the mathematical bounds of the set of the copulas corresponding to the cases of perfect positive (negative) correlation; see [36]. For a complete description of the concept of copulas and on how it serves as modelling stochastic dependence, see e.g., [37,38] and refer to the Sklar's Theorem [39]. Indeed, Sklar's Theorem provides a reading of the copulas as mathematical functions transforming the marginal distributions of a set of random variables into their joint distribution (see also below).

We consider a high-quality dataset of holdings listed in the Italian Stock Market to validate our theoretical proposal. Such a selection, the Italian Stock Market as reference context, has been driven by data availability. Indeed, the phase of data collection has been particularly challenging, with manual collection procedures and matching among different datasets—see the details in Section 4.1. Of course, data availability is the premise of the data collection procedure. This said, even if it is theoretically easy to reproduce the analysis for all the major markets—like the US and the UK ones, the practical implementation in different contexts requires a non-trivial effort and data availability.

We also propose an extension of the analysis to a wide and universal economic system, where in-degrees are assumed to be synthesised by two parametric functions of either power law or exponential types, while the out-degree distribution obeys a power law; see e.g., [40]. In particular, we have included the parameters of such functions in the calibrating quantities set. Such a proposed extension leads to useful discussions about the assessment of missing links in the cross-shareholding matrix, in line with some literature contributions, like e.g., [41–46].

Some results emerge from our study. The obtained outcomes suggest strategies that should be implemented by policymakers if pursuing a highly polarised industrial structure goal—with a company holding the shares of all the other ones and, at the same time, whose shares are included in the portfolios of the others—or a fair joint distribution of diversification and integration. Such policies are built on the basis of the dependence structure between in-degrees and out-degrees and on enforcing the shapes of their distributions in a proper way.

The remaining part of the paper is organised as follows. Section 2 provides some information on the reference literature on cross-shareholding. Section 3 gives the details of the methodological devices used in the analysis. Section 4.1 provides a description of the dataset employed for the methodological validation and, in particular, the network construction in Section 4.2. Section 5 describes and discusses the obtained findings. Conclusions and comments on policy implications are found in Section 6.

2. Brief Review of the Reference Literature on Cross-Shareholdings

This section provides a list of key papers dealing with cross-shareholdings. Such a list is not exhaustive, but the referred contributions are particularly close to the present study—even if they present remarkable differences. As a premise, we have to state that the framework adopted in this paper is quite new when compared to other papers on the cross-shareholdings.

In [47], a complex networks approach is used for identifying the companies that are central in the information flow and for the control. The coupling among in-degree and out-degree is not examined explicitly, although it intervenes in the empirical estimates of the flow-betweenness and of other centrality measures.

The perspective in Abreu et al. [48] is of an empirical nature, without a precise focus on the relationship between in-degrees and out-degrees, i.e., as integration and diversification measures, respectively.

In [49], the possibility to use cross-shareholdings for achieving the control of companies through intermediaries is examined, but there is again no deeper insight on the relationship between integration and diversification as optimal means toward the considered specific targets.

Vitali et al. [50] offer the analysis of the structure and topology of the transnational ownership network of cross-shareholdings. This is a pretty empirical paper, without further steps in the analysis of the stochastic dependence on integration and diversification.

An analysis of the relevance of the cross-shareholdings in the Japanese markets can be found in [51]. The target of the quoted paper is to understand the role of shareholdings in order to reduce the risk/performance ratio. However, the aim is quite different from the one tackled in this present paper.

In [52], Okabe performs an economic analysis on cross-shareholdings in Japan, where this theme is quite relevant. Trends and implications for the Japanese economic system and related public policies are discussed. In [53,54] the focus is on the presence of the power law, and [55] adds more insights typical of complex networks studies. However, such analyses are mostly performed from the perspective of economics and empirical investigation rather than by proposing novel methods.

The framework of the stochastic dependence among integration and diversification considered in the present paper is close to that in [56], but presently under a wider viewpoint; in [56], one uses a rewiring procedure as methodological instrument.

3. Methodology

This section describes the techniques and the tools used for achieving the targets of the analysis.

3.1. Preliminaries and Notations

First, we introduce the main concepts that are used in the paper.

Given a node $j \in V$, the in-degree $k_{in}(j)$ represents the integration, i.e., the number of companies owning shares of company j. It is defined as follows:

$$k_{in}(j) = \sum_{i=1}^{N} a_{ij}$$

In the same line, given $i \in V$, the out-degree $k_{out}(i)$ represents the diversification, i.e., the number of companies in the portfolio of company i. It is defined as follows:

$$k_{out}(i) = \sum_{j=1}^{N} a_{ij}.$$

k_{in} and k_{out} both have to be considered here as random variables, whose empirical distributions are obtained by considering the real data described in Section 4.1.

The cumulative distribution functions of k_{in} and k_{out} is denoted by $F_{k_{in}} : \mathbb{R} \to [0,1]$ and $F_{k_{out}} : \mathbb{R} \to [0,1]$, respectively. Their joint distribution is denoted by $F_{k_{in},k_{out}} : \mathbb{R}^2 \to [0,1]$.

The generic joint distribution function $F_{k_{in},k_{out}}$ is associated to a bivariate density function. It is discrete, in the empirical case we are treating; the distribution is denoted by $\mathbf{p} = (p_{ij} : i = 1,\dots,n; j = 1,\dots,m)$ such that

$$p_{ij} = Prob(k_{in} = i, k_{out} = j), \qquad \forall\, i,\ j, \tag{1}$$

with

$$\sum_{i,j} p_{ij} = 1.$$

The values of the integers n and m will be properly fixed in the subsequent empirical analysis.

In the sequel, for such a bivariate probability distribution, we compute the Tsallis entropy, usually defined as follows:

$$S_q = \frac{1}{q-1}\left(1 - \sum_{i,j} p_{ij}^q\right), \tag{2}$$

where $q \in \mathbb{R}$ is the Tsallis parameter.

A bivariate copula $C : [0,1]^2 \to [0,1]$ (see e.g., [38]) is a special function that is able to describe the dependence structure between two random variables through the classical Sklar's Theorem (see [39]). We enunciate such a crucial result by employing the notation that is used in the present paper.

Theorem 1. *Sklar's Theorem: there exists a copula* $C : [0,1]^2 \to [0,1]$ *such that, for each* $(s,h) \in \mathbb{R}^2$, *one has*

$$F_{k_{in},k_{out}}(s,h) = C(F_{k_{in}}(s), F_{k_{out}}(h)). \tag{3}$$

If $F_{k_{in}}, F_{k_{out}}$ *are continuous, then C satisfying (3) is unique. Conversely, if C is a copula and* $F_{k_{in}}, F_{k_{out}}$ *are distribution functions, then* $F_{k_{in},k_{out}}$ *in (3) is a bidimensional joint distribution function with marginal distribution functions* $F_{k_{in},k_{out}}$.

According to Theorem 1, copulas describe different types of stochastic dependence that could be found between two random variables. In so doing, one is also capable of providing insights on the nature of the stochastic dependence of tis empirical joint distribution.

We denote by $F^C_{k_{in},k_{out}} : [0,1]^2 \to [0,1]$ the joint distribution function resulting from the application of Sklar's Theorem with a generic copula C, according to the previous Formula (3).

Reasoning behind the Tsallis Entropy

This section is devoted to the justification of the selection of Tsallis entropy as a key methodological measurement device. We provide a comparison between Tsallis entropy and the well-known and largely used Gibbs entropy. In fact, Tsallis entropy is known to exhibit substantial strengths when compared to the Gibbs one. To support this statement, we proceed under both technical and applied perspectives.

From a purely mathematical point of view, Tsallis entropy represents a generalisation of the Gibbs entropy. Indeed, Tsallis entropy, formally a fractional exponential approach, depends on an often real (but see [35]) parameter q, introduced in Equation (2); when $q \to 1$, the Tsallis entropy collapses to the Gibbs entropy. Hence, the Tsallis entropy is able to capture several aspects that are not covered by the Gibbs entropy—all of those aspects related to a not unitary parameter q. In our context, the main results will be seen to be related to negative q values. Thus, it is clear that the Gibbs entropy would not allow us to provide a deep understanding of the nature of the stochastic dependence between in-degree and out-degree distributions.

In the context of applied science, we may recall that classical statistical mechanics of macroscopic systems in equilibrium is based on Boltzmann's principle and Gibbs entropy. However, Boltzmann–Gibbs statistical mechanics and standard thermodynamics present serious difficulties or anomalies for non-equilibrium, open, non-ergodic, non-mixing, systems, and for those that exhibit memory retention. Within a long list, we might mention systems that involve long-range interactions (see e.g., [57,58]), non-Markovian stochastic processes, like financial markets (see e.g., [59–64]), dissipative systems in a phase space that has some underlying looking (multi)fractal-like structure (see e.g., [65]), like many open social systems, all hardly having an additive property (see e.g., [66]).

In brief, Tsallis theory provides a better thermo-statistical description than the standard Boltzmann–Gibbs formalism, because the Tsallis fractional exponential approach allows for encompassing cases of non-equilibrium and dissipative systems into hard core statistical mechanics principles.

3.2. Outline of the Analysis

The analysis is carried out in two main directions.

First, we compute and discuss the Tsallis entropy of the joint distribution $F^C_{k_{in},k_{out}}$, which is obtained by applying the Sklar's Theorem with some specific copulas C. In so doing, we provide useful insights on the behaviour of the cross-shareholding system under different scenarios of interactions between in-degrees and out-degrees.

In particular, we address the corner cases of maximal positive and negative dependence, and the case of independence. Such cases correspond to the following copulas:

- Product (independence)
$$C_P(u,v) = uv \tag{4}$$
- Lower Frechet (maximal negative dependence) and Upper Frechet (maximal positive dependence)
$$C_{LF}(u,v) = \max\{u+v-1, 0\}, \qquad C_{UF}(u,v) = \min\{u,v\}. \tag{5}$$

Second, we discuss the sensitivity analysis of the in- and out-degrees distributions when they are properly parametrised, by means of the Tsallis entropy.

In this respect, while the literature points out the ubiquitous presence of a power law for the out-degree distribution, the in-degree is much less studied. However, the main theoretical functions that can be suitably used for approximating the in-degree empirical distribution are either the power law or the exponential law (see [27] and references therein contained). Therefore, on one side, we consider the marginal distribution of the out-degree as following a power law; on the other side, we consider two cases, power law or exponential function for modelling the in-degree empirical distribution.

The power law and the exponential law for a generic discrete random variable X are defined, as follows:

- Power law:
$$Prob(X = x) = ax^{-k}, \tag{6}$$
where $x \geq 0$, $a > 0$ is a normalising constant and $k > 0$.
- Exponential law:
$$Prob(X = x) = ae^{-kx}, \tag{7}$$
where $x \geq 0$, $a > 0$ is a normalising constant and $k > 0$.

Thereafter, we implement the sensitivity analysis in three cases:

(A) under the hypothesis of k_{out} described by a power law as in (6) and k_{in} has its empirical distribution, the power law exponent k is allowed to change and is treated as a parameter;
(B) under the hypothesis of k_{in} power law as in (6) and k_{out} empirical: the power law exponent k is allowed to change and is treated as a parameter; and,
(C) under the hypothesis of k_{in} exponential as in (7) and k_{out} empirical: the parameter k in the exponential is allowed to change, as any parameter does.

Thus, in each case, there are two parameters: q for the Tsallis entropy and k for the power law or exponential. In all cases, we have employed the three copulas C_I, C_{LF}, and C_{UF} introduced in (4) and (5) for deriving the joint probability distribution, according to Theorem 1.

4. The Network

Here, we present the cross-shareholding network that are used in the analysis.

4.1. The Data

We consider the data already used in [27,67]. The dataset gathers data of the Milan Stock Exchange (MIB30) on 10 May 2008. First, data were obtained through the CONSOB database. For each company j, an informative page is shown, which contains the information on the holdings, which is the list of companies i, traded in the same market, which the shares of j are sold to. The set of all of the couples (i, j) constitutes the matrix of cross-holdings. CONSOB is the major surveillance body for the Italian Stock Market. CONSOB verifies the transparency of market operations; it has the power to stop the

market in case of excess of losses/returns; CONSOB controls the proper disclosure of information. Unfortunately, only CONSOB records the holdings above 2%. Therefore, the data were cross-checked through the Bureau Van Dijk platform.

Differently from the database of prices of the shares, there is no command that allows for downloading all of the data at once. The data gathering requires manual opening of each file, and manual storing of the relevant information. Moreover, the way in which the companies are named is not uniform: sometimes, shortcuts are used instead of the original extended names. Therefore, the data collection cannot be done automatically "blind folded". The data also have to be gathered at a selected date: it is like taking a picture of the actual situation of the market on a specific day. The time needed for gathering the data and finalising the sample is quite long, since the data were manually cross-checked with other databases. Notice that the data on banks were cross-checked with the BANKSCOPE database, which, as the name suggests, is specifically focusing on banks, hence not reporting data on other companies.

On the other side, AIDA provides some complementary information, since AIDA contains information on all companies—apart from banks. The cross-checking was necessary to be sure that we include in the database all ownerships due to investments and all cross-relationships among companies—yet excluding some very minor ones due to the management of portfolios by mutual funds. Alas, some companies had very incomplete data. Finally, the resulting sample contains the cross-holdings of 247 stocks of companies. They represent 94% of the total amount of MTA segment (MTA stands for Borsa Italiana's Main Market, that is Italian Main Stock Market. MTA is a regulated market subject to stringent requirements in line with the expectations of professional and private investors.). The sample corresponds to 95.22% of the total capitalisation on that date, May 10th, 2008, which nevertheless makes the analysis quite suitable for a whole outlook about the links among the most relevant traded companies. Notice that the total number of cross-ownership is 243, thus less than the number of companies. In fact, there are companies traded in the Italian Stock Market, which do not buy or sell shares of other companies traded in the Italian market.

The vast majority of holdings is due either to industrial purposes or to an internal organisation of companies: for instance, the energy company ENI owns shares of two other companies, SAIPEM and SNAM RETE GAS, with a specific focus on gas delivery management. Another example is given by the financial company IFIL, which is managing the financial parts of FIAT (now merged in FCA) and JUVENTUS (football club). In turn, IFI PRIV owns the "privileged" part of IFI, belonging to the Agnelli family.

The number of companies holding shares of k other companies decreases sharply as k increases. In fact, there are 72 companies owning shares of only one other company; 16 companies owning shares of two other companies; only seven and six companies are owning shares of three and four other companies, respectively. There are only zero or one companies holding shares of six or more other companies; the maximum ownership in 19 companies is due to the insurance company "Assicurazioni Generali", which uses ownership as part of its institutional mission. The clear prevalence of shareholders who hold shares of only one or a few other companies has been detected in other datasets [68,69].

A symmetric question holds: which is the number h of companies to which a specific company has sold shares? According to the literature on this topic, the question is less popular than the previous one. In our specific dataset, the maximum value of h is 10; there are 84 companies that sell their shares to only one company; 29 companies sell their shares to two companies; 15 are selling to three; only five companies have sold to four other companies, and another five are selling to more than four companies. Therefore, roughly speaking, the very prevailing behaviour is the relation through a sale of shares to only one other company in the market.

4.2. Construction of the Network

The firms are represented by the nodes of an unweighted network. We collect them in a set $V = \{1, \cdots, N\}$. If a company j is held by company i, then there is a directed link from i to j. The links are collected in a set E. In so doing, we create a network (V, E), whose adjacency matrix $A = (a_{ij})_{i,j \in V}$ is a $N \times N$ matrix, such that $a_{ij} = 1$ if $(i, j) \in E$ and $a_{ij} = 0$ otherwise.

The insulated nodes have been removed from the analysis; the giant component and the small connected components being kept, the network is made of 158 nodes, hence the adjacency matrix is 158×158.

5. Results and Discussion

Here, we report the results of the analysis, along with a discussion of these.

As a premise, we set $n = 10$ and $m = 19$, in accord to the maximum values of k_{in} and k_{out}, which are observed in the empirical dataset.

It is immediate to observe that the Tsallis entropy S_q in (2) is strictly decreasing with respect to the parameter $q \in \mathbb{R}$, with an asymptotic behaviour being given by

$$\lim_{q \to -\infty} S_q = +\infty; \qquad \lim_{q \to +\infty} S_q = 0.$$

This said, we restrict our graphical representations of the behaviour of the Tsallis entropy with respect to q to a small interval, including zero, for a better visualisation of the outcomes.

Figure 1 shows the behaviour of the values of the Tsallis entropy as the parameter q varies, in the three cases of joint distributions, $F^C_{k_{in},k_{out}}$ with $C = C_P, C_{LF}, C_{UF}$, as in (4) and (5)—in the upper, middle, and lower panel, respectively.

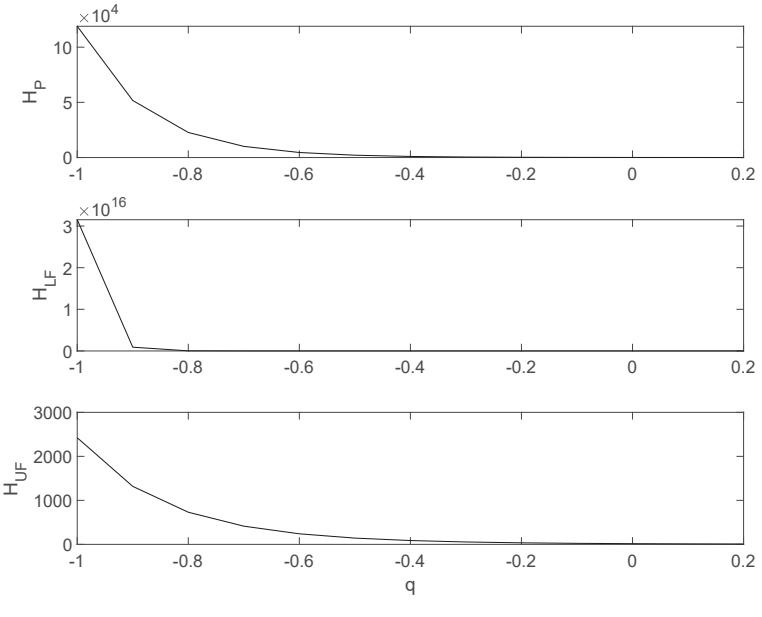


Figure 1. The Tsallis entropy $H = H_P, H_{LF}, H_{UF}$ as a function of q, in the cases of copula $C = C_P, C_{LF}, C_{UF}$ as in (4) and (5)—upper, middle and lower panel, respectively.

The Upper Frechet bound is the one with the slowest decrease; it is substantially flat with respect to the other cases. Moreover, the Lower Frechet bound is associated to very high values of the Tsallis

entropy when q approaches -1; such a case is also the one presenting a very rapid collapse of S_q as q increases.

An interpretation of these results is in order. The predominance—to be intended as the highest values of Tsallis entropy—of the case of copula C_{LF} means that the joint probability between in-degree and out-degree is highly polarised when there is a perfectly negative correlation between such quantities. This is particularly true when q is negative; hence, the fat tails of the distribution do play a key role in determining such an outcome. The results change when moving to the independence and the maximum level of positive dependence. In particular, the Upper Frechet case corresponds to the highest similarity between the uniform case and the considered joint probability distribution.

The policymaker should then force the in-degrees and out-degrees of the companies to exhibit similar patterns—i.e., integration and diversification should coincide—when the target is a homogeneous industrial structure; *a contrario*, integration and diversification should be forced to exhibit a large discrepancy, if the aim of the policymaker is to foster the predominance of a company over the others.

We now deal with the cases (A), (B) and (C) described in the previous section, which are related to different parametrizations of the in- and out-degree marginal distributions.

(A) k_{out} is described by a power law as in (6), while k_{in} is taken with its empirical distribution.

Figure 2 shows the Tsallis entropy as a function of its parameter q and the exponent of the power law k for the cases of copula $C = C_P, C_{LF}, C_{UF}$ as in (4) and (5)—upper, middle and lower panel, respectively.

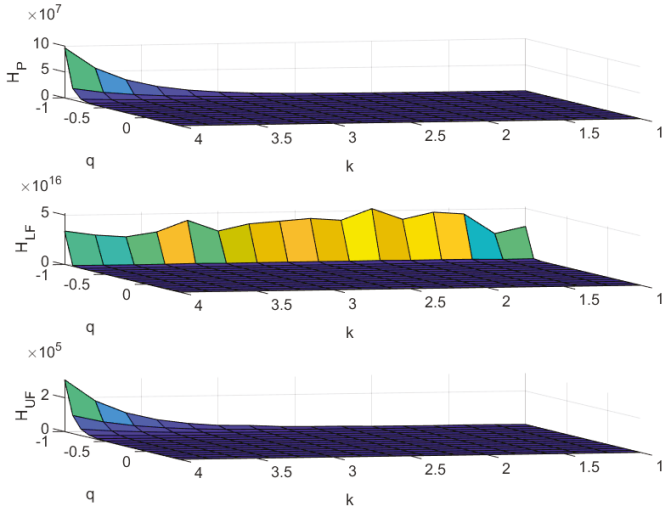


Figure 2. Tsallis entropy as a function of its parameter q and the exponent of the power law k for the out-degree. All the cases of copula $C = C_P, C_{LF}, C_{UF}$, as in (4) and (5)—upper, middle, and lower panel, respectively—are reported.

In all cases, we observe that Tsallis entropy is decreasing as k decreases and q increases. The growth toward infinity is very rapid as q approaches -1. This behaviour is more evident when k assumes large values, i.e., when the probability that k_{out} assumes a large value is particularly small—and when in-degree and out-degree are highly positively correlated or are uncorrelated. If in-degree and out-degree have the maximum level of negative correlation, then the same behaviour seems to be rather independent from the value of the power law parameter. The apparent crests on H_{LF} actually correspond to very high values of H_{LF}; furthermore, the case with C_{LF} is confirmed to have the highest level of Tsallis entropy.

We can read the results by stating that the joint probability of in- and out-degree shows a high level of polarisation in the presence of a perfectly negative correlation. Such a finding does not depend on the specific parametrisation of the out-degree through a power law. Differently, we see polarisation only for k large enough when the cases of stochastic independence or perfectly positive correlations are considered. This behaviour is amplified for negative q values, hence giving credit to the action of the fat tails of the distribution in determining it.

The policymaker has now two devices for shaping the considered industrial structure. Beyond dealing with the dependence between diversification and integration—we refer to the comments stated above for Figure 1—she/he can also force the individual companies to form specific out-degrees distributions. Indeed, in the particular cases of independence and maximum positive correlation, one can obtain some polarisation by shaping the out-degrees in order to obtain a low probability of having large values—i.e., by taking large values of the parameter k. Such an action is not needed when the correlation between in-degree and out-degree is of perfectly positive type.

(B) k_{in} is a power law as in (6) and k_{out} has its empirical distribution.

Figure 3 presents the values of the Tsallis entropy as a function of q and k. Additionally, in this case, copulas C_P, C_{LF}, C_{UF}, as in (4) and (5), are in the upper, middle, and lower panel, respectively. For a better visualisation of the results, we only display when $q < 0$.

The behaviour of the Tsallis entropy is quite similar to that of case (A), with four noticeable exceptions. Firstly, the scales are completely different. The values of the Tsallis entropy are much higher in this case than in case (A). Secondly, to appreciate the decreasing behaviour of the Tsallis entropy, one needs to take q close to -2, instead of $q = -1$, as in the previous case. Thirdly, we observe a deviation in the case of perfectly negative correlation, with two lines of local maxima occurring at $q \simeq -2$, for $k = 2.7$ and $k = 1.8$ (see the arrows in Figure 3). Fourthly, the crest appearing in the case of perfectly negative correlation is much more jagged than in case (A).

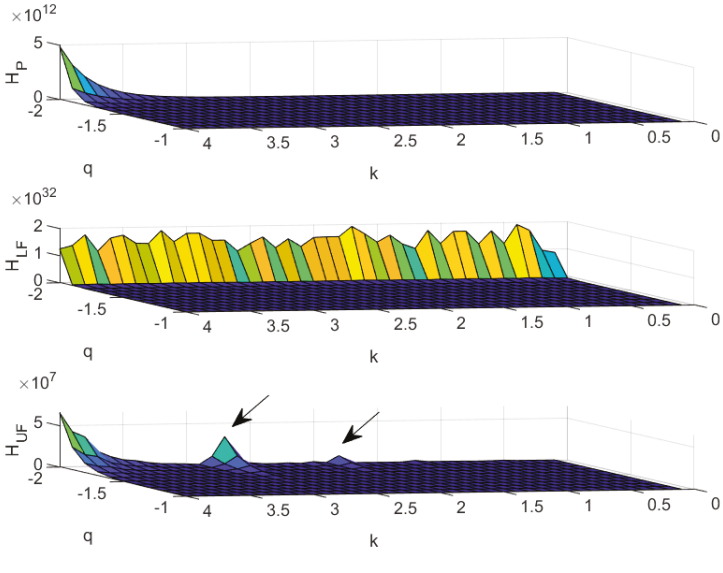


Figure 3. Tsallis entropy as function of its parameter q and the exponent of the power law k for the in-degree. The cases of copulas C_P, C_{LF}, C_{UF} as in (4) and (5) are presented in the upper, middle, and lower panel, respectively.

The similarities between cases (A) and (B) ensure that all of the comments raised for (A) remain valid also for this case (B). The presence of local maxima and the jagged crest do point to the questionability of the power law parameter as a device for controlling the polarisation of the joint

distribution between in-degree and out-degree when the value of q is at its minimum. This is particularly evident for the case of perfectly negative correlation—i.e., in the case of jagged crest—while an action for properly calibrating the parameter $k \simeq -2.7$ and $\simeq 1.8$ remains possible for the case of perfectly positive correlation.

(C) k_{in} has an exponential distribution as in (7) and k_{out} has its empirical distribution.

The upper, middle, and lower panel of Figure 4 display the Tsallis entropy as a function of q and k, for copulas C_P, C_{LF}, C_{UF} as in (4) and (5), respectively; for a clear view of the behaviour of the surface, we only present $q < 0$.

As for (B), the behaviour of Tsallis entropy is also analogous to the one observed for (A), but with three main differences. Indeed, the decreasing behaviour of the Tsallis entropy can be properly visualised for q close to -0.8 (it was -1 and -2 in cases (A) and (B), respectively); moreover, the crest appearing in the middle panel at low values of q is more jagged here than in (A); finally, the minimum value of q appearing in Figure 4 is -0.8 instead of -1 (case A)) and -2 (case B)).

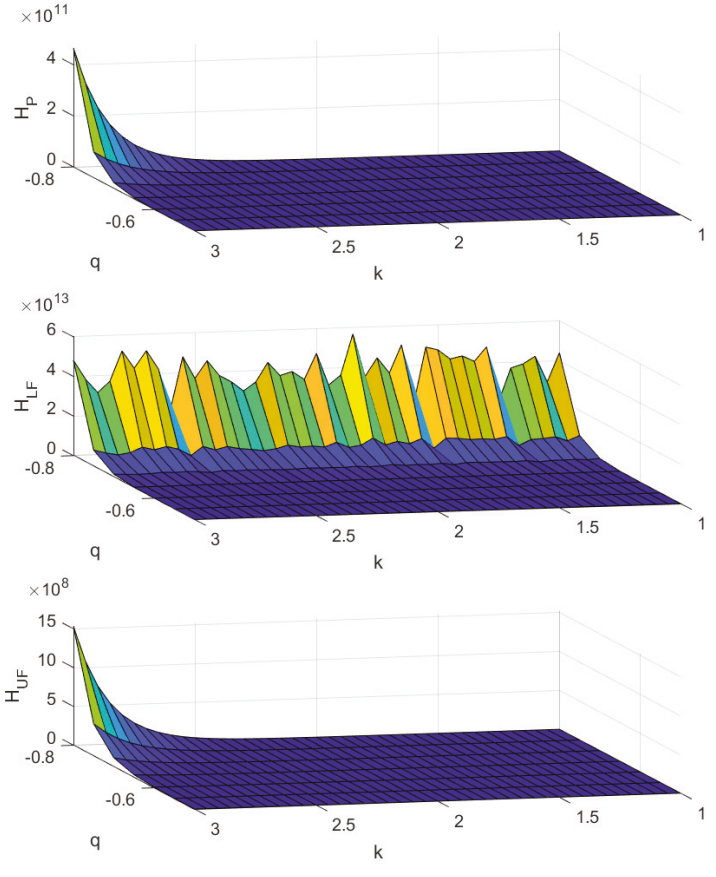


Figure 4. Tsallis entropy as a function of parameter q and k for describing the exponential decrease of the in-degree. The cases of copulas C_P, C_{LF}, C_{UF}, as in (4) and (5), are described in upper, middle and lower panel, respectively.

Some relevant insights can be derived by comparing the three cases (A), (B), and (C). When the desired target is to shape the cross-shareholding network for a polarised situation—with a company holding the widest part of shares of the others and, at the same time, whose shares are in the portfolios of the other companies—then one has to impose a perfectly negative dependence between the in-degree

and the out-degree. Moreover, one has also to shape the distribution of the in-degree as a power law; this means that the probability of having a high in-degree value has to be lower than that of having a low in-degree value. Lastly, the joint distribution between in-degree and out-degree should include also the presence of fat tails, so that one can employ the informative content of the Tsallis entropy in the case of large negative value of q. Under the conditions described above, the Tsallis entropy attains its highest value—see case (B), middle panel. Differently, by imposing the maximum level of positive dependence and a power law behaviour on the out-degree distribution, with a small value of the parameter k, one pursues the objective of shaping the industrial structure towards a more uniform integration and diversification; see case (A), lower panel.

6. Conclusions and Policy Implications

To conclude, we can offer some general remarks on policy implications.

The starting point of the analysis is to describe the industrial structure of a country—in terms of market integration and diversification and, consequently, of concentration. In this respect, the policy makers might aim at fostering the competition in the market or, conversely, at shaping the market for having a leading company.

This theme is of paramount relevance for policy makers. Indeed, the interest of regulatory authorities in the raise of concentration is witnessed by its explicit insertion in official documents. For instance, the study of the classical Herfindahl-Hirschman index (HHi)—which is a relevant measure of market concentration—plays a significant role in the assessment of possible enforcement of US antitrust laws [70]. Since 1982, the Merger Guidelines by the U.S. Department of Justice and the Federal Trade Commission [71] have provided an indication for the identification of post merger markets as "unconcentrated", mildly concentrated, or highly concentrated based on the value of HHi. For a more scientific perspective, we refer e.g., to [56,72]. In this respect, we also mention [22], who have shown that some peculiar combinations of integration and diversification might lead industrial structures to be more vulnerable to financial fluctuations.

Author Contributions: Conceptualization: G.R.; Methodology: R.C.; formal analysis: R.C.; data investigation: G.R.; supervision writing: M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors thank Anna Maria D'Arcangelis for providing data and for many fruitful discussions on the policy implications of the analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Contreras, M.G.A.; Fagiolo, G. Propagation of economic shocks in input-output networks: A cross-country analysis. *Phys. Rev. E* **2014**, *90*, 062812. [CrossRef] [PubMed]
2. Ohnishi, T.; Takayasu, H.; Takayasu, M. Hubs and authorities on Japanese inter-firm network: Characterisation of nodes in very large directed networks. *Prog. Theor. Phys. Suppl.* **2009**, *179*, 157–166. [CrossRef]
3. Luo, J. The power-of-pull of economic sectors: A complex network analysis. *Complexity* **2013**, *18*, 37–47. [CrossRef]
4. Maluck, J.; Donner, R.V. A network of networks perspective on global trade. *PLoS ONE* **2015**, *10*, e0133310. [CrossRef] [PubMed]
5. Ferraro, G.; Iovanella, A. Technology transfer in innovation networks: An empirical study of the Enterprise Europe Network. *Int. J. Eng. Bus. Manag.* **2017**, *9*, 1–14. [CrossRef]
6. Gulati, R.; Westphal, J.D. Cooperative or controlling? The effects of CEO-board relations and the content of interlocks on the formation of joint ventures. *Adm. Sci. Q.* **1999**, *44*, 473–506. [CrossRef]
7. Ceptureanu, S.I.; Ceptureanu, E.G.; Marin, I. Assessing role of strategic choice on organisational performance by Jacquemin- Berry entropy index. *Entropy* **2017**, *19*, 448. [CrossRef]
8. Ferraro, G.; Iovanella, A. Organizing collaboration in inter-organisational innovation networks, from orchestration to choreography. *Int. J. Eng. Bus. Manag.* **2015**, *7*, 7–24. [CrossRef]

9. Bellenzier, L.; Grassi, R. Interlocking directorates in Italy: Persistent links in network dynamics. *J. Econ. Interact. Coord.* **2014**, *9*, 183–202. [CrossRef]
10. Chapelle, A.; Szafarz, A. Controlling Firms Through the Majority Voting Rule. *Physica A* **2005**, *355*, 509–529. [CrossRef]
11. Croci, E.; Grassi, R. The economic effect of interlocking directorates in Italy: New evidence using centrality measures. *Comput. Math. Org. Theory* **2014**, *20*, 89–112. [CrossRef]
12. Rotundo, G.; D'Arcangelis, A.M. Network analysis of ownership and control structure in the Italian Stock market. *Adv. Appl. Stat. Sci.* **2010**, *2*, 255–273.
13. Ceptureanu, E.G.; Ceptureanu S.I. Popescu, D. Relationship between Entropy, Corporate Entrepreneurship and Organisational Capabilities in Romanian Medium Sized Enterprises. *Entropy* **2017**, *19*, 412. [CrossRef]
14. Weber, S.; Weske, K. The joint impact of bankruptcy costs, fire sales and cross-holdings on systemic risk in financial networks. *Probab. Uncertain. Quant. Risk* **2017**, *2*, 9. [CrossRef]
15. Aoyama, H.; Fujiwara, Y.; Ikeda, Y.; Iyetomi, H.; Souma, W. *Econophysics and Companies: Statistical Life and Death in Complex Business Networks*; Cambridge University Press: Cambridge, UK, 2010.
16. Delpini, D.; Battiston, S.; Riccaboni, M.; Gabbi, G.; Pammolli, F.; Caldarelli, G. Evolution of controllability in interbank networks. *Sci. Rep.* **2013**, *3*, 1626. [CrossRef] [PubMed]
17. Clementi, F.; Gallegati, M. Pareto's law of income distribution: Evidence for Germany, the United Kingdom, and the United States. In *Econophysics of Wealth Distributions*; Chatterjee, A., Yarlagadda, S., Chakrabarti, B.K., Eds.; Springer: Milan, Italy, 2005; pp. 3–14.
18. Gao, J.; Barzel, B.; Barabasi, A.L. Universal resilience patterns in complex networks. *Nature* **2016**, *530*, 307–312. [CrossRef] [PubMed]
19. Iori, G.; De Masi, G.; Precup, O.V.; Gabbi, G.; Caldarelli, G. A network analysis of the Italian overnight money market. *J. Econ. Dyn. Control* **2008**, *32*, 259–278. [CrossRef]
20. Newman, M.; Barabasi, A.L.; Watts, D.J. *The Structure and Dynamics of Networks*; Princeton University Press: Princeton, NJ, USA, 2011.
21. Soramaki, K.; Bech, M.L.; Arnold, J.; Glass, R.J.; Beyeler, W.E. The topology of interbank payment flows. *Phys. A Stat. Mech. Appl.* **2007**, *379*, 317–333. [CrossRef]
22. Elliott, M.; Golub, B.; Jackson, M.O. Financial networks and contagion. *Am. Econ. Rev.* **2014**, *104*, 3115–3153. [CrossRef]
23. Tsallis, C. Possible generalisation of Boltzmann-Gibbs Statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
24. Maszczyk, T.; Duch, W. Comparison of Shannon, Renyi and Tsallis Entropy Used in Decision Trees. In *Artificial Intelligence and Soft Computing—ICAISC 2008*; Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5097, pp. 643–651.
25. Rotundo, G.; Ausloos, M. Complex-valued information entropy measure for networks with directed links (digraphs). Application to citations by community agents with opposite opinions. *Eur. Phys. J. B* **2013**, *86*, 169. [CrossRef]
26. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, USA, 1949.
27. Cerqueti, R.; Rotundo, G.; Ausloos, M. Investigating the configurations in cross-shareholding: A joint copula-entropy approach. *Entropy* **2017**, *20*, 134. [CrossRef]
28. Tsallis, C. Economics and Finance: q-Statistical Stylized Features Galore. *Entropy* **2017**, *19*, 457. [CrossRef]
29. Ausloos, M.; Ivanova, K. Dynamical model and nonextensive statistical mechanics of a market index on large time windows. *Phys. Rev. E* **2003**, *68*, 046122. [CrossRef]
30. Ausloos, M.; Miskiewicz, J. Introducing the q-Theil index. *Braz. J. Phys.* **2009**, *39*, 388–395.
31. Devi, S. Financial portfolios based on Tsallis relative entropy as the risk measure. *J. Stat. Mech. Theory Exp.* **2019**, *2019*, 093207. [CrossRef]
32. Maasoumi, E.; Racine, J. Entropy and predictability of stock market returns. *J. Econ.* **2002**, *107*, 291–312. [CrossRef]
33. Vințe, C.; Smeureanu, I.; Furtună, T.-F.; Ausloos, M. An Intrinsic Entropy Model for Exchange-Traded Securities. *Entropy* **2019**, *21*, 1173. [CrossRef]
34. Zhou, R.; Cai, R.; Tong, G. Applications of entropy in finance: A review. *Entropy* **2013**, *15*, 4909–4931. [CrossRef]

35. Wilk, G.; Włodarczyk, Z. Tsallis distribution with complex nonextensivity parameter q. *Phys. A Stat. Mech. Appl.* **2014**, *413*, 53–58. [CrossRef]
36. Frechet, M. Remarques au sujet de la note précédente. *C. R. Acad. Sci. Paris* **1958**, *246*, 2719–2720.
37. Joe, H. *Multivariate Models and Multivariate Dependence Concepts*; CRC Press: London, UK; New York, NY, USA, 1997.
38. Nelsen, R.B. *An Introduction to Copulas*; Springer: New York, NY, USA, 1999.
39. Sklar, M. Fonctions de répartition à *n* dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **1959**, *8*, 229–231.
40. Caldarelli, G. *Scale-Free Networks: Complex Webs in Nature and Technology*; Oxford University Press: Oxford, UK, 2007.
41. Gandy, A.; Veraart, L.A.M. A Bayesian methodology for systemic risk assessment in financial networks. *Manag. Sci.* **2016**, *63*, 4428–4446. [CrossRef]
42. Cimini, G.; Serri, M. Entangling credit and funding shocks in interbank markets. *PLoS ONE* **2016**, *11*, e0161642. [CrossRef] [PubMed]
43. Cinelli, M.; Ferraro, G.; Iovanella, A. Rich-club ordering and the dyadic effect: Two interrelated phenomena. *Phys. A Stat. Mech. Appl.* **2018**, *490*, 808–818. [CrossRef]
44. Cinelli, M.; Ferraro, G.; Iovanella, A. Structural bounds on the dyadic effect. *J. Complex Netw.* **2017**, *5*, 694–711. [CrossRef]
45. Cinelli, M.; Iovanella, A.; Ferraro, G.; Rotundo, G. Assessing the impact of incomplete information on the resilience of financial networks. *Ann. Oper. Res.* **2019**. Available online: https://doi.org/10.1007/s10479-019-03306-y (accessed on 1 January 2020). [CrossRef]
46. D'Arcangelis, A.M.; Rotundo, G. Systemic Risk of Non Performing Loans Market. The Italian case. *J. Appl. Quant. Methods* **2019**, *14*. Available online: http://jaqm.ro/issues/volume-14,issue-1/0_A_G.PHP (accessed on 1 January 2020).
47. D'Errico, M.; Grassi, R.; Stefani, S.; Torriero, A. Shareholding Networks and Centrality: An Application to the Italian Financial Market. In *Networks, Topology and Dynamics*; Naimzada, A.K., Stefani, S., Torriero, A., Eds.; Lecture Notes in Economics and Mathematical Systems; Springer: Berlin/ Heidelberg, Germany, 2009; Volume 613, pp. 215–228.
48. Abreu, M.P.; Grassi, R.; Del-Vecchio, R.R. Structure of control in financial networks: An application to the Brazilian stock market. *Phys. A Stat. Mech. Appl.* **2019**, *522*, 302–314. [CrossRef]
49. Bebchuk, L.A.; Kraakman, R.; Triantis, G. Stock pyramids, cross-ownership, and dual class equity: The mechanisms and agency costs of separating control from cash-flow rights. In *Concentrated Corporate Ownership*; Morck, R.K., Ed.; University of Chicago Press: Chicago, IL, USA, 2017; pp. 295–318.
50. Vitali, S.; Glattfelder, J.B.; Battiston, S. The network of global corporate control. *PLoS ONE* **2011**, *6*, e25995. [CrossRef]
51. Lichtenberg, F.R.; Pushner, G.M. Ownership structure and corporate performance in Japan. *Jpn. World Econ.* **1994**, *6*, 239–261. [CrossRef]
52. Okabe, M. *Cross Shareholdings in Japan. A New Unified Perspective of the Economic System*; Edward Elgar: Cheltenham, UK, 2002.
53. Souma, W.; Fujiwara, Y.; Aoyama, H. Change of ownership networks in Japan. In *Practical Fruits of Econophysics*; Takayasu, H., Ed.; Springer: Tokyo, Japan, 2006; pp. 307–311.
54. Garlaschelli, D.; Battiston, S.; Castri, M.; Servedio, V.; Caldarelli, G. The scale-free topology of market investments. *Phys. A Stat. Mech. Appl.* **2005**, *350*, 491–499. [CrossRef]
55. Chang, X.; Wang, H. Cross-Shareholdings Structural Characteristic and Evolution Analysis Based on Complex Network. *Discrete Dyn. Nat. Soc.* **2017**, *2017*, 5801386.
56. Rotundo, G.; D'Arcangelis, A.M. Network of companies: An analysis of market concentration in the Italian stock market. *Qual. Quant.* **2014**, *48*, 1893–1910. [CrossRef]
57. Barbosa, C.S.; Caraballo, R.; Alves, L.R.; Hartmann, G.A.; Beggan, C.D.; Viljanen, A.; Ngwira, C.M.; Papa, A.R.R.; Pirjola, R.J. The Tsallis statistical distribution applied to geomagnetically induced currents. *Space Weather* **2017**, *15*, 1094–1101. [CrossRef]
58. Petroni, F.; Ausloos, M. High frequency (daily) data analysis of the Southern Oscillation Index. Tsallis nonextensive statistical mechanics approach. *Eur. Phys. J. Spec. Top.* **2007**, *143*, 201–208. [CrossRef]
59. Bila, L.; Grech, D.; Podhajska, E. Methods of Non-Extensive Statistical Physics in Analysis of Price Returns on Polish Stock Market. *Acta Phys. Polonica A* **2016**, *129*, 986–992. [CrossRef]

60. Jakimowicz, A. The Role of Entropy in the Development of Economics. *Entropy* **2020**, *22*, 452. [CrossRef]
61. Queiros, S.D.; Moyano, L.G.; De Souza, J.; Tsallis, C. A nonextensive approach to the dynamics of financial observables. *Eur. Phys. J. B* **2007**, *55*, 161–167. [CrossRef]
62. Rak, R.; Drożdż, S.; Kwapień, J. Nonextensive statistical features of the Polish stock market fluctuations. *Phys. A Stat. Mech. Appl.* **2007**, *374*, 315–324. [CrossRef]
63. Rak, R.; Drożdż, S.; Kwapień, J.; Oswiecimka, P. Stock returns versus trading volume: Is the correspondence more general? *Acta Phys. Pol. B* **2013**, *44*, 2035–2050. [CrossRef]
64. Ruseckas, J.; Gontis, V.; Kaulakys, B. Nonextensive statistical mechanics distributions and dynamics of financial observables from the nonlinear stochastic differential equations. *Adv. Complex Syst.* **2012**, *15*, 1250073. [CrossRef]
65. Pavlos, G.P.; Karakatsanis, L.P.; Iliopoulos, A.C.; Pavlos, E.G.; Tsonis, A.A. Non-extensive statistical mechanics: Overview of theory and applications in seismogenesis, climate, and space plasma. In *Advances in Nonlinear Geosciences*; Tsonis, A., Ed.; Springer: Cham, Switzerland, 2018; pp. 465–495.
66. Gell-Mann, M.; Tsallis, C. (Eds.) *Nonextensive Entropy: Interdisciplinary Applications*; Oxford University Press on Demand: Oxford, UK, 2004.
67. Rotundo, G.; D'Arcangelis, A.M. Ownership and control in shareholding networks. *J. Econ. Interact. Coord.* **2010**, *5*, 191–219. [CrossRef]
68. Li, H.; Fang, W.; An, H.; Yan, L. The shareholding similarity of the shareholders of the worldwide listed energy companies based on a two-mode primitive network and a one-mode derivative holding-based network. *Phys. A Stat. Mech. Appl.* **2014**, *415*, 525–532. [CrossRef]
69. Li, H.; An, H.; Gao, X.; Huang, J.; Xu, Q. On the topological properties of the cross-shareholding networks of listed companies in China: Taking shareholders' cross-shareholding relationships into account. *Phys. A Stat. Mech. Appl.* **2014**, *406*, 80–88. [CrossRef]
70. US Mergers Guidelines. Available online: http://www.stanfordlawreview.org/online/obama-antitrust-enforcement (accessed on 13 June 2020).
71. Horizontal Merger Guidelines. Available online: https://www.justice.gov/atr/horizontal-merger-guidelines-0 (accessed on 13 June 2020).
72. Crane, D.A. Has the Obama Justice Department reinvigorated antitrust enforcement? *Stanf. Law Rev. Online* **2012**, *65*, 13–20.

entropy

MDPI

Article

Long-Range Dependence in Financial Markets: A Moving Average Cluster Entropy Approach

Pietro Murialdo [1,*], Linda Ponta [2,*] and Anna Carbone [1,*]

[1] Institute of Condensed Matter Physics and Complex Systems, DISAT, Politecnico di Torino, 10129 Torino, Italy
[2] School of Industrial Engineering, LIUC-Università Cattaneo, Castellanza, VA 21052, Italy
* Correspondence: pietro.murialdo@polito.it (P.M.); lponta@liuc.it (L.P.); anna.carbone@polito.it (A.C.)

Received: 30 April 2020; Accepted: 4 June 2020; Published: 8 June 2020

Abstract: A perspective is taken on the intangible complexity of economic and social systems by investigating the dynamical processes producing, storing and transmitting information in financial time series. An extensive analysis based on the *moving average cluster entropy* approach has evidenced market and horizon dependence in highest-frequency data of real world financial assets. The behavior is scrutinized by applying the moving average cluster entropy approach to long-range correlated stochastic processes as the Autoregressive Fractionally Integrated Moving Average (ARFIMA) and Fractional Brownian motion (FBM). An extensive set of series is generated with a broad range of values of the Hurst exponent H and of the autoregressive, differencing and moving average parameters p, d, q. A systematic relation between moving average cluster entropy and long-range correlation parameters H, d is observed. This study shows that the characteristic behaviour exhibited by the horizon dependence of the cluster entropy is related to long-range positive correlation in financial markets. Specifically, long range positively correlated ARFIMA processes with differencing parameter $d \simeq 0.05$, $d \simeq 0.15$ and $d \simeq 0.25$ are consistent with moving average cluster entropy results obtained in time series of DJIA, S&P500 and NASDAQ. The findings clearly point to a variability of price returns, consistently with a price dynamics involving multiple temporal scales and, thus, short- and long-run volatility components. An important aspect of the proposed approach is the ability to capture detailed horizon dependence over relatively short horizons (one to twelve months) and thus its relevance to define risk analysis indices.

Keywords: cluster-entropy; Shannon-entropy; financial markets; time series; dynamics

1. Introduction

In recent years, much effort has been spent on studying complex interactions in financial markets by means of information theoretical measures from different standpoints. The information flow can be probed by observing a relevant quantity over a certain temporal range (e.g., price and volatility series of financial assets). Socio-economic complex systems exhibit remarkable features related to patterns emerging from the seemingly random structure in the observed time series, due to the interplay of long- and short-range correlated decay processes. The correlation degree is intrinsically linked to the information embedded in the patterns, whose extraction and quantification add clues to the underlying complex phenomena [1–14].

An information measure $S(x)$ was proposed by Claude Shannon to the aim of quantifying the degree of uncertainty of strings of elementary random events in terms of their probabilities [15]. The elementary stochastic events are related to a relevant variable x whose values are determined by the probability $\{p_i\}$. For example, the size ℓ of a string (block), corresponding to a particular realization within the sequence, can be associated to the probability $p_i(\ell)$ that, for stationary processes, does not depend on the actual position of the string (block) in the sequence. The Shannon measure is

then given by the expectation value $S(\ell) = \sum_i p_i(\ell) \log p_i(\ell)$ and is calculated over all possible strings ℓ. The *entropy density* is defined as $s_\ell = \lim_{\ell \to \infty} S(\ell)/\ell$ and quantifies the rate at which the process produces unexpected information as a function of the size ℓ.

A complexity measure $K(x)$ to quantify the amount of information contained in the string x was proposed by Kolmogorov [16]. The relation between Kolmogorov complexity and Shannon entropy has been extensively investigated, in particular the entropy density s_ℓ for a stationary process corresponds to the Kolmogorov entropy rate [17].

The first step required for the practical implementation of entropy and complexity measures is a suitable partition of the sequence which is critical to unbundle random and deterministic blocks of given length (decryption). The method usually adopted for partitioning a sequence and estimating its entropy is based on a uniform division in blocks with same length [18–21].

The *cluster entropy method* [9–11] implements the partition via a moving average process. The *clusters* correspond to blocks of different sizes, defined as the portion between consecutive intersections of a given time series and moving average. The *cluster entropy method* has been applied to financial markets in [22,23]. Cumulative information measures (indexes) have been worked out with the ability to provide deep insights on heterogeneity and dynamics. In particular:

- **Heterogeneity**. Volatility series have been analysed by using the cluster entropy approach over a constant temporal horizon (six years of tick-by-tick data sampled every minute). An information measure of heterogeneity, the *Market Heterogeneity Index* $I(T, n)$, where T and n are respectively the volatility and moving average windows, has been developed by integrating the cluster entropy curves of the volatility series over the cluster length τ. It has been also shown that the *Market Heterogeneity Index* can be used to yield the weights of an efficient portfolio as a complement to Markowitz and Sharpe traditional approaches for markets not consistent with Gaussian conditions [22].
- **Dynamics**. Prices series have been investigated by using the cluster entropy approach over several temporal horizons (ranging from one to twelve months of tick-by-tick data with sampling interval between 1 up to 20 seconds depending on the specific market). The study has revealed a systematic dependence of the cluster entropy over time horizons in the investigated markets. The *Market Dynamic Index* $I(M, n)$, where M is the temporal horizon and n is the moving average window, defined as the integral of the cluster entropy over τ, demonstrates its ability to quantify the dynamics of assets' prices over consecutive time periods in a single figure [23].

The present study is motivated by the results obtained in [23] showing that cluster entropy of real-world financial markets (NASDAQ, DJIA and S&P500) exhibits significant *market and horizon dependence*. According to classical financial theories, subsequent price deviations are identically and independently distributed (*iid*) and all the information are immediately reflected into markets, thus hampering past observations to predict future outcomes. If that were true, correlation would be negligible and prices would be simply modelled in terms of *fully uncorrelated Brownian motion*. However, several studies have shown that real world markets only partially behave according to the standard theory of perfectly informed and rational agents.

Here, we add further clues to the microscopic origin of the horizon dependence of the cluster entropy in financial markets. To this purpose, the cluster entropy approach is applied to an extensive set of artificially generated series with the aim of shedding light on the characteristic behaviour of real world assets [23]. We report results of the cluster entropy in *Geometric Brownian Motion* (GBM), *Generalized Autoregressive Conditional Heteroscedastic* (GARCH), *Fractional Brownian Motion* (FBM) and *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) processes. Those are well-known processes characterized either by hyperbolically decaying or exponentially decaying correlation functions, features reflected in long-range or short-range dependent dynamics of the elementary random events. The performance of the *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) process and its variants are receiving a lot of attention and are under intense investigation in the financial research community [24–28]. This work clearly demonstrates the relationship between the endogenous dynamics of the time series and their long-range dependence.

It is shown that deviations of the moving average cluster entropy behaviour in comparison to simple Brownian motion is unequivocally related to the long-range dependence of real-world market series. In particular, moving average cluster entropy results obtained on Fractional Brownian Motion with Hurst exponent H in the range $0 \leq H \leq 0.5$ (negatively correlated series) show no time horizon dependence. Conversely, moving average cluster entropy results with Hurst exponent H in the range $0.5 \leq H \leq 1$ (positively correlated series) exhibit some dispersion in the horizon dependence in analogy with the real-world financial markets. Results obtained on ARFIMA series confirm and extend the findings reported for FBMs. Horizon dependence of the cluster entropy is observed for a differencing parameter $0 \leq d \leq 0.5$. Fine tuning of the horizon dependence is obtained by varying the autoregressive p and moving average q components in the ARFIMA series.

The low-frequency volatility has been identified as the long-run component to describe market dynamic fundamentals in recent works [29–33]. The current work demonstrates the ability of the cluster entropy to capture short-range and long-range variability in price returns, thus to identify short-run and long-run factors in volatility and their linkages with macroeconomic variables and asset prices. On account of the dispersion of the *Market Dynamic Index* $I(M, n)$ at increasing values of the horizon M, our findings confirm that the slowest dynamic components (slowly evolving market fundamentals) reflect in the lowest-frequency volatility (large M scales) components of the assets. In this context, volatility can be modelled as a time dependent function, for example through the introduction of a quadratic spline to provide a smooth and nonlinear long-run trend in the volatility time series in the spline-GARCH model [29].

ARFIMA and Spline-GARCH belong to the class of *free-parameters model*, as they require for example the quadratic form of the time-dependence function parameters, or the autoregressive parameters. Conversely, the cluster entropy approach does not require free parameters. The cluster entropy is a *parameter-free model* based on data over some temporal horizons of choice. Hence, the comparison between results obtained by ARFIMA, GARCH models and those obtained by the cluster entropy approaches do not imply redundancy in the outcomes and is robust by design. The ability to extract market dynamic dispersion based only on data could be of relevance to disentangle performance of the different models at short and long horizons. This could be the case of ARFIMA models that tend to perform better on estimating asset variance at long-horizons compared to ARMA models that conversely produce superior results at short-horizons (see for example [24] where results of S&P500 are also reported).

In this work, the method is applied to mainstream financial assets as NASDAQ, DJIA and S&P500 tick-by tick data over the year 2018. The choice of these assets derives mainly from the need to validate the newly proposed cluster entropy approach on widely studied markets whose long range dependence has been quite widely investigated and broadly assessed by several studies. Further to these markets, interesting developments can be envisioned in different sectors that are strongly affected by macroeconomic variables and shock (e.g., time dependent variance and persistence have been observed in Real Estate securities [30] highlighting linkages between real estate stocks and market fundamentals, related to endogenous dynamics and horizon dependence).

The organisation of the work is as follows. The cluster entropy method used for the analysis and the investigated market and artificial data are described in Section 2. Results on cluster entropy and market dynamic index estimated over *Geometric Brownian Motion* (GBM), *Generalized Autoregressive Conditional Heteroskedastic* (GARCH), *Fractional Brownian Motion* (FBM) and *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) series, are reported in Section 3. Finally, results are discussed, conclusions are drawn and a path for future work is suggested in Section 4.

2. Methods and Data

In this section the cluster entropy approach developed in [9,10] is briefly recalled. The second part of this section is devoted to the description of financial market data used in [23]. For the sake of completeness, we also recall the main definitions related to the Geometric Brownian Motion,

Generalized Autoregressive Conditional Heteroskedastic, Fractional Brownian Motion, Autoregressive Fractionally Integrated Moving Average processes.

2.1. Cluster Entropy Method

It is well-known that the general idea behind Shannon entropy is to measure the amount of information embedded in a message to identify the shortest subsequence actually carrying the relevant information and the degree of redundancy which is not necessary to reproduce the initial message. The Shannon functional is written as:

$$S(\tau, n) = \sum P(\tau, n) \log P(\tau, n), \tag{1}$$

where $P(\tau, n)$ is a probability distribution associated with the time series $y(t)$. To estimate the probability distribution $P(\tau, n)$, it is necessary to partition the continuous phase space into disjoints sets. The traditionally adopted methods divide the sequence into segments of equal lengths (blocks). Here, we follow another approach.

In [9,10] the time sequence $y(t)$, is partitioned in *clusters* by the intersection with its moving average $\tilde{y}_n(t)$, with n the size of the moving average. The simplest type of moving average is defined at each t as the average of the n past observation from t to $t - n + 1$,

$$\tilde{y}_n(t) = \frac{1}{n} \sum_{k=0}^{n-1} y(t-k). \tag{2}$$

Note that while the original series is defined from 1 to N, the moving average series is defined from 1 to $N - n + 1$ because n samples are necessary to initialize the series. The original series and the moving average series are indicated as $\{y(t)\}_{t=1}^{N}$ and $\{\tilde{y}_n(t)\}_{t=1}^{N-n+1}$ respectively. Consecutive intersections of the time series and of the moving average series yield a partition of the phase space into a series of *clusters*. Each cluster is defined as the portion of the time series $y(t)$ between two consecutive intersection of $y(t)$ itself and its moving average $\tilde{y}_n(t)$ and has length (or duration) equal to:

$$\tau_j \equiv ||t_j - t_{j-1}||, \tag{3}$$

where t_{j-1} and t_j refers to two subsequent intersections of $y(t)$ and $\tilde{y}_n(t)$. For each moving average window n, the probability distribution function $P(\tau, n)$, i.e., the frequency of the cluster lengths τ, can be obtained by counting the number of clusters $\mathcal{N}_j(\tau_j, n)$ with length τ_j, $j \in \{1, N - n - 1\}$. The probability distribution function $P(\tau, n)$ results:

$$P(\tau, n) \sim \tau^{-D} \mathcal{F}(\tau, n) \quad , \tag{4}$$

where the exponent D indicates the fractal dimension and can be expressed as

$$D = 2 - H \tag{5}$$

with H the Hurst exponent of the sequence. Hence, the fractal dimension ranges between $1 < D < 2$, as the Hurst exponent varies between $0 < H < 1$. In this framework long-range correlation implies that the clusters are organized in a similar way along the time series (self-organized), even for clusters far away in time from each other. The term $\mathcal{F}(\tau, n)$ in Equation (4) takes the form:

$$\mathcal{F}(\tau, n) \equiv e^{-\tau/n} \quad , \tag{6}$$

to account for the drop-off of the power-law behavior for $\tau < n$ and the onset of the exponential decay when $\tau \geq n$ due to the finiteness of n. When $n \to 1$ the lengths τ of clusters tend to be centered around a single value. When $n \to N$, that is when n tends to the length of the whole sequence, only one cluster

with $\tau = N$ is generated. For middle values of n however a broader range of lengths is obtained and therefore the probability distribution spreads all values. When the probability distribution in Equation (4) is fed into the Shannon functional in Equation (1) the result is the following:

$$S(\tau, n) = S_0 + \log \tau^D - \log \mathcal{F}(\tau, n), \tag{7}$$

which, after substituting Equation (6), becomes:

$$S(\tau, n) = S_0 + \log \tau^D + \frac{\tau}{n}, \tag{8}$$

where S_0 is a constant, $\log \tau^D$ accounts for power-law correlated clusters related to τ^{-D} and τ/n accounts for exponentially correlated clusters related to the term $\mathcal{F}(\tau, n)$. The term S_0 can be evaluated in the limit $\tau \sim n \to 1$, which results in $S_0 \to -1$ and $S(\tau, n) \to 0$, that corresponds to the fully deterministic case, where each cluster has size equal to 1. On the other hand, when $\tau \sim n \to N$, the maximum value for the entropy is obtained with $S(\tau, n) = \log N^D$, which corresponds to the case of maximum randomness, where there is one cluster coinciding with the whole series. Equation (8) shows that power-law correlated clusters, characterized by having length $\tau < n$, are described by a logarithmic term as $\log \tau^D$, and their entropy do not depend on the moving average window n. However, for values of $\tau \geq n$, which represent exponentially correlated clusters, the term τ/n becomes predominant. Cluster entropy increases linearly as τ/n, with slope decreasing as $1/n$. Hence, due to the finite size effects introduced by the partitioning method, in $\tau = n$ the behavior of entropy changes and its values exceeds the curve $\log \tau^D$. In other words, clusters that are power-law correlated does not depend on n, are said to be *ordered* and represent deterministic information. Clusters that are exponentially correlated does depend on n, are said to be *disordered* and represent random clusters.

The meaning of entropy in information theory can be related to the corresponding concepts in thermodynamics. In an *isolated system*, the entropy increase dS refers to the irreversible processes spontaneously occurring within the system. In an *open system*, an additional entropy increase dS_{ext} should be taken into account due to the interaction with the external environment.

The term $\log \tau^D$ should be interpreted as the entropy of the isolated system. It is independent on n, that is it is independent on the partitioning method. It takes the form of the Boltzmann entropy, that can be written as $S = \log \Omega$, with Ω the volume of the system. Therefore the quantity τ^D corresponds to the volume occupied by the fractional random walker.

The term τ/n represents the excess entropy caused by the external process of partitioning the sequence. The excess entropy depends on the moving average window n. If same size boxes were chosen, the excess entropy term τ/n would vanish and entropy would reduce to the logarithmic term. When a moving average partition is used, the term τ/n emerges to account for the additional heterogeneity introduced by the randomness of the process. Thence, for exponentially correlated clusters entropy exceeds the logarithmic asymptotic.

In order to increase the sensitivity of the method, the integral of the entropy function over the clusters length τ can be considered:

$$I(n) = \int S(\tau, n) d\tau \quad , \tag{9}$$

which for discrete sets reduces to $I(n) = \sum_\tau S(\tau, n)$. The function $I(n)$ is a cumulative entropy measure able to embed all the information in a single figure.

Equation (9) can be written as:

$$I(n) = \int_1^{\tau(n)} S(\tau, n) d\tau + \int_{\tau(n)}^{\infty} S(\tau, n) d\tau \quad . \tag{10}$$

The first integration is referred to the power law regime of the cluster entropy, the second integration is referred to the linear regime of the cluster entropy (i.e., the excess entropy term).

2.2. Financial Data

The objective of this work is to investigate and shed light on the characteristic features exhibited by cluster entropy of financial markets. In particular here our focus is on the systematic dependence of the cluster entropy of the price series over time horizon M.

In [23] the cluster entropy is applied to a large set of tick-by-tick data of the USA's indexes (S&P500, NASDAQ and DJIA). NASDAQ is an index resulting from all the public firms quoted on the market, DJIA and S&P500 indexes are representative of a selected number of public firms. For each index, investigated data include tick-by-tick prices from January 2018 to December 2018. As the main goal of the paper is to quantify the intrinsic dynamics of prices and to capture the endogenous sources of risk over different temporal horizons, a year of data with no external shocks or crisis have been chosen. More information about the markets can be found at the Bloomberg terminal.

To study the dynamics of financial series different time horizons need to be compared. As explained in the Introduction, entropy is sample-size dependent by definition, thus in order to rule out spurious results the length of the investigated sequences must be the same. Therefore, cluster entropy analysis requires the comparison to be implemented on series with same length. Raw data have been downloaded from the Bloomberg terminal in the form of tick-by-tick data. The lengths of the raw series vary due to different number of trading days and transactions per time unit. It is therefore necessary, as first computational step, to implement a sampling of the raw data to make the length of the series exactly the same. The first raw series ranges from the first transaction of January 2018 to the last one of January 2018; the second ranges from the first transaction in January 2018 to the last of February 2018, . . ., the twelfth ranges from the first transaction in January 2018 to the last of December 2018, a period equivalent to the whole year. Because each raw series ranges from the first tick of 2018 to the last tick of the relative month, the twelve series have very different lengths. The series are sampled to obtain twelve *series* with same length as described in the following.

Twelve sampling time intervals and corresponding frequencies must be defined, i.e., twelve integers indicating for each series the interval of skipped data. Sampling intervals are obtained by dividing the length of each raw series by the length of the shortest raw one and then rounding to the inferior integer. Thence, each raw series is sampled with the relative sampling interval to yield a *sampled series*: for each sample in the sampled series, a number of samples equal to the sampling frequency has been discarded in the raw series. The sampled series obtained are *approximately* of equal lengths. To obtain twelve series of *exactly* equal length, a few observations are cut off, when exceeding the length of the shortest series. The result consists in twelve sampled series that are equal in length and refer to time horizons varying from one month ($M = 1$) to twelve months ($M = 12$). In more details, N_M is the length of the series corresponding to the horizon M (where M ranges from 1 to 12 for one year of data). The shortest monthly series is used to evaluate the minimum value N_M^* and the corresponding sampling frequency. Then, the sampling intervals for the multiple periods is derived by dividing the multiple period lengths (i.e., the sum of multiple consecutive N_M) by the value N_M^*. In Table 1 a few examples of sampling intervals and lengths N_M are shown to clarify the procedure. It is worth noting that the length of sampled series should be at least 10^5 to ensure enough accuracy of the results.

Table 1. Example of lengths and time horizon M for NASDAQ data (2018). The 2nd column corresponds to the number of transactions over the horizon M. These lengths are used as a reference to generate artificial series to allow a direct comparison between results obtained on real and artificial data. The 3rd column corresponds to the sampled lengths used in the calculation of the cluster entropy. The 4th and 5th columns correspond respectively to the raw and rounded time intervals obtained dividing N_M by the series of shortest length N_M^*.

M	N_M	N_M^*	t_S	t_S^*
1	586,866	586,866	1.0000	1
2	1,117,840	586,866	1.9048	1
3	1,704,706	586,866	2.9048	2
4	2,291,572	586,866	3.9048	3
5	2,906,384	586,866	4.9524	4
6	3,493,250	586,866	5.9524	5
7	4,069,315	586,866	6.9340	6
8	4,712,062	586,866	8.0292	8
9	5,243,029	586,866	8.9339	8
10	5,885,781	586,866	10.0292	10
11	6,461,845	586,866	11.0108	11
12	6,982,017	586,866	11.8971	11

2.3. Artificial Data

Artificial series have been generated by using Geometric Brownian Motion, Generalized Autoregressive Conditional Heteroskedastic, Fractional Brownian Motion and Autoregressive Fractionally Integrated Moving Average processes with same temporal structure corresponding to the different horizons of the financial market data reported in [23]. Then the sampling method proceeds analogously from the calculation of the sampling frequency. Such sampling method was applied to series generated by artificial financial models to make sure that the information content would be comparable to that of real-world financial series. In the remainder of this section, we recall the main definitions for the afore mentioned processes.

2.3.1. Geometric Brownian Motion

The Geometric Brownian Motion is the basis of the *Black-Scholes-Merton* model used to price options and is defined by the following difference equation:

$$dX_t = \mu(t)X_t dt + D(t, X_t)\sigma(t)dB_t, \tag{11}$$

where $\mu(t)$ indicates the level of return, $\sigma(t)$ the volatility and dB_t is a simple Brownian motion. Volatility is deterministic and constant and there are no jumps. Increments are independent on previous states.

2.3.2. Autoregressive Conditional Heteroskedasticity Models

We perform simulations by using GARCH(1,1) of the broad family of the autoregressive conditional heteroscedasticity (ARCH) models. It describes the variance of the current error term or innovation as a function of previous values. The GARCH(1,1) model is defined by the following relationships:

$$\mu(t) - E_{t-1}\mu(t) = \sqrt{\sigma_t}\varepsilon_t$$
$$\sigma_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}$$

where $\mu(t)$ represents the return of an asset at time t, $E_{t-1}(\mu_t)$ is the expected return at $t-1$, σ_t characterises the conditional volatility at time t, and ε_t is the innovation term at time t.

2.3.3. Fractional Brownian Motion

The *Fractional Brownian Motion* is a long memory process introduced in [34]:

$$B_H(t) = B_H(0) + \frac{1}{\Gamma(H+1/2)}\Bigg(\int_{-\infty}^{0}\big((t-s)^{H-1/2} - (-s)^{H-1/2}\big)dB(s) + \int_0^t (t-s)^{H-1/2}dB(s)\Bigg). \tag{12}$$

It is also referred to as a *self-similar* process. A stochastic process X_t, with $t \in \mathbb{R}$, is said to be self-similar if there exist $H > 0$ such that for any *scaling factor* $c > 0$,

$$X_{ct} \stackrel{\mathcal{L}}{=} c^H X_t, \tag{13}$$

with H the Hurst exponent and ($\stackrel{\mathcal{L}}{=}$) equivalence in distribution. Self-similar processes are stochastic models where a scaling in time is equivalent, *in term of distribution*, to an appropriate scaling in space. Moreover, if, for any k, the distribution of $(X_{t_1+c} - X_{t_1+c-1}, ..., X_{t_k+c} - X_{t_k+c-1})$ does not depend on c, X_t is said to be self-similar with *stationary increments*. So, a Gaussian process $B_H(t)$ is called a *Fractional Brownian Motion*, if it satisfies: 1. $B_H(t)$ is self-similar with $0 < H < 1$; 2. $B_H(t)$ has stationary increments. When $H = 0.5$ a simple Brownian Motion with independent increments is recovered. When $0 < H < 0.5$ the *Fractional Brownian Motion* is said to be anti-persistent, which means that increments tend to be opposite signed. Conversely, when $0.5 < H < 1$ it is said to be persistent, which means that increments tend to be equally signed.

2.3.4. Autoregressive Fractionally Integrated Moving Average

The *Autoregressive Fractionally Integrated Moving Average* (*ARFIMA*) is one of the most common processes to model long-range correlated asset prices. The Autoregressive Fractionally Integrated Moving Average process of order (p, d, q) with mean μ, may be written, using the lag operator L, as:

$$\Phi(L)(1-L)^d(y_t - \mu) = \Theta(L)\epsilon_t, \tag{14}$$

with ϵ_t *i.i.d.* and $\sim (0, \sigma_\epsilon^2)$. The autoregressive component of the process is represented by the factor:

$$\Phi(L) = 1 - \phi_1 L - ... - \phi_p L^p, \tag{15}$$

where the lag operator of order p shifts the value of y_t back to p observations, so that one obtains:

$$\Phi(L)y_t = (1 - \phi_1 L - ... - \phi_p L^p)y_t = y_t - \phi_1 y_{t-1} - ... - \phi_p y_{t-p}. \tag{16}$$

The moving average component of the process is represented by the factor:

$$\Theta(L)\epsilon_t = (1 + \theta_1 L + ... + \theta_q L^q)\epsilon_t = \epsilon_t + \theta_1\epsilon_{t-1} + ... + \theta_q\epsilon_{t-q} \quad . \tag{17}$$

The fractionally differencing operator $(1-L)^d$ is defined as:

$$(1-L)^d = \sum_{n=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)}. \tag{18}$$

Note that the process is stationary only for $-0.5 < d < 0.5$. For $d < |0.5|$ the ARFIMA process is said to exhibit long memory.

The power spectral representation $f(\lambda)$ of Fractional Brownian Motions and Autoregressive Fractionally Integrated Moving Average Processes provides further details regarding their power law behavior and the relation between the characteristic exponents. It is :

$$\begin{array}{ll} f(\lambda) \sim |\lambda|^{-2d} & \text{(ARFIMA)} \\ f(\lambda) \sim |\lambda|^{1-2H} & \text{(FBM)} \end{array} \tag{19}$$

yielding:

$$H = d + 1/2. \tag{20}$$

3. Results

In this section, the results of the application of the cluster entropy method to several FBM and ARFIMA series are presented. The moving average cluster entropy can be implemented via the MATLAB codes available at [35].

First, a set of benchmark values for the cluster entropy are obtained by implementing the algorithm on Geometric Brownian Motion (GBM) and Generalized Autoregressive Conditional Heteroskedastic (GARCH) series. Geometric Brownian Motion series are generated by means of the MATLAB tool available at [36]. GBM series are analysed with parameters varying in the range $0 \leq \mu \leq 1 \times 10^{-7}$ and $5 \times 10^{-4} \leq \sigma \leq 5 \times 10^{-6}$. GARCH series are generated by using the computational tool provided in MATLAB [37]. Figure 1 reports cluster entropy and market dynamic index results obtained on GBM and GARCH series. The GBM series are generated with the following parameters: $\mu = 1 \times 10^{-7}$ and $\sigma = 5 \times 10^{-4}$; the GARCH series are generated with the following parameters: $\alpha = 0.475$, $\beta = 0.1$ and $\omega = 0.1$. Left and middle panels show cluster entropy curves for time horizons $M = 1$ and $M = 12$, i.e., corresponding respectively to one period (one month) and twelve periods (one year) of data. Right panels show Market Dynamic Index $I(M, n)$ for different horizons M and moving average windows n. $I(M, n)$ does not depend on the temporal horizon M both in GBM and GARCH series.

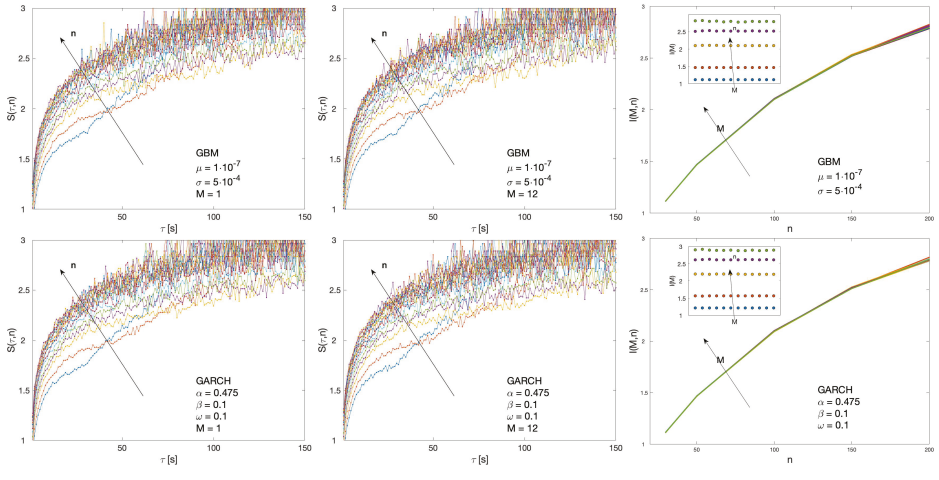


Figure 1. Cluster entropy results for Geometric Brownian Motion (GBM) and Generalized Autoregressive Conditional Heteroskedastic (GARCH) series are reported respectively in the first and in the second row. GBM series are generated with following parameters: $\mu = 1 \times 10^{-7}$ and $\sigma = 5 \times 10^{-4}$. GARCH series are generated with the following parameters: $\alpha = 0.475$, $\beta = 0.1$ and $\omega = 0.1$. Left and middle panels show cluster entropy curves for time horizons $M = 1$ and $M = 12$. Right panels show Market Dynamic Index $I(M, n)$ for different horizons M and moving average windows n. $I(M, n)$ is independent on M.

Results of the cluster entropy approach applied to Fractional Brownian Motion (FBM) are reported in Figure 2. The Fractional Brownian Motion series were generated by means of the FRACLAB tool available at [38]. Several Fractional Brownian Motion series with Hurst exponent varying in the range $0.1 \leq H \leq 0.9$ are analysed. Figure 2 shows the cluster entropy for time horizon $M = 1$ and $M = 12$ for FBM series with $H = 0.3$, $H = 0.5$ and $H = 0.8$.

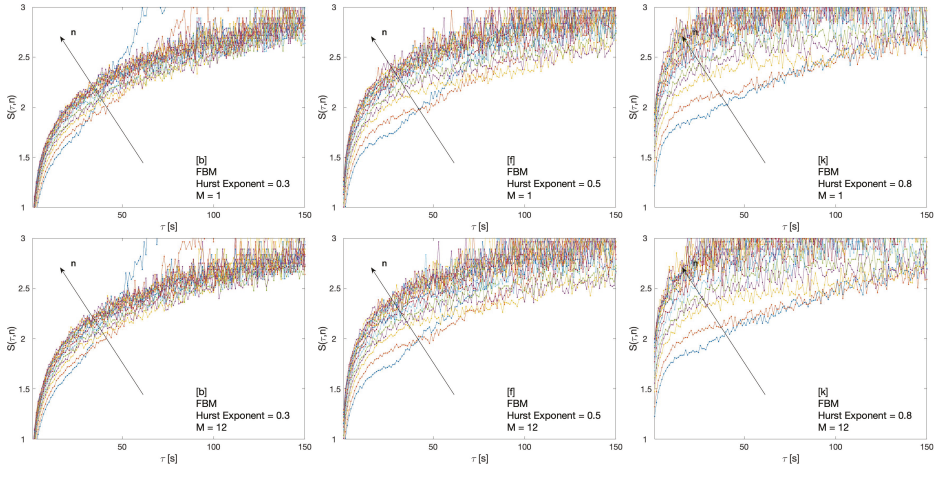


Figure 2. Cluster entropy results for Fractional Brownian Motion (FBM) series with $H = 0.3$, $H = 0.5$, $H = 0.8$. First row shows results for time horizon $M = 1$ (approximately equivalent to the first month (January 2018) of raw data for NASDAQ, S&P500, DIJA). The second row shows results for time horizon $M = 12$ (approximately equivalent to twelve months of data in NASDAQ, S&P500, DIJA, i.e, the whole 2018 year).

In general, cluster entropy calculated at different time horizons M presents a similar behavior. On account of Equation (8), power-law correlated clusters with a smooth logarithmic increase of the entropy for $\tau < n$ can be expected. Conversely, for $\tau \geq n$, the exponentially correlated decay sets the entropy to increase linearly with the term τ/n dominating. However, a quite different behavior is observed for different H. For $H = 0.3$ (anti-correlated FBM series), cluster entropy curves exhibit a very limited dependence on the moving average window n over the range of investigated τ. For $H = 0.5$, cluster entropy curves vary more significantly as the moving average window n changes. For $H = 0.8$, cluster entropy curves vary even more remarkably by taking higher values for increasing n.

The dependence of the cluster entropy on temporal horizon M is reflected in the results of the Market Dynamic Index $I(M, n)$ plotted in Figure 3. The Market Dynamic Index $I(M, n)$ is estimated over several FBM series with different Hurst exponents H. For anticorrelated series $0 \leq H \leq 0.5$, $I(M, n)$ curves overlap for all the moving average windows n and time horizons M. For positively correlated series $0.5 \leq H \leq 0.9$, $I(M, n)$ exhibits slightly different values as a function of time horizons M. It is worth-noting that the magnitude of the marginal increments in $I(M, n)$ at large n increases as H increases for $0 \leq H \leq 0.5$, reaches a maximum for $H = 0.5$ and then decreases again for $0.5 \leq H \leq 0.9$. This effect is evident in the insets of Figure 3.

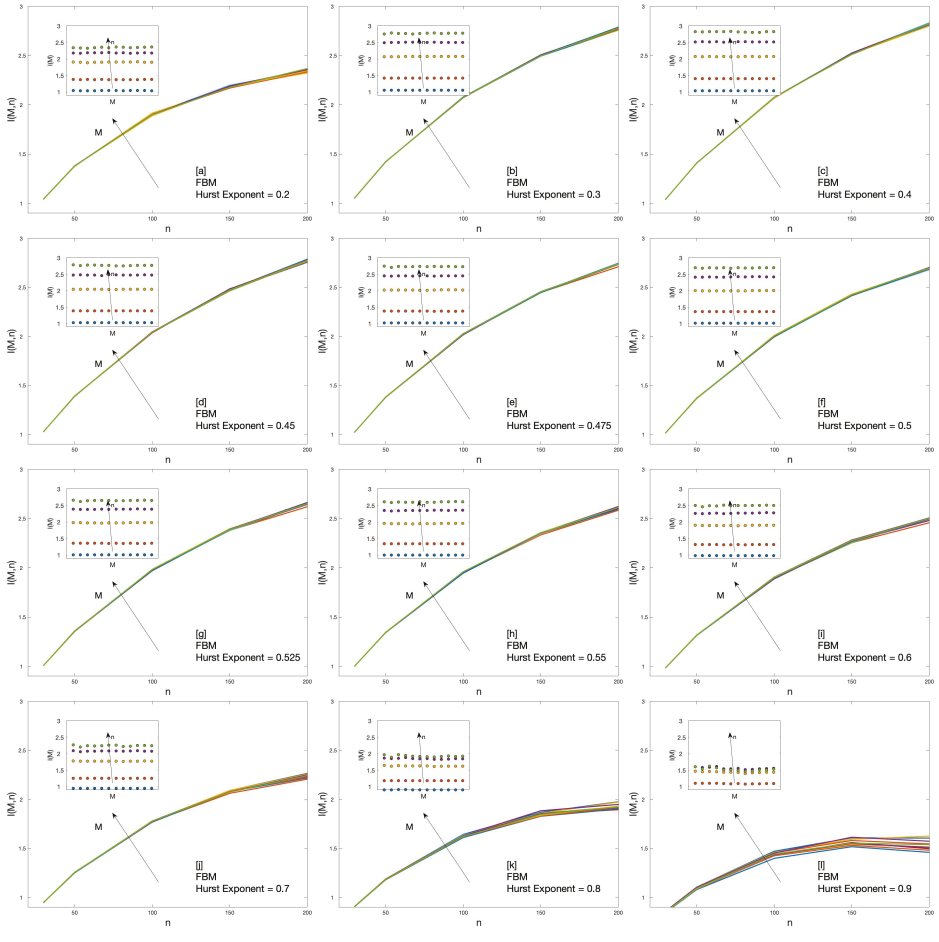


Figure 3. Market Dynamic Index $I(M, n)$ for Fractional Brownian Motion series with Hurst exponent ranging from $H = 0.2$ to $H = 0.9$ respectively from (a) to (l).

The cluster entropy analysis is implemented on *Autoregressive Fractionally Integrated Moving Average* (ARFIMA) series obtained by means of simulations for several combination of parameters [39]. The extent of investigated parameters are marked by alphabet labels in Table 2 for ARFIMA (1,d,1) and in Table 3 for ARFIMA (3,d,2) and ARFIMA(1,d,3).

Table 2. Full set of parameter range for the ARFIMA (1,d,1) processes simulated in this work. Specifically, D is the fractal dimension, H is the Hurst exponent and d is the differencing parameter (1st, 2nd and 3rd columns) which are related by Equation (20), ϕ is the autoregressive parameter (4th column), and θ is the moving average parameter (5th column). *Label* refers to each parameter set (6th column). Specifically, cluster entropy results for the parameter sets: [a1], [b1], [e1], [f1], [i1], [l1] are plotted in Figure 4 (M = 1) and Figure 5 (M = 12), while the Market Dynamic Index is plotted in Figure 6.

D	H	d	ϕ	θ	*Label*
1.45	0.55	0.05	0.20	0.90	a1
			0.90	0.20	b1
1.40	0.60	0.10	0.20	0.90	c1
			0.90	0.20	d1
1.35	0.65	0.15	0.20	0.90	e1
			0.90	0.20	f1
1.30	0.70	0.20	0.20	0.90	g1
			0.90	0.20	h1
1.25	0.75	0.25	0.20	0.90	i1
			0.30	0.40	j1
				0.85	k1
			0.90	0.20	l1
				0.40	m1
				0.85	n1
1.20	0.80	0.30	0.20	0.90	o1
			0.90	0.20	p1
1.02	0.98	0.48	0.30	0.40	q1
				0.85	r1
			0.90	0.40	s1
				0.85	t1

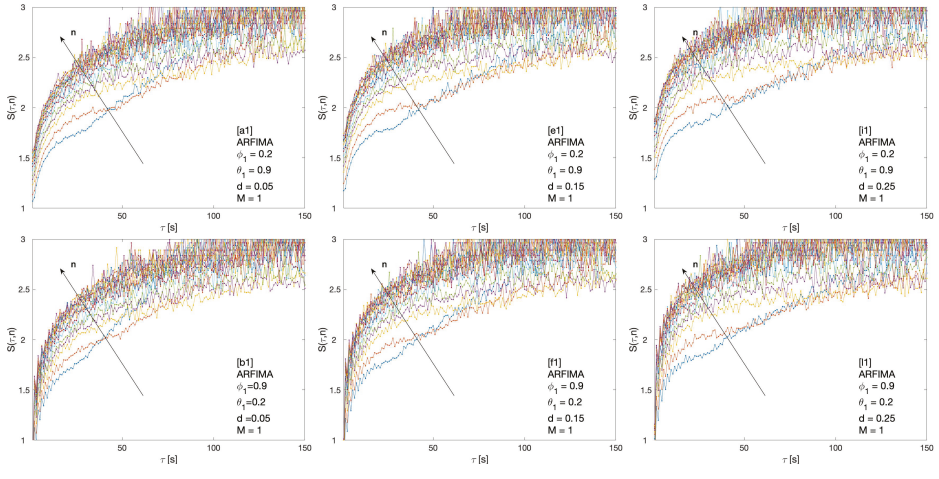


Figure 4. Cluster entropy results for horizon $M = 1$ for ARFIMA series with different combinations of the differencing parameter d, autoregressive parameter ϕ and moving average parameter θ. The differencing parameter takes values $d = 0.05$, $d = 0.15$, $d = 0.25$ with a different combinations of autoregressive and moving average parameter. The full set of analysed values of d, ϕ and θ is reported in Table 2.

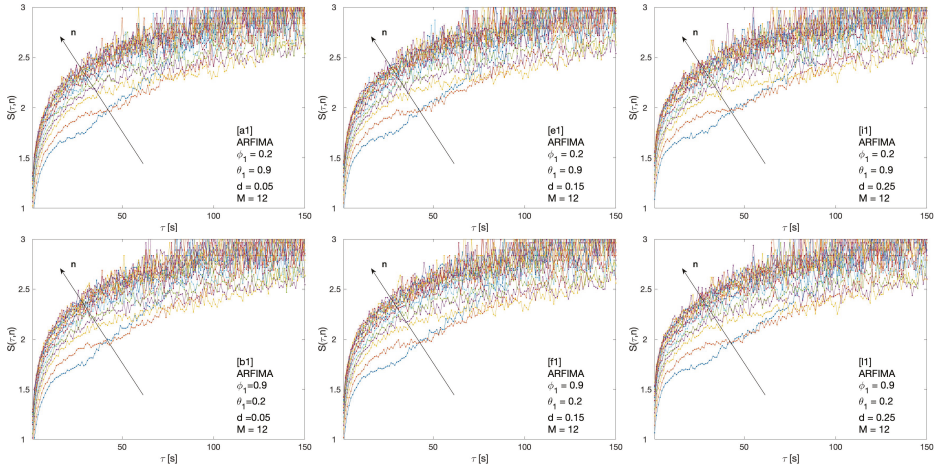


Figure 5. Cluster entropy results for horizon $M = 12$ on ARFIMA series with different combinations of the differencing parameter d, autoregressive parameter ϕ and moving average parameter θ. The differencing parameter takes values $d = 0.05$, $d = 0.15$ and $d = 0.25$ with a different combination of autoregressive and moving average parameters. The full set of analysed values of d, ϕ and θ is reported in Table 2.

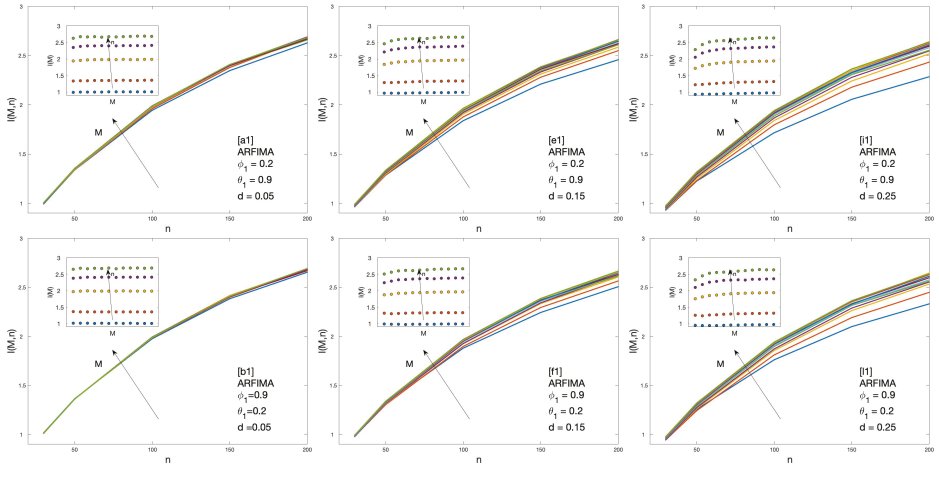


Figure 6. Market Dynamic Index $I(M, n)$ for ARFIMA series with different combinations of the differencing parameter d, autoregressive parameter ϕ, and moving average parameter θ. The differencing parameter takes values $d = 0.05$, $d = 0.15$, $d = 0.25$, with a different combination of autoregressive and moving average parameters. The full set of analysed values of d, ϕ and θ is reported in Table 2.

Table 3. Full set of parameter range for ARFIMA (3,d,2) and ARFIMA(1,d,3) processes simulated in this work. Specifically H is the Hurst exponent and d is the differencing parameter which are related by Equation (20) (1st and 2nd columns); ϕ_1, ϕ_2 and ϕ_3 are the autoregressive parameters (3rd, 4th and 5th columns); θ_1, θ_2 and θ_3 are the moving average parameters (6th, 7th and 8th columns). *Label* refers to each parameter set (10th column). Specifically, cluster entropy results for the parameter sets: [a2], [b2], [e2], [f2], [i2], [j2] are plotted in Figure 7 (M = 1) and Figure 8 (M = 12), while the Market Dynamic Index is plotted in Figure 9.

D	H	d	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	*Label*
1.45	0.55	0.05	0.20	-	-	0.90	0.90	0.90	a2
			0.90	0.90	0.90	0.20	0.20	-	b2
1.40	0.60	0.10	0.20	-	-	0.90	0.90	0.90	c2
			0.90	0.90	0.90	0.20	0.20	-	d2
1.35	0.65	0.15	0.20	-	-	0.90	0.90	0.90	e2
			0.90	0.90	0.90	0.20	0.20	-	f2
1.30	0.70	0.20	0.20	-	-	0.90	0.90	0.90	g2
			0.90	0.90	0.90	0.20	0.20	-	h2
1.25	0.75	0.25	0.20	-	-	0.90	0.90	0.90	i2
			0.90	0.90	0.90	0.20	0.20	-	j2
1.20	0.80	0.30	0.20	-	-	0.90	0.90	0.90	k2
			0.40	0.16	-	0.90	0.81	0.73	l2
			0.90	0.90	0.90	0.20	0.20	-	m2
1.15	0.85	0.35	0.20	-	-	0.90	0.90	0.90	n2
1.02	0.98	0.48	0.40	0.16	-	0.90	0.81	0.73	o2

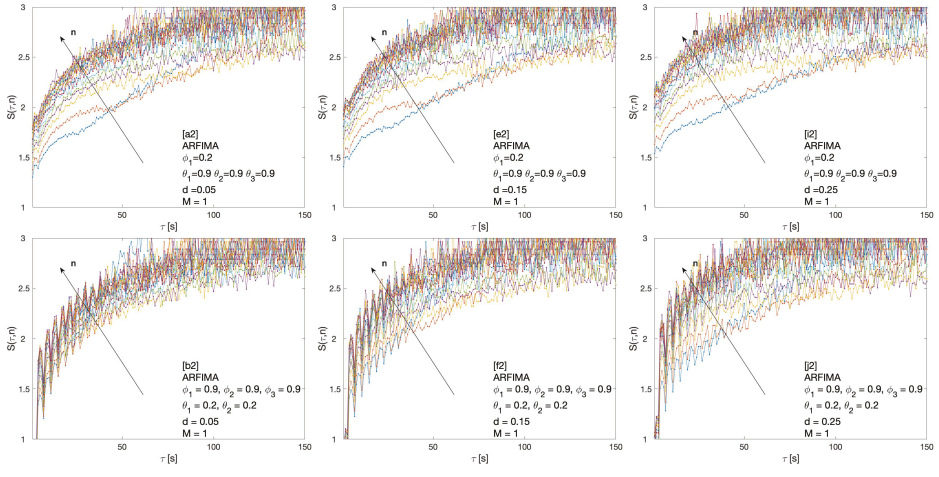


Figure 7. Cluster entropy results for horizon $M = 1$ on ARFIMA series with different combinations of the differencing parameter d, autoregressive parameter ϕ_1, ϕ_2, and ϕ_3 and moving average parameter θ_1, θ_2 and θ_3. The differencing parameter takes values $d = 0.05$, $d = 0.15$, $d = 0.25$, with a different combination of autoregressive and moving average parameters. The full set of analysed values of d, ϕ and θ is reported in Table 3.

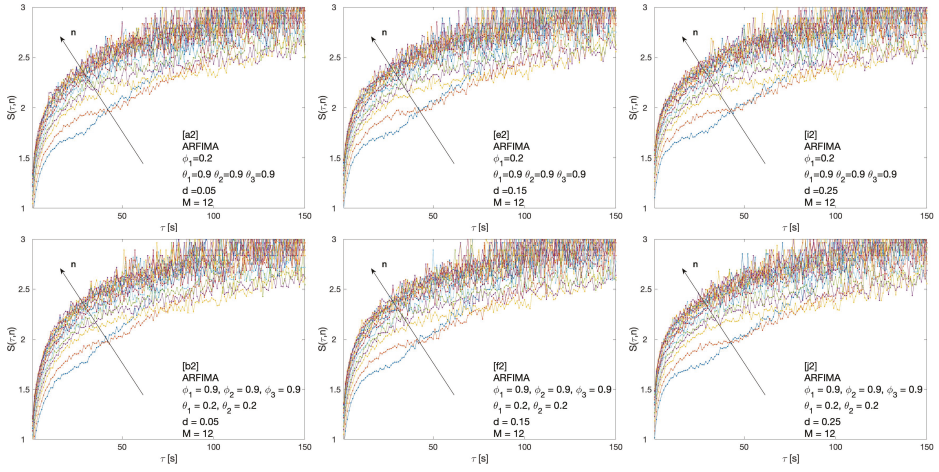


Figure 8. Cluster entropy results for horizon $M = 12$ on ARFIMA series with different combinations of the differencing parameter d, autoregressive parameter ϕ_1, ϕ_2, and ϕ_3 and moving average parameter θ_1, θ_2 and θ_3. The differencing parameter takes values $d = 0.05$, $d = 0.15$ and $d = 0.25$, with a different combination of autoregressive and moving average parameters. The full set of analysed values of d, ϕ and θ is reported in Table 3.

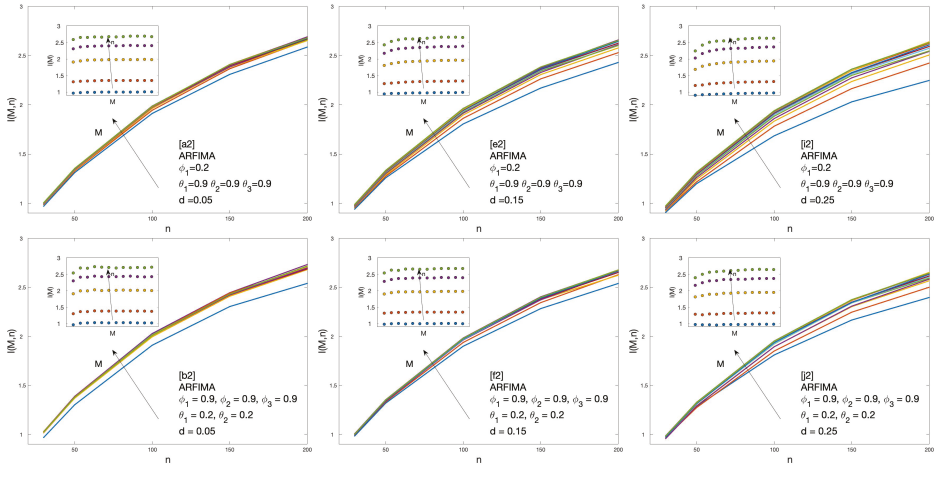


Figure 9. Market Dynamic Index $I(M, n)$ for ARFIMA series with different combinations of the differencing parameter d, autoregressive parameter ϕ_1, ϕ_2, and ϕ_3 and moving average parameter θ_1, θ_2 and θ_3. The differencing parameter takes values $d = 0.05$, $d = 0.15$ and $d = 0.25$, with a different combination of autoregressive and moving average parameters as reported in Table 3.

Cluster entropy results for ARFIMA (1,d,1), corresponding to parameters marked by alphabet labels in Table 2, are shown in Figures 4 and 5. The corresponding market dynamic indexes $I(M, n)$ calculated by using the data of the cluster entropy results on ARFIMA (1,d,1) are shown in Figure 6. Cluster entropy results on ARFIMA (3,d,2) and ARFIMA(1,d,3), corresponding to parameters marked by alphabet labels in Table 3, are shown in Figures 7 and 8. Market Dynamic Index for series generated by ARFIMA (3,d,2) and ARFIMA(1,d,3) processes are reported in Figure 9. With differencing parameter $0 < d < 0.2$, Market Dynamic Index curves are n-invariant for small values of n, but horizon dependence emerges at larger n. When $0.2 < d < 0.5$, Market Dynamic Index curves show a significant

horizon dependence even at small n. Therefore, according to the choice of the differencing parameter d, series generated by ARFIMA processes can reproduce the effect shown by the cluster entropy in real-world financial markets.

4. Discussion and Conclusions

The cluster entropy behavior described by Equation (8) has been replicated by simulations performed on artificially generated series, with results reported in Section 3. Figures show cluster entropy results for the following processes: Geometric Brownian Motion (Figure 1); Generalized Autoregressive Conditional Heteroskedastic processes (Figure 1); Fractional Brownian Motion (Figure 2); Autoregressive Fractionally Integrated Processes (Figures 4, 5, 7 and 8). The focus here is limited to the results shown in Figures 2, 4, 5, 7 and 8 related to FBM and ARFIMA because they are long-range dependent models relevant to the present analysis. The behavior of cluster entropy curves is well represented by Equation (8), while deviations occur in *extreme cases*, as in the case of ARFIMA(5,d,0) models generated by autoregressive parameters $\theta_i \simeq 0.9$, that are far away from those observed in real markets. In general, power-law correlated clusters, characterized by length $\tau < n$, determine the logarithmic behavior and the entropy term $\log \tau^D$, regardless of the moving average window value n. On the other hand, exponentially correlated clusters, characterized by length $\tau > n$, are related to the linear behavior prescribed by the excess entropy term τ/n, which depends on the moving average window n.

Cumulative measures are useful to summarize key information in a single figure. Thus, the Market Dynamic Index $I(M,n)$ is deduced from the cluster entropy results by means of Equation (10). $I(M,n)$ gathers the information present in the FBM series at different time horizons M and moving average windows n as shown in Figure 3. The Market Dynamic Index $I(M,n)$ replicates the characteristic behaviour observed in real world financial markets [23] when estimated in long-range positively correlated sequences. Conversely, one can note that the Market Dynamic Index $I(M,n)$ for Fractional Brownian processes with Hurst exponent $0 < H < 0.5$ (anticorrelated FBMs) does not present any horizon dependence. Conversely, Fractional Brownian Motion series with $0.5 < H < 1$ (positively correlated FBMs) do show horizon dependence. However, as it will be discussed below, Fractional Brownian Motion series fail to fully reproduce the financial markets behavior.

In the case of the ARFIMA processes, a significant horizon dependence emerges, as observed in the Market Dynamic Index curves plotted in Figures 6 and 9. Thus, cluster entropy for ARFIMA processes exhibits horizon dependence as observed in real world financial markets. The extent of long range dependence and its microscopic origin are consistent with findings of previous studies [27,28].

To further validate the findings, statistical significance has been checked by using the *T-paired test* of the null hypothesis h_0 that the cluster entropy values obtained by ARFIMA simulations come from distributions with equal mean, variance and probability p as the simple Brownian Motion ($H = 0.5$), assumed as benchmark. The results of *T-paired test* are reported in Table 4 (for the sake of comparison the results of the *T-paired test* performed on NASDAQ, DJIA and S&P500 markets in Table 5 [23] are also included here).

A qualitative comparison between Tables 4 and 5 suggests an overall similarity between ARFIMA and real world markets behaviour. In particular, p values in column *[f1]* are quite close to those of the S&P500 suggesting a correlation degree with Hurst exponent $H \simeq 0.65$ and differencing parameter $d \simeq 0.15$ for S&P500. Probability values p in column *[e2]* are also close to S&P500, confirming the value $H \simeq 0.65$ and $d \simeq 0.15$. The probability values for DJIA are better approximated by the set of ARFIMA parameters in column *[b1]* and column *[a2]* suggesting lower values of the correlation exponents: $H \simeq 0.55$ and $d \simeq 0.05$. The lower values of the probability p indicate a more complex behavior of the NASDAQ with stronger deviation from the fully uncorrelated Brownian motion. By looking at Table 4, one can relate the NASDAQ behaviour to higher values of the long-range parameters of the ARFIMA model. In particular, the NASDAQ probability values become closer to parameter sets *[i2]* and *[n2]* corresponding to higher correlation degrees and correlation exponents around $H \simeq 0.75$ and $d \simeq 0.25$. The different

horizon dependence of NASDAQ and DJIA, where the former is a diversified stock market with a high degree of heterogeneity and the latter is an index representative of a chosen set of industrial stocks, is consistent with the ability of the cluster entropy index to quantify market heterogeneity.

Table 4. Probability p to reject the null hypothesis that the cluster entropy values for the ARFIMA processes at varying horizons M, have same mean and variance of the Fractional Brownian Motion with $H = 0.5$. The probability p has been estimated by standard *T-paired test*. First column reports the temporal horizon M. The other columns refers to parameter sets *[b1]*, *[f1]*, *[l1]*, *[a2]*, *[e2]*, *[i2]*, *[n2]*, *[o2]* of Tables 2 and 3.

M	*[b1]*	*[f1]*	*[l1]*	*[a2]*	*[e2]*	*[i2]*	*[n2]*	*[o2]*
1	0.9597	0.7938	0.6013	0.8519	0.6779	0.4956	0.3542	0.2314
2	0.9863	0.8429	0.6985	0.9293	0.7883	0.6566	0.5414	0.4304
3	0.9820	0.8789	0.7743	0.938	0.8346	0.7362	0.6468	0.5576
4	0.9848	0.8922	0.8031	0.956	0.8689	0.7827	0.7147	0.6380
5	0.9878	0.9062	0.8325	0.9608	0.8809	0.8102	0.7528	0.6911
6	0.9940	0.9197	0.8517	0.9724	0.9043	0.8417	0.7840	0.7322
7	0.9785	0.9186	0.8633	0.9617	0.9038	0.8521	0.8036	0.7614
8	0.9930	0.9321	0.8775	0.9762	0.9229	0.8710	0.8333	0.7931
9	0.9867	0.9370	0.8890	0.9737	0.9273	0.8809	0.8438	0.8100
10	0.9813	0.9333	0.8952	0.9710	0.9261	0.8880	0.8533	0.8195
11	0.9816	0.9436	0.9011	0.9749	0.9326	0.8965	0.8643	0.8342
12	0.9853	0.9451	0.9072	0.9741	0.9353	0.9019	0.8764	0.8508

Table 5. Probability p to reject the null hypothesis that the cluster entropy values for the NASDAQ, DJIA and S&P500 at varying horizons M have same mean and variance of the Fractional Brownian Motion with $H = 0.5$. First column reports the temporal horizon M. The probability p has been estimated by standard *T-paired test* [23].

M	**NASDAQ**	**S&P500**	**DJIA**
1	0.5154	0.7399	0.8892
2	0.6026	0.8335	0.9257
3	0.6470	0.8588	0.9332
4	0.6631	0.8814	0.9283
5	0.6823	0.9018	0.9417
6	0.7124	0.9246	0.9534
7	0.7162	0.9224	0.9461
8	0.7288	0.9309	0.9618
9	0.7370	0.9479	0.9645

The cluster entropy behavior appears deeply related to positive persistence and long-range correlation. In real-world financial series, horizon dependence deviates from the behaviour of fully uncorrelated series. The Market Dynamic Index, obtained via the integration of the cluster entropy curves, provides this feature in a cumulative, thus more robust, form. In conclusion, contrary to the assumptions of the traditional financial market theories, the hypothesis of efficient markets and rational investor behavior do not hold on account of the horizon dependence of the cluster entropy.

Author Contributions: Conceptualization, A.C.; Data curation, P.M.; Investigation, P.M.; Methodology, L.P. and A.C.; Resources, L.P.; Writing-review & editing, A.C. All authors have read and agreed to the published version of the manuscript.

Funding: Pietro Murialdo acknowledges financial support from FuturICT 2.0 a FLAG-ERA Initiative within the Joint Transnational Calls 2016, Grant Number: JTC-2016_004.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Grassberger, P.; Procaccia, I. Characterization of strange attractors. *Phys. Rev. Lett.* **1983**, *50*, 346. [CrossRef]
2. Crutchfield, J.P. Between order and chaos. *Nat. Phys.* **2012**, *8*, 17–24. [CrossRef]
3. Ormos, M.; Zibriczky, D. Entropy-based financial asset pricing. *PLoS ONE* **2014**, *9*, e115742. [CrossRef] [PubMed]
4. Yang, J. Information Theoretic Approaches in Economics. *J. Econ. Surv.* **2018**, *32*, 940–960. [CrossRef]
5. Ghosh, A.; Julliard, C.; Taylor, A.P. What Is the Consumption-CAPM Missing? An Information-Theoretic Framework for the Analysis of Asset Pricing Models. *Rev. Financ. Stud.* **2017**, *30*, 442–504. [CrossRef]
6. Backus, D.; Chernov, M.; Zin, S. Sources of entropy in representative agent models. *J. Financ.* **2014**, *69*, 51–99. [CrossRef]
7. Zhou, R.; Cai, R.; Tong, G. Applications of entropy in finance: A review. *Entropy* **2013**, *15*, 4909–4931. [CrossRef]
8. Shalizi, C.R.; Shalizi, K.L.; Haslinger, R. Quantifying self-organization with optimal predictors. *Phys. Rev. Lett.* **2004**, *93*, 118701. [CrossRef]
9. Carbone, A.; Castelli, G.; Stanley, H.E. Analysis of clusters formed by the moving average of a long-range correlated time series. *Phys. Rev. E* **2004**, *69*, 026105. [CrossRef]
10. Carbone, A.; Stanley, H.E. Scaling properties and entropy of long-range correlated time series. *Phys. A* **2007**, *384*, 21–24. [CrossRef]
11. Carbone, A. Information Measure for Long-Range Correlated Sequences: The Case of the 24 Human Chromosomes. *Sci. Rep.* **2013**, *3*, 2721. [CrossRef] [PubMed]
12. Zhao, X.; Sun, Y.; Li, X.; Shang, P. Multiscale transfer entropy: Measuring information transfer on multiple time scales. *Commun. Nonlinear Sci. Numer. Simul.* **2018**, *62*, 202–212. [CrossRef]
13. Humeau-Heurtier, A. The multiscale entropy algorithm and its variants: A review. *Entropy* **2015**, *17*, 3110–3123. [CrossRef]
14. Niu, H.; Wang, J. Quantifying complexity of financial short-term time series by composite multiscale entropy measure. *Commun. Nonlinear Sci. Numer. Simul.* **2015**, *22*, 375–382. [CrossRef]
15. Shannon, C.E. A mathematical theory of communication, Part I, Part II. *Bell Syst. Tech. J.* **1948**, *27*, 623–656. [CrossRef]
16. Kolmogorov, A.N. Three approaches to the quantitative definition ofinformation'. *Probl. Inf. Transm.* **1965**, *1*, 1–7.
17. Li, M.; Vitányi, P. *An Introduction to Kolmogorov Complexity and Its Applications*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 3.
18. Marcon, E.; Scotti, I.; Hérault, B.; Rossi, V.; Lang, G. Generalization of the partitioning of Shannon diversity. *PLoS ONE* **2014**, *9*, e90289. [CrossRef] [PubMed]
19. Rubido, N.; Grebogi, C.; Baptista, M.S. Entropy-based generating Markov partitions for complex systems. *Chaos* **2018**, *28*, 033611. [CrossRef]
20. Darbellay, G.A.; Vajda, I. Estimation of the information by an adaptive partitioning of the observation space. *IEEE Trans. Inf. Theory* **1999**, *45*, 1315–1321. [CrossRef]
21. Steuer, R.; Molgedey, L.; Ebeling, W.; Jimenez-Montaño, M.A. Entropy and optimal partition for data analysis. *Eur. Phys. J. B* **2001**, *19*, 265–269. [CrossRef]
22. Ponta, L.; Carbone, A. Information measure for financial time series: Quantifying short-term market heterogeneity. *Phys. A* **2018**, *510*, 132 – 144. [CrossRef]
23. Ponta, L.; Murialdo, P.; Carbone, A. Quantifying horizon dependence of asset prices: A cluster entropy approach. *arXiv* **2019**, arXiv:1908.00257. Available online: https://arxiv.org/abs/1908.00257 (accessed on 20 May 2020).
24. Vera-Valdés, J.E. On Long Memory Origins and Forecast Horizons. *J. Forecast.* **2020**. [CrossRef]
25. Graves, T.; Franzke, C.L.; Watkins, N.W.; Gramacy, R.B.; Tindale, E. Systematic inference of the long-range dependence and heavy-tail distribution parameters of ARFIMA models. *Phys. A* **2017**, *473*, 60–71. [CrossRef]
26. Bhattacharyya, Ranajoy, R.P.D. The Dynamics of India's Major Exchange Rates. *Glob. Bus. Rev.* **2020**. [CrossRef]

27. Bhardwaj, G.; Swanson, N. An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series. *J. Econom.* **2006**, *131*, 539–578. [CrossRef]
28. Baillie, R.T.; Kongcharoen, C.; Kapetanios, G. Prediction from ARFIMA models: Comparisons between MLE and semiparametric estimation procedures. *Int. J. Forecast.* **2012**, *28*, 46–53. [CrossRef]
29. Engle, R.F.; Rangel, J.G. The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Rev. Financ. Stud.* **2008**, *21*, 1187–1222. [CrossRef]
30. Lee, C.L.; Stevenson, S.; Lee, M.L. Low-frequency volatility of real estate securities and macroeconomic risk. *Account. Financ.* **2018**, *58*, 311–342. [CrossRef]
31. Adrian, T.; Rosenberg, J. Stock returns and volatility: Pricing the short-run and long-run components of market risk. *J. Financ.* **2008**, *63*, 2997–3030. [CrossRef]
32. Chernov, M.; Gallant, A.R.; Ghysels, E.; Tauchen, G. Alternative models for stock price dynamics. *J. Econom.* **2003**, *116*, 225–257. [CrossRef]
33. Cotter, J.; Stevenson, S. Modeling long memory in REITs. *Real Estate Econ.* **2008**, *36*, 533–554. [CrossRef]
34. Mandelbrot, B.B.; Van Ness, J.W. Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Rev.* **1968**, *10*, 422–437. [CrossRef]
35. Moving Average Cluster Entropy Code. Available online: https://www.dropbox.com/sh/9pfeltf2ks0ewjl/AACjuScK_gZxmyQ_mDFmGHoya?dl=0 (accessed on 20 May 2020).
36. Geometric Brownian Motion Code. Available online: https://it.mathworks.com/help/finance/gbm.html (accessed on 20 May 2020).
37. Generalized Autoregressive Conditional Hetereskedastic Code. Available online: https://www.mathworks.com/help/econ/garch.html (accessed on 20 May 2020).
38. Fractional Brownian Motion Code. Available online: https://project.inria.fr/fraclab/ (accessed on 20 May 2020).
39. Autoregressive Fractional Integrated Moving Average Code. Available online: https://www.mathworks.com/matlabcentral/fileexchange/25611-arfima-simulations (accessed on 20 May 2020).

entropy

MDPI

Article

Network Analysis of Multivariate Transfer Entropy of Cryptocurrencies in Times of Turbulence

Andrés García-Medina [1,2,*] and José B. Hernández C. [1,3]

[1] Unidad Monterrey, Centro de Investigación en Matemáticas, A.C. Av. Alianza Centro 502, PIIT, Apodaca 66628, Nuevo Leon, Mexico; jose.chaudary@cimat.mx
[2] Consejo Nacional de Ciencia y Tecnología, Av. Insurgentes Sur 1582, Col. Crédito Constructor, Ciudad de México 03940 , Mexico
[3] Escuela de Matemática, Facultad de Ciencias, Universidad Central de Venezuela, Av. Los Ilustres, Los Chaguaramos, Caracas 1020, Venezuela
* Correspondence: andres.garcia@cimat.mx

Received: 10 June 2020; Accepted: 8 July 2020; Published: 11 July 2020

Abstract: We investigate the effects of the recent financial turbulence of 2020 on the market of cryptocurrencies taking into account the hourly price and volume of transactions from December 2019 to April 2020. The data were subdivided into time frames and analyzed the directed network generated by the estimation of the multivariate transfer entropy. The approach followed here is based on a greedy algorithm and multiple hypothesis testing. Then, we explored the clustering coefficient and the degree distributions of nodes for each subperiod. It is found the clustering coefficient increases dramatically in March and coincides with the most severe fall of the recent worldwide stock markets crash. Further, the log-likelihood in all cases bent over a power law distribution, with a higher estimated power during the period of major financial contraction. Our results suggest the financial turbulence induce a higher flow of information on the cryptocurrency market in the sense of a higher clustering coefficient and complexity of the network. Hence, the complex properties of the multivariate transfer entropy network may provide early warning signals of increasing systematic risk in turbulence times of the cryptocurrency markets.

Keywords: cryptocurrencies; multivariate transfer entropy; complex networks

1. Introduction

Cryptocurrencies are new financial instruments based on the technology of blockchains. A coin is defined as a chain of digital signatures. In Bitcoin, each owner transfers the coin by digitally signing a hash of the previous transaction and the public key of the next owner, adding them to the end of the coin [1]. A cryptocurrency exchange platform is a website on which we can buy and sell coins for other digital currency or trust money. Depending on the exchange, they can operate as a stock exchange or as a currency exchange house, which is very effective and safe for users. The easy access of this market through more than 22,000 projects operating within the industry, exchanges with low fees of transactions, more than 5000 virtual coins worldwide, and a daily trade volume of nearly 174 billion dollars have done cryptocurrencies a very attractive instrument of investment for the general population [2].

The interactions of cryptocurrencies are not always identical among all variables, in other words, the variables influence each other with different magnitude. Thus, it is necessary to study the asymmetric dependence structure to understand such interactions. The common measures to estimate dependencies, i.e., linear cross-correlation, cross-spectra, and mutual information; share the characteristic to be symmetric in nature. The usual approach to understanding asymmetric dependencies is through the parametric approach of copulas [3]. Nevertheless, we are interested in

an approach from the point of view of information theory in order to take into account the concept of causality. Even though there exist also attempts to model causality via copula-based methods [4], the information approach works on the broadest sense of free modeling.

The most popular measures of asymmetric dependencies in information theory are related to conditional information or Transfer entropy (TE) based on the concept of Shannon entropy [5]. This quantity was introduced originally with the purpose to quantify the statistical coherence between systems evolving in time [6]. Since then transfer entropy has been used to solve problems of different nature. It has been effective in the study of the neuronal cortex of the brain [7], statistical physics [8], dynamical systems [9], given a thermodynamic interpretation in [10].

In applications to econometrics, transfer entropy can be regarded as a nonlinear generalization of the Granger causality test [11]. There exists a series of results [11–13] that state an exact equivalence between the Granger causality and TE statistics for various approaches and assumptions on the data generating processes, which make it possible to construct TE as a non-parametric test of pure Granger causality. In a previous study [14], we compared a synthetic linear and non-linear models, and an empirical data set of cryptocurrencies, where is highlighted the advantage of the symbolic estimation of TE over traditional Granger causality test. Moreover, in [15] the multivariate version of symbolic transfer entropy has been tested, the authors show that it can be applicable to non-stationary time series in mean and variance and is even unaffected by the existence of outliers and vector autoregressive filtering.

Despite this characteristic, the main deficiency of the general TE as a measure of causality is the possibility of spurious causalities due to indirect influences or common drivers. To explain this point consider the processes X, Y, and Z; if a causal interaction is given by $X \to Y \to Z$, a bivariate analysis would give a significant link between X and Z that is detected as being only indirect in a multivariate analysis [16]. An approach to overcome this issue is proposed in [17] by inferring an effective network structure given multivariate interactions using a greedy algorithm. The authors of [18] improved the methodology by adding a preliminary step to prune the set of sources and implementing a series of rigorous statistical tests to decrease the type I and II errors emerged in the multiple comparisons involved on the computations of multivariate transfer entropy. The approach of these studies employs conditional and collective forms of multivariate transfer entropy [19,20].

On the other hand, complex system theory has a long tradition between physicists. The emergence of regularities have been observed in different systems and theories, such as convection, turbulence, phase transition, nonlinear dynamics, renormalization theory, among other fields of physics [21]. However, there is no strict mathematical definition of what complexity is, rather it is characterized according to the properties presented in the particular system. These properties are commonly referred as scale invariance, self-organized criticality, hierarchical structure, coexistence of collective effects and noise, variability and adaptability, and highly optimized tolerance [22].

Network representation has become a common approach to represent interactions between elements in complex systems. This tool allows characterizing the properties of different phenomenan in a common framework. Graph theory has become an essential piece to the understanding of the structure and behavior of these systems. Based on it has been possible to discover emerging properties having fundamental implications on different areas of knowledge [23]. The representation of the financial markets through networks and the study of their complex structure draw attention in physics since the seminal work [24]. There, it is introduced the ultrametric distance and the Minimum Spanning Tree (MST) to characterize the correlations between the stocks used to compute the Dow Jones Industrial Average (DJIA) index, and the portfolio of stocks used to compute the Standard and Poor's 500 (S&P 500). Also, in [25] it is extended the methodology to study a portfolio of equities at different time horizons, as well as the MST structure of volatilities comparing the network's topological properties of real and artificial markets.

Subsequently, a series of works have emerged that describes an analogy between the foreign exchange market (or Forex) and the cryptocurrency market, using as a framework the tools that

characterize complex networks. To begin with, the work done on [26] studied a large collection of daily time series for world currencies' exchange rates through MST methodology. They find an autoscaling behavior in the degree distribution of the network, and demonstrate the existence of a hierarchical structure in the currency markets by developing an analytical model. In the simultaneous works [27,28] the authors show that the structure of the Forex network depends on the base currency. In addition, they found the network is not stable in time, noting the USD node gradually loses its centrality, while the EUR node starts turning more central during the study period. In relation to cryptocurrencies, the study [29] analyzes the return distribution, volatility autocorrelation, Hurst exponents, and the effects of multiscaling for the Bitcoin market. There, the authors find that this market show signs of maturity during the last months of the analysis, whose characteristics resemble the complex properties of the Forex market. The analysis done in [30], found further evidence of the shared features of Forex and Cryptocurrency markets at high-frequency rate. Besides, it is pointing out the BTC/ETH and EUR/USD exchange rates do not show any significant relationships. Thus, they hypothesize both markets start decoupling. The same authors extend their study to 100 cryptocurrencies, introducing the collective analysis of random matrices [31]. They found that the level of collectivity depends on which cryptocurrency is used as the exchange rate. Moreover, it is detected that the USD begins to disconnect from the network and resemble a fictitious currency, which may imply the cryptocurrencies' autonomy. Last, the work of [32] studies the relationship between both markets from a more econometric perspective, where it is explained the importance of diversifying between them.

The previous results pave the way to carry out extensive studies on cryptocurrencies without having to link them to the forex markets, despite the fact both markets share several similarities. In this spirit we are interested to study the complexity properties that emerge in the induced network by multivariate entropy transfer when considering the price and volume of transaction. In this sense, we are seeking to characterize their asymmetric interactions by applying a series of statistical tests with the intention of considering only significant connections. To the aim of bias reduction, we are looking for multivariate rather than bivariate approach. Likewise, it is desired to describe the turbulence observed during March 2020 by interpreting the temporal behavior of the directed interactions via the complex network's artifacts. In this regard, the turmoil of cryptocurrencies has been analyzed in [33] from the concept of *bull and bear* market. The authors study the three largest cryptocurrencies of Bitcoin, Ethereum, and Litecoin through the technique of Detrended Fluctuation Analysis (DFA) analyzing the Hurst exponent over a different time windows. They find that during the *bull* period the market is efficient, whereas in *bear* times it is inefficient.

Furthermore, there exist literature discussing the spillover effect and systematic risk among the cryptocurrency markets. In [34] it is found that the structural breaks are universally present in seven of the largest cryptocurrencies, whereas it is spreads from the smallest to the largest currencies, in order of capitalization. This finding is done by implementing the Granger causality test, as well as a test for the ARCH and the Dynamic Conditional Correlation MGARCH to the selected coins. Furthermore, the work [35] show evidence that Ethereum is likely to be the independent coin in the cryptocurrency markets, while Bitcoin tends to be the spillover effect recipient. There, the author modeled the system by variants of the Vector Autoregressive Model (VAR), and using jointly distributed Student's-t copulas to measure the risk contagion among cryptocurrencies. Moreover, the study [36] found a risk contagion effect between cryptocurrencies when applying a copula approach. There, it is suggested to perform portfolio diversification to avoid this phenomenon.

On the perspective of transfer entropy, the study of Sandoval [37] uses the information measure to characterize the contagion of institutions in times of crisis. He identifies the companies most vulnerable to be contagious on countries that have suffered sovereign default. A recent study in similar direction is [38], where transfer entropy is estimated by discretizing the return time series into positive and negative values and validated by resampling. The researchers constructed an indicator to measure the systematic risk on the stock market and real state data. They observed the networks manifest strongly connectivity around periods of high volatility around the crash of 2008. Further, the authors of [39]

apply the Rényi Transfer Entropy to investigate the interactions between the crude oil markets and the cryptocurrency markets. Their results suggest that the macroscopic economic value of the US crude oil has an contagion effect on the cryptocurrency markets.

In this study, we analyze the cryptocurrency network induced by the estimation of the multivariate transfer entropy as proposed in [18]. We are especially interested to understand the effects of the financial turbulence of 2020 on the market of cryptocurrencies taking into account the price and volume of transactions as a variable of interest. To obtain deeper insights about the structure of the induced network we quantify the clustering coefficient and estimate the degree distributions of nodes, which are two standard tools from complex networks [40]. Our work follows the line of thought of the literature discussed above, specifically in the sense of studying the systematic risk and contagion between the currencies through the transfer entropy when the cryptocurrency market is in a turbulent situation. Likewise, the March 2020 turmoil is explored with network theory's tools, which have had a fundamental role in the econophysics interpretation of the complex systems that emerge in finance.

However, our work is distinguished in combining a series of elements that give rise to ask and exploring the following questions, which as far as we know have not yet been discussed in the literature: Do the directed networks associated to the multivariate transfer entropy of volume and price of cryptocurrencies present complex properties? Is it possible to characterize and in some sense to anticipate the turmoils on the cryptocurrency markets through the properties of the network induced by the multivariate transfer entropy? Can the clustering coefficients of these networks play the role of an early indicator of turbulence in these markets? Is there self-similarity in the induced networks, and if so, how do we interpret this characteristic in turbulent times?

We hypothesize the complex properties of the multivariate transfer entropy network may provide early warning signals of increasing systematic risk. This is inferring through evidence found during the turbulence of March 2020 for the induced directed networks by the multivariate information measure of the hourly volume and price of 146 coins from December 2019 to April 2020. We hope, these results may help the practitioners who venture to invest in this risky class of financial instruments to have further quantitative tools to assess systematic risk during times of turbulence.

The next Section 2 describes the data under study. Section 3 presents the greedy algorithm and the series of statistical tests involved in the computation of multivariate transfer entropy, as well as preliminary results. In Section 4 the network dynamics characterized by the clustering coefficient and power law fitness are analyzed in the context of the generated network of cryptocurrencies. Finally, Section 5 highlights the implications of the results and future work is proposed.

2. Data

We consider the price and volume in dollars of $p = 146$ cryptocurrencies, which are obtained using the API of CoinMarketCap [2] for the period from 00:00 2019-12-01 to 00:00 2020-04-05, at an hourly frequency, resulting in a total of $n = 3025$ observations (see Supplementary Material). The data acquisition and preprocessing strategy consisted in creating a database on a remote Linux server and automatically make calls to the CoinMarketCap API (in UTC time) to request the price and volume of the first 200 coins in order of capitalization on the day the database was built, i.e., on the 9th November 2019. The frequency and number of currencies were chosen in such a way that it did not exceed the number of requests or credits allowed by the API. The intention was to fulfill the commitment to have the maximum number of coins at the highest possible frequency. The cryptocurrencies are selected under the condition of having each less of 1% of missing values in volume and price during the trading period. In case of no record, a spline interpolation of order three was used to fill the time series gaps. Thus, our set of variables dropped to the reduced variables stated above ($p = 146$). In addition,

we transformed the $2p$ time series into price-returns $r_k^{(price)}(t)$, and volume-returns $r_k^{(volume)}(t)$ by computing the logarithmic difference of consecutive observations

$$r_k^{(price)}(t) = \log(s_k(t)) - \log(s_k(t-1)) \quad (1)$$

$$r_k^{(volume)}(t) = \log(v_k(t)) - \log(v_k(t-1)) \quad (2)$$

where $k = 1, \ldots, p$; $t = 1, \ldots, n$; and $s_k(t)$, $v_k(t)$ stands for the price and volume of cryptocurrency k at time t, respectively. In this way, we deal only with the stochastic part of the time series.

Even though the general approximation of the transfer entropy is non-parametric, we are going to use a Gaussian estimator for the conditional distributions as will be explained in the next section. Then, it is necessary to ensure that the data is stationary. For this purpose, it is a usual practice to take as an input the logarithmic returns to satisfy this requirement. Another approach is not to use excessively long time series, even though it may create a compromise between the bias of the transfer entropy estimator and stationarity. In this study, in addition to using moderately long log-returns time series, we have verified stationarity using the augmented Dickey-Fuller test [41] and the Phillips-Perron test [42]. In both tests it has been obtained a p-value less than 0.001 for all the $2p$ time series considered. Therefore, it is confirmed no evidence of a unit root in any variable.

Further, it is known that the distribution of financial and cryptocurrency data change according to the resolution that we observe them [43–45]. In this sense, it is important to keep in mind the implications of our results are only valid for hourly observations. Our interest in this frequency lies in the high volume of transactions that take place intra-day on the market of cryptocurrencies. However, we did not want to go to higher frequencies (minutes) due to two situations. The first is that volatility increases and it is more difficult to justify stationarity. Second, the justification for using a Gaussian estimator to calculate the entropy transfer would be invalided because the distribution of the logarithmic returns increases its kurtosis and their distribution start resembles one of the Levy family [45]. On the other hand, we are interested to delimit our work for the period of time just before the pandemic effects take place. Consequently, we have the limitation of having not too many days of transactions around this event, so the sample would not be large enough to obtain an unbiased estimation of multivariate transfer entropy if used daily time series.

Finally, it is important to mention that we are working with non-traded prices. Coinmarketcap is known as a *coin-ranking* service because they rank both coins and exchanges by trading volume and market capitalization to weight their cryptocurrencies data. The specific strategy followed by CoinMarketCap can be found in [46]. This type of information would represent a problem if we were interested in proposing an optimal portfolio, hedging strategies, or trading applications, as is properly warned in [47]. Nonetheless, it is not the goal of this study, we are interested in the weighted information provided by CoinMarketCap, since it estimates the formation of prices considering important features of the cryptocurrency data and selected exchanges. The last is more in tune with the vision of complex systems. In other words, the objective of our study is in essence explanatory rather than predictive. Notwithstanding, it should be kept in mind the results are conditioned to the way in which CoinMarketCap weighted the information.

3. Multivariate Transfer Entropy

The Transfer Entropy (TE) from a process X to a process Y measures the amount of uncertainty reduced in future values of X by knowing the past states of Y and X itself. In other words, transfer Entropy (TE) quantifies the amount of information that the past of a source process X provides about the current value of a target process $Y = y_t$, considering the context of Y's own past. In a multivariate setting, a set of sources $X_i, i = 1, \ldots, M$ is provided. The multivariate Transfer Entropy (mTE) from X_i to Y can be defined as the information that the past of X_i provides about $Y = y_t$, in the context of both Y's past and all the other relevant sources in X. The main challenge is to define and identify the relevant sources. In principle, the mTE from X_i to Y is computed by conditioning on

all the other sources in the network, i.e., $X \setminus X_i$. However, in practice, the sizes of the conditioning set must be reduced in order to avoid the curse of dimensionality. The idea is to restrict the conditioning set by finding and including the sources that participate with X_i in reducing the uncertainty about $Y = y_t$, in the context of Y's own past. The set of relevant sources will be denoted as Z [48].

In order to infer Z from X it is followed the greedy algorithm approach suggested by [17,18], where Z is built iteratively by maximizing the conditional mutual information (CMI) criterion. As a first step, a set of candidate variables $c \in C$ is defined from the past values of X; then, Z is built including iteratively the candidate variables c that provides statistically significant maximum information about the current value $Y = y_t$, conditioned on all the variables that have already been selected. More formally, at each iteration, the algorithm selects the past variable $c \in C$ that maximizes the conditional mutual information $I(c, y_t|Z)$ at significance level α of a maximum information test. The set of selected variables forms a multivariate, *non-uniform embedding* of X [49] (See Appendix A).

The implemented greedy algorithm operates in several steps. First, Z is initialized as an empty set and is considered the candidate sets for Y's past C_Y and X's past C_X. Second, variables from C_Y are selected. To this end, for each candidate, $c \in C_Y$, it is estimated the information contribution to $Y = y_t$ as $I(c, y_t|Z)$. Next, it is found the candidate with maximum information contribution, c^*, and is performed a significance test; if it is found significant, must be added c^* to Z and remove it from C_Y. In addition, the *maximum statistic* (see Appendix B.1) is used to test for significance while controlling the Family-Wise Error Rate (FWER). Then, the algorithm stops if $I(c^*, y_t|Z)$ is not significant or C_Y is empty. Third, variables are selected from X's past states, i.e., iteratively candidates in C_X are tested following the procedure described before. Fourth, redundant variables in Z are tested and removed using the *minimum statistic* (see Appendix B.2). The minimum statistic computes the conditional mutual information between every selected variable in Z and the current value, conditional on all the remaining variables in Z. This test is performed to ensure that the variables included in the early steps of the iteration still provide a significant information contribution in the context of the final parent set Z [17]. Fifth, the *omnibus test* is performed (see Appendix B.3) to test the collective information transfer from all the relevant source variables to the target $I(Z_X, y_t|Z_Y)$. The resulting omnibus p-value is later used for the correction of the FWER at the network level. If the omnibus test is significant, a sequential maximum statistic is performed on each selected variable $z \in Z$ to obtain the final information contribution and p-value for each variable.

The mTE between a single source X_i and target Y can be computed from the inferred non-uniform embedding Z. To this end, it is collected from Z all of X_i's selected past variables, X_i, and calculated the mTE as $I(X_i; y_t|Z \setminus X_i)$. Note that time lag between X_i's selected past variables and the current value at time t indicates the information-transfer delay [50]. The delay can be estimated as the lag of the past variable which provides the maximum individual information contribution, where the maximum contribution is indicated either by the maximum raw TE estimate or by the minimum p-value over all variables from the source. Finally, the algorithm must be repeated for every node (or variable) in the network (see [18] for an extensive description of the mTE algorithm and hypothesis testing).

mTE Network of Cryptocurrencies

We consider each time series of price-returns and volume-returns of cryptocurrencies to be a stochastic process in order to detect the causal relations between the variables. The procedure is to fix the target variable $Y = X_i$, the source set as $X \setminus X_i$, and apply the mTE algorithm [48] described above for each time series $i = 1, \ldots, 2p$, where the first p variables represent the price-returns and the last p variables the volume-returns, where $p = 146$ as described in the data section. The intention is to detect if the causal relations of price-returns and volume-returns bring clues to understand the dynamic of cryptocurrency market in times of turbulence. Hence, we design a temporal analysis of time windows of 21 days, sliding by seven days, and using an overlapping of 14 days. Under this procedure, it is obtained $k = 16$ time windows, the first from the 01:00 of 2019-12-01 to 00:00 of 2019-12-22, and the last

from 01:00 of 2020-03-15 to 00:00 of 2020-04-05. Thus, each data frame contain $q = 504$ observations of hourly price-returns and volume-returns, i.e., having dimensions $q \times 2p$.

It is of primary importance to estimate the CMI, represented above as $I(X_i, y_t | Z \setminus X_i)$, in order to quantify mTE, where $i = 1, \ldots, 2p$. We estimate CMI under the assumption that price-returns and volume-returns follow a jointly Gaussian continuous distribution, which is equivalent to assume a linear causal dependency as proposed by Granger [51]. The value added to the original Granger causality test is the use of a multivariate framework to test statistically significant causal relations conditioned to other variables, while the original test only measures bivariate dependencies. We chose Gaussian estimator over a more realistic distribution approach due to the dimension settings as well as the sliding windows turn the computational complexity excessively time demanding. Nevertheless, the results assuming linear dependencies are worthwhile to mention due to the capacity of mTE to detect conditional interactions.

Figure 1 shows the resulting networks for the mTE analysis of cryptocurrencies described above using the Gaussian estimator. Here we set the number of permutations for the surrogate distribution used in the statistical tests (*maximum*, *minimum*, and *omnibus*) to 500, and a significance level of 0.05. The alphabetic order of the subfigures represents the temporal order of the time window of each experiment. Thus, for example, the subfigure (a) shows the results for the time frame 2019-12-01–2019-12-22, the subfigure (b) for 2019-12-08–2019-12-29, and so on. The graph visualization is made using the Kamada-Kawai algorithm [52], where each variable is represented as a node. We have discriminated the price and volume variables separating their corresponding nodes by a fixed distance to the upper right if it corresponds to a price-return node, and to the lower left if it corresponds to a volume-return node. In addition, the price nodes are colored in green, while the volume ones in red. The directed edges represent true causal relations under the mTE methodology, where we have tested for one to three lags, and as a result, it is obtained a binary adjacency matrix A. This matrix has elements $A_{ij} = 1$ if there exit a statistically significant causal relation in this range of lags, and $A_{ij} = 0$ on the contrary, where $i, j = 1, \ldots, 2p$. Finally, the size of each node is drawn according to its clustering coefficient, which is described in the next section. However, here it is already noting the increasing size of the green nodes for the cases (m)–(o), which correspond to the subnetwork of price-returns for the time windows ending at 2020-03-15, 2020-03-22, and 2020-03-29, respectively.

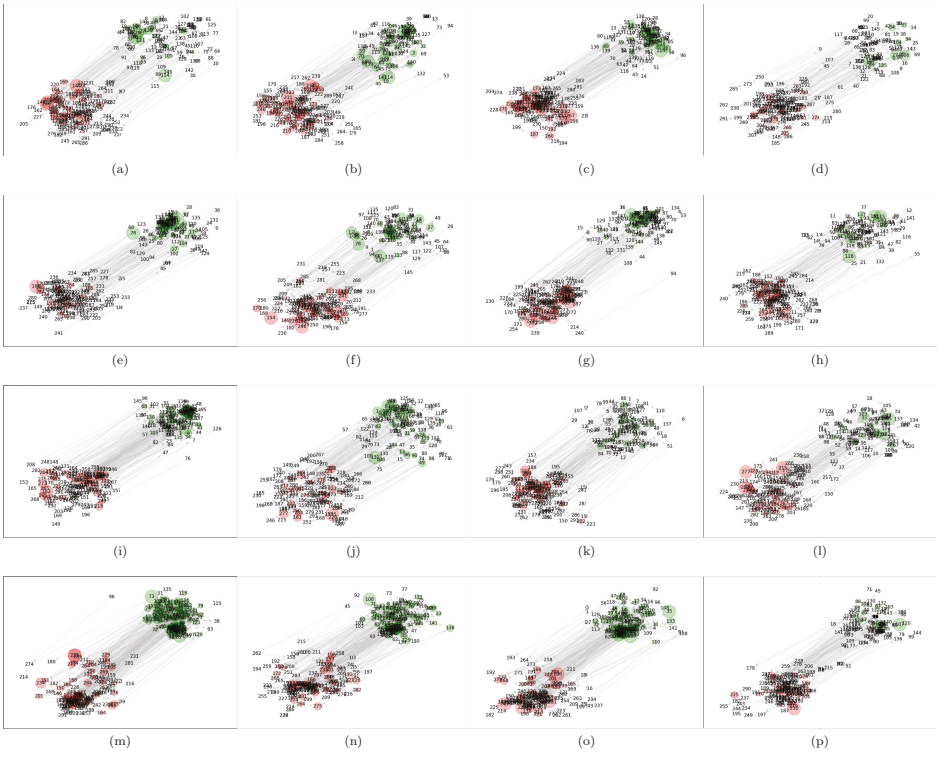


Figure 1. Network representation of mTE results for cryptocurrency variables. The green nodes represent price-return variables, while the red nodes represent volume-return variables. The subfigures (**a**–**p**) show the directed network results for the time window under study. The subfigures are arranged in temporal order from top-left to right-bottom.

4. Complexity Behavior

In this section, we describe and quantify for the data sets under study two essential quantities in complex network theory: clustering coefficient and degree distribution. The former in order to measure cliques or connectivity and the latter to estimate the scale-free property of the network.

4.1. Clustering Coefficient

The tendency of a network to form tightly connected neighborhoods, more than in the random uncorrelated case, can be measured by the clustering coefficient (CC) [53]. Consider a network described by the graph $G = (V, A)$, where V is the number of nodes, and $A = \{a_{ij}\}$ is the $V \times V$ adjacency matrix. The orientation of edges or arrows in directed networks turn the adjacency matrix to be non-symmetrical in general. In a binary directed networks, as are the induced networks by mTE, the node out-degree k_i^{out} is the number of edges pointing out from the node i, while the node in-degree k_i^{in} is the number of edges pointing towards the node i. Formally

$$k_i^{out} = \sum_{j \neq i} a_{ij} = (A)_i \mathbf{1} \quad (3)$$

$$k_i^{in} = \sum_{j \neq i} a_{ji} = (A')_i \mathbf{1}, \quad (4)$$

where A' is the transpose of A, $(A)_i$ is the ith row of A, and $\mathbf{1}$ represents the column vector $(1, 1, \ldots, 1)'$ of dimension V. Similarly, the total degree k_i^{tot} of node i is the sum of its in-degree and out-degree

$$k_i^{tot} = k_i^{in} + k_i^{out} = (A' + A)_i \mathbf{1}. \tag{5}$$

Further, assuming that no self-interactions are present, the bilateral edges $k_i^{\leftrightarrow}$ between i and its neighbors are counted as

$$k_i^{\leftrightarrow} = \sum_{j \neq i} a_{ij} a_{ji} = A_{ii}^2. \tag{6}$$

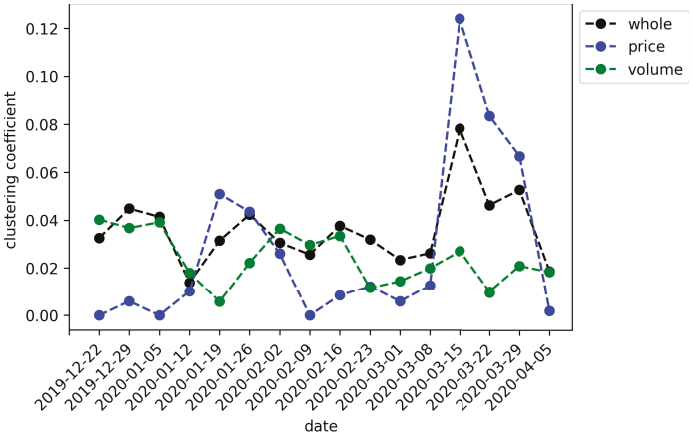
The usual approach to quantify the clustering coefficient of node i is by measuring the ratio of the number of triangles in the graph G with i as one vertex over the number of all possible triangles that i could form. In the directed graph case, the clustering coefficient C_i for node i can be explicitly computed by the expression [54]

$$C_i = \frac{(A + A')_{ii}^3}{2[k_i^{tot}(k_i^{tot} - 1) - 2k_i^{\leftrightarrow}]}, \tag{7}$$

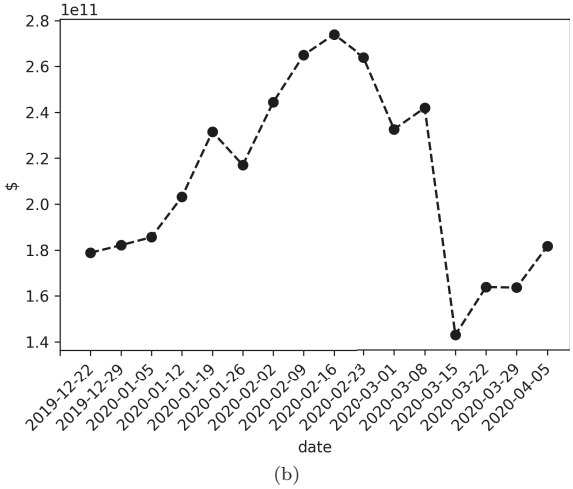
and the overall clustering coefficient C for the directed graph G is obtained by $C = V^{-1} \sum_{i=1}^{V} C_i$, where $C \in [0, 1]$.

Figure 2a shows the overall clustering coefficient as a function of time for the networks associated with each time window of the data set. The black dotted line are the results for the whole networks (i.e., volume-return and price-return nodes), whereas the blue and green dotted lines represent the results for the price-return and volume-return nodes, respectively. It is remarkable the peak in the clustering coefficient for the time window ending at 2020-03-15, especially in the price-return subnetwork and in less amount for the whole network.

In order to gain insight that leads us to a better understanding of this behavior, it is superposed the market capitalization at the end of the time window of each experiment. Thus, Figure 2b shows the average market capitalization in dollars, taking into account the observed values of each cryptocurrency of the data set. It is worth noting the coincidence of the fall in the average of market capitalization with the increase of the overall clustering coefficient for the whole and price-returns networks at date 2020-03-15. This abrupt change in both behaviors is situated at the period of most severe contraction of the global stock markets due to the recent COVID-19 worldwide pandemic. These results suggest that the financial turbulence induce a higher flow of information to the cryptocurrency market on its price-returns features and as a consequence the overall clustering coefficient increase. In the next subsection is analyzed the degree distribution of the networks with the intention to sustain this hypothesis from another angle.



(a)



(b)

Figure 2. (**a**) Overall clustering coefficient as a function of time. Black dotted lines show the results for the whole networks, blue and green dotted lines show results for the price-return and volume-return nodes, respectively. (**b**) Market capitalization averaged over the cryptocurrencies under study at the end of the time window. The units in the y-axis are measured in hundred million dollars.

4.2. Power Law

Power laws are probability distributions with the form $p(x) \propto x^{-\alpha}$, which have the characteristic of being heavy-tailed. This feature may be so extreme that standard deviation only can be defined if $\alpha < 3$, or the mean if $\alpha < 2$. Phenomena following a power law distribution are known as scale-free systems because all values are expected to occur, without a characteristic size or scale. These kinds of distributions have been identified throughout nature, including physics and finance [44,55–57] to cite some relevant examples in the context of this work.

We use the `powerlaw` library [58] to fit a power law to a degree, in-degree, and out-degree distributions for the induced graph by the mTE flow of information of each time window under study. The results are shown in Figure 3 for the whole network, as well as the price-returns and volume-returns

subgraphs, where the vertical lines represent the standard deviation of the adjustments. Also, it is plotted the *p*-value of fitness using the log-likelihood ratio test, wherein all cases the null hypothesis assumes the distributions are characterized by an exponential function. Tables 1–3 show the numerical results associated to Figure 3.

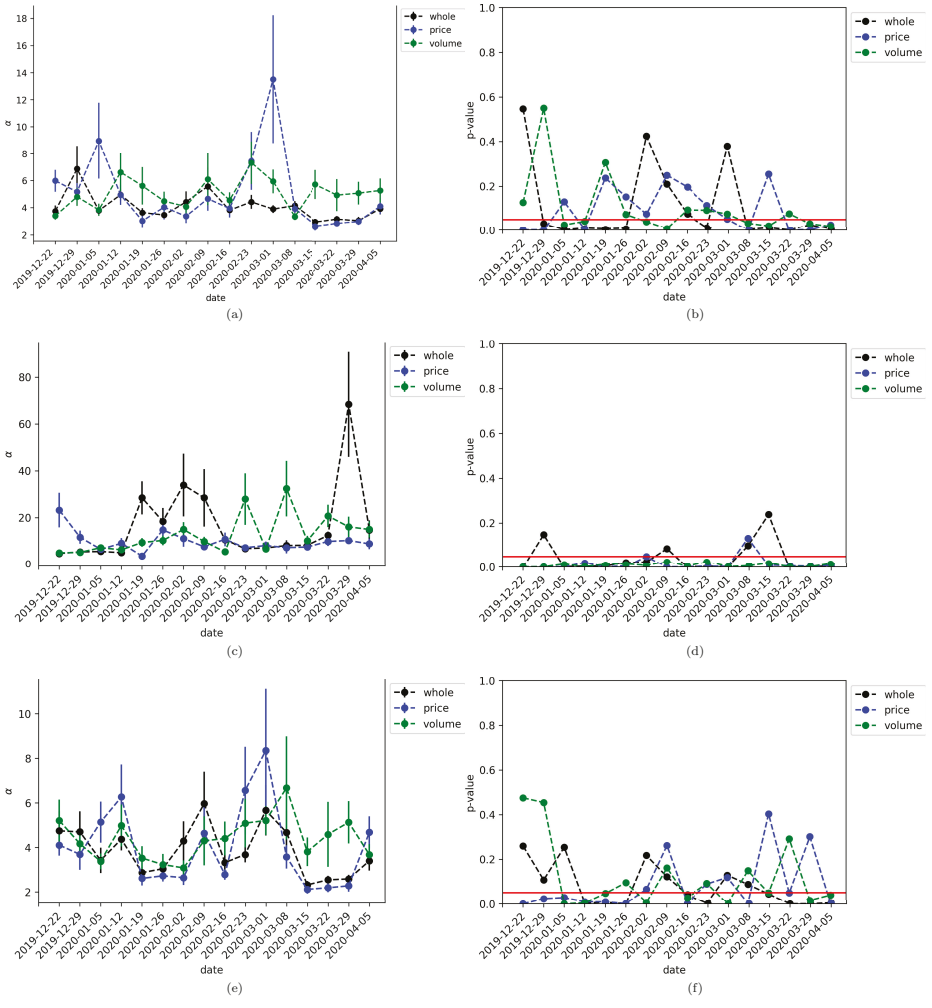


Figure 3. The Behavior of the mTE network structure through time. (**a**,**b**) show the dynamic of the estimated power α and corresponding *p*-value for the degree distribution. (**c**,**d**) show the dynamic of the estimated power α and corresponding *p*-value for the in-degree distribution. (**e**,**f**) show the dynamic of the estimated power α and corresponding *p*-value for the out-degree distribution. In all cases, the black line represents the results for the whole network, the blue line for the price-returns subgraph, and the green line for the volume-returns subgraph. The red straight line is the significance threshold of 0.05 and the vertical lines centered at each point represents the standard deviations.

Table 1. Degree distribution estimation.

Date	α (Whole)	p-Value	α (Price)	p-Value	α (Volume)	p-Value
2019-12-22	3.738	0.5464	5.9778	0.0006	3.3814	0.1268
2019-12-29	6.8615	0.0314	5.1602	0.0006	4.7913	0.5496
2020-01-05	3.7785	0.0011	8.9531	0.1303	3.8275	0.026
2020-01-12	4.9396	0.0085	4.9325	0.0057	6.619	0.0423
2020-01-19	3.6251	0.0057	3.0092	0.2377	5.603	0.3073
2020-01-26	3.4424	0.0081	4.031	0.1531	4.4695	0.0732
2020-02-02	4.4084	0.424	3.3453	0.075	4.0688	0.0401
2020-02-09	5.5648	0.2106	4.6502	0.2502	6.0909	0.0094
2020-02-16	3.8022	0.0759	3.9331	0.1972	4.5331	0.0938
2020-02-23	4.4157	0.0128	7.455	0.1146	7.3266	0.0924
2020-03-01	3.8864	0.3795	13.5086	0.052	5.9385	0.0754
2020-03-08	4.1438	0.0036	3.9098	0.0005	3.3258	0.0334
2020-03-15	2.9365	0.0096	2.6145	0.2548	5.7256	0.0212
2020-03-22	3.1472	0	2.8274	0.0004	4.9333	0.0764
2020-03-29	3.0356	0.0267	2.9792	0.0006	5.0786	0.0313
2020-04-05	3.9247	0.0041	4.1092	0.0255	5.2757	0.0193

Table 2. In-degree distribution estimation.

Date	α (Whole)	p-Value	α (Price)	p-Value	α (Volume)	p-Value
2019-12-22	4.5806	0.0002	23.1967	0.0001	4.8769	0
2019-12-29	5.2993	0.1479	11.592	0.0001	5.076	0.0002
2020-01-05	5.4633	0.0009	6.4081	0	7.1131	0.0132
2020-01-12	4.9525	0.0029	8.9783	0.0203	6.1964	0.0015
2020-01-19	28.4241	0.0029	3.4932	0	9.3489	0.0117
2020-01-26	18.3479	0.0222	14.7033	0.0042	10.2065	0.0141
2020-02-02	33.9089	0.025	11.0187	0.0498	14.9042	0.0068
2020-02-09	28.4241	0.0855	7.5029	0	9.7462	0.0251
2020-02-16	10.5242	0.0002	10.6918	0.0003	5.3289	0.0042
2020-02-23	6.695	0.0025	7.1155	0.003	27.8885	0.0243
2020-03-01	6.9885	0.0015	8.0237	0.0018	6.5015	0.0014
2020-03-08	8.0302	0.0981	7.1173	0.1321	32.3699	0.0056
2020-03-15	8.135	0.2385	7.4099	0.0083	10.02	0.0182
2020-03-22	12.3518	0.0001	9.7665	0.0045	20.6977	0.0003
2020-03-29	68.3999	0.0002	10.1634	0.0035	16.028	0.0014
2020-04-05	14.6396	0.0042	8.732	0.007	14.9042	0.0155

Table 3. Out-degree distribution estimation.

Date	α (Whole)	p-Value	α (Price)	p-Value	α (Volume)	p-Value
2019-12-22	4.7432	0.2594	4.1039	0.0001	5.208	0.4752
2019-12-29	4.698	0.1075	3.6869	0.0229	4.1616	0.4543
2020-01-05	3.4208	0.2542	5.1354	0.0276	3.3738	0.0002
2020-01-12	4.3711	0.0037	6.265	0.009	4.979	0.004
2020-01-19	2.8673	0.0064	2.6074	0.009	3.5196	0.0464
2020-01-26	3.0402	0	2.7225	0.0011	3.2296	0.0955
2020-02-02	4.2886	0.2173	2.6393	0.0657	3.0792	0.0057
2020-02-09	5.9639	0.1216	4.6401	0.2623	4.2951	0.1614
2020-02-16	3.3214	0.0411	2.7803	0.0019	4.3985	0.0263
2020-02-23	3.6758	0.0011	6.5573	0.089	5.0874	0.0921
2020-03-01	5.6642	0.1289	8.3496	0.1186	5.2096	0
2020-03-08	4.6742	0.0873	3.5777	0.0019	6.6722	0.1489
2020-03-15	2.3148	0.0437	2.1134	0.4035	3.8052	0.0488
2020-03-22	2.542	0	2.1833	0.0486	4.5843	0.2916
2020-03-29	2.5827	0.0003	2.2731	0.3013	5.127	0.0152
2020-04-05	3.3965	0.0045	4.6771	0.0022	3.6785	0.0392

Let analyze first the degree distribution case. As can be seen from Figure 3a there is a peak at 2020-03-01 for the price-returns subgraph, where the estimated fit of the degree distribution is $\alpha = 13.5086$, whereas the whole and volume-returns degree distribution is adjusted with a power oscillating below eight. Figure 3b tells us that the adjustment is not significant most of the time, where we can see *p-values* over the threshold of 0.05 (red line). Nevertheless, the peak at 2020-03-01 for the price-returns is at the border of significant fitness. In a similar way, Figure 3c shows the estimated power behavior for the in-degree distribution, where now there exists an abrupt change at 2020-03-29 for the whole network with an estimated $\alpha = 68.3999$. In this scenario, the adjustment is significant most of the time as can be seen in Figure 3d. Finally, Figure 3e,f shows the corresponding results for the out-degree distribution, being all the estimated power below nine, but without a significant pattern.

In short, these results confirm what was found in the previous subsection, in the sense that the change in the estimated power of the degree distribution of the price-returns anticipates the fall of the average market capitalization of cryptocurrencies occurring at 2020-03-15. Particularly, the in-degree distribution shows the most robust results in the sense of statistically significant evidence of a peak at end of March for the whole network. Then, it is a better indicator of the complex properties of the mTE network. It is important to keep in mind that during March the global markets become extremely turbulent because of the COVID-19 pandemic and the subsequent crash of oil prices around the world. As a matter of fact, during the period of time under study the All Country World Index (ACWI) fell around 10% on a single day, which is its largest decline since 2008. This index published by Morgan Stanley Capital International or MSCI inc. is a market capitalization-weighted index designed to provide a broad measure of equity-market performance throughout the world taking into account stocks from 23 developed countries and 24 emerging markets [59]. As such it is an adequate thermometer of the world financial situation. In Figure 4 it is shown the daily log-returns for the time window from 2019-12-02 to 2020-04-06 (data are taken from [60]). There can be seen the financial turbulence start at the end of February and reach its highest volatility at the middle of March, which coincide with the lowest point of the market capitalization of Figure 2b Hence, based on the evidence found in this study, it is argued that periods of economic contractions are preceded by a high power law distribution of node degrees for price-returns. Nevertheless, it is necessary to develop extensive studies between global financial indices and cryptocurrencies to properly justify this hypothesis.

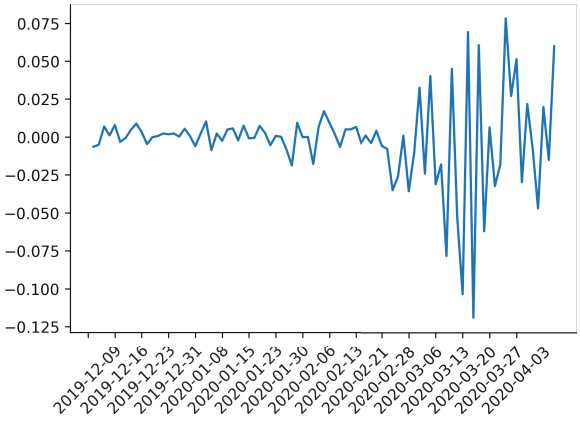


Figure 4. Daily log-returns of ACWI for the trading dates from 2019-12-02 to 2020-04-06.

5. Discussion

The multivariate approach of transfer entropy followed here is an efficient greedy methodology based on multiple hypothesis testing able to detect true causal interactions. Even though the Gaussian specification for the conditional distributions involved in the computations of mTE do not capture non-linear structures in the data, the induced networks of cryptocurrencies exhibit interesting characteristics. Actually, the graph visualization of the interactions of price-returns and volume-returns using the Kamada-Kawai algorithm already displays an informative picture. Under this descriptive representation has been possible to distinguish three time windows where the strength of the attractive forces between nodes stand out to be stronger than the other scenarios. Interestingly, these periods coincide with the most severe fall of the recent worldwide stock markets crash on March 2020.

The clustering coefficient version for directed graphs was also measured. In this case, we first notice that the individual coefficient also increases during the times mentioned above, which motivated us to analyze the temporal behavior of the clustering coefficient for the whole network. Here, our greatest contribution arose finding that the clustering coefficient of the whole network, as well as the price-returns subnetwork, increases dramatically during the same periods of major financial contraction, where we use as an indicator of turbulence the market capitalization of the cryptocurrencies under study.

In addition, we explored deeper the structure of the network through the analysis of the degree distribution of nodes as well as in-nodes and out-nodes. Our intention was to characterize the complexity of the network estimating the power of the associated distributions. Although several of the estimations were not significant, the log-likelihood in all cases bent over a power law distribution, giving evidence of the complex nature of the network. Most importantly, it was found that the power of the distribution has higher estimated values during March 2020, which provides further support to our hypothesis: the structure of the induced cryptocurrency network by mTE changes during times of turbulence in the sense of higher clustering coefficient and complexity.

Future work involves the use of an extensive data set to include the market capitalization, financial indices, sentiment indicators of textual data, as well as volume in a cleaver way in order to verify if the same phenomena are presented in the induced graphs by a more complete data set. The last because what is found here volume does not play a role as relevant as the price is to early-warning signals for future markets turbulence. A forthcoming work is analyzing the same data set from data mining techniques as is the point of view of association rules and the apriori algorithm, where preliminary results already show a rich network structure.

Finally, it is important to clarify that despite the statistical results found here in relation to the power fitness of degree distributions, no economical theory behind power laws has been properly developed yet. At most, we can say the power law degree distributions as well as the clustering coefficient of mTE networks may serve an early warning signal of an increasing systematic risk of the cryptocurrency markets in times of crash. The econophysics community has been put a lot of effort to unravel stylized facts of the network structure of financial markets in general, but it is still necessary to build epistemologically the blocks of the theory jointly with the financial economists to have a common playground to discuss and construct new ideas. As stated in [61]: "*the time is ripe for economists to use those power laws to investigate old and new regularities with renewed models and data*". In this sense, this work only tries to contribute with new evidence in the networks induced by mTE, hoping to soon have an interpretable theory in the same sense as phase transitions, critical points, and scale invariance of turbulent dynamics are in physical statistics.

Supplementary Materials: The following are available online at http://www.mdpi.com/1099-4300/22/7/760/s1, File S1: Cryptocurrency prices, File S2: Cryptocurrency volumes, Table S3: Cryptocurrency tickers.

Author Contributions: Conceptualization, A.G.-M.; methodology, A.G.-M.; formal analysis, A.G.-M.; investigation, A.G.-M.; resources, A.G.-M.; data curation, A.G.-M. and J.B.H.C.; writing—original draft preparation, A.G.-M.; writing—review and editing, A.G.-M. and J.B.H.C.; project administration, A.G.-M.; funding acquisition, A.G.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Consejo Nacional de Ciencia y Tecnología, grant number A1-S-43514.

Acknowledgments: We wish to thank the anonymous reviewers who undoubtedly helped to enrich different aspects of this work.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Non-Uniform Embedding

In order to describe the dynamics of the processes, it is necessary to define composite processes consisting of embedding vectors. Let denote as $x_{i,t}$ the state visited by X_i at time t, where X_i can be characterized by the k past states of the processes $\mathbf{X}_{i,t}^{k} = (x_{i,t-1}, \ldots, x_{i,t-k})$. In the classical uniform embedding framework, the embedding vectors are

$$\mathbf{X}^{(K_j)} = \left(\mathbf{X}_{1,t}^{(k_1)}, \ldots, \mathbf{X}_{j-1,t}^{(k_{j-1})}, \mathbf{X}_{j+1,t}^{(k_{j+1})}, \ldots, \mathbf{X}_{M,t}^{(k_M)}\right) \tag{A1}$$

$$\mathbf{X}^{(K)} = \left(\mathbf{X}_{1,t}^{(k_1)}, \ldots, \mathbf{X}_{M,t}^{(k_M)}\right). \tag{A2}$$

These vectors are used to calculate the conditional entropies involved in the estimation of transfer entropy. The basis of a non-uniform embedding relies on not imposing a priori the form of the embedding vectors. Thus, its shape is determined sequentially by selecting in a progressive manner the terms that contribute most to the explanation of the observed dynamics. These terms are taken from a candidate set, which includes the past states of the processes, $X_1, \ldots, X_M$. To quantify the causality from X_j to X_i it is necessary to contrast the entropy of X_i measured after conditioning on the past of all processes, and after conditioning on the past of all processes except X_j. To accomplish this, it is defined two different sets of initial candidate terms. The first include the past states of $X_j : \Omega = \{x_{i,t-l} | i = 1, \ldots, M; l = 1, \ldots, L\}$; and the second exclude the past states of $X_j : \Omega^{-j} = \{x_{i,t-l} | i = 1, \ldots, M, i \neq j; l = 1 \ldots, L\}$, where L is the maximum lag at which the past of each process is investigated [49].

Appendix B. Statistical Tests

Appendix B.1. Maximum Statistics

Assume that has been chosen the candidate variable C^* from the candidate target past set $\mathbf{Y}_{<t}^{C} = \{y_{t-1}, \ldots, y_{t-L}\}$, which maximizes the CMI contribution. The maximum statistic test reflects this selection process by choosing the maximum value among the surrogates time series. Explicitly, let $I^* := I(C^*, y_t | \mathbf{Y}_{<t}^{S})$ be the maximum contribution. To test I^* for statistical significance the next algorithm is used t [18]:

1. Generate S surrogates time series $C'_{j,1}, \ldots, C'_{j,S}$ for each $C_j \in \mathbf{Y}_{<t}^{C}$, and calculate the corresponding surrogate CMI values $I'_{j,1} = I(C'_{j,1}, y_t | \mathbf{Y}_{<t}^{S}), \ldots, I'_{j,S} = I(C'_{j,S}, y_t | \mathbf{Y}_{<t}^{S})$. The number of surrogates S are chosen in line to the desired significance level α_{max}.
2. For each surrogate $s = 1, \ldots, S$, calculate the maximum CMI value over the candidates $I_s^* := max(I'_{1,s}, \ldots, I'_{n,s})$, where n denotes the number of candidates (the number of comparisons). The values $I_1^*, \ldots, I_S^*$ form the empirical null distribution of the maximum statistic.
3. Compute the p-value for I^* as the proportion of values from the surrogate maximum statistic that are larger than I^*.
4. Finally, I^* is considered significant if the p-value is smaller than α_{max}.

Appendix B.2. Minimum Statistics

The minimum statistic test is used to remove the selected variables that have become redundant in the context of the final set of selected source past variables $\mathbf{X}^S_{<t}$. Also, it controls the FWER due to the multiple comparisons involved in the pruning procedure. If we replace *maximum* with *minimum* in the algorithm described above, then it is obtained the algorithm for the minimum statistic test [18].

Appendix B.3. Omnibus Test

Let $T^* := I(\mathbf{X}^S_{<t}, y_t|\mathbf{Y}^S_{<t})$ be the collective transfer entropy from the selected sources past variables $\mathbf{X}^S_{<t}$ to the target variable y_t. It is necessary to test T^* for statistical significance against the null hypothesis of zero transfer entropy. Thus, using a similar procedure as described in the section of *maximum statistic* test, the null distribution is obtained via shuffling the realizations of the selected sources. The idea behind is to find the set of relevant sources for each node testing all the selected sources collectively [18].

References

1. Nakamoto, S. Bitcoin: A Peer-To-Peer Electronic Cash System. 2008. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 14 May 2020).
2. CoinMarketCap. Available online: https://coinmarketcap.com (accessed on 14 May 2020).
3. Patton, A. Copula methods for forecasting multivariate time series. In *Handbook of Economic Forecasting*; Elliott, G., Timmermann, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2013; Volume 2, pp. 899–960.
4. Hu, M.; Liang, H. A copula approach to assessing Granger causality. *NeuroImage* **2014**, *100*, 125–134. [CrossRef]
5. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
6. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **2000**, *85*, 461–464. [CrossRef]
7. Papana, A.; Kugiumtzis, D.; Larsson, P. Reducing the bias of causality measures. *Phys. Rev. E* **2011**, *83*, 036207. [CrossRef]
8. Barnett, L.; Lizier, J.T.; Harré, M.; Seth, A.K.; Bossomaier, T. Information flow in a kinetic Ising model peaks in the disordered phase. *Phys. Rev. Lett.* **2013**, *111*, 177203. [CrossRef] [PubMed]
9. Liang, X.S. The Liang-Kleeman information flow: Theory and applications. *Entropy* **2013**, *15*, 327–360. [CrossRef]
10. Prokopenko, M.; Lizier, J.T.; Price, D.C. On thermodynamic interpretation of transfer entropy. *Entropy* **2013**, *15*, 524–543. [CrossRef]
11. Barnett, L.; Barrett, A.B.; Seth, A.K. Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys. Rev. Lett.* **2009**, *103*, 238701. [CrossRef]
12. Schindlerova, K. Equivalence of Granger Causality and Transfer Entropy: A Generalization. *Appl. Math. Sci.* **2011**, *5*, 3637–3648.
13. Barnett, L.; Bossomaier, T. Transfer entropy as a log-likelihood ratio. *Phys. Rev. Lett.* **2012**, *109*, 138105. [CrossRef]
14. García-Medina, A.; Farías, G.G. Transfer entropy as a variable selection methodology of cryptocurrencies in the framework of a high dimensional predictive model. *PLoS ONE* **2020**, *15*, e0227269. [CrossRef] [PubMed]
15. Papana, A.; Kyrtsou, C.; Kugiumtzis, D.; Diks, C. Detecting causality in non-stationary time series using partial symbolic transfer entropy: Evidence in financial data. *Comput. Econ.* **2016**, *47*, 341–365. [CrossRef]
16. Runge, J.; Heitzig, J.; Petoukhov, V.; Kurths, J. Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Phys. Rev. Lett.* **2012**, *108*, 258701. [CrossRef] [PubMed]
17. Lizier, J.; Rubinov, M. *Multivariate Construction of Effective Computational Networks from Observational Data*; Technical Report 25; Max Planck Institute for Mathematics in the Sciences: Leipzig, Germany, 2012.
18. Novelli, L.; Wollstadt, P.; Mediano, P.; Wibral, M.; Lizier, J.T. Large-scale directed network inference with multivariate transfer entropy and hierarchical statistical testing. *Netw. Neurosci.* **2019**, *3*, 827–847. [CrossRef] [PubMed]
19. Lizier, J.T.; Prokopenko, M.; Zomaya, A.Y. Local information transfer as a spatiotemporal filter for complex systems. *Phys. Rev. E* **2008**, *77*, 026110. [CrossRef]

20. Lizier, J.T.; Prokopenko, M.; Zomaya, A.Y. Information modification and particle collisions in distributed computation. *Chaos Interdiscip. J. Nonlinear Sci.* **2010**, *20*, 037109. [CrossRef]
21. Goldenfeld, N. *Lectures on Phase Transitions and the Renormalization Group*; Frontiers in Physics; CRC Press: Boca Raton, FL, USA, 1992.
22. Kwapień, J.; Drożdż, S. Physical approach to complex systems. *Phys. Rep.* **2012**, *515*, 115–226. [CrossRef]
23. Barabási, A.L. Scale-free networks: A decade and beyond. *Science* **2009**, *325*, 412–413. [CrossRef]
24. Mantegna, R.N. Hierarchical structure in financial markets. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **1999**, *11*, 193–197. [CrossRef]
25. Bonanno, G.; Caldarelli, G.; Lillo, F.; Micciche, S.; Vandewalle, N.; Mantegna, R.N. Networks of equities in financial markets. *Eur. Phys. J. B* **2004**, *38*, 363–371. [CrossRef]
26. Górski, A.Z.; Drożdż, S.; Kwapień, J. Scale free effects in world currency exchange network. *Eur. Phys. J. B* **2008**, *66*, 91–96. [CrossRef]
27. Kwapień, J.; Gworek, S.; Drożdż, S. Structure and evolution of the foreign exchange networks. *Acta Phys. Pol. B* **2009**, *40*, 175–194.
28. Kwapień, J.; Gworek, S.; Drożdż, S.; Górski, A. Analysis of a network structure of the foreign currency exchange market. *J. Econ. Interact. Coord.* **2009**, *4*, 55–72. [CrossRef]
29. Drożdż, S.; Gębarowski, R.; Minati, L.; Oświęcimka, P.; Wątorek, M. Bitcoin market route to maturity? Evidence from return fluctuations, temporal correlations and multiscaling effects. *Chaos Interdiscip. J. Nonlinear Sci.* **2018**, *28*, 071101. [CrossRef]
30. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Signatures of the crypto-currency market decoupling from the Forex. *Future Internet* **2019**, *11*, 154. [CrossRef]
31. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Competition of noise and collectivity in global cryptocurrency trading: Route to a self-contained market. *Chaos Interdiscip. J. Nonlinear Sci.* **2020**, *30*, 023122. [CrossRef] [PubMed]
32. Baumöhl, E. Are cryptocurrencies connected to forex? A quantile cross-spectral approach. *Financ. Res. Lett.* **2018**, *29*, 363–372. [CrossRef]
33. Zhang, Y.; Chan, S.; Chu, J.; Sulieman, H. On the Market Efficiency and Liquidity of High-Frequency Cryptocurrencies in a Bull and Bear Market. *J. Risk Financ. Manag.* **2020**, *13*, 8. [CrossRef]
34. Canh, N.P.; Wongchoti, U.; Thanh, S.D.; Thong, N.T. Systematic risk in cryptocurrency market: Evidence from DCC-MGARCH model. *Financ. Res. Lett.* **2019**, *29*, 90–100. [CrossRef]
35. Luu Duc Huynh, T. Spillover risks on cryptocurrency markets: A look from VAR-SVAR granger causality and student'st copulas. *J. Risk Financ. Manag.* **2019**, *12*, 52. [CrossRef]
36. Huynh, T.L.D.; Nguyen, S.P.; Duong, D. Contagion risk measured by return among cryptocurrencies. In *International Econometric Conference of Vietnam*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 987–998.
37. Sandoval, L. Structure of a global network of financial companies based on transfer entropy. *Entropy* **2014**, *16*, 4443–4482. [CrossRef]
38. Begušić, S.; Kostanjčar, Z.; Kovač, D.; Stanley, H.E.; Podobnik, B. Information feedback in temporal networks as a predictor of market crashes. *Complexity* **2018**, *2018*, 2834680 . [CrossRef]
39. Huynh, T.L.D.; Shahbaz, M.; Nasir, M.A.; Ullah, S. Financial modelling, risk management of energy instruments and the role of cryptocurrencies. *Ann. Oper. Res.* **2020**, 1–29. [CrossRef]
40. Zenil, H.; Kiani, N.A.; Tegnér, J. A review of graph and network complexity from an algorithmic information perspective. *Entropy* **2018**, *20*, 551. [CrossRef]
41. Dickey, D.A.; Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **1979**, *74*, 427–431.
42. Phillips, P.C.; Perron, P. Testing for a unit root in time series regression. *Biometrika* **1988**, *75*, 335–346. [CrossRef]
43. Mantegna, R.N.; Stanley, H.E. Scaling behaviour in the dynamics of an economic index. *Nature* **1995**, *376*, 46–49. [CrossRef]
44. Mantegna, R.N.; Stanley, H.E. *Introduction to Econophysics: Correlations and Complexity in Finance*; Cambridge University Press: Cambridge, UK, 1999.
45. Begušić, S.; Kostanjčar, Z.; Stanley, H.E.; Podobnik, B. Scaling properties of extreme price fluctuations in Bitcoin markets. *Phys. A Stat. Mech. Its Appl.* **2018**, *510*, 400–406. [CrossRef]

46. CoinMarketCap Methodology. Available online: https://coinmarketcap.com/methodology/ (accessed on 1 July 2020).
47. Alexander, C.; Dakos, M. A critical investigation of cryptocurrency data and analysis. *Quant. Financ.* **2020**, *20*, 173–188. [CrossRef]
48. Wollstadt, P.; Lizier, J.T.; Vicente, R.; Finn, C.; Martínez-Zarzuela, M.; Mediano, P.; Novelli, L.; Wibral, M. IDTxl: The Information Dynamics Toolkit xl: A Python package for the efficient analysis of multivariate information dynamics in networks. *arXiv* **2018**, arXiv:1807.10459.
49. Faes, L.; Nollo, G.; Porta, A. Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique. *Phys. Rev. E* **2011**, *83*, 051112. [CrossRef]
50. Wibral, M.; Pampu, N.; Priesemann, V.; Siebenhühner, F.; Seiwert, H.; Lindner, M.; Lizier, J.T.; Vicente, R. Measuring information-transfer delays. *PLoS ONE* **2013**, *8*, e55809. [CrossRef]
51. Granger, C.W. Investigating causal relations by econometric models and cross-spectral methods. *Econom. J. Econom. Soc.* **1969**, *37*, 424–438. [CrossRef]
52. Kamada, T.; Kawai, S. An algorithm for drawing general undirected graphs. *Inf. Process. Lett.* **1989**, *31*, 7–15. [CrossRef]
53. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **1998**, *393*, 440–442. [CrossRef] [PubMed]
54. Fagiolo, G. Clustering in complex directed networks. *Phys. Rev. E* **2007**, *76*, 026107. [CrossRef]
55. Bouchaud, J.P. Power laws in economics and finance: Some ideas from physics. *Quant. Financ.* **2001**, *1*, 105–112. [CrossRef]
56. Uriel, F. *Turbulence: The Legacy of A. N. Kolmogorov*; Cambridge University Press: Cambridge, UK, 1995.
57. Mandelbrot, B.; Cootner, P.; Gomory, R.; Fama, E.; Morris, W.; Taylor, H. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk. Selecta Volume E*; Springer: Berlin/Heidelberg, Germany, 1997.
58. Alstott, J.; Bullmore, D.P. powerlaw: A Python package for analysis of heavy-tailed distributions. *PLoS ONE* **2014**, *9*, e85777. [CrossRef]
59. MSCI. Available online: https://www.msci.com/ (accessed on 30 June 2020).
60. Yahoo Finance. Available online: https://finance.yahoo.com/quote/ACWI/ (accessed on 30 June 2020).
61. Gabaix, X. Power laws in economics and finance. *Annu. Rev. Econ.* **2009**, *1*, 255–294. [CrossRef]

entropy

MDPI

Article

Complexity in Economic and Social Systems: Cryptocurrency Market at around COVID-19

Stanisław Drożdż [1,2,*,†], Jarosław Kwapień [1,†], Paweł Oświęcimka [1,3,†], Tomasz Stanisz [1,†] and Marcin Wątorek [2,†]

[1] Complex Systems Theory Department, Institute of Nuclear Physics, Polish Academy of Sciences, ul. Radzikowskiego 152, 31-342 Kraków, Poland; jaroslaw.kwapien@ifj.edu.pl (J.K.); pawel.oswiecimka@ifj.edu.pl (P.O.); tomasz.stanisz@ifj.edu.pl (T.S.)

[2] Faculty of Computer Science and Telecommunication, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland; marcin.watorek@pk.edu.pl

[3] Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, ul. prof. Stanisława Łojasiewicza 11, 30-348 Kraków, Poland

* Correspondence: Stanislaw.drozdz@ifj.edu.pl

† These authors contributed equally to this work.

Received: 25 August 2020; Accepted: 15 September 2020; Published: 18 September 2020

Abstract: Social systems are characterized by an enormous network of connections and factors that can influence the structure and dynamics of these systems. Among them the whole economical sphere of human activity seems to be the most interrelated and complex. All financial markets, including the youngest one, the cryptocurrency market, belong to this sphere. The complexity of the cryptocurrency market can be studied from different perspectives. First, the dynamics of the cryptocurrency exchange rates to other cryptocurrencies and fiat currencies can be studied and quantified by means of multifractal formalism. Second, coupling and decoupling of the cryptocurrencies and the conventional assets can be investigated with the advanced cross-correlation analyses based on fractal analysis. Third, an internal structure of the cryptocurrency market can also be a subject of analysis that exploits, for example, a network representation of the market. In this work, we approach the subject from all three perspectives based on data from a recent time interval between January 2019 and June 2020. This period includes the peculiar time of the Covid-19 pandemic; therefore, we pay particular attention to this event and investigate how strong its impact on the structure and dynamics of the market was. Besides, the studied data covers a few other significant events like double bull and bear phases in 2019. We show that, throughout the considered interval, the exchange rate returns were multifractal with intermittent signatures of bifractality that can be associated with the most volatile periods of the market dynamics like a bull market onset in April 2019 and the Covid-19 outburst in March 2020. The topology of a minimal spanning tree representation of the market also used to alter during these events from a distributed type without any dominant node to a highly centralized type with a dominating hub of USDT. However, the MST topology during the pandemic differs in some details from other volatile periods.

Keywords: complex systems; cryptocurrencies; multifractal analysis; detrended cross-correlations; minimal spanning tree

1. Introduction

Whether complexity of a system is viewed in the purely intuitive sense of a nontrivial order that emerges spontaneously from an overall disorder or it is grasped more formally using one of several dozen mathematical, physical, and information-theoretic measures, we are surrounded by its signatures and face its manifestations almost everywhere. We are complex ourselves: We live in a society that

is complex and we interact with others in a complex way. There is no exaggeration in a statement that our society is the most complex structure known to us in the universe. Social phenomena like the emergence of communication and cooperation, build-up of hierarchies and organizations, opinion formation, the emergence of political systems, and the structure and dynamics of financial markets are all among the iconic examples of the real-world complexity [1–3].

Specialists from such disciplines like mathematics, physics, information theory, and data science working together with econometrists, sociologists, quantitative linguists, and psychologists for more than a quarter century have already been dealing with such phenomena trying to describe them in a language of exact science, and to model and explain them using methods and tools that had earlier been applied successfully to natural systems. while much has already been done and much has been achieved, the complexity of the social and economic systems is still far from being properly understood. This is why every possible effort and every meaningful contribution is welcome as it can bring us closer to the ultimate goal of understanding complexity both in reference to these systems in particular and as a physical phenomenon in general. It is also important to approach the problem from different angles by collecting many interdisciplinary works and views in one place like this Special Issue as human society eludes any narrow-scope, single-discipline analysis.

1.1. Money, Fiat Currencies, and Cryptocurrencies

Among a variety of emergent phenomena that we observe in human society, one of the most important is money. It appeared spontaneously and independently in many cultures and, although it used to have different material forms in different regions, it always served the same purpose: To facilitate trade by avoiding a problem of double coincidence of needs that restricts barter trading severely and inherently. According to economical models, a status of money is acquired in a process of the spontaneous symmetry breaking by a commodity that is the most easily marketable or, in other words, that is the most liquid one [4,5]. After receiving such a status by some commodity, its liquidity is amplified by a kind of self-propelling mechanism, because everybody desires to have an asset that is considered as the most desirable by others. However, there is another condition for a commodity to be used as money: Its value expressed in other assets has to be viewed as stable. Sometimes it happens that current money loses its value which causes people to withdraw themselves from using it and to replace it with some other, more stable asset. Thus, for a given asset its status of money may either be durable or temporary. This is an important issue in contemporary economy based on fiat money that does not have any intrinsic value unlike the assets that used to play a role of money earlier in history. Value of the fiat currencies depends crucially on policies of the central banks, which can be subject to change. Moreover, the central banks may increase money supply at any time, which can lead to inflation rate increase. This undermines confidence in the official currencies and became the ignition to introduce cryptocurrencies over a decade ago.

The first cryptocurrency was proposed in 2008—Bitcoin (BTC) [6]. The idea behind it was to decouple a currency from any institution or government, while preserving its status of a universal means of exchange, and to base a trust in this currency solely on a technology that supports it. Such a currency had to combine the advantages of both cash and electronic money: Anonymity of use (like cash) and capability of being transferred immediately to any place in the world (like electronic money). The already-existing technologies of asymmetric cryptography and distributed database (with a new consensus mechanism—"proof of work") were linked into a decentralized secure register—blockchain [7] that forms a staple of BTC. Unlike traditional currencies, Bitcoin has inherently limited supply to prevent any loss of its value due to inflation.

The first widely recognized exchange enabling bitcoin to be exchanged for traditional currencies, Mt. Gox, was launched in July 2010 followed by the first online (black) market—Silk Road. The latter was a place where one could anonymously buy anything and pay with bitcoins, which was the first practical application of a cryptocurrency. It significantly increased the demand and contributed to the first speculative bubble on BTC [8]. A subsequent crash occurred after closing Silk Road and

suspending trade on Mt. Gox between October 2013 and February 2014. As Bitcoin's recognition increased, the use of blockchain technology became more popular and it turned out that it can also be used for trustful processing of computer codes in a decentralized way. In 2015 the Ethereum distributed computing network was launched [9], which allows one to issue private tokens through a so-called Initial Coin Offer (ICO) and to raise capital in a simplified way for various projects. An ICO boom that contributed to next speculative bubble on cryptocurrencies that occurred in 2017 (the ICO-mania [10]). At that time the number of issued cryptocurrencies doubled from 700 to 1400 and the market capitalization reached 800 billion USD. A crash in January 2018, in which BTC lost over 80% of its value and other cryptocurrencies lost even 99%, may be compared with the dot-com bubble crash in 2000 that ended the most euphoric phase of investor attitude towards the Internet-related companies. At present the market is more consolidated and shows signatures of maturity [11].

1.2. Basic Information on the Blockchain Technology

In order to create an electronic "currency" that can easily be exchanged for goods and operated without any central authority, while at the same time that cannot be multiplied indefinitely like electronic files, it is required that all transactions involving that "currency" have to be registered publicly, which ensures that no registry can be modified afterwards. The Bitcoin network register consists of a sequence of block files built one upon another (a blockchain) containing information about past transactions and the instances of new Bitcoin unit creation. A new network participant has to enter the network directly via a network client or via an external wallet and must send information about the client's address and a specified Bitcoin sum it owns. This information is then distributed to all other network nodes but, in return, the new participant is granted access to the complete information about other network node addresses and how many BTC units belong to these addresses. Thus, credibility of the system is provided by the technology itself by imposing certain set of rules each network participant must obey and by allowing the network participants to control each other. However, since the Bitcoin blockchain is public, one can trace the transaction history of each unit, which in theory might compromise transaction anonymity.

The transaction correctness is guaranteed with the help of the asymmetric cryptography. Private keys of a sender and a receiver are used to encode and to decode a transaction (i.e., to send and to receive coins), while their public keys are used as their public addresses allowing for their network identity verification. while such a transaction is visible to any other network participant, nobody can effectively alter and re-encode it as they do not know the private keys of the involved parties. For the network, in order to function correctly, the key implemented feature is a consensus mechanism that ensures that all participants agree upon ownership of the cryptocurrency units and how many units total circulate. while collecting information from many transactions taking place on the Bitcoin network, the consensus mechanism has to overcome a problem that some information sources can be unreliable. It is done by the so-called proof-of-work (PoW) protocol used by miners, i.e., the network nodes with dedicated software that collect transactions, verify their correctness, and integrate them into blocks. This is a resource-consuming task so the miners are got to perform it by receiving new coins in exchange for sharing their resources with the network. The new Bitcoin unit is generated only after majority of the miners agree upon correctness of the new block and it has been distributed over the network. The block has to meet relevant criteria expressed by a specific form of the hash function to be considered as a valid one and included in the blockchain. Each miner decides to include a given block into its own blockchain copy individually and the consensus is settled in a kind of game with a Nash equilibrium state. One has to believe that majority of other miners agrees on the specific block's validity and adds it to its own blockchain or will not receive the profit otherwise.

Mining a new Bitcoin unit requires much energy to be spent so the very process demands optimization of the resources used and discourage padding the blocks with fictitious information as a rejection probability for such a block by other miners is too large. Therefore it serves as a proof of work that a participant made the effort of maintaining the network, indeed. The employed solution

that each new block contains a header of the previous one practically eliminates a problem of potential modifying the past transactions—it is not viable economically since it would require rebuilding of the entire chain. The Bitcoin protocol was designed in such a way that new blocks are formed with constant frequency, which is achieved by adjusting the amount of the corresponding calculations needed to the network's actual computing power. Moreover, the reward for forming a new block is halved every 210,000 blocks in order to approach quasi-asymptotically an impassable limit of 21 million Bitcoin units.

The Bitcoin protocol is not static and undergoes constant modifications. A reason for this is that the protocol in its original design has some drawbacks that can challenge its security and lower comfort of its use. Among the pivotal issues is low performance (the network can handle only 5 transactions per second on average, compared to 1700 transactions per second in the Visa network), high operating costs that equal the amount of electric energy consumed by small industrialized countries (like Ireland or Denmark [12]), and formidable computer facility. Moreover, one of the blockchain technology advantages—the inability of making changes—may sometimes be viewed as its disadvantage if one considers the protocol correcting since it requires cloning of the entire network and abandoning the original blockchain. Up to now a mechanism of reducing transaction size and allowing to pack more transactions in a single block (called "segregated witness", SegWit) has already been implemented and work on another mechanism—"Lightning Network"—that allows for micropayments outside the main blockchain and increasing the bandwidth, is currently underway.

However, such changes are viewed by inefficient by many who prefer building alternative networks from scratch or by using only certain features of the Bitcoin protocol, while replacing other with better solutions. Thus, over the last decade, a multitude of different protocols were proposed and implemented, which led to introduction of new cryptocurrencies. Most of them still exploit the PoW protocol, but its the most popular alternative is the proof-of-stake (PoS) [13]. In this algorithm miners do not exist and the block validation process is granted to some randomly chosen network nodes. Consistently, the block formation is not rewarded with new units but rather the validator nodes are rewarded with transaction fees. Fraud is discouraged by excluding the fraudulent participants from the network and securing that, in such a case, the reward for forming a new block is smaller than possible loss in already owned units. The main advantage of PoS is efficiency: Because of a lack of the complicated and long calculations, no specialized user group is needed to confirm blocks and everything can be done faster than in the case of PoW. There are various versions of the PoS protocol, like the "delegated proof of stake" (DPoS) based on voting system engaging trustful delegated network nodes or the "proof-of-authority" (PoA) based on granting reputation to the validator nodes instead of cryptocurrency units and abandoning the decentralization paradigm. Main advantage of both protocols (together with their hybrid versions) is scalability—more participants mean larger transaction capacity of the related network.

1.3. Other Applications of the Blockchain Technology

The first Bitcoin alternative that was introduced in 2011 and managed to survive until today was Litecoin (LTC). Basically, this is a Bitcoin's clone that differs from its parent in that it has a higher average creation frequency (4 min) and a higher prospected total number of units (84 million) as well as it uses different hash function (script instead of dSHA-256 used by Bitcoin). These changes allowed LTC for much smaller resource demand than BTC and made LTC be computable on standard CPUs. The first cryptocurrency that was not based on the Bitcoin's PoW protocol was Ripple (XRP) [14] introduced in 2012. It was intended to be used as a method of transferring money between banks and stock markets in real time even outside national borders. In August 2020 XRP was the third cryptocurrency in terms of capitalization. A related cryptocurrency, Stellar (XLM), also offers transactions between financial institutions, but unlike Ripple based on a proprietary code its code is open source. Both XRP and XLM do not have a fixed supply limit and, thus, they are subject to inflation.

A separate group of cryptocurrency protocols was designed to ensure user anonymity. The corresponding cryptocurrencies are called "private coins": Dash (DASH), Monero (XMR), Zcash (ZEC), and many others. Dash uses a two-layer network with PoW and miners in the first layer and PoS and "masternodes" in the second one. Monero, being considered as the most secure private coin and often used by the criminal world [15], provides anonymity thanks to a Ring Confidential Transactions (RingCT) where the public keys (addresses) are hidden in the blockchain [16]. Zcash is based on a solution that allows one to confirm information without having to disclose it. Zcash allows for perfect anonymity of both the sender and the recipient as well as transaction size. Since the anonymous addresses are compatible with the public ones, transactions can be made between public and hidden wallets and vice versa. DASH and ZEC have a maximum supply set in advance, while XMR does not.

Apart from the cryptocurrencies, another important category of blockchain applications is cryptocommodities (together with the former called cryptoassets). They are automatically executed computer codes that perform certain actions if certain conditions are met. Cryptocommodities enable payments for using a decentralized computing network. The first such cryptoasset was Ethereum—an open-source computing platform designed for programming decentralized applications and smart contracts that was launched in 2015 [9]. This platform has its own programming language and its own cryptocurrency, Ethereum (ETH), that serves as a payment unit for carrying out computational operations on the platform. Ethereum is based on PoW consensus mechanism, but it uses another hash function (Ethash) supporting use of GPUs in the mining process and there is no upper limit on mining. Instead of fixed block size, here each block requires a specific number of "Gas" units related to the computing power needed to complete the transactions it contains. The average block-completion frequency is 15 s and the maximum transaction number per second is around 25. The Ethereum concept gained quickly high popularity among the cryptocurrency community and, currently, ETH is the second crytocurrency in terms of capitalization. The success of smart contracts (i.e., computer codes allowing for automatic execution and control of transaction agreement actions) and possibility of collecting funds under Initial Coin Offers on the Ethereum platform, gave a boost to the emergence of similar platforms offering possibility of creating applications in a decentralized environment. Major projects of this type include EOS and Cardano; both have their own cryptocurrencies and both allow for collecting funds under ICOs.

Yet another group of cryptoassets are tokens, which are means of payment in decentralized applications built on platforms like Ethereum or contracts that are issued within ICOs for development of blockchain ventures. They usually don't have their own blockchain. In general, the blockchain technology, thanks to elimination of the need to trust individual participants of a given system and ensuring security, can satisfactorily be used wherever there is a central intermediary connecting sellers and buyers who earns on commissions (for example, Uber and Airbnb). Some of the already introduced applications in the token form are Augur (a platform enabling creation and participation in plants from any thematic range), Filecoin (a decentralized file storage system based on the PoW system that rewards users for sharing their computer storage devices), and IOTA (a project of a partially decentralized, open settlement platform for the needs of the so-called "Internet of things"), Basic Attention Token (a project designed to connect advertisers and content creators with users that rewards the creators for attracting users with the content they provide). Finally, the so-called "stable coins"—a combination of the token and cryptocurrency assets—allow one to relate their value to some other, more conventional asset like US dollar (e.g., USDT, USDC, TUSD, or PAX).

1.4. Cryptocurrency Market

Cryptocurrency trading is possible, because they are easily convertible to traditional currencies like USD or EUR and to other cryptocurrencies. This possibility is provided by 330 trading platforms (August 2020) open 24 h a day, seven days a week. This, together with a fact that the most investors are individuals, distinguishes the cryptocurrency market from Forex, where trading takes place from Monday to Friday essentially on the OTC market where mainly banks and other financial

institutions participate in. Another peculiarity of the cryptocurrency market is that there is no reference exchange rate unlike Forex, where such reference rates are provided by Reuters. The sole exception is Bitcoin, whose exchange rate to USD is given by futures quoted on Chicago Mercantile Exchange [17]. Decentralization of the market means that the same cryptocurrency pairs are traded on different platforms, which—if accompanied by limited liquidity—can lead to sizeable valuation differences between platforms that produce arbitrage opportunities, both the dual and triangluar ones [11,18,19].

The entire cryptocurrency market capitalization is around 350 billion USD, which is close to the capitalization of a middle-size stock exchange and also comparable with the capitalization of the largest American companies. There are 6500 different cryptocurrencies on the market right now, which gives a total of nearly 26,500 cryptocurrency pairs [20]. Founded in 2017, Binance [21] is currently one of the largest cryptocurrency exchange in terms of volume. Binance offers trading on approximately 650 cryptocurrency pairs including pairs with its own cryptocurrency called binance (BNB), used to pay commissions on this exchange.

The spectacular development of a cryptocurrency market has attracted much interest of the scientific community. The first Bitcoin-related papers were published already in 2013–2015 [22,23], but a real boom on cryptocurrency-related publications occurred after 2017. Initially, only bitcoin was of significant interest [24–26], but soon also other cryptocurrencies went under investigation [27–29]. Then there appeared studies reporting on correlations within the market [30–38], and its relationship with regular markets [39–43]. Recently, some researchers focused their attention on possible use of BTC as a hedging instrument for Forex [44], for gold and other commodities [45], as well as for the stock markets [46,47]. There is also a few review papers devoted to the cryptocurrency markets:[11,48,49].

The cryptocurrency market has already gone through a long route from a mere curiosity and a playground for the technology enthusiasts, via an emerging-market stage characterized by a relatively small capitalization, poor liquidity, large price fluctuations, short-term memory, frequent arbitrage opportunities, and weak complexity, to a more mature form characterized by medium capitalization, improved liquidity, inverse-cubic power-law fluctuations [50,51], long-term memory, sparse arbitrage opportunities, and increasing complexity. This is the most interesting aspect of the cryptocurrency market route to maturity: The signatures of complexity that are best quantified in terms of the multifractal analysis. See Ref. [11] for a comprehensive study of this transition started in 2012 and ended essentially in 2018, as viewed from the multifractality perspective. Here we shall consider a more recent period of 2019–2020, which comprises, among others, two significant events, i.e., the bull market between April and July 2019 and the Covid-19 pandemics (from March 2020). Based on high-frequency data covering a large number of cryptocurrency pairs and a few principal traditional-market assets, we investigate a potential impact of these events on the cryptocurrency market structure and its relation to the traditional markets.

2. Methods and Results

2.1. Data

For this study we collected high-frequency recordings of X/BTC and BTC/USDT exchange rates, where X is one of 128 cryptocurrencies traded on Binance platform [21] and USDT is related to USD by a 1:1 peg [52]. The exchange rates $P(t)$ were sampled every 1 min. We calculated their normalized logarithmic returns $r_{\Delta t}$ defined by

$$r_{\Delta t} = (R_{\Delta t} - \mu_R)/\sigma_R, \quad R_{\Delta t}(t) = \log(P(t+\Delta t)) - \log(P(t)), \tag{1}$$

where μ_R and σ_R are mean and standard deviation of $R_{\Delta t}(t)$, respectively, and Δt is sampling interval. We also collected 1-min quotes of several conventional assets expressed in US dollar—13 currencies: AUD, EUR, GBP, NZD, CAD, CHF, CNH, JPY, MXN, NOK, PLN, TRY, ZAR, three stock market indices: Dow Jones Industrial Average (DJI), Nasdaq100, S&P500, and four commodities: XAU (gold), CL (crude oil), XAG (silver), and HG (copper). They all come from Dukascopy platform [53],

so do the BTC/USD and ETH/USD exchange rates. These quotes were also transformed into time series of returns.

2.2. Multifractal Formalism

Multifractal analysis is one of the most promising methods of studying empirical data representing natural and social systems as it is able to quantify complexity of such systems and express it in a relatively simple way with a small set of associated quantities. It has already been applied in many works to univariate and multivariate data sets from a number of different systems: Physics [54], biology [55], chemistry [56], geophysics [57], hydrology [58], atmospheric physics [59], quantitative linguistics [60], behavioral sciences [61], cognitive structures [62], music [63], songbird rhythms [64], physiology [65], human behaviour [66], social psychology [67] and even ecological sciences [68], but especially financial markets [69–77].

Let us consider two time series of the same length: x_i, y_i, where $i = 1, ..., T$ (T has to be large enough to overcome statistical uncertainties). Signal profiles are created from these time series by integrating and subtracting their mean:

$$X(j) = \sum_{i=1}^{j} [x_i - \langle x \rangle], \quad Y(j) = \sum_{i=1}^{j} [y_i - \langle y \rangle]. \tag{2}$$

These signal profiles are then divided into segments ν of length s. They may be separate or partially overlapping; if they are separate, their number is $M_s = \lfloor T/s \rfloor$. A local trend is then removed from each segment by fitting the data with polynomials $P^{(m)}_{X,\nu}$, $P^{(m)}_{Y,\nu}$ of degree m (typically, it is $m = 2$ [78–80]). Now covariance F^2_{xy} is determined from the residual signals for each segment [81,82]:

$$F^2_{xy}(\nu, s) = \frac{1}{s} \sum_{k=1}^{s} \{ \left[X((\nu-1)s + k) - P^{(m)}_{X,\nu}(k) \right] \left[Y((\nu-1)s + k) - P^{(m)}_{Y,\nu}(k) \right] \} \tag{3}$$

and then it is used to calculate the q-th order fluctuation function [83]:

$$F^q_{xy}(s) = \frac{1}{M_s} \sum_{\nu=1}^{M_s} \mathrm{sign}(F^2_{xy}(\nu, s)) |F^2_{xy}(\nu, s)|^{q/2}, \tag{4}$$

where $\mathrm{sign}(F^2_{xy}(\nu, s))$ means a sign function. If $F^2_{xy}(\nu, s)$ is considered as a value of a random variable, the parameter q resembles an exponent specifying the order of the moment: Its large positive values favour segments characterized by large variance by increasing their relative magnitude with respect to small-variance segments, while negative values of q do the opposite. Thus, by applying different values of q, one can construct effective filters that select the segments of a certain variance range.

The fluctuation function (4) has to be calculated for different segment lengths s. If $F^q_{xy}(s)$ is of a power-law form, i.e.,

$$F^q_{xy}(s)^{1/q} = F_{xy}(q, s) \sim s^{\lambda(q)}, \tag{5}$$

where $q \neq 0$, the original time series x_i and y_i are fractally cross-correlated. If $\lambda(q) = \mathrm{const}$, this is monofractal cross-correlation, otherwise it is multifractal one.

A special case is $x_i \equiv y_i$ for all i (one signal). In this case we have

$$F(q, s) = \left[\frac{1}{M_s} \sum_{\nu=1}^{M_s} [F^2(\nu, s)]^{\frac{q}{2}} \right]^{\frac{1}{q}} \tag{6}$$

and the fractal case corresponds to

$$F(q, s) \sim s^{h(q)}, \tag{7}$$

where $h(q)$ is the generalized Hurst exponent. For $h(q) = \text{const}$ the signal is monofractal, otherwise it is multifractal [80]. A useful measure of fractal properties is singularity spectrum $f(\alpha)$ defined by

$$\alpha = h(q) + qh'(q), \quad f(\alpha) = q[\alpha - h(q)] + 1, \tag{8}$$

where α is the Hölder exponent. $f(\alpha)$ can be interpreted as a fractal dimension of the singularities characterized by a given α. In the monofractal case it consists of a single point, while in the multifractal case it can have a shape of inverted parabola or some asymmetric concave function.

Width of $f(\alpha)$ can be interpreted as a measure of a signal's complexity, because the wider it is, the more singularity types can be identified in this signal. This width depends on a range of q and it is quantified by

$$\Delta\alpha = \alpha_{\max} - \alpha_{\min}, \tag{9}$$

where $\alpha_{\min} = \alpha(q_{\max})$ and $\alpha_{\max} = \alpha(q_{\min})$ are the minimum and maximum value of α that have been calculated for different values of q. Another important feature of the $f(\alpha)$ its left-right asymmetry [84]. A left-hand-side asymmetry corresponds to more diverse multifractality (stronger correlations) at the large amplitude level, while a right-hand-side asymmetry indicates that signal parts with small amplitude are a dominant source of multifractality.

As $F^q_{xy}(s)$ denotes the qth-order detrended covariance, one can define the qth-order detrended correlation coefficient [85,86]:

$$\rho(q,s) = \frac{F^q_{xy}(s)}{\sqrt{F^q_{xx}(s)F^q_{yy}(s)}}, \tag{10}$$

in analogy to the qth-order Pearson correlation coefficient. Here F_{xx} and F_{yy} are calculated from Equation (6). The coefficient $\rho(q,s)$ can assume values in a range $[-1,1]$ provided $q > 0$. For $q \leq 0$ a situation becomes more complicated as $\rho(q,s)$ may fall outside that range, which requires more delicate interpretation [85]. Therefore, many studies in which $\rho(q,s)$ is used are carried out with a restriction $q > 0$. The coefficient $\rho(q,s)$ describes detrended cross-correlations between two signals on different scales s after amplifying data points within a given amplitude range. This filtering ability of $\rho(q,s)$ is its advantage over more standard correlation measures, because the cross-correlation strength among empirical time series can be size-dependent [87]. The coefficient $\rho(q,s)$ may be used for any two signals without a requirement that they have to be fractal.

2.3. Multifractal Properties of the Cryptocurrency Market

We start our analysis by taking a look at the BTC/USDT exchange rate from 01/2019 to 06/2020. This period shown in Figure 1 (top panel) starts near the lowest point of the bear market ($\sim$3200 USDT) that begun in 12/2017 and deprived BTC over 80% of its maximum value. During the subsequent one a half year BTC/USDT rate experienced a growth to a local maximum of 12,800 USDT in 07/2019 (+300%), a local minimum in 03/2020 at 4400 USDT (over 60% loss) related to a Covid-19 pandemic onset, and a recent growth to a present price of 12,200 USDT (+170%). The BTC price expressed in USDT was highly unstable over the considered period. This observation is supported by Figure 1 (bottom panel) that shows BTC/USDT 1 min returns.

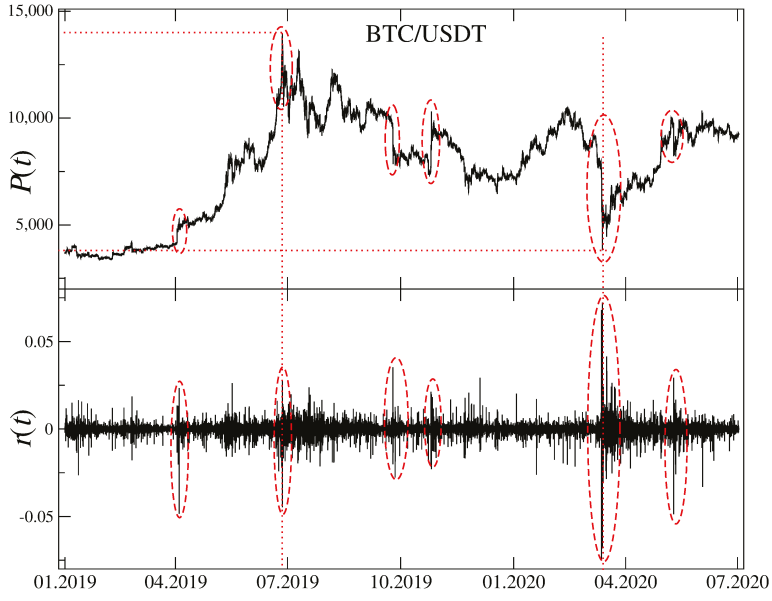


Figure 1. Time evolution of the BTC/USDT exchange rate (top) together with the corresponding logarithmic returns (bottom). Several interesting events can be distinguished like start of a bull market in April 2019 and its end in July 2019, a sudden decrease and then an equally sudden increase in October and November 2019, the Covid-19 pandemic outbreak and related panic in March 2020 and the pandemic's 2 wave in June 2020. Local extrema of $P(t)$ are indicated by the vertical (time) and horizontal (price) dotted lines.

Despite of the fact that BTC/USDT rate is the most important observable on the cryptocurrency market since BTC has the largest capitalization, it cannot be used as a proxy allowing one to describe dynamics of the whole market, which is in fact much richer. Thus, in order to express the evolution of a significant part of the market in terms of a single quantity, a market index was created from the exchange rates X/USDT (with X standing for a cryptocurrency) for 8 the most capitalized cryptocurrencies: BTC, ETH, XRP, BCH, LTC, ADA, BNB, and EOS. In 2020, these assets stand for 88% of the market capitalization. In order to create the index, the exchange rates were summed with the same weight despite the difference in capitalization. A parallel, weighted index would predominantly reflect the dynamics of BTC, ETH, and XRP, so we prefer the unweighted version as more a diversified one.

Figure 2 shows results of the multifractal analysis of the cryptocurrency index returns and the BTC/USDT returns performed by using a moving window of 30 days with a 5-day step. Instead of presenting the singularity spectra $f(\alpha)$ for each window position, temporal evolution of the key quantities describing shape of these spectra is shown: $\alpha_{\min}(t)$, $\alpha_0(t)$, and $\alpha_{\max}(t)$ (see right panel of Figure 3 for the examples). These quantities allow for inferring about the singularity spectrum localization, width, and possible asymmetry of its shoulders [88]. We restricted the applied values of q to $[-3,3]$ for a reason that will be explained later. By looking at the spectra for BTC/USDT (the second topmost panel in Figure 2), one sees that a difference $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$ describing the spectrum width is sufficient to infer about multifractality of the data under study. This agrees with results of our previous study [11].

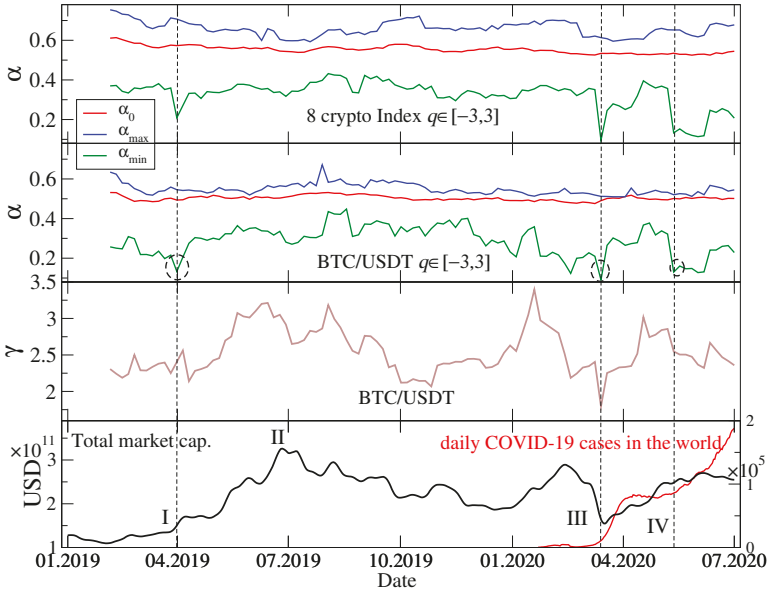


Figure 2. (Top) Characteristic values of the Hölder exponent: $\alpha_{\min}$ (green line, bottom), α_0 (red line, middle), and $\alpha_{\max}$ (blue line, top)—see Equation (9) in Section 2.2 and Figure 3—describing the singularity spectra $f(\alpha)$ for the index returns representing 8 the most capitalized cryptocurrencies, calculated in a 30-day-long moving window with a step of five days and for $-3 \leq q \leq 3$. Each date represent a window that ends on that day. (Upper middle) The same quantities as in the top panel, but here calculated for the BTC/USDT exchange rate returns. Three interesting cases of small $\alpha_{\min}$ are indicated by dashed circles. (Lower middle) Scaling exponent γ of the cumulative distribution function fitted to tails of the empirical cdf in each moving window position. Values equal or below $\gamma = 2$ correspond to Lévy-stable distributions. (Bottom) Total cryptocurrency market capitalization and new Covid-19 cases in the world as function of time. Characteristic events are indicated by vertical dashed lines and Roman numerals: Start of a bull market in April 2019 (event I), its end in July 2019 (event II), the Covid-19 panic in March 2020 (event III), and start of the 2nd wave of the pandemic in May-June 2020 (event IV).

Except for July-August 2019, when $f(\alpha)$ is left-right symmetric ($\alpha_{\max} - \alpha_0 \approx \alpha_0 - \alpha_{\min}$), throughout the remaining part of the analyzed period there is significant asymmetry with the left-hand shoulder ($q > 0$) being much longer than the right-hand one ($q < 0$). In a few instances, i.e., in April 2019, January 2020, March 2020, and May-June 2020, this asymmetry of $f(\alpha)$ became extreme and revealed a bifractal-like shape (see also [89]). Mathematical bifractals are characterized by the existence of only 2 singularity types with $\alpha_1 = 0$ and $0 < \alpha_2 < 1$. However, in practical situations, the finite-size effects smear the spectra so that in such a case there is a continuous transition between both singularity types and a spectrum consists of a long left shoulder reaching a vicinity of $\alpha_1 = 0$ and a residual right shoulder near α_2 [84,90]. Two characteristic cases of $f(\alpha)$ (symmetry and bifractal-like asymmetry) are shown in Figure 3 (right panel).

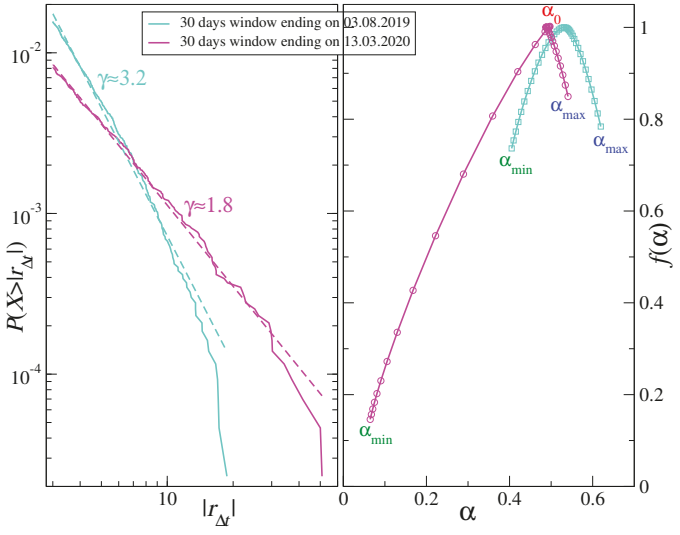


Figure 3. (Left) Cumulative distribution function $P(X > |r_{\Delta t}|)$ calculated in 30-day windows. Two extreme cases of power-law tail are shown with the scaling exponent $\gamma \approx 1.8$ (mid February–mid March 2020) and $\gamma \approx 3.2$ (July 2019) representing stable and unstable distributions, respectively. (Right) Singularity spectra $f(\alpha)$ calculated in the same windows as above. An example of asymmetric, bifractal-like spectrum (mid February - mid March 2020) and an example of symmetric spectrum (July 2019) are shown together with characteristic values of the Hölder exponent: $\alpha_{\min}$, α_0, and $\alpha_{\max}$ (see Equation (9) in Section 2.2).

On the probability distribution function level, the actual bifractal spectra occur if a signal under study has a heavy-tailed pdf in the Lévy-stable regime ($p(|r_{\Delta t}|) \sim 1/|r_{\Delta t}|^{\gamma+1}$, where $\gamma \leq 2$), but empirically one can sometimes obtain a strong left-hand-side asymmetry even for a signal with an unstable pdf provided it is substantially leptokurtic [90]. Figure 3 illustrates this connection between a cumulative distribution function (cdf) $P(X > |r_{\Delta t}|) \sim 1/|r_{\Delta t}|^{\gamma}$ (left panel) and $f(\alpha)$ (right panel) for two time windows that show clearly different properties of both cdf and $f(\alpha)$—a symmetric $f(\alpha)$ corresponding to a steep cdf with $\gamma \approx 3.2$ (a window covering July and August 2019) and an asymmetric $f(\alpha)$ corresponding to a heavy-tail cdf with $\gamma \approx 1.8$ (a window covering February and March 2020). These values of γ point to the aforementioned restriction $-3 \leq q \leq 3$ applied to the calculation of F^q_{xy} and $\Delta\alpha$: For $|q| > 3$ the moments of the distribution $p(|r_{\Delta t}|)$ can diverge, so can $F^q_{xy}(s)$ especially for small scales s.

The BTC/USDT return distribution function reflects a combination of two factors: (1) How fast the information spreads over the market—the heavier tails are, the slower this spreading proceeds, and (2) how volatile is the market—periods that cover turmoils with high volatility also result in heavier tails of pdf/cdf. It was documented in Ref. [11] that along with a process of the cryptocurrency market maturation the scaling exponent γ increases with time. This happens because as recognition of the market and its capitalization increase, more and more transactions take place, which decreases the average inter-transaction waiting time and allows the market participants to react faster. Faster reactions are crucial for the market to become efficient, which means more Gaussian-like fluctuations (larger γ). On the other hand, extremely large fluctuations and amplified volatility are characteristic for the periods with negative events, which decrease γ. Figure 2 (the 3rd panel from top) shows a scaling exponent γ obtained by fitting a power-law function to the BTC/USDT returns cdf in each position of the 30-day moving window. Indeed, such events like a bear market after July 2019 and the Covid-19 outbreak in March 2020 resulted in relatively small values of γ, while a bull market

between April and July 2019 and an escape from conventional assets to alternative ones observed between January and February 2020 led to larger values of γ.

Even if BTC is only one of many actively traded cryptocurrencies on the Binance platform, its strongest position due to the largest capitalization (between 50% and 70% of total market capitalization in the considered period) causes other cryptocurrencies to evolve accordingly. This'observation comes from the topmost panel of Figure 2 presenting $\alpha_{\min}$, α_0, and $\alpha_{\max}$ for the 8-cryptocurrency index. Qualitatively, the temporal course of these quantities does not differ much from the temporal course of their counterparts for BTC/USDT (the 2nd panel from top). The only significant difference is that for the index a transition to a bifractal-like $f(\alpha)$ spectrum in March 2020 was sharp and it was not preceded by its gradual change starting from January 2020 as it was the case with BTC/USDT.

By looking at the bottom panel of Figure 2, where total market capitalization is plotted as a function of time together with the Covid-19 pandemic severity parametrized by the number of daily new cases, and by comparing this plot with the remaining three, one can infer about how various market events and the pandemic influenced complexity of the market dynamics. The main events are denoted by Roman numerals: The beginning of the bull market in April 2019 (event I), its end in July 2019 (event II), the Covid-19 panic in March 2020 (event III), and the second pandemic wave that started in May 2020 (event IV). These events could be distinguished because they were associated with particularly large fluctuations (Figure 1). Among them, the events I, III, and IV had a significant impact on the multifractal properties of the exchange rate fluctuations by sizeable decreasing of $\alpha_{\min}$ (visible both for the cryptocurrency index and the BTC/USDT exchange rate). However, the event II did not have such an impact. In contrast, the pdf/cdf tails reflected overall market phase more than specific events except for the Covid-19 panic in March 2020.

2.4. Cryptocurrency Market Versus Standard Markets

From a practical point of view, among the most interesting issues related to any asset and any market is how much it is related to other assets or markets, and, in other words, whether it can be exploited for portfolio diversification and hedging [91–93]. As the investors may be interested in different time horizons and may want to hedge against events of different magnitude, the q-dependent detrended cross-correlation coefficient $\rho(q,s)$ defined by Equation (10) is a measure that is particularly useful in this context since it is sensitive to both scale and amplitude of the asset price returns. We choose the BTC/USD exchange rate as a representative of the whole cryptocurrency market—it is the most frequently traded asset, the most capitalized asset, and the most mature one (based on our previous results [11]). We calculate $\rho(q,s)$ for this rate and each of the remaining conventional assets listed in Section 2.1. However, we observe that this measure behaves similar for S&P500, Nasdaq100, and DJI, so we abandon the latter two indices and show only the results for S&P500. In parallel, we neglect AUD, NZD, ZAR, CHN, MXN, EUR, GBP, NOK, TRY, and PLN as their correlations with BTC were close to zero throughout the period under consideration. We consider two temporal scales that correspond to different horizons: $s = 10$ min, which is the shortest scale available provided we use 1-min returns, and $s = 360$ min that represents approximately a trading day in the US stock market. The latter value means that in a moving widow there was only 10 segments over which the averaging was carried out in $F^q_{xy}(s)$ (see Equation (4)), so we could not look at longer scales. As regards the parameter q, we focused on $q > 0$ in order to avoid the interpretation subtleties that could occur otherwise (see Section 2.2). We carried out our analysis for different values of q, but here we shall report only the results for $q = 1$ and $q = 4$. The former choice did not favour any value range of the fluctuation function F^2_{xy} since, for each segment ν in Equation (4), it was counted with the same weight. Therefore $q = 1$ allowed us for considering all time periods in the same way irrespective of whether the market was quiet or turbulent. On the other hand, $q = 4$ corresponds to favouring the segments with the largest return covariance and degrading the other segments. Thus, this case is interesting from a perspective of the investors that want to hedge against the largest price movements and the

largest-impact events. The intermediate values of q were also investigated, but the related results fell between these two cases and, thus, they are not presented here. Moreover, the calculations for $q > 4$ were progressively less interesting with increasing q as the event statistics became poor.

Figure 4 displays temporal course of $\rho(q,s)$ for a combination of the above-described cases of s and q. In each panel the cross-correlation coefficients for BTC and each of the 8 other assets are shown. Curiously, we do not observe any statistically significant values of $\rho(q,s)$ during the whole year 2019 even though there were then important events on the cryptocurrency market, like the bull and the subsequent bear market. However, these events were not related to any of the conventional assets considered here. We see that even the periods of high volatility in April and July-August 2019 did not cause any action that could potentially be sensed by the regular markets. We can explain this lack of reaction by a relatively small capitalization of the cryptocurrency market—far too low for the other markets to detect a possible influx of a capital withdrawn from cryptocurrencies (if such an influx actually took place). In 2019 there was no turmoil in the conventional markets, thus nothing could correlate the cryptocurrencies with the conventional assets from this direction, too.

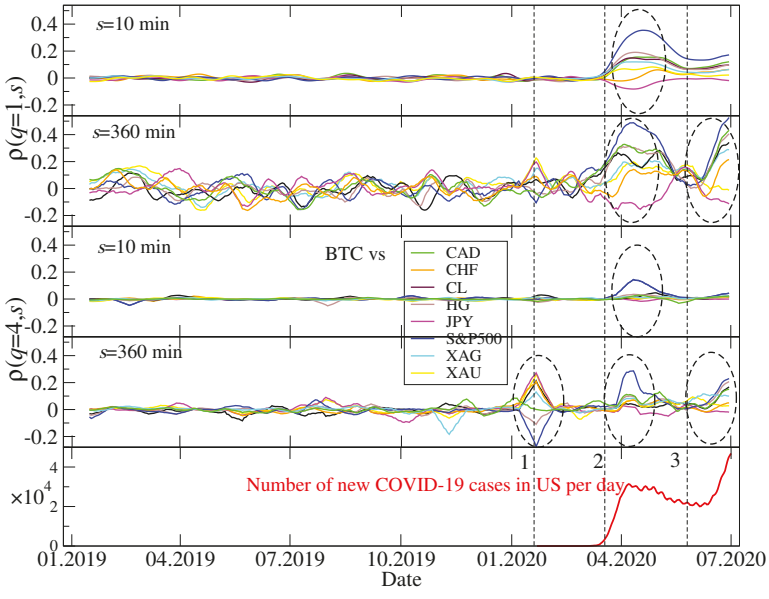


Figure 4. Temporal evolution of the detrended cross-correlation coefficient $\rho(q,s)$ calculated for the BTC/USD exchange rate and the conventional assets expressed in US dollar: Japanese yen (JPY), Canadian dollar (CAD), Swiss franc (CHF), crude oil (CL), silver (XAG), gold (XAU), copper (HG), and the S&P500 index. The $\rho(q,s)$ coefficient was calculated in a moving 10-day-long window with a step of 1 day and its s and q parameters are represented by $s = 10$ min (the shortest scale), $s = 360$ min (approximately a trading day in the US stock market), $q = 1$ (all data points are considered), and $q = 4$ (only the data points with large amplitude are considered). In each panel events with the statistically significant, genuine cross-correlations are marked with dashed ellipses. The daily number of new Covid-19 cases in the United States is also shown for a comparison (bottom). The particular market events are indicated: (1) A sharp drop of the US stock market indices after the first case of Covid-19 had been identified in the United States; (2) a Covid-19 outburst related panic on the financial markets; (3) a bear market return on risky assets that was related to the 2nd wave of the pandemic.

In contrast, there were 1 to 3 periods of significant inter-market cross-correlations in the first half of 2020, dependent on s and q. The first important period in the end of January and the begin of February was associated with a sharp drop of S&P500 and other US stock market indices triggered by the first identified local case of Covid-19. The cryptocurrency market reacted rather moderately with only a short period of large and delayed fluctuations. This is why there are no significant elevation of $|\rho(q,s)|$ for short scales for any return size. The cryptocurrency market must have been calm enough to delay reaction so long that it is identifiable only on large scales (like $s = 360$ min). The cross-correlation is positive with the fiat currencies, while negative with the US stock markets. As all the considered assets are expressed in USD, the positive correlations of BTC with the fiat currencies in January/February 2020 mean that there was a global flee from US dollar to other major currencies that increased the corresponding exchange rates as well as a flee from the US stock markets to the cryptocurrency market.

Opposite situation took place during the pandemic's 2nd wave in June 2020 (and, possibly, beyond that month): The cross-correlations are stronger for $q = 1$ than for $q = 4$. On the one hand, for $s = 10$ min moderate values of $\rho(q,s)$, mainly positive ones, are seen for $q = 1$, but they are not seen for $q = 4$. On the other hand, for $s = 360$ min large values of $\rho(q,s)$ are observed for $q = 1$ and slightly smaller, but also significant, for $q = 4$. Therefore we still see that the correlations cannot be built in their full magnitude on short scales and they need some time to develop completely. However, a larger $\rho(q,s)$ for $q = 1$ indicates that the cross-correlations affect all the returns irrespective of their amplitude (small and moderate returns dominate in number, thus more segments in $F^q_{xy}(s)$ are correlated in this case than in the case of $q = 4$). Majority of assets are correlated positively with an exception for JPY that is anticorrelated with BTC.

The third and the most important interval of the inter-market cross-correlations happened between the two above discussed events and it covers the pandemic outbreak and a financial market panic in March 2020. The mutual coupling of the different markets (including the BTC/USD exchange rate for the first time) was especially evident in this case. We observe the same rule here as in the two previous cases that the cross-correlations need time to build up, thus they are stronger for $s = 360$ min than for $s = 10$ min. However, they are clearly evident even for $s = 10$ min. Interestingly, if we look at the largest returns ($q = 4$), we see that BTC is (positively) correlated mainly with S&P500, while it is more independent as regards other assets. If we take a look at the results for $q = 1$, the cross-correlations appear strong between BTC and all other assets except for CHF (only small negative correlation) and gold (XAU). The corresponding values of $\rho(q,s)$ are positive for S&P500, CAD, copper (HG), crude oil (CL), and silver (XAG), while they are negative for JPY. This cannot be viewed as a surprise since the Swiss franc and Japanese yen are considered safe assets together with gold and their pricing in USD behave differently than the remaining assets' pricing did.

For a comparison, Figure 5 shows $\rho(q,s)$ calculated for the ETH/USDT exchange rate and the same conventional assets as in the BTC/USDT case above. We see that the only qualitative difference between Figures 4 and 5 is a much smaller detrended cross-correlation coefficient value for $q = 4$ in the case of ETH/USDT and Event 1 (the first US Covid-19 case). In fact, the principal interest was then directed towards BTC and not the other cryptocurrencies. We did not analyze other cryptocurrencies, because only BTC and ETH are traded on the Dukascopy platform, which all the conventional asset quotes used here came from. If other cryptocurrencies were taken into consideration, they must have been taken from Binance and synchronized additionally, which might have introduced some spurious correlations. This is why we restricted this analysis to BTC and ETH only.

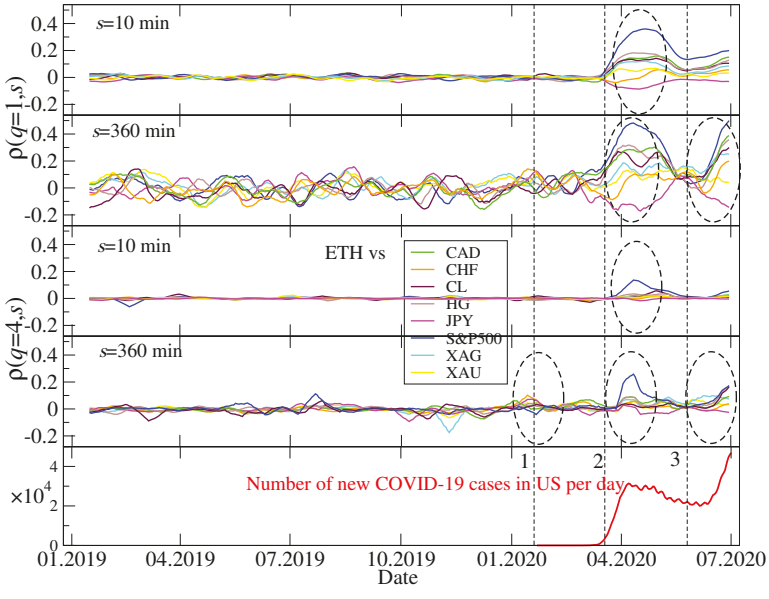


Figure 5. Temporal evolution of $\rho(q,s)$ calculated for the ETH/USDT exchange rate and the conventional assets expressed in US dollar: Japanese yen (JPY), Canadian dollar (CAD), Swiss franc (CHF), crude oil (CL), silver (XAG), gold (XAU), copper (HG), and the S&P500 index. For more description see caption to Figure 4.

2.5. Cryptocurrency Market Structure

We have already discussed the fractal autocorrelations of the cryptocurrency exchange rates with respect to US dollar and the cross-correlations between bitcoin and the assets representing conventional markets. Now its time to look at the inner correlation structure of the cryptocurrency market itself. Our data set consists of 128 cryptocurrencies expressed in BTC. This effectively removes the impact of BTC on any other coin, so we have some insight into the market's finer, secondary correlation structure (the primary structure is such that all the cryptocurrencies are correlated with BTC and form the market as a connected whole [11]). In our earlier work we identified that throughout short history of the market, there were only two cryptocurrencies that played the role of the market's center (in terms of the network centrality): BTC for the most time and ETH in the first half of 2018. ETH, sometimes together with USDT, was also identified as the most frequent secondary hub of the market, after BTC [11]. Here we study the market's structure between January 2019 and June 2020—a period that was not a subject of the previous study.

Minimal spanning tree is an acyclic spanning subset of a complete weighted network that is minimal in terms of the total length of its edges. In a typical MST construction, the Pearson correlation coefficient [94] is used to form a correlation matrix that defines a complete network. Here we follow Refs. [11,87] and define the network based on the $\rho(q,s)$ matrix. This matrix has entries equal to $\rho(q,s)$ calculated for all possible pairs of the exchange rates X/BTC and Y/BTC, where X,Y denote any cryptocurrency from our $N = 128$ element set. By doing this, we obtain $N(N-1)/2 = 128 * 127/2 = 8128$ coefficients $\rho(q,s)$ for each choice of q and s (as before, here we restrict our discussion to $q = 1$, $q = 4$, $s = 10$ min, and $s = 360$ min). In order to move to a metric space, we recalculate the coefficients in a form of a distance:

$$d_{XY}(q,s) = \sqrt{2\Big(1 - \rho_{XY}(q,s)\Big)}. \tag{11}$$

Since $-1 \leq \rho(q,s) \leq 1$ for $q > 0$, we obtain limiting values for distance: $0 \leq d_{XY}(q,s) \leq 2$, where $d_{XY} = 0$ means perfect cross-correlation between X and Y, $d_{XY} = 2$ means perfect anticorrelation, and $d_{XY} = \sqrt{2}$ means perfect statistical independence. Based on all values of $d_{XY}(q,s)$ we construct MST by using the Prim's algorithm [95].

Figure 6 shows qMSTs calculated for $q = 1$ for three specific periods (from top to bottom): January 2019, July 2019, and March 2020. The first one was distinguished because it overlaps with a period when ETH was a hub with the highest network centrality (the largest number of connections or the largest degree) for all scales. For both other periods, a role of the central hub was played by another cryptocurrency: USDT—in July 2019 (all scales) and in March 2020 for short scales (represented by $s = 10$ min in Figure 6). However, no overwhelmingly dominant node was observed in MST corresponding to March 2020 and $s = 360$ min. In fact, the structure of the latter MST differs substantially from the structure of the remaining 5 trees in Figure 6: It can be categorized as a distributed network in contrast to the generally centralized form of the rest, where there is a clearly identifiable center (ETH or USDT) and the peripheries. There is a possible explanation why USDT becomes a central hub in turbulent periods, especially the sudden dropdowns: Investors that want to close the cryptocurrency positions change them primarily to USDT, which is a stable coin pegged to USD [52] and only then to the proper US dollar. This manoeuvre can mutually correlate most cryptocurrencies via USDT.

Figure 6. Minimal spanning trees (MSTs) calculated based on the q-dependent detrended correlation coefficient $\rho(q,s)$ for the exchange rates of a form X/BTC, where X stands for one of 128 cryptocurrencies traded on Binance [21]. Each node is labeled by the corresponding cryptocurrency ticker. All trees correspond to $q = 1$. On the left there are MSTs obtained for $s = 10$ min, while on the right there MSTs obtained for $s = 360$ min. Each row shows MSTs calculated in a different period (a 7-day-long moving window with a step of 1 day): January 2019 (top), July 2019 (middle), and March 2020 (bottom).

Now let us consider MSTs constructed from the filtered signals, in which the largest returns were amplified by taking $q = 4$ (see Figure 7). In this case MSTs show a richer pool of forms. Only one tree shows a centralized topology: For $s = 10$ min and March 2020, though its central hub (USDT) does not dominate the networks unlike it was for $q = 1$ (Figure 6). Moreover, there is only one tree that can be categorized as distributed: For $s = 360$ min and March 2020. All the remaining trees reveal intermediate form between the centralized and distributed ones: There are several nodes that can be called local hubs. This is the case of the hierarchical networks that sometimes are scale-free. For $s = 10$ min, such a situation was present in January 2019 (ETH and USDT) and July 2019 (USDT, ONT, XLM, THETA, BCPT, and RVN), while for $s = 360$ min similar situations also occurred in January 2019 (ETH, EOS, and LTC) and in July 2019 (XLM, THETA, LOOM, USDT, ADA, AION, and DAX).

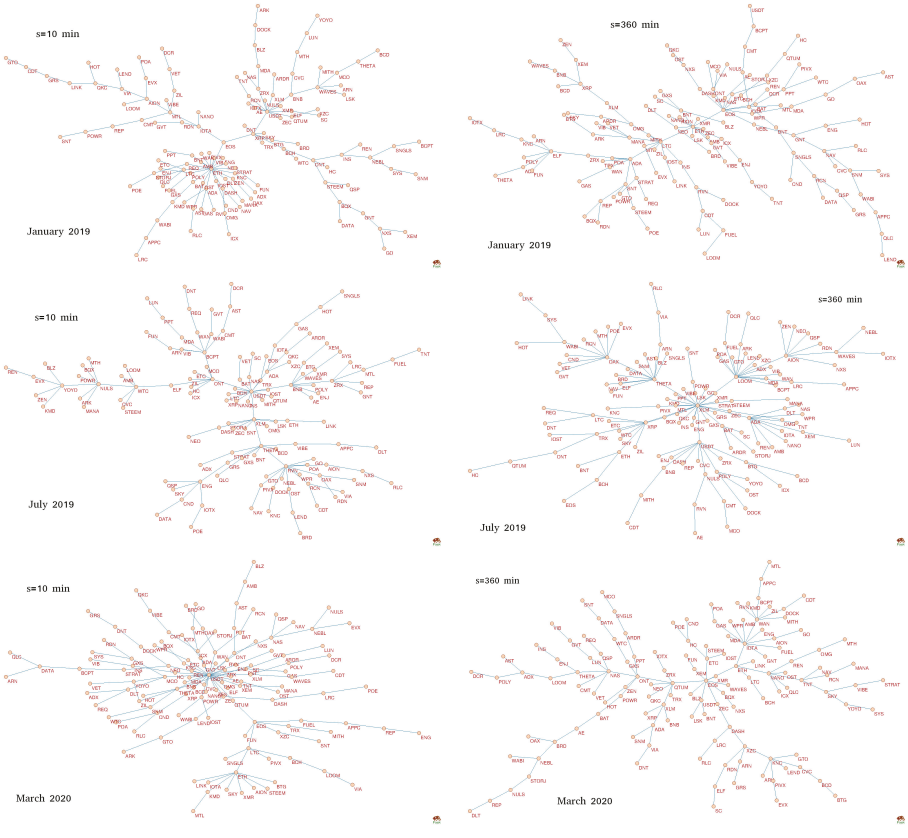


Figure 7. Minimal spanning trees (MSTs) calculated based on the q-dependent detrended correlation coefficient $\rho(q, s)$ for the exchange rates of a form X/BTC, where X stands for one of 128 cryptocurrencies traded on Binance [21]. Each node is labeled by the corresponding cryptocurrency ticker. All trees correspond to $q = 4$. On the left there are MSTs obtained for $s = 10$ min, while on the right there MSTs obtained for $s = 360$ min. Each row shows MSTs calculated in a different period (a 7-day-long moving window with a step of 1 day): January 2019 (top), July 2019 (middle), and March 2020 (bottom).

The trees shown in Figures 6 and 7 represent only a few periods, but in order to look at the market structure evolution over the whole considered interval of time, it is not convenient to look at the trees for individual windows. Therefore, we calculated a few network characteristics that grasp the essential properties of the MST topology in each window. These are the mean path length $\langle L(q, s) \rangle$

between a pair of the MST nodes (the averaging is carried out over all possible pairs) describing how distributed (large $\langle L(q,s)\rangle$) or concentrated (small $\langle L(q,s)\rangle$) is a tree, the mean q-dependent detrended cross-correlation coefficient $\langle \rho(q,s)\rangle$ (the averaging is carried out over all possible cryptocurrency pairs), describing how strong are typical network edges, and the maximum node degree $k_{\max}(q,s)$, describing how central is the main hub. Time evolution of these quantities is shown in Figure 8 for $q = 1$ and in Figure 9 for $q = 4$. Apart from two scales considered in Figures 6 and 7, i.e., $s = 10$ min and $s = 360$ min, we added a medium scale of $s = 60$ min.

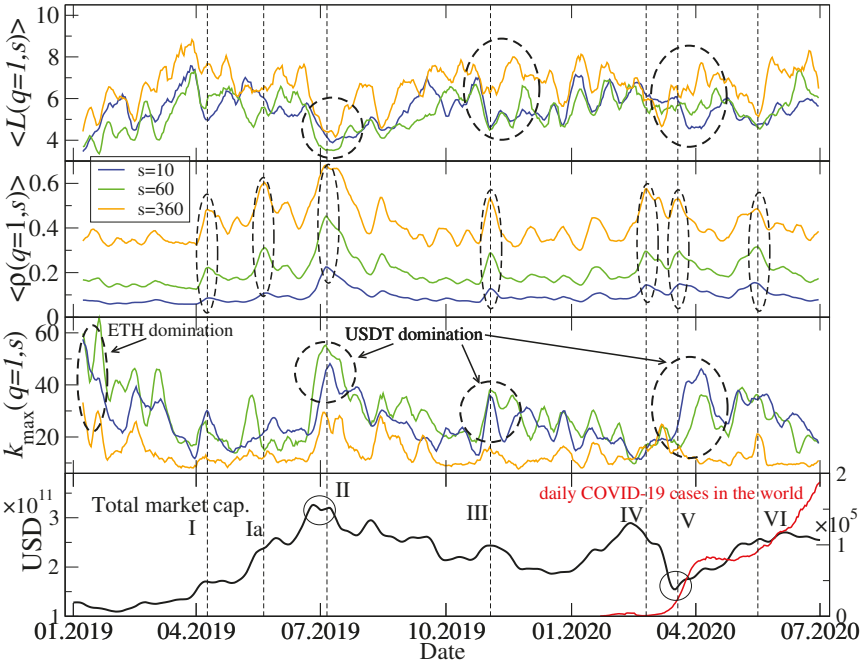


Figure 8. Network characteristics describing minimal spanning trees (MSTs) calculated for $q = 1$ and for the following scales: $s = 10$ min, $s = 60$ min, and $s = 360$ min. The average path length $\langle L(q,s)\rangle$ between a pair of MST nodes (top), the average q-dependent detrended cross-correlation coefficient $\langle \rho(q,s)\rangle$ (upper middle), the maximum node degree $k_{\max}(q,s)$ (lower middle), together with the total market capitalization in US dollars and the daily number of new Covid-19 cases in the world (bottom). Several events related to a relatively strong cross-correlations are marked with vertical dashed lines, Roman numerals, and dashed ellipses: Start of a bull market in April 2019 (event I) and its continuation in May 2019 (event Ia), a peak of the bull market in July 2019 (event II), a local peak followed by a sharp drop of the market capitalization in November 2019 (event III), the Covid-19 panic in mid March 2020 (events IV-V), and the 2nd Covid-19 wave from May 2020 (event VI).

While $\langle \rho(q,s)\rangle$ is largely a different measure than the two other ones, $\langle L(q,s)\rangle$ and $k_{\max}(q,s)$ can be related to each other: If $k_{\max}(q,s)$ is large, a majority of the nodes is connected to it and $\langle L(q,s)\rangle$ can be small; the opposite relation is also true. Figure 8 confirms these observations for $q = 1$: Typically, the elevated values of $\langle L(1,s)\rangle$ (top panel) are associated with the suppressed values of $k_{\max}(1,s)$ (lower middle panel) no matter what was a particular cause of such a change of the MST structure. The most important topological changes detectable by $\langle L(1,s)\rangle$ and $k_{\max}(1,s)$ occurred after the end of ETH domination in the market in January–February 2019 (topolgy changed from a highly centralized one with ETH being the hub to a rather distributed one), after the end of the bull phase in July–August 2019 (topology returned temporarily to a centralized form but with USDT as the central hub), during a local

peak and the subsequent decline of the market in November 2019 (another short period of a centralized topology with USDT domination), and during and after the Covid-19 outbreak March–May 2020 (another phase of USDT domination, but longer than the preceding ones). On the level of $\langle \rho(1,s) \rangle$, there can be 7 interesting periods pointed out (upper middle panel of Figure 8). As one might expect, the longer scale, the stronger are the mean cross-correlations; this is a systematical relation throughout the whole analyzed time interval. This is a typical effect observed on many financial markets, which is related to the liquidity and capitalization differences among the assets. Since the cryptocurrencies with small capitalization are traded less frequently than those with large capitalization, it takes more time for a piece of market information to spread over such cryptocurrencies. Thus, the cross-correlations among them can only be built and detected on longer scales.

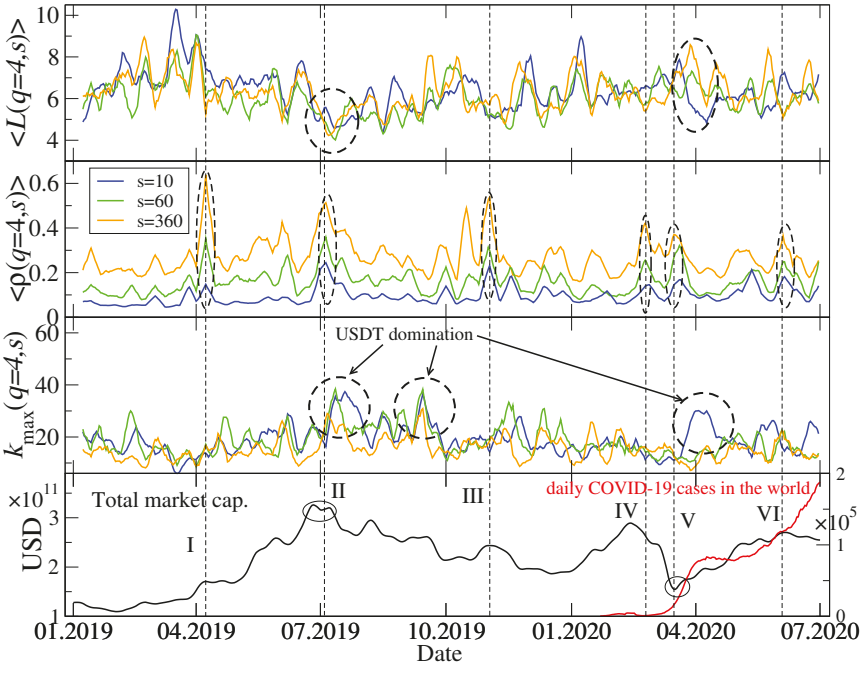


Figure 9. The same network characteristics describing MSTs as in Figure 8, but here calculated for $q = 4$.

A more interesting situation as regards the different scales s can be found if one compares, on the one hand, $k_{\max}(1,s)$ between these scales and, on the other hand, $\langle L(1,s) \rangle$. Let us look at two events: A peak and decline of the bull market in July 2019 (event II) and the Covid-19 pandemic (events IV–VI). During the former, $k_{\max}(1,s)$ shows a standard behaviour, i.e., for $s = 10$ min and $s = 60$ min it is significantly larger than for $s = 360$ min; the same can be said for the events IV-VI. According with what it has been said above, we might expect that in both cases $\langle L(1, 360 \text{ min}) \rangle$ should be larger than $\langle L(1, 10 \text{ min}) \rangle$ and $\langle L(1, 60 \text{ min}) \rangle$. while this was the case, indeed, during the pandemic outbreak in March 2020, nothing like this happened during the bull market peak in July 2019, when $\langle L(1,s) \rangle$ was comparable for all the scales. Such a deviation from the overall rule that a longer scale is associated with a better-developed hierarchical or a more distributed MST topology (smaller $k_{\max}(q,s)$) and a shorter scale is associated with either a more centralized network topology (larger $k_{\max}(q,s)$) was rather unusual as for the whole studied period.

Figure 9 differs from Figure 8 only in that it shows the same quantities but for $q = 4$ (mainly the cross-correlations for the large-amplitude returns are considered now). The above-discussed relation

between $k_{\max}(q,s)$ for different scales s is less clear for $q=4$ than it is for $q=1$. Only in July 2019 and in March-April 2020 there can be distinguished some characteristic structures in time evolution of $k_{\max}(4,s)$ and $\langle L(4,s)\rangle$. For the events that took place in July 2019, a relation between values of the maximum node degree for different scales and a relation between values of the mean path length also for different scales resemble those identified for $q=1$. For the Covid-19 outbreak period, the small difference is that now $k_{\max}(4,60\text{ min})$ is comparable to its counterpart for $s=360$ min instead of $s=10$ min as for $q=1$. There is no difference between $q=1$ and $q=4$ as regards $\langle\rho(4,s)\rangle$: The longer the scale is, the stronger are the cross-correlations.

To summarize observations related to the MST topology, in the analyzed period from January 2019 to June 2020 this topology used to change substantially during periods of large volatility in such a way that from a hierarchical or distributed network structure that was typical outside these periods it used to transform itself to a more centralized structure with a dominating hub and much stronger cross-correlations between the nodes (see also [38]). The most interesting period was the Covid-19 pandemic, during which on short and moderate scales for $q=1$ one observed first a significant increase of the MST centralization (large $k_{\max}(q,s)$) and a subsequent slow return to a more distributed form (moderate $k_{\max}(q,s)$) but still with a distinguished central hub. However, on the longest scale this effect was not observed and $k_{\max}(1,360\text{ min})$ was elevated only once in May 2020. This suggests that the most sudden and nervous movements that correlate the market and centralize its topology on short time scales tend to be blurred as time passes and we go from short to long scales, where topology becomes much more of a distributed or hierarchical type. Such a behaviour observed recently during the pandemic, which can be considered as an external perturbation to the market, differs from the behaviour observed during the peak and collapse of the bull market in July 2019, which was no doubt a result of the internal evolution of the market. Whether this internal/external events may be source of the observed peculiarities of the Covid-19 period, one cannot state for sure as both events were unique during the analyzed time interval and cannot be confirmed by other events of similar type.

3. Summary

In our work we focused on dynamical and structural properties of the cryptocurrency market. We analyzed empirical data representing the exchange rates of 129 cryptocurrencies traded on the Binance platform, including BTC. The analysis comprised three parts, each of which was intended for investigating a different aspect of the market structure. We started from a multifractal analysis of the BTC/USDT exchange rate as the most important one together with a similar analysis of an artificial cryptocurrency index based on 8 the most capitalized coins. This analysis may be considered as an extension of the analysis reported in Ref. [11] on the most recent time interval from January 2019 to June 2020. The results showed that throughout this interval the cryptocurrency dynamics produces multifractal fluctuations (returns) with some intermittent signatures of bifractality that can be assigned to specific volatile periods like the Covid-19 outburst in March 2020 or a bull market start in April 2019. Moreover, on a level of the return distributions such bifractal-like singularity spectra can be accounted for by the pdf/cdf power-law tails that fall into the Lévy-stable regime [90]. Outside these volatile periods spectra are wide but with much smaller left-right asymmetry.

The analysis of the cross-correlations between the cryptocurrency market represented by BTC/USD or ETH/USD and the conventional markets represented by the major fiat currencies, the most important commodities (e.g., crude oil and gold), and the US stock market indices brought us to an observation that the cryptocurrency market was decoupled from the remaining markets throughout the whole year 2019, but it used to couple temporarily to those markets during some events in the first half of 2020, like in January when the first Covid-19 case was reported in the United States, in March during the pandemic outbreak, and in May-July during the pandemic's 2nd wave. In the first case, BTC was anticorrelated with the major stock market indices like S&P500 and Nasdaq100, but in the second and the third cases the analogous cross-correlations were positive. Positive were then also the cross-correlations between BTC and several fiat currencies and commodities. A lack

of the statistically valid cross-correlations in 2019, when the conventional assets did not experience anything turbulent, was supposedly caused by the asymmetry in market capitalization between the cryptocurrency market and the conventional markets to the disadvantage of the cryptocurrency market, which was too small to have any sizeable impact on the other markets. However, the conventional markets can easily influence the cryptocurrency market if they are turbulent. This is exactly what was observed in March 2020 and June 2020. Except for January 2020, when, unlike BTC/USD, ETH/USD was not correlated with the conventional assets, both the exchange rates reveal a similar relation with these assets.

A network representation of the cryptocurrency market can shed light on the market's inner cross-correlation structure. Our analysis based on the exchange rates of 128 coins with respect to BTC revealed that turbulent periods on the market result in a sudden transition between different network topology types. During the periods of normal dynamics, the market has a distributed-network topology or a hierarchical-network topology, in which no node dominates the network and there is a hierarchy of hubs with decreasing centrality (e.g., node degree). Typically, for long scales the hierarchical-network topology is more pronounced than for short scales, where a centralized-network topology prevails. This is because the cryptocurrencies of small capitalization are less liquid, so a piece of information needs more time to be fully processed by them and the cross-correlations, especially those more subtle, sector-like, and related to less prominent cryptocurrencies, can only build up on sufficiently long time scales. This picture is altered if there comes a volatile period. During such periods the network becomes highly centralized with one dominating hub for all the scales. The most often the role of such a hub is played by USDT, because it is pegged to the US dollar and, thus, considered as more stable than other cryptocurrencies. If investors flee the cryptocurrency market, they first change their assets to USDT, and only then to USD, which can correlate a majority of the cryptocurrencies together via USDT. However, a compact, star-like topological form exists shortly and soon it returns to a more distributed, more branched form.

We also noticed that the most significant events as regards their impact on the market topology—the transition from a bull market to a bear market in July 2019 and the Covid-19 pandemics that started in March 2020—differ in some details of that impact. During the pandemics, a transition from a centralized form to a distributed form occurred predominantly on short and medium scales, while on long scales it was less pronounced. Contrary to that, in July 2019 the topology shift was visible on all the scales. It is a matter of future analyses to address a question whether this difference can be related to endogenous (a trend reversal) vs. exogenous (the pandemic) origin of both events.

Author Contributions: Conceptualization, S.D., P.O., and M.W.; methodology, S.D., J.K., P.O., T.S., and M.W.; software, T.S. and M.W.; validation, S.D., J.K., P.O., T.S., and M.W.; formal analysis, P.O., T.S., and M.W.; investigation, P.O., T.S., and M.W.; resources, T.S. and M.W.; data curation, T.S. and M.W.; writing–original draft preparation, J.K. and M.W.; writing–review and editing, J.K. and M.W.; visualization, T.S. and M.W.; supervision, S.D. and P.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kwapień, J.; Drożdż, S. Physical approach to complex systems. *Phys. Rep.* **2012**, *515*, 115–226. [CrossRef]
2. Jakimowicz, A. The role of entropy in the development of economics. *Entropy* **2020**, *22*, 452. [CrossRef]
3. Klamut, J.; Kutner, R.; Struzik, Z.R. Towards a universal measure of complexity. *Entropy* **2020**, *22*, 866. [CrossRef]
4. Bak, P.; Norrelykke, S.F.; Shubik, M. Money and Goldstone modes. *Quant. Financ.* **2001**, *1*, 186–190. [CrossRef]
5. Oświęcimka, P.; Drożdż, S.; Gębarowski, R.; Górski, A.Z.; Kwapień, J. Multiscaling Edge Effects in an Agent-based Money Emergence Model. *Acta Phys. Pol. B* **2015**, *46*, 1579–1592. [CrossRef]
6. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008. Available online: https://git.dhimmel.com/bitcoin-whitepaper/ (accessed on 20 August 2020).

7. Wattenhofer, R. *The Science of the Blockchain*, 1st ed.; CreateSpace Independent Publishing Platform: Scotts Valley, CA, US, 2016.
8. Gerlach, J.-C.; Demos, G.; Sornette, D. Dissection of Bitcoin's multiscale bubble history from January 2012 to February 2018. *R. Soc. Open Sci.* **2019**, *6*, 180643. [CrossRef]
9. Ethereum. Available online: https://www.ethereum.org/ (accessed on 20 August 2020).
10. Aste, T. Cryptocurrency market structure: Connecting emotions and economics. *Digit. Financ.* **2019**, *1*, 5–21. [CrossRef]
11. Wątorek, M.; Drożdż, S.; Kwapień, J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale Characteristics of the Emerging Global Cryptocurrency Market, to Be Published. 2020.
12. Cambridge Bitcoin Electricity Consumption Index. Available online: https://www.cbeci.org/ (accessed on 20 August 2020).
13. Proof-of-Stake. Available online: https://academy.binance.com/blockchain/proof-of-stake-explained (accessed on 20 August 2020).
14. Ripple. Available online: https://ripple.com/ (accessed on 20 August 2020).
15. "Monero Ransom". Available online: https://www.nytimes.com/2019/01/10/world/europe/norway-kidnapping-monero.html (accessed on 20 August 2020).
16. Monero. Available online: https://web.getmonero.org/library/Zero-to-Monero-2--0-0.pdf (accessed on 20 August 2020).
17. CME Group. Available online: https://www.cmegroup.com/ (accessed on 20 August 2020).
18. Makarov, I.; Schoar, A. Trading and arbitrage in cryptocurrency markets. *J. Financ. Econ.* **2020**, *135*, 293–319. [CrossRef]
19. Gębarowski, R.; Oświęcimka, P.; Wątorek, M.; Drożdż, S. Detecting correlations and triangular arbitrage opportunities in the Forex by means of multifractal detrended cross-correlations analysis. *Nonlinear Dyn.* **2019**, *98*, 2349–2364. [CrossRef]
20. CoinMarketCap. Available online: https://coinmarketcap.com (accessed on 20 August 2020).
21. Binance. Available online: https://www.binance.com/ (accessed on 20 August 2020).
22. Kristoufek, L. Bitcoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Sci. Rep.* **2013**, *3*, 3415. [CrossRef]
23. Kristoufek, L. What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PLoS ONE* **2015**, *10*, 1–15. [CrossRef] [PubMed]
24. Bariviera, A.F.; Basgall, M.J.; Hasperué, W.; Naiouf, M. Some stylized facts of the Bitcoin market. *Phys. A* **2017**, *484*, 82–90. [CrossRef]
25. Drożdż, S.; Gębarowski, R.; Minati, L.; Oświęcimka, P.; Wątorek, M. Bitcoin market route to maturity? Evidence from return fluctuations, temporal correlations and multiscaling effects. *Chaos* **2018**, *28*, 071101. [CrossRef] [PubMed]
26. Garnier, J.; Solna, K. Chaos and order in the bitcoin market. *Phys. A* **2019**, *524*, 708–721. [CrossRef]
27. Wu, K.; Wheatley, S.; Sornette, D. Classification of cryptocurrency coins and tokens by the dynamics of their market capitalizations. *R. Soc. Open Sci.* **2018**, *5*, 180381. [CrossRef]
28. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Signatures of crypto-currency market decoupling from the Forex. *Future Internet* **2019**, *11*, 154. [CrossRef]
29. Kristoufek, L.; Vosvrda, M. Cryptocurrencies market efficiency ranking: Not so straightforward. *Phys. A* **2019**, *531*, 120853. [CrossRef]
30. Stosić, D.; Stosić, D.; Ludermir, T.B.; Stosić, T. Collective behavior of cryptocurrency price changes. *Phys. A* **2018**, *507*, 499–509. [CrossRef]
31. Bariviera, A.F.; Zunino, L.; Rosso, O.A. An analysis of high-frequency cryptocurrencies prices dynamics using permutation-information-theory quantifiers. *Chaos* **2018**, *28*, 075511. [CrossRef]
32. Bouri, E.; Gupta, R.; Roubaud, D. Herding behaviour in cryptocurrencies. *Financ. Res. Lett.* **2019**, *29*, 216–221. [CrossRef]
33. Zięba, D.; Kokoszczyński, R.; Śledziewska, K. Shock transmission in the cryptocurrency market. Is Bitcoin the most influential? *Int. Rev. Financ. Anal.* **2019**, *64*, 102–125. [CrossRef]
34. Drożdż, S.; Minati, L.; Oświęcimka, P.; Stanuszek, M.; Wątorek, M. Competition of noise and collectivity in global cryptocurrency trading: Route to a self-contained market. *Chaos* **2020**, *30*, 023122. [CrossRef] [PubMed]

35. Ferreira, P.; Kristoufek, L.; Johnson de Area Leao Pereira, E. DCCA and DMCA correlations of cryptocurrency markets. *Phys. A* **2020**, *545*, 123803. [CrossRef]
36. Papadimitriou, T.; Gogas, P.; Gkatzoglou, F. The evolution of the cryptocurrencies market: A complex networks approach. *J. Comp. Appl. Math.* **2020**, *376*, 112831. [CrossRef]
37. Polovnikov, K.; Kazakov, V.; Syntulsky, S. Core-periphery organization of the cryptocurrency market inferred by the modularity operator. *Phys. A* **2020**, *540*, 123075. [CrossRef]
38. García-Medina, A.; Hernández, J.B. Network analysis of multivariate transfer entropy of cryptocurrencies in times of turbulence. *Entropy* **2020**, *22*, 760. [CrossRef]
39. Corbet, S.; Meegan, A.; Larkin, C.; Lucey, B.; Yarovaya, L. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Econ. Lett.* **2018**, *165*, 28–34. [CrossRef]
40. Corelli, A. Cryptocurrencies and exchange rates: A relationship and causality Analysis. *Risks* **2018**, *6*, 111. [CrossRef]
41. Ji, Q.; Bouri, E.; Gupta, R.; Roubaud, D. Network causality structures among Bitcoin and other financial assets: A directed acyclic graph approach. *Quart. Rev. Econ. Financ.* **2018**, *70*, 203–213. [CrossRef]
42. Kristjanpoller, W.; Bour, E. Asymmetric multifractal cross-correlations between the main world currencies and the main cryptocurrencies. *Phys. A* **2019**, *523*, 1057–1071. [CrossRef]
43. Manavi, S.A.; Jafari, G.; Rouhani, S.; Ausloos, M. Demythifying the belief in cryptocurrencies decentralized aspects. A study of cryptocurrencies time cross-correlations with common currencies, commodities and financial indices. *Phys. A* **2020**, *556*, 124759. [CrossRef]
44. Urquhart, A.; Zhang, H. Is Bitcoin a hedge or safe haven for currencies? An intraday analysis. *Int. Rev. Financ. Anal.* **2019**, *63*, 49–57. [CrossRef]
45. Shahzad, S.J.H.; Bouri, E.; Roubaud, D.; Kristoufek, L.; Lucey, B. Is Bitcoin a better safe-haven investment than gold and commodities? *Int. Rev. Financ. Anal.* **2019**, *63*, 322–330. [CrossRef]
46. Shahzad, S.J.H.; Bouri, E.; Roubaud, D.; Kristoufek, L. Safe haven, hedge and diversification for G7 stock markets: Gold versus bitcoin. *Econ. Model.* **2019**, *87*, 212–224. [CrossRef]
47. Wang, P.; Zhang, W.; Li, X.; Shen, D. Is cryptocurrency a hedge or a safe haven for international indices? A'comprehensive and dynamic perspective. *Financ. Res. Lett.* **2019**, *31*, 1–18. [CrossRef]
48. Fang, F.; Ventre, C.; Basios, M.; Kong, H.; Kanthan, L.; Li, L.; Martinez-Regoband, D.; Wu, F. Cryptocurrency Trading: A Comprehensive Survey. *arXiv* **2020**, arXiv:2020.11352.
49. Corbet, S.; Lucey, B.; Urquhart, A.; Yarovaya, L. Cryptocurrencies as a financial asset: A systematic analysis. *Int. Rev. Financ. Anal.* **2019**, *62*, 182–199. [CrossRef]
50. Gopikrishnan, P.; Meyer, M.; Amaral, L.A.N.; Stanley, H.E. Inverse cubic law for the distribution of stock price variations. *Eur. Phys. J. B* **1998**, *3*, 139–140. [CrossRef]
51. Gabaix, X.; Gopikrishnan, P.; Plerou, V.; Stanley, H.E. A theory of power-law distributions in financial market fluctuations. *Nature* **2003**, *423*, 267–270. [CrossRef]
52. Tether. Available online: https://tether.to (accessed on 20 August 2020).
53. Dukascopy. Available online: https://www.dukascopy.com (accessed on 20 August 2020).
54. Subramaniam, A.R.; Gruzberg, I.A.; Ludwig, A.W.W. Boundary criticality and multifractality at the two-dimensional spin quantum Hall transition. *Phys. Rev. B* **2008**, *78*, 245105. [CrossRef]
55. Ivanov, P.C.; Amaral, L.A.N.; Goldberger, A.L.; Havlin, S.; Rosenblum, M.G.; Struzik, Z.R.; Stanley, H.E. Multifractality in human heartbeat dynamics. *Nature* **1999**, *399*, 461–465. [CrossRef]
56. Stanley, H.E.; Meakin, P. Multifractal phenomena in physics and chemistry. *Nature* **1988**, *335*, 405–409. [CrossRef]
57. Witt, A.; Malamud, B.D. Quantification of long-range persistence in geophysical time series: Conventional and benchmark-based improvement techniques. *Surv. Geophys.* **2013**, *34*, 541–651. [CrossRef]
58. Koscielny-Bunde, E.; Kantelhardt, J.W.; Braund, P.; Bunde, A.; Havlin, S. Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies. *J. Hydrol.* **2006**, *322*, 120–137. [CrossRef]
59. Kantelhardt, J.W.; Koscielny-Bunde, E.; Rybski, D.; Braun, P.; Bunde, A.; Havlin, S. Long-term persistence and multifractality of precipitation and river runoff records. *J. Geophys. Res. Atmos.* **2006**, *111*, D01106. [CrossRef]
60. Drożdż, S.; Oświęcimka, P.; Kulig, A.; Kwapień, J.; Bazarnik, K.; Grabska-Gradzińska, I.; Rybicki, J.; Stanuszek, M. Quantifying origin and character of long-range correlations in narrative texts. *Inf. Sci.* **2016**, *331*, 32–44 [CrossRef]

61. Ihlen, E.A.F.; Vereijken, B. Multifractal formalisms of human behavior. *Hum. Mov. Sci.* **2013**, *32*, 633–651. [CrossRef]
62. Dixon, J.A.; Holden, J.G.; Mirman, D.; Stephen, D.G. Multifractal dynamics in the emergence of cognitive structure. *Top. Cogn. Sci.* **2012**, *4*, 51–61. [CrossRef]
63. Jafari, G.R.; Pedram, P.; Hedayatifar, L. Long-range correlation and multifractality in Bach's Inventions pitches. *J. Stat. Mech.* **2007**, *2007*, P04012. [CrossRef]
64. Roeske, T.C.; Kelty-Stephen, D.; Wallot, S. Multifractal analysis reveals musiclike dynamic structure in songbird rhythms. *Sci. Rep.* **2018**, *8*, 4570. [CrossRef]
65. Nagy, Z.; Mukli, P.; Herman, P.; Eke, A. Decomposing multifractal crossovers. *Front. Physiol.* **2017**, *8*, 533. [CrossRef] [PubMed]
66. Domański, P.D. Multifractal properties of process control variables. *Int. J. Bifurc. Chaos* **2017**, *27*, 1750094. [CrossRef]
67. Krawczyk, M.J.; Oświęcimka, P.; Kułakowski, K.; Drożdż, S. Ordered avalanches on the Bethe lattice. *Entropy* **2019**, *21*, 968. [CrossRef]
68. Kelty-Stephen, D.G.; Palatinus, K.; Saltzman, E.; Dixon, J.A. A tutorial on multifractality, cascades, and interactivity for empirical time series in ecological science. *Ecol. Psychol.* **2013**, *25*, 1–62. [CrossRef]
69. Ausloos, M.; Ivanova, K. Multifractal nature of stock exchange prices, Comput. *Phys. Commun.* **2002**, *147*, 582–585. [CrossRef]
70. Oświęcimka, P.; Kwapień, J.; Drożdż, S. Multifractality in the stock market: Increments versus waiting times. *Phys. A* **2005**, *347*, 626–638 [CrossRef]
71. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Rak, R. The foreign exchange market: return distributions, multifractality, anomalous multifractality and the Epps effect. *New J. Phys.* **2010**, *12*, 105003. [CrossRef]
72. Grech, D. Alternative measure of multifractal content and its application in finance. *Chaos Solitons Fractals* **2016**, *88*, 183–195. [CrossRef]
73. Zhao, L.; Li, W.; Fenu, A.; Podobnik, B.; Wang, Y.; Stanley, H.E. The q-dependent detrended cross-correlation analysis of stock market. *J. Stat. Mech.* **2018**, *2*, 023402. [CrossRef]
74. Rak, R.; Drożdż, S.; Kwapień, J.; Oświęcimka, P. Detrended cross-correlations between returns, volatility, trading activity, and volume traded for the stock market companies. *EPL* **2015**, *112*, 48001. [CrossRef]
75. Jiang, Z.Q.; Xie, W.J.; Zhou, W.X.; Sornette, D. Multifractal analysis of financial markets: A review. *Rep. Prog. Phys.* **2019**, *82*, 125901. [CrossRef] [PubMed]
76. Wątorek, M.; Drożdż, S.; Oświęcimka, P.; Stanuszek, M. Multifractal cross-correlations between the World Oil and other Financial Markets in 2012–2017. *Energy Econ.* **2019**, *81*, 874–885. [CrossRef]
77. Kwapień, J.; Oświęcimka, P.; Drożdż, S. Components of multifractality in high-frequency stock returns. *Phys. A* **2005**, *350*, 466–474. [CrossRef]
78. Oświęcimka, P.; Drożdż, S.; Kwapień, J.; Górski, A.Z. Effect of detrending on multifractal characteristics. *Acta Phys. Pol. A* **2013**, *123*, 597–603. [CrossRef]
79. Oświęcimka, P.; Kwapień, J.; Drożdż, S. Wavelet versus detrended fluctuation analysis of multifractal structures. *Phys. Rev. E* **2006**, *74*, 016103. [CrossRef] [PubMed]
80. Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Bunde, A.; Havlin, S.; Stanley, H.E. Multifractal detrended fluctuation analysis of nonstationary time series. *Phys. A* **2002**, *316*, 87–114. [CrossRef]
81. Podobnik, B.; Stanley, H.E. Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series. *Phys. Rev. Lett.* **2008**, *100*, 1–4. [CrossRef]
82. Zhou, W.-X. the components of empirical multifractality in financial returns. *EPL* **2009**, *88*, 28004. [CrossRef]
83. Oświęcimka, P.; Drożdż, S.; Forczek, M.; Jadach, S.; Kwapień, J. Detrended cross-correlation analysis consistently extended to multifractality. *Phys. Rev. E* **2014**, *89*, 023305. [CrossRef]
84. Drożdż, S.; Oświęcimka, P. Detecting and interpreting distortions in hierarchical organization of complex time series. *Phys. Rev. E* **2015**, *91*, 030902. [CrossRef]
85. Kwapień, J.; Oświęcimka, P.; Drożdż, S. Detrended fluctuation analysis made flexible to detect range of cross-correlated fluctuations. *Phys. Rev. E* **2015**, *92*, 052815. [CrossRef] [PubMed]
86. Zebende, G.F. DCCA cross-correlation coefficient: Quantifying level of cross-correlation. *Phys. A* **2011**, *390*, 614–618. [CrossRef]
87. Kwapień, J.; Oświęcimka, P.; Forczek, M.; Drożdż, S. Minimum spanning tree filtering of correlations for varying time scales and size of fluctuations. *Phys. Rev. E* **2017**, *95*, 052313. [CrossRef] [PubMed]

88. Drożdż, S.; Kowalski, R.; Oświęcimka, P.; Rak, R.; Gębarowski, R. Dynamical variety of shapes in financial multifractality. *Complexity* **2018**, *2018*, 7015721. [CrossRef]
89. Mnif, E.; Jarboui, A.; Mouakhar, K. How the cryptocurrency market has performed during COVID-19? A multifractal analysis. *Financ. Res. Lett.* **2020**, in press. [CrossRef] [PubMed]
90. Drożdż, S.; Kwapień, J.; Oświęcimka, P.; Rak, R. Quantitative features of multifractal subtleties in time series. *EPL* **2009**, *88*, 60003. [CrossRef]
91. Demir, E.; Bilgin, M.H.; Karabulut, G.; Doker, A.C. The relationship between cryptocurrencies and COVID-19 pandemic. *Eurasian Econ. Rev.* **2020**, *10*, 349–360. [CrossRef]
92. Conlon, T.; Corbet, S.; McGee, R.J. Are cryptocurrencies a safe haven for equity markets? An international perspective from the COVID-19 pandemic. *Res. Int. Bus. Financ.* **2020**, *54*, 101248. [CrossRef]
93. Kristoufek, L. Grandpa, grandpa, tell me the one about Bitcoin being a safe haven: New evidence from the COVID-19 pandemic. *Front. Phys.* **2020**, *8*, 296. [CrossRef]
94. Pearson, K. Note on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* **1895**, *58*, 240–242.
95. Prim, R.C. Shortest connection networks and some generalizations. *Bell Syst. Tech. J.* **1957**, *36*, 1389–1401. [CrossRef]

entropy

MDPI

Article

Information Transfer between Stock Market Sectors: A Comparison between the USA and China

Peng Yue [1], Yaodong Fan [2,3], Jonathan A. Batten [4] and Wei-Xing Zhou [1,5,6,*]

[1] School of Business, East China University of Science and Technology, Shanghai 200237, China; pyue@mail.ecust.edu.cn
[2] School of Business, University of Technology Sydney, Sydney NSW 2007, Australia; yaodong.fan@student.uts.edu.au
[3] College of Business, Shanghai University of Finance and Economics, Shanghai 200433, China
[4] School of Economics, Finance and Banking, College of Business, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia; jabatten@gmail.com
[5] Department of Mathematics, East China University of Science and Technology, Shanghai 200237, China
[6] Research Center for Econophysics, East China University of Science and Technology, Shanghai 200237, China
* Correspondence: wxzhou@ecust.edu.cn

Received: 13 January 2020; Accepted: 5 February 2020; Published: 7 February 2020

Abstract: Information diffusion within financial markets plays a crucial role in the process of price formation and the propagation of sentiment and risk. We perform a comparative analysis of information transfer between industry sectors of the Chinese and the USA stock markets, using daily sector indices for the period from 2000 to 2017. The information flow from one sector to another is measured by the transfer entropy of the daily returns of the two sector indices. We find that the most active sector in information exchange (i.e., the largest total information inflow and outflow) is the *non-bank financial* sector in the Chinese market and the *technology* sector in the USA market. This is consistent with the role of the non-bank sector in corporate financing in China and the impact of technological innovation in the USA. In each market, the most active sector is also the largest information sink that has the largest information inflow (i.e., inflow minus outflow). In contrast, we identify that the main information source is the *bank* sector in the Chinese market and the *energy* sector in the USA market. In the case of China, this is due to the importance of net bank lending as a signal of corporate activity and the role of energy pricing in affecting corporate profitability. There are sectors such as the *real estate* sector that could be an information sink in one market but an information source in the other, showing the complex behavior of different markets. Overall, these findings show that stock markets are more synchronized, or ordered, during periods of turmoil than during periods of stability.

Keywords: information transfer; transfer entropy; stock markets; econophysics

1. Introduction

Complex systems, such as financial markets, are usually composed of many subsystems; in the case of financial markets, information flows and interactions within the market itself are rarely investigated even though they are critical to driving the complex dynamics of the complex system as a whole. Many methods have been proposed to unveil these different relationships among subsystems, such as correlations including simple correlation analysis [1,2], Granger causality [3], nonparametric approaches such as the thermal optimal path method [4–6], and mutual information analysis [7–9]. These different approaches have their own advantages and limitations. Importantly, while Granger causality is commonly used to identify time-varying single or bidirectional causality in economics, it is sensitive to sample period selection and complexity in the underlying time series, as well as having other issues [10,11].

In this paper, we use an alternative approach termed transfer entropy to identify the information transfers between industrial sectors in the world's two largest economies: the USA and China. Transfer entropy, as a kind of log-likelihood ratio [12], is a measure that quantifies information flow based on the probability density function (PDF). Better than correlations or Granger causality, transfer entropy not only identifies the direction of the information flow but also quantifies the flows between different subsystems. In other words, it is capable of quantifying the strength and direction of the interaction between different subsystems at the same time. This approach has found wide application [13–21]. Furthermore, variation and extensions of transfer entropy have been developed that are suitable for different situations [22], such as symbolic transfer entropy [23].

There are many studies adopting the concept of transfer entropy to economic systems such as financial time series [18,24,25], stock market indices [26,27], composite index and the constituent stocks [28,29], and indices of industry sectors of a stock market [30].

Stock price fluctuations reflect both global and local news as well as news within a subsystem. There are also well-known calendar anomalies related to business cycle and market participants sector rotations [31]. In a related work, Oh et al. investigated the information flows among different sectors of the Korean stock market [30]. They measured the amount of information flow and the degree of information flow asymmetry between industry sectors around the subprime crisis and identified the insurance sector as the key information source after the crisis. Although the authors do not attribute a economic basis for this finding, it is likely linked to the insurance sector acting as a leading indicator of risk in the economy. In this work, their analysis is extended and a comparative study is performed on the information transfer among different industry sectors of the Chinese and the USA stock markets. These two markets are respectively the largest emerging and developed stock markets associated with the two largest economies in the world.

The rest of this paper is organized as follows. Section 2 describes the method for calculating symbolic transfer entropy and the sector indices time series for the Chinese and the USA stock markets. Section 3 presents the empirical results about the information flows between stock market sectors and its relationship with market states. Section 4 concludes this work.

2. Method and Data

2.1. Symbolic Transfer Entropy

Schreiber was the first to use transfer entropy to measure information transfer and detect asymmetry in the interactions among subsystems [13]. He treated a sleeping human's breath rate time series and heart rate time series as two subsystems and found that the information flow from the heart to the breath signal is dominant. To explore the transfer entropy between two time series, there are various approaches in the literature. We need to briefly summarize what the other approaches are and why the symbolic transfer method is used. We use the symbolic transfer entropy introduced by Staniek and Lehnertz [23]. Consider two different daily closing prices time series $\{X_t\}$ and $\{Y_t\}$, $t = 1, 2, \ldots, L$, which have the same length L. Closing prices are used to ensure that prices factor in local market news as well as intramarket news from the various sectors. Transfer entropy $T^S_{y\to x}$ assumes that X_t is influenced by the previous l states of the same variable and by the m previous states of variable Y, for financial markets, only the day before is important [32]. Hence, we use $l = m = 1$ in this study. The procedure to calculate the symbolic transfer entropy $T^S_{y\to x}$ from time series $\{y_t\}$ to $\{x_t\}$ is briefly described in the following five steps:

First, we adopt the log returns $\{x_t\}$ instead of the original price time series $\{X_t\}$ by

$$x_t \equiv \ln(X_t) - \ln(X_{t-1}) \tag{1}$$

where X_t is the closing price of the index on the tth trading day.

Second, the returns are discretized into q nonoverlapping windows of equal length Δ. If there are too many windows, the chance of having particular combinations drops very quickly, making the calculation of probabilities slower and less informative [32]. Hence, it is irrational if q is too large or too small. Marschinski and Kantz consider $q = 2$ and $q = 3$ in their research [24]; Sandoval uses $q = 24$ and $q = 6$ [32]. We aim at finding the optimal q to maximize the transfer entropy difference between two time series meanwhile minimizing the calculation cost. In our comparative investigations, the parameter q varies from 2 to 22 with a moving step of 1. We find that when $q \geq 15$, the difference becomes significantly nonzero. Considering the calculation cost and the strength of transfer entropy, in this work, we use $q = 15$. We obtain the maximum value $x_{\max}$ and minimum value $x_{\min}$ of the time series x_t under investigation. The length of each interval is $\Delta_x = [x_{\max} - x_{\min}]/q$ and the kth interval is $[x_{\min} + (k-1)\Delta_x, x_{\min} + k\Delta_x)$. Similarly, we repeat the procedure for y_t and its $\Delta = \Delta_y$ is usually different from Δ_x.

Third, the log return time series $\hat{x}$ and $\hat{y}$ are described as

$$\hat{x}_t = f(x_t) = k_x \text{ and } \hat{y}_t = f(y_t) = k_y, \quad k_x, k_y = 1, 2, \cdots, q, \tag{2}$$

where $x_t \in [x_{\min} + (k_x - 1)\Delta_x, x_{\min} + k_x\Delta_x)$ and $y_t \in [y_{\min} + (k_y - 1)\Delta_y, y_{\min} + k_y\Delta_y)$.

Fourth, the number of elements in the qth interval are denoted by $\hat{x}_t^q$ and $\hat{y}_t^q$, respectively, and then calculate the probabilities $p(\hat{x}_t) = \hat{x}_t^q/(L-1)$ and $p(\hat{y}_t) = \hat{y}_t^q/(L-1)$ and the joint probabilities $p(\hat{x}_t, \hat{y}_t)$, $p(\hat{x}_t, \hat{x}_{t+1})$ and $p(\hat{x}_{t+1}, \hat{x}_t, \hat{y}_t)$.

Fifth, in information theory, different bases of entropy lead to different units of entropy. Base 2 is the most widely applied in transfer entropy for most of empirical works. Therefore, in our study, we use Base 2 to calculate transfer entropy. The symbolic transfer entropy from time series $\{y_t\}$ to time series $\{x_t\}$ is calculated as

$$T_{y\to x}^S = \sum_{\hat{x}_{t+1}, \hat{x}_t, \hat{y}_t} p(\hat{x}_{t+1}, \hat{x}_t, \hat{y}_t) \log_2 \frac{p(\hat{x}_{t+1}|\hat{x}_t, \hat{y}_t)}{p(\hat{x}_{t+1}|\hat{x}_t)}, \tag{3}$$

where the joint probability $p(\hat{x}_{t+1}, \hat{x}_t, \hat{y}_t)$ means the probability that the combination of $\hat{x}_{t+1}$, $\hat{x}_t$ and $\hat{y}_t$ occurs, while $p(\hat{x}_{t+1}|\hat{x}_t, \hat{y}_t)$ and $p(\hat{x}_{t+1}|\hat{x}_t)$ are the conditional probabilities that $\hat{x}_{t+1}$ has a particular value when the values of previous samples $\hat{x}_t$ and $\hat{y}_t$ are known and $\hat{x}_t$ is known, respectively. Since

$$p(\hat{x}_{t+1}|\hat{x}_t, \hat{y}_t) = \frac{p(\hat{x}_{t+1}, \hat{x}_t, \hat{y}_t)}{p(\hat{x}_t, \hat{y}_t)} \text{ and } p(\hat{x}_{t+1}|\hat{x}_t) = \frac{p(\hat{x}_{t+1}, \hat{x}_t)}{p(\hat{x}_t)}, \tag{4}$$

we can simplify Equation (3) and obtain

$$T_{y\to x}^S = \sum_{\hat{x}_{t+1}, \hat{x}_t, \hat{y}_t} p(\hat{x}_{t+1}, \hat{x}_t, \hat{y}_t) \log_2 \frac{p(\hat{x}_{t+1}, \hat{x}_t, \hat{y}_t) p(\hat{x}_t)}{p(\hat{x}_{t+1}, \hat{x}_t) p(\hat{x}_t, \hat{y}_t)}. \tag{5}$$

This expression is used for the estimation of the symbolic transfer entropy.

2.2. Data Description

To conduct our analysis, we selected two sets of data from two major stock markets: the Chinese and the USA stock market. The Chinese stock market is the largest emerging market, while the US stock market is the worlds largest developed stock market.

For the Chinese stock market, we retrieved and analyzed the SWS sector indices issued by Shenyin & Wanguo Securities Co., Ltd. (http://www.swsresearch.com). In total, this gave 28 sector indices of the Chinese stock market, and covered 3508 individual stocks. For each sector index series, there were 4359 daily prices from 4 January 2000 to 29 December 2017. The various sectors with their corresponding six-digit codes include: agriculture and forestry (801010), mining (801020), chemical (801030), steel (801040), non-ferrous metals (801050), electronic (801080), household appliances (801110),

food and drink (801120), textile and apparel (801130), light manufacturing (801140), biotechnology (801150), utilities (801160), transportation (801170), real estate (801180), commercial trade (801200), leisure and services (801210), composite (801230), building materials (801710), building and decoration (801720), electrical equipment (801730), national defense (801740), computer (801750), media (801760), communications (801770), bank (801780), non-bank financial (801790), automobile (801880), and mechanical equipment (801890).

For the US stock market, we chose 16 sector indices composed by Thompson Reuters Co., Ltd. (http://www.thomsonreuters.cn). For each sector index time series, there were 4695 daily prices from 3 January 2000 to 29 December 2017. The differences in day count in the two series are due to differences in holidays in the two countries. The stock sectors are appliances resources (M3L), banking/investment services (BIL), cyclical construction producers (YPL), cyclical consumer services (CRL), energy (E2L), food/beverages (FBL), healthcare services (HSL), industrial and commercial services (IVL), industrial goods (IGL), mineral resources (MRL), pharmaceuticals/MD research (PHL), real estate (REL), retailers (RTL), technology (TEL), transportation (TRL), and utilities (U2$). We use the concept of symbolic transfer entropy to detect and measure the information flows among the return time series of the 28 sector indices of the Chinese stock market and the 16 sector indices of the USA stock market.

3. Results and Discussion

3.1. Symbolic Transfer Entropy and Degree of Asymmetric Information Flow of the Whole Samples

As mentioned in Section 1, symbolic transfer entropy can proxy for the strength and direction of the information flow between two time series. Following Oh et al. [30], we used the degree of asymmetric information flow to measure the information effect between two stock sectors, which is defined as

$$\Delta T^S_{i\to j} = T^S_{i\to j} - T^S_{j\to i}. \tag{6}$$

It follows that $\Delta T^S_{j\to i} = -\Delta T^S_{i\to j}$. We show the calculation results of our datasets in four heat maps (top row for the Chinese sectors and bottom row for the US sectors) of $T^S_{i,j}$ and $\Delta T^S_{i,j}$ in Figure 1, in which each cell shows the value of T^S (left plot) or ΔT^S (right plot) from sector i to sector j. We observe that the values in the diagonal matrices $T^S_{i,j}$ and $\Delta T^S_{i,j}$ are zeros, which is trivial and can be understood from the concept of symbolic transfer entropy. We find that the *non-bank financial* sector (code 790) has roughly the highest T^S values for both inflows and outflows among the Chinese sectors, and the *technology* sector (code TEL) has roughly the highest T^S values for both inflows and outflows among the US sectors; the *non-bank financial* sector comprises three Level 2 sectors in the SWS index system which are *security*, *insurance*, and *multivariate financial*. This suggests that during the sample period from 2000 to 2017, the *non-bank financial* sector and the *technology* sector were respectively the most active in the two stock markets. That is, there was more information exchange between these sectors with the other sectors in their own stock markets than between other sectors in the same stock market.

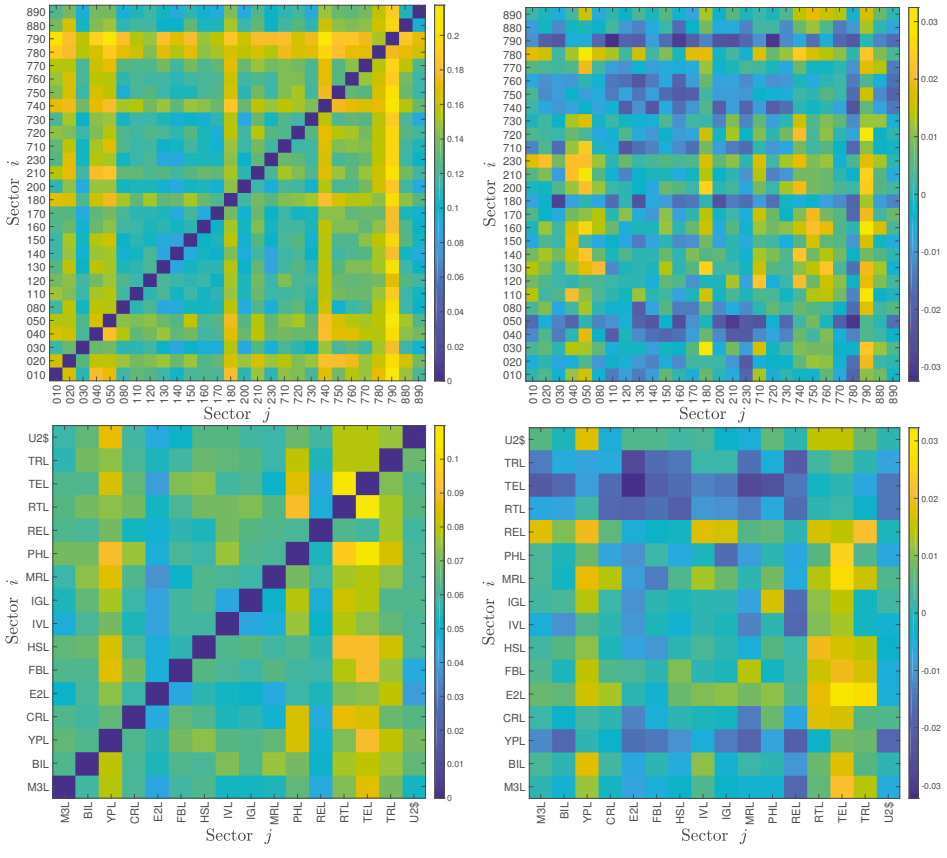


Figure 1. Heat maps of the symbolic transfer entropy matrix $T^S_{i,j}$ (left matrices) and the degree of asymmetric information flow $\Delta T^S_{i,j}$ (right matrices) between the 28 Chinese stock market sectors (2000-2017, top matrices) and the 16 USA stock market sectors (2000-2017, bottom matrices). To simplify the label, we use the last three digits of each 6-digit code to represent the corresponding Chinese stock market sector.

3.2. Average Inflow and Outflow

For each sector i, the average outflow $F_{\text{out},i}$ and inflow $F_{\text{in},i}$ of information can be calculated as follows [30]:

$$F_{\text{out},i} = \frac{1}{n-1} \sum_{p \neq i} T^S_{i \to p} \tag{7a}$$

and

$$F_{\text{in},i} = \frac{1}{n-1} \sum_{p \neq i} T^S_{p \to i}, \tag{7b}$$

where the points with $i = j$ are not included. Figure 2a,c show the bar charts of the average information inflows and outflows for all the sectors. This figure confirms that the *non-bank financial* sector (code 790) and the *technology* sector (code TEL) were the most active sectors in information exchange, respectively. We also find that the more information a sector sends out to other sectors, the more information it receives from others generally. Therefore, the outflow and inflow are positively related to each other. We present in Figure 2b,d the scatter plot of $F_{\text{out},i}$ against $F_{\text{in},i}$, which confirms a significant positive

correlation. The least-squares regression results in the following linear relationship for the Chinese market are as follows:

$$F_{\text{out},i} = 0.724 F_{\text{in},i} + 0.037, \tag{8a}$$

where the p-values of the two coefficients are respectively 3×10^{-15} and 2×10^{-6} and the adjusted R^2 is 0.908. Similarly, for the USA market we have

$$F_{\text{out},i} = 0.291 F_{\text{in},i} + 0.046, \tag{8b}$$

where the p-values of the two coefficients are 6×10^{-4} and 5×10^{-8}, respectively, and the adjusted R^2 is 0.548. It is clear from this simple estimation that the linear relationship is more significant for the Chinese stock market. We argue that the linearity reflects the degree of traders' actions on the idiosyncratic traits of market sectors. The higher linearity of the Chinese stock market implies that the traders in the Chinese market are more irrational, such that their behavior is less reflected in the idiosyncratic traits of market sectors in their decision-making process.

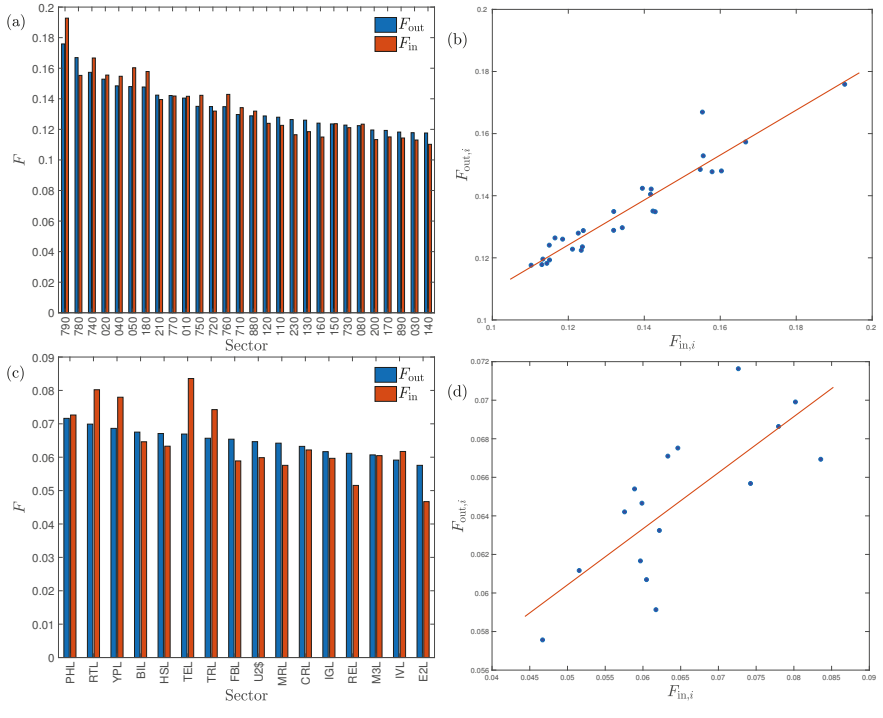


Figure 2. Average outflow $F_{\text{out},i}$ and inflow $F_{\text{in},i}$ of information for stock market sectors. (**a**) Bar chart for the Chinese stock market. (**b**) Scatter plot for the Chinese stock market. (**c**) Bar chart for the USA stock market. (**d**) Scatter plot for the USA stock market.

We also use the average degree of asymmetric information flow ΔF_i to measure the net information of sector i being sent to other sectors, which is defined as follows [30]:

$$\Delta F_i = F_{\text{out},i} - F_{\text{in},i}. \tag{9}$$

We illustrate in Figure 3 the average degree of asymmetric information flow ΔF of the sectors in descending order for the two stock markets.

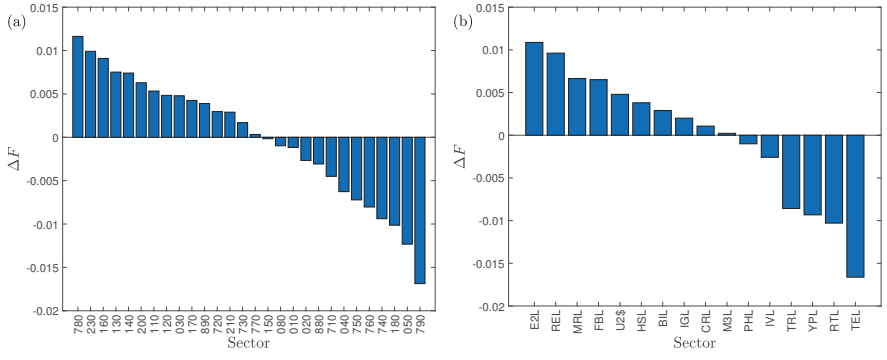


Figure 3. Degree of asymmetric information flow ΔF in descending order of industry sectors. (**a**) Chinese stock market sectors. (**b**) USA stock market sectors.

Among all the 28 Chinese sectors, the *bank* sector (code 780) has the highest ΔF value, while the ΔF value of the *non-bank financial* sector is the lowest. This finding suggests that the *bank* sector has the highest net outflow of information and is thus the most influential sector, while the *non-bank financial* sector is the most influenced sector. If we regard the Chinese stock market as an information transfer system, the *bank* sector is a big information source, influencing other sectors, while the *non-bank financial* sector is a big information sink, influenced by other sectors. Concerning the absolute ΔF value, we find that the *biotechnology* (code 150) is the closest one to zero, which indicates that the strength of information outflows and inflows are approximately equal and there is little net information transferred between the *biotechnology* sector and the whole market.

Among all the 16 US sectors, the *energy* sector (code E2L) has the highest ΔF value, while the ΔF value of the *technology* sector (code TEL) is the lowest. This suggests that the *energy* sector has the highest net outflow of information and is thus the most influential sector, while the *technology* sector is the most influenced sector. Therefore, the *energy* sector is a big information source, influencing other sectors, while the *technology* sector is a big information sink, influenced by other sectors. When we consider the absolute ΔF value, we find that the *appliances* sector (code M3L) is the closest one to zero, which indicates that the strength of information outflows and inflows are approximately equal and there is little net information transferred between the *appliances* sector and the whole market.

Although the sectors in both markets are similar, they play different roles in the two information transfer processes. For instance, the *real estate* sector is an information sink in the Chinese market but an information source in the US market. These results highlight the importance of the *real estate* sector in driving economic output in China and its less significant role in the US.

3.3. Yearly Evolution of Symbolic Transfer Entropy and Degree of Asymmetric Information Flow

Economic sectoral relationships are known to be unstable and change over time. For example, Bernanke (2016) highlighted the changing correlation between the energy and industrial sectors in the US over the last decade. To qualify the evolution of information flows over time, we calculated the symbolic transfer entropy matrix $T^S(t)$ and the asymmetric average information flow $\Delta T^S(t)$ for each year t. The four $T^S(t)$ heat maps of the Chinese stock market for years 2000, 2003, 2007, and 2011 are illustrated in Figure 4 , and the four $T^S(t)$ heat maps of the US stock market for years 2000, 2003, 2007, and 2011 are illustrated in Figure 5, respectively. For the Chinese stock market, it is found that the heat maps share some pattern of similarity. For instance, some relative bright lines emerge vertically and horizontally, echoing the pattern in Figure 4. However, these heat maps also exhibit remarkable differences. The most significant feature is that the heat maps become brighter over time, indicating that there are more information transfers among different sectors with the development of the stock market. The corresponding four heat maps of the asymmetric information flow $\Delta T^S(t)$ are shown

in Figure 6. A similar evolution of patterns is observed in the US stock markets, which is shown in Figures 5 and 7. However, we do not observe a monotonic increase in information flows among the US sectors, in which the information flows among sectors were smaller in 2011.

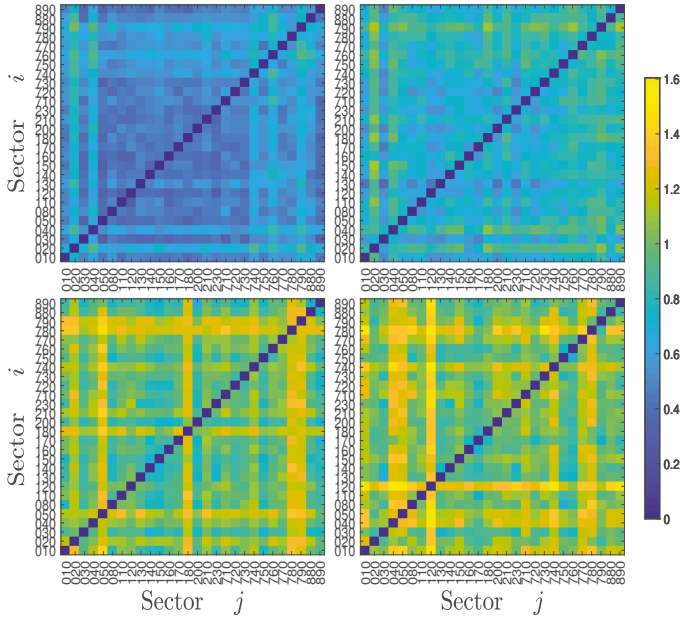


Figure 4. Heat maps of the symbolic transfer entropy matrix $T^S_{i,j}$ for four years (2000, 2003, 2007, and 2011) of the Chinese stock market.

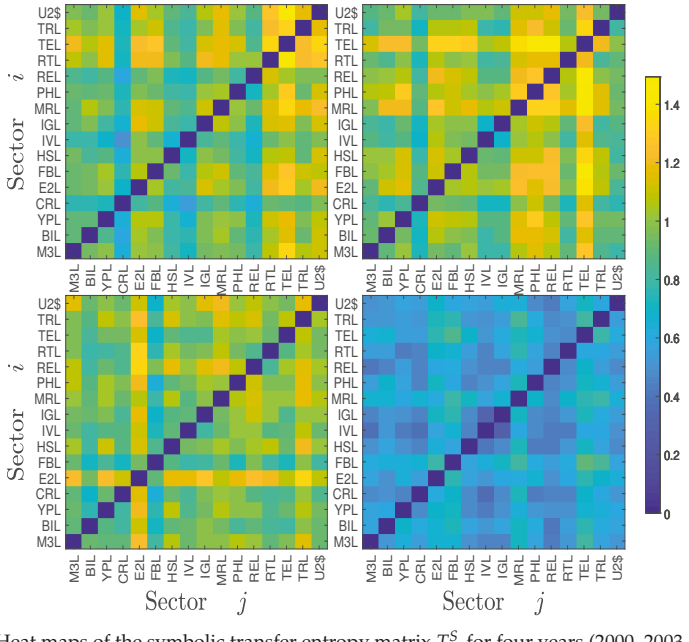


Figure 5. Heat maps of the symbolic transfer entropy matrix $T^S_{i,j}$ for four years (2000, 2003, 2007, and 2011) of the USA stock market.

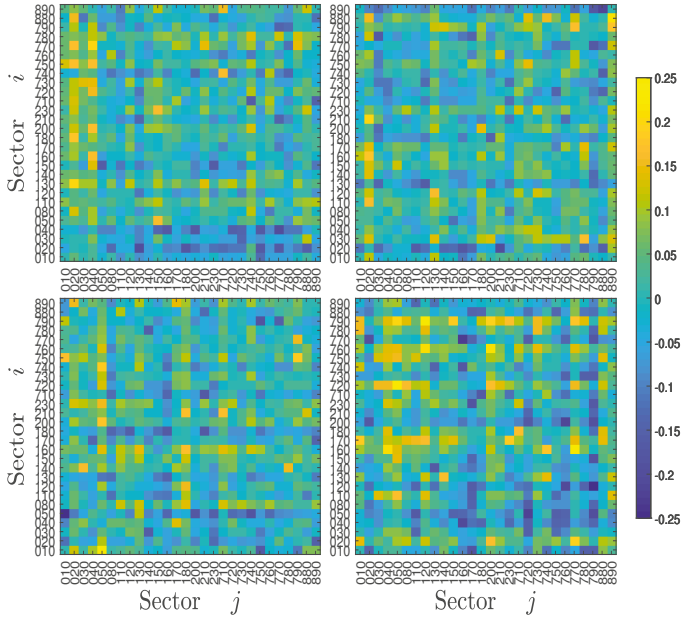


Figure 6. Heat maps of the symbolic transfer entropy matrix $\Delta T^S_{i,j}$ for four years (2000, 2003, 2007, and 2011) of the Chinese stock market.

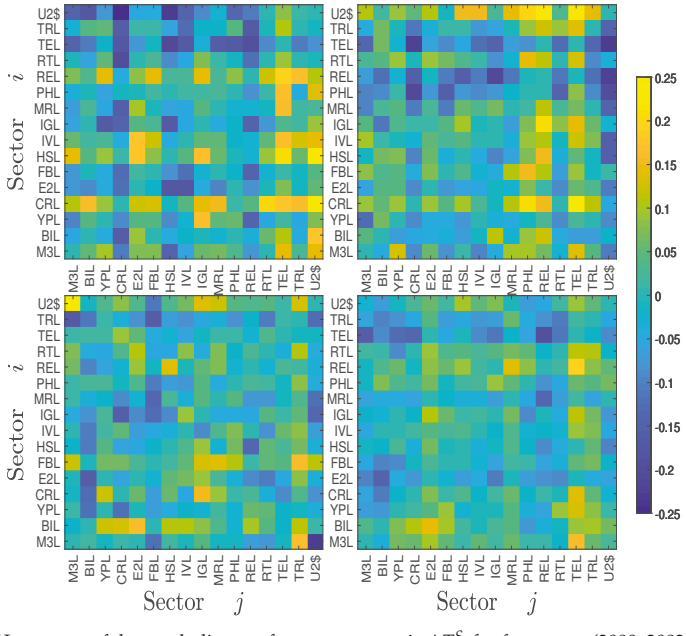


Figure 7. Heat maps of the symbolic transfer entropy matrix $\Delta T^S_{i,j}$ for four years (2000, 2003, 2007, and 2011) of the USA stock market.

To further quantify the evolution of information flows, we calculated the average of the symbolic transfer entropy matrix $T^S(t)$ for each year t as follows:

$$\left\langle T^S(t) \right\rangle = \frac{1}{n(n-1)} \sum_{i \neq j} T^S_{i \to j}(t), \tag{10a}$$

where the diagonal with $i = j$ is not included, and the average asymmetric information flow $\Delta T^S(t)$ for each year t is measured as follows:

$$\left\langle \Delta T^S(t) \right\rangle = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{i} \left| \Delta T^S_{i \to j}(t) \right|, \tag{10b}$$

where the lower triangle (i.e., the part with $i \leq j$) is not included. We note that there are no objective criteria to determine the window size. Too long windows will result in too few data points and vague evolution paths, while too short windows lead to less statistics and more noise [33]. The choice of one year is a trade-off.

The evolutionary trajectories of the average symbolic transfer entropy $\langle T^S(t) \rangle$ and the average asymmetric information flow $\langle \Delta T^S(t) \rangle$ from 2000 to 2017 of the Chinese stock market are presented in Figure 8a,b, respectively, while the results for the US stock market are presented in Figure 8c,d. For the Chinese stock market, we observe two local minima around 2001 and 2016 for $\langle T^S(t) \rangle$ and three local minima around 2001, 2008, and 2016 for $\langle \Delta T^S(t) \rangle$. This observation is of particular interest, because the three periods correspond to key periods of market volatility associated with the market crashes in June 2001 [34], December 2007 [35], June 2009 [35], June 2015 [36], and January 2006 [37]. For the US stock market, we observe four local minima around 2001, 2008, 2011, and 2016 for $\langle T^S(t) \rangle$ and three local minima around 2001, 2011, and 2015 for $\langle \Delta T^S(t) \rangle$, which correspond to the 9/11 terrorist attack in 2001 [38], the subprime mortgage crisis in 2007 [39], the July–August 2011 stock market crash [40], and the 2015–16 stock market selloff beginning in the United States on 18 August 2015. It is documented for other types of networks that the structure of networks usually changes around large market movements (see [41] and the references therein).

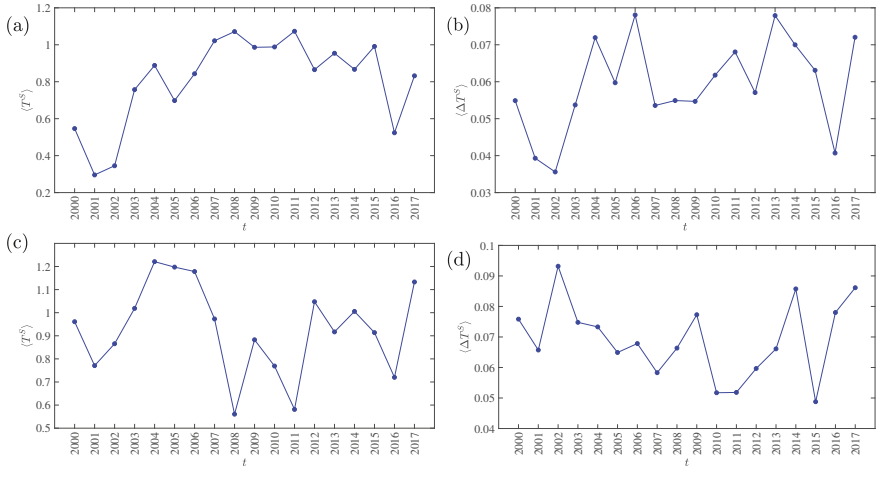


Figure 8. (**a**) Evolution of the average symbolic transfer entropy $\langle T^S(t) \rangle$ from 2000 to 2017 of the Chinese stock market. (**b**) Evolution of the average asymmetric information flow $\langle \Delta T^S(t) \rangle$ from 2000 to 2017 of the Chinese stock market. (**c**) Evolution of the average symbolic transfer entropy $\langle T^S(t) \rangle$ from 2000 to 2017 of the USA stock market. (**d**) Evolution of the average asymmetric information flow $\langle \Delta T^S(t) \rangle$ from 2000 to 2017 of the USA stock market.

We conclude that, during market turmoil periods, both the average information transfer and the average asymmetric information flow are lower than in stable states. This conclusion is not surprising. During bubbles and antibubbles, investors exhibit stronger convergence in decision making. The majority of investors buy stocks during bubbles and sell stocks during antibubbles. Although stock markets have higher volatility during periods of turmoil, investors' actions are more synchronized. In other words, stock markets are more integrated during periods of turmoil than during periods of stability.

4. Conclusions

In this work, we compared the information transfer between industry sectors in the Chinese and US stock markets based on their symbolic transfer entropy. We used daily returns of key sector indices from 2000 to 2017. The results of this work offer several important insights into information flows between industry sectors. First, we find that the most active sector in information exchange is the *non-bank financial* sector in the Chinese market and the *technology* sector in the US market. Second, concerning the net information flow of individual sectors, we find that the main information source is the *bank* sector in the Chinese market and the *energy* sector in the US market, while the information sink is the *non-bank financial* in the Chinese market and the *technology* sector in the US market. The two information sinks with the largest net information inflow in the two markets are exactly the two most active sectors with the largest information transfer. Third, the same sector may play different roles in the two markets. For example, the *real estate* sector is an information sink in the Chinese market but an information source in the US market. Thus, the US stock market is expected to react to demand related to news originating from the housing sector, such as building approvals, whereas in China this is not the case since the markets are driven by supply side factors such as changes in bank lending.

We also investigated the evolution of the yearly information transfer for both markets. It is found that the local minima of the average symbolic transfer entropy $\langle T^S(t)\rangle$ and the average asymmetric information flow $\langle \Delta T^S(t)\rangle$ correspond to periods of market turmoil. We argue that stock markets are more integrated during periods of turmoil than in stable periods, which results in smaller entropy.

Note that while there have been several studies that use entropy-based techniques to predict market fluctuations and crashes [42–47] or measures [48–50], in this study we argue that the average symbolic transfer entropy $\langle T^S(t)\rangle$ and the average asymmetric information flow $\langle \Delta T^S(t)\rangle$ do not have a direct predictive power for market crashes. Further research is required to better understand the dynamics of market crashes, which are likely not driven by historical correlations but rather by behavioral factors.

Author Contributions: Data curation, P.Y.; Formal analysis, P.Y.; Investigation, P.Y., Y.F., J.A.B. and W.-X.Z.; Methodology, P.Y.; Supervision, W.-X.Z.; Writing—original draft, P.Y.; Writing—review & editing, Y.F., J.A.B. and W.-X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by the National Natural Science Foundation of China (U1811462, 71532009), the Shanghai Philosophy and Social Science Fund Project (2017BJB006), the Program of Shanghai Young Top-notch Talent (2018), and the Fundamental Research Funds for the Central Universities.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mantegna, R.N.; Stanley, H.E. *An Introduction to Econophysics: Correlations and Complexity in Finance*; Cambridge University Press: Cambridge, UK, 2000.
2. Zhang, Z.C.; Zhang, Y.J.; Shen, D.H.; Zhang, W. The dynamic cross-correlations between mass media news, new media news, and stock returns. *Complexity* **2018**, *2018*, 7619494. [CrossRef]
3. Shan, J.; Pappas, N. The relative impacts of Japanese and US interest rates on local interest rates in Australia and Singapore: A Granger causality test. *Appl. Financ. Econ.* **2000**, *10*, 291–298. [CrossRef]
4. Sornette, D.; Zhou, W.X. Non-parametric determination of real-time lag structure between two time series: The "optimal thermal causal path" method. *Quant. Financ.* **2005**, *5*, 577–591. [CrossRef]

5. Meng, H.; Xu, H.C.; Zhou, W.X.; Sornette, D. Symmetric thermal optimal path and time-dependent lead-lag relationship: Novel statistical tests and application to UK and US real-estate and monetary policies. *Quant. Financ.* **2017**, *17*, 959–977. [CrossRef]
6. Xu, H.C.; Zhou, W.X.; Sornette, D. Time-dependent lead-lag relationship between the onshore and offshore Renminbi exchange rates. *J. Int. Financ. Mark. Inst. Money* **2017**, *49*, 173–183. [CrossRef]
7. Dionisio, A.; Menezes, R.; Mendes, D.A. Mutual information: a measure of dependency for nonlinear time series. *Physica A* **2004**, *344*, 326–329. [CrossRef]
8. Abigail, J. A complex network model for seismicity based on mutual information. *Physica A* **2013**, *392*, 2498–2506. [CrossRef]
9. Fiedor, P. Networks in financial markets based on the mutual information rate. *Phys. Rev. E* **2014**, *89*, 052801. [CrossRef]
10. Ghysels, E.; Hill, J.B.; Motegi, K. Testing for Granger causality with mixed frequency data. *J. Econom.* **2016**, *192*, 207–230. [CrossRef]
11. Gotz, T.B.; Hecq, A.; Smeekes, S. Testing for Granger causality in large mixed-frequency VARs. *J. Econom.* **2016**, *193*, 418–432. [CrossRef]
12. Barnett, L.; Bossomaier, T. Transfer entropy as a log-likelihood ratio. *Phys. Rev. Lett.* **2012**, *109*, 138105. [CrossRef] [PubMed]
13. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **2000**, *85*, 461–464. [CrossRef] [PubMed]
14. Ai, X.B. Inferring a drive-response network from time series of topological measures in complex networks with transfer entropy. *Entropy* **2014**, *16*, 5753–5776. [CrossRef]
15. Hu, Y.Z.; Zhao, H.Y.; Ai, X.B. Inferring weighted directed association networks from multivariate time series with the small-shuffle symbolic transfer entropy spectrum method. *Entropy* **2016**, *18*, 328. [CrossRef]
16. Yook, S.H.; Chae, H.; Kim, J.; Kim, Y. Finding modules and hierarchy in weighted financial network using transfer entropy. *Physica A* **2016**, *447*, 493–501. [CrossRef]
17. Borge-Holthoefer, J.; Peera, N.; Goncalves, B.; Gonzalez-Bailon, S.; Arenas, A.; Moreno, Y.; Vespignani, A. The dynamics of information-driven coordination phenomena: A transfer entropy analysis. *Sci. Adv.* **2016**, *2*, e1501158. [CrossRef]
18. Zhang, N.N.; Lin, A.J.; Shang, P.J. Multiscale symbolic phase transfer entropy in financial time series classification. *Fluct. Noise Lett.* **2017**, *16*, 1750019. [CrossRef]
19. Toriumi, F.; Komura, K. Investment index construction from information propagation based on transfer entropy. *Comput. Econ.* **2018**, *51*, 159–172. [CrossRef]
20. Servadio, J.L.; Convertino, M. Optimal information networks: Application for data-driven integrated health in populations. *Sci. Adv.* **2018**, *4*, e1701088. [CrossRef]
21. Zhang, Y.P.; Shang, P.J.; Xiong, H.; Xia, J.A. Multiscale analysis of time irreversibility based on phase-space reconstruction and horizontal visibility graph approach. *Fluct. Noise Lett.* **2018**, *17*, 1850006. [CrossRef]
22. He, J.Y.; Shang, P.J. Comparison of transfer entropy methods for financial time series. *Physica A* **2017**, *482*, 772–785. [CrossRef]
23. Staniek, M.; Lehnertz, K. Symbolic transfer entropy. *Phys. Rev. Lett.* **2008**, *100*, 158101. [CrossRef] [PubMed]
24. Marschinski, R.; Kantz, H. Analysing the information flow between financial time series. *Eur. Phys. J. B* **2002**, *30*, 275–281. [CrossRef]
25. Mao, X.G.; Shang, P.J. Transfer entropy between multivariate time series. *Commun. Nonlinear Sci. Numer. Simul.* **2017**, *47*, 338–347. [CrossRef]
26. Dimpfl, T.; Peter, F.J. Using transfer entropy to measure information flows between financial markets. *Stud. Nonlinear Dyn. Econom.* **2013**, *17*, 85–102. [CrossRef]
27. Kwon, O.; Yang, J.S. Information flow between stock indices. *EPL (Europhys. Lett.)* **2008**, *82*, 68003. [CrossRef]
28. Kwon, O.; Yang, J.S. Information flow between composite stock index and individual stocks. *Physica A* **2008**, *387*, 2851–2856. [CrossRef]
29. Kwon, O.; Oh, G. Asymmetric information flow between market index and individual stocks in several stock markets. *EPL (Europhys. Lett.)* **2012**, *97*, 28007. [CrossRef]
30. Oh, G.; Oh, T.; Kim, H.Y.; Kwon, O. An information flow among industry sectors in the Korean stock market. *J. Korean Phys. Soc.* **2014**, *65*, 2140–2146. [CrossRef]
31. Leibon, G.; Pauls, S.; Rockmore, D.; Savell, R. Topological structures in the equities market network. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20589–20594. [CrossRef]

32. Sandoval, L., Jr. Structure of a global network of financial companies based on transfer entropy. *Entropy* **2014**, *16*, 4443–4482. [CrossRef]
33. Song, D.M.; Tumminello, M.; Zhou, W.X.; Mantegna, R.N. Evolution of worldwide stock markets, correlation structure, and correlation-based graphs. *Phys. Rev. E* **2011**, *84*, 026108. [CrossRef] [PubMed]
34. Zhou, W.X.; Sornette, D. Antibubble and prediction of China's stock market and real-estate. *Physica A* **2004**, *337*, 243–268. [CrossRef]
35. Jiang, Z.Q.; Zhou, W.X.; Sornette, D.; Woodard, R.; Bastiaensen, K.; Cauwels, P. Bubble diagnosis and prediction of the 2005-2007 and 2008-2009 Chinese stock market bubbles. *J. Econ. Behav. Org.* **2010**, *74*, 149–162. [CrossRef]
36. Sornette, D.; Demos, G.; Zhang, Q.; Cauwels, P.; Filimonov, V.; Zhang, Q.Z. Real-time prediction and post-mortem analysis of the Shanghai 2015 stock market bubble and crash. *J. Investig. Strateg.* **2015**, *4*, 77–95. [CrossRef]
37. Wei, L.J.; Zhang, W.; Xiong, X. The mechanism and solution for the liquidity stampede crisis in stock markets. *J. Manag. Sci. China* **2017**, *13*, 1–23. (In Chinese)
38. Charles, A.; Darne, O. Large shocks and the September 11th terrorist attacks on international stock markets. *Econ. Model.* **2006**, *23*, 683–698. [CrossRef]
39. Demyanyk, Y.; Hemert, O.V. Understanding the subprime mortgage crisis. *Rev. Financ. Stud.* **2011**, *24*, 1848–1880. [CrossRef]
40. Jayech, S. The contagion channels of July-August-2011 stock market crash: A DAG-copula based approach. *Eur. J. Oper. Res.* **2016**, *249*, 631–646. [CrossRef]
41. Han, R.Q.; Xie, W.J.; Xiong, X.; Zhang, W.; Zhou, W.X. Market correlation structure changes around the Great Crash: A random matrix theory analysis of the Chinese stock market. *Fluct. Noise Lett.* **2017**, *16*, 1750018. [CrossRef]
42. Maasoumi, E.; Racine, J. Entropy and predictability of stock market returns. *J. Econom.* **2002**, *107*, 291–312. [CrossRef]
43. Eom, C.; Oh, G.; Jung, W.S. Relationship between efficiency and predictability in stock price change. *Physica A* **2008**, *387*, 5511–5517. [CrossRef]
44. Lahmiri, S. Entropy-based technical analysis indicators selection for international stock markets fluctuations prediction using support vector machines. *Fluct. Noise Lett.* **2014**, *13*, 1450013. [CrossRef]
45. Zhou, R.X.; Zhan, Y.; Cai, R.; Tong, G.Q. A mean-variance hybrid-entropy model for portfolio selection with fuzzy returns. *Entropy* **2015**, *17*, 3319–3331. [CrossRef]
46. Zou, Y.C.; Yu, L.A.; He, K.J. Wavelet entropy based analysis and forecasting of crude oil price dynamics. *Entropy* **2015**, *17*, 7167–7184. [CrossRef]
47. Benedetto, F.; Giunta, G.; Mastroeni, L. On the predictability of energy commodity markets by an entropy-based computational method. *Energy Econ.* **2016**, *54*, 302–312. [CrossRef]
48. Hou, Y.F.; Liu, F.Y.; Gao, J.B.; Cheng, C.X.; Song, C.Q. Characterizing complexity changes in Chinese stock markets by permutation entropy. *Entropy* **2017**, *19*, 514. [CrossRef]
49. Gu, R.B. Multiscale Shannon entropy and its application in the stock market. *Physica A* **2017**, *484*, 215–224. [CrossRef]
50. Begušić, S.; Kostanjćar, Z.; Kovač, D.; Stanley, H.E.; Podobnik, B. Information feedback in temporal networks as a predictor of market crashes. *Complexity* **2018**, *2018*, 2834680. [CrossRef]

entropy

MDPI

Article

Unexpected Information Demand and Volatility Clustering of Chinese Stock Returns: Evidence from Baidu Index

Gang Chu [1], Xiao Li [2,*], Dehua Shen [1] and Yongjie Zhang [1]

[1] College of Management and Economics, Tianjin University, Tianjin 300072, China; chugang@tju.edu.cn (G.C.); dhs@tju.edu.cn (D.S.); yjz@tju.edu.cn (Y.Z.)
[2] School of Finance, Nankai University, Tianjin 300350, China
* Correspondence: xiaoli@nankai.edu.cn

Received: 20 November 2019; Accepted: 26 December 2019; Published: 28 December 2019

Abstract: This paper employs the Baidu Index as the novel proxy for unexpected information demand and shows that this novel proxy can explain the volatility clustering of Chinese stock returns. Generally speaking, these findings suggest that investors in China could take advantage of the Baidu Index to obtain information and then improve their investment decision.

Keywords: volatility clustering; Baidu Index; information demand; generalized autoregressive conditional heteroscedasticity model (GARCH); mixture of distribution hypothesis

1. Introduction

Recently, scholars have begun to employ Internet news as the proxy for information flow to explain the volatility clustering of stock returns. For example, Zhang et al. firstly employed the number of news which appeared in Baidu News as the novel proxy for information arrival and showed that this proxy could explain the volatility clustering of the SME PRICE INDEX [1]. Based on the Mixture of Distribution Hypothesis (MDH), Shen et.al. [2] further showed that this novel proxy could also explain the volatility clustering for individual stocks. This Internet news-based proxy has gained increasing popularity and is used in various empirical studies, such as in [3–6], among others. However, the key drawback of this Internet news-based proxy is that media outlets may not play a role in diffusing information, and the observed phenomenon, such as reduced volatility clustering, is driven by investor sentiment or psychological biases [4,7,8]. In this paper, we construct a novel proxy for unexpected information demand based on the search frequency of the Baidu Index, and show that this novel proxy could explain the volatility clustering of stock returns in the Chinese stock market.

In that sense, the contribution of this paper is twofold. Firstly, unlike prevailing studies employing trading volume as the proxy for information flow [9–13], we advocate a novel proxy by calculating the unexpected information demand with the Baidu Index. The rationale to employ the Baidu Index is that: Zhang et al. [14] claims that compared with Google Trends, the Baidu Index provides more scientific, authentic, and objective data, and the search results are given at daily frequency. In particular, our results show that the Baidu Index explains more volatility clustering compared to the studies relying on trading volume as the proxy for information flow, such as those by [10] and [12]. Secondly, we provide stock-level evidence that Internet information could explain volatility clustering by focusing on 40 stocks in the Chinese stock market. To the best of our knowledge, we are the first to employ the Baidu Index to explain the volatility clustering at stock level.

The remainder of this paper is structured as follows: Section 2 describes the data and variables construction; Section 3 gives the research methodology; Section 4 gives the empirical results; and Section 5 presents the conclusions.

2. Data and Variables Construction

We used the daily stock closing price over the whole period of 1 January 2015 to 31 December 2018 from the China Stock Market & Accounting Research Database (CSMAR). The following model was used to calculate the daily return of stocks:

$$Return_{i,t} = \frac{(ClosingPrice_{i,t} - ClosingPrice_{i,t-1})}{ClosingPrice_{i,t-1}} \quad (1)$$

where the $Return_{i,t}$ represents the return of stock i on day t, and the $ClosingPrice_{i,t}$ represents the closing price of stock i on day t.

In this paper, we used keyword search volume data from the Baidu Index instead of Google Trend. The Baidu search engine is the biggest search engine in China, and we collected the search volume time series data from the website (https://index.baidu.com). The abnormal change of the Baidu search volume represents the unexpected information demand. We followed Drake, Roulstone, and Thornock [15] to define the abnormal search volume:

$$\overline{BSVI}_{i,t} = \frac{1}{10}\sum_{k=1}^{10} BSVI_{i,t-7\times k} \quad (2)$$

$$AbSearch_{i,t} = \frac{BSVI_{i,t} - \overline{BSVI}_{i,t}}{\overline{BSVI}_{i,t}} \quad (3)$$

where, $BSVI_{i,t}$ represents the Baidu search volume of stock i on day t. $AbSearch_{i,t}$ represents the abnormal search volume of stock i on day t. We defined $AbSearch_{i,t}$ as that Baidu search volume ($BSVI$) on day t for stock i less the average $BSVI$ for the same stock and weekday over the previous 10 weeks, and divided it by the average $BSVI$ for the same stock and weekday over the previous 10 weeks.

We randomly selected 40 stocks from the whole stock market which have a significant autoregressive conditional heteroscedasticity model (ARCH) effect. Figure 1 illustrates the daily return, the autocorrelation coefficient, and the partial correlation coefficient of SHENZHEN ZHENYE (GROUP) CO., LTD (Shenzhen, China) (000006.SZ). We found that the autocorrelation coefficient and the partial correlation coefficient are significantly different from zero (the value exceeds the confidence level), which show that the return time-series of stock 000006.SZ has a significant correlation.

To examine the ARCH effect in residuals, we used two different tests—The Ljung-Box-Pierce Q squared residual correlation diagram and the ARCH Lagrange Multiplier (LM) test. We used Ljung-Box-Pierce Q to investigate the autocorrelation and partial correlation for the squared residuals of the mean equation. Table 1 reports that the Ljung-Box-Pierce Q test is statistically significant at the 5% level in 5-order, 10-order, 15-order, and 20-order lags for the 40 stocks. It denotes that there is significant autocorrelation for all 40 stocks and indicates a significant ARCH effect in the residuals of the mean equation.

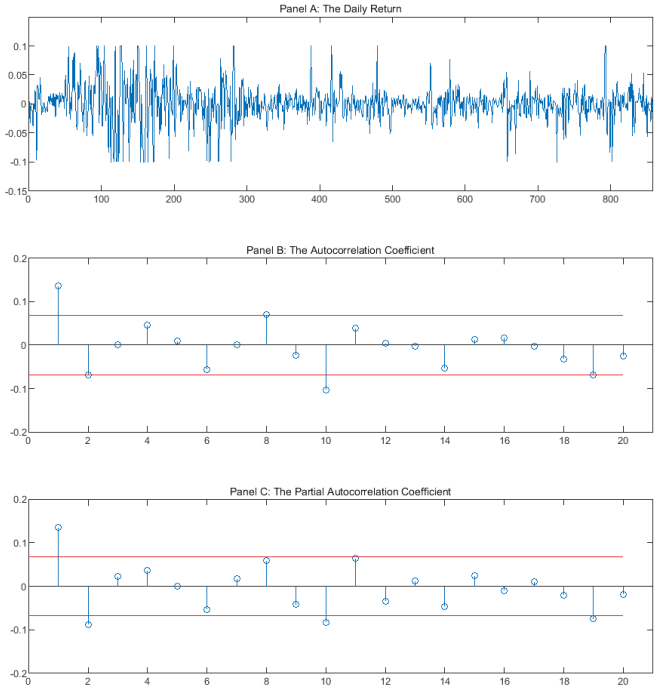


Figure 1. This figure shows the daily return, the autocorrelation coefficient, and the partial correlation coefficient of SHENZHEN ZHENYE(GROUP) CO., LTD (000006.SZ).

Table 1. The results of Ljung-Box-Pierce Q.

Stock Code	5-Lags	10-Lags	15-Lags	20-Lags
000006.SZ	21.6051 ***	38.4482 ***	42.4545 ***	48.2823 ***
000009.SZ	8.0681	20.2608 **	50.1010 ***	57.5135 ***
000012.SZ	5.9939 ***	22.6637 **	30.2325 **	39.3211 ***
000014.SZ	24.7229 ***	39.5542 ***	60.2213 ***	68.8523 ***
000032.SZ	18.5736 ***	35.9589 ***	41.3653 ***	55.0598 ***
000055.SZ	16.1696 ***	25.1449 ***	35.4292 ***	43.4041 ***
000058.SZ	20.4609 ***	31.4845 ***	36.7648 ***	42.5812 ***
000060.SZ	13.4690 **	14.2572	32.2536 ***	33.5518 **
000062.SZ	24.7848 ***	28.1216 ***	36.2232 ***	50.3635 ***
000063.SZ	15.2538 ***	25.4355 ***	28.4957 **	33.0139 **
000070.SZ	30.8077 ***	46.3809 ***	49.2339 ***	52.8709 ***
000088.SZ	22.7924 ***	52.1273 ***	56.3863 ***	65.1081 ***
000417.SZ	5.0400	26.7173 ***	40.2973 ***	43.4603 ***
000151.SZ	18.5001 ***	43.9111 ***	61.21 ***	63.2111 ***
000428.SZ	18.8814 ***	39.0454 ***	54.2291 ***	58.0627 ***
000488.SZ	10.6168 *	21.6755 **	31.5148 **	42.4101 ***
000501.SZ	10.5518 *	43.9884 ***	59.9645 ***	72.0285 ***
000506.SZ	83.1071 ***	96.4226 ***	109.1136 ***	110.5655 ***
000572.SZ	2.5366	25.0778 ***	28.9072 **	38.4491 ***
000544.SZ	16.2461 ***	28.9105 ***	39.0724 ***	52.2869 ***
000545.SZ	15.6275 ***	20.2738 **	30.3052 **	49.6051 ***
000548.SZ	7.1463	28.3473 ***	42.6825 ***	49.8242 ***
000551.SZ	10.6251 *	22.0099 **	29.1647 **	31.4479 **
000552.SZ	9.4039 *	11.0429	39.3816 **	42.3156 ***

Table 1. *Cont.*

Stock Code	5-Lags	10-Lags	15-Lags	20-Lags
000554.SZ	10.2074 *	14.6459	19.2328	21.6216
000559.SZ	11.7471 **	35.2579 ***	40.7503 ***	42.0554 ***
000561.SZ	13.8186 **	18.5989 **	29.6349	39.4647 ***
000565.SZ	16.7915 ***	19.1473 **	30.4736 **	37.2379 **
000570.SZ	16.6161 ***	33.7508 ***	52.1117 ***	58.3116 ***
000576.SZ	61.1017 ***	68.8473 ***	72.7982 ***	89.2181 ***
000597.SZ	6.4010	21.4204 **	31.0142 **	34.1832 **
000600.SZ	9.1283 *	28.6727 ***	37.5731 ***	48.2757 ***
000603.SZ	24.6331 ***	36.1251 ***	47.0366 ***	68.8977 ***
000619.SZ	23.0051 ***	34.4507 ***	47.1044 ***	49.5425 ***
000628.SZ	18.5710 ***	23.5099 ***	27.5565 **	35.9639 ***
000629.SZ	12.6282 **	17.6407 **	20.1427	30.2449 *
000631.SZ	5.3062	21.3837 **	31.8666 **	39.0597 ***
000635.SZ	12.1342 **	22.4461 ***	35.1346 ***	38.7178 ***
000639.SZ	21.8042 ***	41.4061 ***	49.9885 ***	62.4539 ***
000652.SZ	10.0423 *	22.7582 ***	24.5094 *	35.6892 **

Notes: *, ** and *** denotes statistical significance at the 10%, 5%, and 1% levels, respectively.

The ARCH Lagrange Multiplier LM test was calculated by an auxiliary test regression and used to test the heteroscedasticity of the time-series. Table 2 reports that the LM values are statistically significant at the 5% level in 5-order, 10-order, 15-order, and 20-order lags for the 40 stocks. It indicates the existence of an ARCH effect in the residuals sequence. Hence, the Generalized ARCH (GARCH) model is appropriate to use for all the 40 stocks.

Table 2. The results of the GARCH test.

Stock Code	5-Lags	10-Lags	15-Lags	20-lags
000006.SZ	127.3155 ***	149.9881 ***	169.4464 ***	178.2135 ***
000009.SZ	206.7346 ***	219.8304 ***	221.5407 ***	224.5281 ***
000012.SZ	161.4745 ***	179.4592 ***	193.8611 ***	199.8437 ***
000014.SZ	231.9872 ***	238.4752 ***	243.2365 ***	250.4378 ***
000032.SZ	228.9215 ***	239.2971 ***	244.2169 ***	248.8801 ***
000055.SZ	163.0790 ***	183.5448 ***	186.8118 ***	189.9978 ***
000058.SZ	185.5711 ***	209.1581 ***	231.3122 ***	237.3947 ***
000060.SZ	181.5004 ***	189.1091 ***	221.4234 ***	226.1046 ***
000062.SZ	293.3655 ***	318.0177 ***	330.9509 ***	341.3959 ***
000063.SZ	144.8571 ***	156.4018 ***	163.2234 ***	170.6981 ***
000070.SZ	261.1783 ***	248.7214 ***	262.1619 ***	267.1721 ***
000088.SZ	199.2880 ***	219.7765 ***	234.2186 ***	241.0107 ***
000151.SZ	236.1484 ***	254.6980 ***	259.5269 ***	268.1101 ***
000417.SZ	254.2967 ***	275.6469 ***	285.6959 ***	309.9711 ***
000428.SZ	289.8030 ***	297.0388 ***	300.2933 ***	304.8884 ***
000488.SZ	209.8855 ***	227.1245 ***	234.4095 ***	239.7541 ***
000501.SZ	166.3387 ***	208.5034 ***	213.3474 ***	216.3341 ***
000506.SZ	178.4966 ***	186.1122 ***	189.4991 ***	194.8627 ***
000544.SZ	245.3693 ***	253.3165 ***	267.4441 ***	271.1274 ***
000545.SZ	118.8800 ***	131.7289 ***	149.6220 ***	168.4821 ***
000548.SZ	272.6703 ***	296.1373 ***	303.4039 ***	310.8151 ***
000551.SZ	232.2056 ***	241.4726 ***	245.8694 ***	245.7574 ***
000552.SZ	231.5626 ***	241.2432 ***	274.7245 ***	285.2567 ***
000554.SZ	208.7359 ***	223.8271 ***	231.5750 ***	233.5482 ***
000559.SZ	311.8441 ***	321.3317 ***	334.5868 ***	351.0071 ***
000561.SZ	268.0727 ***	277.8340 ***	286.8747 ***	292.8237 ***
000565.SZ	229.6917 ***	242.0975 ***	263.9275 ***	265.7864 ***

Table 2. *Cont.*

Stock Code	5-Lags	10-Lags	15-Lags	20-lags
000572.SZ	179.4836 ***	195.8057 ***	200.9513 ***	205.6801 ***
000570.SZ	272.4294 ***	303.6636 ***	320.8354 ***	323.7563 ***
000576.SZ	254.1562 ***	264.0926 ***	279.6719 ***	281.7115 ***
000597.SZ	248.8782 ***	255.9566 ***	263.5923 ***	271.2491 ***
000600.SZ	214.9866 ***	218.9253 ***	239.0820 ***	242.9311 ***
000603.SZ	221.6153 ***	247.1927 ***	251.0414 ***	260.9005 ***
000619.SZ	338.7260 ***	345.4322 ***	369.5048 ***	369.2384 ***
000628.SZ	239.2112 ***	252.4262 ***	255.7817 ***	258.4453 ***
000629.SZ	51.1104 ***	51.9229 ***	57.4383 ***	65.2544 ***
000631.SZ	206.7708 ***	225.3945 ***	241.8617 ***	250.3248 ***
000635.SZ	135.8456 ***	139.1217 ***	142.8570 ***	142.3168 ***
000639.SZ	182.9437 ***	187.7796 ***	196.4492 ***	206.4381 ***
000652.SZ	176.4028 ***	206.0478 ***	214.2256 ***	213.9942 ***

Notes: *** denotes statistical significance at the 1% level.

The results of the Ljung-Box-Pierce Q and ARCH Lagrange Multiplier (LM) tests show that there is serious heteroscedasticity and autocorrelation on returns of the stock, and the GARCH(1,1) model fits the data well. We used GARCH(1,1) to calculate the daily return volatility. The GARCH(1,1) model is as follows:

$$\varepsilon_t = \sqrt{h_t} v_t \tag{4}$$

$$h_t = \alpha_0 + \beta_1 h_{t-1} + \alpha_1 \varepsilon_{t-1}^2 \tag{5}$$

Table 3 reports the Pearson and Spearman correlation coefficients between daily return volatility and the Baidu search volume. This table suggests that there is positive significant contemporaneous correlation between daily return volatility and Baidu search volume in all 40 stocks. Furthermore, the mean of the Pearson correlation coefficients is 0.6428, and the mean of the Spearman correlation is 0.6398, which denote that these two variables are highly correlated.

Table 3. Contemporaneous correlations between daily return volatility and the logarithm value of Baidu search volume index (BSVI). This table represents the contemporaneous correlation coefficients between daily return volatility and the logBSVI. The daily return volatility was evaluated by GARCH(1,1), and the BSVI was downloaded from Baidu website (http://index.baidu.com/). "Pearson" denotes the Pearson correlation coefficients and "Spearman" denotes the Spearman correlation coefficients.

Stock Code	Pearson	Spearman	Stock Code	Pearson	Spearman
000006.SZ	0.6350 ***	0.6116 ***	000548.SZ	0.6844 ***	0.6774 ***
000009.SZ	0.5438 ***	0.5759 ***	000551.SZ	0.5720 ***	0.5755 ***
000012.SZ	0.5695 ***	0.6016 ***	000552.SZ	0.4721 ***	0.5402 ***
000014.SZ	0.7483 ***	0.7943 ***	000554.SZ	0.6373 ***	0.6269 ***
000032.SZ	0.6620 ***	0.5754 ***	000559.SZ	0.7577 ***	0.6282 ***
000055.SZ	0.5939 ***	0.6426 ***	000561.SZ	0.6496 ***	0.6368 ***
000058.SZ	0.6933 ***	0.6983 ***	000565.SZ	0.7033 ***	0.5856 ***
000060.SZ	0.7179 ***	0.7161 ***	000572.SZ	0.2658 ***	0.2312 ***
000062.SZ	0.7943 ***	0.7083 ***	000570.SZ	0.6686 ***	0.6296 ***
000063.SZ	0.7191 ***	0.7783 ***	000576.SZ	0.8218 ***	0.9019 ***
000070.SZ	0.7676 ***	0.6209 ***	000597.SZ	0.6126 ***	0.6302 ***
000088.SZ	0.5323 ***	0.5504 ***	000600.SZ	0.6494 ***	0.7218 ***
000151.SZ	0.7625 ***	0.6843 ***	000603.SZ	0.6619 ***	0.6735 ***
000417.SZ	0.6989 ***	0.7001 ***	000619.SZ	0.7267 ***	0.7798 ***
000428.SZ	0.6507 ***	0.5203 ***	000628.SZ	0.6563 ***	0.6311 ***

Table 3. *Cont.*

Stock Code	Pearson	Spearman	Stock Code	Pearson	Spearman
000488.SZ	0.5604 ***	0.5268 ***	000629.SZ	0.5891 ***	0.7649 ***
000501.SZ	0.5668 ***	0.5873 ***	000631.SZ	0.7011 ***	0.6506 ***
000506.SZ	0.7013 ***	0.7260 ***	000635.SZ	0.4228 ***	0.3719 ***
000544.SZ	0.6417 ***	0.6558 ***	000639.SZ	0.6920 ***	0.7639 ***
000545.SZ	0.6415 ***	0.6238 ***	000652.SZ	0.5663 ***	0.6709 ***

Notes: *** denotes statistical significance at the 1% levels.

To further consider the relation between the daily return volatility and Baidu search volume, we introduced another direct measure, namely, mutual information. Mutual information is a useful indicator in information theory to measure relative information, and it is widely used to measure the correlation between two different variables. To measure the correlation between two equal length time series $\{x_t\}$ and $\{y_t\}$, $t = 1, 2, 3, \ldots, N$, we computed the mutual information between these two time series, as follows:

$$MI(X, Y) = \int_Y \int_X p(x, y) \log\left(\frac{p(x, y)}{p(x)p(y)}\right) dxdy \quad (6)$$

where $p(x, y)$ is the joint probability density distribution function of X and Y; $p(x)$ is the marginal probability density distribution function of X, and $p(y)$ is the marginal probability density distribution function of Y.

Table 4 represents the mutual information between the daily return volatility and abnormal Baidu search volume. All 40 stocks showed a positive value of mutual information, and the mean of mutual information is 0.7106, which denotes that these two variables are highly correlated. The empirical results clearly support that there is a significant correlation between the Baidu index and daily return volatility.

Table 4. The mutual information between the daily return volatility and abnormal Baidu search volume. This table reports the mutual information between the daily return volatility and abnormal Baidu search volume. The daily return volatility is the GARCH(1,1) volatility of Bollerslev [16], and the abnormal Baidu search volume (*AbSearch*) is calculated by Model 3.

Stock Code	Mutual Information	Stock Code	Mutual Information
000006.SZ	0.7427	000548.SZ	0.6864
000009.SZ	0.7552	000551.SZ	0.6625
000012.SZ	0.5988	000552.SZ	0.7403
000014.SZ	0.6822	000554.SZ	0.7728
000032.SZ	0.7541	000559.SZ	0.7279
000055.SZ	0.6711	000561.SZ	0.6869
000058.SZ	0.7179	000565.SZ	0.6817
000060.SZ	0.7687	000572.SZ	0.8038
000062.SZ	0.7136	000570.SZ	0.7769
000063.SZ	0.7878	000576.SZ	0.6755
000070.SZ	0.6836	000597.SZ	0.6933
000088.SZ	0.6865	000600.SZ	0.6991
000151.SZ	0.7326	000603.SZ	0.7440
000417.SZ	0.7225	000619.SZ	0.7093
000428.SZ	0.7077	000628.SZ	0.6846
000488.SZ	0.7558	000629.SZ	0.6222
000501.SZ	0.7073	000631.SZ	0.7017
000506.SZ	0.7357	000635.SZ	0.7117
000544.SZ	0.6833	000639.SZ	0.6755
000545.SZ	0.6853	000652.SZ	0.6752

3. Methodology

In the time series financial model, the disturbance variance is often found to be less stable. The conditional variance of the error term usually varies with time and relies on the magnitude of the previous errors. In order to solve the heteroscedasticity issue, Bollerslev [16] proposed the generalized autoregressive conditional heteroscedasticity model (GARCH), which is designed to deal with the volatility persistence and describe how the amplitude of return varies over time. In this paper, the GARCH(1,1) model was adopted due to the fact that it has been shown to be suitable to deal with conditional variance that fits many financial time series quite well [16,17]. The GARCH model can be described by the following models:

$$R_t = \mu + \varepsilon_t, \ where \ \varepsilon_t \left| \Omega_{t-1} \right. \sim \left(0, h_t^2\right) \tag{7}$$

$$h_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 \tag{8}$$

where R_t represents the stock return at day t. μ is a constant, ε_t represents the serially uncorrelated errors, and h_t^2 represents the conditional variance of the ε_t. The sum of the coefficients α and β indicates the degree of volatility persistence.

The Baidu search volume index (SVI) is an ideal proxy for information demand because this variable reflects effort by the investor to obtain firm-specific financial information. The abnormal search volume (*AbSearch*) represents investors' demand to search for information. Clark [18] proposed the Mixture of Distributions Hypothesis (MDH), and believes that the price time varying conditional volatility is associated with the information flow. According to the MDH, we made a rational assumption that introducing a proxy of information arrival into the variance model will decrease the observed volatility clustering. Therefore, we proposed an extended model that contains an abnormal Baidu search volume, which can be written as follows:

$$h_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 + \lambda AbSearch_t \tag{9}$$

If the assumption is correct, the volatility persistence, represented by $\alpha + \beta$, should be significantly reduced in comparison with the benchmark model, that is, the original GARCH(1,1) model.

4. Empirical Results

We firstly focus on the estimation results of the benchmark GARCH(1,1) model. In an unreported table, both the coefficients α and β are statistically significant at the 1% level. The sum of the coefficients $\alpha + \beta$ range from 0.998455 to 0.769458 with a mean value of 0.904028. Figure 2 illustrates the residuals, standardized conditional variance, and standardized residuals of SHENZHEN ZHENYE (GROUP) CO., LTD (000006.SZ). We find that the benchmark model fits the volatility dynamic quite well. Table 5 presents the estimation results of the extended model that contains *AbSearch*. All the coefficients α, β, and γ of the extended GARCH(1,1) model are statistically significant at the 1% level. The sum of the coefficients $\alpha + \beta$ range from 0.87521 to 0.489454 with a mean value of 0.698305. Table 6 reports the summarized results for the degree of volatility clustering, indicating that $\alpha + \beta$ is significantly decreased. In particular, we found that after incorporating the proxy for the unexpected information demand, the sum of the coefficients $\alpha + \beta$ dropped significantly with an average of 0.205723. All these findings suggest that the GARCH(1,1) model captures the volatility clustering well, and the unexpected information demand was able to explain the volatility clustering.

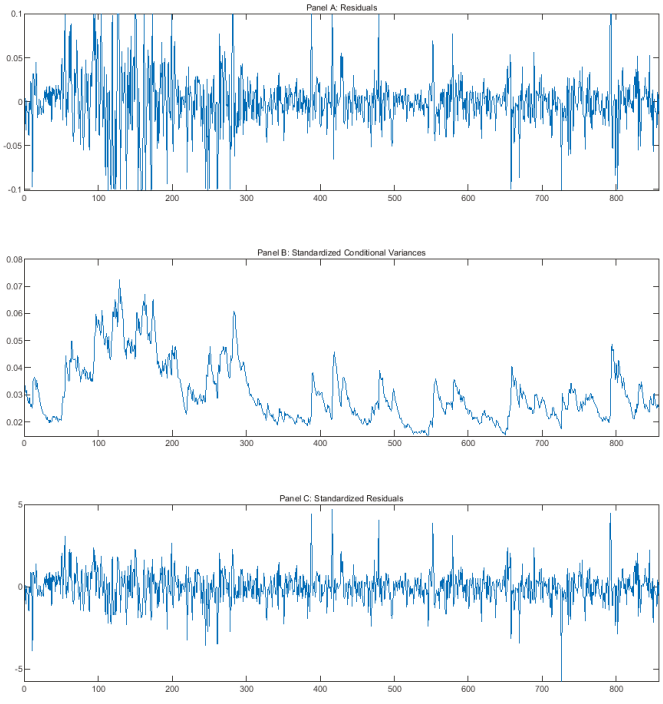


Figure 2. This figure shows the residuals, standardized conditional variance, and standardized residuals of SHENZHEN ZHENYE(GROUP) CO., LTD (000006.SZ).

Table 5. Estimates of extended GARCH (1,1) model.

Stock Code	α	*t*-Value	β	*t*-Value	λ	*t*-Value	$\alpha + \beta$
000006.SZ	0.3275 ***	6.76	0.4355 ***	8.58	1.4196 ***	7.08	0.7630
000009.SZ	0.2766 ***	7.31	0.4708 ***	11.45	2.5797 ***	8.77	0.7474
000012.SZ	0.4678 ***	9.39	0.1253 ***	2.81	0.6176 ***	9.51	0.5931
000014.SZ	0.4026 ***	7.23	0.1878 ***	3.93	1.1363 ***	11.29	0.5904
000032.SZ	0.4138 ***	6.00	0.3026 ***	3.98	1.4328 ***	6.77	0.7164
000055.SZ	0.5191 ***	9.01	0.0861 *	1.68	1.4104 ***	10.15	0.6052
000058.SZ	0.3105 ***	5.85	0.1790 ***	3.64	1.1872 ***	11.96	0.4895
000060.SZ	0.3959 ***	7.21	0.3349 ***	5.33	1.6069 ***	6.91	0.7308
000062.SZ	0.4127 ***	6.84	0.2809 ***	4.84	1.6888 ***	9.97	0.6936
000063.SZ	0.2923 ***	6.17	0.4882 ***	8.99	1.6272 ***	6.45	0.7805
000070.SZ	0.3485 ***	5.9	0.2478 ***	3.78	1.0664 ***	7.99	0.5963
000088.SZ	0.3229 ***	6.94	0.4895 ***	8.71	1.3952 ***	7.88	0.8124
000151.SZ	0.3279 ***	5.66	0.2568 ***	3.70	1.3112 ***	9.61	0.5847
000417.SZ	0.5668 ***	10.32	0.2694 ***	7.34	1.8860 ***	8.27	0.8362
000428.SZ	0.6182 ***	9.23	0.2467 ***	11.19	2.2509 ***	7.32	0.8649
000488.SZ	0.3507 ***	6.10	0.2504 ***	3.42	1.5869 ***	6.08	0.6011
000501.SZ	0.2791 ***	6.56	0.4980 ***	10.03	2.5483 ***	5.70	0.7771
000506.SZ	0.4896 ***	6.46	0.3344 ***	7.39	0.9747 ***	7.12	0.8240
000544.SZ	0.5815 ***	10.42	0.2937 ***	7.42	1.5176 ***	8.42	0.8752
000545.SZ	0.2750 ***	5.24	0.3463 ***	4.28	0.7839 ***	7.54	0.6213
000548.SZ	0.4858 ***	8.63	0.2981 ***	5.91	1.0613 ***	9.50	0.7839

Table 5. *Cont.*

Stock Code	α	*t*-Value	β	*t*-Value	λ	*t*-Value	α + β
000551.SZ	0.3274 ***	6.91	0.5260 ***	6.87	0.8253 ***	4.12	0.8534
000552.SZ	0.4378 ***	9.33	0.2068 ***	4.04	1.9028 ***	8.63	0.6446
000554.SZ	0.3346 ***	6.32	0.3040 ***	4.55	1.9121 ***	8.52	0.6386
000559.SZ	0.4647 ***	9.44	0.3338 ***	10.7	2.7574 ***	11.15	0.7985
000561.SZ	0.4894 ***	8.23	0.1391 **	2.72	1.2136 ***	8.36	0.6285
000565.SZ	0.4156 ***	6.23	0.2874 ***	4.69	1.0779 ***	8.71	0.7030
000572.SZ	0.2432 ***	5.88	0.2796 ***	3.71	1.0002 ***	10.64	0.5228
000570.SZ	0.4308 ***	8.62	0.4063 ***	8.92	3.4746 ***	7.05	0.8371
000576.SZ	0.7076 ***	10.99	0.1333 ***	5.40	1.1049 ***	16.31	0.8409
000597.SZ	0.4523 ***	7.68	0.1750 **	2.97	1.2753 ***	6.96	0.6273
000600.SZ	0.4189 ***	7.20	0.2645 ***	5.81	1.6152 ***	12.89	0.6834
000603.SZ	0.3137 ***	5.42	0.4218 ***	6.8	1.6865 ***	6.40	0.7355
000619.SZ	0.5090 ***	8.75	0.2387 ***	5.44	2.0084 ***	7.01	0.7477
000628.SZ	0.3317 ***	5.99	0.1746 **	2.72	1.1340 ***	8.64	0.5063
000629.SZ	0.3255 ***	4.43	0.3858 ***	4.42	0.6925 ***	5.59	0.7113
000631.SZ	0.4233 ***	8.51	0.2695 ***	5.93	1.6231 ***	11.47	0.6928
000635.SZ	0.3423 ***	5.70	0.1835 **	2.63	1.1718 ***	9.39	0.5258
000639.SZ	0.3883 ***	6.74	0.2836 ***	5.46	0.8062 ***	6.57	0.6719
000652.SZ	0.4641 ***	8.29	0.2115 ***	4.58	1.0943 ***	10.74	0.6756

Notes: *and *** denotes statistical significance at the 10% and 1% level, respectively.

Table 6. Improvement by the extended model.

α + β	Min	Max	Mean	SD
Basic GARCH (1,1)	0.7695	0.9985	0.9040	0.06697
Extended GARCH (1,1)	0.4895	0.8752	0.6983	0.1051
Basic GARCH (1,1)-Extended GARCH (1,1)	0.06535 ***	0.4309 ***	0.2057 ***	0.08369 ***

Notes: *** denotes statistical significance at the 1% level.

5. Conclusions

This paper employed the Baidu search volume index (*BSVI*) as the novel proxy for unexpected information demand and validates the MDH. *BSVI* represents investors' searching behavior through the channel of Baidu, which is the largest search engine in China. In that sense, *BSVI* is a suitable proxy for the information demand. To test the contemporaneous correlation, we employed the Pearson and Spearman correlation coefficients, as well as the mutual information between BSVI and returns and volatiles. The empirical results based on the GARCH(1,1) model reveal a positive and significant impact of the abnormal Baidu Search volume on the conditional volatility of stock return. Generally speaking, these findings suggest that investors in China could take advantage of the Baidu Index to gather information about the stock market and then improve their financial decision-making. For example, investors could employ the high-frequency news to "nowcast" the return volatility, and thus make the optimal investment decision.

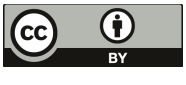
Author Contributions: Conceptualization, D.S.; Formal analysis, Y.Z.; Funding acquisition, D.S. and Y.Z.; Methodology, G.C. and X.L.; Software, G.C.; Writing—Original Draft, X.L.; Writing—Review & Editing, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (71771170, 71801136 and 71701150) and the Young Elite Scientists Sponsorship Program by Tianjin (TJSQNTJ-2017-09).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, Y.; Feng, L.; Jin, X.; Shen, D.; Xiong, X.; Zhang, W. Internet information arrival and volatility of SME PRICE INDEX. *Phys. A* **2013**, *399*, 70–74. [CrossRef]
2. Shen, D.; Zhang, W.; Xiong, X.; Li, X.; Zhang, Y. Trading and non-trading period Internet information flow and intraday return volatility. *Phys. A* **2016**, *451*, 519–524. [CrossRef]
3. Da, Z.; Engelberg, J.; Gao, P. In search of attention. *J. Financ.* **2011**, *66*, 1461–1499. [CrossRef]
4. Fang, J.; Gozgor, G.; Lau, C.-K.M.; Lu, Z. The Impact of Baidu Index Sentiment on the Volatility of China's Stock Markets. Available online: https://www.sciencedirect.com/science/article/abs/pii/S1544612318305609 (accessed on 29 January 2019).
5. Shen, D.; Li, X.; Zhang, W. Baidu news coverage and its impacts on order imbalance and large-size trade of Chinese stocks. *Financ. Res. Lett.* **2017**, *23*, 210–216. [CrossRef]
6. Shen, D.; Li, X.; Zhang, W. Baidu news information flow and return volatility: Evidence for the Sequential Information Arrival Hypothesis. *Econ. Model.* **2018**, *69*, 127–133. [CrossRef]
7. Darrat, A.F.; Zhong, M.; Cheng, L.T.W. Intraday volume and volatility relations with and without public news. *JBF* **2007**, *31*, 2711–2729. [CrossRef]
8. Hirshleifer, D.; Hsu, P.-H.; Li, D. Innovative efficiency and stock returns. *J. Financ. Econom.* **2013**, *107*, 632–654. [CrossRef]
9. Andersen, T.G. Return volatility and trading volume: An information flow interpretation of stochastic volatility. *J. Financ.* **1996**, *51*, 169–204. [CrossRef]
10. Kalev, P.S.; Liu, W.-M.; Pham, P.K.; Jarnecic, E. Public information arrival and volatility of intraday stock returns. *JBF* **2004**, *28*, 1441–1467. [CrossRef]
11. Lamoureux, C.G.; Lastrapes, W.D. Heteroskedasticity in stock return data: Volume versus garch effects. *J. Financ.* **1990**, *45*, 221–229. [CrossRef]
12. Wagner, N.; Marsh, T.A. Surprise volume and heteroskedasticity in equity market returns. *Quant. Financ.* **2005**, *5*, 153–168. [CrossRef]
13. Wang, P.; Zhang, W.; Li, X.; Shen, D. Trading volume and return volatility of Bitcoin market: Evidence for the sequential information arrival hypothesis. *J. Econ. Interact. Coord.* **2019**, *14*, 377–418. [CrossRef]
14. Zhang, W.; Shen, D.; Zhang, Y.; Xiong, X. Open source information, investor attention, and asset pricing. *Econ. Model.* **2013**, *33*, 613–619. [CrossRef]
15. Drake, M.S.; Roulstone, D.T.; Thornock, J.R. Investor information demand: Evidence from Google searches around earnings announcements. *J. Account. Res.* **2012**, *50*, 1001–1040. [CrossRef]
16. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [CrossRef]
17. Wang, R.; Chen, J.J. ARCH effects, trading volume and the information flow interpretation: Empirical evidence from the Chinese stock markets. *J. Chin. Econ. Bus. Stud.* **2012**, *10*, 169–191. [CrossRef]
18. Clark, P.K. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* **1973**, *41*, 135–155. [CrossRef]



entropy

MDPI

Article

Pricing Constraint and the Complexity of IPO Timing in the Stock Market: A Dynamic Game Analysis

Zhiqiang Hu [1], Yuan Hu [1], Yushan Jiang [1] and Zhen Peng [2,*]

[1] Economics and Management School, Wuhan University, Wuhan 430072, China; huzq@whu.edu.cn (Z.H.); yuanhuems@whu.edu.cn (Y.H.); gai.peng@whu.edu.cn (Y.J.)
[2] School of Business, Hubei University, Wuhan 430062, China
* Correspondence: pengzhen@hubu.edu.cn

Received: 9 April 2020; Accepted: 12 May 2020; Published: 13 May 2020

Abstract: The timing of an initial public offering (IPO) is a complex dynamic game in the stock market. Based on a dynamic game model with the real option, this paper investigates the relationship between pricing constraint and the complexity of IPO timing in the stock market, and further discusses its mechanism. The model shows that the IPO pricing constraint reduced the exercise value of the real option of IPO timing, thus restricting the enterprise's independent timing and promoting an earlier listing. The IPO price limit has a stronger effect on high-trait enterprises, such as technology enterprises. Lowering the upper limit of the pricing constraint increases the probability that enterprises are bound by this restriction during IPO. A high discount cost and stock-market volatility are also reasons for early listing. This paper suggests a theoretical explanation for the mechanism of the pricing constraint on IPO timing in the complex market environment, which is an extension of IPO timing theory, itself an interpretation of the IPO behavior of Chinese enterprises. These findings provide new insights in understanding the complexity of IPOs in relation to the Chinese stock market.

Keywords: pricing constraint; IPO timing; dynamic game model; real option; complexity of IPOs

1. Introduction

The discussion of the initial public offering (IPO) timing mechanism is one of the current hot topics in the field of corporate finance and financial market. Existing discussions on the IPO timing of enterprises are mostly derived from the research on IPO anomalies [1], which explained the internal logic of IPO timing from the perspective of enterprises themselves and market volatility [2]. The financial markets are complex systems [3]; however, the timing of an IPO is a complex game, which is influenced by many factors, both at the market level and macro level, and the game in IPO timing constitutes one part of the complexity in the stock market. When discussing the factors of IPO timing at the market level, the existing literature often talks about the role of information asymmetry and investor sentiment. There is little literature on the possible market regulation of IPOs, especially on the effect of price constraints on IPO timing.

Unlike overseas IPO markets, the issuance of new shares in the Chinese stock market is still at an emerging stage, and the corresponding issuance system is being gradually built and improved. In order to maintain the order of the issuing market and alleviate the "three highs" phenomenon (high issue price, high price-to-earnings (P/E) ratio, and high issue financing quota) in the actual issuance of new shares, the China Securities Regulatory Commission launched a series of reforms on the pricing of new shares, including setting the P/E ratio control on the pricing of new shares. In view of this, we studied the pricing constraint in IPO timing, which is a valuable expansion of the IPO timing theory, and complements the lack of attention given to market regulation in the previous literature.

On the other hand, the financing for technology companies is also a hot topic in the field of corporate finance. With the establishment of the China STAR Market (the Sci-Tech innovation board of the Shanghai Stock Exchange (SSESTAR), was established in the Shanghai Stock Market on the November 5, 2018. It is a new board independent of the existing main board market. Its main purpose is to pilot IPO registration reform and provide support for technological innovation of enterprises), many Chinese enterprises will choose to make an IPO in this market in future. Compared with the past, the SSESTAR Market has abolished the regulation of the P/E ratio and relieved the IPO price limit. This paper provides a theoretical foundation for that policy design. Specifically, compared with other enterprises, the pricing constraint has a stronger constraint on the independent IPO timing of technology IPOs. The pricing constraint will encourage them to go public earlier, which may further lead to insufficient financing, and thus weaken the IPO's attraction of the domestic market for enterprises in the long term. Therefore, this research suggested that the release of restrictions on the IPO pricing may help to encourage companies to make an IPO in the Chinese stock market, enhancing the positive impact of capital market financing on the development of China's innovation economy.

In this paper, we studied the influence of the pricing constraint on IPO timing based on a dynamic game model with the real option. The timing of an IPO is a complex multi-factor dynamic game, which is not only influenced by the enterprises themselves and the stock market fluctuations but also market regulation policy. The innovations of this research are as follows: first, it is research on the pricing constraint in IPO timing. The previous literature focused on the analysis of enterprise and market factors but overlooked the regulatory facts. This paper analyzes the influence of market regulation, especially for technology enterprises, and this is an expansion of the IPO timing theory. Second, our model incorporates market control factors into the dynamic mechanism of IPO timing, making the modeling more complex. By the introduction of multiple influencing factors and the real option, the model is much closer to the actual situation of IPO timing in the Chinese market. Third, we investigate the impact of the pricing constraint on technology enterprises in IPO timing. Our results prove that the price constraint can significantly advance the IPO timing of technology enterprises, and the resulting insufficiency of financing may be one of the reasons why Chinese enterprises seek to make an IPO in the overseas market, which provides new insights for understanding the complexity of the IPOs in relation to the Chinese stock market.

A breakdown of the paper is structured as follows: Section 2 is the literature review, where we review those studies related to IPO timing and the modeling of IPOs. Section 3 illustrates the parameters, settings, and important mechanisms of our model. Section 4 contains the derivation of the IPO timing dynamic game model. We first studied the IPO market timing equilibrium with no constraints, and then extended to an analysis of price limit constraints. Section 5 is the conclusion, in which we review the conclusions of this paper and indicate possible research directions for future study.

2. Literature Review

2.1. IPO Timing

The previous literature focused on the following two aspects to explore the motivation and results of IPO timing. On the one hand, the existence of information asymmetry in the stock market may lead to the distortion of IPO pricing, thus leading enterprises to choose the time of high financing to go public, "in light of the time conditions". To increase the amount of listed capital, the enterprises to be listed tend to raise capital by listing when the market values the new shares highly, and delay the execution when the market values the new shares poorly. This theoretical hypothesis is supported by the empirical data.

Ibbotson et al. [4] and Korajczyk [5] found through empirical studies on the data of IPO enterprises in the market that enterprises would choose to go public when they could provide more accurate pricing of new shares, and that there was a positive correlation between IPO earnings and the number of new shares issued in subsequent markets. Lucas and McDonald [6] further believed that the existence

of information asymmetry would result in the IPO price of enterprises being overestimated by the market. Then, the issuing of new shares when the IPO price was overestimated and suspension of the issuance when the IPO price was underestimated, verifying the motivation of enterprises in actively choosing the listing time.

On the other hand, some researchers believed that investor sentiment in the stock market is also a potential factor influencing the IPO timing of enterprises. The discussion in this respect comes from the expansion of the theory of investor sentiment. When the investor sentiment in the stock market keeps rising, it is accompanied by a general overvaluation of stock prices. At this point, the enterprises to be listed will issue IPO shares at a high price by virtue of high investor sentiment [7–9]. In addition, in the new issue, a large number of optimistic investors tend to gather in the primary market, and the pursuit of optimistic investors for a new issue will lead to a herd effect in the market. Therefore, when there are a large number of optimistic investors in the market, the overall optimistic market sentiment will push up the IPO price and underpricing rate of the IPO companies, stimulate the issuance of new shares, and then induce an IPO wave in the industry.

2.2. IPO Game Modeling

For the modeling of the IPO timing mechanism, the theoretical modeling of information asymmetry is the first step. Chemmanur and Fulghieri [10] proposed a two-phase IPO game decision-making model. They suggested that enterprises would waver between an external IPO financing or selling part of the equity. Their model concluded that the longer an enterprise has been operating, the more information it has publicly disclosed to the stock market, and the more likely it is to IPO. Therefore, the subsequent literatures discussed the information asymmetry in IPO from the information spillover effect.

Hoffmann-Burchardi [11] introduced the information spillover effect into the IPO timing model, believing that the existing IPO events can provide market investors with information about the industry and macro economy of the enterprise, thus affecting the IPO pricing of those subsequent IPO enterprises. The model shows that enterprises will choose to delay the IPO timing until the market obtains information from the IPOs of enterprises so as to price the enterprises that are about to go public more accurately. Alti [12], whose model is similar to Hoffmann-Burchardi [11], analyzed the information spillover effect of IPO events from the perspective of cost and benefit. The model analysis suggested that the cost of listing is caused by the existence of informed traders, which leads to adverse pricing in the market. However, the occurrence of IPO events in the past will reduce the degree of information asymmetry in the market and reduce the listing cost of subsequent listed companies.

Colak and Gunay [13] analyzed the macroeconomic factors in the spillover of IPO information. They got two Bayesian–Nash equilibriums for IPO timing. The first hypothesis is that the decision-making is independent, the quality of the enterprises listed in the early stage is often poor, and the enterprises with good quality will wait strategically and go public only after receiving the signal of the improvement of the macro economy. The second is that enterprises' market timing will affect each other, and thus the probability of high-quality enterprises listing after the IPO of low-quality enterprises is higher in this situation.

In the theoretical modeling of IPO timing of investor sentiment, the model of Pastor and Veronesi [14] emphasizes the close relationship between IPO timing and investor sentiment. By introducing changes in investor sentiment towards the market as a whole, they assumed that companies had US-style call options because they could flexibly choose the timing of their listings and could obtain excess returns by exercising the options, at the cost of forgoing the cost of investment if the market sentiment continued to deteriorate. The model indicates that, when the expected market return rate falls and the future profit margin and operational risk of enterprises rise, the probability of enterprises choosing to make an IPO increases.

After the extension of the modeling terms to the multi-period game model, scholars represented by Spielgel and Tookes [15] built more complex multi-period game models of IPO timing and incorporated

the influences of technological innovation and product market competition into the IPO timing of enterprises. The results of the model analysis demonstrate that the time an enterprise chooses to go public is related to the degree of technological innovation, and the market share owned by the enterprise also affects the timing of the IPO.

Compared with the above model research, Aghamollay and Guttmanz [16] attempted to evaluate an enterprise's internal value (that is, the enterprise characteristics), information asymmetry, and investor sentiment, to consider the main body of the enterprise in different periods of the IPO timing dynamic game. Their model provided a relatively complete framework for describing the intrinsic mechanism of IPO timing.

2.3. Literature Comment

There are many studies on the mechanism of IPO timing; however, the mechanism of IPO timing discussed in those literatures is incomplete. Between January 2005 and December 2008, the China Securities Regulatory Commission (CSRC) imposed a P/E ratio of 30 times on domestic IPOs to maintain market stability. Since the resumption of IPO issuance in June 2014, the CSRC has introduced a series of reforms on IPO pricing, including setting the most critical constraint on IPO pricing—the P/E ratio of 23 times. However, there are not many studies on the IPO pricing constraints, and the analysis of IPO pricing constraints and the IPO timing mechanism are relatively rarely discussed. This paper adds to the growing literature on the complexity in the stock market by including the investigation of IPO pricing constraints in the study of the IPO timing mechanism.

3. Model Specification

3.1. Parameters

The trait factor ψ_i, which reflects the business behavior, financial characteristics, technology research, and development of the enterprise, represents the intrinsic value of the enterprise. Before $t = 1$, each company can only observe their own trait factors, which are independent of each other. Companies are risk-neutral individuals and are bound to go public at some point in the period mentioned above. The enterprise discloses its idiosyncratic trait factor ψ_i when it goes public, and this information is true without cost.

The market factor c_t represents the information asymmetry, investor sentiment, and rational expectation adjustment in the stock market. We set the factor c_t to follow the random first order autoregressive process of mean reversion, the AR (1) process [17,18]. This process is not only related to the adjustment of investors to rational expectations [19,20], but also related to investor sentiment and market information asymmetry. The specific settings of the AR (1) process are shown in Equation (1).

$$c_t = \lambda c_{t-1} + \varepsilon_t \tag{1}$$

where c_t is the value of the common factor in the period. c_1 is unknown in $t = 1$, unless there is at least one IPO in the market, then the rest can observe c_1 at the end of this period.

The discounted cost r represents a series of potential costs, such as market share, revenue decline, and debt increase, that would be lost if the company delayed the IPO under other conditions.

Last, enterprises are advance homogeneous; that is, the trait factors of enterprises have the same idiosyncratic distribution. They also face the same discount rate r, and common factor c_t effects in the same way.

All parameters included in this model are described as shown in Table 1.

Table 1. Description of the parameters in the model.

Parameter	Definition of the Parameters
ψ_i	The trait factors of enterprise i, $\psi_i \epsilon (0, +\infty)$, $i \in (1, \mathrm{N})$, $\mathrm{N} \geq 2$.
ψ_t^*	The critical condition of whether the enterprise delays the IPO in period t. t = 1,2.
c_t	The common market factor in period t. $c_t \epsilon (-\infty, +\infty)$, $c_0 = 0$, $c_1 = \widetilde{\varepsilon_1}$. ε_t is the stochastic error term, $\varepsilon_t \sim N\left(0, \sigma^2\right)$.
ω_t	The market environment variable at period t, which is closely related to the common market factor c_t in the same period.
λ	The mean reversion rate of the common market factors, $\lambda \in (0, 1)$.
r	A collection of the discounted costs considered in an IPO. The discount rate stays the same for each period.
t	The decision time point of IPO timing in dynamic game model, t = 0, 1, 2, 3.
$u_{i,t}$	The expected utility of enterprises listing in period t, which reflects the market valuation of those enterprises.
$F(\cdot)$,$f(\cdot)$	The cumulative distribution function and density function of the common market factor c_t.
$G(\cdot)$	The cumulative distribution function of the trait factor ψ_i.
$Pr\left(NI_{i \neq j}^t\right)$	The probability of no IPO in the market at t period.
$Pr\left(I^t\right)/Pr\left(NI^t\right)$	The probability of an IPO for enterprise i at period t, the probability that enterprise i will not IPO.
$NS(\psi_i)$	The non-strategic enterprises. The expected utility of an IPO in a fixed period.
$S(\psi_i)$	The strategic enterprise. The expected utility of strategic IPO.
$V_t(\psi_i)$	The real option of IPO timing. The expected utility difference of an IPO between strategic and non-strategic enterprises.

3.2. Assumptions

There are N (N ≥ 2) companies to be listed. The timing of the IPO is a three-period decision ($t \in$ {0,1,2,3}). The valuation of the IPO depends on the combined effect of three factors: enterprise trait factor ψ_i, the factor representing the characteristics of the industry and even the market volatility c_t and the discounted cost r.

When an enterprise conducts an IPO at any time t, the expected utility obtained from its listing (that is, the valuation level) is determined by Equation (2).

$$u_{i,\eta} = \frac{E\left(\psi_i + c_\eta \middle| \omega_\eta\right)}{(1+r)^{\eta-1}} \tag{2}$$

The IPO timing process in this model is shown below.

t = 0, enterprises can only observe their own characteristics.

t = 1, all the companies decide whether to make an IPO at the same time. If at least one company is listed in t = 1, then the IPO company will obtain its market valuation, and common factor c_1 is known for all companies.

t = 2, all the enterprises not listed in phase t = 1 will decide whether to make an IPO at this time. If at least one company is listed in phase t = 2, the enterprises carrying out the IPO will also obtain their market valuation, and the common factor c_2 will be observed by all enterprises.

t = 3, all companies not listed in t = 1 and t = 2 will be listed in t = 3.

The IPO behavior of enterprises will affect each other, and the IPO events of the previous period in the market will affect the IPO decisions of the next period.

3.3. The Real Option Method

Referring to the research of Myers [21] and Busaba et al. [22] on the real option method in valuation, this paper introduces the options for enterprises to choose IPO timing. The IPO timing option is defined as the right of enterprises to choose IPO by observing the common factors of the previous period.

The model defines the expected utility of an enterprise (hereinafter referred to as a non-strategic enterprise) that does not execute the above strategic IPO timing and always chooses to list in $t = 2$ (hereinafter referred to as a non-strategic enterprise) as $NS(\varphi_i)$. The expected utility of an enterprise (hereinafter referred to as a strategic enterprise) that is not listed at $t = 1$ but at the time of a strategic IPO at $t = 2$ is defined as $S(\varphi_i)$. Therefore, the IPO timing real option $V_t(\varphi_i)$ is the difference between the expected utility of an IPO of strategic enterprises and non-strategic enterprises, as shown in Equations (3)–(5).

$$NS(\psi_i) = E\left[\frac{\psi_i}{1+r}\right] = \frac{\psi_i}{1+r} \tag{3}$$

$$S(\psi_i) = Pr\left(c_2 < c_2^*(\psi_i)\right)E\left[\frac{\psi_i + c_3}{(1+r)^2}\right] + \left(1 - Pr\left(c_2 < c_2^*(\psi_i)\right)\right)E\left[\frac{\psi_i + c_2}{1+r}\right] \tag{4}$$

$$V_t(\psi_i) \equiv S(\psi_i) - NS(\psi_i) = \int_{-\infty}^{c_2^*(\psi_i)} \left(\frac{\psi_i + \lambda^2 c_1}{(1+r)^2} - \frac{\psi_i + \lambda c_1}{1+r}\right) f(c_t) dc_t \tag{5}$$

4. The IPO Timing Model

Compared with the discussion of market factors in the previous literature, this model places emphasis on the analysis of the regulatory factors of the stock market. Specifically, on the basis of Aghamollay and Guttmanz [16], this paper introduces the setting of IPO pricing constraints to explore the impact mechanism of pricing control measures on the enterprise's IPO decision.

Enterprises maximize the expected utility of IPO (valuation). For any enterprise i, the IPO will be conducted at $t = 1$ only when its trait factors ψ_i are higher than the critical trait factors ψ_i^* and the expected utility of the current IPO is higher. Similarly, at $t = 2$, if an existing enterprise j is listed at $t = 1$, the enterprise $i \neq j$ will make an IPO at the stage if, and only if, its characteristic factors are above the critical value ψ_2^*. At the same time, when there is at least one IPO in $t = 1$, the enterprise i may delay the IPO to observe the specific information of the common factors c_1 at the end of $t = 1$. If the common factor is low enough at this time, considering that the volatility of the common factor follows the mean reversion process AR (1), its value at $t = 3$ will be greater than that at $t = 2$. In that case, the enterprise can delay its IPO to obtain the real option income generated by it.

Therefore, enterprise i will weigh the real option obtained by delaying the IPO at $t = 1$ against the discounted costs it faces, to finally determine the IPO period. After this section, the critical equilibrium of the dynamic game of IPO market timing will be solved. First, the equilibrium of the IPO market timing of the enterprise i in $t = 2$ will be discussed, and then the convergence in $t = 1$ will be discussed.

4.1. The Unconstrained Model

According to the setting, if there is no IPO in $t = 1$, the dominant choice of any enterprise i at this time is to make an IPO in $t = 2$.

If there is at least one IPO in stage $t = 1$, for the enterprises i that have not yet had an IPO in stage $t = 1$, if, and only if, their expected utility in a stage $t = 3$ and stage $t = 2$ IPO is equal, the enterprises will remain neutral to IPO in $t = 2$ (strategic delaying); otherwise, the enterprise should make an IPO in the current period of $t = 2$, as shown in Equation (6).

$$\frac{\psi_i + E(c_2|\omega_1)}{1+r} = \frac{\psi_i + E(c_3|\omega_1)}{(1+r)^2} \tag{6}$$

The critical condition of IPO timing at $t = 2$ is shown in Equation (7).

$$\psi_i \geq \psi_2^*(c_1) = -c_1(1 + r - \lambda)\left(\frac{\lambda}{r}\right) \tag{7}$$

Furthermore, if the enterprise i makes an IPO in $t = 1$, its expected benefits are fixed as $\psi_i + E(c_1) = \psi_i$. If the enterprise delays the IPO at $t = 1$, its expected utility is related to the IPO situation in the stock market during the same period.

To be specific: (1) when there is no IPO in the stock market of $t = 1$, the enterprise will be listed in $t = 2$, the expected utility is $\frac{E(\psi_i+c_2)}{1+r} = \frac{\psi_i}{1+r}$. (2) If there is at least one IPO in the stock market of $t = 1$ and the characteristics of the enterprise are higher than the critical condition, $\psi_i > \psi_i^*(c_2)$, the enterprise will make an IPO in $t = 2$, and the expected utility is $\frac{E\left(\psi_{i+}c_2|\psi_i > \psi_i^*(c_2)\right)}{1+r}$. (3) IPOs happened at $t = 3$ to obtain the real option proceeds from delay; in this case, the IPO expected utility is $\frac{E\left(\psi_i+c_3|\psi_i<\psi_i^*(c_2)\right)}{(1+r)^2}$. Thus, the expected utility of enterprise i obtained from the delay of the IPO at $t = 1$ is shown in Equation (8).

$$Pr\left(NI_{i\neq j}^1\right)\left(\frac{\psi_i}{1+r}\right) + \left(1 - Pr\left(NI_{i\neq j}^1\right)\right)\left(\begin{array}{c} Pr\left(I_{i\neq j}^1\right)E\left[payoff\ at\ t=2|\psi_i, I_i^2\right] \\ +Pr\left(NI_{i\neq j}^1\right)E\left[payoff\ at\ t=3|\psi_i, NI_i^2\right] \end{array}\right) \tag{8}$$

Because the trait factor ψ_i is independent of each enterprise, the probability of an IPO is the same for any enterprise i. By introducing the cumulative density function correlation of the above variable, the critical condition of IPO in $t = 1$ is shown in Equation (9), the derivation is shown in Appendix A.

$$\begin{array}{l} \psi_1^* \\ = \left[G\left(\psi_1^*\right)\right]^{N-1}\left(\frac{\psi_1^*}{1+r}\right) \\ +(1 \\ -\left[G\left(\psi_1^*\right)\right]^{N-1}\left(\begin{array}{c} F\left(\frac{\psi_1^*}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)\frac{\psi_1^*}{1+r} + \frac{1}{1+r}\lambda\sigma^2 f\left(-\frac{\psi_1^*}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right) \\ +\left(1 - F\left(\frac{\psi_1^*}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)\right)\frac{\psi_1^*}{1+r} - \frac{1}{(1+r)^2}\lambda^2\sigma^2 f\left(-\frac{\psi_1^*}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right) \end{array}\right) \end{array} \tag{9}$$

Comparing the critical conditions of the IPO timing in $t = 2$ and $t = 1$ without pricing constraints, we can give Proposition 1 as below.

Proposition 1. *The higher the characteristics of an enterprise, the earlier the optimal listing time.*

4.2. IPO Timing with a Pricing Constraint

In order to limit the three high issues in the process of new issues, the CSRC has introduced a variety of measures to limit the IPO issue price of enterprises, and then to restrict the amount of financing obtained by enterprises. When there is an issue pricing constraint in new issue pricing, this paper sets the expected utility of enterprises to make a new issue at the upper limit of pricing. For any enterprise making an IPO, the expected utility function for the enterprises' IPO is shown in Equation (10).

$$u_{i,\eta} = \frac{min\left\{E\left(\psi_i + c_\eta|\omega_\eta\right), P^*\right\}}{(1+r)^{\eta-1}} \tag{10}$$

When enterprise i is unable to observe the information of market common factors, its decision is similar to that of unconstrained pricing. On the contrary, the utility of enterprises in delaying an IPO to $t = 2$ is reduced, so all enterprises will choose to go public at $t = 1$.

If enterprise i can observe the market common factors of the period and the market common factors of the period is greater than zero or equals to zero ($c_1 \geq 0$), due to the mean recovery feature of the market common factors, the enterprise cannot obtain the real option by delaying. Therefore, all unlisted companies in $t = 1$ will be listed as soon as possible in $t = 2$.

If the common factor, observed in $t = 1$, is $c_1 \leq 0$, for all companies not making an IPO on $t = 1$ but at $t = 2$, the expected utility for the IPO is shown in Equation (11).

$$\frac{min\left\{E(\psi_i + c_2 | c_1), P^*\right\}}{1+r} = \frac{min\{\psi_i + \lambda c_1, P^*\}}{1+r} \tag{11}$$

When the IPO is postponed to $t = 3$, the expected utility for the IPO is shown in Equation (12).

$$\frac{min\left\{E(\psi_i + c_3 | c_1), P^*\right\}}{1+r^2} = \frac{min\{\psi_i + \lambda c_1, P^*\}}{1+r^2} \tag{12}$$

where P^* is the ceiling of the IPO's expected utility. Due to the existence of the maximum price ceiling in IPO pricing, enterprises will face the following three situations: (1) in $t = 2$ and $t = 3$, they will be subject to the issue pricing constraints, and all of them will be priced according to the price ceiling to obtain the expected utility. (2) They are not subject to the issue pricing constraint in term $t = 2$, whereas they are subject to the issue pricing constraint in term $t = 3$, and this is issued according to the upper limit of the pricing. (3) The IPO in $t = 2$ and $t = 3$ will not be subject to the issue pricing.

4.2.1. Critical Conditions

Referring to the paradigm of this basic model, this section first analyzes the IPO timing equilibrium of $t = 2$, and then discusses the IPO timing equilibrium of $t = 1$. In addition, the common market factors of $t = 1$ in this part should be $c_1 \leq 0$ (according to the above analysis, if $c_1 \geq 0$, the enterprise i, should choose to go public in $t = 2$ as early as possible).

The critical equilibrium of IPO timing at $t = 2$ is shown below.

(1) Pricing constrained in both $t = 2$ and $t = 3$

In this case, the trait factor ψ_i of enterprise i is in the range shown in Equation (13).

$$\psi_i + \lambda^2 c_1 > \psi_i + \lambda c_1 > P^* \tag{13}$$

The trait factor $\psi_i > P^* - \lambda c_1$, the expected utility of IPOs in $t = 2$ is $\frac{P^*}{(1+r)^2}$, and the probability of an IPO is $F\left(\frac{\psi_i - P^*}{\lambda}\right)$. Due to the pricing constraint, enterprise i could not obtain the proceeds from the exercise of IPO timing; therefore, it must make an IPO as soon as possible in $t = 2$.

(2) Pricing unconstrained in $t = 2$ but pricing constrained in $t = 3$

In this case, the trait factor ψ_i of the enterprise is in the range shown in Equation (14).

$$\psi_i + \lambda^2 c_1 > P^* > \psi_i + \lambda c_1 \tag{14}$$

The trait factor $P^* - \lambda c_1 > \psi_i > P^* - \lambda^2 c_1$, the expected utility of IPOs in $t = 2$ is $\frac{\psi_i + \lambda c_1}{1+r}$, and the probability of an IPO is $F\left(\frac{P^* - \psi_i}{\lambda^2}\right) + F\left(\frac{P^* - \psi_i}{\lambda}\right) - 1$. After delay to $t = 3$, the IPO's expected utility is $\frac{P^*}{(1+r)^2}$. The critical condition of IPO in $t = 2$ is shown in Equation (15).

$$\psi_i = \psi_2^*(c_1) = \frac{P^* - \lambda(1+r)c_1}{1+r} \tag{15}$$

Thus, only if the trait-factor $\psi_i \geq \psi_2^*(c_1)$, the enterprise i chooses to make an IPO in $t = 2$.

(3) Pricing unconstrained in both $t = 2$ and $t = 3$

In this case, the trait factor ψ_i of the enterprise is in the range shown in Equation (16).

$$P^* > \psi_i + \lambda^2 c_1 > \psi_i + \lambda c_1 \tag{16}$$

The trait factor follows $\psi_i < P^* - \lambda^2 c_1$, and the probability is $F\left(\frac{P^*-\psi_i}{\lambda^2}\right)$. We skip the analysis here, as it is consistent with the unconstrained IPO timing.

The critical equilibrium of IPO timing at $t = 1$ is shown below.

(1) Pricing constrained in both $t = 2$ and $t = 3$

Whether there is an IPO in $t = 1$ or not, the IPO's expected utility obtained by enterprise i in $t = 2$ is $\frac{P^*}{1+r}$, lower than the expected utility P^* at $t = 1$. As a result, the enterprise i will choose to make an IPO at $t = 1$ rather than delaying.

(2) Pricing unconstrained in $t = 2$ but pricing constrained in $t = 3$

If there is no IPO in the market of $t = 1$, then the enterprise i will be listed at $t = 2$, and the expected utility is $\frac{P^*}{1+r}$; if there is at least one IPO in $t = 1$, when the trait factor is higher than the critical condition of $t = 2$, the expected utility of the IPO in $t = 2$ is $\frac{E\left(\psi_{i+}c_2|\psi_i > \psi_i^*(c_2)\right)}{1+r}$; otherwise, the IPO will be postponed to $t = 3$ and the expected utility is $\frac{P^*}{(1+r)^2}$.

The expected utility obtained by the enterprise in the IPO of $t = 1$ is P^*. The critical condition for the IPO at $t = 1$ is the trait factor of ψ_1^*, which equals the expected utility P^* at $t = 1$, as shown in Equation (17), the derivation is shown in Appendix A.

$$\begin{array}{l} P^* \\ = \left[G\left(\psi_1^*\right)\right]^{N-1}\left(\frac{P^*}{1+r}\right) \\ +\left(1-\left[G\left(\psi_1^*\right)\right]^{N-1}\right)\left\{\begin{array}{c} F\left(\frac{(1+r)\psi_1^*-P^*}{\lambda(1+r)}\right)\frac{\psi_1^*}{1+r}+\frac{\lambda}{1+r}\sigma^2 f\left(\frac{P^*-(1+r)\psi_1^*}{\lambda(1+r)}\right) \\ +\left(1-F\left(\frac{(1+r)\psi_1^*-P^*}{\lambda(1+r)}\right)\right)\frac{P^*}{(1+r)^2} \end{array}\right\} \end{array} \tag{17}$$

4.2.2. Comparative Static Analysis

(1) Pricing constraint in both $t = 2$ and $t = 3$

In this case, all enterprises will choose to make an IPO as soon as possible in phase $t = 1$, so there is no comparative static analysis of critical conditions in $t = 2$.

(2) Pricing unconstrained in $t = 2$ but pricing constrained in $t = 3$

The critical condition of the IPO in $t = 2$ is shown in Equation (18).

$$\psi_2^*(c_1) \equiv \frac{P^* - \lambda(1+r)c_1}{1+r} \tag{18}$$

With the other conditions unchanged, in this section, we discuss the influence of the changes of the following three factors on the critical conditions of $t = 2$ IP: (a) the discount rate r; (b) the common factors recovery rate λ; (c) the upper limit P^* of IPO utility with a pricing constraint, as shown in Equations (19)–(21).

$$\frac{\partial}{\partial\lambda}\psi_2^*(c_1) = -\frac{P^*}{1+r^2} < 0 \tag{19}$$

$$\frac{\partial}{\partial\lambda}\psi_2^*(c_1) = -c_1 > 0 \tag{20}$$

$$\frac{\partial}{\partial P^*}\psi_2^*(c_1) = \frac{1}{1+r} > 0 \tag{21}$$

These calculated results show that with the increase in discount cost, the critical condition of an IPO will decrease, which is consistent with the conclusion when there is no pricing constraint, and this also reflects the early listing tendency of the enterprise itself. There is only one path for the influence of the common factor mean recovery rate λ on the critical condition of an IPO in $t = 2$. The increase in λ indicates that the accuracy of predicting the future stock market conditions will be improved, which will increase the extra profit that companies can obtain from delaying their IPO. The raising of the upper limit of pricing constraint P^* will raise the critical condition of the IPO decision, if the IPO pricing limit is set in the market, the increase in the IPO pricing limit can relax the restrictions on the IPO timing, and the enterprises may gain additional expected utility from this IPO delay.

On the influence on the critical condition of an IPO decision in $t = 1$, the discount rate is negative. As for the recovery speed of common factors, in this section, we deduce its influence by analyzing the real option and the critical condition of the timing of $t = 2$ IPOs. According to the model, the IPO real option is shown in Equation (22), the derivation is shown in Appendix B.

$$V_2(\psi_i) = \int_{-\infty}^{c_2^*(\psi_i)} \left(\frac{P^*}{(1+r)^2} - \frac{\psi_i + \lambda c_1}{1+r} \right) f(c_1) dc_1 \tag{22}$$

Then, we solve for the partial derivatives, as shown in Equation (23).

$$\frac{\partial V_2(\psi_i)}{\partial \lambda} = \int_{-\infty}^{c_2^*(\psi_i)} \left(\frac{1}{1+r} \right) f(c_1) dc_1 > 0 \tag{23}$$

As the common factor mean recovery rate λ grows, enterprises can take advantage of the real option by delaying their IPOs. Thus, the higher the recovery speed λ, the higher the critical condition of an IPO in $t = 1$, for the partial differential of pricing upper limit P^*, as shown in Equation (24).

$$\frac{\partial}{\partial P^*} V_2(\psi_i) = \frac{1}{(1+r)^2} F\left(\frac{P^*}{1+r} - \psi_i \right) + \left(\frac{(1-\lambda)P^*}{1+r} - (1-\lambda)\psi_i \right) f\left(\frac{P^*}{1+r} - \psi_i \right) \tag{24}$$

The calculations suggest that, although there may be a critical value of P^*, the further derivation shows that the reduction in the upper limit of IPO pricing P^* will reduce the value of the IPO timing option, and reduce the critical condition of IPO timing of both $t = 1$ and $t = 2$, thus leading to the early listing of enterprises.

4.3. The Effects of the Pricing Constraint

Facing the IPO pricing constraint and regulation in the stock market, if the expected price of an enterprise exceeds that ceiling, it has to re-price according to the ceiling and obtain the corresponding expected utility P^*. Enterprises will face different situations of the pricing constraints according to their own trait factors. This section will expand in two aspects: the restricted probability and IPO timing critical condition.

4.3.1. Restricted IPO Probability

Keeping the trait factors of enterprises unchanged, when the IPO price ceiling P^* is lowered, the probability of IPO $F\left(\frac{\psi_i - P^*}{\lambda} \right)$ grows, the IPO probability $\left[F\left(\frac{P^* - \psi_i}{\lambda^2} \right) + F\left(\frac{P^* - \psi_i}{\lambda} \right) - 1 \right]$ goes down, and the probability $F\left(\frac{P^* - \psi_i}{\lambda^2} \right)$ declines. On the other hand, if the price ceiling P^* is fixed, our research shows that the higher the trait factors ψ_i, the higher the probability $F\left(\frac{\psi_i - P^*}{\lambda} \right)$; however, probability $\left[F\left(\frac{P^* - \psi_i}{\lambda^2} \right) + F\left(\frac{P^* - \psi_i}{\lambda} \right) - 1 \right]$ and probability $F\left(\frac{P^* - \psi_i}{\lambda^2} \right)$ decline.

4.3.2. IPO Timing Critical Condition

After the implementation of price control on the new issue pricing, the enterprises are faced with different pricing constraints due to differences in the trait factors. Our model analysis showed that pricing constraint reduces the critical condition of IPO timing when $t = 2$. At this time, the pricing constraints lead to those enterprises with high trait factor can only price according to the price ceiling at $t = 3$, thus reducing the value of the real option obtained in the delayed IPO, so they are more inclined to go public at $t = 2$ in advance.

The above analysis leads to Propositions 2–4, as below.

Proposition 2. *Lowering the ceiling on prices increases the probability that enterprises will be affected by IPO pricing constraints.*

Proposition 3. *The price constraint has a higher probability of influencing the IPO of high trait factor enterprises.*

Proposition 4. *The existence of a pricing constraint will reduce the critical condition of the IPO, which promotes enterprises going public in advance.*

4.4. Discussion

In this section, we discuss the conclusions with the previous studies. In previous literatures, the studies of enterprises' IPO decision focused on the factor of enterprises' intrinsic value [2,23], information asymmetry [11], and market investor sentiment [7–9] individually, but they failed to explain the complexity of IPOs. This paper integrates the three factors of enterprises' intrinsic value, information asymmetry and market investor sentiment into the united theoretical model of enterprises' IPO decision, and studies the complexity of the mechanisms.

The price behavior in the stock market is complex [24,25]. Our model provided new insights in understanding the relationship between IPOs and market trend. If the enterprise expects that the future stock market trend will be negative (meaning the common factor c_t is negative in the future), whether there is IPO price control in the market or not, the dominant decision of the company is to go public as soon as possible at $t = 1$, which avoids delaying the listing and reduces the expected utility obtained by the IPO. The real option on this IPO timing will not be exercised yet.

If the enterprise expects the future stock market to be positive (the common factor c_t will rise), then enterprises can exercise the real option of IPO timing to delay their IPO period to get a higher expected utility (valuation). However, when there is a pricing constraint, the value of the real option will decrease or even disappear, and the expected utility of IPOs is reduced and the IPO time has to be brought forward.

This paper can explain the complexity of IPOs [26–28]. First, it is expected that IPOs with a large financing scale will not happen in the Chinese stock market, as the issuance price constraint directly limits the financing quota of companies, and if the companies go public, they cannot obtain sufficient financing. Secondly, the valuation of technological innovation is also included in the trait factors. Due to the IPO pricing constraints, enterprises with high valuations of technological innovation will be restricted and listed earlier, which could further limit the amount of IPO capital and prompt them to stay private or change their IPO location.

The core point of this paper is that pricing constraints will reduce the value of IPO real options, which can theoretically explain the Chinese companies' IPO behavioral characteristics. Obviously, when an enterprise faces current IPO price constraints or future pricing constraints, listing as soon as possible to maximize its utility for the IPO should be the dominant choice for the enterprise. Figure 1 shows that the higher the characteristics of an enterprise (ψ), the greater the negative effect of pricing constraints (P^*) on the value of IPO real options ($\frac{\partial V_2}{\partial P^*}$).

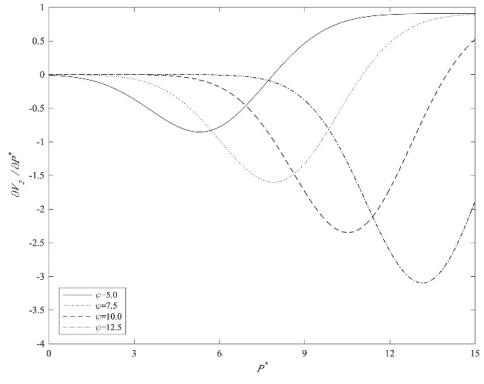


Figure 1. The effect of pricing constraints on the value of IPO real options. Note: The other parameters in the simulation are set as follows: $r = 0.05$, $\lambda = 0.5$, F(X) is the CDF of X ~ N (0,2). The result is robust when other values of the parameters are selected.

5. Conclusions

Based on the characteristics of the IPO limit in the Chinese stock market, this paper discussed the influence of pricing constraints on IPO timing, which is an extension of the IPO timing theory. In the process of IPO timing, enterprises not only need to comprehensively consider their own characteristics, the market fluctuations, and macroeconomic factors, but also must consider how the market regulation of the IPO will have a significant and different impact on their timing decision, which significantly increases the complexity of IPO decision modeling.

The IPO pricing constraint will short the waiting period of enterprises, that is, promote the enterprises going public in advance. This restriction has a more significant impact on those enterprises with high trait factors, such as technology enterprises. The early IPO of companies may hamper their ability to raise sufficient capital in the stock market and reduce their incentive to make an IPO on the Chinese stock market. Generally speaking, these findings provide new insights into understanding the complexity of IPOs in relation to the Chinese stock market.

However, this paper has the limitation that requires future study. The model assumes that the information disclosure of IPO enterprises is true and cost-free, whereas in the actual stock market, the cost of information disclosure, which is caused by information asymmetry and selective disclosure, exists. Therefore, clarifying the mechanism of information disclosure cost in the IPO market timing game should be the future development direction for the model.

Author Contributions: Conceptualization, Z.H. and Y.H.; methodology, Y.H. and Y.J.; formal analysis, Y.H. and Z.P.; validation, Z.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A.

Derivation: The critical condition of IPO timing in $t = 1$ is shown in Equation (A1)

$$\begin{aligned} &Pr\left(NI^1_{i\neq j}\right)\left(\tfrac{\psi_i}{1+r}\right)+\left(1-Pr\left(NI^1_{i\neq j}\right)\right)\left(\begin{array}{c} Pr\left(I^1_{i\neq j}\right)E\left[payoff\ at\ t=2\middle|\psi_i,I^2_i\right] \\ +Pr\left(NI^1_{i\neq j}\right)E\left[payoff\ at\ t=3\middle|\psi_i,NI^2_i\right]\end{array}\right) \\ &\quad=\left[G\left(\psi^*_1\right)\right]^{N-1}\left(\tfrac{\psi^*_1}{1+r}\right)+ \\ &\left(1-\left[G\left(\psi^*_1\right)\right]^{N-1}\right)\left(\begin{array}{c} Pr\left(\psi_i>-c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right)E\left[\frac{\psi_i+c_2}{1+r}\middle|\psi_i>c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right] \\ +\left(1-Pr\left(\psi_i>-c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right)\right)E\left[\frac{\psi_i+c_3}{(1+r)^2}\middle|\psi_i\leq-c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right]\end{array}\right) \end{aligned} \tag{A1}$$

At any point in time, the enterprise knows its own trait factors, as shown in Equation (A2).

$$Pr\left(\psi_i>-c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right)=Pr\left(c_1>-\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)=F\left(\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right) \tag{A2}$$

We can derive the Equation (A3).

$$\begin{aligned} E\left[\tfrac{\psi_i+c_2}{1+r}\middle|\psi_i>c_1(1+r-\lambda)\left(\tfrac{\lambda}{r}\right)\right]&=E\left[\tfrac{\psi_i+c_2}{1+r}\middle|c_1>\tfrac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right] \\ &=\frac{1}{F\left(\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)}\int_{-\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}}^{\infty}\frac{\psi_i+E(c_2|c_1)}{1+r}f(c_1)dc_1 \\ &=\tfrac{\psi_i}{1+r}+\tfrac{1}{1+r}\frac{1}{F\left(\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)}\int_{-\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}}^{\infty}\frac{E(c_2|c_1)}{1+r}f(c_1)dc_1 \\ &=\tfrac{\psi_i}{1+r} \\ &+\tfrac{1}{1+r}\frac{1}{F\left(\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)}\left[\int_{-\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}}^{\infty}(\lambda c_1+\varepsilon_2)f(c_2)dc_2\right]f(c_1)dc_1 \\ &=\tfrac{\psi_i}{1+r}+\tfrac{1}{1+r}\frac{1}{F\left(\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)}\int_{-\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}}^{\infty}\lambda c_1f(c_1)dc_1 \\ &=\tfrac{\psi_i}{1+r}+\tfrac{\lambda}{1+r}E\left[c_1|c_1>\tfrac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right] \end{aligned} \tag{A3}$$

As $X\sim N\left(\mu_X,\sigma^2\right)$, its expectation of truncated normal distribution is shown in Equation (A4).

$$E[X|X\in[a,b]]=\mu_X-\sigma^2\frac{f(a)-f(b)}{F(a)-F(b)} \tag{A4}$$

If $a\approx-\infty$, $E[X|X<b]=\mu_X-\sigma^2\frac{f(b)}{F(b)}$ and if $b\approx\infty$, $E[X|X>a]=\mu_X+\sigma^2\frac{f(a)}{1-F(a)}$.
Thus, we have Equation (A5).

$$\begin{aligned} E\left[\tfrac{\psi_i+c_2}{1+r}\middle|\psi_i>c_1(1+r-\lambda)\left(\tfrac{\lambda}{r}\right)\right]&=\tfrac{\psi_i}{1+r}+\tfrac{\lambda}{1+r}E\left[c_1|c_1>\tfrac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right] \\ &=\tfrac{\psi_i}{1+r}+\tfrac{\lambda}{1+r}\left(-\sigma^2\frac{-f\left(-\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)}{1-F\left(-\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)}\right) \end{aligned} \tag{A5}$$

Similarly, we have Equation (A6).

$$\begin{aligned} E\left[\frac{\psi_i+c_3}{1+r}|\psi_i \le -c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right] \\ = \frac{\psi_i}{(1+r)^2} + \frac{\lambda^2}{(1+r)^2}E\left[c_1|c_1 < -\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right] \\ = \frac{\psi_i}{(1+r)^2} + \frac{\lambda^2}{(1+r)^2}\left(-\sigma^2\frac{-f\left(-\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)}{1-F\left(\frac{\psi_i}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)}\right) \end{aligned} \tag{A6}$$

The expected utility of enterprises in delaying their IPO at $t = 1$ is shown in Equation (A7).

$$\begin{aligned} &\left[G\left(\psi_1^*\right)\right]^{N-1}\left(\frac{\psi_1^*}{1+r}\right) + \left(1-\left[G\left(\psi_1^*\right)\right]^{N-1}\right) \\ &*\left(\begin{array}{c} Pr\left(\psi_i > -c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right)E\left[\frac{\psi_i+c_2}{1+r}\middle|\psi_i > c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right] \\ +\left(1-Pr\left(\psi_i > -c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right)\right)E\left[\frac{\psi_i+c_3}{(1+r)^2}|\psi_i \le -c_1(1+r-\lambda)\left(\frac{\lambda}{r}\right)\right] \end{array}\right) = \left[G\left(\psi_1^*\right)\right]^{N-1}\left(\frac{\psi_1^*}{1+r}\right) \\ &+\left(1-\left[G\left(\psi_1^*\right)\right]^{N-1}\right) \\ &*\left(\begin{array}{c} F\left(\frac{\psi_1^*}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)\frac{\psi_1^*}{1+r} + \frac{\lambda\sigma^2}{1+r}f\left(-\frac{\psi_1^*}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right) \\ +\left(1-F\left(\frac{\psi_1^*}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right)\right)\frac{\psi_1^*}{1+r} - \frac{\lambda^2\sigma^2}{(1+r)^2}f\left(-\frac{\psi_1^*}{(1+r-\lambda)\left(\frac{\lambda}{r}\right)}\right) \end{array}\right) \end{aligned} \tag{A7}$$

Therefore, the critical condition of IPO timing in $t = 1$ without the pricing constraint can be achieved. If the pricing constraint exists, as shown in Equation (A8),

$$\begin{aligned} &Pr\left(NI_{i\neq j}^1\right)\left(\frac{P^*}{1+r}\right) + (1- \\ &Pr\left(NI_{i\neq j}^1\right)\left(\begin{array}{c} Pr\left(I_{i\neq j}^1\right)E\left[payoff\ at\ t=2\middle|\psi_i, I_i^2\right] \\ +Pr\left(NI_{i\neq j}^1\right)E\left[payoff\ at\ t=3\middle|\psi_i, NI_i^2\right] \end{array}\right) = \left[G\left(\psi_1^*\right)\right]^{N-1}\left(\frac{P^*}{1+r}\right) + \\ &(1- \\ &\left[G\left(\psi_1^*\right)\right]^{N-1}\left(\begin{array}{c} F\left(\frac{(1+r)\psi_1^*-P^*}{\lambda(1+r)}\right)E\left[payoff\ at\ t=2\middle|\psi_i, I_i^2\right] \\ +\left(1-F\left(\frac{(1+r)\psi_1^*-P^*}{\lambda(1+r)}\right)\right)E\left[payoff\ at\ t=3\middle|\psi_i, NI_i^2\right] \end{array}\right) = \\ &\left[G\left(\psi_1^*\right)\right]^{N-1}\left(\frac{P^*}{1+r}\right) + (1- \\ &\left[G\left(\psi_1^*\right)\right]^{N-1}\left(\begin{array}{c} Pr\left(\psi_i > \frac{P^*-(1+r)c_1}{\lambda(1+r)}\right)E\left[\frac{\psi_i+c_2}{1+r}\middle|\psi_i > \frac{P^*-(1+r)c_1}{\lambda(1+r)}\right] \\ +Pr\left(\psi_i \le \frac{P^*-(1+r)c_1}{\lambda(1+r)}\right)\frac{P^*}{(1+r)^2} \end{array}\right) \end{aligned} \tag{A8}$$

and as shown in Equation (A9),

$$\begin{aligned} E\left[\frac{\psi_i+c_2}{1+r}\middle|\psi_i>\frac{P^*-(1+r)c_1}{\lambda(1+r)}\right] &= E\left[\frac{\psi_i+c_2}{1+r}|c_1>\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}\right] \\ &= \frac{1}{F\left(\frac{(1+r)\psi_i-P^*}{\lambda(1+r)}\right)}\int_{\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}}^{\infty}\frac{\psi_i+E[c_2|c_1]}{1+r}f(c_1)dc_1 \\ &= \frac{\psi_i}{1+r}+\frac{1}{1+r}\frac{1}{F\left(\frac{(1+r)\psi_i-P^*}{\lambda(1+r)}\right)}\int_{\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}}^{\infty}\frac{E[c_2|c_1]}{1+r}f(c_1)dc_1 \\ &= \frac{\psi_i}{1+r}+\frac{1}{1+r}E\left[c_2|c_1>\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}\right] \\ &= \frac{\psi_i}{1+r}+\frac{1}{1+r}\left(0-\sigma^2\frac{-f\left(\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}\right)}{1-F\left(\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}\right)}\right) \\ &= \frac{\psi_i}{1+r}+\frac{1}{1+r}\sigma^2\frac{f\left(\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}\right)}{1-F\left(\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}\right)} \end{aligned} \tag{A9}$$

thus, the expected utility of enterprises in postponing the IPO at $t = 1$ is shown in Equation (A10).

$$\begin{aligned} &Pr\left(NI^1_{i\neq j}\right)\left(\frac{P^*}{1+r}\right) \\ &+\left(1-Pr\left(NI^1_{i\neq j}\right)\right)\begin{pmatrix} Pr\left(I^1_{i\neq j}\right)E\left[payoff\ at\ t=2\middle|\psi_i,I_i^2\right] \\ +Pr\left(NI^1_{i\neq j}\right)E\left[payoff\ at\ t=3\middle|\psi_i,NI_i^2\right]\end{pmatrix} = \left[G\left(\psi_1^*\right)\right]^{N-1}\left(\frac{P^*}{1+r}\right) \\ &+\left(1-\left[G\left(\psi_1^*\right)\right]^{N-1}\right)\begin{pmatrix} F\left(\frac{(1+r)\psi_1^*-P^*}{\lambda(1+r)}\right)\frac{\psi_i}{1+r}+\frac{1}{1+r}\sigma^2 f\left(\frac{P^*-(1+r)\psi_i}{\lambda(1+r)}\right) \\ +\left(1-F\left(\frac{(1+r)\psi_1^*-P^*}{\lambda(1+r)}\right)\right)\frac{P^*}{(1+r)^2}\end{pmatrix} \end{aligned} \tag{A10}$$

Appendix B. Derivation: The Real Option of IPO Timing

If $c_2 < c_2^*(\psi_i)$, the strategic enterprises will choose to delay their IPO until $t = 3$, and we get Equation (A11).

$$(\psi_i) = Pr\left(c_2<c_2^*(\psi_i)\right)E\left[\frac{\psi_i+c_3}{(1+r)^2}\right]+\left(1-Pr\left(c_2<c_2^*(\psi_i)\right)\right)E\left[\frac{\psi_i+c_2}{1+r}\right] \tag{A11}$$

The value of the IPO timing real option is shown in Equation (A12).

$$\begin{aligned} V_2(\psi_i)\equiv S(\psi_i) \quad & -NS(\psi_i) \\ &= Pr\left(c_2<c_2^*(\psi_i)\right)E\left[\frac{\psi_i+c_3}{(1+r)^2}\right]+\left(1-Pr\left(c_2<c_2^*(\psi_i)\right)\right)E\left[\frac{\psi_i+c_2}{1+r}\right] \\ &-E\left[\frac{\psi_i+c_2}{1+r}\right] \\ &= Pr\left(c_2<c_2^*(\psi_i)\right)E\left[\frac{\psi_i+c_3}{(1+r)^2}\right]-Pr\left(c_2<c_2^*(\psi_i)\right)E\left[\frac{\psi_i+c_2}{1+r}\right] \\ &= Pr\left(c_2<c_2^*(\psi_i)\right)E\left[\frac{\psi_i+c_3}{(1+r)^2}-\frac{\psi_i+c_2}{1+r}\middle|c_2<c_2^*(\psi_i)\right] \\ &= F\left(c_2^*(\psi_i)\right)\frac{1}{F\left(c_2^*(\psi_i)\right)}\int_{-\infty}^{c_2^*(\psi_i)}\left(\frac{\psi_i+E[c_3|c_1]}{(1+r)^2}\right. \\ &\left.-\frac{\psi_i+E[c_2|c_1]}{1+r}\right)f(c_1)dc_1 \\ &= \int_{-\infty}^{c_2^*(\psi_i)}\left(\frac{\psi_i+E[c_3|c_1]}{(1+r)^2}-\frac{\psi_i+E[c_2|c_1]}{1+r}\right)f(c_1)dc_1 \end{aligned} \tag{A12}$$

and as shown in Equations (A13) and (A14).

$$E[c_2|c_1] = \int_{-\infty}^{\infty} (\lambda c_1 + \varepsilon_2) f(\varepsilon_2) d\varepsilon_2 = \lambda c_1 \tag{A13}$$

$$E[c_3|c_1] = \int_{-\infty}^{\infty} \left(\lambda \left(\int_{-\infty}^{\infty} (\lambda c_1 + \varepsilon_2) f(\varepsilon_2) d\varepsilon_2 \right) + \varepsilon_2 \right) f(\varepsilon_2) d\varepsilon_2 = \lambda^2 c_1 \tag{A14}$$

thus, we have Equation (A15).

$$V_2(\psi_i) = \int_{-\infty}^{c_2^*(\psi_i)} \left(\frac{\psi_i + \lambda^2 c_1}{(1+r)^2} - \frac{\psi_i + \lambda c_1}{1+r} \right) f(c_1) dc_1 \tag{A15}$$

This derivation does not consider the pricing constraint, when the IPO is faced with the issue pricing constraint, the IPO timing real option is shown in Equation (A16).

$$\begin{aligned} V_2(\psi_i) \equiv S(\psi_i) \quad & -NS(\psi_i) \\ & = Pr\left(c_2 < c_2^*(\psi_i)\right) \tfrac{P^*}{(1+r)^2} + \left(1 - Pr\left(c_2 < c_2^*(\psi_i)\right)\right) E\left[\tfrac{\psi_i + c_2}{1+r}\right] \\ & - E\left[\tfrac{\psi_i + c_2}{1+r}\right] \\ & = Pr\left(c_2 < c_2^*(\psi_i)\right) \tfrac{P^*}{(1+r)^2} - Pr\left(c_2 < c_2^*(\psi_i)\right) E\left[\tfrac{\psi_i + c_2}{1+r}\right] \\ & = Pr\left(c_2 < c_2^*(\psi_i)\right) E\left[\tfrac{P^*}{(1+r)^2} - \tfrac{\psi_i + c_2}{1+r} | c_2 < c_2^*(\psi_i)\right] \\ & = F\left(c_2^*(\psi_i)\right) \tfrac{1}{F\left(c_2^*(\psi_i)\right)} \int_{-\infty}^{c_2^*(\psi_i)} \left(\tfrac{P^*}{(1+r)^2} - \tfrac{\psi_i + E[c_2|c_1]}{1+r} \right) f(c_1) dc_1 \\ & = \int_{-\infty}^{c_2^*(\psi_i)} \left(\tfrac{P^*}{(1+r)^2} - \tfrac{\psi_i + \lambda c_1}{1+r} \right) f(c_1) dc_1 \end{aligned} \tag{A16}$$

References

1. Ritter, J.R.; Welch, I. A Review of IPO Activity, Pricing and Allocations. *J. Financ.* **2002**, *57*, 1795–1828. [CrossRef]
2. Ritter, J.R. The Long-Run Performance of Initial Public Offerings. *J. Financ.* **1991**, *46*, 3–27. [CrossRef]
3. Kwapień, J.; Drożdż, S. Physical Approach to Complex Systems. *Phys. Rep.* **2012**, *515*, 115–226. [CrossRef]
4. Ibbotson, R.G.; Sindelar, J.L.; Ritter, J.R. Initial Public Offerings. *J. Appl. Corp. Financ.* **1988**, *1*, 37–45. [CrossRef]
5. Korajczyk, R.A.; Lucas, D.J.; Mcdonald, R.I. Understanding Stock Price Behavior around the Time of Equity Issues. *J. Financ. Quant. Anal.* **1992**, *27*, 397–418. [CrossRef]
6. Lucas, D.J.; Mcdonald, R.L. Equity Issues and Stock Price Dynamics. *J. Financ.* **1990**, *45*, 1019–1043. [CrossRef]
7. Lowry, M.B. Why does IPO Volume Fluctuate So Much? *J. Financ. Econ.* **2003**, *67*, 3–40. [CrossRef]
8. Derrien, F. IPO Pricing in "Hot" Market Conditions: Who Leaves Money on the Table? *J. Financ.* **2005**, *60*, 487–521. [CrossRef]
9. Ljungqvist, A.P.; Nanda, V.K.; Singh, R. Hot Markets, Investor Sentiment and IPO Pricing. *J. Bus.* **2006**, *79*, 1667–1702. [CrossRef]
10. Chemmanur, T.J.; Fulghieri, P. A Theory of the Going-Public Decision. *Rev. Financ. Stud.* **1999**, *12*, 249–279. [CrossRef]
11. Hoffmann-Burchardi, U. Clustering of Initial Public Offerings, Information Revelation and Underpricing. *Eur. Econ. Rev.* **2001**, *45*, 353–383. [CrossRef]
12. Alti, A. IPO Market Timing. *Rev. Financ. Stud.* **2005**, *18*, 1105–1138. [CrossRef]
13. Çolak, G.; Günay, H. Strategic Waiting in the IPO markets. *J. Corp. Financ.* **2011**, *17*, 555–583. [CrossRef]
14. Pastor, L.; Veronesi, P. Rational IPO Waves. *J. Financ.* **2005**, *60*, 1713–1757. [CrossRef]

15. Spiegel, M.I.; Tookes, H. Dynamic Competition, Innovation and Strategic Financing. *Soc. Sci. Electron. Publ.* **2008**, 1–64. [CrossRef]
16. Aghamolla, C.; Guttman, L. Strategic Timing of IPO and Disclosure: A Dynamic Model of Multiple Firms. In *Working Paper*; Columbia University: New York, NY, USA, 2015.
17. Fama, E.F.; French, K.R. Permanent and Temporary Components of Stock Prices. *J. Political Econ.* **1988**, *96*, 246–273. [CrossRef]
18. Bessembinder, H.; Coughenour, J.F.; Seguin, P.J.; Smoller, M.M. Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure. *J. Financ.* **1995**, *50*, 361–375. [CrossRef]
19. Campbell, J.Y.; Hentschel, L. No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. *J. Financ. Econ.* **1992**, *31*, 281–318. [CrossRef]
20. Veronesi, P. Stock Market Overreactions to Bad News in Good Times: A Rational Expectations Equilibrium Model. *Rev. Financ. Stud.* **1999**, *12*, 975–1007. [CrossRef]
21. Myers, S.C. Determinants of Corporate Borrowing. *J. Financ. Econ.* **1977**, *5*, 147–175. [CrossRef]
22. Benveniste, L.M.; Busaba, W.Y.; Wilhelm, W.J. Information Externalities and the Role of Underwriters in Primary Equity Markets. *J. Financ. Intermediation* **2002**, *11*, 61–86. [CrossRef]
23. Jenkinson, T.; Ljungqvist, A.P. *Going Public: The Theory and Evidence on How Companies Raise Equity Finance*; Oxford University Press on Demand: Oxford, UK, 2001.
24. Anton, S.G.; Afloarei Nucu, A.E. Sovereign Credit Default Swap and Stock Markets in Central and Eastern European Countries: Are Feedback Effects at Work? *Entropy* **2020**, *22*, 338. [CrossRef]
25. Peng, Z.; Hu, C.S. The Threshold Effect of Leveraged Trading on the Stock Price Crash Risk: Evidence from China. *Entropy* **2020**, *22*, 268. [CrossRef]
26. Lang, M.; Lundholm, R. Cross-Sectional Determinants of Analyst Ratings of Corporate Disclosures. *J. Account. Res.* **1994**, *31*, 246–271. [CrossRef]
27. Lerner, J. Venture Capitalists and the Decision to Go Public. *J. Financ. Econ.* **1994**, *35*, 293–316. [CrossRef]
28. Pagano, M.; Panetta, F.; Zingales, L. Why Do Companies Go Public? An Empirical Analysis. *J. Financ.* **1998**, *53*, 27–64. [CrossRef]

entropy

MDPI

Article

A Comprehensive Framework for Uncovering Non-Linearity and Chaos in Financial Markets: Empirical Evidence for Four Major Stock Market Indices

Lucia Inglada-Perez

Department of Statistics and Operational Research, Complutense University, Plaza Ramón y Cajal, s/n Ciudad Universitaria, 28040 Madrid, Spain; lucia.inglada.perez@ucm.es

Received: 9 November 2020; Accepted: 14 December 2020; Published: 18 December 2020

Abstract: The presence of chaos in the financial markets has been the subject of a great number of studies, but the results have been contradictory and inconclusive. This research tests for the existence of nonlinear patterns and chaotic nature in four major stock market indices: namely Dow Jones Industrial Average, Ibex 35, Nasdaq-100 and Nikkei 225. To this end, a comprehensive framework has been adopted encompassing a wide range of techniques and the most suitable methods for the analysis of noisy time series. By using daily closing values from January 1992 to July 2013, this study employs twelve techniques and tools of which five are specific to detecting chaos. The findings show no clear evidence of chaos, suggesting that the behavior of financial markets is nonlinear and stochastic.

Keywords: nonlinear dynamics; chaos; time series analysis; stock exchange market; Lyapunov; recurrence plots; BDS; correlation dimension; GARCH model

1. Introduction

Research on modeling financial time series has traditionally assumed linear patterns. Unfortunately, these models are incapable of explaining specific phenomena or market events, such as bubbles and recessions. Therefore, there is interest in and a need to introduce alternative methods, most of which come from other scientific disciplines such as mathematics, physics and engineering [1], to model the dynamics of financial series, and to detect a possible nonlinear and determinist chaotic behavior. This opens up the range of alternatives for analyzing and predicting financial series and particularly stock market price series, traditionally anchored in methods such as technical and fundamental analysis or in assumptions like the efficient market, defined by Fama [2] and based on the fact that asset prices reflect all available information, causing asset returns to be unpredictable [3].

Many scholars have empirically studied the existence of nonlinear dynamics in financial series [4], meanwhile several theoretical models consistent with the presence of nonlinearity in asset prices have arisen [5]. The nonlinear approach can capture the characteristics of the financial series and their sudden fluctuations, and, therefore, plays an important role in economic modeling [6,7]. The initial interest in the application of nonlinear models has been extended to solve the question of whether the nonlinear dynamics have a stochastic or deterministic behavior. This issue constitutes a key point in the process of modeling and forecasting financial time series. The emergence of new models that deal with the volatility present in financial time series, such as the autoregressive conditional heteroskedasticity ARCH [8], generalized (GARCH) [9] and exponential GARCH (EGARCH) [10] models, in consonance with the development of the chaos theory that can explain the effect of shocks on the stock markets as part of the endogenous dynamics of the series itself, constitute a fundamental contribution to solving this issue.

Although the literature on the evidence of chaos in economic and financial time series is extensive, there are no incontrovertible results [11]. There are two divergent positions concerning the underlying dynamics of financial markets. Whilst several studies have found evidence of chaotic patterns in financial markets e.g., [12,13], the most recent studies support a non-chaotic but rather stochastic behavior. This is due to the use of an incorrect specification and to the approach applied [14,15]. Indeed, the former studies applied a methodology, such as the correlation dimension method, that is not sensitive to noise, an intrinsic feature of financial series. However, studies that have used new and more robust methods, such as the test proposed by Matilla-Garcia and Ruiz Marin (MGRM test) [6] and the Lyapunov test based on BenSaïda and Litimi [14], do not find evidence for chaos. In fact, to ensure that conclusions are sound and reliable, traditional tests should be applied to time series that consist of many observations and are noise-free.

The existence of chaos in the stock market remains a relevant issue. Its importance lies in the fact that finding the chaos of low dimension could allow a reliable forecast in the short term, but not in the medium and long term, since a chaotic system is unstable [16]. However, there is no standard methodology for the analysis of the presence of nonlinearity and the existence of a chaotic pattern. Therefore, it is mandatory to explore the existence of a chaotic deterministic structure considering different approaches [17].

The main goal of this research is to shed light on the behavior of stock markets by examining their dynamics and volatility. Through a comprehensive methodological approach, using a wide range of procedures, this study investigates the existence of deterministic chaos and nonlinearity in a series of daily stock indexes of four major stock markets: namely the Dow Jones Industrial Average, the Nasdaq-100, the Ibex 35 and the Nikkei 225. In addition, the paper examines other issues such as the existence of volatility clustering. By using daily closing values from January 1992 to July 2013, this study employs twelve techniques and tools of which five are specific to detecting chaos.

Overall, this work contributes to filling the gap in the international literature on nonlinear and chaotic behavior in stock markets as it uses a variety of tools including the most powerful methods to uncover the underlying dynamics of these markets. Even more so, a review of the international literature (see following section) shows that there are numerous works that aim to detect the existence of nonlinear and chaotic behavior in stock markets, but to the best of our knowledge none of them apply such a high number of techniques as this research.

The methodology applied consisted of the following steps. First, all linear dependencies were removed from the data by applying autoregressive integrated moving average (ARIMA) filters. A wide range of procedures were then applied to the residuals obtained to detect the existence of nonlinearity. If nonlinear dependence was detected, since linear structures had already been removed using the best fit ARIMA model, it was indicative of some type of nonlinear dependence in the returns series resulting from a nonlinear stochastic system or a nonlinear chaotic system. Then it might be caused by the existence of a volatility cluster. In this case, the appropriate generalized autoregressive conditional heteroskedasticity (GARCH) and exponential GARCH models (EGARCH) were applied. Afterwards, the existence of chaotic motion was explored by means of five techniques i.e., 0/1 test, the correlation dimension, Lyapunov exponent, MGRM test [6] and recurrence plots, including those that are ideally suitable for noisy time series analysis.

Empirical findings suggested that although there has been found a dominant nonlinear structure in financial markets, determinism cannot be assumed and hence chaos cannot be inferred.

The rest of the paper is organized as follows: in the second section, a review of the international literature on stock market modelling is carried out, with special emphasis on those works that contrast the possible existence of nonlinear and chaotic behavior. The third section describes the methodological process adopted in this research. In the fourth section, the data sources used as well as the main results obtained for the four stock market index series considered are presented and commented on, and the existing evidence for and against the hypotheses of nonlinearity and a chaotic regime is analyzed

and discussed. Finally, the last section draws a series of conclusions on the research carried out and indicates possible future lines of research.

2. Literature Review

When studying the existing literature on research aimed at analyzing and contrasting the existence of nonlinear and chaotic dynamics in the stock markets, a first conclusion is that the results obtained depend on the data and methods used in each study. As shown in Table 1, most studies, especially since the appearance of the BDS test [18], suggest the existence of nonlinear dependencies in stock markets.

Table 1. Results of research on nonlinearity and chaos in the stock market.

Authors (Year)	Series Considered	Tools Used	Results
Scheinkman and LeBaron (1989) [12]	They study several weekly and daily series of the US stock market.	BDS statistics. Correlation Dimension.	Evidence of nonlinearity. Evidence of chaos.
Hsieh (1991) [19]	He studies the weekly series of returns for a stock market portfolio.	BDS test. Correlation Dimension.	No evidence of IID. No evidence of chaos.
Blank (1991) [20]	He analyzes the behavior of the S&P 500 stock index	Correlation Dimension. Lyapunov exponents test.	Evidence of chaos (presence of deterministic nonlinearity). This is a necessary but not sufficient condition for chaos, as there are no tests of statistical significance of the estimates.
Mayfield and Mizrach (1992) [21]	They analyze the time series of the high frequency returns in the S&P 500 index.	Correlation Dimension. Lyapunov test.	Evidence of chaos.
Yang and Brorsen (1993) [22]	They study several futures markets, including the S&P 500 stock index.	BDS. Brock's residual test.	Evidence of nonlinearity. Evidence of chaos in about half of the cases studied.
Abhyankar et al. (1995) [23]	They study the presence of nonlinearity and chaos in the FTSE-100 index yield series.	Hinich test. BDS test. Lyapunov exponent test.	Evidence of nonlinearity. No evidence of chaos.
Sewell et al. (1996) [24]	They investigate the weekly changes in six major stock indices (the US, Korea, Taiwan, Japan, Singapore and Hong Kong) and the World Index.	Spectral analysis. Nonlinear dynamics techniques.	Evidence of nonlinearity in some of the time series. No evidence of chaos.
Abhyankar et al. (1997) [25]	They examine nonlinear dependence and chaos in the returns of the stock market indices: FTSE 100, S&P 500, NIKKE 225 and DAX.	BDS. Lee, White and Granger neural-network-based tests. Lyapunov exponent.	Evidence of nonlinearity. No evidence of chaos.
Barkoulas and Travlos (1998) [3]	They study the presence of a chaotic behavior in the Athens stock market.	BDS test. Correlation Dimension. Kolmogorov Entropy.	Evidence of nonlinearity. No evidence of chaos.
Pandey et al. (1998) [26]	They investigate the presence of chaos in five major European stock markets and the United States.	BDS. Rescaled Range (R/S) analysis.	Evidence of nonlinearity. No IID. No evidence of chaos.
Gao and Wang (1999) [27]	They study among others the daily futures series of the S&P 500.	BDS statistic. TAR-F Statistic. Q2 Test.	Evidence of nonlinearity. No evidence of chaos.
McKenzie (2001) [15]	He tests for the presence of chaos in 12 national stock market indices.	BDS test. Close returns test.	Evidence of nonlinearity. No evidence of chaos.
Kyrtsou and Terraza (2002) [28]	They examine the dynamics of the French stock market (CAC 40 Index).	Fractional integration test. Correlation dimension. Lyapunov exponents.	Evidence of nonlinearity. Evidence of chaos.

Table 1. *Cont.*

Authors (Year)	Series Considered	Tools Used	Results
Antoniou and Vorlow (2005) [29]	They examine the noise-free versions of a set of FTSE 100 stock returns time series.	BDS statistic. Surrogate data analysis.	Evidence of nonlinearity. Inconclusive evidence of chaos.
Yousefpoor et al. (2008) [30]	They study the possible chaotic behavior of some selective stocks from the Tehran stock market.	BDS test. Lyapunov exponent test. Close returns test.	Evidence of nonlinearity. No evidence of chaos.
Matilla-Garcia and Ruiz Marin (2010) [6]	They study the chaotic behavior of the daily series of the Dow Jones Industrial Average stock market index.	0–1 test based on permutation entropy.	No evidence of chaos.
Mishra et al. (2011) [31]	This study tests for the presence of nonlinear dependence and deterministic chaos in the rate of returns series for six Indian stock market indices.	Test of independence. Variance ratio test. Hurst exponent. BDS test. Lyapunov exponent test.	Evidence of nonlinearity. Evidence of chaos in two out of six cases.
BenSaïda (2012) [32]	He investigates the existence of chaotic dynamics in the S&P 500, Nikkei 225 and CAC 40 stock index during the period 1999–2008.	A new methodology to apply Lyapunov's exponent method.	No evidence of chaos.
Webel (2012) [33]	He analyzes chaotic behavior daily log returns of the 30 DAX members.	0–1 test.	Evidence of chaos.
BenSaïda and Litimi (2013) [14]	They examine nonlinearity and chaotic behavior in the following indices S&P 500, NASDAQ composite, Nikkei 225, CAC 40, FTSE 100 and DAX.	Lyapunov exponent (Neural network architecture is used in this test).	Evidence of nonlinearity. No evidence of chaos.
BenSaïda (2014) [34]	He investigates the existence of chaotic dynamics in the Standard and Poor's 500 index returns over 4 different frequencies: weekly, daily, 30 min and 5 min basis.	Lyapunov exponent test.	No evidence of chaos.
Tiwari and Gupta (2019) [35]	They test for chaos in the historical daily and monthly datasets spanning over one century of stock returns for G7 countries.	0–1 test. Lyapunov exponent.	Evidence of chaos.

IId stands for independent and identically distributed. Source: Own work.

Whether this nonlinear dependence is from chaotic deterministic origin, a sample of the most representative publications on this issue (Table 1) leads to conclude that it is not possible to give a single answer to this question, although there is a clear tendency in the most recent research towards the alternative of no chaotic behavior in the stock market series. Thus, in the first stage, initiated by the work of Scheinkman and LeBaron [12], who, using daily and weekly data of returns, found evidence of chaotic dynamics in the US stock markets, works that obtained a chaotic behavior were frequent. However, more recently, once new methods and more powerful tests on the detection of chaotic dynamics were incorporated, it can be observed that most research suggests that the data are not chaotic, that is, they present stochastic behavior.

In this regard, many of the studies carried out do not reach a categorical conclusion on the presence of chaos in financial markets due to an erroneous specification of the methods and tests used [14,15]. In fact, the traditional tests used should be applied to time series with a high number of observations and free of noise in order to obtain unambiguous conclusions, but the latter is not usually the case in financial time series, and in particular in stock markets [14,15].

In relation to the development of new tests to detect chaotic behavior, among others, BenSaïda and Litimi [14] incorporated a new test that has been shown to be more powerful in detecting chaos

than the method based on the Lyapunov exponent sign, very frequently used in previous studies. Using this test, they analyzed the behavior of ten financial series of returns and concluded that there is nonlinearity in the examined series but in none of them is there a chaotic behavior, suggesting that the chosen financial series are stochastic. Matilla-Garcia and Ruiz Marin [6] proposed a new test to determine whether the dynamics of a series is deterministic (including low dimensional chaos) rather than stochastic and applied it to the daily series of the Dow Jones index.

3. Methodology

3.1. Methodological Framework

For the study of the existence of nonlinearity and chaotic behavior, a comprehensive methodological framework has been adopted. Table 2 presents all the techniques used in this research. Their main characteristics are described below, with special emphasis on the tools for detecting chaotic dynamics in a time series.

Table 2. Tools used.

Tool	Reference	Feature Tested
Augmented Dickey Fuller (ADF)	[36]	Stationarity (unit roots)
Phillips–Perron (PP)	[37]	Stationarity (unit roots)
Kwiatkowski–Phillips–Schmidt–Shin (KPSS)	[38]	Stationarity (unit roots)
Runs	[39]	Randomness
Keenan	[40]	Nonlinearity
Tsay	[41]	Nonlinearity
Teräsvirta	[42]	Nonlinearity
White	[43]	Nonlinearity
BDS	[18]	Independence; Randomness
	[44]	IID; Nonlinearity
Kaplan	[45]	Nonlinearity
0/1	[46]	Chaos
Correlation Dimension	[47]	Chaos
Lyapunov Exponent	[14]	Chaos
MGRM	[6]	Chaos
Recurrence Plots	[48]	Chaos

Before proceeding with the description of the methodological process, it is appropriate to define the concepts of linear and nonlinear series. A stochastic process $(X_t,\ t \in Z)$ is said to be a linear process if for every $t \in Z$ $X_t = \sum_{j=0}^{\infty} \beta_j \epsilon_{t-j}$ where $a_0 = 1$, $\epsilon_t,\ t \in Z$ is a process of independent and identically-distributed random variables (iid) with $E[X_t] = 0$, $E\left[X_t^2\right] = \sigma^2$ and $\sum_{j=0}^{\infty} \left|\beta_j\right| < \infty$. Any process that does not satisfy this condition is said to be nonlinear. Nonlinear models display features that cannot be modelled by linear processes: e.g., higher-moment structures, time-changing variance, volatility clustering, breaks, thresholds, and asymmetric cycles. These traits are very often present in the study of financial time series.

MGRM stands for the test proposed by Matilla-Garcia and Ruiz Marin ([6] Source: Own work)

The methodological framework includes the following steps.

(i) Stationarity testing

Stationarity was studied by means of the following tests:

- Augmented Dickey Fuller Test (ADF)
- Phillips–Perron Test (PP)
- Kwiatkowski–Phillips–Schmidt–Shin Test (KPSS)

After testing that each series under study was not stationary, the log-returns (from now on returns) from each series were obtained by means of the expression: $Ln(R_{t-1}) = \ln(P_t) - \ln(P_{t-1})$ such that P_t is the time series data at time t.

This new series called "return" has important properties such as stationarity. This property is essential, since it is a prerequisite in a large percentage of both the modelling techniques and the analysis of nonlinearity or chaotic behavior that will be used in this research.

(ii) Linearity modelling

Subsequently, all linear dependence is removed from the return series by applying autoregressive moving average (ARMA) filters. The choice of the model was made as previously described [10], which basically consists of choosing the ARMA (p,q) model which presents the lowest value according to the Schwarz Information Criterion [49].

Following the process described by Barkoulas et al. [50], the residuals have been studied. In the case that the residuals were correlated, the order of the ARMA model has been increased, increasing the magnitude of p and q. Once the optimal linear model has been determined, the residuals are extracted from it. The new residuals are called the ARMA series.

(iii) Nonlinearity testing

The presence of nonlinearity is a necessary but not sufficient condition for the existence of a chaotic deterministic component in the system. Moreover, it is an inherent characteristic of complex systems. To test the existence of nonlinearity and randomness, a wide range of methods has been used, thus managing to supplement the possible limitations of each of them. The following techniques were applied (Table 2).

- The Runs test [39] to evaluate the randomness of the series.
- Keenan's test [40], a Portmanteau's test that contrasts the hypothesis of a linear model versus a nonlinear one.
- Tsay's test [41], which assumes a generalization of the previous test.
- The test of Teräsvirta [42], which is based on the methodology of neural networks.
- White's test [43], which is also based on the neural network methodology.
- BDS [18,44], one of the most powerful tests for nonlinearity and other types of data dependence.
- Kaplan's test [5] that has also demonstrated statistical power in determining whether data are linear.

The properties of the tests used in this research are well known in the literature so only the main differences between them are described below.

Keenan's test [40] examines the hypothesis of a linear model versus a nonlinear one. It is a particular case of the RESET test. The starting hypothesis for the RESET test is that data correspond to a linear model versus the alternative hypothesis that the model is nonlinear. It is based on the idea that if the residuals from the linear model are independent, they should not be correlated with the regressors used in the estimated equation or with the adjusted values and therefore, the regression in the residuals of these values should not be statistically significant. The null hypothesis of linearity is rejected if the value of the F-statistic for the chosen sample exceeds the theoretical value of the F-standard. The RESET test is easy to implement and does not require fitting many parameters and also has a reasonable power to detect some types of nonlinearity.

Keenan [40] proposes a nonlinearity test changing the RESET test to avoid multi-collinearity. Specifically, Keenan assumes that every stationary time series $\{Y_t\}$ can be approximated (Volterra expansion) as follows:

$$Y_t = \mu + \sum_{u=-\infty}^{\infty} \theta_u \varepsilon_{t-u} + \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \theta_{uv} \varepsilon_{t-u} \varepsilon_{t-v} \quad (1)$$

such that μ is the mean of Y_t and $\{\varepsilon_t\}$ is a stationary, independent and identically distributed sequence with a mean of zero. The process is linear, if the term containing the double sum

$\sum_{u=-\infty}^{\infty}\sum_{v=-\infty}^{\infty}\theta_{uv}\varepsilon_{t-u}\varepsilon_{t-v}$ is null, that is if all the coefficients θ are zero. The method consists of testing whether the coefficient of the double-sum term is zero, to affirm whether the process is linear or not. The test statistic has as a null hypothesis that the time series is linear, and is asymptotically distributed as $F(1,\ n-2p-2)$ distribution, where n is the sample size, and p is the order of linear AR(p) model, that is the number of lags involved in the regression.

The Tsay's test [41] is a generalization of the previous test. Tsay's test improved on the power of the Keenan's test by allowing for disaggregated nonlinear variables, thus generalizing Keenan's test by explicitly looking for quadratic serial dependence in the data. Tsay [41] showed that corresponding test statistic is asymptotically distributed as Snedecor's F distribution with the following degrees of freedom: $F(m,\ n-p-m-1)$ where n is the sample size, $m = p(p-1)/2$, and p is the order of AR fitted model, that is the number of lags involved in the regression.

The Teräsvirta test [42] is constructed using neural network models. This test models the original series by means of a Taylor series. It assumes a model of the shape:

$$y_t = y_t\beta + \sum_{j=1}^{q}\sum_{j=1}^{q}\delta_j y_{t-i}y_{t-j} + \sum_{j=1}^{q}\sum_{j=1}^{q}\sum_{j=1}^{q}\delta_{i,j,k}y_{t-i}y_{t-j}y_{t-k} + \varepsilon_t \quad (2)$$

The null hypothesis is the following: $H_0 = \delta_j = \delta_{ijk} = 0$.

The contrast, and therefore the choice of the optimal model, is made by means of a statistic similar to that used in the previous tests and which attends to an $F(p_2-p_1,\ n-p_2)$ distribution, where p_1, p_2 and n are the number of parameters of the first model, the number of parameters of the second model and the sum of both, respectively.

White's test [43] also uses neural net methods to test for nonlinearity. In this test, the time series is fitted by a single hidden-layer feed-forward neural network, which is used to determine whether any nonlinear structure remains in the residuals of an AR(p) process fitted to the same time series. White's test has power to test against various types of nonlinearity in the mean, thus it can be used to distinguish among those nonlinear processes that are nonlinear in the mean and those that are not (such as ARCH and GARCH) [51]. The null hypothesis for the test is linearity in the mean. This methodology has the advantage that a pre-filtering of the conditional variance is unnecessary. A fitted neural net is used to produce the measurable function of the process's history and an AR(p) process as the linear filter. The hypothesis that the fitted function does not correlate with the residuals of the AR(p) process is then tested. The resulting test statistic has an asymptotic chi squared distribution under the null of linearity in the mean.

BDS [18,44] is one of the most powerful tests for nonlinearity and other types of data dependence. The test uses the correlation integral, a measure of the number of times that temporal patterns are repeated in the data as the test statistic. It was originally formulated to study independence and nonlinear structure (iid) in a time series. However, the test also has statistical power to detect a large number of linear and nonlinear processes. In particular, the authors of the method showed that when applying the test to the residuals of a linear model, if the test rejects the initial hypothesis (i.e., they are iid) it will indicate the existence of linear dependence in the data. Consider a time series $\{y_t\}_{t=1}^{T}$ and define its m-history as $y_t^m = \left(y_t, y_{t-1}, \ldots,\ y_{t-(m-1)}\right)$. The BDS tests the null hypothesis that the variable of interest is independently and identically distributed (iid). Under the null hypothesis, the BDS statistic is obtained by

$$V(T,m,\varepsilon) = \sqrt{T}\frac{C(T,m,\varepsilon) - C(T,1,\varepsilon)^m}{\hat{\sigma}(T,m,\varepsilon)} \quad (3)$$

The correlation integral asymptotically follows standard normal distribution. $\hat{\sigma}(T,m,\varepsilon)$ is the standard sample deviation of $C(T,m,\varepsilon) - C(T,1,\varepsilon)^m$. Moving from the hypothesis that a time series is IID, the BDS tests the null hypothesis that $C(T,m,\varepsilon) = C(T,1,\varepsilon)^m$, which is equivalent to the null hypothesis of iid against an unspecified alternative [11].

Kaplan's test [51] has demonstrated statistical power in determining whether data are linear. This test is based on the concept of continuity in deterministic systems, which establishes that after an

iteration, two points that are initially close will remain close, while if the underlying model is stochastic two points initially close, they may have images that are very distant from each other. The null hypothesis of the test is the linearity of the system versus the nonlinearity. Kaplan's test can be used to test either for nonlinearity or for more focused special cases of nonlinearity [51]. In a general and summarized way, given a vector $y_t \equiv \left(y_t, y_{t-\tau}, \ldots, y_{t-(m-1)\tau}\right)$ embedded in an m-dimensional phase space and obtained from the set of observed data $\{y_t\}_{t=1}^{T}$ and be the image of the point y_t for an fixed positive integer named delay τ, $y_{t+\tau} = f(y_t)$.

Kaplan's technique examines for an embedding dimension m and a delay τ all the distances between two pairs of points $\delta_{i,j} = \left|y_j - y_k\right|$ and their respective images, $\varepsilon_{i,j} = \left|y_{j+\tau} - y_{k+\tau}\right|$ to later calculate the average of the values of $\varepsilon_{j,k}$, conditioned to the corresponding value of $\delta_{i,j}$, that is $E(r) \equiv \overline{\varepsilon_{j,k}}$ for j and k such that $\delta_{j,k} < r$.

$E(r)$ is therefore an average of all the images, whose initial points are very close. For every deterministic system it is obvious that $\lim_{r\to 0} E(r) \to 0$, but for chaotic systems this convergence is not so clear. It is defined then the so-called Kaplan's (K) statistic as: $K = \lim_{r\to 0} E(r)$.

The value of K is expected to be higher in non-deterministic systems, compared to deterministic systems. As mentioned, the null hypothesis of the test is that the underlying model of the data is a linear dynamic system. To perform the test, it needs to work with a K statistic that is calculated for the original series (K test) and a stochastic linear generating model with the same properties (with the same histogram and autocorrelation function) as the original system (KS). If the value of the statistic of the original series exceeds that of the series under the alternative hypothesis of stochastic linearity, the null hypothesis of linearity is accepted. The distribution of the statistic is not tabulated, but Kaplan proposes two maximum levels for testing the null hypothesis of linearity. The first is the minimum KS estimated from the subrogated series, and the second is the mean minus two or three times the standard deviation of all the estimated KS. The null hypothesis of linearity is rejected when the K value calculated for the original reconstructed series is greater than at least one of the two dimensions (called, for simplicity, KS and KSmin respectively).

Despite its simplicity, Kaplan's test has shown to be able to detect a wide spectrum of nonlinearity classes and has also demonstrated statistical power in determining whether data are linear [51].

(iv) Volatility modelling

A wide range of procedures are then applied to the residuals of the selected ARMA model obtained to detect the existence of nonlinearity. If nonlinear dependence is detected, then it might be caused by the existence of a volatility cluster. This being the case, the conditional variance is modeled by fitting ARCH family models (GARCH and EGARCH), using the same method as above to select the best model. After the estimation of both models, the residuals obtained were standardized by means of their conditional standard deviations (from now on GARCH and EGARCH series).

(v) Study of chaotic behavior

Finally, the existence of the chaotic component in each series was studied. To do this, the following methods were applied (Table 2).

- The Correlation Dimension, a measure of the complexity of a dynamic system that allows a deterministic system to be distinguished from a stochastic one [47].
- Lyapunov's Exponent, which analyzes in the system the property of sensitivity to initial conditions [14].
- The 0/1 test [46] that provides a value close to zero or one, indicating the latter value existence of chaos in the series.
- MGRM test [6], which has the main advantage of not needing to build the attractor to assess the existence of chaos.
- The Recurrence Plots, which is a visual tool that allows to analyze the existence of periodic patterns, among other aspects of the time series [48].

3.2. Tools to Detect Chaotic Regime

To carry out many of the techniques used in this research it is necessary to reconstruct the attractor. Following Barkoulas et al. [50], given a discrete dynamic system of the form

$$x_t = F(x_{t-1}), x \in \mathbb{R}^n \tag{4}$$

such that $F : U \to R^n$ is a function and U an open subset of $\mathbb{R}^n$ [52], a closed invariant set $A \subset U$ is an attracting limit set of U if there is an open neighborhood V of A, such that the limit set of iterates is A, $\forall x \epsilon V$ when $t \to \infty$.

Empirically, it is often observed as a series of scalar observations y, which represent the multidimensional system in Equation (4). In order to recover the dynamics of the system (i.e., original trajectory) by analyzing the observed time series y_t, the Takens [53] embedding theorem is used, which is presented below.

An m-dimensional vector is defined and constructed from the observed time series.

$$y_t^m = (y_t, \ldots, y_{t+m-1}) = \left(g(x_t), \ldots, g\left(F^{m-1}(x_t)\right)\right) \equiv I_m \tag{5}$$

where F^{m-1} is the composition of F with itself $m-1$ times. The idea is to reconstruct the state dimension space by expanding the one dimensional signal y_t into an m-dimensional phase space, where each observation in the signal y_t is replaced by the vector y_t^m in Equation (5).

Takens' embedding theorem states that for each pair (F, g) the map $I_m : R^n \to R^m$ will be an embedding for $m \geq 2n + 1$. This guarantees the existence of difeomorphisms between the original and the reconstructed attractor as long as the embedding dimension m is sufficiently large with respect to the dimension of the attractor. In short, Takens' theorem assures that both attractors can be considered to represent the same dynamical system in different coordinate systems when $m \geq 2n + 1$.

The main characteristics of the tools for detecting a chaotic regime used in this research are described below.

3.2.1. Correlation Dimension

The Correlation Dimension arises as a response to the problem of estimating the dimensions to characterize a chaotic phenomenon. When reconstructing, it is necessary to know the number of dimensions, m, or the embedding dimension, both to make its representation in the phase diagram and to estimate a simple model of the phenomenon. An important characteristic of chaotic attractors is their dimension, which is defined as the lower limit of the number of state variables (degrees of freedom) needed to describe steady-state behavior. The correlation dimension test is a topologic one which measures a quantity called correlation dimension. It distinguishes chaotic series from random series by investigating the correlation dimension behavior of the data [30].

Let us consider the series $\{y_t\}_{t=1}^n$ and, from this, the sequence of $N = n - m + 1$ m-dimensional vectors, $y_t^m = (y_t, \ldots, y_{t+m-1})$ that gives the reconstructed series.

To estimate the dimension of the reconstructed attractor, the algorithm of Grassberger and Procaccia [47] is used, which is based on the correlation integral given by the following expression:

$$C(\epsilon) = \frac{2}{N(N-1)} \sum_{t<s} H\left\{\epsilon - \left\|X_t^m - X_s^m\right\|\right\} \tag{6}$$

where $\|.\|$, m and H represent respectively a norm operator, the embedding dimension and the Heaviside function. There are several norms that can be used to measure the distance between two different state vectors, such as the Euclidean norm or the maximum norm.

To determine the correlation dimension, it is needed to determine how $C(\varepsilon)$ changes as ϵ changes. As ϵ grows, the value of $C(\varepsilon)$ grows because the number of near points to be included increases. Grassberger and Procaccia [47] showed that for sufficiently small ε, $C(\varepsilon)$ can be well

approximated by $C(\varepsilon) \sim \varepsilon^{v}$). In other words, when $\varepsilon \to 0$, $C(\varepsilon)$ grows at rate v where v is the value of the correlation dimension (CD). The estimate of v when $m \to \infty$, provides the correlation dimension (CD). The dimension of a dynamic system is determined by estimating the slope of the regression of $ln\ C(\varepsilon)$ versus log ϵ and an intercept for small values of ε and depends on the chosen embedding dimension. If the data are purely stochastic, the correlation dimension will equal m for all m. If the data are deterministic, the estimated slope will stabilize at one point, not increasing as m increases. This "saturation" of the slope is the estimated correlation dimension for the unobserved process, which underlies the process that generated the data. That is, if the dynamic system is chaotic, $C(\varepsilon)$ will stabilize at some point D, as m grows.

3.2.2. Lyapunov

Lyapunov's exponents measure the average rate of divergence or convergence between two adjacent orbits, i.e., they quantify the sensitivity to initial conditions in the phase space, identifying the basic attribute of deterministic chaos. It is approached through an exponential function, in which the exponent determines the rate of divergence of adjacent orbits that start from close points in an infinitesimal way. Chaotic systems exhibit a positive coefficient λ and systems that are stable, a negative coefficient.

The Lyapunov exponent is one of the most employed techniques to assess the presence of chaotic behavior in time series [16]. Specifically, the largest Lyapunov exponent can be used to measure the rate of separation of closed trajectories and estimate the overall degree of chaos of a nonlinear dynamical system [54].

There are many algorithms in the literature for the estimation of this coefficient. The algorithm proposed by BenSaïda and Litimi [14] is followed in this research. Specifically, it has been empirically proven that the methodology adopted in this article to determine the Lyapunov exponent (λ) is the one that behaves best for noisy series [55]. Indeed, to estimate λ from experimental or observational data, there are two main classes of methods, both of which are based on reconstructing the space state by the delay coordinates methods. The direct methods are based on the calculation of the growth rate of the difference between two trajectories with an infinitesimal difference in their initial conditions [56]. Among the limitations of this method is that it cannot accept measurement errors or noise [57]. On the contrary, the Jacobian-based approach can give consistent estimates of the Lyapunov exponents even in the presence of noise [58,59]. The Jacobian method is based on nonparametric regression to estimate the Jacobians and λ. It consists of computing the Jacobian matrix of the chaotic map. However, for a scalar time series, the map generating the process is usually unknown; as a result, the Jacobian matrix could not be estimated, and the Lyapunov exponent cannot be computed. For that purpose, it is needed to approximate the unknown chaotic map with a known function. McCaffrey et al. [58] compared several alternatives: thin-plate splines, neural nets, radial basis functions and projection pursuit. Based on the simulations performed by them, neural net was the best regression method for chaotic systems with noise. In this line, Bailey et al. [59] proposed a regression method which involves the use of neural networks. Simulation results for a noisy Henon system suggest that the neural net regression method yields accurate estimates values of the Lyapunov exponent.

In order to shrink the noise in a dynamical system, a wavelet-based denoising method to filter the data has been employed by several researchers. In particular, the theory of signal denoising using wavelets has been developed by Donoho and Johnstone [60]. Garcin and Guégan [61] adapted the theory for signals in which the noise influence is nonlinear and the wavelet transform-based detection of chaos has been proposed by Rubežić et al. [62]. While this approach could be appropriate for physical systems where noise is an intruder of the real pure signal, for financial data, where noise is an inherent property to markets, denoising the data could modify some of the stylized financial facts that have been discussed earlier in the paper and alter the true dynamics that underlie the time series to be tested [63]. Hence, following this reasoning, a neural network approach has been chosen in this research.

Briefly, the methodology consists in the following steps: let us consider a time series $\{y_t\}_{t=1}^{T}$ represented as follows:

$$y_t = f(y_{t-L}, y_{t-2L}, \ldots, y_{t-mL}) + \varepsilon_t \tag{7}$$

where L, m, ε and f stand for the time delay, the embedding dimension, noise added to the series and an unknown chaotic map, respectively and t is the time script. The Lyapunov exponent (LE) is defined as:

$$\hat{\lambda} = \frac{1}{2M} \ln(v_1) \tag{8}$$

where the "block length" M is the number of evaluation points used for estimating the Lyapunov exponent which stands for an arbitrarily selected number of observations and v_1 is the largest eigenvalue of the matrix $(T_M U_0)'(T_M U_0)$, with $T_M = \prod_{t=1}^{M-1} J_{M-t}\, J_t$, the Jacobian matrix of the chaotic map and $U_0 = (1, 0, \ldots, 0)'$.

Because f is usually unknown, it is needed to approximate the Jacobian matrix. The authors employ a single-layer feed-forward neural network using nonlinear least squares for different values of $m = 1 \ldots 8$ and later calculate the LE spectrum. Hence, the chaotic system is estimated by the following equation:

$$y_t \approx \alpha_0 + \sum_{j=1}^{q} \alpha_j \tan h\left(\beta_{o,j} + \sum_{i=1}^{m} \beta_{i,j} y_{t-iL}\right) + \varepsilon_t \tag{9}$$

(L, m q) are selected as the triplet that provides the highest value for λ and are associated with the complexity of the system. The test for chaos is then constructed based on the asymptotic distribution of λ [56].

3.2.3. 0/1 Method

The 0/1 method for chaos was developed to distinguish between regular and chaotic dynamics in deterministic dynamical systems. This tool for chaos was proposed by Gottwald and Melbourne [46] and has a series of advantages over other tests, such as the fact that it is not necessary to reconstruct the phase space or that the result, which is 0–1, is very easy to interpret. Also, rather than requiring phase space reconstruction which is necessary to apply standard Lyapunov exponent methods to the analysis of discretely sampled data, the technique works directly with the time series and does not involve any preprocessing of the data.

The input of the test is a one-dimensional time series $y(t)$ for $t = 1, 2, 3, \ldots$ The data $y(t)$ is used to drive the following two-dimensional system:

$$p(t) = \sum_{j=1}^{t} y(j) \cos(y(j))\ t = 1, 2, 3, \ldots \tag{10}$$

$$q(t) = \sum_{j=1}^{nt} y(j) \sin(y(j))\ t = 1, 2, 3, \ldots \tag{11}$$

Define the (time-averaged) mean square displacement (MSD)

$$M(t) = \lim \frac{1}{N} [p(j+t) - p(j)]^2\ t = 1, 2, 3, \ldots \tag{12}$$

If the system is chaotic, then $M(t)$ will grow linearly over time. If the system is not chaotic, $M(t)$ will be bounded. The asymptotic growth rate of MSD is defined as:

$$K = \lim_{t \to \infty} \frac{\log M(t)}{\log (t)} \tag{13}$$

K can be determined numerically by a linear regression of log ($M(t)$) versus log (t). Under general conditions, the limits $M(t)$ and K can be shown to exist, and K takes either the value $K = 0$ signifying regular dynamic or the value $K = 1$ signifying chaotic dynamic.

3.2.4. MGRM Test

The MGRM Test is explained in detail in [6]. As the authors of the method argue, it is a test for the deterministic process as opposed to the stochastic one which is based on symbolic dynamics and entropy. The MGRM test is based on the concept of permutation entropy, which has its roots in symbolic dynamics. The basic idea behind symbolic dynamics consists of dividing the phase space into a finite number of regions and labeling each region with an alphabetical symbol. A relevant property of symbolic dynamics is that essential features of the underlying dynamics, such as its deterministic or stochastic nature or its complexity, are preserved. Likewise, entropy accounts for the unpredictability of the system under study, which is a crucial feature of complex systems.

In the first instance the notation is introduced: let $\{y_t\}_{t\in T}$ be a stationary time series, $\{y_t\}_{t\in I}$ an observation with $I = \{1,\ldots,T\}$ and m the embedding dimension with $m \geq 2$. Ordinal patterns will be defined for "m". To that end, the scalar time series is embedded to an m-dimensional space: $Y_m(t) = \left(y_t, y_{t+1}, \ldots, y_{t+(m-1)}\right)$ *for* $t \in T$. The ordinal pattern of embedding dimension m, at a given time t is defined as the unique permutation $\pi_m(t) \equiv (r_0\, r_1 \ldots r_{m-1})$ of the set {0, 1, ..., $m-1$} satisfying:

$$y_{t+r_0} \leq y_{t+r_1} \leq \cdots \leq y_{t+r_{m-1}} \tag{14}$$

$$r_{s-1} < r_s \text{ if } y_{t+r_{s-1}} = y_{t+r_s} \tag{15}$$

By means of (15) it is guaranteed the uniqueness of the permutation defined by (14). So, the vector or m-history $Y_m(t)$ is converted into a unique *symbol* $\pi_m(t)$. In fact, $\pi_m(t)$ describes how the order of the dates: $t+0 < t+1 < \ldots < t+(m-1)$ is turned into the order of the corresponding analyzed values. The basic idea is to divide naturally the state space in which the dynamics takes place into a finite number of partitions using the time-dependent information contained in the m-history $Y_m(t) \in R^m$. According to the previous definition, partitions depend on the ordinal structure of the m-history. In particular, $\pi_m(t) = \pi_m(s)$, $s \neq t$, if and only if for all $k, I\, \epsilon\, \{0,1,\ldots,m-1\}$ with $k \neq I$ it holds that $y_{t+1} \leq y_{t+k} \leftrightarrow y_{s+1} \leq y_{s+k}$.

In general, given a time series $\{y_t\}_{t\in T}$, all $m!$ permutations of order m are considered here as possible order types of m different numbers. Then the relative frequency or unconditional success probability $p(\pi)$ of each symbol or permutation π for a given time series and an embedding dimension parameter m exists and it can be defined as:

$$p(\pi) = \frac{card\{t \,\big|\, 0 \leq t \leq T-(m-1),\ Y_m(t)\ has\ type\ \pi\ \}}{T-m+1} \tag{16}$$

Given an embedding dimension $m \geq 2$, modified Shannon entropy that stands for the $m!$ distinct symbols is defined as follows:

$$h(m) = -\sum_{i=1}^{m!} p(\pi_i) \log p(\pi_i) \tag{17}$$

The test is constructed as follows: Let m be embedding dimension and T be the number of observations and fix $w, k \in \mathbb{N}$ such that $w = \frac{m!}{k}$. Next, the subsets W_j that must verify $W_1 \subseteq W_2 \subseteq \cdots \subseteq W_n$ in the following way $W_j = W_{j-1} \cup \left\{ w\ symbols\ chosen\ at\ random\ in\ S_m \setminus W_{j-1} \right\}$ for $j = 2, \ldots, k$. are constructed.

Subsequently the next modified permutation entropy's function is calculated as:

$$h^{W_j}(m) = -\sum_{\pi_j \in w_j} p(\pi) \log p(\pi) \tag{18}$$

The authors of the method showed that by analyzing the modified permutation entropy's function values it is possible to distinguish and identify deterministic systems. In this process no more information is gained by increasing the number of symbols under consideration. In contrast, in non-deterministic process, this information or complexity is increased.

This latter property is tested in the following way: Let $dh^{W_j}(m)$ be the $h^{W_j}(m)$ *slope*:

$$dh^{W_j}(m) = \frac{h^{W_{j+1}}(m) - h^{W_j}(m)}{\log\left(\frac{j+1}{j}\right)} \tag{19}$$

When considering random process, the numerical slope of permutation entropy will increase with *(log(jw))*, while this will not hold for chaotic or regular processes.

The property described above that identifies deterministic systems $h^{W_j}(m)$ is checked by carrying out the following regression:

$$dh^{W_j}(m) = \alpha_0 + \alpha_1 j + \varepsilon_j \; for \; j = 1, 2, \ldots, \; k-1 \tag{20}$$

where ε_j is white noise. As a result, the estimated parameter $\hat{\alpha}_1$ can be used to evaluate $dh^{W_j}(m)$ increases with j. In mathematic notation, the test is the following one:

$$\begin{array}{l} H_0 \equiv \alpha_1 = 0 \\ H_1 \equiv \alpha_1 > 0 \end{array} \tag{21}$$

Indeed, regression (20) can be considered as a simple symbol-trend model. As in the simple time-trend model, the ordinary least squares (OLS) estimate $\hat{\alpha}_1$ is so that asymptotically the usual t-test of H_0: $\hat{\alpha}_1 = 0$ is valid. If the coefficient obtained is null, then the series is deterministic; otherwise (greater than 0), it is stochastic [6].

3.2.5. Recurrence Plots

In addition to the analytical methods described above, there are other methods for detecting chaos and other patterns of a more visual nature, the so-called recurrence plot (RP) being one of the most widely used.

The Recurrence Plot (RP) is an analysis tool that reveals the existence of recurrent and intermittent patterns in time series. First proposed by Ekmann et al. [64], it has been widely applied in the characterization of dynamic systems. This topological method shows the hidden structures of time series from a qualitative point of view. The plots are constructed by assuming mutual distances that belong to the same path in the reconstructed phase space [17].

The construction is done as follows:

Let m be the embedding dimension and y_t^m the vector m-dimensional in the reconstructed phase space in time t = 1, 2, . . . , k

The recurrence matrix is generated by comparing each embedded vector y_i^m with the other y_j^m. A point is drawn if this comparison is less than a value ε for a specific distance. That is, if the condition is met: $\left\| y_i^m - y_j^m \right\| < \varepsilon$.

Recurrence of a moment of state i at a different time j occurs when y_j^m is close enough to y_i^m. The RP can therefore show which vectors are close together and which are far apart.

The diagonal structures of the RP identify the range in which a fraction of the path is relatively close to another at a different time. For a deterministic chaos system, small lines are observed parallel to the main diagonal. However, in random systems, small line segments are absent, and evenly distributed points are shown.

This technique is independent of some constraints such as sample size, noise and stationarity [11]. It also provides additional information about the structure of the attractor, since the plot preserves the temporal order of the series, allowing the place and periodicity of the periodic orbits to be known.

4. Empirical Results

Applying the methodological framework described above, the main results obtained are shown below.

4.1. Data

Four different daily time series are analyzed, using the closing prices of the following major stock market indices: the Dow Jones Industrial Average (Dow Jones), the Ibex 35 (Ibex), the Nasdaq 100 (Nasdaq) and the Nikkei 225 (Nikkei). All the data were obtained from Yahoo Finance and covers the period from 1 January 1992 until 31 July 2013. Figure 1 shows the evolution of the series over the sample period.

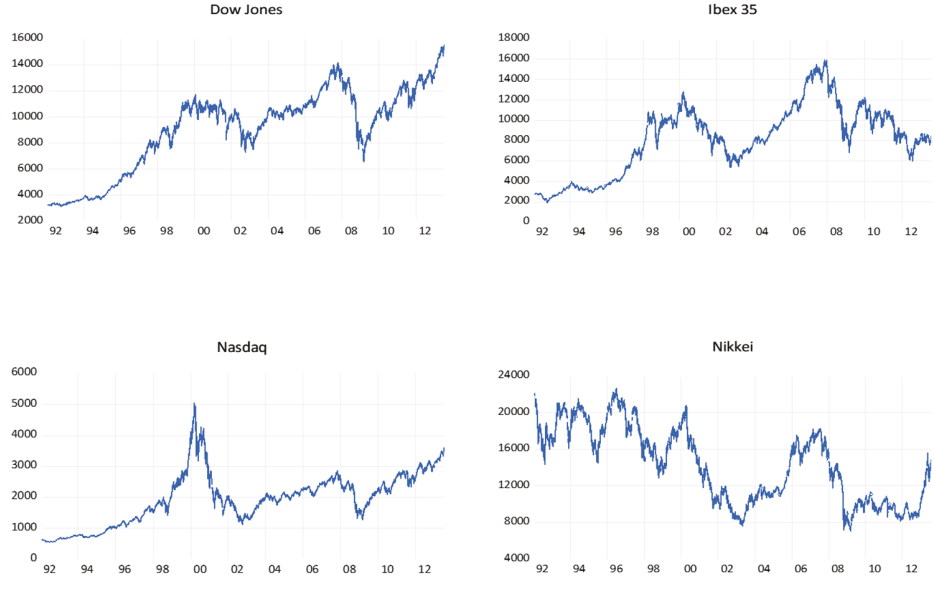


Figure 1. Plot of stock indices series. Time points (year) are on x-axis and observations are on y-axis. Source: Yahoo finance.

Table 3 presents the main descriptive statistics of the series studied. Both the series of the main stock market index of the U.S. Stock Exchange and that of the Spanish Stock Exchange show a negative asymmetry coefficient (−0.51 and −0.05, respectively), unlike the other financial series considered in the study. Moreover, it can be concluded that the data do not come from a normal distribution, as reflected by the Jarque–Bera test (p-value < 0.05).

Table 3. Descriptive statistics.

	Original Series	Returns	ARMA	GARCH	EGARCH
			DOW JONES		
Mean	9108.07	0.0003	-5.8×10^{-7}	−0.0356	0.0028
Standard Deviation	3147.61	0.0112	0.0112	0.9994	0.9999
Median	10,080.3	0.0005	0.0004	0.0017	0.0350
Minimum	3136.58	−0.0821	−0.0790	−6.5989	−6.0801
Maximum	15,567.7	0.1051	0.0993	3.4852	3.7538
Skewness Coefficient	−0.5103	−0.1364	−0.2810	−0.4213	−0.3715
Kurtosis Coefficient	2.2721	11.4338	10.7387	4.4806	4.2585
Jarque–Bera Test	354.59 *	16,062.4 *	13,575.8 *	654.4 *	481.64 *
N	5415	5414	5412	5412	5412
			IBEX		
Mean	8124.38	0.0002	4.86×10^{-8}	−0.0407	−0.0063
Standard Deviation	3354.74	0.0146	0.0145	0.9989	1.0001
Median	8362.9	0.0007	0.0005	−0.0158	0.0166
Minimum	1873.58	−0.0959	−0.0965	−6.5618	−7.7301
Maximum	15,945.7	−0.1348	0.1335	5.9752	5.2568
Skewness Coefficient	−0.0526	−0.0054	−0.0428	−0.2057	−0.1960
Kurtosis Coefficient	2.3074	7.8791	7.7376	4.4174	4.4775
Jarque–Bera Test	110.96 *	5382.03 *	5076.16 *	492.48 *	528.23 *
N	5427	5426	5426	5426	5426
NASDAQ					
Mean	1950.235	-7.1910^{-5}	-2.30×10^{-7}	−0.0316	0.0043
Standard Deviation	842.569	0.0149	0.0156	0.0351	0.9999
Median	1989.22	-9.3410^{-5}	0.0009	−6.1386	0.0706
Minimum	547.84	−0.1195	−0.1018	4.1826	−6.1143
Maximum	5048.62	0.1366	0.1265	0.9994	3.5622
Skewness Coefficient	0.3527	−0.0584	−0.0979	−0.3825	−0.4038
Kurtosis Coefficient	3.0323	12.7954	8.4806	4.0381	4.0070
Jarque–Bera Test	112.519 *	22,759.76 *	6782.03 *	374.96 *	375.74 *
N	5415	5414	5412	5412	5412
NIKKEI					
Mean	14,304.3	-9.122×10^{-5}	-5.22×10^{-8}	−0.0357	−0.0045
Standard Deviation	4073.77	0.0154	0.0154	0.9997	1.0001
Median	14,416.6	6.3110^{-5}	0.0002	−0.0160	0.0129
Minimum	7054.98	−0.1211	−0.1204	−6.4953	−6.7925
Maximum	22,667	−0.1323	0.1282	5.6944	5.7235
Skewness Coefficient	0.1026	−0.2147	−0.2397	−0.1943	−0.1597
Kurtosis Coefficient	1.7032	8.2470	8.2517	4.4910	4.5000
Jarque–Bera Test	379.75 *	6104.3 *	6125.1 *	522.85 *	518.01 *
N	5287	5286	5286	5286	5286

Notes: Statistics for the time series considered in this study: original series, return series, and residuals of autoregressive moving average (ARMA), generalized autoregressive conditional heteroskedasticity (GARCH) and exponential GARCH (EGARCH) models. Skewness and kurtosis coefficient corresponds to the Fisher asymmetry coefficient and the kurtosis coefficient, respectively. *: The asterisk denotes significance at 5% level. Source: Own work.

4.2. Returns

After testing that the series are not stationary (see Table 4), the returns from each series were obtained, by considering the logarithms of the ratio of two consecutives prices. Figure 2 shows the time development of the returns. They show an excess of kurtosis, a negative skewness coefficient (Table 3) and more fat-tails in comparison with those from a normal distribution. These characteristics are typical of financial returns [2]. Excessive kurtosis means that returns far from the average are more common than in a normal distribution and therefore the investor is subject to greater risk. The Jarque–Bera test

indicates departure from the normal distributions for all the cases. According to the results from the Augmented Dickey–Fuller, Philips–Perron and KPSS tests, all the returns series are stationary (Table 4).

Table 4. Stationarity analysis.

	Index	**Returns**
DOW JONES		
1. Augmented Dickey Fuller Test		
Constant	−1.0288 (0.7451)	**−56.0490 (0.0001)**
Constant and Linear Trend	−2.2115 (0.4825)	**−56.0525 (0.0000)**
2. Phillips–Perron Test		
Constant	−0.9623 (0.7685)	**−78.6723 (0.0001)**
Constant and Linear Trend	−2.1111 (0.5389)	**−78.6832 (0.0001)**
3. Kwiatkowski–Phillips–Schmidt–Shin Test		
Constant	7.1098	**0.1793**
Constant and Linear Trend	1.1008	**0.0832**
IBEX		
1. Augmented Dickey Fuller Test		
Constant	−1.8643 (0.3496)	**−53.5682 (0.0001)**
Constant and Linear Trend	−1.6676 (0.7655)	**−53.5897 (0.0000)**
2. Phillips–Perron Test		
Constant	−1.8167 (0.3727)	**−71.1395 (0.0001)**
Constant and Linear Trend	−1.5767 (0.8023)	**−71.1609 (0.0000)**
3. Kwiatkowski–Phillips–Schmidt–Shin Test		
Constant	5.1229	**0.2791**
Constant and Linear Trend	1.0088	**0.0543**
NASDAQ		
1. Augmented Dickey Fuller Test		
Constant	−1.3869 (0.5904)	**−54.9507 (0.0001)**
Constant and Linear Trend	−2.0781 (0.5574)	**−54.9485 (0.0000)**
2. Phillips–Perron Test		
Constant	−1.3869 (0.5904)	**−74.6870 (0.0001)**
Constant and Linear Trend	−2.0781 (0.5574)	**−74.6836 (0.0001)**
3. Kwiatkowski–Phillips–Schmidt–Shin Test		
Constant	0.0860	**0.1198**
Constant and Linear Trend	0.6694	**0.0861**
NIKKEI		
1. Augmented Dickey Fuller Test		
Constant	−2.3032 (0.1710)	**−75.6134 (0.0001)**
Constant and Linear Trend	−2.5121 (0.3222)	**−75.6142 (0.0001)**
2. Phillips–Perron Test		
Constant	−2.3485 (0.1568)	**−75.7549 (0.0001)**
Constant and Linear Trend	−2.5440 (0.3066)	**−75.7590 (0.0001)**
3. Kwiatkowski–Phillips–Schmidt–Shin Test		
Constant	5.5059	**0.0950**
Constant and Linear Trend	0.4156	**0.0491**

Augmented Dickey–Fuller *p*-values correspond to MacKinnon [65] one-sided *p*-values. For the Phillips–Perron test, lags were based on bandwidth Newey–West using Bartlett kernel. Critical values for the Kwiatkowski, Phillips, Schmidt and Shin test are 0.463 and 0.146 respectively for the constant and linear plus linear trend model. Significant values at the 95% confidence level are in bold.

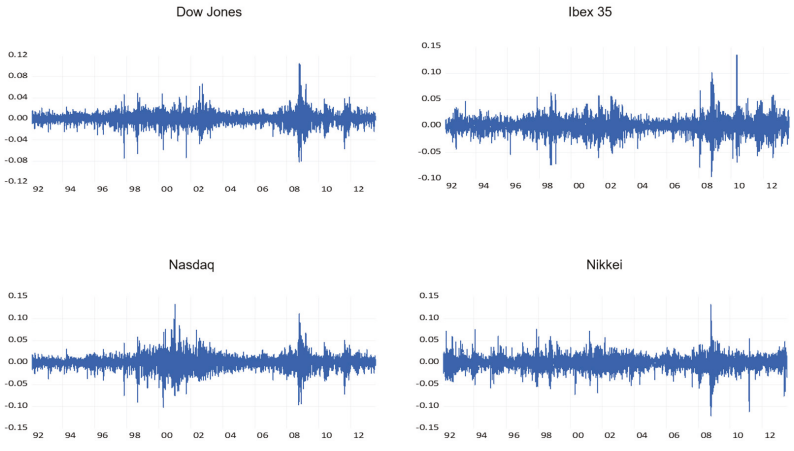


Figure 2. Plot of stock indices returns series. Time points (year) are on x-axis and observations are on y-axis. Source: Own work.

4.3. Linear Dependence

As some methodology (for example the BDS test) is not robust to the presence of linear relationships, the linear dependence was removed using ARMA models. Following the criterion described by Barkoulas et al. [51], the best ARMA model was selected for each series (Table 5). The selected fitted models for each of the series (ARMA (2,5); ARMA (0,3); ARMA (2,2) and ARMA (0,1), for the series Dow Jones, Ibex, Nasdaq and Nikkei respectively) are shown in Table 5 that shows all the models fitted in this research. Details of each model are presented in Supplementary Materials A (Tables SA1, SA4, SA7 and SA10). According to the results from the Augmented Dickey–Fuller, Philips–Perron and KPSS tests, applied on the residuals of the selected model and shown in Table 4, all the residuals series (ARMA) are stationary.

Table 5. Models.

Index	Linear Model	Nonlinear Model	
	[ARMA]	[GARCH]	[EGARCH]
Dow Jones	ARMA (2,5)	GARCH (2,1)	EGARCH (2,1)
Ibex	ARMA (0,3)	GARCH (2,1)	EGARCH (2,1)
Nasdaq	ARMA (2,2)	GARCH (2,1)	EGARCH (2,3)
Nikkei	ARMA (0,1)	GARCH (1,1)	EGARCH (1,1)

4.4. Nonlinear Dependence

Next, Keenan, Tsay, Teräsvirta, White, BDS and Kaplan methods were used to study nonlinearity. The principal advantage of using a broad set of instruments is to obtain the most information possible about the nature of the series, since to date none of the methods has proved successful in detecting all types of dependence. Nevertheless, some tests, such as the BDS, White or Kaplan, are more efficient in detecting dependency [51].

The BDS test was applied once the linear dependence was eliminated, as previously suggested by Brock et al. [51]. In this way the BDS test served as an indirect method of analyzing nonlinearity: if it rejected the null hypothesis, then the series was proven to be nonlinear. This procedure is theoretically feasible, since the calculations on the residuals of an autoregressive model do not lose the relevant information derived from the original series if the latter comes from a nonlinear chaotic system [51].

The BDS test is applied taking into account the different embedding dimensions (range: 2–9). The results are presented in Supplementary Materials B (Tables SB3–SB14).

Kaplan's method is applied to a series in which the functional form that generated the series is unknown. The method's aim is to determine whether there is evidence of an underlying deterministic mechanism, i.e., that the hypothesis that is tested is the stochastic linearity of the process. It is considered that the null hypothesis is verified if the value of the statistic K, calculated for the original series (K test), is smaller than the values obtained for the surrogates of each series, defined as the smallest value (Kmin) between the minimum and the mean, minus two standard deviations (KS). Several different values for both the delay and the embedding dimension parameters were taken into account, and the conclusions did not differ (see Supplementary Materials B).

Table 6 presents a summary of results for the nonlinear analysis. All the procedures used, except for Kaplan, suggest the existence of nonlinearity in the Dow Jones Index ARMA series. A similar pattern is observed for the Nasdaq Index. Further, as all the tests indicate some nonlinear dependence in the Ibex and Nikkei indexes, the data seems to present some type of dependence, since the BDS test shows significant results for the different embedding dimensions and the epsilons considered. For the latter, the White and Teräsvirta tests confirm the existence of dependence.

Table 6. Summary of results for the nonlinear analysis.

Dow Jones					
Test	Ho	Returns	ARMA (2,5)	GARCH (2,1)	EGARCH (2,1)
Runs	Randomness	R	NR	R	R
Keenan	La	NR	R	NR	NR
Tsay	La	R	R	NR	NR
White	La	R	R	NR	NR
Teräsvirta	La	R	R	R	R
BDS	Lb	-	R	NR	Mixture
Kaplan	L	NR	Mixture	Mixture	NR
Ibex					
Test	Ho	Returns	ARMA (0,3)	GARCH (2,1)	EGARCH (2,1)
Runs	Randomness	NR	NR	NR	NR
Keenan	La	NR	NR	NR	NR
Tsay	La	R	NR	R	NR
White	La	NR	NR	NR	NR
Teräsvirta	La	R	R	NR	NR
BDS	Lb	-	R	NR	Mixture
Kaplan	L	Mixture	Mixture	NR	Mixture
Nasdaq					
Test	Ho	Returns	ARMA (2,2)	GARCH (2,1)	EGARCH (2,3)
Runs	Randomness	R	R	R	R
Keenan	La	R	NR	NR	NR
Tsay	La	R	R	R	R
White	La	R	R	R	R
Teräsvirta	La	R	R	R	R
BDS	Lb	R	R	Mixture	Mixture
Kaplan	L	R	Mixture	NR	NR
Nikkei					
Test	Ho	Returns	ARMA (0,1)	GARCH (1,1)	EGARCH (1,1)
Runs	Randomness	R	NR	R	R
Keenan	La	NR	NR	R	NR
Tsay	La	R	R	NR	NR
White	La	R	R	NR	NR
Teräsvirta	La	R	R	NR	NR
BDS	Lb	-	R	Mixture	Mixture
Kaplan	L	NR	NR	Mixture	Mixture

Ho: Null hypothesis. La: Linear in mean. Lb: This test can be applied as a linear test once the linear dependence has been eliminated. R represents that the null hypothesis is rejected; NR represents that the null hypothesis is not rejected. Mixture represents that the test reported no clear results. For the Kaplan method a total of 30 surrogates were generated from the original series, considering the same embedding dimensions and delay parameters.

4.5. Volatility Clustering

The next step consists of modelling the conditional variance by fitting ARCH family models (GARCH and EGARCH). In first instance, the best GARCH (p,q) model was selected so as to adhere to the previously described criterion [51]. The selected models were GARCH (2,1) for the Dow Jones, Ibex and Nasdaq indexes and the GARCH (1,1) for the Nikkei Index (Table 5). After fitting each model, standardized residuals were obtained (from now on GARCH and EGARCH series). Unlike GARCH models, EGARCH models can be used to estimate the conditional variance, taking into consideration the sign of the innovation of the previous period. These types of models successfully capture the asymmetric response in the conditional variance and, hence, are suitable candidates for modeling financial processes. The chosen models were EGARCH (2,1), EGARCH (2,3), and EGARCH (1,1) for the Dow Jones and Ibex indexes, the Nasdaq Index and the Nikkei series, respectively (Table 5 and Supplementary Materials A). The standardized residuals of all the series are less leptokurtic (with an average value of 4) than those obtained from the ARMA models (Table 3).

Since nonlinearity is a necessary, but not sufficient, condition for chaotic behavior, its existence was first analyzed in residuals of the volatility models (Table 6 and Supplementary Materials B). In general, linear hypothesis is not rejected, but nonlinearity was found in some cases. Indeed, the BDS test shows some remaining dependence in the Dow Jones Index, which is compatible with the existence of chaos in the EGARCH model residuals. The Tsay and BDS tests reject linearity for the Ibex Index in the GARCH and EGARCH models, respectively. Likewise, some procedures show evidence of nonlinearity in the case of the GARCH and EGARCH models for the Nasdaq series. Finally, for the Nikkei Index, the BDS rejects the hypothesis of independence for all the cases considered.

4.6. Chaotic Behavior Analysis

The following methods were used to test the existence of chaotic dynamics. Results are presented in Supplementary Materials C.

Correlation Dimension

The correlation dimension (CD) [48] quantifies the degree of complexity of a system and distinguishes a deterministic system from a stochastic one. The results suggest that, in all cases except for the EGARCH model in the Dow Jones Index, the CD increases as the embedding dimension increases, however, the CD is below the expected value for a random process (Supplementary Materials Table SC1). Moreover, except in the latter series, the saturation of the slope as the embedding dimension increases is not observed. Thus, it is not possible to assure that the series are chaotic. Overall, there is sufficient evidence against the existence of a strange attractor; and if this is the case, it might be of a high dimension.

Lyapunov Test

The Lyapunov exponent reflects the average rate of convergence or divergence of two paths that are, initially, points that are very close in the phase space. Positive values indicate the existence of chaotic dynamics. Here, the algorithm described by BenSaïda and Litimi [14] was used. Negative significant Lyapunov exponents were obtained in all cases (Supplementary Materials C, Table SC2). Thus, the assumption of chaotic behavior was rejected in all cases.

0/1 method

Following the procedure described by the authors of the method, which consists of using several frequencies to increase the degree of robustness of the test, the test was carried out a total of 6000 times and the median of all realizations was taken [46]. The results of the 0/1 method are shown in Supplementary Materials Table SC3. The test provides a value close to 0 or 1. If the value is 0 it is concluded that the series is clearly stochastic, however if the result is 1, it cannot be stated with certainty that the series has a chaotic component, since the result may be due to noise. All the results obtained are close to value 1 (Supplementary Materials Table SC3), so it is concluded that the series may have a lot of noise or may have a significant chaotic component.

MGRM Test

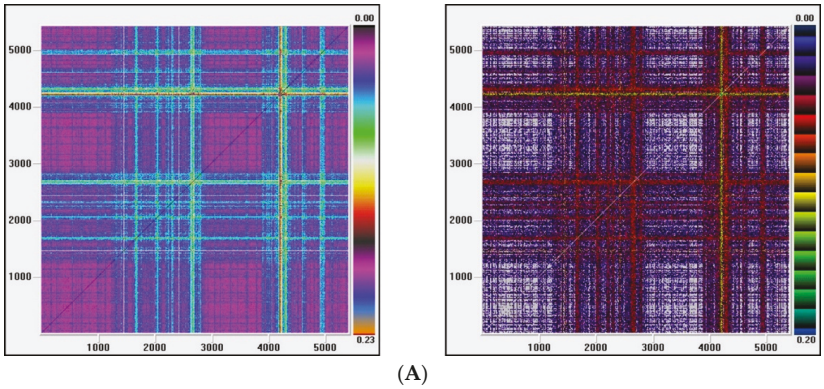
This test is constructed as follows. First the Shannon modified entropy values are calculated. Then, the information or complexity is estimated by a linear regression that used the entropies as the explanatory variable. The coefficient of the entropies is then analyzed as in deterministic series the complexity derived from the entropy of permutation does not increase when the number of symbols increases, once the saturation point is reached. To conduct the test, the parameters were fixed such that m = 4, k = 2, and w = 12. The results were positive in all the cases (Supplementary Materials Table SC4). Thus, the hypothesis of determinism is rejected.

Recurrence plots

Visual Recurrence Analysis is based on Eckmann's [48] definition of a recurrence graph. The degree of complexity and the existence of chaos is analyzed by generating the recurrence plots. This kind of analysis can capture the recurrence property of states, one of the essential properties observed in chaotic systems. The points above the main diagonal, representing the distance between the same vector and each embedded vector in the phase space and therefore segments parallel to the diagonal, would indicate a chaotic behavior [50].

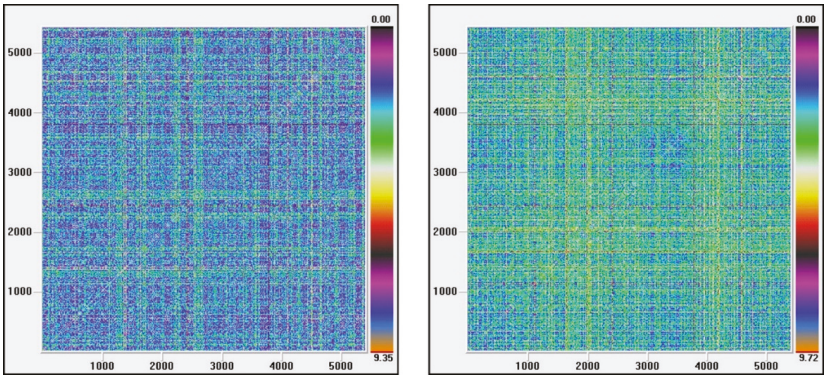
Figures 3–6 show the recurrence plots for the four series considered. For their elaboration, the optimal values of the two parameters, m—the embedding dimension and τ—the considered delay, as previously described, have been considered. Likewise, the Euclidean distance has been chosen and the radius cut-off point (ε) is defined as 10% of the maximum distance between all the points of the reconstructed phase space. This value is adopted in several works, for example in Barkoulas et al. [50].

The recurrence plots can present different patterns, among them the presence of short segments parallel to the main diagonal are related with the existence of chaos. Except the one corresponding to the EGARCH model of the Nikkei stock, none of these patterns is observed in the series evaluated, concluding that the underlying system that generated the series apparently does not have a significant chaotic component.



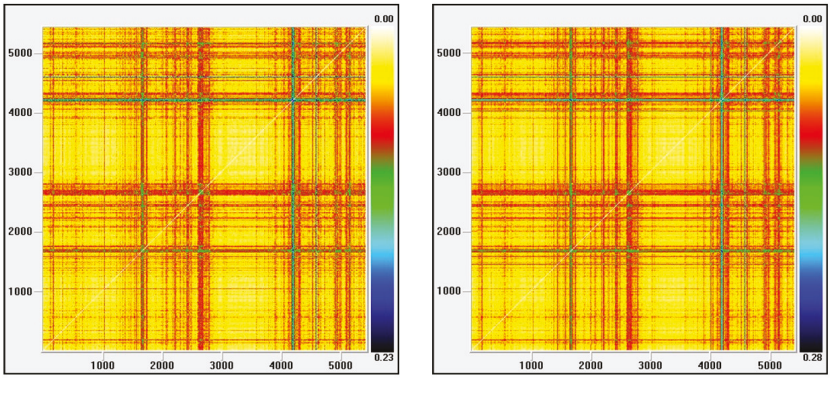
(**A**)

Figure 3. *Cont.*

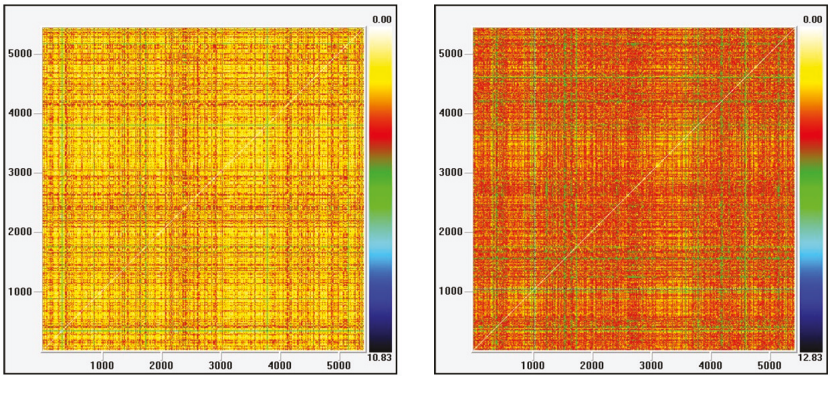


(**B**)

Figure 3. (**A**) Recurrence plot for the returns and ARMA series from the Dow Jones index; (**B**) Recurrence plot for the GARCH and EGARCH series from the Dow Jones index.

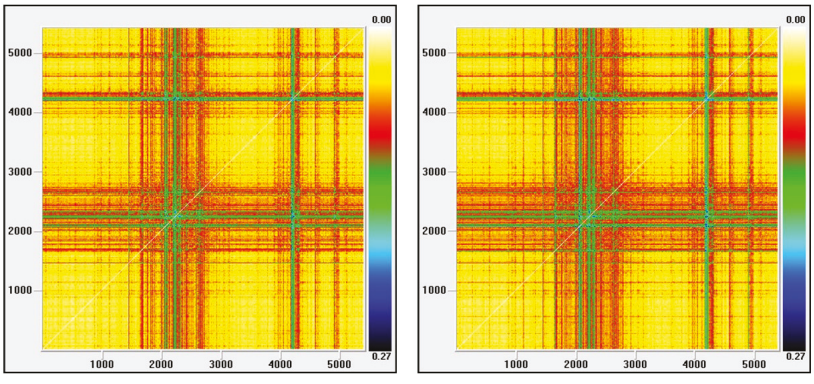


(**A**)

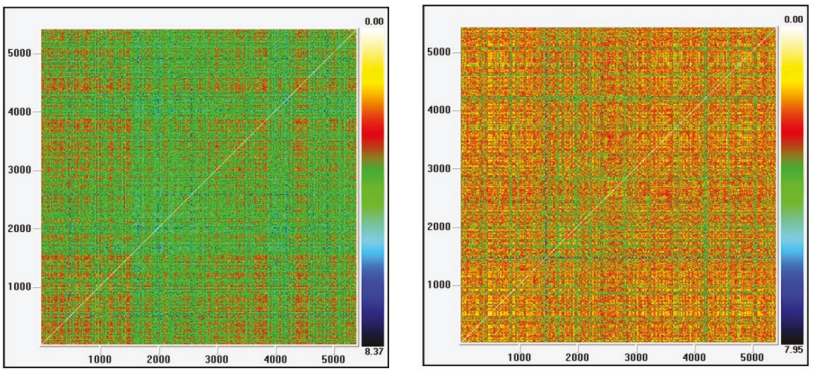


(**B**)

Figure 4. (**A**) Recurrence plot for the returns and ARMA series from the Ibex index; (**B**) Recurrence plot for the GARCH and EGARCH series from the Ibex index.

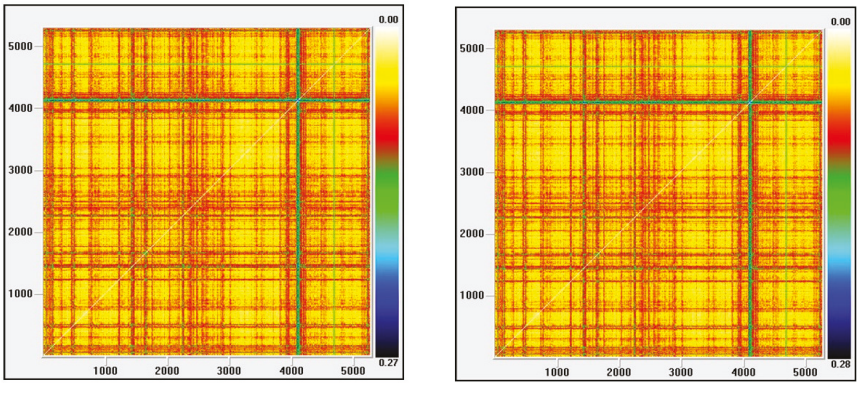


(**A**)



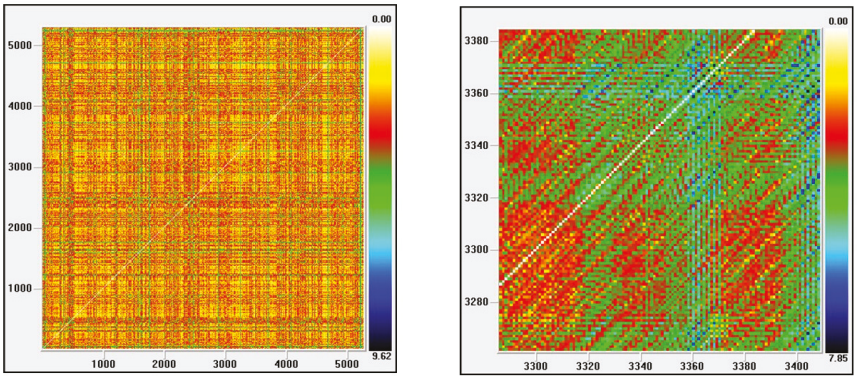
(**B**)

Figure 5. (**A**) Recurrence plot for the returns and ARMA series from the Nasdaq index; (**B**) Recurrence plot for the GARCH and EGARCH series from the Nasdaq Index.



(**A**)

Figure 6. *Cont*.



(**B**)

Figure 6. (**A**) Recurrence plot for the returns and ARMA series from the Nikkei index; (**B**) Recurrence plot for the GARCH and EGARCH series from the Nikkei Index.

5. Conclusions

This study examines the underlying dynamics of four of the principal stock exchange indexes and the asset markets that they represent. To determine the existence of a nonlinear and chaotic regime in the analyzed time series, a comprehensive methodological framework has been adopted that integrates a great number of tools. Thus, seven techniques have been used for nonlinear analysis and five for the analysis of chaotic behavior, including those most suitable for noisy time series. Results support the existence of nonlinearity, which is not consistent with chaos. In addition, GARCH/EGARCH models explain a significant part of the nonlinear structure that is found in the four stock markets analyzed.

The findings are in concordance with the conclusions of other researchers who use suitable procedures for noisy series to detect chaos in stock markets (e.g., [14]). On the other hand, our conclusion contradicts the findings of previous studies that found evidence of low dimensional chaos but used several techniques that do not account for noise (e.g., [66]).

The analysis of the stock exchange indexes carried out in this research is of great interest because the markets they represent are currently a faithful barometer of the evolution of activity in the most developed economies. Likewise, the modelling and study of the existence of nonlinearity and the chaotic dynamics of stock markets is particularly useful for agents involved in capital markets: investors, financial intermediaries, credit institutions, regulators, etc. Knowledge of their dynamics is essential for making accurate predictions of their future evolution and of crucial importance for properly managing the level of risk in the capital market and correctly and efficiently applying prices in the derivatives and futures markets.

Further research is needed to examine other types of both parametric and nonparametric nonlinear models, as well the relationships between the different stock market indices, using a multivariate GARCH model (MGARCH). Also, a new research would consist of the study of the effects of the collapse of stock prices caused by the COVID-19 outbreak on the dynamics of the series analyzed and see how this phenomenon would have affected the results obtained. Other further work could be to carry out the study dividing the total sample into several periods.

Supplementary Materials: The following are available online at http://www.mdpi.com/1099-4300/22/12/1435/s1, A: Details of the Models Applied to the Series; B: Results of the Randomness and NonLinearity tests and C: Chaotic Component Study.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Su, X.; Wang, Y.; Duan, S.; Ma, J. Detecting chaos from agricultural product price time series. *Entropy* **2014**, *16*, 6415–6433. [CrossRef]
2. Fama, D.F. Efficient Capital Markets: A Review of Theory and Empirical Work. In Proceedings of the Twenty-Eighth Annual Meeting of the American Finance Association, New York, NY, USA, 28–30 December 1970; Volume 25, pp. 383–417.
3. Barkoulas, J.; Travlos, N. Chaos in an Emerging Capital Market? The Case of the Athens Stock Exchange. *Appl. Financ. Econ.* **1998**, *8*, 231–243. [CrossRef]
4. Lahmiri, S.; Bekiros, S. Nonlinear analysis of Casablanca Stock Exchange, Dow Jones and S&P500 industrial sectors with a comparison. *Phys. A* **2020**, *539*, 122923.
5. Sarantis, N. Nonlinearities, cyclical behaviour and predictability in stock markets: International evidence. *Int. J. Forecast.* **2001**, *17*, 459–482. [CrossRef]
6. Matilla-Garcia, M.; Ruiz Marin, M. A new test for chaos and determinism based on symbolic dynamics. *J. Econ. Behav. Organ.* **2010**, *76*, 600–614. [CrossRef]
7. Granger, C.W.J. Non-linear Models: Where Do We Go Next—Time Varying Parameter Models? *Stud. Nonlinear Dyn. Econ.* **2008**, *12*, 1–9. [CrossRef]
8. Engle, R.F. Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K. Inflation. *Econometrica* **1982**, *50*, 987–1007. [CrossRef]
9. Bollerslev, T. Generalized Autoregressive Conditional Heteroskedasticity. *J. Econ.* **1986**, *31*, 307–327. [CrossRef]
10. Nelson, D.B. Condicional Heteroskedasticity in Asset Returns: A New Aprproach. *Econometrica* **1991**, *59*, 347–370. [CrossRef]
11. Faggini, M. Chaotic time series analysis in economics: Balance and perspectives. *Chaos* **2014**, *24*, 042101. [CrossRef]
12. Scheinkman, J.; LeBaron, B. Nonlinear Dynamics and Stock Returns. *J. Bus.* **1989**, *62*, 311–337. [CrossRef]
13. Caraiani, P. Is the Romanian Business Cycle Characterized by Chaos? *Rom. J. Econ. Forecast.* **2012**, *3*, 142–151.
14. BenSaïda, A.; Litimi, H. High level chaos in the exchange and index markets. *Chaos Solitons Fractals* **2013**, *54*, 90–95. [CrossRef]
15. McKenzie, M.D. Chaotic behavior in national stock market indices: New evidence from the close return test. *Glob. Financ. J.* **2001**, *12*, 35–53. [CrossRef]
16. Lahmiri, S. On fractality and chaos in Moroccan family business stock returns and volatility. *Phys. A* **2017**, *473*, 29–39. [CrossRef]
17. Inglada-Perez, L. Uncovering nonlinear dynamics in air transport demand. *Int. J. Transp. Econ.* **2016**, *XLIII*, 33–66.
18. Brock, W.A.; Dechert, W.D.; Sheinkman, J.A. *A Test of Independence Based on the Correlation Dimension*; SSRI no. 8702 1987; Department of Economics, University of Wisconsin: Madison, WI, USA, 1987.
19. Hsieh, D.A. Chaos and Nonlinear Dynamics: Application to Financial Markets. *J. Financ.* **1991**, *46*, 1839–1877. [CrossRef]
20. Blank, S.C. Chaos in Futures Markets: A Nonlinear Dynamically Analysis. *J. Futures Mark.* **1991**, *11*, 711–728. [CrossRef]
21. Mayfield, E.S.; Mizrach, B. On Determining the Dimension of Real-time Stock Price Data. *J. Bus. Econ. Stat.* **1992**, *10*, 367–374.
22. Yang, S.; Brorsen, B.W. Nonlinear Dynamics of Daily Futures Prices: Conditional Heteroskedasticity or Chaos? *J. Futures Mark.* **1993**, *13*, 175–191. [CrossRef]
23. Abhyankar, A.H.; Copeland, L.S.; Wong, W. Non-linear dynamics in real-time equity market indices: Evidence from the UK. *Econ. J.* **1995**, *105*, 864–880. [CrossRef]
24. Sewell, S.P.; Stansell, S.R.; Lee, I.; Below, D. Using chaos measures to examine international capital market integration. *Appl. Financ. Econ.* **1996**, *6*, 91–101. [CrossRef]
25. Abhyankar, A.H.; Copeland, L.S.; Wong, W. Uncovering nonlinear structure in real time, stock-market indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100. *J. Bus. Econ. Stat.* **1997**, *15*, 1–14.
26. Pandey, V.; Kohers, T.; Kohers, G. Deterministic nonlinearity in the stock returns of major European markets and the United States. *Financ. Rev.* **1998**, *33*, 45–63. [CrossRef]

27. Gao, A.H.; Wang, G.H.K. Modeling Nonlinear Dynamics of Daily Futures Price Changes. *J. Futures Mark.* **1999**, *19*, 325–351. [CrossRef]
28. Kyrtsou, C.; Terraza, M. Stochastic chaos or ARCH effects in stock series? A comparative study. *Int. Rev. Financ. Anal.* **2002**, *11*, 407–431. [CrossRef]
29. Antoniou, A.; Vorlow, C.E. Price Clustering and Discreteness: Is there Chaos behind the Noise? *Phys. A* **2005**, *348*, 389–403. [CrossRef]
30. Yousefpoor, P.; Esfahani, M.S.; Nojumi, H. Looking for systematic approach to select chaos tests. *Appl. Math. Comput.* **2008**, *198*, 73–91. [CrossRef]
31. Mishra, R.K.; Sehgal, S.; Bhanumurthy, N.R. A search for long-range dependence and chaotic structure in Indian stock market. *Rev. Financ. Econ.* **2011**, *20*, 96–104. [CrossRef]
32. BenSaïda, A. Are financial markets stochastic: A test for noisy chaos. *Am. Int. J. Contemp. Res.* **2012**, *2*, 57–68.
33. Webel, K. Chaos in German stock returns—New evidence from the 0—1 test. *Econ. Lett.* **2012**, *115*, 487–489. [CrossRef]
34. BenSaïda, A. Noisy chaos in intraday financial data: Evidence from the American index. *Appl. Math. Comput.* **2014**, *226*, 258–265. [CrossRef]
35. Tiwari, A.H.; Gupta, R. Chaos in G7 stock markets using over one century of data: A note. *Res. Int. Bus. Financ.* **2019**, *47*, 304–310. [CrossRef]
36. Dickey, D.; Fuller, W. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **1979**, *74*, 426–431.
37. Phillips, P.C.B.; Perron, P. Testing for Unit Roots in Time Series Regression. *Biometrika* **1988**, *75*, 335–346. [CrossRef]
38. Kwiatkowski, D.; Phillips, P.C.B.; Schmidt, P.; Shin, Y. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *J. Econ.* **1992**, *54*, 159–178. [CrossRef]
39. Wald, A.; Wolfowitz, J. On a test whether two samples are from the same population. *Ann. Math. Stat.* **1940**, *11*, 147–162. [CrossRef]
40. Keenan, D.M.A. Tukey nonadditivity-type test for time series non-linearity. *Biometrika* **1985**, *72*, 39–44. [CrossRef]
41. Tsay, R.S. Nonlinearity tests for time series. *Biometrika* **1986**, *73*, 461–466. [CrossRef]
42. Teräsvirta, T.; Lin, C.F.; Granger, C.W.J. Power of the Neural Network Linearity Test. *J. Time Ser. Anal.* **1993**, *14*, 209–220. [CrossRef]
43. White, H. An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks. In Proceedings of the International Joint Conference on Neural Networks, New York, NY, USA, 1–3 July 1989; IEEE Press: New York, NY, USA, 1989; pp. 451–455.
44. Brock, W.A.; Scheinkman, J.A.; Dechert, W.D.; LeBaron, B. A test for independence based on the correlation dimension. *Econom. Rev.* **1996**, *15*, 197–235. [CrossRef]
45. Kaplan, D.T. Exceptional events as evidence for determinism. *Physica D* **1994**, *73*, 38–48. [CrossRef]
46. Gottwald, G.A.; Melbourne, I. A new test for chaos in deterministic systems. *Proc. Math. Phys. Eng. Sci.* **2004**, *460*, 603–611. [CrossRef]
47. Grassberger, P.; Procaccia, I. Dimensions and entropies of strange attractors from a fluctuating dynamics approach. *Physica* **1984**, *13*, 34–54. [CrossRef]
48. Eckmann, J.; Ruelle, D. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **1985**, *57*, 617–656. [CrossRef]
49. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
50. Barkoulas, J.T.; Chakraborty, A.; Ouandlous, A. A metric and topological analysis of determinism in the crude oil spot market. *Energy Econ.* **2012**, *34*, 584–591. [CrossRef]
51. Barnett, W.A.; Gallant, A.R.; Hinich, M.J.; Jungeilges, J.A.; Kaplan, D.T.; Jensen, M.J. A single-blind controlled competition among tests for nonlinearity and chaos. *J. Econ.* **1997**, *82*, 157–192. [CrossRef]
52. Fraser, A.M.; Swinney, H.L. Independent Coordinates for Strange Attractors from Mutual Information. *Phys. Rev. A* **1986**, *33*, 1134–1140. [CrossRef]
53. Takens, F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence*; Rand, D., Young, L., Eds.; Springer: Berlin/Heidelberg, Germany, 1981; pp. 366–381.
54. Gu, R.; Chen, X.; Li, X. Chaos recognition and fractal analysis in the term structure of Shanghai Interbank Offered Rate. *Phys. A* **2014**, *412*, 101–112. [CrossRef]

55. Lahmiri, S. Investigating existence of chaos in short and long term dynamics of Moroccan exchange rates. *Phys. A* **2017**, *465*, 655–661. [CrossRef]
56. Wolf, A.; Swift, J.B.; Swinney, H.L.; Vastano, J.A. Determining Lyapunov exponents from a time series. *Phys. D* **1985**, *16*, 285–317. [CrossRef]
57. Shintani, M.; Linton, O. Nonparametric neural network estimation of Lyapunov exponents and direct test for chaos. *J. Econ.* **2004**, *120*, 1–33. [CrossRef]
58. McCaffrey, D.; Ellner, S.; Gallant, A.; Nychka, D. Estimating the Lyapunov exponent of a chaotic system with nonparametric regression. *J. Am. Stat. Assoc.* **1992**, *87*, 682–695. [CrossRef]
59. Bailey, B.A.; Ellner, S.; Nychka, D.W. Chaos with confidence: Asymptotics and applications of local Lyapunov exponents. In *Nonlinear Dynamics and Time Series: Building a Bridge Between the Natural and Statistical Sciences*; Cutler, C.D., Kaplan, D.T., Eds.; Fields Institute Communications: American Mathematical Society: Providence, RI, USA, 1997; pp. 115–133.
60. Donoho, D.; Johnstone, I. Ideal spatial adaptation via wavelet shrinkage. *Biometrika* **1994**, *81*, 425–455. [CrossRef]
61. Garcin, M.; Guégan, D. Wavelet shrinkage of a noisy dynamical m with non-linear noise impact. *Phys. D* **2016**, *325*, 126–145. [CrossRef]
62. Rubežić, V.; Djurović, I.; Sejdić, E. Average wavelet coefficient-based detection of chaos in oscillatory circuits. *COMPEL* **2017**, *36*, 188–201. [CrossRef]
63. Litimi, H.; BenSaida, A.; Belkacem, L.; Abdallah, O. Chaotic behavior in financial market volatility. *J. Risk* **2019**, *21*, 27–53. [CrossRef]
64. Eckmann, J.P.; Kamphorst, S.O.; Ruelle, D. Recurrence plots of dynamical systems. *Europhys. Lett.* **1987**, *4*, 973–977. [CrossRef]
65. MacKinnon, J.G. Numerical distribution functions for unit root and cointegration tests. *J. Appl. Econom.* **1996**, *11*, 601–618. [CrossRef]
66. Caraiani, P. Characterizing emerging European stock markets through complex networks: From local properties to self-similar characteristics. *Phys. A* **2012**, *391*, 3629–3637. [CrossRef]

entropy

MDPI

Article

Do Liquidity Proxies Based on Daily Prices and Quotes Really Measure Liquidity?

Barbara Będowska-Sójka [1,*,†] and Krzysztof Echaust [2,†]

[1] Department of Econometrics, Poznań University of Economics and Business, al. Niepodległości 10, 61-875 Poznań, Poland

[2] Department of Operations Research, Poznań University of Economics and Business, al. Niepodległości 10, 61-875 Poznań, Poland; krzysztof.echaust@ue.poznan.pl

* Correspondence: barbara.bedowska-sojka@ue.poznan.pl

† These authors contributed equally to this work.

Received: 22 May 2020; Accepted: 15 July 2020; Published: 17 July 2020

Abstract: This paper examines whether liquidity proxies based on different daily prices and quotes approximate latent liquidity. We compare percent-cost daily liquidity proxies with liquidity benchmarks as well as with realized variance estimates. Both benchmarks and volatility measures are obtained from high-frequency data. Our results show that liquidity proxies based on high-low-open-close prices are more correlated and display higher mutual information with volatility estimates than with liquidity benchmarks. The only percent-cost proxy that indicates higher dependency with liquidity benchmarks than with volatility estimates is the Closing Quoted Spread based on the last bid and ask quotes within a day. We consider different sampling frequencies for calculating realized variance and liquidity benchmarks, and find that our results are robust to it.

Keywords: liquidity proxy; liquidity benchmark; volatility estimate; correlation coefficient; partial determination; mutual information

1. Introduction

Liquidity is unobservable and elusive concept, which encompasses many transaction properties observed on the markets [1]. Various definitions of liquidity are proposed in the literature related to the bid-ask spreads [2], focused on the price impact of trading volumes [3], or referring to the market depth and dynamics of the order book [4,5]. Liquidity studies are performed on the basis of different information sets with different data frequency and on different markets. In order to maintain a uniform approach we focus on the bid-ask spread as a measure of transaction costs and follow the definition of liquidity as the ability to trade in a reasonable time and at a low cost [6].

Two types of liquidity measures are widely recognized: benchmarks and proxies [7]. In a calculation of benchmarks high-frequency data are required. These data are gathered in big datasets. Dealing with them is highly challenging and time-consuming. In the past decade a number of researchers have sought to determine which liquidity proxy based on low-frequency (daily) data is the best one to represent unobserved liquidity. The competition for the best liquidity proxy relies on the examination of the strength of dependency between proxies and benchmarks [2,7–9]. There is no single answer, which measure is the best approximation for the unobserved liquidity and thus its proper measurement is a very demanding process [9–11].

Proxies for bid-ask spreads, the so-called percent-cost proxies, are based either on bid and ask quotes (the closing quoted spread of Chung and Zhang [12]) or on high-low-open-close *HLOC* prices (the effective spread of Corwin and Schultz [8], Abdi and Ranaldo measure [13], or high-low range [14]). The application of the high and low prices is justified by the fact that high prices are usually buyer-initiated prices, while low prices are usually the seller-initiated [8]. However, these prices are

also commonly used for non-parametric volatility estimators as e.g., Garman and Klass estimators [15]. Moreover, there is the evidence in the literature that liquidity is related to volatility [16].

To our best knowledge it has not been verified yet, whether daily proxies based on the range of prices and quotes measure unobserved liquidity or volatility. In order to address this gap we examine to what extent liquidity proxies measure liquidity and/or volatility. We employ four benchmarks based on high-frequency data as well as four percent-cost proxies based on daily data [7,17]. Volatility is approximated by two realized variance measures [18] as well as downside and upside realized semivariance [19].

Three approaches are applied: firstly, we investigate the correlation coefficients for proxies and either benchmarks or volatility estimates. Secondly, through the partial determination analysis we examine which of these two, liquidity benchmark or volatility estimate, explains variability of liquidity proxies [20]. Thirdly, we apply mutual information to measure inherent dependencies between any proxy and either liquidity benchmark or volatility estimate [21]. All approaches are conducted within the cross-section and the portfolio time-series settings.

This paper makes a unique contribution to the literature. We find that proxies proposed in the literature based on high-low-open-close prices measure volatility rather than liquidity. The closing quoted spread proposed by Chung and Zhang [12] is the only daily proxy which shows higher dependence with liquidity benchmark than with any volatility estimate. This measure uses the bid and ask quotes observed at the end of the day. Other percent-cost liquidity proxies based on four prices (applied in [8,13,14]) approximate volatility, not liquidity. These conclusions are robust to the changes of an approach undertaken, the cross-section or the portfolio time-series, a method of the dependency measurement and the aggregation of liquidity measures, daily or monthly. They also remain unchanged when high liquidity or low liquidity periods are considered.

Liquidity and volatility are the key factors in price formation process, which are as important in the case of emerging markets as in the case of developed ones. Our study is conducted on the biggest emerging market in the Central and East European countries, on the Warsaw Stock Exchange (WSE). Thus this study extends the understanding of the nature of those relations also on relatively less liquid markets. Our findings are important for both practitioners who seeks for the best liquidity proxies as for academics who deliberate on such measures.

The rest of the paper is organized as follows: Section 2 presents the literature review on volatility and liquidity relation, Section 3 shows the research methodology, Section 4 presents empirical results, Section 5 investigates the robustness of the results in sub-periods and Section 6 concludes.

2. Volatility and Liquidity—The Literature Review

Discussion on the relationship between volatility and liquidity has a long history [22,23]. Obviously, these two are of the highest importance to regulators and practitioners. As both volatility and liquidity are latent and both are closely related to the process governing prices, the task of complete distinction of these two is challenging. Karpoff [24] shows the evidence that the large volume and price changes have common sources in the information flow process. Thus the dissemination of information among market participants seems to play crucial role in shaping these two. However, Karpoff did not use a notion of "liquidity". His seminal paper is on volume, but volume itself might be perceived as a liquidity measure.

The relation between liquidity and volatility in the microstructure theory is not unambiguously defined. In the inventory models this relation is negative [25,26]: higher liquidity implies lower volatility and vice versa. In the information-based models this relation could be also positive [27]: higher liquidity might be accompanied by higher volatility.

The empirical studies show different results in this area. On the one hand Chung and Zhang find that a market uncertainty represented by the Volatility Index, VIX, is a crucial determinant for stock liquidity in the US [12]. Also Ma et al. [28] show that liquidity on the stock markets is lower when investor risk perception reflected by VIX is higher. On the emerging markets Girard et al. [29] find that

the relationship between expected volume and volatility is negative and relate it to market inefficiencies. There is the evidence that the transaction costs are higher on the emerging markets [9,30,31]. Also, liquidity tends to decrease when volatility on a domestic market or the market uncertainty measured by VIX increase [32]. On the other hand, Chordia et al. [33] indicate that market volatility induces lower spreads, which means that there is a positive relationship between liquidity and volatility. The evidence from the Chinese stock market is that although market volatility reduces trading activity, it has mixed effects on market liquidity [34] .

The difference between the best buy and sell prices, the bid-ask spread, has been historically the most popular measure of liquidity [35]. Domowitz et al. [30] differentiate between liquidity (approximated by trading volume), transaction costs (spreads) and volatility, and consider the relationship between these three variables. They show that higher volatility tends to reduce turnover. In their approach liquidity is separated from transaction costs, while in majority of studies the transaction costs (namely bid-ask spreads) are used to measure liquidity (e.g., [2,25,36,37]).

Summing up, there is a clear distinction between liquidity and volatility in the literature, even if the exact definitions of both concepts vary from one study to another [38,39]. It seems that the liquidity estimates from different dimensions should be interrelated, and they should express stronger dependency with each other than with any volatility estimate.

3. Data and Methodology

A vast number of papers is driven by the need of obtaining the best proxy of liquidity at the possible lowest cost. The liquidity measures usually require the access to databases with intraday quotations. The existence of the simple, easy-to-calculate and widely available measure would be appreciated by the market participants. Thus many attempts of creating such a measure on the basis of daily data are made (e.g., [2,7,8,13,36]). We focus on measures based on daily prices or quotations, and examine how strongly are these liquidity proxies related to liquidity benchmarks as well as to well-known volatility estimates.

We consider one market, the Warsaw Stock Exchange (WSE), which is an order-driven market without market makers. This market has been considered as the emerging one [40–43]. We use long sample of 11 years (2737 days or 133 months); the sample period starts from January 2006 and ends up in December 2016. This 11-year period is long enough to capture different market regimes. Although within this time the WSE was considered as an emerging market, previous studies show that the coherence of liquidity measures is similar to one observed on the developed markets [14]. We take into account quotations of 73 stocks that have been constantly listed within this period and are considered as either big or medium in terms of capitalization. In the case of the WSE it means that they have market value over 50 mln euro. Stocks which experienced splits within sample period were removed from the study. The list of stocks is available upon request. Our primary data come from tick-by-tick database and are cleared from the errors such as multiple records, entries with negative spread, entries for which the spread is more than 50 times the median spread on that day etc. [44]. Finally they are aggregated into equally sampled intraday data.

The empirical framework is conducted on the basis of methodology presented in [7]. Both the cross-section approach for the levels, as well as the portfolio time-series for differences of liquidity measures are applied. The novelty of our approach lies in the examination of interdependence of proxies with benchmarks and with volatility estimates at the same time. Additionally to the calculation of correlation coefficients, we also conduct regression analysis and calculate partial determination coefficients—it allows us to decompose the impact of both benchmarks and volatility measures on variation of proxies. Finally, we examine the dependence between variables using the mutual information measure.

Since the aim of the paper is to examine the relationship between proxies and both benchmarks and volatility estimates, we employ different measures for each of these categories. Starting with

proxies we use only percent-cost proxies that are based on the daily values, either *HLOC* prices, or bid and ask quotes. The following proxies are considered:

- effective spread estimator of Corwin and Schultz [8], which rely on the empirical observation, that the highest price within the day t, H_t, is the buyer-initiated price, whereas the lowest price, L_t, is the seller-initiated price:
$$CSA_t = \frac{2(e^{\alpha_t} - 1)}{1 + e^{\alpha_t}}, \tag{1}$$
where $\alpha_t = \frac{\sqrt{2\beta_t} - \sqrt{\beta_t}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma_t}{3-2\sqrt{2}}}, \beta_t = (\log(\frac{H_t}{L_t}))^2 + (\log(\frac{H_{t+1}}{L_{t+1}}))^2$, and $\gamma_t = \log\frac{max(H_t, H_{t+1})}{min(L_t, L_{t+1})}$. We adjusted the high-low ratio spread estimator for overnight returns [8];
- the closing percent quoted spread proposed by Chung and Zhang [12]:
$$CQS_t = 2(P_{A_t} - P_{B_t})/(P_{A_t} + P_{B_t}), \tag{2}$$
where P_{A_t} and P_{B_t} are ask and bid quotes, respectively, observed at the end of the day t. It is the only one among our proxies that is based on quotes instead of prices;
- the high-low range which is a reformulation of the closing percent quoted spread of Chung and Zhang where the bid and ask quotes are replaced with the high and low prices:
$$HLR_t = 2(H_t - L_t)/(H_t + L_t); \tag{3}$$
- the measure of Abdi and Ranaldo [13] defined as follows:
$$AR_t = 2\sqrt{(C_t - 0.5(H_t + L_t))(C_t - 0.5(H_{t+1} + L_{t+1}))}, \tag{4}$$
where C_t is the closing price on day t.

All daily proxies are interpreted in the same way—the higher the value, the less liquidity is provided.

We also consider different benchmarks, which control for several aspects of the transaction costs. Assume the following notation: there are K equally sampled observations within a day, $k = 0, 1, \ldots, K$. The benchmarks are defined as follows:

- proportional effective spread
$$PES_k = ES_k/MP_k, \tag{5}$$
and ES is an effective spread, obtained as
$$ES_k = 2D_k(P_k - MP_k), \tag{6}$$
where P_k is price of the last transaction in an equally spaced time interval (e.g., 5-min), while MP_k is the mid price of the best ask quote, P_A, and the best bid quote, P_B, within specified interval; $MP_k = 0.5(P_{A_k} + P_{B_k})$. D_k is a variable indicating the direction of the k-th trade with 1 and -1 for buy and sell orders, respectively. In order to indicate the direction of a trade, Lee and Ready algorithm is applied [45].
- proportional quoted spread
$$PQS_k = 2(P_{A_k} - P_{B_k})/(P_{A_k} + P_{B_k}). \tag{7}$$

This spread is based on the quotes only and does not take into account the direction of orders measure [46]. Next two measures are based on the transaction (trade) prices or quotes:

- squared log return on trade prices

$$SRTP_k = (\log(P_k) - \log(P_{k-1}))^2, k \neq 0, \tag{8}$$

- midquote squared return

$$MSR_k = (\log(MP_k) - \log(MP_{k-1}))^2, k \neq 0. \tag{9}$$

When aggregating over period (day or month) a stock's liquidity benchmark is calculated as volume-weighted average of its values computed over all k observations in the period.

Volatility is approximated by estimates which are based on the logarithmic high-frequency returns:

- realized variance [47]

$$RV_t = \sum_{k=1}^{K} r_{t,k}^2, \tag{10}$$

where k represents an interval and t is for a given day. It is assumed that at a sufficiently high frequency and in the absence of jumps, the realized variance can be a good approximation of the unobservable volatility. Thus we also consider minimum RV as a measure that is known to be robust to jumps:

- minRV [18]

$$minRV_t = \frac{\pi}{\pi - 2} \frac{K}{K-1} \sum_{k=1}^{K-1} min(|r_{t,k}|; |r_{t,k+1}|)^2. \tag{11}$$

Additionally, we also consider two realized semivariances that allow to focus on the particular risk of long or short position [19]:

- downside realized semivariance

$$sRVd_t = \sum_{k=1}^{K} r_k^2 \times I(r_{t,k} < 0), \tag{12}$$

  where I is an indicator variable conditioning calculation of the variance only on the basis of negative returns.
- upside realized semivariance

$$sRVu_t = \sum_{k=1}^{K} r_k^2 \times I(r_{t,k} > 0), \tag{13}$$

  where I is an indicator of positive returns.

All benchmarks and volatility estimates are calculated with the highfrequency R package [48].

4. Empirical Research

The empirical research is divided into three parts. Firstly, we apply cross-section analysis for the levels of liquidity and volatility measures as well as portfolio time-series analysis for the first difference of time-series. Secondly, we provide results for the partial determination coefficient analysis, which enables us to differentiate between the relation of a proxy with a liquidity benchmark and volatility estimate. In the last step we calculate mutual information which quantifies the amount of information about a proxy obtained through observing the benchmarks or volatility estimates.

Before examining the dependency between considered variables, we present averages of our proxies, benchmarks and volatility estimates aggregated into a monthly frequency. Figure 1 shows that the dynamics of the four proxies based on daily prices and quotes are similar. The average values of proxies are increasing (and liquidity is decreasing) in the time of global financial crisis in 2008 as

well as in mid 2011 as a result of the sovereign debt crisis in Europe. This pattern is common across all daily proxies.

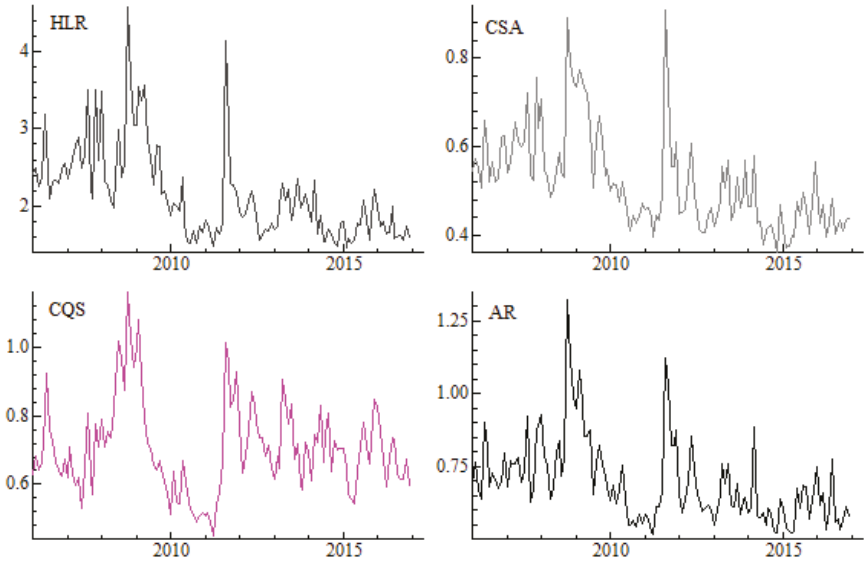


Figure 1. Monthly proxies calculated in the cross-section approach. Note: The following liquidity proxies based on the daily data aggregated to monthly frequency are presented: *HLR* stands for the high-low range [14], *CSA* is spread estimator of [8], *CQS* is the closing quoted spread of [12], while *AR* is the spread estimator of [13]. All graphs are prepared in OxMetrics [49].

Figure 2 presents liquidity benchmarks considered in the study. We find that the overall trends and comovements of measures are quite similar. The higher the benchmarks' values, the less liquidity is provided. Finally, the dynamics of four volatility estimates in the cross-section approach is presented in Figure 3. There are no substantial visual differences in behaviour of the series, besides the fact that realized variance, *RV*, displays the highest values. Both realized semivariances measure risk of either positive returns (upside realized semivariance *sRVu*) or negative returns (downside realized semivariance *sRV*d), while *minRV* is robust to jumps measure of volatility.

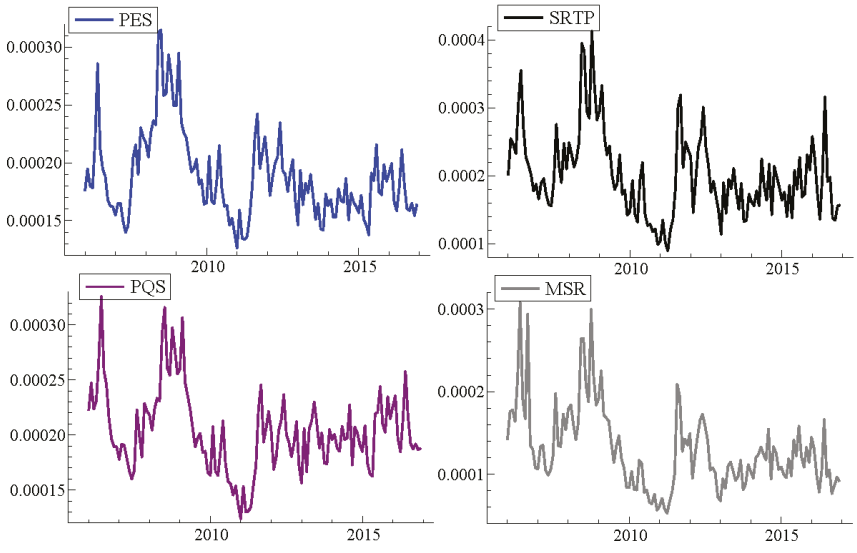


Figure 2. Monthly liquidity benchmarks in the cross-section approach. Note: This graphs shows liquidity benchmarks obtained on the basis of 5-min data and aggregated to monthly values in a cross-section approach. *SRTP* is squared log return on trade prices [48], *PES* is the proportional effective spread [35], *PQS* is the proportional quoted spread [50], and *MSR* is the midquote squared return [48]. For the sake of comparison, *SRTP* and *MSR* are multiplied by 100.

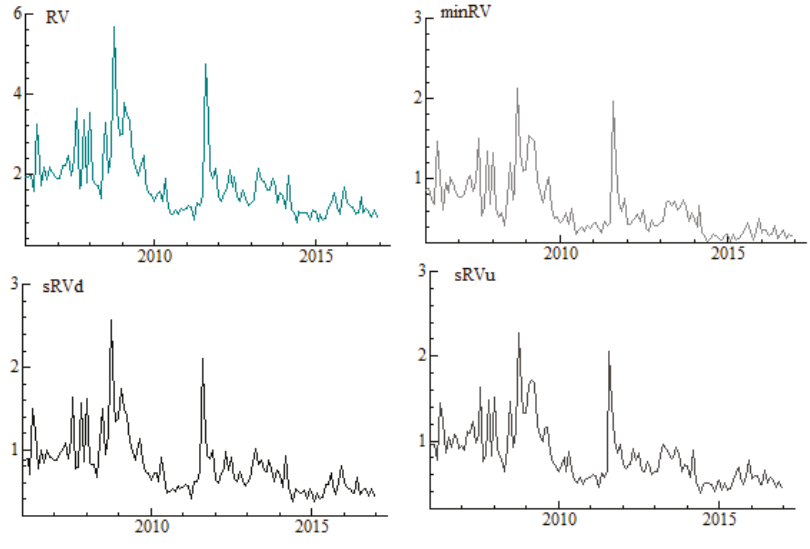


Figure 3. Monthly volatility estimates in the cross-section approach. Note: The following estimators of volatility calculated in 5-min frequency and aggregated to monthly values are considered: *RV* is realized variance [18], *minRV* is minimum realized variance [18], *sRVd* and *sRVu* are downside and upside realized semivariances, respectively [19].

4.1. Cross-Section Analysis

The average cross-sectional correlations between liquidity proxies and liquidity benchmarks or volatility estimates are computed according to the research methodology presented in [7]: for each day (and each month) we calculate the cross-sectional correlations across all firms in the sample, and then the average correlation is calculated over all days (or months). Spearman rank correlations are applied. This analysis is provided in four different frequencies and with the respect to all considered measures of either liquidity or volatility, for daily and monthly measures separately. We check if the correlations are different between each proxy-benchmark and proxy-volatility pairs using t-test on the time-series of correlations in the spirit of Fama-MacBeth. Following Fong et al. [7], we calculate the cross-sectional correlations and then regress the correlations of one pair on the correlations of another pair. We assume that the time series of correlations of each proxy is IID over time, and examine if the regression intercept is zero and the slope is one. The Newey-West standard errors are applied in order to adjust for autocorrelation [51].

Table 1 presents the Spearman rank ρ coefficients for proxies and liquidity benchmarks calculated on the basis of four frequencies: 5-, 10-, 30-, and 60-minute data, and volatility estimates calculated in the same frequencies. Benchmarks and volatility estimates are presented in pairs for each sampling frequency. Columns 2–5 present the correlations obtained for the series aggregated into daily data, while columns 6–9 display the correlations for series aggregated into monthly data.

For estimates calculated on the basis of 5-min data, we find the evidence that for HLR, CSA and AR correlations with benchmarks are rather low (in absolute values) and definitely weaker than these observed with volatility estimates. It means that proxies based on daily prices ($HLOC$), are closer to volatility estimates than to any benchmark. The mostly striking example is HLR, which is characterized by strong correlations (higher than 55%) with any estimate of volatility, and simultaneously is weakly correlated with benchmarks as PES or PQS. The only opposite case is observed for CQS, for which the correlations with liquidity benchmarks are stronger than with volatility estimates. This finding holds true also when other frequencies are examined.

When focusing on the monthly aggregates, the main result holds: correlations with volatility estimates are stronger than with benchmarks for all proxies, and the only exception is CQS. Our findings show that correlations of CQS with PES and PQS are around 79% and thus are very close to correlation coefficients reported in [7]—79.9% and 91.5%, respectively. For the remaining sampling frequencies similar results are observed. For CQS the lower the frequency of calculating benchmarks or volatility estimates is, the higher the correlation with volatility, but still the correlations with benchmarks remain high.

4.2. Portfolio Time-Series Approach

The portfolio time-series approach is based on equally-weighted portfolios across all stocks for a day or a month. We compute a benchmark or volatility estimate in a specified interval by taking the average of detrended benchmarks and volatility estimates over all stocks in a day or a month. The detrending is done by calculating first differences of the time-series. As detrended series are stationary Pearson correlations are calculated. Table 2 shows that independently of the sampling frequency for HLR, CSA and AR the correlations with any estimate of volatility are higher than with any liquidity benchmark. The same situation applies to monthly portfolios.

The remarkable exception among proxies is CQS again, for which in daily portfolios the correlations with both spreads PES and PQS are 47% and 50%, respectively, while the correlations with volatility estimates ranges from 12% to 29% (these numbers apply to the case in which both benchmarks and volatility estimates are based on 5-min data). In monthly portfolios there is no clear answer, which of these two, liquidity benchmarks or volatility estimates, are highly correlated with CQS. As we examine two overlapping correlations with a common variable (proxy), first with a benchmark and second with a volatility estimate, we apply Zou's test [52]. This test calculates the

confidence interval of differences between two correlations. If the confidence interval includes zero the null hypothesis that these correlations are equal cannot be rejected.

Table 1. The Spearman rank correlations between liquidity proxies and liquidity benchmarks or volatility estimates—the cross-section analysis.

	Daily				Monthly			
Proxies	**HLR**	**CSA**	**CQS**	**AR**	**HLR**	**CSA**	**CQS**	**AR**
Liquidity benchmark: 5 min								
PES	9.13	3.57	66.75 *	10.68	−1.63	−1.45	79.82 *	28.42
PQS	12.15	4.62	68.35 *	10.85	2.92	2.66	78.47 *	29.45
SRTP	40.66	13.53	46.29 *	15.81 •	18.39	10.87	72.49 *	35.82
MSR	36.62	8.02	37.80 *	9.84	23.70	11.36	66.60 *	34.36
Volatility estimate: 5 min								
RV	73.11	24.38	29.82	19.01	65.90	49.69	52.42	56.19
sRVu	68.07	23.19	25.00	17.17	68.49	52.08	47.27	55.02
sRVd	68.79	25.20	25.93	18.64	66.47	51.53	49.80	55.37
minRV	55.55	19.01	−18.83	6.27	77.69	58.96	−23.75	25.57
Liquidity benchmark: 10 min								
PES	8.30	3.07	66.72 *	10.50	−6.31	−5.05	84.38 *	28.76
PQS	12.26	4.53	68.75 *	8.75	−0.63	0.10	83.70 *	30.48
SRTP	45.53	14.59	42.09 *	15.90	23.03	13.20	67.04 *	36.77
MSR	42.35	9.29	32.01 *	9.83	30.07	14.58	58.15 •	34.52
Volatility estimate: 10 min								
RV	71.35	23.32	31.09	19.05	60.77	44.70	58.84	56.88
sRVu	65.14	22.00	25.67	17.08	62.48	46.27	55.71	56.34
sRVd	66.04	24.37	27.04	18.83	59.62	45.24	58.81	56.59
minRV	57.49	20.04	−13.60	7.63	78.53	58.52	−9.60	31.27
Liquidity benchmark: 30 min								
PES	8.52	2.92	66.19 *	10.57	−10.22	−8.40	87.67 *	29.18
PQS	13.04	4.89	69.02 *	11.08	−2.94	−1.53	87.33 *	31.32
SRTP	52.85	14.94	34.49 *	15.05	34.85	18.68	54.92	36.19
MSR	50.44	10.22	22.75	9.03	42.66	21.09	41.56	32.33
Volatility estimate: 30 min								
RV	69.17	21.24	30.11	18.52	57.27	38.95	60.69	57.01
sRVu	59.17	19.45	23.47	16.09	57.42	39.34	58.37	56.08
sRVd	60.57	22.79	25.26	18.23	54.28	38.16	61.49	29.18
minRV	58.50	20.34	−6.21	9.88	75.26	55.43	12.56	41.84
Liquidity benchmark: 60 min								
PES	9.25	2.91	64.37 *	10.35	−10.63	−8.93	87.26 *	29.05
PQS	13.21	5.14	68.37 *	11.18	−2.95	−1.16	87.26 *	31.27
SRTP	55.96	14.65	29.53 *	29.53 *	40.89	21.44	48.29	36.59
MSR	54.14	10.61	17.80	8.61	48.79	24.25	34.93	32.79
Volatility estimate: 60 min								
RV	67.93	19.43	27.17	17.45	58.44	36.79	56.43	55.48
sRVu	53.56	17.05	19.76	14.32	57.51	36.59	53.80	53.95
sRVd	55.71	21.32	21.62	17.06	54.52	35.44	56.40	54.03
minRV	55.99	19.37	−4.58	10.67	70.03	50.06	21.26	44.66

Note: The Spearman rank correlations (presented in percentage) are calculated for levels of liquidity proxies and liquidity benchmarks or volatility estimates. The following liquidity proxies are considered: *HLR* stands for the high-low range [14], *CSA* is spread estimator of [8], *CQS* is the closing quoted spread of [12], while *AR* is the spread estimator of [13]. Among liquidity benchmarks *PES* is the proportional effective spread [35], *PQS* is the proportional quoted spread [50], *SRTP* and *MSR* are squared return on trade prices and the midquote squared return, respectively [48]. For volatility estimates *RV* is realized variance, [18], *sRVu* and *sRVd* are upside and downside realized semivariances, respectively [19], while *midRV* is minimum realized variance [18]. * signifies that the correlation is statistically significantly higher at the 5% level than the correlation between the same proxy and *RV*, • signifies that the correlation is insignificantly different from the correlation between the same proxy and *RV* at the 5% level.

We find that considered proxies based on *HLOC* prices are much more related to volatility estimates than to liquidity benchmarks. The only measure which in daily frequency pronounces higher correlation coefficient with two liquidity benchmarks, *PES* and *PQS*, is the closing quoted spread, *CQS*. These results are consistent for all frequencies in which calculation of benchmarks or volatility estimates has been done.

Table 2. The Pearson correlation coefficients between liquidity proxies and liquidity benchmarks or volatility estimates—the portfolio time-series approach.

	Daily				Monthly			
	HLR	**CSA**	**CQS**	**AR**	**HLR**	**CSA**	**CQS**	**AR**
Liquidity benchmark: 5 min								
PES	17.53	6.34	47.11 *	3.81	38.25	37.26	63.21 •	45.23
PQS	20.52	8.87	50.39 *	5.88	51.92	51.92	43.38	44.93
SRTP	18.66	10.32	13.64	5.79	56.65	54.43	70.32 •	64.88
MSR	6.61	4.60	4.81	1.74	55.62	48.43	62.12 •	59.38
Volatility estimate: 5 min								
RV	81.42	43.11	24.26	26.51	94.86	84.54	69.40	79.65
sRVu	63.16	27.13	12.10	19.11	94.16	82.52	65.68	76.00
sRVd	78.42	44.45	29.00	25.42	93.97	85.13	69.72	80.46
minRV	73.97	39.41	17.69	24.80	94.28	83.58	62.63	76.33
Liquidity benchmark: 10 min								
PES	18.74	5.34	44.00 *	46.62 *	41.90	40.49	65.24 •	48.32
PQS	22.10	8.60	49.98 *	6.35	48.24	47.52	75.34 •	56.27
SRTP	29.55	15.39	14.18	9.65	64.25	59.42	67.27 •	61.45
MSR	7.80	5.18	3.59	2.30	69.96	61.18	66.27 •	71.16
Volatility estimate: 10 min								
RV	79.79	42.29	23.56	26.49	94.92	84.08	70.20	79.74
sRVu	53.29	20.67	9.16	15.77	94.22	81.75	66.71	75.56
sRVd	76.29	44.00	28.44	25.23	93.81	84.81	71.43	81.11
minRV	75.46	38.49	20.99	25.55	93.88	83.36	61.45	75.62
Liquidity benchmark: 30 min								
PES	17.64	5.13	48.17 *	2.96	41.75	40.35	69.44 •	48.39
PQS	21.88	8.31	52.37 *	6.97	47.11	46.73	77.69 •	54.88
SRTP	29.24	15.99	9.43	8.65	72.02	63.45	58.69	71.57
MSR	17.52	11.89	6.17	6.63	70.95	58.82	57.21	68.03
Volatility estimate: 30 min								
RV	77.23	43.84	21.71	26.91	94.46	82.25	70.08	79.81
sRVu	38.84	13.71	2.81	11.36	93.14	78.39	65.68	73.70
sRVd	72.32	44.42	26.91	24.93	92.91	83.26	71.95	81.78
minRV	74.60	34.40	21.79	24.29	94.43	82.87	64.27	76.83
Liquidity benchmark: 60 min								
PES	17.37	4.72	47.87 *	2.56	40.46	37.54	69.88 •	45.85
PQS	20.96	8.24	52.59 *	6.63	44.41	44.56	76.86 •	51.96
SRTP	33.67	17.27	9.37	9.88	68.37	55.49	47.97	63.67
MSR	24.05	13.47	8.20	7.46	67.38	53.70	44.44	62.58
Volatility estimate: 60 min								
RV	76.04	43.39	20.86	26.77	93.99	80.80	69.08	78.43
sRVu	30.77	8.45	−0.87	7.84	91.20	74.61	62.62	69.52
sRVd	69.95	44.39	26.50	25.15	91.97	81.89	71.45	80.95
minRV	72.50	29.13	26.03	20.74	92.99	81.10	64.23	73.32

Note: The Pearson correlation coefficients (presented in percentage) are calculated for differences of liquidity proxies and liquidity benchmarks or volatility estimates. The following liquidity proxies are considered: *HLR* stands for the high-low range [14], *CSA* is spread estimator of [8], *CQS* is the closing quoted spread of [12], while *AR* is the spread estimator of [13]. Among liquidity benchmarks *PES* is the proportional effective spread [35], *PQS* is the proportional quoted spread [50], *SRTP* and *MSR* are squared return on trade prices and the midquote squared return, respectively [48]. For volatility estimates *RV* is realized variance, [18], *sRVu* and *sRVd* are upside and downside realized semivariances, respectively [19], while *midRV* is minimum realized variance [18]. * signifies that the correlation is statistically significantly higher at the 5% level than the correlation between the same proxy and *RV*, • signifies that the correlation is insignificantly different from the correlation between the same proxy and *RV* at the 5% level.

4.3. Partial Determination Coefficients

The previous sub-sections are devoted to the correlations between proxies and benchmarks or volatility estimates separately. Here we propose to apply the regression analysis and investigate partial determination coefficients. The idea is the following: we consider linear regression for liquidity and volatility measures, that have been used in the cross-section (in Section 4.1) and the portfolio time-series analysis (in Section 4.2). For the former the equation has a following form:

$$Proxy_i = \alpha_0 + \alpha_1 \cdot Bench_i + \alpha_2 \cdot RV_i + \epsilon_i, \tag{14}$$

where $Proxy_i$ denotes liquidity proxy for stock i, $Bench_i$ denotes liquidity benchmark, RV_i stays for volatility estimate.

For the latter, the portfolio time-series analysis, the equation is following:

$$\Delta Proxy_t = \beta_0 + \beta_1 \cdot \Delta Bench_t + \beta_2 \cdot \Delta RV_t + \epsilon_t, \tag{15}$$

where Δ is the first difference, and t is a time index. The regressions are estimated for both daily and monthly portfolios.

The coefficient of partial determination is the proportion of variation, that can be described by the predictors used in the full model, but cannot be explained in a reduced model [53]. The formula to compute the coefficient of partial determination, PR^2, is as follows:

$$PR^2 = \frac{SS_{reduced} - SS_{full}}{SS_{reduced}}, \tag{16}$$

where $SS_{reduced}$ is the sum of squares of residuals from the model with only one independent variable, and SS_{full} is the sum of squares of residuals from the full model. In our case the reduced model is a model with either a liquidity benchmark or a volatility estimate, while the full model takes into account both variables simultaneously. Since RV among all volatility estimates displays the highest correlation coefficient with proxies, we further show the results for this estimate. As the changes in frequency have no impact on the correlation coefficients and 5-min frequency of observation is usually used as a rule of thumb [54–56], henceforth we present results for benchmarks and volatility estimates based on 5-min frequency only (the calculations for remaining proxies and frequencies are available upon request). In calculations rsq R package [57] is applied.

Firstly, we provide results for the cross-section approach. Table 3 presents the determination coefficients, R^2, as well as partial determination coefficients for both variables, the liquidity benchmark PR^2_{Bench} and the volatility estimate PR^2_{RV}. We find that both for daily and monthly proxies the value of determination coefficient R^2 varies from 5% to 62%. The comparison of partial determination coefficients, PR^2_{Bench} and PR^2_{RV}, shows that for HLR, CSA and AR both in daily and monthly data partial determination coefficients are much higher for the volatility estimate than for the liquidity benchmark. The only proxy for which we obtain higher partial determination coefficient for liquidity benchmarks than for volatility proxy is CQS. Here in the case of proportional effective spread, PES, in daily data the partial determination coefficient for liquidity is 40% versus 2% for volatility. For PQS the impact of benchmark is 42%, while for volatility it is less than 2% (1.6%). In the case of monthly data, the conclusions are nearly the same.

Secondly, we repeat the procedure for the portfolio time-series approach. Table 4 shows the determination and partial determination coefficients. As in the previous case in daily data CQS is the only proxy for which the partial determination coefficients for liquidity benchmarks, namely PES, PQS and MSR, are higher than for volatility estimates. For monthly data the partial determination coefficients are more balanced: for PES we obtain 30% for liquidity benchmark versus 40% for volatility estimate, while for PQS we get 39% versus 36%. Still CQS seems to represent liquidity, while the other proxies measure volatility.

Table 3. Partial determination coefficients—The cross-section approach.

	Daily			**Monthly**		
	R^2	PR^2_{Bench}	PR^2_{RV}	R^2	PR^2_{Bench}	PR^2_{RV}
			HLR			
PES	0.586	0.099	0.573	0.620	0.082	0.613
PQS	0.582	0.091	0.566	0.612	0.063	0.596
SRTP	0.569	0.062	0.474	0.602	0.040	0.496
MSR	0.549	0.019	0.466	0.594	0.020	0.485
			CSA			
PES	0.096	0.023	0.077	0.082	0.024	0.061
PQS	0.095	0.022	0.076	0.077	0.019	0.057
SRTP	0.093	0.020	0.061	0.076	0.018	0.049
MSR	0.091	0.017	0.070	0.074	0.016	0.053
			CQS			
PES	0.465	0.403	0.020	0.393	0.338	0.029
PQS	0.485	0.424	0.016	0.409	0.355	0.016
SRTP	0.241	0.153	0.019	0.202	0.130	0.018
MSR	0.186	0.091	0.035	0.153	0.075	0.025
			AR			
PES	0.070	0.020	0.042	0.052	0.020	0.027
PQS	0.070	0.019	0.042	0.051	0.019	0.026
SRTP	0.068	0.017	0.029	0.049	0.017	0.019
MSR	0.065	0.014	0.042	0.048	0.016	0.025

Note: The partial determination coefficients are from linear regression in a form: $Proxy = \alpha_0 + \alpha_1 \cdot Bench + \alpha_2 \cdot RV + \epsilon$ estimated both for daily data (columns 2-4) and monthly data (columns 5-7). Volatility is proxied by the realized variance calculated from the 5-min data. The same frequency is applied to the calculation of different liquidity benchmarks. Among liquidity proxies HLR stands for the high-low range [14], CSA is spread estimator of [8], CQS is the closing quoted spread of [12], while AR is the spread estimator of [13]. Among liquidity benchmarks PES is the proportional effective spread [35], PQS is the proportional quoted spread [50], $SRTP$ and MSR are squared return on trade prices and the midquote squared return, respectively [48].

4.4. Mutual Information

So far we have used dependency measures which assume linearity of the relation. As a robustness check we also apply a mutual information which could be considered as a nonparametric dependency measure. Mutual information is an estimate of inherent dependence between two random variables. It specifies the "amount of information" that is shared by two variables and is expressed in terms of the joint probability distribution. The MI concept comes from the Information Theory and is closely related to that of entropy [58–60]. The entropy of random variable X, $H(X)$, is expressed in the following way:

$$H(X) = E[-\log_2(p(X))] = -\sum_{i=1}^{L} p(x_i)\log_2(p(x_i)), \tag{17}$$

where $p(X)$ is a probability mass function, while L is the length of the time series.

The joint entropy for two random variables, X and Y, is defined as:

$$H(X,Y) = -\sum_{i=1}^{L} p(x_i,y_i)\log_2(p(x_i,y_i)), \tag{18}$$

where $p(x_i,y_i)$ is the joint probability that $X = (x_i)$ and $Y = (y_i)$. The mutual information between X and Y is then defined:

$$MI(X,Y) = H(X) + H(Y) - H(X,Y). \tag{19}$$

MI can be then normalized:

$$MI^* = MI(X,Y)/\sqrt{H(X)H(Y)}. \tag{20}$$

The normalized values of MI^* are within $[0,1]$ interval, with 0 denoting that both random variables are independent, and 1 denoting they share the same information.

Table 4. Partial determination coefficients—The portfolio time-series approach.

	Daily			**Monthly**		
	R^2	PR^2_{Bench}	PR^2_{RV}	R^2	PR^2_{Bench}	PR^2_{RV}
			HLR			
PES	0.663	0.000	0.652	0.900	0.003	0.883
PQS	0.663	0.001	0.648	0.901	0.007	0.875
SRTP	0.665	0.006	0.653	0.902	0.016	0.855
MSR	0.663	0.001	0.662	0.903	0.027	0.859
			CSA			
PES	0.186	0.001	0.183	0.717	0.009	0.672
PQS	0.186	0.000	0.179	0.718	0.013	0.653
SRTP	0.186	0.000	0.177	0.722	0.024	0.604
MSR	0.186	0.000	0.184	0.716	0.004	0.629
			CQS			
PES	0.350	0.320	0.014	0.637	0.301	0.396
PQS	0.372	0.343	0.008	0.684	0.390	0.356
SRTP	0.047	0.003	0.036	0.625	0.276	0.258
MSR	0.159	0.120	0.035	0.567	0.165	0.295
			AR			
PES	0.071	0.000	0.069	0.659	0.067	0.571
PQS	0.070	0.000	0.067	0.667	0.090	0.544
SRTP	0.071	0.000	0.067	0.693	0.161	0.471
MSR	0.070	0.000	0.070	0.672	0.103	0.494

Note: The partial determination coefficients are from linear regression in a form: $\Delta Proxy = \alpha_0 + \alpha_1 \cdot \Delta Bench + \alpha_2 \cdot \Delta RV + \epsilon$ estimated both for daily data (columns 2–4) and monthly data (columns 5–7). Volatility is proxied by the realized variance calculated from the 5-min data. The same frequency is applied to the calculation of different liquidity benchmarks. Among liquidity proxies HLR stands for the high-low range [14], CSA is spread estimator of [8], CQS is the closing quoted spread of [12], while AR is the spread estimator of [13]. Among liquidity benchmarks PES is the proportional effective spread [35], PQS is the proportional quoted spread [50], $SRTP$ and MSR are squared return on trade prices and the midquote squared return, respectively [48].

In the study the mutual information measures are calculated both for the cross-section approach and the portfolio time-series approach (we applied the infotheo R package [61]). Table 5 shows the results for the cross-section on daily and monthly data. For proxies versus benchmarks relation in daily data the highest values are obtained for CQS and either PES or PQS. In monthly data all benchmarks have the highest mutual information with CQS. For proxy versus volatility relation, HLR is characterized by the highest mutual information with all volatility estimates both for daily and for monthly data.

Table 6 presents the average mutual information for the portfolio time-series approach. For proxies versus benchmarks the highest mutual information is observed between CQS and PES or PQS. These dependencies are even more pronounced in the case of monthly data.

For volatility versus proxies relation, HLR is featured by the highest amount of mutual information with any volatility estimate, while CQS shows the lowest mutual information with volatility measures. These results hold for both daily and monthly data. Summing up, the results obtained in this Section do not differ significantly from the previous findings.

Table 5. The mutual information—The cross-section approach.

	Daily				Monthly			
	HLR	**CSA**	**CQS**	**AR**	**HLR**	**CSA**	**CQS**	**AR**
Benchmarks								
PES	5.96	5.51	22.96	5.68	6.76	6.95	35.57	8.21
PQS	6.25	5.55	24.21	5.65	6.71	6.99	35.32	8.49
SRTP	10.82	6.06	13.31	6.33	8.56	7.46	28.27	10.07
MSR	11.28	5.74	10.17	5.82	9.54	7.19	23.27	9.32
Volatility estimates								
RV	26.59	8.14	8.51	7.06	21.78	14.04	16.60	16.27
sRVu	23.52	7.88	8.43	6.61	23.77	14.89	15.03	15.66
sRVd	24.10	8.19	8.63	6.74	22.10	14.34	15.76	16.16
minRV	17.79	7.31	9.25	5.35	30.70	17.33	10.58	7.57

Note: The numbers in table denotes the averages of normalized mutual information (in percentage) between proxies and benchmarks or volatility estimates. The latter two are calculated on the basis of 5-min frequency. The following liquidity proxies are considered: *HLR* stands for the high-low range [14], *CSA* is spread estimator of [8], *CQS* is the closing quoted spread of [12], while *AR* is the spread estimator of [13]. Among liquidity benchmarks *PES* is the proportional effective spread [35], *PQS* is the proportional quoted spread [50], *SRTP* and *MSR* are squared return on trade prices and the midquote squared return, respectively [48]. For volatility estimates *RV* is realized variance, [18], *sRVu* and *sRVd* are upside and downside realized semivariances, respectively [19], while *midRV* is minimum realized variance [18].

Table 6. The mutual information—the portfolio time-series approach.

	Daily				Monthly			
	HLR	**CSA**	**CQS**	**AR**	**HLR**	**CSA**	**CQS**	**AR**
Benchmarks								
PES	1.63	1.35	7.48	1.16	6.58	7.14	17.78	10.15
PQS	1.99	1.08	8.54	1.39	8.03	7.47	21.23	11.42
SRTP	2.30	1.49	1.90	1.48	9.20	9.64	18.13	16.25
MSR	2.40	1.70	1.79	1.67	10.79	9.41	17.10	14.16
Volatility estimates								
RV	15.34	3.73	1.99	2.23	31.51	21.95	14.13	20.12
sRVu	8.71	2.48	1.53	1.89	27.70	18.07	12.33	16.62
sRVd	12.21	4.31	2.60	2.75	30.20	25.48	16.15	22.56
minRV	10.00	3.33	1.49	2.21	29.79	22.48	15.12	15.41

Note: The numbers in table denotes normalized mutual information (in percentage) between proxies and benchmarks or volatility estimates. The latter two are calculated on the basis of 5-min frequency. Mutual information is obtained for daily and monthly data separately. The following liquidity proxies are considered: *HLR* stands for the high-low range [14], *CSA* is spread estimator of [8], *CQS* is the closing quoted spread of [12], while *AR* is the spread estimator of [13]. Among liquidity benchmarks *PES* is the proportional effective spread [35], *PQS* is the proportional quoted spread [50], *SRTP* and *MSR* are squared return on trade prices and the midquote squared return, respectively [48]. For volatility estimates *RV* is realized variance [18], *sRVu* and *sRVd* are upside and downside realized semivariances, respectively [19], while *midRV* is minimum realized variance [18].

5. Sub-Period Analysis

A potential drawback in our approach is that empirical results may be sensitive to the number of observations taken into considerations. Moreover, some statistical properties may depend on the specifics of the time-series and may be sensitive to the choice of the sample period. This section is devoted to validate the results and assess their consistency. Instead of the whole 11-year period we have chosen two specific two-year sub-periods (484 days) that are closely related to the market liquidity level. The sub-periods are 2007.07.02–2009.06.08. and 2013.01.02–2014.12.30 and relate to the low liquidity and high liquidity regimes, respectively (see Figures 1 and 2). Following previous results we conduct an analysis using all three approaches, the calculation of correlation coefficients, partial

determination coefficients and the mutual information for both cross-section and portfolio time-series analyses. Since in the chosen sub-periods there are only 24 months we carry out these computations for the daily aggregation and 5-minute frequency only. The results are shown in the Appendix A in Tables A1–A6.

Generally in sub-periods we find the same relations as in the whole sample. In both periods volatility estimates show higher correlations with proxies than with liquidity benchmarks. The only exception is *CQS* which demonstrates much higher correlation with benchmarks than with volatility estimates in the cross-section analysis. In the portfolio time-series approach we find high dependence only between *CQS* and *PES* or *PQS* benchmarks. The regression analysis highlights very weak impact of any intraday measures on *CSA* and *AR* (low determination coefficient). However, *HLR* seems to be entirely explained by volatility measures. The mutual information results confirm those obtained in two previous approaches. The highest mutual information is observed for *CQS* and both *PES* and *PQS*. Also we find high mutual information for *HLR* in association with volatility estimates. These results hold for both sub-periods.

When the comparison between two sub-periods is performed, *CQS* as the best proxy for liquidity indicates higher correlation with *PES* and *PQS* in high liquidity subperiod than in low liquidity time. Hovewer, according to Zou's test [52] this difference in correlations between high and low liquidity periods are significant only for *CQS*-*PQS* pair. The regression analysis confirms this finding, i.e., the determination coefficient as a goodness of fit measure is higher in the high liquidity period when the impact of benchmarks in the bivariate relationship is stronger. The mutual information allows to formulate the same conclusions but only in the cross-section approach. In the portfolio time-series the results are ambiguous.

6. Conclusions

This paper investigates whether liquidity proxies based on daily data commonly used in the literature indeed approximate latent liquidity. The relations between stock market volatility and liquidity have been the subject of much recent investigation both from the academics' and practitioners' point of view. Our research question is driven by the fact that some liquidity proxies, similarly to some volatility measures, apply four prices, that is high-low-open-close prices [15]. In such circumstances, there arises a question, what exactly is measured by a given liquidity proxy. This is an important issue, as there is a need for easy-to-obtain and calculate liquidity measure and many horse races for finding the best proxy are run. The proxies based on range of prices are often examined in such races.

Our results show that measures based on high and low prices capture rather unobserved volatility than liquidity. Both the effective spread estimator of Corwin and Schultz [8] and the spread of Abdi and Ranaldo [13] that have been proposed recently are closer related to different volatility estimates than to any liquidity benchmarks used in the study. Also the high-low range used as a reformulation of the closing quoted spread of Chung and Zhang is less correlated with benchmarks than with volatility estimates. These findings are confirmed by partial determination coefficients from the regression analysis as well as the non-parametric approach based on the mutual information calculation. They hold for the cross-section approach as well as the portfolio time-series.

The only measure based on closing bid and ask quotes, the closing quoted spread of Chung and Zhang [12], has higher dependency with liquidity benchmarks than with volatility estimates. This is confirmed within the correlation analysis, the partial determination coefficients' analysis and through application of mutual information as a measure of non-linear dependency. All these approaches unanimously indicate that among daily proxies *CQS* is mostly related to liquidity benchmarks. This proxy is also indicated as the best one in Fong et al. [7].

The answer to the question in the title, "do liquidity proxies based on daily prices and quotes really measure liquidity?" is 'yes' for proxies based on daily quotes and 'no' for proxies based on daily prices as the latter approximate volatility rather than liquidity. According to our results the proper

measurement of liquidity based on daily data requires knowledge of bid and ask prices at the end of the day. Unfortunately this information is not offered in the widely available databases.

Author Contributions: Conceptualization, B.B.-S. and K.E.; methodology, B.B.-S. and K.E.; software, B.B.-S. and K.E.; formal analysis, K.E.; resources, B.B.-S. and K.E.; data curation, K.E.; supervision, B.B.-S., writing—original draft preparation, B.B.-S. and K.E.; writing—review and editing B.B.-S. and K.E.; funding acquisition, B.B.-S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Science Centre in Poland under the grant no.UMO-2017/25/B/HS4/01546.

Acknowledgments: We are grateful for the fruitful discussion during 4th International Workshop on "Financial Markets and Nonlinear Dynamics" (FMND) in Paris, Vietnam Symposium in Banking and Finance in Ha Noi, and 10th Annual Financial Market Liquidity Conference in Budapest. We would like to thank anonymous Reviewers for outlining very detailed list of remarks. It was a great help in improving our paper. All remaining errors are ours.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. The Spearman rank correlations between liquidity proxies and liquidity benchmarks or volatility estimates—The cross-section approach in two subperiods.

	Low Liquidity Period				High Liquidity Period			
	HLR	**CSA**	**CQS**	**AR**	**HLR**	**CSA**	**CQS**	**AR**
Liquidity benchmark								
PES	6.72	2.96	64.42 *	9.06	12.71	3.41	68.84 *	11.10 •
PQS	8.39	3.27	65.12 *	8.84	15.64	4.56	71.48 *	11.84 •
SRTP	37.44	11.28	43.39 *	12.64	42.34	13.78	50.38 *	16.73
MSR	32.44	6.51	37.99 *	7.80	39.45	7.85	41.15 *	10.39
Volatility estimate								
RV	73.00	22.16	26.37	15.88	72.25	24.50	34.77	20.18
sRVu	67.77	21.79	22.50	14.71	67.58	23.43	28.78	18.43
sRVd	69.80	22.08	23.21	14.93	67.71	25.76	29.37	19.58
minRV	59.31	17.69	−16.22	5.03	52.83	19.57	−19.47	7.04

Note: The Spearman rank correlations (presented in percentage) are calculated for levels of liquidity proxies and liquidity benchmarks or volatility estimates. Two periods are taken into account: low liquidity period (2007.07.02–2009.06.08) and high liquidity period (2013.01.02–2014.12.30). For the remaining abbreviations please refer to the notes below Table 1.

Table A2. The Pearson correlation coefficients between liquidity proxies and liquidity benchmarks or volatility estimates—The portfolio time-series approach in two subperiods.

	Low Liquidity Period				High Liquidity Period			
	HLR	**CSA**	**CQS**	**AR**	**HLR**	**CSA**	**CQS**	**AR**
Liquidity benchmark								
PES	13.84	4.46	40.68 •	1.60	12.53	6.89	44.04 *	8.55
PQS	15.63	7.43	42.94 *	4.02	18.26	12.40	54.76 *	11.67 •
SRTP	18.18	10.87	7.20	5.95	5.66	4.23	10.33	−2.01
MSR	1.97	2.08	2.87	−1.06	14.17	4.34	8.43	−4.67
Volatility estimate								
RV	81.76	44.57	31.74	25.65	79.51	40.41	26.55	20.74
sRVu	62.54	24.49	19.23	14.07	59.31	16.78	10.97	6.99
sRVd	77.03	46.59	33.48	27.46	77.68	47.40	31.67	25.80
minRV	78.55	40.50	28.34	24.27	68.01	36.20	19.29	21.89

Note: The Pearson rank correlations (presented in percentage) are calculated for levels of liquidity proxies and liquidity benchmarks or volatility estimates. Two periods are taken into account: low liquidity period (2007.07.02–2009.06.08) and high liquidity period (2013.01.02–2014.12.30). For the remaining abbreviations please refer to the notes below Table 2.

Table A3. Partial determination coefficients—The cross section approach in subperiods.

	Low Liquidity Period			**High Liquidity Period**		
	R^2	PR^2_{Bench}	PR^2_{RV}	R^2	PR^2_{Bench}	PR^2_{RV}
			HLR			
PES	0.590	0.110	0.578	0.574	0.096	0.558
PQS	0.586	0.102	0.574	0.570	0.088	0.550
SRTP	0.571	0.070	0.492	0.555	0.056	0.449
MSR	0.548	0.021	0.483	0.536	0.017	0.440
			CSA			
PES	0.090	0.024	0.071	0.095	0.023	0.077
PQS	0.090	0.023	0.071	0.093	0.022	0.075
SRTP	0.088	0.022	0.059	0.092	0.020	0.060
MSR	0.086	0.019	0.066	0.089	0.017	0.070
			CQS			
PES	0.433	0.381	0.016	0.492	0.411	0.023
PQS	0.441	0.390	0.014	0.525	0.449	0.017
SRTP	0.214	0.143	0.018	0.276	0.164	0.016
MSR	0.176	0.101	0.023	0.220	0.099	0.045
			AR			
PES	0.062	0.022	0.035	0.070	0.017	0.044
PQS	0.062	0.022	0.035	0.070	0.016	0.042
SRTP	0.060	0.020	0.027	0.069	0.016	0.029
MSR	0.057	0.017	0.034	0.067	0.013	0.043

Note: Two periods are taken into account: low liquidity period (2007.07.02–2009.06.08) and high liquidity period (2013.01.02–2014.12.30). The partial determination coefficients are from linear regression in a form: $Proxy = \alpha_0 + \alpha_1 \cdot Bench + \alpha_2 \cdot RV + \epsilon$ estimated both for low liquidity period (columns 2–4) and high liquidity period (columns 5–7). Volatility is proxied by the realized variance calculated from the 5-min data. The same frequency is applied to the calculation of different liquidity benchmarks. For the remaining abbreviations please refer to the notes below Table 3.

Table A4. Partial determination coefficients—The portfolio time-series approach in subperiods.

	Low Liquidity Period			**High Liquidity Period**		
	R^2	PR^2_{Bench}	PR^2_{RV}	R^2	PR^2_{Bench}	PR^2_{RV}
			HLR			
PES	0.669	0.002	0.662	0.637	0.012	0.631
PQS	0.669	0.002	0.661	0.634	0.004	0.621
SRTP	0.673	0.013	0.662	0.633	0.001	0.631
MSR	0.669	0.003	0.669	0.633	0.001	0.625
			CSA			
PES	0.199	0.000	0.197	0.164	0.001	0.160
PQS	0.199	0.000	0.194	0.164	0.000	0.150
SRTP	0.199	0.001	0.190	0.163	0.000	0.162
MSR	0.199	0.000	0.198	0.164	0.000	0.162
			CQS			
PES	0.235	0.149	0.083	0.221	0.162	0.034
PQS	0.247	0.163	0.077	0.314	0.262	0.021
SRTP	0.101	0.001	0.097	0.077	0.006	0.067
MSR	0.101	0.000	0.100	0.072	0.002	0.066
			AR			
PES	0.066	0.000	0.066	0.044	0.001	0.037
PQS	0.066	0.000	0.064	0.047	0.004	0.034
SRTP	0.066	0.000	0.063	0.045	0.002	0.044
MSR	0.066	0.001	0.066	0.049	0.007	0.047

Note: Two periods are taken into account: low liquidity period (2007.07.02–2009.06.08) and high liquidity period (2013.01.02–2014.12.30). The partial determination coefficients are from linear regression in a form: $Proxy = \alpha_0 + \alpha_1 \cdot Bench + \alpha_2 \cdot RV + \epsilon$ estimated both low liquidity period (columns 2–4) and high liquidity period (columns 5–7). Volatility is proxied by the realized variance calculated from the 5-min data. The same frequency is applied to the calculation of different liquidity benchmarks. For the remaining abbreviations please refer to the notes below Table 4.

Table A5. The mutual information—The cross-section approach in two subperiods.

	Low Liquidity Period				High Liquidity Period			
	HLR	**CSA**	**CQS**	**AR**	**HLR**	**CSA**	**CQS**	**AR**
Liquidity benchmark								
PES	6.17	5.33	21.53	5.29	6.38	5.41	24.17	5.90
PQS	6.25	5.36	21.62	5.26	6.84	5.57	26.41	5.81
SRTP	10.37	5.81	12.17	5.83	11.37	6.01	14.83	6.52
MSR	10.54	5.43	10.18	5.51	12.23	5.60	11.13	6.06
Volatility estimate								
RV	26.05	7.61	7.60	6.43	26.38	8.03	9.89	7.28
sRVu	22.98	7.65	7.72	6.37	23.37	7.60	9.68	6.63
sRVd	24.45	7.67	7.68	6.16	23.51	8.22	9.73	6.88
minRV	19.25	7.32	8.72	5.29	16.81	7.16	9.89	5.32

Note: The numbers in table denotes normalized mutual information (in percentage) between proxies and benchmarks or volatility estimates. The latter two are calculated on the basis of 5-min frequency. Two periods are taken into account: low liquidity period (2007.07.02–2009.06.08) and high liquidity period (2013.01.02–2014.12.30). For the remaining abbreviations please refer to the notes below Table 5.

Table A6. The mutual information—The portfolio time-series approach in two subperiods.

	Low Liquidity Period				High Liquidity PerioCQSd			
	HLR	**CSA**	**CQS**	**AR**	**HLR**	**CSA**	**CQS**	**AR**
Liquidity benchmark								
PES	1.95	1.75	11.97	1.02	2.78	2.26	6.21	2.38
PQS	2.77	1.18	12.68	1.75	2.07	3.47	8.13	2.37
SRTP	4.00	1.81	1.80	2.37	2.33	1.86	2.51	2.24
MSR	3.33	2.19	2.64	2.22	3.40	1.65	3.05	2.40
Volatility estimate								
RV	17.50	5.50	2.84	3.38	17.00	3.75	3.84	1.99
sRVu	9.34	3.26	1.77	3.48	11.09	1.61	2.59	1.97
sRVd	12.62	5.72	3.23	3.44	16.27	3.38	4.18	2.26
minRV	14.29	5.23	4.21	3.95	10.77	2.08	3.38	1.89

Note: The numbers in table denotes normalized mutual information (in percentage) between proxies and benchmarks or volatility estimates. The latter two are calculated on the basis of 5-min frequency. Two periods are taken into account: low liquidity period (2007.07.02–2009.06.08) and high liquidity period (2013.01.02–2014.12.30). For the remaining abbreviations please refer to the notes below Table 6.

References

1. Kyle, A.S. Continuous auctions and insider trading. *Econometrica* **1985**, *53*, 1315–1335. [CrossRef]
2. Goyenko, R.Y.; Holden, C.W.; Trzcinka, C.A. Do liquidity measures measure liquidity? *J. Financ. Econ.* **2009**, *92*, 153–181. [CrossRef]
3. Amihud, Y. Illiquidity and stock returns: Cross-section and time-series effects. *J. Financ. Mark.* **2002**, *5*, 31–56. [CrossRef]
4. Hautsch, N. *Econometrics of Financial High-Frequency Data*; Springer: Berlin/Heidelberg, Germany, 2012.
5. Hautsch, N.; Jeleskovic, V. Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models. Available online: https://ssrn.com/abstract=1292493 (accessed on 15 July 2020).
6. Pastor, L.; Stambaugh, R.F. Liquidity risk and expected stock returns. *J. Political Econ.* **2003**, *111*, 642–685. [CrossRef]
7. Fong, K.Y.L.; Holden, C.W.; Trzcinka, C.A. What Are the Best Liquidity Proxies for Global Research? *Rev. Financ.* **2017**, *21*, 1355–1401. [CrossRef]
8. Corwin, S.A.; Schultz, P. A simple way to estimate bid-ask spreads from daily high and low prices. *J. Financ.* **2012**, *67*, 719–759. [CrossRef]

9. Ahn, H.J.; Cai, J.; Yang, C.W. Which Liquidity Proxy Measures Liquidity Best in Emerging Markets? *Economies* **2018**, *6*, 67. [CrossRef]
10. Coughenour, J.F.; Saad, M.M. Common market makers and commonality in liquidity. *J. Financ. Econ.* **2004**, *73*, 37–69. [CrossRef]
11. Batten, J.A.; Vo, X.V. Liquidity and Return Relationships in an Emerging Market. *Emerg. Mark. Financ. Trade* **2014**, *50*, 5–21. [CrossRef]
12. Chung, K.H.; Zhang, H. A simple approximation of intraday spreads using daily data. *J. Financ. Mark.* **2014**, *17*, 94–120. [CrossRef]
13. Abdi, F.; Ranaldo, A. A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices. *Rev. Financ. Stud.* **2017**, *30*, 4437–4480. [CrossRef]
14. Będowska-Sójka, B. The coherence of liquidity measures. The evidence from the emerging market. *Financ. Res. Lett.* **2018**, *27*, 118–123. [CrossRef]
15. Garman, M.; Klass, M. On the Estimation of Security Price Volatilities from Historical Data. *J. Bus.* **1980**, *53*, 67–78. [CrossRef]
16. Karolyi, G.A.; Lee, K.H.; van Dijk, M.A. Understanding Commonality in Liquidity around the World. *J. Financ. Econ.* **2012**, *105*, 82–112. [CrossRef]
17. Będowska-Sójka, B.; Echaust, K. What is the best proxy for liquidity in the presence of extreme illiquidity? *Emerg. Mark. Rev.* **2020**. [CrossRef]
18. Andersen, T.G.; Bollerslev, T.; Diebold, F.X. Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility. *Rev. Econ. Stat.* **2007**, *89*, 701–720. [CrossRef]
19. Barndorff-Nielsen, O.E.; Kinnebrock, S.; Shephard, N. Measuring downside risk? Realised semivariance. In *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*; Bollerslev, T.R.J., Watson, M., Eds.; Oxford University Press: New York, NY, USA, 2010.
20. Kenett, D.Y.; Huang, X.; Vodenska, I.; Havlin, S.; Stanley, H.E. Partial correlation analysis: Applications for financial markets. *Quant. Financ.* **2015**, *15*, 569–578. [CrossRef]
21. Dionisio, A.; Menezes, R.; Mendes, D.A. Mutual information: A measure of dependency for nonlinear time series. *Physica A* **2004**, *344*, 326–329. [CrossRef]
22. Epps, T.W.; Epps, M.L. The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis. *Econometrica* **1976**, *44*, 305–321. [CrossRef]
23. Cohen, K.J.; Ness, W.L.; Okuda, H.; Schwartz, R.A.; Whitcomb, D.K. The determinants of common stock returns volatility: An international comparison. *J. Financ.* **1976**, *31*, 733–741. [CrossRef]
24. Karpoff, J.M. The relation between price changes and trading volume: A survey. *J. Financ. Quant. Anal.* **1987**, *22*, 109–126. [CrossRef]
25. Amihud, Y.; Mendelson, H. Asset pricing and the bid-ask spread. *J. Financ. Econ.* **1986**, *17*, 223–249. [CrossRef]
26. Copeland, T.E. A model of asset trading under the assumption of sequential information arrival. *J. Financ.* **1976**, *31*, 1149–1168. [CrossRef]
27. Amihud, Y.; Mendelson, H.; Pedersen, L.H. Liquidity and asset prices. *Found. Trends Financ.* **2005**, *1*, 269–364. [CrossRef]
28. Ma, R.; Anderson, H.D.; Marshall, B.R. Market Volatility, Liquidity Shocks, and Stock Returns: Worldwide Evidence. *Pac.-Basin Financ. J.* **2018**, *49*, 164–199. [CrossRef]
29. Girard, E.; Biswas, R. Trading Volume and Market Volatility: Developed vs. Emerging Stock Markets. *Financ. Rev.* **2007**, *42*, 429–459. [CrossRef]
30. Domowitz, I.; Glen, J.; Madhavan, A. Liquidity, volatility and equity trading costs across countries and over time. *Int. Financ.* **2001**, *4*, 221–255. [CrossRef]
31. Będowska-Sójka, B. The dynamics of low-frequency liquidity measures: The developed versus the emerging market. *J. Financ. Stab.* **2019**, *42*, 136–142. [CrossRef]
32. Będowska-Sójka, B.; Echaust, K. Commonality in Liquidity Indices: The Emerging European Stock Markets. *Systems* **2019**, *7*, 24. [CrossRef]
33. Chordia, T.; Roll, R.; Subrahmanyam, A. Market Liquidity and Trading Activity. *J. Financ.* **2001**, *56*, 501–530. [CrossRef]
34. Ma, R.; Anderson, H.D.; Marshall, B.R. Stock market liquidity and trading activity: Is China different? *Int. Rev. Financ. Anal.* **2018**, *56*, 32–51. [CrossRef]

35. Blau, B.M.; Griffith, T.G.; Whitby, R.J. The maximum bid-ask spread. *J. Financ. Mark.* **2018**, *41*, 1–16. [CrossRef]
36. Lesmond, D.A.; Ogden, J.P.; Trzcinka, C.A. A new estimate of transaction costs. *Rev. Financ. Stud.* **1999**, *12*, 1113–1141. [CrossRef]
37. Lesmond, D.A. Liquidity of emerging markets. *J. Financ. Econ.* **2005**, *77*, 411–452. [CrossRef]
38. Capelle-Blancard, G.; Havrylchyk, O. The impact of the French securities transaction tax on market liquidity and volatility. *Int. Rev. Financ. Anal.* **2016**, *47*, 166–178. [CrossRef]
39. Galariotis, E.C.; Krokida, S.I.; Spyrou, S.I. Herd behavior and equity market liquidity: Evidence from major markets. *Int. Rev. Financ. Anal.* **2016**, *48*, 140–149. [CrossRef]
40. Rak, R.; Drożdż, S.; Kwapień, J. Nonextensive statistical features of the Polish stock market fluctuations. *Phys. A Stat. Mech. Its Appl.* **2007**, *374*, 315–324. [CrossRef]
41. Rak, R.; Drożdż, S.; Kwapień, J.; Oświęcimka, P. Stock Returns Versus Trading Volume: Is the Correspondence More General? *Acta Phys. Pol. B* **2013**, *44*, 2035. [CrossRef]
42. Kozłowska, M.; Denys, M.; Wiliński, M.; Link, G.; Gubiec, T.; Werner, T.; Kutner, R.; Struzik, Z. Dynamic bifurcations on financial markets. *Chaos Solitons Fractals* **2016**, *88*, 126–142. [CrossRef]
43. Będowska-Sójka, B. Commonality in liquidity measures. The evidence from the Polish stock market. In Proceedings of the International Scientific Conference Hradec Economic Days 2019, Hradec Kralove, Czech Republic, 25–26 March 2019; University of Hradec Kralove: Hradec Kralove, Czech Republic, 2019; pp. 29–40.
44. Barndorff-Nielsen, O.E.; Hansen, P.R.; Lunde, A.; Shephard, N. Realized kernels in practice: Trades and quotes. *Econom. J.* **2009**, *12*, C1–C32. [CrossRef]
45. Lee, C.M.C. Market integration and price execution for NYSE-listed securities. *J. Financ.* **1993**, *48*, 1009–1038. [CrossRef]
46. Boudt, K.; Petitjean, M. Intraday liquidity dynamics and news releases around price jumps: Evidence from the DJIA stocks. *J. Financ. Mark.* **2014**, *17*, 121–149. [CrossRef]
47. Andersen, T.G.; Bollerslev, T. Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *Int. Econ. Rev.* **1998**, *39*, 885–905. [CrossRef]
48. Boudt, K.; Cornelissen, J.; Payseur, S.; Nguyen, G.; Schermer, M. Highfrequency: Tools for Highfrequency Data Analysis. R Package Version 0.5.3. 2018. Available online: https://cran.r-project.org/web/packages/highfrequency/highfrequency.pdf (accessed on 15 July 2020).
49. Doornik, J.; Hendry, D. *Empirical Econometric Modelling*; PcGiveTM11; Timberlake Consultants: London, UK, 2005.
50. Bessembinder, H. Issues in assessing trade execution costs. *J. Financ. Mark.* **2003**, *6*, 233–257. [CrossRef]
51. Newey, W.K.; West, K.D. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* **1987**, *55*, 703–708. [CrossRef]
52. Zou, G.Y. Toward using confidence intervals to compare correlations. *Psychol. Methods* **2007**, *12*, 399–413. [CrossRef]
53. Lipsitz, S.R.; Leong, T.; Ibrahim, J.; Lipshultz, S. A partial correlation coefficient and coefficient of determination for multivariate normal repeated measures data. *J. R. Stat. Soc. Ser. Stat.* **2001**, *50*, 87–95. [CrossRef]
54. Giot, P. Market risk models for intraday data. *Eur. J. Financ.* **2005**, *11*, 309–324. [CrossRef]
55. Będowska-Sójka, B. Is intraday data useful for forecasting VaR? The evidence from EUR/PLN exchange rate. *Risk Manag.* **2018**, *20*, 326–346. [CrossRef]
56. Chevallier, J.; Sévi, B. On the volatility-volume relationship in energy futures markets using intraday data. *Energy Econ.* **2012**, *34*, 1896–1909. [CrossRef]
57. Zhang, D. Package rsq: R-Squared and Related Measures. R Package 1.1. 2018. Available online: https://cran.r-project.org/web/packages/rsq/rsq.pdf (accessed on 15 July 2020).
58. Cover, T.; Thomas, J. *Elements of Information Theory*; Wiley: New York, NY, USA, 1991.
59. Yue, P.; Fan, Y.; Batten, J.A.; Zhou, W.X. Information Transfer between Stock Market Sectors: A Comparison between the USA and China. *Entropy* **2020**, *22*, 194. [CrossRef]

60. Hu, S.; Zhong, M.; Cai, Y. Impact of Investor Behavior and Stock Market Liquidity: Evidence from China. *Entropy* **2019**, *21*, 1111. [CrossRef]
61. Meyer, P.E. infotheo: Information-Theoretic Measures. R Package Version 1.2.0. 2014. Available online: https://cran.r-project.org/web/packages/infotheo/infotheo.pdf (accessed on 15 July 2020).

entropy

MDPI

Article

The Evolution Characteristics of Systemic Risk in China's Stock Market Based on a Dynamic Complex Network

Yong Shi [1,2,3,4], Yuanchun Zheng [1,2,3], Kun Guo [2,3,5,*], Zhenni Jin [2,3,6] and Zili Huang [7]

1 School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100190, China; yshi@ucas.ac.cn (Y.S.); zhengyuanchun14@mails.ucas.ac.cn (Y.Z.)
2 Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, Beijing 100190, China; jinzhenni18@mails.ucas.ac.cn
3 Research Center on Fictitious Economy & Data Science, Chinese Academy of Sciences, Beijing 100190, China
4 College of Information Science and Technology, University of Nebraska at Omaha, Omaha, NE 68182, USA
5 School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China
6 Sino-Danish College, University of Chinese Academy of Sciences, Beijing 100049, China
7 Geisel School of Medicine, Dartmouth College, Hanover, NH 03755, USA; zili.huang.gr@dartmouth.edu
* Correspondence: guokun@ucas.ac.cn; Tel.: +86-138-1043-9286

Received: 9 April 2020; Accepted: 29 May 2020; Published: 2 June 2020

Abstract: The stock market is a complex system with unpredictable stock price fluctuations. When the positive feedback in the market amplifies, the systemic risk will increase rapidly. During the last 30 years of development, the mechanism and governance system of China's stock market have been constantly improving, but irrational shocks have still appeared suddenly in the last decade, making investment decisions risky. Therefore, based on the daily return of all a-shares in China, this paper constructs a dynamic complex network of individual stocks, and represents the systemic risk of the market using the average weighting degree, as well as the adjusted structural entropy, of the network. In order to eliminate the influence of disturbance factors, empirical mode decomposition (EMD) and grey relational analysis (GRA) are used to decompose and reconstruct the sequences to obtain the evolution trend and periodic fluctuation of systemic risk. The results show that the systemic risk of China's stock market as a whole shows a downward trend, and the periodic fluctuation of systemic risk has a long-term equilibrium relationship with the abnormal fluctuation of the stock market. Further, each rise of systemic risk corresponds to external factor shocks and internal structural problems.

Keywords: complex network; systemic risk; structural entropy; stock market; EMD

1. Introduction

The stock market is a typical complex system with multiple stock prices fluctuating from equilibrium to deviation and to equilibrium again. A large number of heterogeneous investors buy and sell stocks frequently, making the relationships between different stocks unpredictable. In most scenarios, owing to some factors like herd effect, investors' investment strategies converge [1,2]; when some investors buy a stock, other investors tend to buy the same one, and furthermore, when the vast majority of investors buy or sell stocks, other investors usually follow this action. At the same time, different listed companies are another heterogeneous agent in the stock market. On the one hand, the economic exchanges between listed companies will lead to the linkage of their stock prices. On the other hand, similar actions by investors on similar stocks can cause herd behavior between different stock prices. When the prices of a large number of stocks in the market tend to be consistent, it means that the herd effect in the market is higher, and the stock market is more likely to fluctuate excessively and consistently, leading to higher market systemic risks [3,4].

In former studies, the capital asset pricing model (CAPM) framework was usually used to analyze financial systematic risks as a basic theory [5–8]. According to CAPM, risks can be divided into systematic risk (or market risk) and non-systematic risk, while the latter can be diminished through investment portfolios. The systematic risk often refers to pervasive, far-reaching, perpetual market risk, which can be measured by the variance of the portfolio (Beta) altogether. Therefore, most studies on systematic risk are based on Beta values. Although this theory is widely adopted, it usually comes with a number of hypotheses, such as homogenous investors in capital markets. However, in modern financial markets, different investors generally have different degrees of rationality, ability to obtain information, and sensitivity to prices, that is, investors are usually heterogenous. Hence, CAPM may not be a reasonable model in the real complex world [9,10]. More importantly, this paper focuses on the systemic risk, which reflects the stability of the system and the characteristics of risk transmission among individuals in a certain complex system.

A complex network, which is based on physics and mathematics theory, can tackle complicated practical problems [11]. It is especially suitable for modeling, analysis, and calculation in complex finance systems [12]. Nowadays, the literature on applying complex networks to finance is growing in size, and complex networks have become important tools in the finance field [13]. After 30 years of development, China's stock market is growing in scale and vitality, while the market operation mechanism and management system are constantly improving. Nevertheless, there have been several typical bear and bull markets in recent years, and systemic risk in the stock market has risen periodically. Therefore, a dynamic complex network of individual stocks in China's stock market is constructed in this paper to measure the dynamic systemic risk of China's stock market. Then, the tendency evolution and cycle change characteristics of systemic risk are explored.

The structure of this paper is as follows. Section 2 summarizes the applications of complex networks in the field of economy and finance; Section 3 introduces the data and methodology used in this paper; Section 4 proposes the empirical results and analysis; and the conclusions and some discussion are given in Section 5.

2. Related Works

Construction of the network consists of two important steps, defining nodes and defining edges. In previous studies, nodes are usually represented by different agents in the financial market, that is, stocks or bonds, and edges are symbolized by the relationship between such agents. Pearson's correlation coefficient is the most common and easiest way to measure the correlation between two entities in the financial market [14–20]. For example, McDonald et al. used Pearson correlation coefficient to construct a currency-related network in the global foreign exchange market and obtained temporary dominant or dependent currency information [16]. In addition, other correlation coefficients, such as Spearman rank-order correlation coefficient [21], multifractal detrended cross-correlation analysis (MFCCA) [22,23], multifractal detrended fluctuation analysis (MFDCA) [24], and cophenetic correlation coefficient (CCC) [25], have also been put forward. Furthermore, correlation can also be defined by some econometric methods, such as the Granger causality test [26,27], cointegration test [28], dynamic correlation coefficient with GARCH (DCC-GARCH) [29], and so on.

After the definition of edges, some filter methods for choosing the important edges should be applied. Otherwise, the complex network will be very large and complicated, which is not conducive to subsequent analysis. Minimum spanning tree (MST) can be used for this purpose. After MST operation, the complex network will retain only $N - 1$ edges, where N is the number of the nodes, which greatly facilitates the study of the network topology. At present, MST is most commonly used to simplify the financial complex network [14,15,18–20,22,29–31]. For example, in 1999, Mantegna first proposed that MST could be used to search for important edges in the stock market network, and a stock market topology with economic significance could be obtained [14]. Except for MST, other greedy algorithms similar to MST, such as planar maximally filtered graph (PMFG) [28], can also be

used as filter methods. Furthermore, setting thresholds and only retaining edges with a correlation coefficient greater than the threshold can work as a filter too [17].

After construction of the financial network with the above methods, some structural characteristic information can be obtained by analyzing the network indicators. Previous scholars have found that stock markets in different countries have similar topologies. Zhuang et al. used stock data from the Shanghai stock market from 2002 to 2004 to build an undirected and unweighted network. It was found that China's stock networks had the typical statistical features of complex networks, that is, small world and scale-free property [17]. Tu also selected Chinese stock data and built a stock market complex network. The author calculated some network indicators like degree centrality, PageRank, hyperlink induced topic search (HITS), local clustering coefficient, K-shell, and so on, finding the driving forces of China's financial market [28]. Chi et al. selected the U.S. stock market data to build several weighted networks (using price, price returns, and trading volumes as edges separately), finding that the U.S. stock price networks are scale-free networks, which means the variation of stock prices is strongly influenced by a relatively small number of stocks [29]. Caraiani used the complex network to study the returns of major European emerging countries' stock markets, which also had the characteristic of being scale-free [24]. Furthermore, complex network studies were also conducted on the Japanese stock market [32], Korean stock market [33], Hong Kong stock market [34], and so on.

On the basis of the topological structure of the financial network, the systemic risk in the financial market can be measured and the impact of crisis on the financial system can be analyzed. In general, crisis will change the topology of the financial network, and a more complex topology (or a larger value of structural entropy) often comes with a greater systemic risk [35]. Onnela et al., for example, built a U.S. stock market network with MST and conducted a tree structure analysis on it, finding that the financial crisis caused the tree length to shrink, which means the topological structure of the stock market experienced a downfall when facing systemic risk [15]. Long et al. selected CSI 300 data in China, established a dynamically correlated stock industry network with MST, and studied its connection characteristics and topological structure. The results reveal that industries with large betweenness centrality, closeness centrality, clustering coefficient, and small node occupancy layer are associated with greater systemic risk contribution [29]. He and Deem found that the systemic risk and recession will lead to a more hierarchical structure of global trade networks [25]. Bardoscia et al. also found that when the connectedness in bank networks increases, the financial system's ability to deal with systemic risk will deteriorate [36].

To sum up, according to various distance definitions and filter methods, different financial complex networks can be constructed, thus obtaining different measures of systemic risk. Pearson's correlation coefficients are simple and intuitive, and can effectively measure the correlation between different stocks in the Chinese stock market. Since its establishment in 1990, China's stock market has been rapidly developing for nearly 30 years. While the internal operating mechanism adjusts constantly, China's stock market is also subject to external impacts, so the systemic risk of the market is dynamic. This paper constructs a dynamic complex network based on the stocks daily return data; symbolizes the systemic risk of China's stock market by complex network indicators; decomposes and reconstructs the sequence into three components including trend, cycle, and high frequency disturbance; and, finally, examines the evolution of China's stock market systemic risk.

3. Data and Methodology

This paper establishes a dynamic complex network model utilizing the daily data of all a-shares in Shanghai and Shenzhen stock markets. Because of the small number of stocks at the beginning of the establishment of the Chinese stock market, we selected 90 trading days before the first trading day in 1997 as the starting point, and 29 February 2020 as the end point.

The methodology is shown in Figure 1. First, the correlation coefficients between return rates of stocks were used as edge weights to build dynamic complex networks with a window period of 90 days and 1 day for step. Then, the average weight and structural entropy of the network in each

day can be obtained. The ratio of the stocks with weight in the top 10% to the average weight in each window period can also be calculated and defined as the concentration ratio of important stock. Therefore, four network indexes could be derived. Next, these four network indexes were combined with the stock market index and 0–1 standardization before empirical mode decomposition (EMD) was performed. Through the above process, the original sequences were divided into a number of intrinsic mode functions (IMFs). Then, the results were reconstructed with grey relational analysis (GRA), making each sequence have three items, that is, tendency, cycle, and disturbance. Finally, the statistical analysis of the three components was conducted in order to explore the development of China's stock market and the evolution characteristics of systemic risk. Modeling of the complex network, EMD, and GRA is introduced as follows.

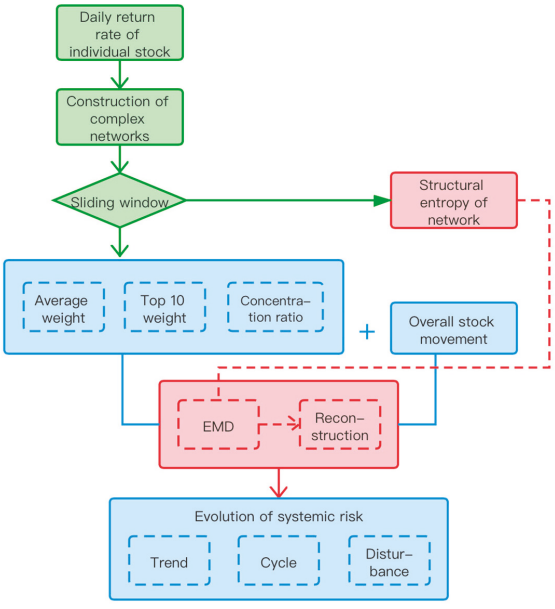


Figure 1. Methodology. EMD, empirical mode decomposition.

3.1. Construction of the Complex Network

A complex network consists of several nodes and edges linking them. The node is the basic element of a complex network, which is the abstract expression of an "individual" in the real world. The edge is an expression of the relationship between the elements and can be given weight according to the extent of the relationships. Here, w_{ij} represents the weight of the edge linking node i and node j, where $i,\ j = 1, 2, 3, \ldots, n$ and n is the number of nodes in a certain network. For an undirected network,

$$w_{ij} = w_{ji} \tag{1}$$

We can also use the weighted degree to represent the importance of nodes, which is defined as

$$dw_i = \sum_{j \in v(i)} w_{ij} \tag{2}$$

where $v_{(i)}$ is the set of nodes linking to node i. The larger the weighted degree, the stronger the degree of correlation with other nodes and the more important the node.

We use the return rates of a-share stocks on China's stock market as the network nodes and construct the network using correlation coefficient ρ_{ij} as the edge weight.

$$w_{ij} = \rho_{ij} = < X_{it} \cdot X_{jt} > \tag{3}$$

Here, $\{X_{it},\ i = 1, 2, \cdots, n; t = 1, 2, \cdots, T\}$ is the original stock return rates data and $< \cdots >$ indicates a time-average over the T data points for each time series.

After we get w_{ij}, we calculate the average weight, top 10 nodes weight, and concentration ratio below:

$$average\ weight = \frac{1}{n} \sum_{i=1}^{n} dw_i \tag{4}$$

$$top\ 10\ nodes = \frac{1}{10} \sum_{i \in top(i)} dw_i \tag{5}$$

$$concentration\ ratio = \frac{top\ 10\ nodes}{average\ weight} \tag{6}$$

where $top_{(i)}$ means the nodes i with the top 10 weights (dw_i).

Furthermore, we calculate the network's structural entropy, which is often used to measure the complexity of the complex network system [37]. However, as the structural entropy of the all-connected network is constant, it is meaningless for our analysis, so we need to remove the edge of weak correlation to get a non-all-connected network for calculating the structural entropy.

The threshold value of the correlation coefficient is set at 0.4. If the absolute value of the correlation coefficient ρ_{ij}, that is, w_{ij}, is less than 0.4, this edge will be cut off, and we will get a non-fully connected network to calculate the structural entropy E_{deg} under each window [37]:

$$E_{deg} = -k \sum_{i=1}^{N} p_i \log p_i \tag{7}$$

where N is the total number of nodes in the network; k is Boltzmann's constant; and p_i can be calculated by the number of edges connecting to node i, namely, the degree of node i:

$$p_i = \frac{degree(i)}{\sum_{i=1}^{N} degree(i)} \tag{8}$$

3.2. Empirical Mode Decomposition

Combining the three network indexes with China's stock market index gives four input data, named as $\{Y_{kt},\ k = 1, 2, 3, 4; t = 1, 2, \cdots, T\}$. Y_{kt} have to be 0–1 standardized, owing to significant differences at the numerical level, that is,

$$Z_{kt} = \frac{Y_{kt} - \overline{Y}_k}{std(Y_k)} \tag{9}$$

For the signal $Z(t)$, the upper and lower envelopes are determined by local maximum and minimum values of the cubic spline interpolation. m_1 is the mean of envelopes. Subtracting m_1 from $Z(t)$ yields a new sequence h_1. If h_1 is steady (does not have a negative local maximum or positive local minimum), it is denoted as the intrinsic mode function (imf_1). If h_1 is not steady, it is decomposed again, until steady series is attained, which is denoted as imf_1. Then, m_1 replaces the original $Z(t)$ and m_2 is the mean of the envelopes of m_1, and m_1 is similarly decomposed. Repeating these processes K times gives imf_k, that is,

$$imf_k = imf_{(k-1)} - m_k \tag{10}$$

Finally, let *res* denote the residual of $Z(t)$ and all imfs:

$$res = X(t) - imf_1 - imf_2 - \cdots - imf_K \tag{11}$$

where imfs and *res* could be extracted for the GRA process.

3.3. Grey Relational Analysis

The grey relational analysis was first put forward by Deng J L in 1989 [38]. His grey relational degree model, which is usually called the grey relative correlation degree, mainly focused on the influence of distance between points in the system.

The grey relative correlation degree formula is given by Equation (12).

$$r_{ij}^1 = \frac{1}{N}\sum_{t=1}^{N} \frac{\underset{j}{min}\ \underset{t}{min}\left|d_i(t) - d_j(t)\right| + \rho\ \underset{j}{max}\ \underset{t}{max}\left|d_i(t) - d_j(t)\right|}{\left|d_i(t) - d_j(t)\right| + \rho\ \underset{j}{max}\ \underset{t}{max}\left|d_i(t) - d_j(t)\right|} \tag{12}$$

where $d_i(t)$ is the reference series; $d_j(t)$ is the compared series; and ρ is the distinguishing coefficient, which is usually equal to 0.5.

In order to overcome the weakness of the grey relative correlation degree, the absolute correlation degree was proposed by Mei (1992) [39]. The formula is given by Equation (13).

$$r_{ij}^2 = \frac{1}{N-1}\sum_{t=1}^{N-1} \frac{1}{1 + \left|d_i(t+1) - d_i(t) + d_j(t+1) - d_j(t)\right|} \tag{13}$$

Considering the weakness and strength, we used the grey comprehensive relational degree to classify the noise terms and market fluctuation terms. The formula of the grey comprehensive relational degree is given by Equation (14):

$$r_{ij} = \beta r_{ij}^1 + (1-\beta) r_{ij}^2 \tag{14}$$

where β is the weight of the grey relative relational degree, which is valued as 0.5.

4. Empirical Analysis

4.1. Dynamic Characteristics of Complex Networks

Figure 2a compares the three average weight related indicator of the dynamic complex network with the dynamic evolution of the Shanghai composite index standardized by setting it as 1000 on the first trading day of 1997. It can be seen that the average weight of the complex network and the average weight of the top 10 stocks have strong synchronization, with a high correlation of 0.9896. Therefore, both of them can be used as proxy indicators of systemic risk. However, the concentration ratio is not consistent with the overall systemic risk. The concentration of risk is relatively low when the systemic risk is high, which means the risk is relatively decentralized. Furthermore, the concentration ratio and the average weight are significantly negatively correlated with a correlation coefficient of −0.91329. In this way, we will focus on using the index of the average weight to measure the systemic risk of the Chinese stock market.

It can also be seen from Figure 2a that, although there is a correlation between two average weight indexes (all and top 10) and the stock index, the coefficients, −0.1370 and −0.0829, are relatively small. This proved that the level of systemic risk is not determined by the move of overall price trend.

In order to further investigate the relationship between the systemic risk represented by the average weight, the Beta value (β) obtained by the CAPM model, and the stock average variance (V), we estimated β_t and V_t as follows:

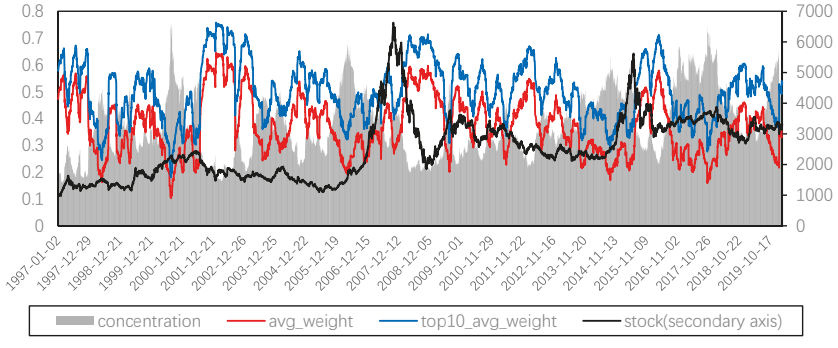
$$X_{kt} = r_f + \beta_{kt}(Y_t - r_f) + e_{kt} \tag{15}$$

$$\beta_t = \frac{1}{N}\sum_{k=1}^{N}\beta_{kt} \tag{16}$$

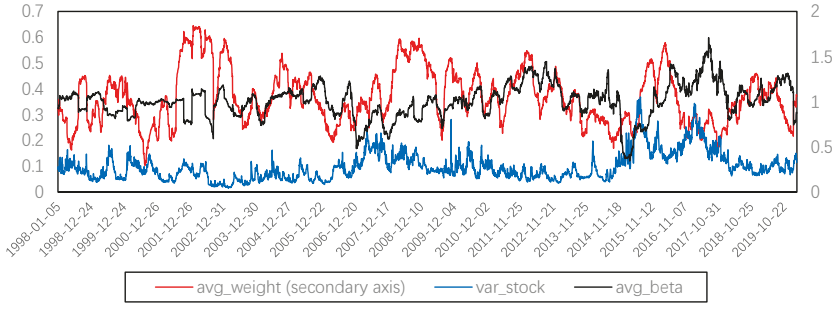
$$V_t = \frac{1}{N}\sum_{k=1}^{N}[\frac{1}{T}\sum_{m=1}^{T}(X_{km} - \overline{X_{kt}})^2] \tag{17}$$

where N is the total number of stocks; T is the length of the sliding window; r_f is the risk-free interest rate, which was set to 3%; X_{kt} is the return of the kth stock in the sliding window t; Y_t is the return of the stock index, which is symbolized for market return and is represented by 000001.SH; β_{kt} is calculated by MLS with Y_t and X_{kt}; e_{kt} is the error term; β_t is the average of all individual stocks' Beta; and V_t is the average variance of all stocks in sliding window t.

In Figure 2b, we compare the systemic risk with Beta and stock variance, finding that these three have different moving trends, which shows that our systemic risk index can catch unique market fluctuations. Furthermore, the systemic risk index was ahead of Beta in several stages, such as from June 2006 to July 2008 or from July 2015 to August 2017, which shows that our systemic risk index has a certain risk pre-warning ability.



(**a**) Dynamic average weight of complex networks.

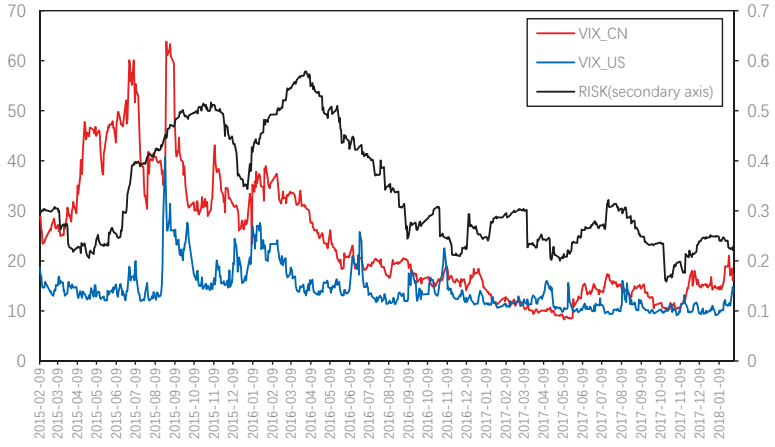


(**b**) Comparing average weight with stock variance and Beta.

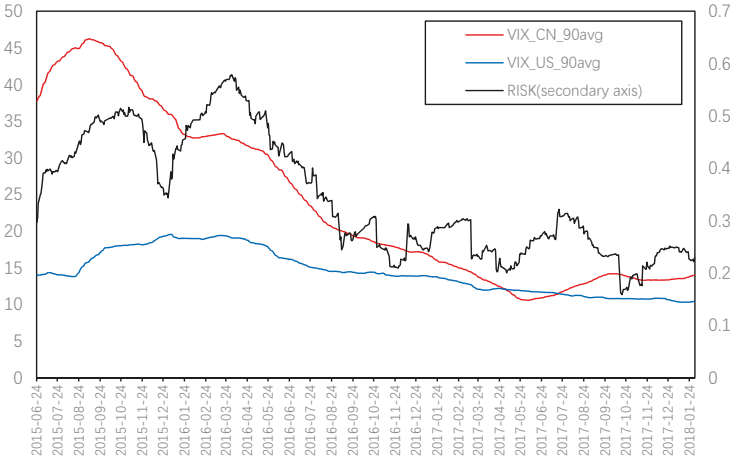
Figure 2. Comparison of complexity measures of systemic risk with other measures.

We further compared the systemic risk represented by average weight with the volatility index (VIX) of China and the U.S. stock market. Considering the Chinese VIX cannot cover the above research range, the U.S. VIX was selected for comparison purposes. The correlation coefficient between the two VIX in this range is significantly positive, but the coefficient is only 0.5626.

Figure 3a presents the great differences in the trend of VIX between China and the United States. It can be seen that the correlation coefficient between average weight and Chinese VIX is 0.4763 during the interval since the Chinese VIX launched. It is noteworthy that the volatility index leads the systemic risk index to a certain extent. This is confirmed by the results obtained from the cross-correlation analysis with the maximum coefficient of 0.7469, corresponding to lags of 55 days (which means current systemic risk is highly related to the VIX from 55 days prior). However, this is mainly because the systemic risk index constructed in this paper was compiled using the sliding window method, with the window length of 90 days, so the systemic risk index of a certain time, t actually represents the systemic risk of the previous 90 days.



(**a**) Volatility index (VIX) and systemic risk.



(**b**) Ninety-day average for VIX and systemic risk.

Figure 3. Comparison of systemic risk with VIX.

In fact, the complex network characteristics of individual stocks are effective at reflecting the systemic risk of the market. To verify this, we calculated the 90-day averages for VIX, which are shown in Figure 3b. It can be seen that the systemic risk index constructed in this paper is consistent with the 90-day average trend for China's VIX, and the systemic risk is ahead of China's VIX after 2017 and is more sensitive, which proves the effectiveness of the systemic risk index derived from the complex network.

Figure 4 shows the comparison between the structural entropy and the number of nodes in a complex network. It can be seen that the structural entropy is highly correlated with the number of nodes, and the correlation coefficient reaches 0.9302. In other words, the increase in system complexity of China's stock market is mainly caused by the increase in the number of listed companies. Nevertheless, we can also find that, in addition to the overall upward trend, structural entropy also has periodic fluctuations. Therefore, multi-scale analysis is required to determine whether the system complexity represented by structural entropy is related to systemic risk.

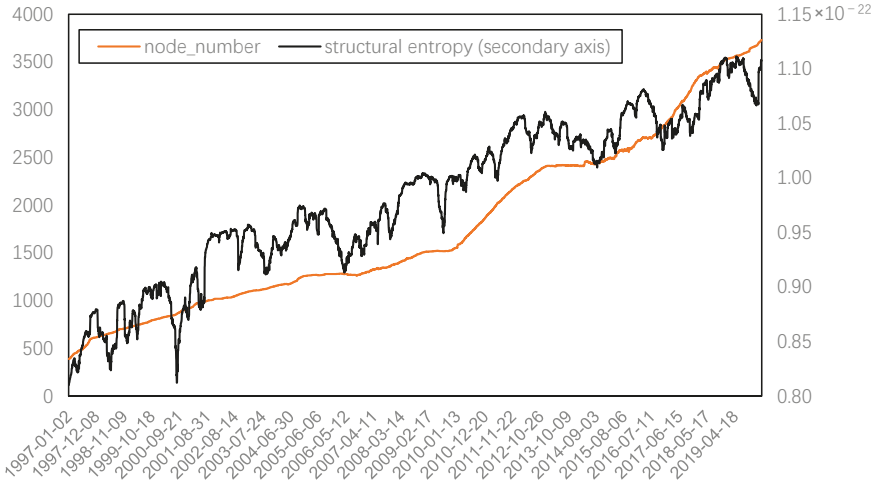
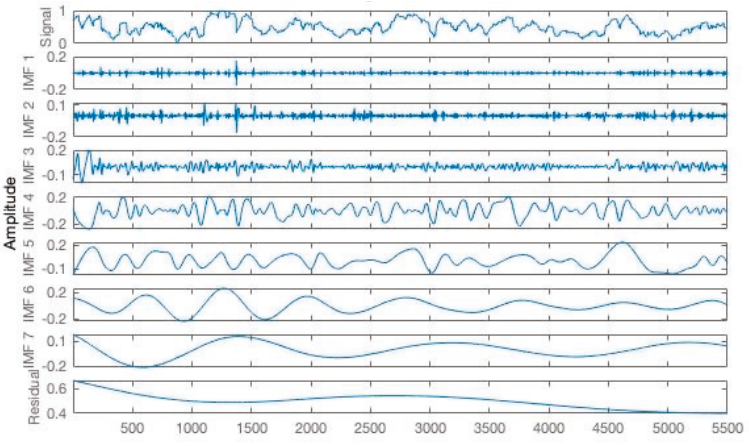


Figure 4. Dynamic structural entropy of complex networks.

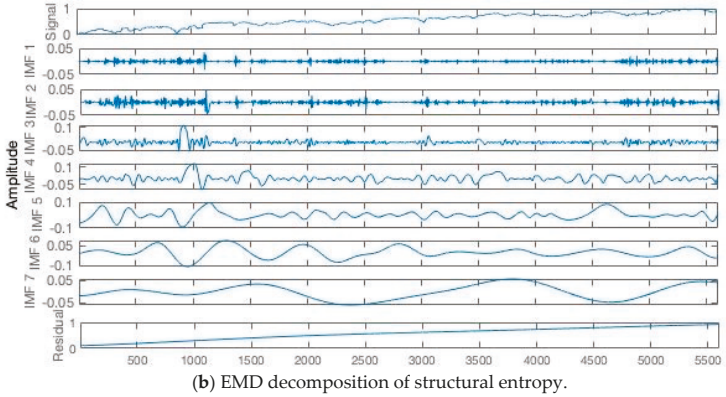
4.2. Decomposition and Reconstructione

Figure 5 presents the EMD results of the standardized systemic risk index, structural entropy, and stock price index, respectively. It can be seen that the two original sequences are divided into seven IMFs and one residual term, among which the residual term can represent the overall trend of indexes' evolution to a certain extent, while the IMF of lower frequency can describe the periodic fluctuation of indexes in different time scales, and the IMF of highest frequency represents the stochastic perturbation.

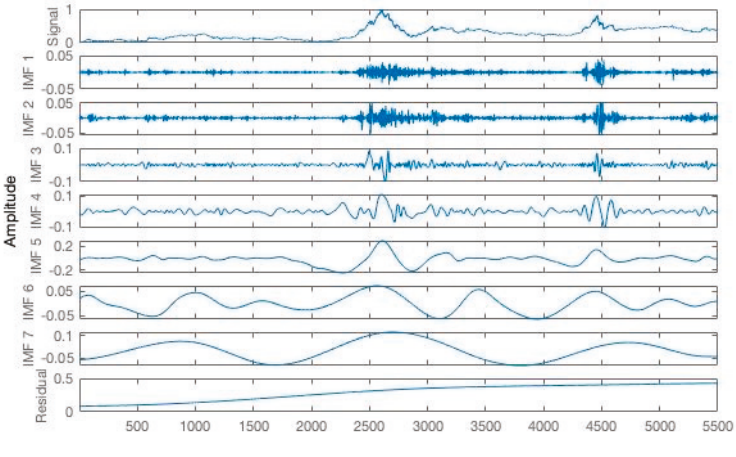
Through EMD, it can be found that the residual term, also known as the trend term, decomposed by structural entropy, represents the growth in the number of network nodes, and the correlation coefficient between this residual term and the number of network nodes can be further improved to 0.9428. When removing the trend term from the original sequence and comparing it to the systemic risk series represented by the average weight, as shown in Figure 6, the highly consistent fluctuations between the two series can be seen, and the correlation coefficient of the two reaches 0.7572. Therefore, adjusted structural entropy, that is, removing the trend term of the network size, can also measure the systemic risk. Nevertheless, owing to the high correlation between these two series, the following analysis only focuses on the systemic risk represented by average weight.



(**a**) EMD decomposition of systemic risk.



(**b**) EMD decomposition of structural entropy.



(**c**) EMD decomposition of stock price index.

Figure 5. Results of EMD decomposition. IMF, intrinsic mode function.

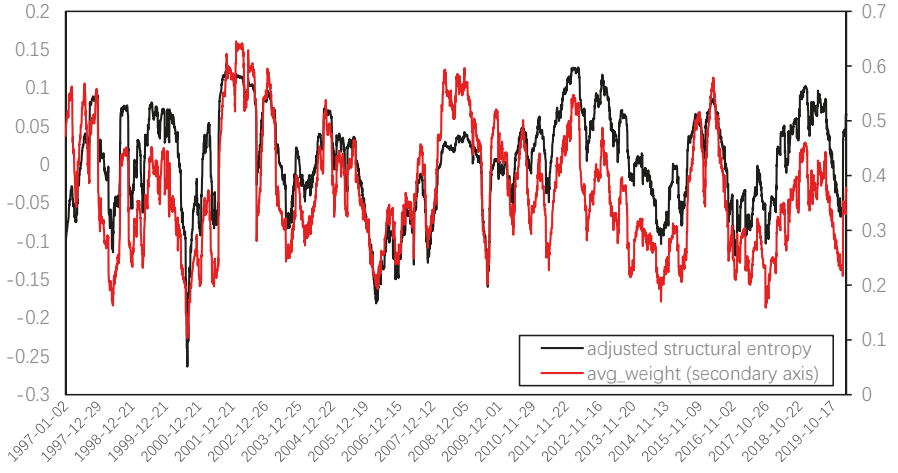


Figure 6. Adjusted structure entropy and average weight.

In order to further observe the systemic risk evolution of the Chinese stock market, several IMFs and residual terms obtained from EMD decomposition were combined using the method of grey correlation degree. Figures 7 and 8 present the trend term, cycle term, and random term of systemic risk (average weight) and the stock price index. Then, we focused on the overall trend change and cycle fluctuation of systemic risk in China's stock market.

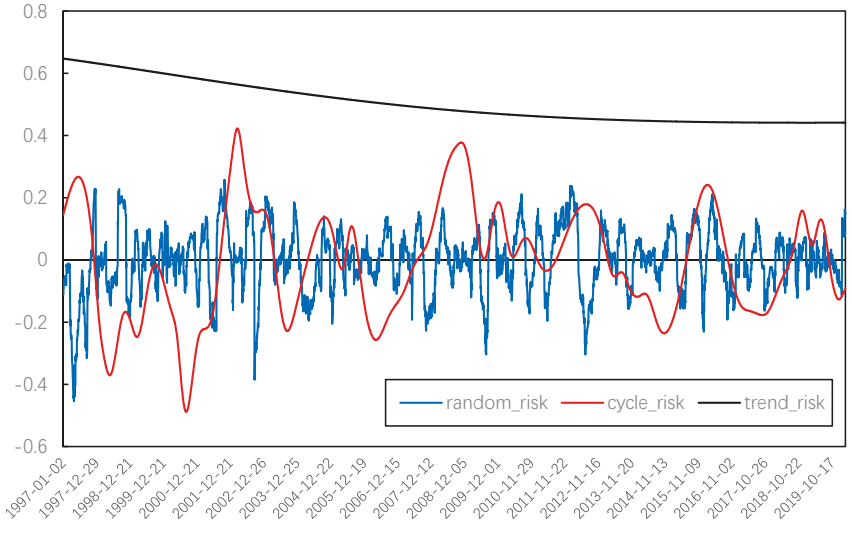


Figure 7. IMFs reconstruction of systemic risk.

For the long-term tendency, we found that the overall trend of the stock price rose steadily, while the systemic risk has been declining slowly throughout the evolution of the Chinese stock market since 1997. This means that, although there is still phased systemic risk in the Chinese stock market, the overall level of systemic risk is declining as the operating mechanism and related regulations are constantly improving.

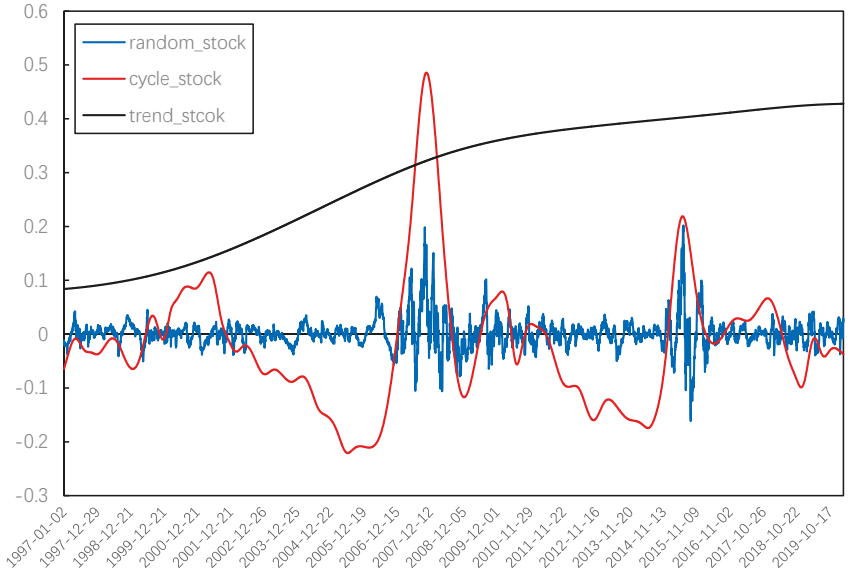


Figure 8. IMFs reconstruction of stock price index.

For the cycle fluctuation, the rise of systemic risk is usually caused by the joint action of external shocks and internal operations, which is manifested in the excessive rise and fall in the stock market. Therefore, the cyclical characteristics of systemic risk have no direct relationship with the fluctuations of the stock market. Thus, we converted the cycle fluctuation of the stock market into the difference from the price mean using (18).

$$cycle\_abs\_stock = abs[cycle_{stock} - average(cycle_{stock})] \tag{18}$$

Considering that *cycle_abs_stock* and *cycle_risk* are both non-stationary, we calculated their first-order differences. The results of Augmented Dickey–Fuller (ADF) tests show that both variables are an integrated of order one. Therefore, cointegration tests can be proposed on the original sequences. The results of Johnson Trace tests show that there are at least two cointegration relationships between the two variables, which confirms that there is a long-term equilibrium relationship between stock price volatility and systemic risk. The equilibrium equation is

$$cyclel\_abs\_stock = 0.0660 cycle\_risk + 0.0625. \tag{19}$$

All the coefficients are significant at the 5% significance level, so the volatility of the stock market is positively related to systemic risk from the perspective of long-term equilibrium, which means that, while the stock price deviates from the theoretical value of equilibrium, the systemic risk will be at a high level.

In Figure 9, when the blue line is above 0, the systemic risk is large, while when the blue line is below 0, the systemic risk is small. The red line represents the absolute value of stock price movements, and the red line is clearly ahead of the above-zero parts of the blue line.

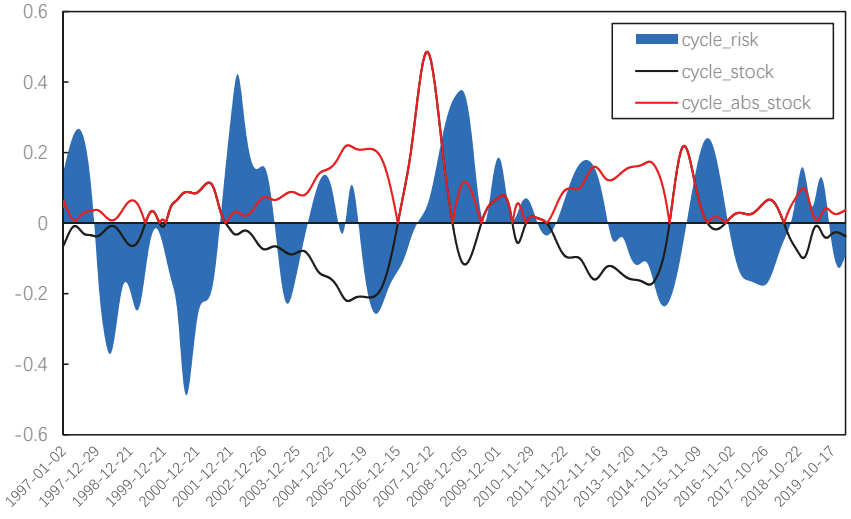


Figure 9. Cycle evolution of systemic risk.

Figure 9 shows the cycle evolution of systemic risk, the lead-lag relationship between systemic risk and stock volatility is dynamic owing to the sliding window processing. From the perspective of the whole cycle evolution, we found that there were several periods of high systemic risk in the Chinese stock market since 1997, as described below.

1997–1998: The stock market was in a shock stage during this period. On the one hand, the Chinese stock market was impacted by external factors such as the Asian financial crisis; on the other hand, the operating mechanism at that time was not perfect enough, with frequent insider trading and market manipulation. The systemic risk was at a high level during this period and, therefore, the stock market began to comprehensively reform its trading mechanism in 1998. Although the market fluctuation was not violent from the current perspective, it actually contained many factors causing systemic risk.

2001–2002: The stock market was in a declining bear stage during this period. Owing to the poor performance of high-tech companies, resulting from the burst of the global Internet bubble and the launch of the policy reducing the state-owned shares holding of listed companies, the stock market had a big crash in China. Related departments issued a series of favorable strategies such as reducing interest rates and trading commissions; however, the imperfection of the market led to a number of "black markets", which brought a high systemic risk.

2007–2008: This period includes both excessive rise and fall of the market. The reform of non-tradable shares in 2005, together with a series of positive policies such as the entry of insurance funds and the appreciation of renminbi (RMB), promoted the rise of the Chinese stock market. However, a lot of speculation by inexperienced individual investors caused a more and more serious herding effect, and the systemic risk was maintained at a high level for a long time. Followed by the global financial crisis brought by the U.S. subprime crisis, with the launch of stock index futures, the Chinese stock market began to reverse to a bear stage, and the systemic risk in this stage also remained at a high level.

2011–2012: This stage was another volatile bear market; the fluctuation of stock price was much smaller than that of the previous stage, but the systemic risk still remained at a similar level. Even though the Chinese economy maintained a high growth rate during this period, the stock market was influenced by the global financial markets, as well as the European debt crisis. The low volatility of the stock market still contained large systemic risks, which were reinforced by the frequent occurrence of black swan events such as a rear-end collision of bullet trains, clenbuterol, and so on.

2015–2016: The market price was rising rapidly in 2015 and the systemic risk was also in a climbing stage. However, a high level of risk still appeared in 2016, which was a stage of rapid and frequent fluctuations. The issuing scale of new stocks increased significantly, driving frequent market shocks such as thousands of shares rising or falling together, two triggering circuit breaker events in a day, and so on. Thus, the overall capital presents a large-scale net outflow, and the investor sentiment fluctuates abnormally.

To summarize, the systemic risk of the stock market will significantly increase in the irrational stages of rise, fall, and frequent shocks. However, extremely high systemic risk is more likely in the cases of collapse and frequent shocks.

5. Discussion

Complex networks have been widely used in the field of socio-economic analysis. Most of them focus on the risk contagion of banks and international economic or trade exchanges; however, studies on the stock market are limited. In fact, a complex network provides an important tool for the study of the stock market, which is a self-organizing complex system with multi-agent interactions. The average weight of the complex network can be used to measure the aggregation of positive feedback in the market, so as to measure the overall systemic risk.

On the basis of the data of all a-shares in China, this paper constructs a dynamic complex network of stock correlation, and the change of average weight as well as adjusted structural entropy of the network are used to measure the evolution of systemic risk in China's stock market. Although, owing to the use of a sliding window, the average weight or structural entropy in fact presents the average systemic risk level in the past 90 days, it also reflects the evolution of systemic risk in China's stock market for more than 20 years as a whole. The results show that the systemic risk of China's stock market shows a downward trend on the whole, which is closely related to the continuous improvement of the management system and operation mechanism of the financial market. In addition, there is a long-term equilibrium relationship between the cycle fluctuation of systemic risk and the excessive fluctuation of the stock market. Since 1997, the stages with high systemic risk have appeared with excessive increases, excessive falls, and frequent fluctuations of the stock market. Meanwhile, it can also be seen from Figure 1 that the global stock market began to fluctuate significantly under the influence of the novel coronavirus pneumonia. The Chinese stock market is relatively stable at present, but the systemic risk has been climbing rapidly since the beginning of February. Therefore, we must be alert to the further expansion of the systemic risk of the Chinese stock market under the double impact of internal and external factors.

Author Contributions: Data curation, Y.Z.; Formal analysis, K.G.; Funding acquisition, Y.S. and K.G.; Investigation, Y.Z., K.G., Z.J., and Z.H.; Methodology, K.G.; Software, Y.Z.; Supervision, Y.S.; Writing—original draft, Z.J. and Z.H.; Writing—review & editing, K.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the University of Chinese Academy of Sciences, the National Natural Science Foundation of China No. 71501175 and No. 91546201. National Science Foundation of China, Key Project, #71932008.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Scharfstein, D.S.; Stein, J.C. Herd behavior and investment. *Am. Econ. Rev.* **1990**, *80*, 4654–4679.
2. Bikhchandani, S.; Sharma, S. Herd behavior in financial markets. *IMF Staff Pap.* **2000**, *47*, 279–310.
3. Li, X.; Sullivan, R.N.; Garcia, F.L. The low-volatility anomaly: Market evidence on systematic risk vs. mispricing. *Financ. Anal. J.* **2016**, *72*, 36–47. [CrossRef]
4. De, S.R.A. Unobservable systematic risk, economic activity and stock market. *J. Bank. Financ.* **2018**, *97*, 51–69.

5. Price, K.; Price, B.; Nantell, T.J. Variance and lower partial moment measures of systematic risk: Some analytical and empirical results. *J. Financ.* **1982**, *37*, 843–855. [CrossRef]
6. Estrada, J. Systematic risk in emerging markets: The D-CAPM. *Emerg. Mark. Rev.* **2002**, *3*, 365–379. [CrossRef]
7. Caporale, T. Time varying CAPM betas and banking sector risk. *Econ. Lett.* **2012**, *115*, 293–295. [CrossRef]
8. Iqbal, M.J.; Shah, S.Z.A. Determinants of systematic risk. *J. Commer.* **2012**, *4*, 47.
9. Mantegna, R.N.; Stanley, H.E. *An Introduction to Econophysics: Correlations and Complexity in Finance*; Cambridge University: Cambridge, UK, 1999.
10. Kwapień, J.; Drożdż, S. Physical approach to complex systems. *Phys. Rep.* **2012**, *515*, 115–226. [CrossRef]
11. Costa, L.F.; Oliveira, J.O.N.; Travieso, G.; Rodrigues, F.A.; Boas, P.R.V.; Antiqueira, L.; Viana, M.R.; Enrique, L. Analyzing. and modeling real-world phenomena with complex networks: A survey of applications. *Adv. Phys.* **2011**, *60*, 329–412. [CrossRef]
12. D'Arcangelis, A.M.; Rotundo, G. Complex networks in finance. *Lect. Notes Econ. Math. Cham.* **2016**, *683*, 209–235.
13. Nagurney, A. Networks in economics and finance in Networks and beyond: A half century retrospective. *Networks* **2019**, 14–15. [CrossRef]
14. Mantegna, R.N. Hierarchical structure in financial markets. *Eur. Phys. J. B* **1999**, *11*, 193–197. [CrossRef]
15. Onnela, J.P.; Chakraborti, A.; Kaski, K.; Kertesz, J.; Kanto, A. Dynamics of market correlations: Taxonomy and portfolio analysis. *Phys. Rev. E* **2003**, *68*, 056110. [CrossRef]
16. McDonald, M.; Suleman, O.; Williams, S.; Howison, S.; Johnson, N.F. Detecting a currency's dominance or dependence using foreign exchange network trees. *Phys. Rev. E* **2005**, *72*, 046106. [CrossRef]
17. Zhuang, X.; Min, Z.; Chen, S. Characteristic analysis of complex network for Shanghai Stock Market. *J. Northeast. Univ. Nat. Sci.* **2007**, *28*, 1053.
18. Chi, K.; Liu, T.J.; Lau, C.M.F. A network perspective of the stock market. *J. Empir. Financ.* **2010**, *17*, 659–667.
19. Nobi, A.; Lee, J.W. Systemic risk and hierarchical transitions of financial networks. *Chaos* **2017**, *27*, 063107. [CrossRef]
20. Liao, Z.; Wang, Z.; Guo, K. The dynamic evolution of the characteristics of exchange rate risks in countries along "The Belt and Road" based on network analysis. *PLoS ONE* **2019**, *14*. [CrossRef]
21. Miccichè, S.; Bonanno, G.; Lillo, F.; Mantegna, R.N. Degree stability of a minimum spanning tree of price return and volatility. *Physica A* **2003**, *324*, 66–73.
22. Kwapień, J.; Oświęcimka, P.; Forczek, M.; Drożdż, S. Minimum spanning tree filtering of correlations for varying time scales and size of fluctuations. *Phys. Rev. E* **2017**, *95*, 052313. [CrossRef]
23. Wang, Y.; Zheng, S.; Zhang, W.; Wang, G.; Wang, J. Fuzzy entropy complexity and multifractal behavior of statistical physics financial dynamics. *Physica A* **2018**, *506*, 486–498. [CrossRef]
24. Caraiani, P. Characterizing emerging European stock markets through complex networks: From local properties to self-similar characteristics. *Physica A* **2012**, *391*, 3629–3637. [CrossRef]
25. He, J.; Deem, M.W. Structure and response in the world trade network. *Phys. Rev. Lett.* **2010**, *105*, 198701. [CrossRef]
26. Wang, Z.; Yan, Y.; Chen, X. Time and frequency structure of causal correlation networks in the China bond market. *Eur. Phys. J. B* **2017**, *90*, 137. [CrossRef]
27. Billio, M.; Getmansky, M.; Lo, A.W.; Pelizzon, L. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *J. Financ. Econ.* **2012**, *104*, 535–559. [CrossRef]
28. Tu, C. Cointegration-based financial networks study in Chinese stock market. *Physica A* **2014**, *402*, 245–254. [CrossRef]
29. Long, H.; Zhang, J.; Tang, N. Does network topology influence systemic risk contribution? A perspective from the industry indices in Chinese stock market. *PLoS ONE* **2017**, *12*. [CrossRef]
30. Kwapień, J.; Gworek, S.; Drożdż, S.; Górski, A. Analysis of a network structure of the foreign currency exchange market. *J. Econ. Interact. Coord.* **2009**, *4*, 55. [CrossRef]
31. Bonanno, G.; Caldarelli, G.; Lillo, F.; Mantegna, R.N. Topology of correlation-based minimal spanning trees in real and model markets. *Phys. Rev. E* **2003**, *68*, 046130. [CrossRef]
32. Konishi, M. A global network of stock markets and home bias puzzle. *Appl. Financ. Econ. Lett.* **2007**, *3*, 197–199. [CrossRef]
33. Lee, K.E.; Lee, J.W.; Hong, B.H. Complex networks in a stock market. *Comput. Phys. Commun.* **2007**, *177*, 186. [CrossRef]

34. Li, P.; Wang, B. An approach to Hang Seng Index in Hong Kong stock market based on network topological statistics. *Chin. Sci. Bull.* **2006**, *51*, 624–629. [CrossRef]
35. Yalamova, R.; McKelvey, B. Explaining what leads up to stock market crashes: A phase transition model and scalability dynamics. *J. Behav. Financ.* **2011**, *12*, 169–182. [CrossRef]
36. Bardoscia, M.; Battiston, S.; Caccioli, F.; Caldarelli, G. Pathways towards instability in financial networks. *Nat. Commun.* **2017**, *8*, 14416. [CrossRef]
37. Lu, K.; Yang, Q.; Chen, G. Singular cycles and chaos in a new class of 3D three-zone piecewise affine systems. *Chaos J. Nonlin. Sci.* **2019**, *29*, 043124. [CrossRef]
38. Deng, J. Introduction to grey system theory. *J. Grey Syst.* **1989**, *1*, 1–24.
39. Mei, Z. The concept and computation method of grey absolute correlation degree. *Syst. Eng.* **1992**, *10*, 43–44.

entropy

MDPI

Article

The Threshold Effect of Leveraged Trading on the Stock Price Crash Risk: Evidence from China

Zhen Peng [1] and Changsheng Hu [2,*]

[1] School of Business, Hubei University, Wuhan 430062, China; pengzhen@hubu.edu.cn
[2] Economics and Management School, Wuhan University, Wuhan 430072, China
* Correspondence: hcs_xj@whu.edu.cn

Received: 4 February 2020; Accepted: 24 February 2020; Published: 26 February 2020

Abstract: The stock price crash constitutes one part of the complexity in the stock market. We aim to verify the threshold effect of leveraged trading on the stock price crash risk from the perspective of feedback trading. We empirically demonstrate that leveraged trading has a threshold effect on the stock price crash risk on the basis of monthly data on leveraged trading in the Chinese stock market from January 2014 to December 2016. At a low leverage ratio, leveraged trading reduces the stock price crash risk; however, as the leverage ratio increases and exceeds a certain threshold, leveraged trading asymmetrically increases the stock price crash risk. These findings provide new insights in understanding the complexity in the Chinese stock market.

Keywords: leveraged trading; stock price crash risk; threshold effect; complexity in stock market

1. Introduction

The financial markets are very complex systems; factors of both the internal and the external origin are strongly interrelated by a largely unknown network of connections and feedbacks (positive and negative) [1]. Behavioral finance theory thinks that the abnormal volatility (according to the Efficient Market Hypothesis [2], changes in asset prices should be driven entirely by fundamental information, and the volatility of asset prices should equal the volatility of fundamental factors. Therefore, abnormal volatility is defined as the portion of the volatility of asset prices that exceeds the boundary of the volatility of fundamental factors [3]) is closely related to investors' feedback trading [3]. In addition, Feedback trading is a very common phenomenon in the financial markets and it can have a significant impact on the complexity of asset prices behavior [4].

The stock price crash constitutes one part of the complexity in the stock market [1]. Controversy surrounding the role of leveraged trading in the stock price crash is ongoing. Some scholars think that leveraged trading will mitigate the impact of arbitrage restrictions and reduce stock price crashes in the stock market [5–7]. However, other scholars assert that leveraged trading might expand the influence of private information in the stock market, thus inducing price speculation and increasing stock price crashes [8]. With the establishment of a leveraged trading system in the Chinese stock market in 2010, the Chinese academic community has begun to examine this issue. Some scholars provided evidence that the introduction of short selling can eliminate stock price bubbles and improve market pricing efficiency [9,10]; others reported that leveraged trading will cause the stock price to be overvalued during a bull market and increase the stock price crash risk [11,12].

The severe asymmetry of leveraged trading in the Chinese stock market is an important factor affecting the stock price crash risk. Being constrained by insufficient securities, the high cost of short selling, and the over-optimism of investors during the bull market, the volume of short selling is vastly inferior to the volume of margin trading in the Chinese stock market. The severe asymmetry between margin trading and short selling has weakened the effect of short selling on suppressing

price overvaluing [13]. "Strong margin trading and weak short selling" thus characterize leveraged trading. However, leveraged trading plays a role in eliminating arbitrage restrictions and asymmetry also causes more restrictions, thereby forming a complex positive feedback loop, resulting in drastic price volatility, and possibly causing stock prices to crash.

We aim to study the effects of leveraged trading on the stock price crash risk from the perspective of feedback trading. Sentiment feedback trading, changes in asset prices affect investor sentiment, and investor sentiment in turn affects asset prices, is the most representative and well-known in all the types of feedback trading. The premise of sentiment feedback trading is that investor sentiment has a systematic effect on asset prices [14]. Sentiment feedback trading is characterized by the fact that the fundamental factors in the feedback are exogenous variables, and investor sentiment and asset prices have no direct impact on the fundamental factors. Lots of psychological experimental evidence verifies the sentiment feedback trading [15,16].

The motivation of the research is focused on testing whether the leverage ratio has a threshold effect on the stock price crash risk. We empirically demonstrate that leveraged trading has a threshold effect on the stock price crash risk on the basis of monthly data on leveraged trading in the Chinese stock market from January 2014 to December 2016. Specifically, changes in the leverage ratio are closely positively correlated with the stock price crash risk. Under a low leverage ratio, leveraged trading reduces price volatility and the stock price crash risk. However, leveraged trading asymmetrically increases the stock price crash risk, as the leverage ratio increases and exceeds a certain threshold.

This paper adds to the growing literature on the complexity in the stock market by testing whether leveraged trading has a threshold effect on the stock price crash risk, which will provide new insights in understanding the complexity in the Chinese stock market.

A breakdown of the paper is structured, as follows: Section 2 gives the introduction for hypothesis and methodology, Section 3 shows data preprocessing and descriptive statistics, Section 4 gives results presentation and discussion, Section 5 performs some robustness checks, and Section 6 provides the conclusions.

2. Hypothesis and Methodology

2.1. Hypothesis

The following relationship between the leverage ratio and investors' feedback trading is proposed on the basis of the sentiment mechanisms of the preferences and beliefs of investors, such as "myopic loss aversion" [17], "mental accounting" [18], "gambling preferences" [19], "realization utility" [20], "herd behavior" [21], "regret aversion bias" [22], and "heterogeneous beliefs" [7,23]: at low leverage ratios, investors tend to exhibit the positive feedback trading behavior of "chasing up and down" to gain higher returns, and at high leverage ratios, investors are more inclined to show the negative feedback trading behavior of "selling high and buying low" to "lock in" profits and avoid losses [24]. Therefore, after the introduction of leveraged trading, the feedback trading pattern of investors is no longer completely random, but somewhat certain. In the model of Hu and Peng [25], it showed that leveraged trading could affect investor sentiment and investors' feedback trading behavior. When the leverage ratio exceeds a certain threshold, the strong shift of feedback trading from positive to negative will cause stock price crashes. Therefore, we propose the hypothesis, as follows.

Hypothesis 1. *The leverage ratio has a threshold effect on the stock price crash risk.*

2.2. Sample Data

July 2014 is generally thought to be the starting point for the abnormal volatility in the Chinese stock market, after which the market entered a bull market and the prices rose rapidly [13]. The prices began to crash in mid-June 2015. Notably, from 12 June to 9 July 2015, the Shanghai Composite Index fell by nearly 35%. At the beginning of January 2016, the Shanghai Composite Index continued to fall.

We selected monthly sample data from January 2014 to December 2016 to comprehensively study the mechanism of the leverage ratio on the stock price crash risk under abnormal volatility. The stocks that were eligible for leveraged trading during this period were then selected (the initial sample size is 950 stocks). These data used in this paper were all obtained from the CSMAR (China Stock Market & Accounting Research) database and the Wind database in China.

The sample period of 36 months from January 2014 to December 2016 includes the whole period of abnormal volatility in the recent Chinese stock market, including the phase of stock prices skyrocketing and the phase of stock prices crash. We cannot only study the relationship between leveraged trading and investor sentiment based on this sample, but also the relationship between leveraged trading and abnormal volatility. This is also the main objective of this work. What is more, after 2016, leveraged trading in the Chinese stock market was, to some extent, restricted by the regulations, and was not a completely spontaneous market behavior.

2.3. Definition of Core Variables

2.3.1. Leverage Ratio

In this work, market-level data on leveraged trading were used to measure the leverage ratio. The leverage ratio is defined, as shown in Equation (1):

$$tradingleverage_t = \frac{financing_t - shorting_t}{marketvalue_t} \tag{1}$$

where *tradingleverage* represents the leverage ratio, *financing* represents the total margin trading balance, *shorting* represents the total short selling balance, and *marketvalue* represents the total market value. The time horizon is monthly. Equation (1) reflects the proportion of the net margin trading in the market at month t to its total tradable market value. The larger the ratio, the higher the proportion of margin trading in the market, which results in a higher leverage ratio.

2.3.2. Stock Price Crash Risk

At present, many types of indicators are used to measure the stock price crash risk. The first is a volatility indicator, which includes volatility, amplitude, and cumulative volatility. The more severe the price volatility, the greater the possibility of a price crash. The second is the distribution of a stock return; if the distribution of stock return is extremely negative (leftward), then the stock return has a large tail risk and it is likely to crash. We measured the stock price crash risk in terms of the skewness of the distribution of a stock return, specifically *DUVOL* (down-to-up volatility) [26,27]. We also used other indicators, such as *NCSKEW* (negative coefficient of skewness) [26,27] and *volatility*, for the empirical test. However, *DUVOL* is consistent with the reality of the stock price crash risks in the Chinese stock market.

We use the model that is shown in Equation (2) to estimate the return of an individual stock after market risk adjustment:

$$r_{i,k} = \alpha + \beta_{1,i} \times r_{m,k-2} + \beta_{2,i} \times r_{m,k-1} + \beta_{3,i} \times r_{m,k} + \beta_{4,i} \times r_{m,k+1} + \beta_{5,i} \times r_{m,k+2} + \varepsilon_{i,k} \tag{2}$$

where $r_{i,k}$ is the return of stock i on day k and $r_{m,k}$ is the average return of all stocks on day k while using the weighted market value. In Equation (2), in addition to adding $r_{m,k}$, we add two lagged terms ($r_{m,k-1}$ and $r_{m,k-2}$) of market return and two leading terms ($r_{m,k+1}$ and $r_{m,k+2}$) of market return to eliminate the effects of asynchronous stock trading. $\varepsilon_{i,k}$ represents the residual after regression. $W_{i,k} = ln\left(1 + \varepsilon_{i,k}\right)$ determines the idiosyncratic return of stock i on day k after market risk adjustment.

Subsequently, used the regression result of Equation (2) to construct *DUVOL*. The daily return data of stock i was divided into the up phase (where $W_{i,k}$ was more than $\overline{W}$) and the down phase (where $W_{i,k}$ is less than $\overline{W}$), depending on whether $W_{i,k}$ is greater than the monthly average return

$\overline{W}$. Subsequently, we calculated the standard deviation of the returns in these two phases. *DUVOL* is defined, as shown in Equation (3):

$$DUVOL_{i,k} = ln\{[(n_u - 1)\sum_{down} W_{i,s}^2]/(n_d - 1)\sum_{up} W_{i,s}^2] \quad (3)$$

where $n_u(n_d)$ represents the days, in which is $W_{i,k}$ more or less than $\overline{W_i}$. The larger the value of *DUVOL*, the more serious the negative bias (leftward) of the stock's idiosyncratic return, which also indicates a higher risk of a stock price crash.

A price crash across the entire market is unavoidable when the risk of a price crash for the main stocks in the market is high. We first calculated the monthly *DUVOL* at the individual stock level and then used the average trend of all stocks' monthly *DUVOL* values to represent the stock price crash risk at the market level.

2.4. Empirical Model

We studied the threshold effect of leverage on the stock price crash risk at the market level. The regression model is defined, as shown in Equation (4):

$$\begin{aligned} CrashRisk_t = \alpha + \beta_1 \times tradingleverage_{t-1} + \beta_2 \times dummies \times tradingleverage_{t-1} \\ +\varphi \times ControlVariables_{t-1} + \sum timevariables + \varepsilon_t \end{aligned} \quad (4)$$

where the explained variable *CrashRisk* represents the stock price crash risk at the market level (the average value of all stocks' monthly *DUVOL*). The explanatory variable *tradingleverage* represents the market leverage ratio and *dummies* represents the dummy variables that are used for setting the time threshold and leverage ratio threshold in this work. *dummies* examine whether the impact of the leverage ratio on the stock price crash risk varies with different time thresholds and different leverage ratio thresholds of the Chinese stock market.

In this study, the following control variables were added to the regression model: (1) *retn* represents the stock index return; (2) *lnsize* represents a natural logarithm of the total market value; (3) *turnover* represents the stock turnover rate, which reflects investor sentiment [28]; and, (4) *illiquidity* represents market illiquidity. The ratio of the absolute value of the market return to trading volume is expressed as $illiquidity_t = \left|\frac{retn_t}{tradingvolume_t}\right|$ [29]. Per common practice, we used a first-order lag of all the explanatory variables on the right side of the model [11]. Time dummy variables are also included as controls for the time trend, which could eliminate the risk of spurious regression caused by non-stationarity with time series. Including the robust standard error in the regression controls heteroscedasticity. The time horizon is monthly.

3. Data Preprocessing and Descriptive Statistics

3.1. Data Preprocessing

We screened the sample stocks and removed the following sample stocks: (1) financial stocks; (2) stocks with abnormal trading status, including ST, * ST, and delisted stocks; (3) stocks in their IPO month; (4) stocks with a long-term suspension (less than 220 trading days in one year); and, (5) stocks that have been disqualified from leveraged trading by the stock exchanges in China. In the Chinese stock market, ST denotes a stock that is specially treated due to bad financial issues; * ST denotes a stock that is specially treated to warn of the risk of the listing being terminated. 815 stocks remained in the sample after the above screening. All continuous variables were subjected to winsorizing according to the 1% and 99% quantiles to control for the influence of extreme values. The missing values in the data were filled while using the linear interpolation method. We then calculated the average of all continuous variables to obtain the average of all sample stocks, which represented the variables at the market level.

3.2. Descriptive Statistics

Since the launch of leveraged trading in the Chinese stock market on 31 March 2010, the leveraged trading business has developed rapidly. The balance of margin trading and short selling in the Chinese stock market increased from 12,772 billion RMB in 2010 to 967,961 billion RMB at the end of 2016 and it has continued to grow very rapidly, as shown in Table 1. However, the structural imbalance is serious in China's margin trading and short selling business. In Table 1, generally speaking, a notable feature is that the size of margin trading accounts for more than 99% of the total balance of leveraged trading, whereas short selling always accounts for less than 1%. Margin trading occupies most of leveraged trading, and the function of short selling is seriously limited. Leveraged trading exhibits are characterized by strong margin trading and weak short selling, which strengthens the arbitrage restrictions in the Chinese stock market. Table 1 shows that the size of margin trading has increased rapidly since 2010; margin trading peaked during the peak of the bull market in 2015 and then began to decline.

Table 1. Annual changes in the balance of Chinese leveraged trading.

Year	Total Balance of Leveraged Trading (Billion RMB)	Balance of Margin Trading (Billion RMB)	The Balance of Short Selling (Billion RMB)	Proportion of Margin Trading in Total Leveraged Trading	Proportion of Short-Selling in Total Leveraged Trading
2010	12.7720	12.7610	0.0110	0.999	0.001
2011	38.2070	37.5480	0.0650	0.983	0.002
2012	89.5160	85.6940	3.8210	0.957	0.043
2013	346.5270	343.700	3.0570	0.991	0.009
2014	1025.6560	1017.3730	8.2830	0.992	0.008
2015	1174.2670	1171.3070	2.9600	0.997	0.003
2016	967.9610	963.6710	4.2900	0.996	0.004

Data source: The Wind database in China.

Table 2 lists the descriptive statistical results of the variables. The stock price crash risk and leverage ratio show a significantly high volatility at the market level.

Table 2. Descriptive statistics of variables.

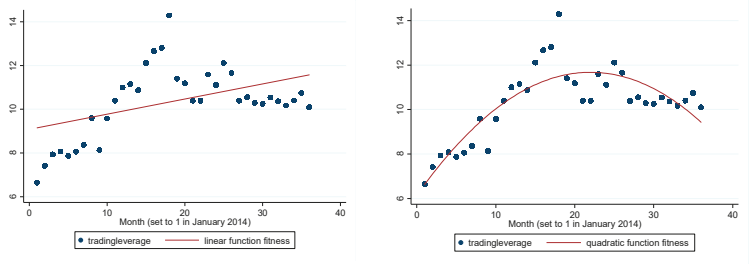
Variable	Sample Size	Mean	Standard Deviation	Min	Max
Explained variable: Stock price crash risk					
CrashRisk	36	−10.4478	1.6882	−13.4223	−6.4126
Core explanatory variable: Leverage ratio					
tradingleverage	36	10.4236	1.5542	6.6421	14.2860
Control variables					
retn	36	2.2086	9.5035	−27.7085	19.6621
lnsize	36	5.0406	0.2104	4.6640	5.5149
turnover	36	84.0669	36.5198	34.1426	162.0876
illiquidity	36	0.2606	0.1470	0.1097	0.8251

3.3. Time Trend Fitness of Variables

In Section 3.3, we analyze the time trend fitness of variables, and empirically demonstrate the potential threshold effects of leveraged trading on the stock price crash risk. Although the non-linear effects of the stock price crash risk may vary [30,31]; however, the quadratic effect of leveraged trading on the stock price crash risk has been confirmed by many studies [23,24,32]. Therefore, we compare the linear trend and quadratic trend fitness of variables based on the literatures mentioned above.

Figure 1a,b show the time trend of the market leverage ratio in terms of linear function fitness and quadratic function fitness. The effect of the quadratic function fitness is better than that of linear fitness because the adjusted R^2 of the quadratic function fitness is higher than the adjusted R^2 of linear

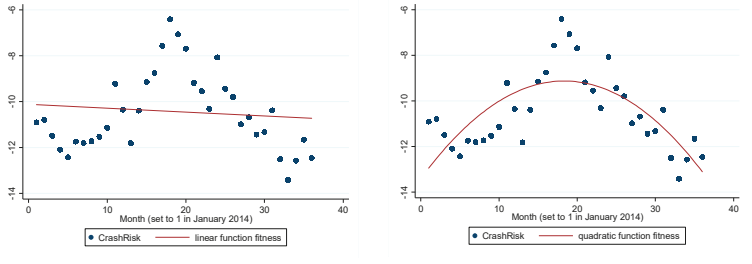
fitness. *tradingleverage* shows an inverted U-shape distribution over time, increasing during the bull market (July 2014–May 2015), but decreasing during a bear market (July 2014–December 2016). The demarcation point (around June 2014) is in line with the period when the China Securities Regulatory Commission implemented strong measures to deleverage.



(**a**) (**b**)

Figure 1. (**a**) Linear function fitness for *tradingleverage* (adj $R^2 = 0.21$) and (**b**) quadratic function fitness for *tradingleverage* (adj $R^2 = 0.69$).

Figure 2a,b show the time trend for the market price crash risk. The effect of quadratic function fitness is also better than that of linear fitness, because the adjusted R^2 of quadratic function fitness is higher than that of linear fitness. *CrashRisk* has an inverted U-shape distribution over time, similar to the distribution of the leverage ratio. *CrashRisk* is relatively low before the start of the bull market, but it increases during the bull market and decreases during the bear market. The demarcation point (around June 2014) is also in line with the date of the Chinese stock market price crash.



(**a**) (**b**)

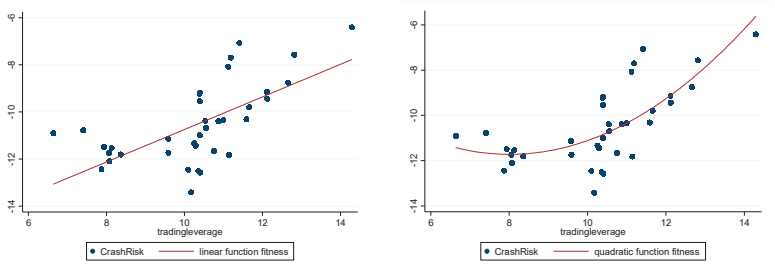
Figure 2. (**a**) Linear function fitness for *CrashRisk* (adj $R^2 = 0.0101$) and (**b**) quadratic function fitness for *CrashRisk* (adj $R^2 = 0.5452$).

3.4. Correlation Fitness of Tradingleverage and CrashRisk

Figure 3a,b show the fitness of *tradingleverage* and *CrashRisk*. Here, the effect of the quadratic function fitness is also better than that for linear fitness because the adjusted R^2 of quadratic function fitness is higher than that of linear fitness.

Figure 3b shows that the leverage ratio has a non-linear effect on the stock price crash risk, and the leverage ratio threshold point is about 10%. When *tradingleverage* is less than 10%, an increase in *tradingleverage* does not significantly increase *CrashRisk* and it might have reduced *CrashRisk*. When *tradingleverage* exceeds 10%, an increase in *tradingleverage* significantly increases *CrashRisk*. When comparing Figures 1b and 3b, the time point corresponding to the leverage ratio threshold is around October 2014. When *tradingleverage* reaches its maximum value (about 14%), *CrashRisk* also reaches its maximum (about −6), as shown in Figure 3b. Comparing Figures 1b and 2b shows that June 2015

has the highest *tradingleverage* and highest *CrashRisk*, which is the date of the Chinese stock market price crash.



(**a**) (**b**)

Figure 3. (**a**) Linear function fitness for *tradingleverage* and *CrashRisk* (adj $R^2 = 0.4063$) and (**b**) quadratic function fitness for *tradingleverage* and *CrashRisk* (adj $R^2 = 0.5708$).

Figure 4 shows the relationship between *tradingleverage*, *month*, and *CrashRisk* to further depict the relationship of the leverage ratio and the stock price crash risk. The lighter the color, the greater the *CrashRisk*. As per Figure 4, as leverage ratio increases, the color becomes lighter, and the stock price crash risk increases; over time, the color lightens and then turns dark, which means that the stock price crash risk initially increases but decreases after a period of time. Figure 4 confirms the correlation between the leverage ratio and the stock price crash risk.

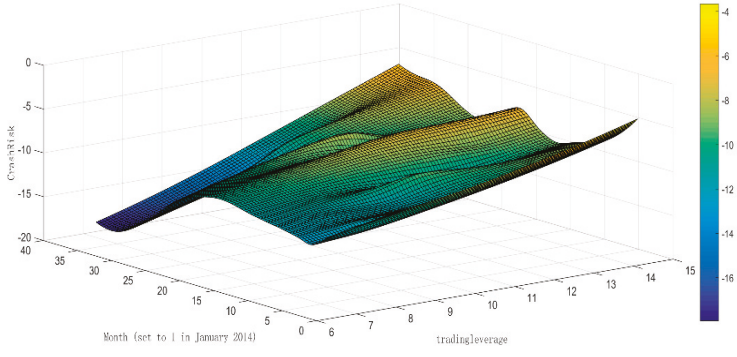


Figure 4. The relationship among *tradingleverage*, *month*, and *CrashRisk*. The lighter the color, the greater the *CrashRisk*.

The time dummy variables are set, as follows: *dummyt1* (take a value of 1 if the time is after September 2014; otherwise, take 0); *dummyt2* (take a value of 1 if the time is after May 2015; otherwise, take 0). A dummy variable of the leverage ratio is then set: *dummylev* (take a value of 1 when the leverage ratio is greater than 10%; otherwise, take 0). The second leverage ratio threshold (potential threshold) could only be artificially set to 14% for June 2015, which is the date of the Chinese stock market price crash, since the maximum leverage ratio in the sample is about 14%.

4. Results Presentation and Discussion

We examined the threshold effect of leveraged trading on the stock price crash risk according to Equation (4). Table 3 shows the results.

Table 3. Regression results of the leverage ratio to the stock price crash risk.

	Model 1	**Model 2**	**Model 3**	**Model 4**
Variable		$CrashRisk_t$		
$tradingleverage_{t-1}$	0.181 **	−0.112 ***	0.254 ***	−0.0897 ***
	(0.00858)	(0.00655)	(0.00706)	(0.00576)
$dummyt1 \times tradingleverage_{t-1}$		0.223 ***		
		(0.00345)		
$dummyt2 \times tradingleverage_{t-1}$			0.148 ***	
			(0.00166)	
$dummylev \times tradingleverage_{t-1}$				0.289 ***
				(0.00204)
$retn_{t-1}$	−0.0530 ***	−0.0552 ***	−0.0196	−0.0564 ***
	(0.0154)	(0.0152)	(0.0150)	(0.0151)
$turnover_{t-1}$	0.0213 ***	0.0200 ***	0.0148 **	0.0194 **
	(0.00704)	(0.00717)	(0.00678)	(0.00739)
$illiquidity_{t-1}$	0.196 ***	0.202 ***	0.259 ***	−0.513 ***
	(0.0365)	(0.0340)	(0.0205)	(0.0323)
$lnsize_{t-1}$	3.462	3.662	3.237	3.490
	(2.391)	(2.473)	(1.947)	(2.502)
constant	−30.33 ***	−29.44 ***	−29.37 ***	−28.32 **
	(9.528)	(10.09)	(8.157)	(10.60)
Time trend	control	control	control	control
N	35	35	35	35
Adj R^2	0.731	0.735	0.779	0.739
F statistic	49.55 ***	47.78 ***	26.54 ***	54.49 ***
Portmanteau (Q) statistic	14.16	15.88	23.91	12.67

Note: The standard error in brackets is a robust standard error. **, and *** denote significance at 10%, 5%, and 1% significance levels, respectively. Adj R^2 measures the goodness of fit of the model. F statistic is used to test significance of the model. The F statistics are significant at the 10% significance level in Model 1 to Model 4. Therefore, we can accept the hypothesis of significance of the model. The Portmanteau (Q) statistic is used to test time series autocorrelation. The Portmanteau (Q) statistics are not significant at the 10% significance level in Model 1 to Model 4. Therefore, we can accept the null hypothesis of no autocorrelation.

The endogeneity issue deserves attention in the regression model. It is necessary to control the variables related to *tradingleverage* that also affect *CrashRisk* in the regression model, so as to reduce the impact of the endogeneity issue as much as possible in order to ensure that the thresholds of leveraged trading on the stock price crash risk are unbiased in the empirical results. In the recent studies, the variables related to *tradingleverage* that also affect *CrashRisk* are the following four categories: (1) the stock index return (*retn*) [26,33]; (2) investor sentiment (*turnover*) [24,28]; (3) market liquidity (*illiquidity*) [29,32]; and, (4) the total market value (*lnsize*) [33]. In Table 3, we controlled the above variables, which can effectively reduce the interference of the endogeneity issue on the results.

Table 3 shows the regression result of the leverage ratio to the stock price crash risk. This result is subdivided into four models. Model 1 tested whether a linear relationship existed between the leverage ratio to the stock price crash risk; Model 2–Model 4 further examined whether a threshold effect of the leverage ratio to the stock price crash risk existed. Model 2 tested whether October 2014 was a time threshold and if exceeding this time threshold increased the stock price crash risk. Model 3 tested whether June 2015 was another time threshold and whether exceeding this time threshold caused the stock price to crash. Model 4 tested whether 10% of the leverage ratio was the leverage ratio threshold and whether exceeding this leverage ratio threshold increased the stock price crash risk.

Model 1 shows that *tradingleverage* is significantly positive, and an increase in *tradingleverage* significantly increases *CrashRisk*; *turnover* is also significantly positive, which means that high investor

sentiment promotes the stock price crash risk. *illiquidity* is significantly positive, which indicates that a deterioration in market liquidity significantly increases the stock price crash risk, which is consistent with Wei et al. [34].

Model 2 shows that *dummyt1* × *tradingleverage* is significantly positive, and the magnitude of the effect is relatively large, thus further verifying that October 2014 is one of time thresholds. Before October 2014, the increase in the leverage ratio reduces the stock price crash risk and maintains stock market stability during this period. However, after October 2014, the leverage ratio asymmetrically increases the stock price crash risk. A high investor sentiment and deterioration of market liquidity will significantly increase the stock price crash risk.

Model 3 shows that *dummyt2* × *tradingleverage* is significantly positive, and the magnitude of the effect is relatively large, thus further verifying that June 2015 is another time threshold. After June 2015, the increase in the leverage ratio boosts price volatility and eventually asymmetrically increases the stock price crash risk.

Model 4 shows that *dummylev* × *tradingleverage* is significantly positive, and the magnitude of the effect is relatively large, thus indicating that the threshold effect of leveraged trading on the stock price crash risk occurs at around the 10% of leverage ratio. 10% of the leverage ratio is a significant threshold. When the leverage ratio is less than 10%, an increase in the leverage ratio reduces the stock price crash risk and maintains the stock market stability. However, the leverage ratio asymmetrically increases the stock price crash risk when the leverage ratio exceeds 10%. The leverage ratio threshold reveals that leveraged trading has an asymmetric effect on stock market stability. High investor sentiment will significantly increase the stock price crash risk, but the deterioration of market liquidity will significantly decrease the stock price crash risk. Of the leverage ratio, 10%, is also the turning point of feedback trading pattern of investors shift from the positive feedback trading behavior of "chasing up and down" to the negative feedback trading behavior of "selling high and buying low" [24]. The strong shift of feedback trading from positive to negative will indeed cause stock price crashes when the leverage ratio exceeds a certain threshold [25].

We discuss the conclusions with the previous studies. First, the empirical results verify the threshold effect of leveraged trading on the stock price crash risk. We verified that October 2014 and June 2015 are two significant time thresholds, and 10% of the leverage ratio is a significant leverage ratio threshold, which is consistent with the models [7,23,25] and the empirical results [24,34]. The important roles of investor behavior and market liquidity on the stock price crash risk have been confirmed. Second, the conclusions are different from the rational financial theory. The rational financial theory asserts that leveraged trading is a rational behavior of arbitrage, and leveraged trading maintains the stock market as being sustainable [6–8]. However, we find that leveraged trading plays a positive role in mitigating arbitrage restrictions and maintaining market stability in the stage with a low leverage ratio, but it has a negative role in exacerbating sentiment feedback trading, and increasing stock price crashes under a high leverage ratio. The effect of leveraged trading on the stock price crash risk is two-edged.

Thus far, the hypothesis is verified.

5. Robustness Test

The discussions in this work have limitations. Before examination of the threshold effect, the possible threshold values are pre-judged based on the fitness of the variables' correlation, which is somewhat subjective. We re-collected the market level data and used the threshold regression model for an in-depth analysis to test the robustness of the results [35].

5.1. Methodology

Threshold regression rolls the sample to test all potential thresholds and, therefore, requires a sufficient sample size. We re-collected the daily data of the stock index from 1 January 2014 to 31 December 2016. The threshold effect regression model was constructed, as shown in Equation (5).

$$\begin{aligned} CrashRisk_t = \alpha + \beta_1 \times tradingleverage_{t-1} \times \mathbb{I}(q_i \le \gamma) \\ +\beta_2 \times tradingleverage_{t-1} \times \mathbb{I}(q_i \ge \gamma) \\ +\varphi \times ControlVariables_{t-1} + \sum timevariables + \varepsilon_t \end{aligned} \quad (5)$$

where q_i represents a series of potential thresholds (including time thresholds and leverage ratio thresholds). $\mathbb{I}q_i \le \gamma$ is an illustrative function, which takes 1 if the expression in the parentheses is true; otherwise, it takes 0. The time horizon is daily.

5.2. Threshold Regression Results and Discussion

We used a threshold regression model to re-examine the thresholds (including the time thresholds and leverage ratio thresholds). Table 4 shows the regression results of the time thresholds and Table 5 provides the results of the leverage ratio thresholds.

Table 4. The regression results of the time thresholds.

Variables	Coefficient	Robust Standard Error	Z Statistic	*p* Value	95% Confidence Interval	
Explained Variable: *CrashRisk*						
$retn_{t-1}$	0.023 ***	0.006	4.1	0.000	0.012	0.035
$turnover_{t-1}$	0.002 ***	0.0003	6.380	0.000	0.001	0.002
$illiquidity_{t-1}$	−0.059	0.084	−0.710	0.479	−0.22	0.105
$lnsize_{t-1}$	0.007 **	0.003	2.040	0.041	0.000	0.013
Regime 1			Before 20 October 2014			
$tradingleverage_{t-1}$	0.006 ***	0.0008	7.010	0.000	0.004	0.008
constant	−0.066 *	0.037	−1.790	0.074	−0.138	0.006
Regime 2			Between 20 October 2014 and 28 May 2015			
$tradingleverage_{t-1}$	0.003 ***	0.001	4.970	0.000	0.018	0.004
constant	−0.040	0.040	−1.010	0.311	−0.118	0.038
Regime 3			After 28 May 2015			
$tradingleverage_{t-1}$	0.018 ***	0.001	26.150	0.000	0.017	0.020
constant	−0.089 **	0.039	−2.270	0.023	−0.166	−0.012
time thresholds	(1)	20 October 2014		SSE	0.009	
	(2)	28 May 2015		SSE	0.006	

Note: The standard error in brackets is a robust standard error. *, **, and *** denote significance at 10%, 5%, and 1% significance levels, respectively. SSE represents sum squared residual.

Table 4 verifies that 20 October 2014 and 28 May 2015 are the two significant time thresholds, which is consistent with the results presented in Section 4. Table 5 also verifies that 10.2% and 13.6% are the two significant leverage ratio thresholds. The magnitudes of the threshold effect are relatively large. When compared with Figure 1a, 10.2% of the leverage ratio corresponds to around 4 September 2014, and 13.6% of the leverage ratio corresponds to around 3 July 2015. These time points are slightly later than the results that are presented in Section 4.

Table 5. The regression results of the leverage ratio thresholds.

Variables	Coefficient	Robust Standard Error	Z Statistic	*p* Value	95% Confidence Interval	
Explained Variable: *CrashRisk*						
$retn_{t-1}$	0.003	0.010	0.310	0.76	−0.016	0.022
$turnover_{t-1}$	0.002 ***	0.0004	4.820	0.00	0.001	0.002
$illiquidity_{t-1}$	0.149	0.123	1.210	0.23	−0.093	0.391
$lnsize_{t-1}$	0.041 ***	0.002	17.270	0.00	0.036	0.045
Regime 1	Leverage ratio is below 10.2%					
$tradingleverage_{t-1}$	−0.02 ***	0.002	−12.98	0.00	−0.027	−0.020
constant	−0.43 ***	0.026	−16.65	0.00	−0.475	−0.375
Regime 2	Leverage ratio is between 10.2% and 13.6%					
$tradingleverage_{t-1}$	0.005 ***	0.0007	6.110	0.00	0.003	0.006
constant	−0.48 ***	0.028	−17.06	0.00	−0.539	−0.428
Regime 3	Leverage ratio is higher than 13.6%					
$tradingleverage_{t-1}$	0.010 ***	0.002	4.46	0.00	0.006	0.015
constant	−0.51 ***	0.028	−18.47	0.00	−0.567	−0.458
leverage ratio	(1)	10.2%		SSE	0.0124	
thresholds	(2)	13.6%		SSE	0.0112	

Note: The standard error in brackets is a robust standard error. *, **, and *** denote significance at 10%, 5%, and 1% significance levels, respectively. SSE represents sum squared residual.

6. Conclusions

The stock price crash constitutes one part of the complexity in the stock market, and we studied the effect of leveraged trading on the stock price crash risk from the perspective of feedback trading. The findings that are presented here confirm that leveraged trading has the threshold effects (both in the time dimension and the leverage ratio dimension) on the stock price crash risk. We found that leveraged trading not only plays a positive role in mitigating arbitrage restrictions and maintaining market stability in the stage with a low leverage ratio, but it also has a negative role in exacerbating sentiment feedback trading and increasing stock price crashes under a high leverage ratio. Generally speaking, these findings provide new insights in understanding the complexity in the Chinese stock market.

However, this paper has several limitations that require future study. First, we only consider "the whole period of abnormal volatility", and this leads to a risk of data snooping. Therefore, high frequency data and more recent data should be applied for further robustness test and sensitivity analysis of the results. Second, conducting country comparisons and analyzing the differences between the stock markets in different countries will make sense. For example, developed financial markets and emerging markets significantly differ in terms of investor structure and transaction institution. By country comparisons, we can analyze the impact of the factors mentioned above on leveraged trading and the stock price crash risk. Third, the method of computational finance can be applied in order to analyze the micro mechanism of leveraged trading affecting investors' behavior. For example, we can use the method of computational finance to build an artificial stock market where investors use leveraged trading to conduct behavioral gaming, which helps to search for more complex dynamic patterns in the relationship between leveraged trading and the stock price crash risk. Therefore, more investigations should be conducted for enhancing the reliability and applicability of the research results.

Author Contributions: Formal analysis, Z.P.; Methodology, Z.P.; Supervision, C.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Postdoctoral Science Foundation of China (2019M652605) and the National Natural Science Foundation of China (71671134).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kwapień, J.; Drożdż, S. Physical Approach to Complex Systems. *Phys. Rep.* **2012**, *515*, 115–226. [CrossRef]
2. Fama, E.F. Efficient Capital Market: A Review of Theory and Empirical Work. *J. Financ.* **1970**, *25*, 382–417. [CrossRef]
3. Shiller, R.J. The Use of Volatility Measures in Assessing Market Efficiency. *J. Financ.* **1981**, *36*, 291–304.
4. Hu, C.S.; Peng, Z.; Chi, Y.C. Feedback Trading, Trading Inducement and Asset Price Behavior. *Econ. Res. J.* **2017**, *5*, 189–202.
5. Grossman, S.J.; Miller, M.H. Liquidity and Market Structure. *J. Financ.* **1988**, *43*, 617–633. [CrossRef]
6. Hsieh, D.A.; Miller, M.H. Margin Regulation and Stock Market Volatility. *J. Financ.* **1990**, *45*, 3–29. [CrossRef]
7. Hong, H.; Stein, J.C. Differences of Opinion, Short-Sales Constraints, and Market Crashes. *Rev. Financ. Stud.* **2003**, *16*, 487–525. [CrossRef]
8. Chowdhry, B.; Nanda, V. Leverage and Market Stability: The Role of Margin Rules and Price Limits. *J. Bus.* **1998**, *71*, 179–210. [CrossRef]
9. Xiao, H.; Kong, A.G. A Study on the Mechanism of the Securities Margin Trading on the Fluctuations of the Special Nature of the Stock Price: A test Based on the Difference-in-Difference Model. *Manag. World* **2014**, *8*, 30–43.
10. Li, Z.S.; Chen, C.; Li, B.X. Does Short Selling Improve Price Efficiency in the Chinese stock market? Evidence from Natural Experiments. *Econ. Res. J.* **2015**, *5*, 165–177.
11. Chu, J.; Fang, J.X. Margin-trading, Short-selling and the Deterioration of Crash Risk. *Econ. Res. J.* **2016**, *5*, 143–158.
12. Wang, Z.Y.; Wang, Z.X. Price Limit, Margin Trading and Stock Price Volatility: A Comparative Study between A-share and H-share. *Econ. Res. J.* **2017**, *4*, 151–165.
13. Xu, C.S.; Ma, K. The Effect of Margin Trading on the Stock Overvaluation in the Bull Market: An Empirical Analysis based on the Data of 2014 Shanghai A-shares. *Econ. Rev.* **2017**, *1*, 40–52.
14. De Long, J.B.; Shleifer, A.; Summers, L.H.; Waldmann, R. Noise Trader Risk in Financial Markets. *J. Political Econ.* **1990**, *98*, 703–738. [CrossRef]
15. Tversky, A.; Kahneman, D. Availability: A Heuristic for Judging Frequency and Probability. *Cogn. Psychol.* **1974**, *4*, 207–232.
16. De Bondt, W.P.M. Betting on Trends: Intuitive Forecasts of Financial Risk and Return. *Int. J. Forecast.* **1993**, *9*, 355–371. [CrossRef]
17. Benartzi, S.; Thaler, R.H. Myopic Loss Aversion and the Equity Premium Puzzle. *Q. J. Econ.* **1995**, *110*, 73–92. [CrossRef]
18. Thaler, R.H. Mental Accounting Matters. *J. Behav. Decis. Mak.* **1999**, *12*, 183–206. [CrossRef]
19. Barberis, N.; Huang, M. Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *Am. Econ. Rev.* **2008**, *98*, 2066–2100. [CrossRef]
20. Barberis, N.; Xiong, W. Realization Utility. *J. Financ. Econ.* **2012**, *104*, 251–271. [CrossRef]
21. Lee, C.; Shleifer, A.; Thaler, R.H. Investor Sentiment and the Closed-End Fund Puzzle. *J. Financ.* **1991**, *46*, 75–109. [CrossRef]
22. Kahneman, D.; Riepe, M.W. Aspects of Investor Psychology. *J. Portf. Manag.* **1998**, *24*, 52–65. [CrossRef]
23. Guo, W.C.; Wang, F.Y.; Wu, H.M. Financial Leverage and Market Volatility with Diverse Beliefs. *Econ. Theory* **2011**, *47*, 337–364. [CrossRef]
24. Peng, Z.; Hu, C.S. Leveraged Trading, Irrational Sentiment and Sustainability in the Stock Market: Evidence from China. *Sustainability* **2020**, *12*, 1310. [CrossRef]
25. Hu, C.S.; Peng, Z. Margin Trading, Sentiment Feedback and the Volatility of Asset Prices. In Proceedings of the Conference Paper for the 15th China Finance Conference, Guangzhou, China, 10–11 November 2018.
26. Chen, J.; Hong, H.; Stein, J.C. Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices. *J. Financ. Econ.* **2001**, *61*, 345–381. [CrossRef]
27. Kim, J.B.; Li, Y.; Zhang, L. CFOs versus CEOs: Equity Incentives and Crashes. *J. Financ. Econ.* **2011**, *101*, 713–730. [CrossRef]
28. Baker, M.; Wurgler, J. Investor Sentiment and the Cross-Section of Stock Return. *J. Financ.* **2006**, *61*, 1645–1680. [CrossRef]

29. Amihud, Y. Illiquidity and Stock Return: Cross-section and Time-series Effects. *J. Financ. Mark.* **2002**, *5*, 31–56. [CrossRef]
30. Gençay, R.; Gradojevic, N. The Tale of Two Financial Crises: An Entropic Perspective. *Entropy* **2017**, *19*, 244. [CrossRef]
31. Wang, X.D.; Hui, X.F. Cross-Sectoral Information Transfer in the Chinese Stock Market around Its Crash in 2015. *Entropy* **2018**, *20*, 663. [CrossRef]
32. Brunnermeier, M.K.; Pedersen, L.H. Funding liquidity and market liquidity. *Rev. Financ. Stud.* **2009**, *22*, 2201–2238. [CrossRef]
33. Fama, E.F.; French, K.R. Common Risk Factors in the Return on Stocks and Bonds. *J. Financ. Econ.* **1993**, *33*, 3–56. [CrossRef]
34. Wei, L.J.; Zhang, W.; Xiong, X. The Mechanism and Solution for the Liquidity Stampede Crisis in Stock Markets. *J. Manag. Sci. China* **2017**, *20*, 1–23.
35. Hansen, B.E. Sample Splitting and Threshold Estimation. *Econometrica* **2000**, *68*, 575–603. [CrossRef]

entropy

MDPI

Article

Looking at Extremes without Going to Extremes: A New Self-Exciting Probability Model for Extreme Losses in Financial Markets

Katarzyna Bień-Barkowska

Institute of Econometrics, Warsaw School of Economics, Madalińskiego 6/8, 02-513 Warsaw, Poland; katarzyna.bien@sgh.waw.pl

Received: 14 June 2020; Accepted: 16 July 2020; Published: 20 July 2020

Abstract: Forecasting market risk lies at the core of modern empirical finance. We propose a new self-exciting probability peaks-over-threshold (SEP-POT) model for forecasting the extreme loss probability and the value at risk. The model draws from the point-process approach to the POT methodology but is built under a discrete-time framework. Thus, time is treated as an integer value and the days of extreme loss could occur upon a sequence of indivisible time units. The SEP-POT model can capture the self-exciting nature of extreme event arrival, and hence, the strong clustering of large drops in financial prices. The triggering effect of recent events on the probability of extreme losses is specified using a discrete weighting function based on the at-zero-truncated Negative Binomial (NegBin) distribution. The serial correlation in the magnitudes of extreme losses is also taken into consideration using the generalized Pareto distribution enriched with the time-varying scale parameter. In this way, recent events affect the size of extreme losses more than distant events. The accuracy of SEP-POT value at risk (VaR) forecasts is backtested on seven stock indexes and three currency pairs and is compared with existing well-recognized methods. The results remain in favor of our model, showing that it constitutes a real alternative for forecasting extreme quantiles of financial returns.

Keywords: forecasting market risk; value at risk; extreme returns; peaks over threshold; self-exciting point process; discrete-time models; generalized Pareto distribution

1. Introduction

Forecasting extreme losses is at the forefront of quantitative management of market risk. More and more statistical methods are being released with the objective of adequately monitoring and predicting large downturns in financial markets, which is a safeguard against severe price swings and helps to manage regulatory capital requirements. We aim to contribute to this strand of research by proposing a new self-exciting probability peaks-over-threshold (SEP-POT) model with the prerogative of being adequately tailored to the dynamics of real-world extreme events in financial markets. Our model can capture the strong clustering phenomenon and the discreteness of times between the days of extreme events.

Market risk models that account for catastrophic movements in security prices are the focal point in the practice of risk management, which can clearly be demonstrated by repetitive downturns in financial markets. The truth of this statement cannot be more convincing nowadays as global equity markets have very recently reacted to the COVID-19 pandemic with a plunge in prices and extreme volatility. The coronavirus fear resulted in panic sell-outs of equities and the U.S. S&P 500 index plummeted 9.5% on 12 March 2020, experiencing its worst loss since the famous Black Monday crash in 1987. Directly 2, 4, 6, and 7 business days later, the S&P 500 index registered additional huge price drops amounting to, correspondingly, 12%, 5.2%, 4.3%, and 2.9%, respectively. At the same time,

the toll that the COVID-19 pandemic took on European markets was also unprecedented. For example, the German bluechip index DAX 30 plunged 12.2% on 12 March 2020, which was followed by a further 5.3%, 5.6%, 2.1% losses, correspondingly, 2, 4, and 7 business days later. The COVID-19 aftermath is a real example that highlights the strong clustering property of extreme losses.

One of the most well-recognized and widely used measures of exposure to market risk is the value at risk (VaR). VaR summarizes the quantile of the gains and losses distribution and can be intuitively understood as the worst expected loss over a given investment horizon at a given level of confidence [1]. VaR can be derived as a quantile of an unconditional distribution of financial returns, but it is much more advisable to model VaR as the conditional quantile, so that it can capture the strongly time-varying nature of volatility inherent to financial markets. The volatility clustering phenomenon provides the reason for using the generalized autoregressive conditional heteroskedasticity (GARCH) models to derive the conditional VaR measure [2]. However, over the last decade, the conventional VaR models have been subject to massive criticism, as they failed to predict huge repetitive losses that devastated financial institutions during the global crisis of 2007–2008. Therefore, special focus and emphasis is now placed on adequate modeling of extreme quantiles for the conditional distribution of financial returns rather than the distribution itself.

One of the relatively recent and intensively explored approaches to modeling extreme price movements is a dynamic version of the POT model which relies on the concept of the marked self-exciting point process. Unlike the GARCH-family models, POT-family models do not act on the entire conditional distribution of financial returns. Instead, their focus moves to the distribution tails where—in order to account for their heaviness—the probability mass is usually approximated with the generalized Pareto distribution. Early POT models described the occurrence of extreme returns as realizations of an independent and identically distributed (i.i.d.) variable, which led to VaR estimates in the form of unconditional quantiles. One of the first dynamic specifications of POT models that took into account the volatility clustering phenomenon and allowed economists to perceive VaR as a conditional quantile was a two-stage method developed in [3]. This method required estimating an appropriately specified GARCH-family model in the first stage and fitting the POT model to GARCH residuals. A new avenue for forecasting VaR was opened up when the point-process approach to POT models was released in [4]. This methodology was later extended in several publications [5–14]. The benefit of this model is that it does not require prefiltering returns using GARCH-family estimates while at the same time it can capture the clustering effects of extreme losses and maintain the merits of the extreme value theory. The point-process POT model approximates the time-varying conditional probability of an extreme loss over a given day with the help of a conditional intensity function that characterizes the arrival rate of such extreme events. The intensity function can either be formulated in the spirit of the self-exciting Hawkes process [4,5,10–12] (which is extensively used in geophysics and seismology), in the form of the observation-driven autoregressive conditional intensity (ACI) model [13], or using the autoregressive conditional duration (ACD) models [6–8] (the last two methodologies were very popular in the area of market microstructure and the modeling of financial ultra-high-frequency data [15–17]). In all cases, the timing of extreme losses depends on the timing of extreme losses observed in the past.

This study does not strictly rely on the above mentioned point process approach to POT models. The discrete-time framework of our SEP-POT model is motivated by observation of real-world financial data measured daily, which is the most common frequency used in POT models of risk. The empirical analysis put forward in this paper is based on the daily log returns of seven international stock indexes (i.e., CAC 40 (France), DAX 30 (Germany), FTSE 100 (United Kingdom), Hang Seng (Hong Kong), KOSPI (Korea), NIKKEI (Japan), and S&P 500 (U.S.)) as well as the daily log returns of three currency pairs (JPY/USD, USD/GBP, USD/NZD). The daily log returns for the equity market were calculated from the adjusted daily closing prices downloaded from the Refinitiv Datastream database. The foreign exchange (FX) rates were obtained from the Federal Reserve Economic Data repository and are measured in following units: Japanese Yen to one U.S. Dollar (JPY/USD), U.S. Dollars to one British

Pound (USD/GBP), U.S. Dollars to one New Zealand Dollar (USD/NZD). Extreme losses are defined as the daily negated log returns (log returns pre-multiplied by -1) whose magnitudes (in absolute terms) are larger than a sufficiently large threshold, u. Figure 1 shows that for u corresponding to the 0.95-quantile of the unconditional distribution of negated log returns, the daily measurement frequency, and the broad set of financial instruments, the relative frequency mass of the time interval between subsequent extreme losses is concentrated on small integer values. Indeed, about 45% of all such durations is distributed on distinct discrete values of 1–5 days, and the most frequent time span between subsequent extreme losses is one day (about 12–13% of cases).

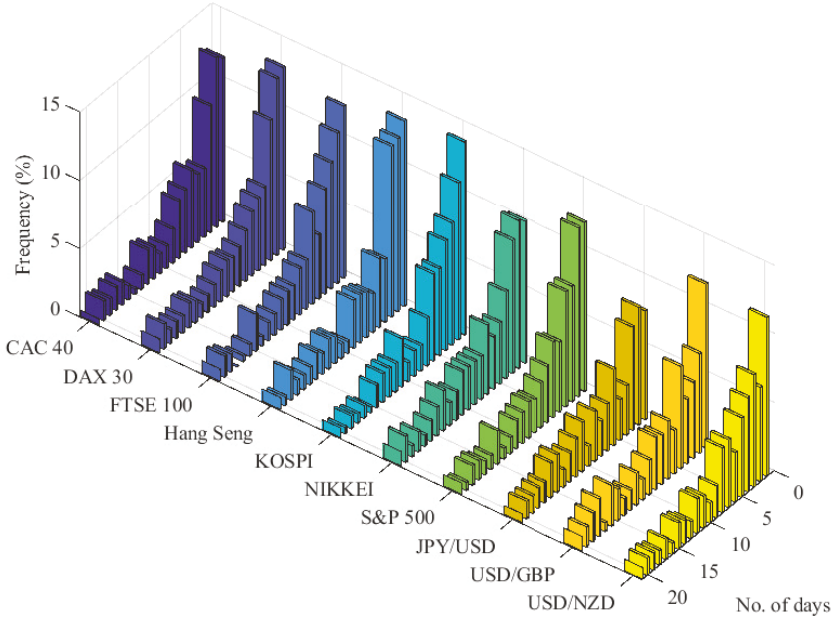


Figure 1. Frequency histogram for the time intervals (in number of days) between subsequent extreme losses for seven equity indexes and three FX rates between January 1981 and March 2020.

The SEP-POT model relates to the published work on the point-process approach to POT models but is consistent with the observed discreteness of threshold exceedance durations. Thus, in our model, the values of the time variable are treated as indivisible time units upon which extreme losses can be observed. As a result that the extreme losses are clustered, the model incorporates the self-exciting component. Accordingly, the extreme loss probability is affected by the series of time spans (in number of days) that have elapsed since all past extreme loss events. We apply the weighting function in the form of the at-zero-truncated Negative Binomial (NegBin) distribution that allows the influence of previous extreme losses to decay over time. The functional form of the extreme loss probability in our SEP-POT model is drawn from [18], where a very similar specification was proposed to depict the self-exciting nature of terrorist attacks in Indonesia and forecasted the probability of future terrorist attacks as a function of attacks observed in the past. Inspired by this work, we check the adequacy of such a discrete-time approach in the framework of POT models of risk. To this end, we perform an extensive validation of the SEP-POT model both in and out of sample and compare it with three widely-recognized VaR measures: one based on the self-exciting intensity

(Hawkes) POT model, one derived from the exponential GARCH model with skewed Student's t distribution (skewed-t-EGARCH) model, and the last one was delivered by the Gaussian GARCH model. The results for VaR at high confidence levels (>99%) remain in favor of the SEP-POT model, and hence, the model constitutes a real alternative for measuring the risk of large losses.

Section 2 outlines the point process approach to POT models, introduces the SEP-POT model, and outlines the backtesting methods used for model validation. Section 3 presents the empirical findings and discusses the extensive backtesting results. Finally, Section 4 concludes the paper and proposes areas for future research.

2. Methods

2.1. Self-Exciting Intensity POT Model

Consider $\{Y_t\}$ ($Y_t \in \mathbb{R}$) denoting the stochastic process that characterizes the evolution of negated daily log returns on a financial asset, being the daily log returns pre-multiplied by -1. The convention of using negated log returns legitimizes treating extreme losses as observations that fall into the right tail of distribution. More precisely, the extreme losses are defined as such positive realizations of Y_t that are larger than a sufficiently large threshold u. The magnitudes of extreme losses over a threshold u, (i.e., $\tilde{Y}_t = Y_t - u$) will be referred to as the threshold exceedances. The time intervals between subsequent threshold exceedances will be referred to as threshold exceedance durations.

Let $\{t_i, Y_{t_i}\}_{i\in\{1,2,\ldots,n\}}$ denote an observed sample path of (1) the times when extreme losses are observed (i.e., $0 < t_i < t_{i+1}$) and (2) the corresponding magnitudes of such losses (i.e., Y_{t_i}). If one pursued a continuous-time approach (i.e., assuming $t \in \mathbb{R}_+$), the realized sequence $\{t_i, Y_{t_i}\}_{i\in\{1,2,\ldots,n\}}$ of extreme returns with their locations in time can be treated as an observed trajectory of the marked point process. Treating these instances of threshold exceedance as realizations of a random variable allows us to model the occurrence rate of extreme losses Y_{t_i} at different time points $\{t_i\}$, for example, days. An excellent introduction to the theory and statistical properties of point processes can be found in [19].

The crucial concept in the point process theory is the conditional intensity function that characterizes the time structure of event locations, and hence, the evolution of the point process. The conditional intensity function is defined as follows:

$$\lambda(t|\mathcal{F}_t) \quad = \quad \lim_{\Delta\downarrow 0} \frac{\Pr\left[\left(N(t, t+\Delta]\right) > 0|\mathcal{F}_t\right]}{\Delta}, \tag{1}$$

where $N(t, s]$ denotes a number of events in $(t, s]$. Note that the conditional intensity function can intuitively be treated as the instantaneous conditional probability of an event (per unit of time) immediately after time t. To account for the clustering of extreme losses, $\lambda(t|\mathcal{F}_t)$ depends on $\mathcal{F}_t$ being an information set available at t, consisting of the complete history of event time locations and their marks, (i.e., $\mathcal{F}_t \equiv \sigma\{(t_i, Y_{t_i}), \forall i : t_i \leq t\}$). If $\lambda(t|\mathcal{F}_t)$ was constant over time (i.e., $\lambda(t|\mathcal{F}_t) = \lambda$) then for $t_i \in [0, \infty)$ the point process would correspond to a homogeneous Poisson point process with an arrival rate λ.

The notion of the conditional intensity facilitates the derivation of the conditional VaR measure. The VaR at a confidence level $1 - q$, (i.e., $q \in (0,\ 1)$ denotes a VaR coverage level), represents a qth quantile in the conditional distribution of financial returns. After taking advantage of working with the negated log returns and based on the notation introduced so far, the VaR (for a coverage level q) estimated for a day $t + 1$ can be derived from the following equation:

$$\Pr(Y_{t+1} > y_{q,t+1}|\mathcal{F}_t) = q. \tag{2}$$

Hence, the VaR for a coverage rate q is equal to $y_{q,t+1}$, because the probability that a (negated) return exceeds the threshold value $y_{q,t+1}$ over a day $t + 1$ is equal to q. This probability can be further

rewritten as a product of: (1) the probability of an extreme loss arrival (i.e., a threshold exceedance) over day $t+1$ (given $\mathcal{F}_t$), and (2) the conditional probability that the size of this extreme loss is larger than $y_{q,t+1}$ (given that an extreme loss was observed over day $t+1$):

$$\Pr(Y_{t+1} > u|\mathcal{F}_t) \Pr(Y_{t+1} > y_{q,t+1}|Y_{t+1} > u; \mathcal{F}_t) = q. \tag{3}$$

The early, classical POT model of the extreme value theory (EVT) (The EVT offers two major classes of models for extreme events in finance: (1) the block maxima method, which uses the largest observations from samples of i.i.d. data, and (2) the POT method, which is more efficient for practical application because it uses all large realizations of variables, provided that they exceed a sufficiently high threshold. A detailed exposition of these methods can be found in [20].) assumes that the financial return data is i.i.d. Hence, threshold exceedances are also i.i.d homogeneous Poisson distributed in time. Accordingly, the probability of observing a threshold exceedance over given day t is constant and can be estimated as a proportion of threshold exceedances in the sample (i.e., n/T, where n is the number of threshold exceedances and T denotes the length of time series of financial returns). By this logic, the standard POT model neglects repeated episodes of increased volatility and therefore also ignores the clustering property of extreme losses. As noted in [20], the standard POT model is not directly applicable to financial return data.

The more recent dynamic versions of the classical POT model found in several studies (i.e., [4–14]), are directly motivated by the behavior of the non-homogeneous Poisson point process, where the intensity rate of threshold exceedances, $\lambda(t|\mathcal{F}_t)$, can vary over time due to the temporal bursts in volatility. According to such a point process approach to POT models, the first factor on the left-hand side of Equation (3) (i.e., the conditional probability of a threshold exceedance over day $t+1$) can be derived based on the (time varying) conditional intensity function as follows:

$$\begin{aligned} \Pr(Y_{t+1} > u|\mathcal{F}_t) &= \Pr\left[N(t,t+1] > 0|\mathcal{F}_t\right] \\ &= 1 - \Pr\left[N(t,t+1] = 0|\mathcal{F}_t\right] \\ &= 1 - \exp\left(-\int_t^{t+1} \lambda(v|\mathcal{F}_v)\, \mathrm{d}v\right), \end{aligned} \tag{4}$$

because the probability of no events in $(t,s]$ (i.e., $N(t,s] = 0$) can be given as $\Pr(N(t,s] = 0|\mathcal{F}_t) = \exp\left(-\int_t^s \lambda(v|\mathcal{F}_v)\, \mathrm{d}v\right)$ [21].

The POT models use the Pickands–Balkema–de Haan's theorem of EVT, which allows us to approximate the second factor on the left-hand side of Equation (3) (i.e., the conditional probability that Y_{t+1} exceeds $y_{q,t+1}$, given that it surpassed the threshold u) using the generalized Pareto distribution, as follows:

$$\begin{aligned} \Pr(Y_{t+1} - u > y_{q,t+1} - u|Y_{t+1} > u; \mathcal{F}_t) &= 1 - \Pr(Y_{t+1} - u \leq y_{q,t+1} - u|Y_{t+1} > u; \mathcal{F}_t) \\ &\approx 1 - F_{GP}(y_{q,t+1} - u|Y_{t+1} > u; \mathcal{F}_t) \\ &= \left(1 + \xi \frac{y_{q,t+1} - u}{\sigma}\right)_+^{-1/\xi}, \end{aligned} \tag{5}$$

where $F_{GP}(\cdot)$ denotes the cumulative distribution function of the generalized Pareto (GP) distribution with the scale parameter $\sigma \in \mathbb{R}_{>0}$ and the shape parameter $\xi \in \mathbb{R}_{\neq 0}$. If $\xi \to 0$, $F_{GP}(\cdot)$ tends to the cumulative distribution function of an exponential distribution.

Equations (3)–(5) provide the grounds for the derivation of $VaR_{q,t+1}$, as follows:

$$VaR_{q,t+1} = \left[\left(\frac{q}{1 - \exp\left(-\int_t^{t+1} \lambda(v|\mathcal{F}_v)\mathrm{d}v\right)}\right)^{-\xi} - 1\right] \cdot \frac{\sigma}{\xi} + u. \tag{6}$$

The dynamic versions of the POT models benefit from both (1) the point process theory, which allows for the time-varying intensity rate of threshold exceedances, and hence, the clustering of extreme losses, and (2) the EVT, which allows us to account for the tail risk of financial instruments. Thus, the forecasts of daily VaR can be time-varying and react to the new information. (The early, classical POT models of EVT assume a constant intensity, λ, and a constant scale parameter of the GP distribution for threshold exceedances, σ. Accordingly, the VaR level is constant over time: $VaR_q = \left[\left(\frac{qT}{n}\right)^{-\xi} - 1\right] \cdot \frac{\sigma}{\xi} + u$.) In empirical applications, appropriate dynamic specifications of selected components in Equation (6) are needed. One possible way of specifying the time-varying conditional intensity function $\lambda(t|\mathcal{F}_t)$ is provided by the Hawkes process [19]. The Hawkes process belongs to the class of so called self-exciting processes where past events can accelerate the occurrence of future events. Accordingly, the conditional intensity function is defined as follows:

$$\begin{aligned} \lambda(t|\mathcal{F}_{t-}) &= \mu + \int_{-\infty}^{t} w(t-v)\, \mathrm{d}N(v) \\ &= \mu + \sum_{t_i<t} w(t-t_i), \end{aligned} \tag{7}$$

where $\mu \in \mathbb{R}_{>0}$ denotes a constant and $w(\cdot)$ refers to a non-negative weighting function that captures the impact of past events, (i.e., extreme-loss days). Accordingly, each threshold exceedance at $t_i < t$ contributes an amount $w(t-t_i)$ to the risk of an extreme loss at t. This is necessary to provide a convenient parametric functional form for $w(\cdot)$. The well-recognized weighting function that we apply in the empirical portion of this paper is an exponential kernel function, given as follows:

$$w(x) = \alpha \exp(-\beta\, x), \tag{8}$$

where $\alpha \in \mathbb{R}_{\geq 0}$, $\beta \in \mathbb{R}_{\geq 0}$ are the parameters to be estimated. Accordingly, $\lambda(t|\mathcal{F}_{t-})$ is based on the summation of exponential kernel functions evaluated at the time intervals that start at the times of previous extreme losses (i.e., t_i) and last up to time t. The parameters α and β capture, correspondingly, the scale (i.e., the amplitude) and the rate of decay characterizing an influence of past events on the current intensity. The point process features the self-excitation property because the conditional intensity function rises instantaneously after an extreme loss is registered, which, in the aftermath, triggers the arrival of next events. This mechanism results in the clustering effect, characterizing the location of extreme losses in time. The time-varying nature of the conditional intensity function results in the fluctuations of VaR (see Equation (6)). On top of the clustering feature, the self-exciting intensity POT (i.e., SEI-POT) model for VaR (c.f., Equation (6)) can be further extended to account for the serial correlation in the magnitudes of the threshold exceedances. This can be achieved by providing an appropriate dynamic model for the scale parameter of the GP distribution in Equation (5). In the empirical portion of this paper we use the following specification:

$$\sigma_t = \sigma(\bar{Y}_t|\mathcal{F}_{t-}) = \mu_s + \sum_{t_i<t} \alpha_s \bar{Y}_{t_i} \exp(-\beta_s(t-t_i)), \tag{9}$$

where $\mu_s \in \mathbb{R}_{>0}$, $\alpha_s \in \mathbb{R}_{\geq 0}$, $\beta_s \in \mathbb{R}_{\geq 0}$ denote the parameters to be estimated. Accordingly, the threshold exceedance magnitude is affected by the sizes and times of past threshold exceedances.

Unlike standard POT models, where the times of threshold exceedances are assumed to be i.i.d distributed according to the homogeneous Poisson process and the magnitudes of threshold exceedances are assumed to be i.i.d. GP distributed, the dynamic point-process-based variants of the POT models allow for a time-varying intensity rate of threshold exceedances and a time-varying expected magnitude of these threshold exceedances. Accordingly, the VaR is also time-varying. The interplay of fluctuations in $\lambda(t|\mathcal{F}_t)$ and in the scale parameter of the GP distribution for the

threshold exceedances, σ_t, elevates VaR in turbulent periods of financial turmoil and decreases its level during calm periods. Hence, the VaR adjusts to observed market conditions.

2.2. Self-Exciting Probability POT Model

In this section we introduce the self-exciting probability POT model that obeys the natural distinction between the processes defined in discrete and continuous time. The structure of our model still draws from Equation (3), but we treat time as if it was composed of indivisible distinct units (days). Therefore, we refrain from approximating the conditional extreme loss probability using the conditional intensity function that characterizes the evolution of point process in continuous time. As a result that we formulate our model in discrete time, we directly describe the conditional probability of an extreme loss over day t, as follows:

$$p_t \quad = \quad \Pr(Y_t > u|\mathcal{F}_{t-1}) = g(\tilde{\lambda}_t), \tag{10}$$

where $g(\cdot) \in (0,1)$ denotes a link function. One possible choice of specifying $g(\cdot)$ (cf., [18]) is:

$$p_t = 1 - \exp(-\tilde{\lambda}_t), \tag{11}$$

where $p_t \in (0,1)$ if $\tilde{\lambda}_t > 0$.

Based on [18], the conditional probability of an extreme loss arrival over day t can be described in a dynamic fashion that exposes the self-exciting nature of the SEP-POT model as follows:

$$\tilde{\lambda}_t = \mu + \alpha \sum_{t_i<t} g(t - t_i), \tag{12}$$

where $\mu \in \mathbb{R}_{>0}$ denotes a constant determining a baseline probability, $\alpha \in \mathbb{R}_{\geq 0}$ determines the scale (amplitude) of the impact that the time location of the ith past extreme-loss event exerts on p_t, and $g(\cdot) \geq 0$ denotes the weighting function (i.e., discrete kernel function) that makes the past extreme-loss events less impactful than the more recent events. We specify $g(\cdot)$ as the probability function of the at-zero-truncated negative binomial (NegBin) distribution.

A probability function of a NegBin distribution is:

$$f(x;\omega,\kappa) = \frac{\Gamma(\kappa+x)}{\Gamma(\kappa)\Gamma(x+1)} \left(\frac{\kappa}{\kappa+\omega}\right)^{\kappa} \left(\frac{\omega}{\omega+\kappa}\right)^{x}, \quad x = 0,1,2,\cdots, \tag{13}$$

where $\omega \in \mathbb{R}_{>0}$ and $\kappa \in \mathbb{R}_{>0}$ are the parameters of the distribution and $E(u) = \omega$ and $Var(u) = \omega + \omega^2/\kappa$. For $\kappa \to \infty$, the NegBin distribution converges to a Poisson distribution. For $\kappa = 1$, the geometric distribution is obtained.

The at-zero-truncated NegBin distribution was formerly used in high-frequency-finance for modeling the non-zero price changes of financial instruments [22,23]. The probability function of at-zero-truncated NegBin distribution is given as $g(x;\omega,\kappa) = f(x;\omega,\kappa)/(1 - f(0;\omega,\kappa))$ (for $x = 1,2,...$), where $f(0;\omega,\kappa) = (\kappa/(\kappa+\omega))^{\kappa}$:

$$g(x;\omega,\kappa) = \frac{\Gamma(\kappa+x)}{\Gamma(\kappa)\Gamma(x+1)} \left[\left(\frac{\kappa+\omega}{\kappa}\right)^{\kappa} - 1\right]^{-1} \left(\frac{\omega}{\omega+\kappa}\right)^{x}, \quad x = 1,2,\cdots, \tag{14}$$

Figure 2 illustrates the self-exciting property of the SEP-POT model. The plots shown in the upper row depict the at-zero-truncated NegBin kernel functions evaluated at the time distances to previously observed events (i.e., $g(t - t_i)\ \forall i :\ t_i < t$). The impact of past events on p_t diminishes with time and the shape of decay is determined by parameters ω and κ. The scale of this impact is determined by α. The resulting conditional probability function of an extreme loss arrival is therefore based on the summation of the weighted kernel functions based on all the backward recurrence times. The choice

of an at-zero-truncated NegBin distribution guarantees flexibility in feasible shapes of the weighting function to properly reflect the dynamic properties of the data.

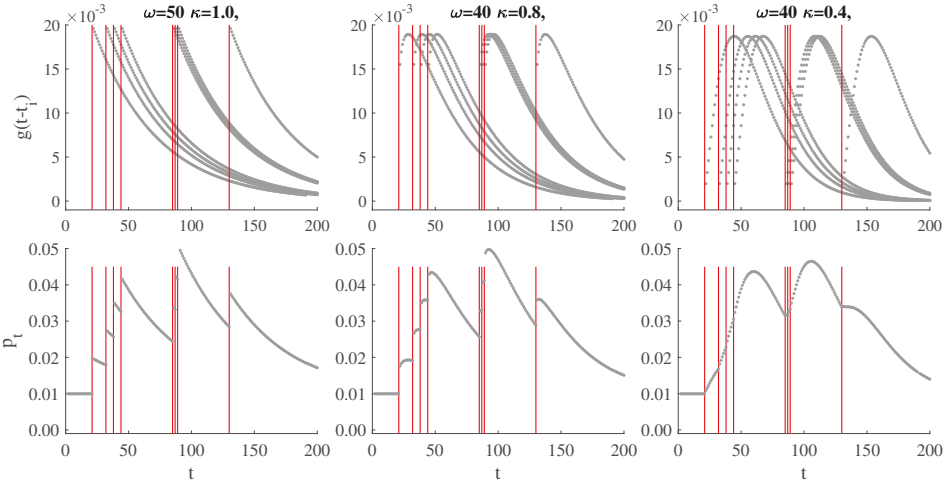


Figure 2. Illustration of the self-exciting probability model for eight events. Upper rows: Feasible shapes of the weighting functions $g(t-t_i), \ \forall i \ t_i < t$, at $\mu = 0.01$, $\alpha = 0.5$ (red lines indicate times of events). Lower row: The resulting probability p_t.

Like in existing dynamic extensions of the POT methodology, the threshold exceedance magnitudes in the SEP-POT model are described using the generalized Pareto distribution with the time-varying scale parameter. We specify this parameter as follows:

$$\sigma_t = \sigma(\bar{Y}_t|\mathcal{F}_{t-1}) = \mu_s + \alpha_s \sum_{t_i<t} \bar{Y}_{t_i} g_s(t-t_i;\omega_s), \tag{15}$$

where $\mu_s \in \mathbb{R}_{>0}$ is a constant, $\alpha_s \in \mathbb{R}_{\geq 0}$ is a scale parameter, and $g_s(x;\omega_s)$ (for $x = 1, 2, ...,$) denotes the nonnegative discrete weighting (kernel) function. For this purpose, we use a PDF of a geometric distribution with parameter $\omega_s \in \mathbb{R}_{>0}$, because it constitutes a natural discrete counterpart to an exponential distribution used in the continuous-time framework of the SEI-POT model (see Equation (9)). Hence, the magnitude of the threshold exceedance awaited at t is affected by the times and sizes of all previously observed threshold exceedances. The monotonically decaying weighting function allows distant events to affect the magnitudes of losses less than the recent events do.

The SEP-POT model assumes that the density function $f^u_{Y_t}(y_t|\mathcal{F}_{t-1})$, depicting the right tail of the distribution of the negated financial returns, has the following form:

$$\begin{aligned} f^u_{Y_t}(y_t|\mathcal{F}_{t-1}) \quad &= \quad p_t^{\mathbf{1}_{\{t=t_i\}}} \cdot (1-p_t)^{(1-\mathbf{1}_{\{t=t_i\}})} \\ &\cdot \quad \left(\frac{1}{\sigma_t} \left(1 + \xi \frac{y_t - u}{\sigma_t} \right)_+^{-1/\xi - 1} \right)^{\mathbf{1}_{\{t=t_i\}}}, \end{aligned} \tag{16}$$

which means that Y_t either surpasses the threshold u, i.e., belongs to the right tail of distribution ($\mathbf{1}_{\{t=t_i\}} = 1$), and hence, is drawn from the generalized Pareto distribution with probability p_t, or does not belong to the distribution tail ($\mathbf{1}_{\{t=t_i\}} = 0$) with probability $1 - p_t$.

This reasoning allows us to formulate the log-likelihood function of the SEP-POT model as the sum of two log-likelihoods as follows:

$$\ln \mathcal{L} = \ln \mathcal{L}_1 + \ln \mathcal{L}_2, \tag{17}$$

where:

$$\begin{aligned} \ln \mathcal{L}_1 &= \sum_{t=1}^{T} \left[\mathbf{1}_{\{t=t_i\}} \ln(p_t) + (1 - \mathbf{1}_{\{t=t_i\}}) \ln(1-p_t) \right] \\ &= \sum_{t=1}^{T} \left[\mathbf{1}_{\{t=t_i\}} \ln(\exp(\tilde{\lambda}_t) - 1) - \tilde{\lambda}_t \right], \end{aligned} \tag{18}$$

and

$$\ln \mathcal{L}_2 = -(1/\xi + 1) \sum_{i=1}^{n} \ln \left(1 + \xi \frac{y_{t_i} - u}{\sigma_t} \right) - \sum_{i=1}^{n} \ln(\sigma_t). \tag{19}$$

The VaR for a coverage rate q forecasted for day t (based on the information up to and including day $t-1$) can be derived from the SEP-POT model as follows:

$$\begin{aligned} q &= \Pr(Y_t > u | \mathcal{F}_{t-1}) \Pr(Y_t > y_{q,t} \big| Y_t > u; \mathcal{F}_{t-1}) \\ &= p_t \left(1 + \xi \frac{y_{q,t} - u}{\sigma_t} \right)^{-1/\xi}. \end{aligned} \tag{20}$$

Hence:

$$VaR_{q,t} = \left[\left(\frac{q}{1 - \exp(-\tilde{\lambda}_t)} \right)^{-\xi} - 1 \right] \cdot \frac{\sigma_t}{\xi} + u. \tag{21}$$

The SEP-POT model provides the grounds not only to derive the VaR, but also the expected shortfall (ES). Unlike the VaR, the ES is a coherent risk measure. It represents the conditional expectation of loss given that the loss lies beyond the VaR [24]. Accordingly, the ES corresponding to a coverage rate q, forecasted for a day t based on the information set up to and including day $t-1$ is defined as following:

$$ES_{q,t} = E(Y_t \big| Y_t > VaR_{q,t}; \mathcal{F}_{t-1}). \tag{22}$$

Equation (22) can be also rewritten as follows:

$$ES_{q,t} = VaR_{q,t} + E(Y_t - VaR_{q,t} \big| Y_t > VaR_{q,t}; \mathcal{F}_{t-1}). \tag{23}$$

The ES can be derived based on the standard definition of the mean excess function for the GP distribution. For $u' > u$, the mean excess function $e(u')$ corresponding to the GP distribution (where $\sigma > 0$, $0 < \xi < 1$) is defined as:

$$e(u') = E(Y_t - u' \big| Y_t > u') = \frac{\sigma + \xi(u' - u)}{1 - \xi}. \tag{24}$$

Hence, the expected size of losses exceeding the threshold u' is a linear function of $u' - u$. The ES forecasts from the SEP-POT model can be derived by applying the definition of $e(u')$ to Equation (23)

and by specifying the scale parameter of the GP distribution, σ, according to Equation (15). This leads to the formula for ES as follows:

$$\begin{aligned} ES_{q,t} &= VaR_{q,t} + \frac{\sigma_t + \xi(VaR_{q,t} - u)}{1-\xi} \\ &= \frac{VaR_{q,t} + \sigma_t - \xi u}{1-\xi}. \end{aligned} \tag{25}$$

2.3. Backtesting Methods

We use four backtesting procedures to assess the accuracy of the VaR delivered by the SEP-POT model. Each of these methods refer to the notion of a VaR exceedance or a VaR violation, being a binary indicator function, I_t, defined as follows:

$$I_t = \begin{cases} 1, & \text{for} \quad Y_t > VaR_{q,t} \\ 0, & \text{for} \quad Y_t \leq VaR_{q,t}. \end{cases}$$

The backtesting is based on the comparison of forecasted daily VaR numbers with observed daily returns over a given period. A VaR exceedance occurs when an actual loss is larger than the VaR predicted for that day. If the SEP-POT model were a true data generating process, than, $\forall t \ \Pr(I_t = 1|\mathcal{F}_{t-1}) = q$, which implies that the VaR violations would be i.i.d.

2.3.1. Unconditional Coverage Test

The first test that we consider is a widely used unconditional coverage (UC) test [25] where the null hypothesis states that the proportion of VaR exceedances according to a risk model (i.e., π) matches with the assumed coverage level for VaR (i.e., q): $H_0: \ \pi = q$. The UC test is formulated as a likelihood ratio test which compares two Bernoulli likelihood functions. Asymptotically, as the number of observations T goes to infinity, the test statistic is distributed as χ^2 with one degree of freedom:

$$LR_{UC} = -2\ln\left\{q^{T_1}(1-q)^{1-T_1} / \left[(T_1/T)^{T_1}(1-T_1/T)^{1-T_1}\right]\right\} \ \sim \ \chi^2_1, \tag{26}$$

where T_1 denotes the number of VaR violations in the sample of T returns.

2.3.2. Conditional Coverage Test

The second test is the conditional coverage (CC) that not only verifies the correct coverage but also sheds light on the independence of VaR violations over time [26]. This test was established in such a way that it aims to reject the VaR models when a risk model produces either the incorrect proportions or the clusters of exceedances. To this end, the process of VaR violations is described by a first-order Markov model and the CC test is based on the estimated transition matrix, as follows:

$$\begin{bmatrix} \hat{\pi}_{00} & \hat{\pi}_{01} \\ \hat{\pi}_{10} & \hat{\pi}_{11} \end{bmatrix} = \begin{bmatrix} T_{00}/(T_{00}+T_{01}) & T_{01}/(T_{00}+T_{01}) \\ T_{10}/(T_{10}+T_{11}) & T_{11}/(T_{10}+T_{11}), \end{bmatrix} \tag{27}$$

where π_{00} and π_{01} denote, correspondingly, the conditional probability of no VaR violation and a VaR violation (today), given that yesterday there was no VaR violation. Analogously, π_{11} and π_{10} denote, correspondingly, the conditional probability of a VaR violation and no VaR violation (today) directly after a VaR violation yesterday. As given in Equation (27), the elements of the transition matrix are estimated with the actual proportions of VaR violations, where T_{ij}, for $i \in \{0,1\}, j \in \{0,1\}$ is the number of (negated) returns with the indicator function I_t equal to j directly following an indicator's value i. The CC null hypothesis states that the conditional probability of a VaR violation directly after another VaR violation is the same as the conditional probability of a VaR violation after no violation and, at the same time, it is equal to the assumed coverage level for VaR (i.e., $H_0: \ \pi_{01} = \pi_{11} = q$).

Asymptotically, as the number of observations T goes to infinity, the test statistic LR_{CC} is distributed as a χ^2 with two degrees of freedom:

$$LR_{CC} = -2\ln\left\{q^{T_1}(1-q)^{1-T_1} / \left[(1-\hat{\pi}_{01})^{T_{00}}\ \hat{\pi}_{01}^{T_{01}}\ (1-\hat{\pi}_{11})^{T_{10}}\ \hat{\pi}_{11}^{T_{11}}\right]\right\}\ \sim\ \chi^2_2 \tag{28}$$

However, because the CC test is established on the Markov property of the violation process, it is sensitive to the dependence of order one only. Therefore, the CC test cannot be used to verify whether the current VaR exceedance depends on the sequence of states that preceded the last one.

2.3.3. Dynamic Quantile Conditional Coverage Test

The next two backtesting methods shed more light on the higher-order autocorrelation in the process of VaR violations. They also allow us to conclude whether the violations are affected by some previously observed explanatory variables. The first test is a dynamic quantile (DQ) test [27] that is based on a hit function, as follows:

$$Hit_t = I_t - q. \tag{29}$$

The correctly specified VaR model should form the Hit_t sequence with a mean value insignificantly different from 0, because Hit_t equals $1-q$, each time Y_t is larger than the daily VaR and $-q$, otherwise. Moreover, there should be no correlation between the current and the lagged values of the Hit_t sequence or between the current values of the Hit_t sequence and the current VaR. If the risk model corresponds to the true data generating process, the conditional expectation of Hit_t should be 0 given any information known at $t-1$. The DQ test that we use in the empirical section of our paper can be derived as the Wald statistic from an auxiliary regression, as follows:

$$Hit_t = \phi_0 + \sum_{j=1}^{4}\phi_j Hit_{t-j} + \phi_5 VaR_{q,t} + \varepsilon_t. \tag{30}$$

The null hypothesis states that the current value of a hit function (i.e., Hit_t) is not correlated with its four lags and the forecasted VaR (i.e., $VaR_{q,t}$ which is based on information known at $t-1$). Thus $H_0:\ \phi_j = 0\ \forall j \in \{0,...,5\}$. Hence, the null hypothesis states that the coverage probability produced by a risk model is correct (i.e., $\phi_0 = 0$) and none of the five explanatory variables affects Hit_t. Hence, the DQ test statistic is asymptotically χ^2 distributed with six degrees of freedom:

$$DQ = \frac{\mathbf{Hit}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Hit}}{q(1-q)}\ \sim\ \chi^2_6, \tag{31}$$

where **Hit** denotes a $T \times 1$ vector with observations of Hit_t variable and $\mathbf{X}$ denotes the standard $T \times 6$ matrix containing a column of ones and observations on the five explanatory variables at times $t = 1, ..., T$, according to the regression given in Equation (30).

2.3.4. Dynamic Logit Conditional Coverage Test

The dynamic logit test of conditional coverage might be treated as an extension of the DQ conditional coverage test [28]. This method considers the dichotomous nature of VaR violations. Accordingly, instead of the linear regression given by Equation (30), this test is established based on the dynamic logit model for I_t: $E[I_t|\mathcal{F}_{t-1}] = \Pr(I_t|\mathcal{F}_{t-1}) = F(a_t)$, where $F(\cdot)$ denotes the cumulative distribution function of a logistic distribution and a_t is specified as follows:

$$a_t = \phi_0 + \phi_1 a_{t-1} + \phi_2 I_{t-1} + \phi_3 VaR_{q,t}, \tag{32}$$

The autoregressive structure of Equation (32) allows us to better capture the dependence of a VaR violation probability upon possible explanatory factors. The null hypothesis is $H_0:\ \phi_0 = F^{-1}(q)$, $\phi_1 = \phi_2 = \phi_3 = 0$. Thus, the null states that the coverage probability delivered by a risk model

corresponds to the assumed coverage rate for VaR (i.e., $\phi_0 = F^{-1}(q)$) and none of regressors used in Equation (32) causes an incidence of VaR violation. The test statistic can be established as a likelihood ratio test statistic. Accordingly, it requires estimating the model given by Equation (32) and comparing its empirical log likelihood, $\ln \mathcal{L}_F$, with the restricted log likelihood under the null $\ln \mathcal{L}_R$. Under the null, the LR test statistic is χ^2 distributed with four degrees of freedom:

$$LR_{DL} = -2\{\ln \mathcal{L}_R - \ln \mathcal{L}_F\} \sim \chi^2_4. \tag{33}$$

3. Results and Discussion

In our empirical study we use daily log-returns from seven major stock indexes worldwide (CAC 40, DAX 30, FTSE 100, Hang Seng, KOSPI, NIKKEI, and S&P 500) and three currency pairs (JPY/USD, USD/GBP, USD/NZD). The CAC 40, DAX 30, and FTSE 100 are the major equity indexes in France, Germany, and U.K., respectively, and they are often perceived as the proxies or the real-time indicators for a much broader European stock market. The Hang Seng, KOSPI, and NIKKEI demonstrate the investment opportunity on the largest Asian equity markets in Hong Kong, South Korea, and Japan, respectively. S&P 500 constitutes a widely-investigated benchmark stock index reflecting the state of the overall U.S. economy. These seven indices monitor the state of the international equity market in its three global financial centers—western Europe, eastern Asia, and the U.S. As far as selection of the FX rates is concerned, according to [29], the JPY/USD and USD/GBP are the second and third most traded currency pair in the world, after EUR/USD (We did not investigate the EUR/USD currency pair due to a much smaller number of observations when comparing to the other time series; the euro was launched on 1 January 1999). The NZD/USD, often nicknamed as the Kiwi by FX traders, is a classical example of the commodity currency pair that co-fluctuates with the world prices of primary commodities (i.e., New Zealand exports oil, metals, dairy, and meat products). The New Zealand Dollar is also treated by international investors as a carry trade currency—therefore, it is very sensitive to interest rate risk. For each of these financial instruments we split the data spanning over a four-decades-long period into: (1) the in-sample data (i.e., 2 January 1981–31 December 2014) dedicated to the estimation and evaluation of our models and (2) the out-of-sample data (i.e., 2 January 2015–31 March 2020) which is reserved for VaR backtesting purposes. For each of the time series, the initial threshold u was set as the 95%-quantile of the in-sample unconditional distribution of negated log returns. Hence, the 5% largest negated returns were defined as extreme losses, which means that, on average, an extreme loss can be observed with probability 0.05. The selection of the threshold value u was a compromise between (1) the desired number of observations in the tail of the distribution to reduce noise and to ensure stability in parameter estimates (i.e., the lower the u, the more observations used for estimation) and (2) the goodness-of-approximation of the threshold exceedance distribution with the GP distribution (i.e., the higher the u, the better the approximation with the GP distribution). The latter issue was solved using two diagnostic tools, that confirmed the adequate goodness-of-fit of the conditional GP distribution. We used the D-test proposed in Ref. [30] and the χ^2 test for uniformity of probability integral transforms (PIT) based on the GP density estimates. Figure 3 illustrates extreme losses corresponding to the German DAX 30 index between January 1981 and March 2020. The examination of panels [a] and [b] allows us to conclude that the periodic volatility bursts are paralleled with the strong clustering effects for both (1) the magnitudes of extreme losses and (2) the days that they occur. Indeed, the quantile-quantile (QQ) plot (panel [c]) comparing empirical quantiles of the time intervals between subsequent extreme-loss days against the quantiles of an exponential distribution proves that the times of extreme losses are not distributed according to the homogeneous Poisson point process. Clustering of extreme events is also demonstrated by the shape of the autocorrelation function (ACF), indicating significant positive autocorrelations in both time intervals between successive threshold exceedances and the observed magnitudes of such exceedances.

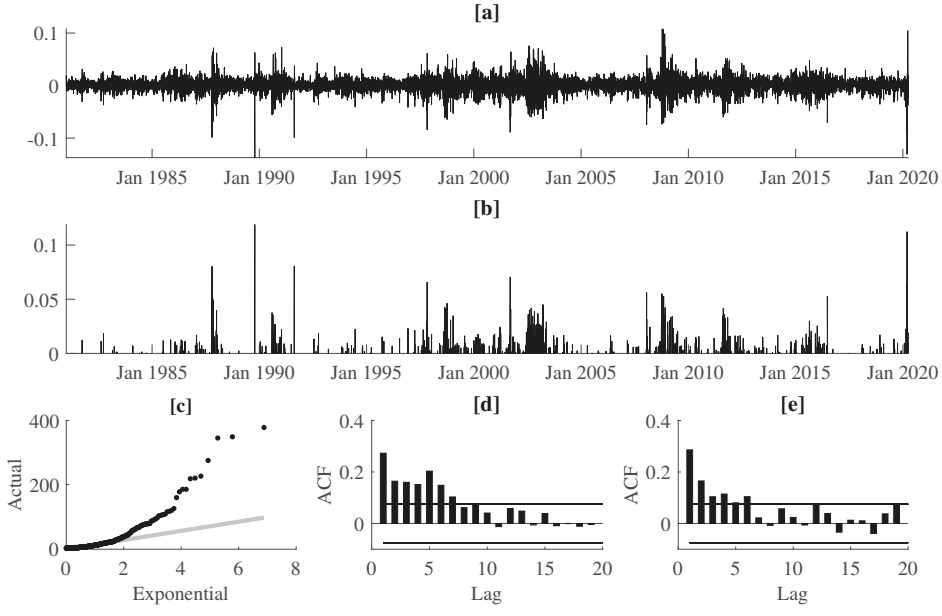


Figure 3. Panel (**a**) presents the daily log returns for the DAX index between Jan. 1981 and March 2020, panel (**b**) shows the corresponding ground-up threshold exceedances (i.e., the magnitudes of losses over the threshold u), panel (**c**) illustrates the quantile-quantile plot of inter-exceedance durations (in number of days) against the exponential distributions, and panels (**d**,**e**) present the autocorrelation functions for the inter-exceedance durations and the threshold exceedances, respectively.

The descriptive statistics of the CAC 40, DAX 30, and FTSE 100 data are summarized in Table 1 (analogical results for the remaining time series can be obtained from the author upon request). We see that for the CAC 40, DAX 30, and FTSE 100, the threshold exceedances were obtained as the losses surpassing u that is equal to 0.021, 0.021, and 0.017, respectively. Out of 8574 (CAC 40), 8563 (DAX 30), and 7826 (FTSE 100) daily log returns in-sample, these threshold values allow us to expose, correspondingly, 429, 428, and 391 extreme losses that were used for the model estimation purposes. For the FTSE 100 index, we have less observations (corresponding to three years: 1981–1983), because the in-sample period starts on 3 January 1984, when the FTSE 100 index was established. Although the official base date for the DAX 30 index is 31 December 1987, the DAX 30 index was linked with the former DAX index which dates back to 1959. The official base date for the CAC 40 also begins on 31 December 1987, but between 2 January 1981 and 30 December 1987 it could be measured as the "Insee de la Bourse de Paris." The threshold-exceedance durations cover a very wide range of observed values. For example, for the FTSE 100 index, the range spans from one day (with the relative frequency equal to 12.8% in-sample and 11.3% out-of-sample) up to 304 days in-sample or 205 days out-of-sample. In-sample, the largest threshold exceedance, equal to 0.114, was observed on the Black Monday of 20 October 1987 and it corresponded to a 12.22% decrease of the index. Out-of-sample, the maximum threshold exceedance is equal to 0.099 (a 10.87% plunge in the index) and was observed on the Black Thursday of 12 March 2020, being a single day in a chain of stock market crashes induced by the COVID-19 pandemic.

Realized gains and losses are measured over distinct days, and hence, the time spans between extreme losses are comprised of discrete time units (i.e., days). The scale of this phenomenon can be seen by looking at the considerable proportion of threshold exceedance durations equal to one, two, or three (business) days. Moreover, about 45% of such durations is less than or equal to five days and

over 60% are less than or equal to ten days. Another striking observation from Table 1 is the clustering of extreme losses. Large losses tend to occur in waves, which is seen from the Ljung-Box test statistics $Q(k)$ (where $k \in \{5, 10, 15\}$) for the lack of up to kth-order serial correlation. These test statistics are significantly different from zero, and hence, the null hypothesis of no autocorrelation in threshold exceedance durations must be rejected. Indeed, due to the COVID-19 outbreak, between 24 Februry and 31 March 2020 (i.e., over 27 business days) the CAC 40, DAX 30, and FTSE 100 suffered from as many as 10 (CAC40 and DAX 30) or 11 (FTSE 100) extreme losses (with the shortest and the longest threshold exceedance durations equal to one and five business days only, respectively). Extreme loss days tended to occur very close to each other, but this phenomenon is paralleled by the significant autocorrelation in the magnitudes of observed threshold exceedances. Based on the Ljung-Box test results, the null hypothesis of no autocorrelation in the threshold exceedance sizes needs to be rejected. The observed threshold exceedance durations are by their very nature discrete and feature strong positive autocorrelation. Therefore, our SEP-POT model is suitably tailored to this data.

Table 1. Descriptive statistics for the threshold exceedance durations and the threshold exceedance magnitudes for the CAC 40, DAX 30, and FTSE 100 indexes. ($Q(k)$ denotes the Ljung-Box test statistics for the lack of autocorrelation up to k-th order; $Q(k)$ ***, $Q(k)$ **, and $Q(k)$ * denote the statistics significant at the 1%, 5%, and 10% levels).

	CAC 40		**DAX 30**		**FTSE 100**	
	In-Sample	**Out-of-Sample**	**In-Sample**	**Out-of-Sample**	**In-Sample**	**Out-of-Sample**
no. of daily returns	8574	1342	8563	1326	7826	1327
threshold value	0.021	0.021	0.021	0.021	0.017	0.017
no. of exceedances (n)	429	48	428	59	391	53
			threshold exceedance durations			
Min	1	1	1	1	1	1
Max	397	414	378	345	304	205
$\frac{\#1_{\{t_i - t_{i-1} = 1\}}}{n}$	0.121	0.104	0.135	0.119	0.128	0.113
$\frac{\#1_{\{t_i - t_{i-1} = 2\}}}{n}$	0.128	0.104	0.138	0.085	0.110	0.132
$\frac{\#1_{\{t_i - t_{i-1} = 3\}}}{n}$	0.086	0.145	0.096	0.153	0.087	0.132
$\frac{\#1_{\{t_i - t_{i-1} <= 5\}}}{n}$	0.441	0.438	0.486	0.424	0.455	0.453
$\frac{\#1_{\{t_i - t_{i-1} <= 10\}}}{n}$	0.629	0.625	0.645	0.610	0.652	0.566
Mean	19.965	27.917	19.988	22.441	19.992	25
SD	40.002	66.572	41.786	49.212	40.245	42.287
Q(5)	83.307 ***	10.304 *	101.787 ***	5.688	69.617 ***	15.757 ***
Q(10)	91.822 ***	11.844	125.821 ***	9.206	108.4010 ***	20.790 **
Q(15)	95.692 ***	15.182	131.185 ***	11.796	134.1770 ***	31.453 ***
			threshold exceedance magnitudes			
Min	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Max	0.130	0.110	0.116	0.109	0.114	0.099
Mean	0.011	0.015	0.012	0.012	0.010	0.012
SD	0.014	0.020	0.014	0.017	0.012	0.016
Q(5)	92.293 ***	14.309 **	51.668 ***	16.47251 ***	174.289 ***	27.924 ***
Q(10)	95.931 ***	16.093 *	61.467 ***	17.225 *	198.963 ***	28.375 ***
Q(15)	97.429 ***	17.579	66.683 ***	18.845	200.149 ***	28.954 **

The SEP-POT model was estimated by maximizing the log likelihood function given in Equations (17)–(19). To this end, we used the constrained maximum likelihood (CML) library of the Gauss mathematical and statistical system. The standard errors of the parameter estimates were derived from the asymptotic covariance matrix based on the (inverse) of a computed Hessian. Table 2 presents the estimation results for the CAC 40, DAX 30, and FTSE 100 (analogical results for the remaining time series can be obtained from the author upon request). The parameter estimates

responsible for the self-excitement mechanism, both in the probability of threshold exceedances (i.e., $\hat{\alpha}$, $\hat{\omega}$, $\hat{\kappa}$) and the magnitudes of these exceedances (i.e., $\hat{\alpha}_s$, $\hat{\omega}_s$) are highly statistically significant. The parameter estimates for DAX 30 and CAC 40 indices look very much alike, especially for the conditional probability of threshold exceedances, which means that these two stock markets are closely related to each other.

Table 2. Maximum likelihood (ML) parameter estimates of the self-exciting probability peaks-over-threshold (SEP-POT) for the CAC 40 and DAX 30 indices. Standard errors given in brackets.

Parameter	CAC 40		DAX 30		FTSE 100	
model for the probability of threshold exceedances						
μ	0.017	(0.002)	0.014	(0.002)	0.012	(0.002)
α	0.710	(0.052)	0.776	(0.053)	0.823	(0.063)
ω	13.452	(1.583)	13.919	(1.531)	20.923	(4.831)
κ	0.719	(0.086)	0.751	(2.806)	1.655	(0.490)
model for the sizes of threshold exceedances						
μ_s	0.006	(0.000)	0.006	(0.001)	0.005	(0.001)
α_s	2.225	(0.389)	2.242	(0.388)	2.583	(0.457)
ω_s	7.161	(2.272)	10.439	(3.793)	12.624	(4.075)
ξ	0.122	(0.044)	0.110	(0.042)	0.070	(0.038)
AIC	16.001		15.997		15.978	
BIC	72.454		72.439		71.701	

Obtained series for $\hat{p}_t$, $\hat{\sigma}_t$, and $\hat{VaR}_{0.01,t}$ are illustrated in Figure 4. The extreme loss probability (i.e., $\hat{p}_t$) features a strong self-excitation property because it reacts to extreme-loss days with abrupt increases and, if there are no further intervening events, it slowly wanders in the downward direction. In calm and prosperous periods of the stock market history, the path of $\hat{p}_t$ rests on very low levels. However, in turbulent periods, when the location of extreme-loss days is very dense, $\hat{p}_t$ tends to involve very high numbers. More specifically, persistently elevated $\hat{p}_t$ levels can be seen during the market downturn of 2002–2003 and the global crisis of 2008–2009. For the CAC 40 and FTSE 100, the highest in-sample $\hat{p}_t$ level, equal to 0.2834 (CAC 40) and 0.3082 (FTSE 100), was reached on Monday, 24 November 2008. Both maximum values were triggered by a self-excitation mechanism during the prevailing stock market turmoil. Directly before 24 November 2008 the market suffered three consecutive extreme-loss days–November 19. (Wednesday), 20. (Thursday) and 21. (Friday). For the DAX 30 index, the in-sample $\hat{p}_t$ peaked to its highest level (0.3126) on 11 November 1987, in the aftermath of 10 steep losses that started on the Black Monday of 19 October. The last three were observed on three business days, 6–10 November 1987. Out-of-sample, the highest $\hat{p}_t$ levels of 0.2298 (CAC 40), 0.2416 (DAX 30), and 0.2339 (FTSE 100) corresponded to 24 March 2020 (CAC 40 and DAX 30) and 19 March 2020 (FTSE 100). COVID-19-induced anxiety before 24 March, resulted in the concentration of six threshold exceedances for CAC 40 and DAX 30 in March 2020 alone, where the last of these threshold exceedances took place just one day before the highest $\hat{p}_t$ level was reached on 23 March 2020.

Observed fluctuations of $\hat{p}_t$ are accompanied with the strongly time-varying behavior of $\hat{\sigma}_t$ (i.e., the estimate of the dispersion parameter in the conditional distribution of threshold exceedances). The losses exceeding u trigger upward jumps in both numbers, boosting the awaited probability and the size of a threshold exceedance. For the CAC 40 index, $\hat{\sigma}_t$ peaked to its highest level (0.059) on 15 May 1981, due to enormous panic and sell-offs on the Paris Bourse just days before Francois Mitterand announced hostile reforms for the stocks quoted at the Bourse. Indeed, the preceding days saw the CAC 30 index plunge by over 30%. The UK and German markets were mostly untouched by these French policy-oriented events, and the highest $\hat{\sigma}_t$ was registered on 27 October 1987 (FTSE 100) and 29 October 1987 (DAX 30) at the levels of 0.051 (FTSE 100) and 0.042 (DAX 30), just after a few huge price drops were observed including the famous Black Monday on 19 October 1987. Note, that

the maximum $\hat{\sigma}_t$ levels do not have to coincide with those of $\hat{p}_t$. This is because $\hat{\sigma}_t$ is also affected by the magnitude of past threshold exceedances. For all data in this study, the highest out-of-sample $\hat{\sigma}_t$ levels were registered in the second half of March 2020.

The self-triggering nature of $\hat{p}_t$ and $\hat{\sigma}_t$ give rise to variations in daily VaR, as shown in the panel [c] of Figure 4. What catches special attention is that the obtained path of VaR estimates tends to adjust to both periods of calm and turmoil in the history of equity markets—it quickly reacts to price jumps and bursts in volatility and accounts for persistent swings in stock prices.

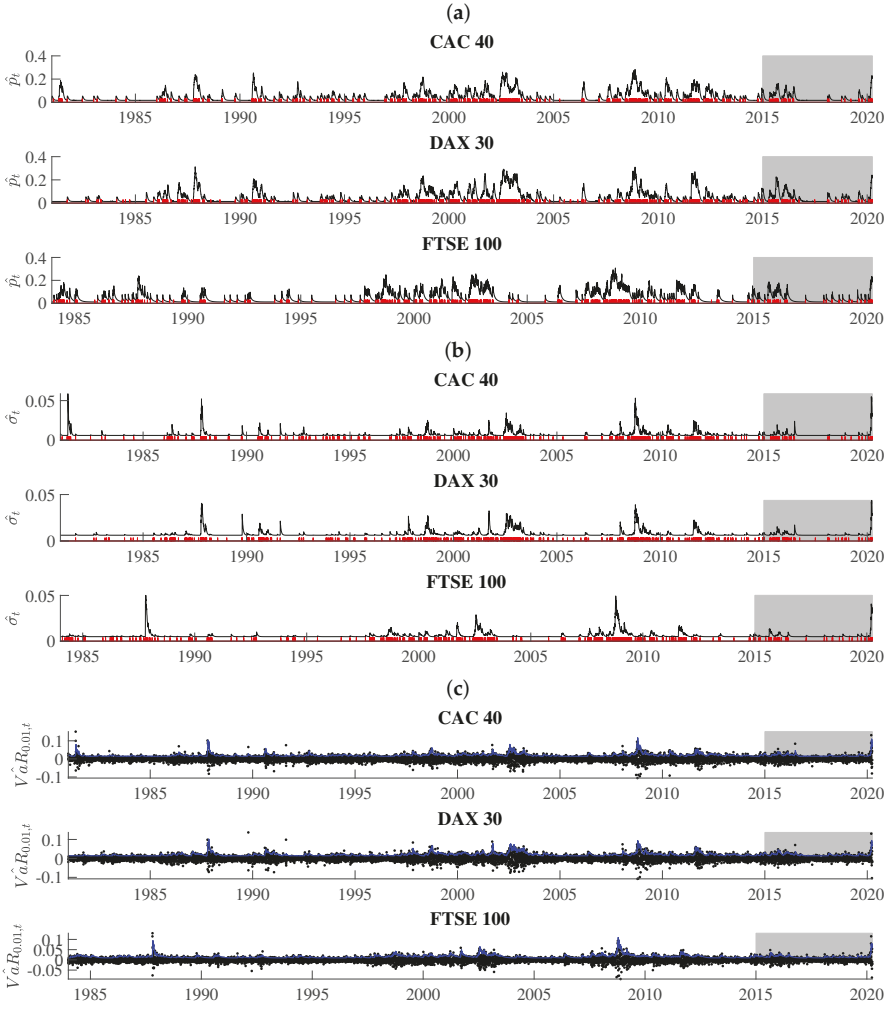


Figure 4. Estimation results from the SEP-POT models: the conditional probability of a threshold exceedance (i.e., p_t, panel (**a**)); the time-varying scale parameter of the generalized Pareto (GP) distribution for the magnitudes of threshold exceedances (i.e., s_t panel (**b**)); the daily value at risk (VaR) at the confidence level 99 % (in blue color) that overlays the (negated) log returns (panel (**c**)). The days of extreme losses were marked in red. The shadowed area corresponds to the out-of-sample period.

We verified whether the SEP-POT model is appropriate for forecasting the daily VaR. To ensure a big-picture perspective over its usefulness in diverse practical applications, we derived the daily VaR

levels for six assumed theoretical coverage rates (i.e., for $q \in \{0.05, 0.025, 0.01, 0.005, 0.0025, 0.001\}$), and compared them with corresponding VaR numbers from three competing risk models (i.e., the self-exciting intensity (Hawkes) POT model (SEI-POT), the EGARCH(1,1) model with the skewed-t distributed innovations and the standard GARCH(1,1) model with normally-distributed innovations). For the sake of fair comparison between the four risk models under study, the accuracy of VaR forecasts was validated with four backtesting procedures. Moreover, each of these statistical routines was distinctly applied to examine the following: (1) the in-sample goodness-of-fit and (2) the out-of-sample accuracy. Considering ten financial instruments under study, six coverage levels for VaR (q) and four models (SEP-POT VaR, SEI-POT VaR, skewed-t-EGARCH VaR, and Gaussian GARCH VaR), we ended up with 240 VaR series in-sample and 240 series out-of-sample. Therefore, for clarity of exposition, the backtesting results were summarized in the form of heatmap graphs (cf., Figures 5–8). Heatmaps use a grid of colored rectangles across two axes where the horizontal axis corresponds to the assumed VaR coverage level and the vertical axis corresponds to the financial instrument under study. The color of each little rectangle (in shades of red and green) reflects the p-value of a backtesting procedure. The white colour corresponds to a p-value equal to 0.05. The darker the red color indicates an increasingly smaller p-value, one that it is less than 0.05. The darker the tone of green indicates an increasingly higher p-value, one that it is larger than 0.05. For example, panel [a] of Figure 5 presents the p-values corresponding to the UC test statistics. Each of the four heatmaps in panel [a] refers to the VaR delivered from a different model: the SEP-POT, SEI-POT, skewed-t-EGARCH, and Gaussian GARCH.

According to the UC test results, the VaR based on the SEP-POT, SEI-POT, and skewed-t-EGARCH models produce, in-sample, a rather accurate proportion of violations. The best in-sample results were delivered by the skewed-t-EGARCH model; however, its superiority diminishes out-of-sample, where the skewed-t-EGARCH model failed in 13 out of 60 instances. Out-of-sample, the SEP-POT VaR and SEI-POT VaR models rejected the null of correct coverage only three times. The EGARCH model seems to produce good VaR forecasts for high coverage levels (i.e., $q = 0.05$). For $q < 0.05$, the EGARCH VaR model is left behind the SEI-POT VaR model and SEP-POT VaR model. As expected, the advantage of VaR models based on POT methodology is most visible for the extreme quantiles. As far as the Gaussian GARCH VaR model is concerned, its performance is dramatically worse than other risk models both in-sample and out-of-sample. The model produces incorrect VaR forecasts for small q (i.e., $q \leq 0.025$), which can be explained by insufficient probability mass in the tails of Gaussian distribution.

The results of the CC test checking both the correct coverage and the lack of dependence of order one in VaR violations seem to support the SEP-POT VaR model (cf., Figure 6). The poorest fit corresponds to the highest q levels (i.e., $q = 0.05$) because in such cases, the null of proper specification had to be rejected both in-sample and out-of-sample for FTSE 100, KOSPI, NIKKIEI, and S&P 500. However, the SEP-POT VaR model seems to be slightly superior than the SEI-POT VaR model. In sample, only in six instances out of 60 did the SEP-POT VaR model fail. For the SEI-POT VaR model, the number of failures was 10 and for the skewed-t-EGARCH VaR model it was nine. As in the case of the UC test, the CC test results indicate that the Gaussian GARCH VaR model rendered the worst fit—the null was not rejected in only seven cases, mainly for the lowest quantiles (i.e., for $q = 0.05$). Out-of-sample, the SEP-POT and the SEI-POT models deliver the similar quality of daily VaR forecasts and both win over GARCH-family models.

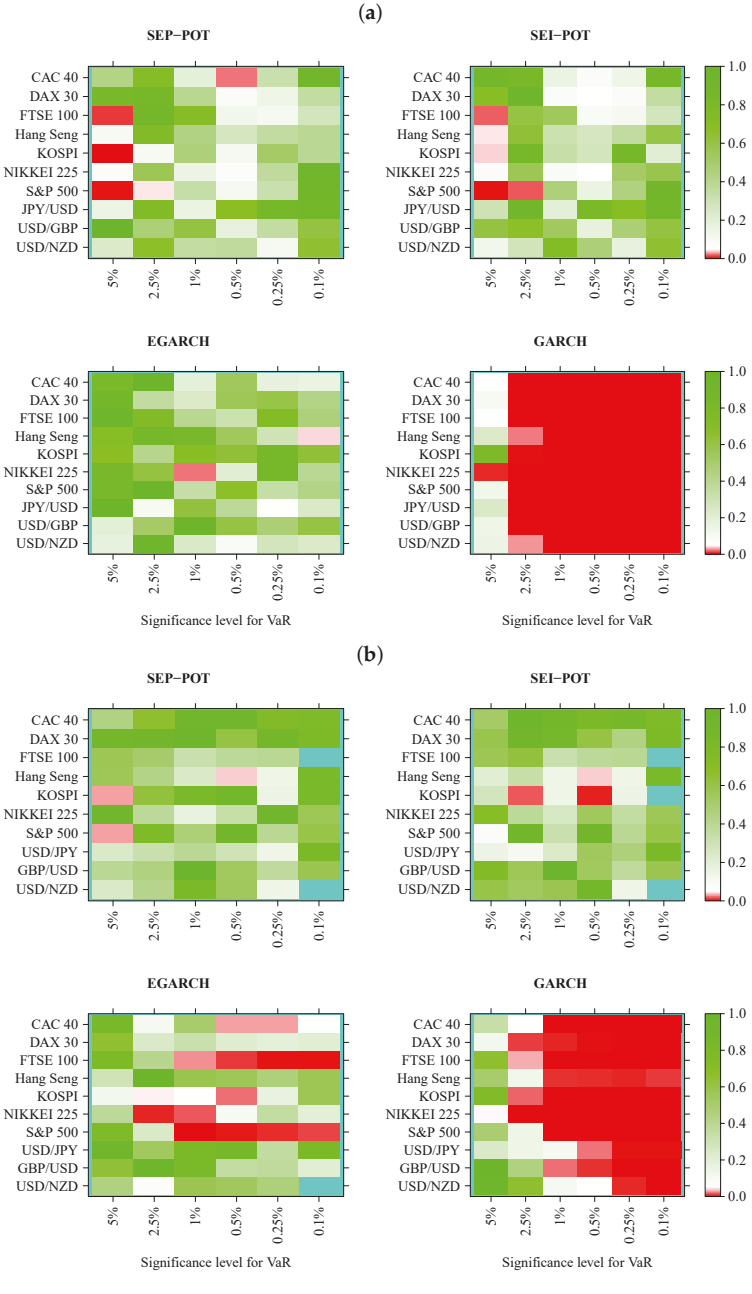


Figure 5. Heatmap charts showing p-value for the in-sample (panel (**a**)) and out-of-sample (panel (**b**)) for unconditional coverage (UC) tests. VaR series was calculated from the self-exciting probability POT model (SEP-POT), self-exciting intensity (Hawkes) POT model (SEI-POT), the EGARCH(1,1) model with the skewed-t distribution (EGARCH), and standard GARCH(1,1) model with normally-distributed innovations (GARCH). The squares of the heatmaps in the shades of red correspond to p-value < 0.05. The rectangles in turquoise color correspond to no VaR exceedances.

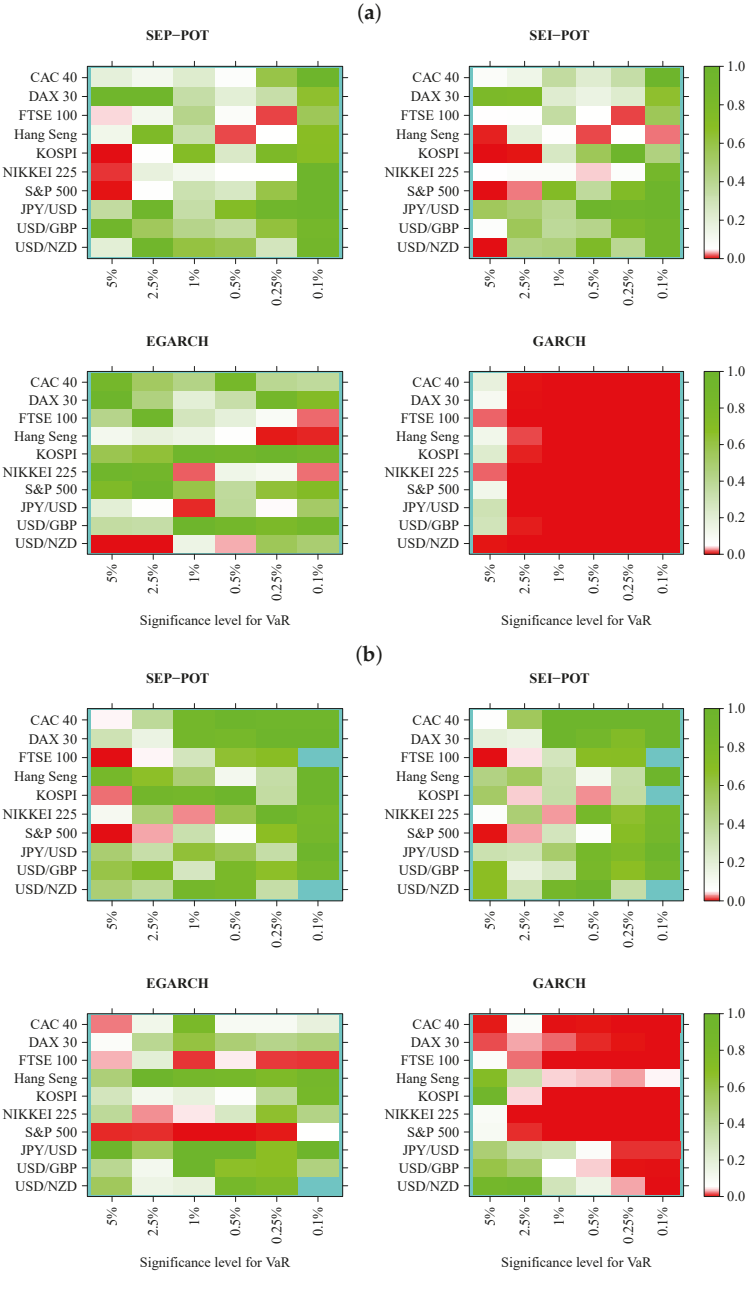


Figure 6. Heatmap charts showing p-value for the in-sample (panel (**a**)) and out-of-sample (panel (**b**)) for conditional coverage (CC) tests. VaR series was calculated from the self-exciting probability POT model (SEP-POT), self-exciting intensity (Hawkes) POT model (SEI-POT), the EGARCH(1,1) model with the skewed-t distribution (EGARCH), and standard GARCH(1,1) model with normally-distributed innovations (GARCH). The squares of the heatmaps in the shades of red correspond to p-value < 0.05. The rectangles in turquoise color correspond to no VaR exceedances.

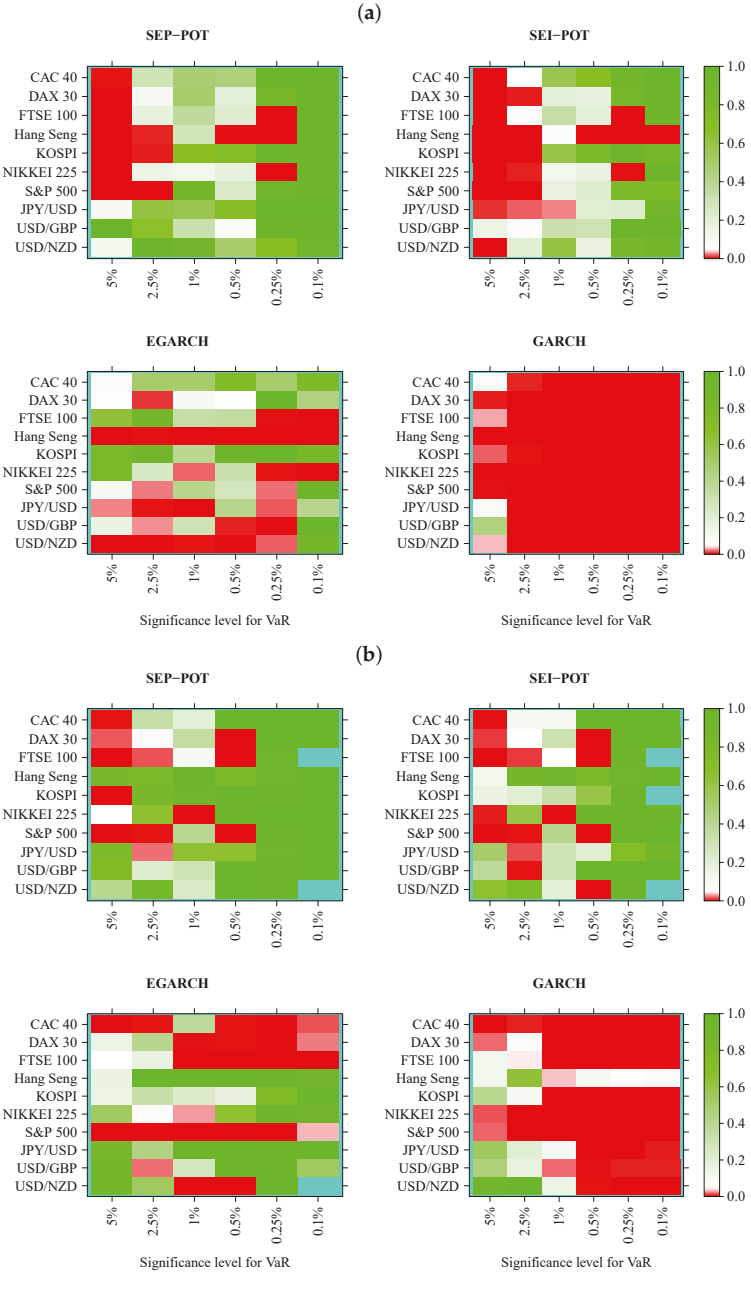


Figure 7. Heatmap charts showing p-value for the in-sample (panel (**a**)) and out-of-sample (panel (**b**)) for dynamic quantile (DQ) conditional coverage tests. VaR series was calculated from the self-exciting probability POT model (SEP-POT), self-exciting intensity (Hawkes) POT model (SEI-POT), the EGARCH(1,1) model with the skewed-t distribution (EGARCH), and standard GARCH(1,1) model with normally-distributed innovations (GARCH). The rectangles of the heatmaps in the shades of red correspond to p-value < 0.05. The rectangles in turquoise color correspond to no VaR exceedances.

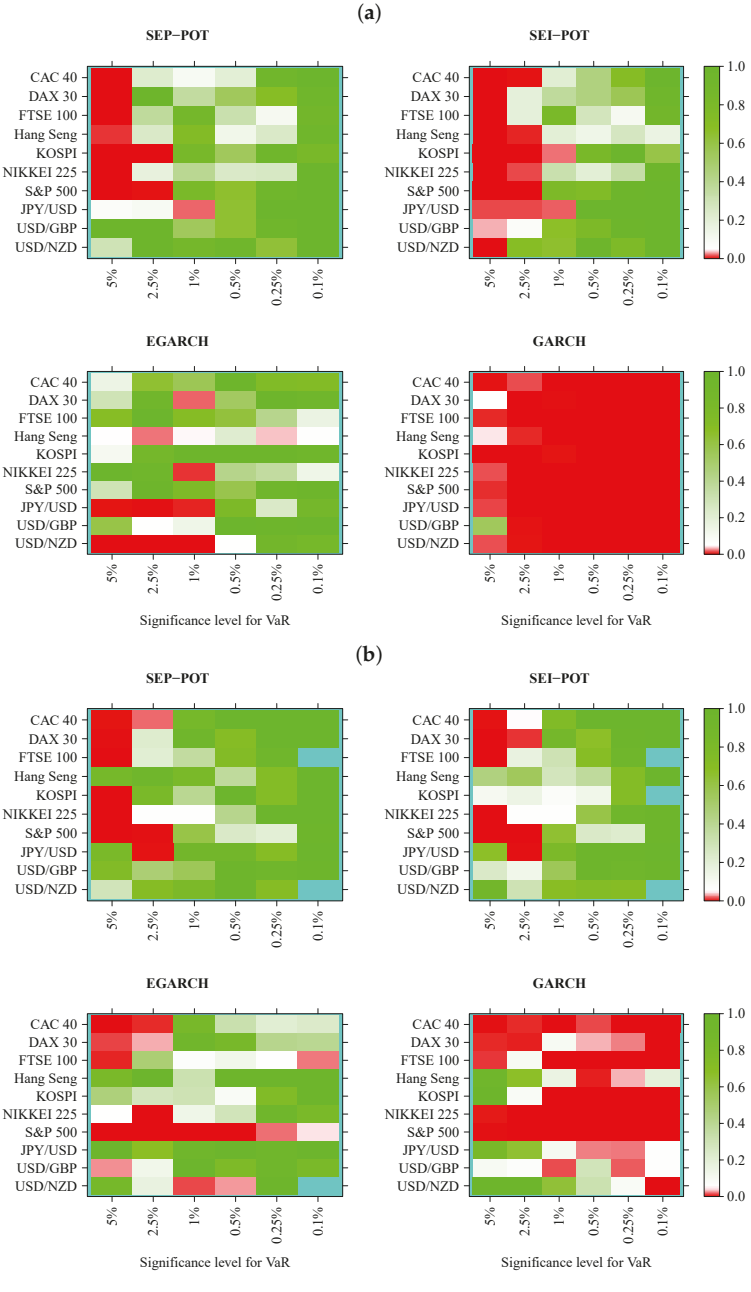


Figure 8. Heatmap charts showing p-value for the in-sample (panel (**a**)) and out-of-sample (panel (**b**)) for dynamic logit conditional coverage tests. VaR series was calculated from the self-exciting probability POT model (SEP-POT), self-exciting intensity (Hawkes) POT model (SEI-POT), the EGARCH(1,1) model with the skewed-t distribution (EGARCH), and standard GARCH(1,1) model with normally-distributed innovations (GARCH). The rectangles of the heatmaps in the shades of red correspond to p-value < 0.05. The rectangles in turquoise color correspond to no VaR exceedances.

Turning our attention to Figure 7, which illustrates the results of the DQ test, the first striking observation is that a much larger area of all heatmaps is marked with shades of red when compared to the results of the CC tests. Indeed, the DQ test is more demanding than the CC test because checks not only whether a VaR violation today is uncorrelated with the fact of a VaR violation yesterday but it also checks whether VaR violations are affected by some covariates from a wider information set, where we used the current VaR and the *Hit* variable observations from one to four days ago (as in original work [27]). The superiority of the SEP-POT VaR model over its competitors is clearly visible. Although the SEP-POT VaR model has a clear tendency to mis-specify VaR at the highest q levels (i.e., $q = 0.05$), the DQ test results for the SEI-POT VaR and the VaR based on the GARCH family models are inferior. In-sample, the DQ test rejected 14 SEP-POT VaR models, 21 SEI-POT VaR models, 26 skewed-t-EGARCH VaR and 57 (i.e., nearly all) Gaussian GARCH VaR models. Out-of-sample, the advantage of the SEP-POT VaR model over the SEI-POT VaR model is less vivid—the first model failed in 12 instances and the latter failed in 14.

Figure 8 illustrates the results of the dynamic logit CC test. We can observe a systematic pattern as far as the SEP-POT VaR and SEI-POT VaR models are concerned. The area marked in red concentrates on the left-hand side of the heatmaps both in and out-of-sample, which means that VaR is mis-specified if derived for high coverage rates (i.e., $q = 0.05$). This deficit of POT VaR models is recouped by their accuracy at low q levels. Indeed, for $q \leq 0.005$ in-sample and for $q \leq 0.01$ out-of-sample, both POT models are not able to reject the null. The SEP-POT VaR model was still slightly more successful than the remaining risk models. In-sample, it failed only 10 times (mainly for $q = 0.05$), whereas the SEI-POT VaR model failed 18 times, the skewed-t-GARCH model failed ten times, yet the Gaussian GARCH VaR model managed to pass this test only two times. Out-of-sample, both POT VaR models were equally correct. For the SEP-POT and SEI-POT VaR model, the null of correct conditional coverage was rejected nine times. The dynamic logit CC test rejected the skewed-t-EGARCH model in 16 and the Gaussian GARCH in majority of cases.

The practical implications of the SEP-POT model stem from its suitability to provide adequate VaR and ES predictions. The VaR forecasts can be used by financial institutions as internal control measures of market risk. The adequacy of risk models used by financial institutions is of utmost importance for the market regulator. Commercial banks have used VaR models for several years to calculate regulatory capital charges using the internal model-based approach of the Basel II regulatory framework. According to the more recent recommendations of the Basel Committee on Banking Supervision (BCBS), banks should use ES to ensure a more prudent capture of "tail risk" and capital adequacy during periods of significant stress in the financial markets [31]. This attitude remains in line with the core objective of the dynamic POT models (including the SEP-POT model), as they focus on the quantification of both the forecasted probability and the awaited size of huge losses, also producing the time-varying ES forecasts. The recent Basel III accord, comprising a set of regulations developed by the BCBS, further reinforces the role of bank units responsible for internal model validations. For more about the current regulatory framework of market risk management see [32]. Despite the recent shift from VaR to ES models in the calculation of capital requirements, ES forecasts remain highly sensitive to the quality of VaR predictions.

All in all, our findings pinpoint that the SEP-POT model constitutes a reasonable promising alternative for forecasting extreme quantiles of financial returns and the daily VaR, especially for very small coverage rates. Undoubtedly, further examination of the theoretical properties of the SEP-POT model and its forecasting accuracy is needed. The model should be backtested using other classes of financial instruments and compared against other extreme risk models. However, there is a plethora of VaR models in the literature—therefore, there are no two or three candidate specifications against which the SEP-POT model should be benchmarked and compared. Only among the point process-based POT models there have been variants put forward, including the ACD-POT model (which is based on the dynamic specifications of time, i.e., duration, that elapses between consecutive extreme losses [6–8]) or the ACI-POT model (with its multivariate extensions) that provides an explicit autoregressive

specification for the intensity function [13]. All these dynamic versions of POT models exploit both strands of the literature: the point process theory and the EVT, accounting for the clustering of extreme losses and the heavy-tailness of the loss distribution. The SEP-POT model is also suitable tailored to these features but also explicitly accounts for the discreteness of times between extreme losses. The empirical findings in this paper provide much support for our SEP-POT model. However, further efforts should be focused on benchmarking and comparison with a broader range of methods under the same settings (i.e., the same data and the same period).

4. Conclusions

We proposed a new self-exciting probability POT model for forecasting the risk of extreme losses. Existing methods within the point process approach to POT models pursue a continuous-time framework and therefore involve specification of an intensity function. Our model is inspired by leading research in this area but is based on observation of the real-world data as we built our model for discrete time. Hence, our model is a dynamic version of a POT model where extreme losses might occur upon a sequence of indivisible time units (i.e., days). Instead of delivering a new functional form for a conditional intensity of the point process, we propose its natural discrete counterpart being the conditional probability of experiencing an extreme event on a given day. This conditional probability is described in a dynamic fashion, allowing the recent events to have a greater effect than the distant ones. Thus, extreme losses arrive according to a self-exciting process, which allows for a realistic capturing of their clustering properties. The functional form of the conditional probability in the SEP-POT model resembles the conditional intensity function used in ETAS models. However, we rely on discrete weighting functions based on at-zero-truncated negative binomial (NegBin) distribution to provide a weight for the influence of past events.

Our move toward the discrete-time setup is backed up by the descriptive analysis of the data. On average, the probability mass for nearly 45% of the time intervals between extreme-loss days is distributed upon a set of discrete values ranging from one up to five days, and the shortest one-day-long duration has a relative frequency of 12% (for the threshold u set equal to the 95%-quantile of the unconditional distribution for negated returns). Accordingly, the motivation of the SEP-POT model lies in allowing the data to speak for itself. Using the at-zero-truncated NegBin distribution as a weighting function in the equation for the conditional probability of extreme loss, we try to tailor the method to the data specificity. The conditional distribution for the magnitudes of threshold exceedances also remain in line with this approach. We specify the evolution of the threshold exceedance magnitudes in a self-exciting fashion utilizing the weighting scheme based on the geometric probability density function. Accordingly, the sizes of more distant threshold exceedances have less effect on the current magnitudes of extreme losses than the more recent events do.

The backtesting results stay in favour of the SEP-POT VaR model. We used four backtesting procedures to check the practical utility of our approach for seven major stock indexes and three currency pairs both in- and out-of-sample. The out-of-sample period covered as much as over five years involving the series of catastrophic downswings in equity prices due to the COVID-19 pandemic in March 2020. We compared VaR forecasts delivered by the SEP-POT model with three widely recognized alternatives: self-exciting intensity (Hawkes) POT-VaR, skewed-t-GARCH VaR and Gaussian GARCH VaR model. Outcomes of backtesting procedures pinpoint that the SEP-POT model for VaR is a good alternative to existing methods.

The standard structure of the SEP-POT model offers several interesting generalizations. For example, it is possible to explain the conditional probability of an extreme loss with some covariates. Some potential candidate explanatory variables include price volatility measures such as high-low price ranges and measures of realized volatility. For stock indexes, some valuable information can be found in volatility indexes such as the CBOE volatility (VIX) index for the U.S. equity market. Unlike existing point process-based POT models, the merits of the SEP-POT model seem to lie in its discrete-time nature. Indeed, the Bernoulli log-likelihood function given in Equation (18) makes it easy

to update an information set in the SEP-POT model on a regular, day-by-day basis. Another interesting generalization of the SEP-POT model could be to add the multi-excitation effect caused by different types of events. For example, the conditional probability of an extreme loss on one market could be additionally co-triggered by crashes observed in another market. Finally, the contemporaneous spillover effect between different markets can be captured using multivariate extensions of the SEP-POT model, for example based on extreme copula functions. These issues are left for further research.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jorion, P. *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd ed.; McGrow-Hill Companies: New York, NY, USA, 2006.
2. Christoffersen, P.F. *Elements of Financial Risk Management*; Elsevier: Amsterdam, The Netherlands, 2012.
3. McNeil, A.; Frey, R. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *J. Empir. Financ.* **2000**, *7*, 271–300. [CrossRef]
4. Chavez-Demoulin, V.; Davidson, A.C.; McGill, J.A. Estimating value-at-risk: A point process approach. *Quant. Financ.* **2005**, *5*, 227–234. [CrossRef]
5. Chavez-Demoulin, V.; McGill, J.A. High-frequency financial data modeling using Hawkes processes. *J. Bank. Financ.* **2012**, *36*, 3415–3426. [CrossRef]
6. Hamidieh, K.; Stoev, S.; Michailidis, G. Intensity-based estimation of extreme loss event probability and value at risk. *Appl. Stoch Model Bus.* **2013**, *29*, 171–186. [CrossRef]
7. Herrera, R.; Schipp, B. Value at risk forecasts by extreme value models in a conditional duration framework. *J. Empir. Financ.* **2013**, *23*, 33–47. [CrossRef]
8. Pyrlik, V. Autoregressive conditional duration as a model for financial market crashes prediction. *Physica A* **2013**, *392*, 6041–6051. [CrossRef]
9. Herrera, R.; González, N. The modeling and forecasting of extreme events in electricity spot markets. *Int. J. Forecast.* **2014**, *30*, 477–490. [CrossRef]
10. Grothe, O.; Korniichuk, V.; Manner, H. Modeling multivariate extreme events using self-exciting point processes. *J. Economet.* **2014**, *182*, 269–289. [CrossRef]
11. Clements, A.E.; Herrera, R.; Hurn, A.S. Modelling interregional links in electricity price spikes. *Energy Econ.* **2015**, *51*, 383–393. [CrossRef]
12. Herrera, R.; Clements, A.E. Point process models for extreme returns: Harnessing implied volatility. *J. Bank. Financ.* **2018**, *88*, 161–175. [CrossRef]
13. Hautsch, N.; Herrera, R. Multivariate dynamic intensity peaks-over-threshold models. *J. Appl. Econ.* **2020**, *35*, 248–272. [CrossRef]
14. Stindl, T.; Chen, F. Modeling extreme negative returns using marked renewal Hawkes processes. *Extremes* **2019**, *22*, 705–728. [CrossRef]
15. Hautsch, N. *Econometrics of Financial High-Frequency Data*; Springer: Berlin, Germany, 2012.
16. Engle, R.F.; Russell, J.R. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* **1998**, *66*, 1127–1162. [CrossRef]
17. Pacurar, M. Autoregressive conditional duration (ACD) models in finance: A survey of the theoretical and empirical literature. *J. Econ. Surv.* **2008**, *22*, 711–751. [CrossRef]
18. Porter, M.D.; White, G. Self-exciting hurdle models for terrorist activity. *Ann. Appl. Stat.* **2012**, *6*, 106–124. [CrossRef]
19. Daley, D.J.; Vere-Jones, D. *An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods*, 2nd ed.; Springer: New York, NY, USA, 2003. [CrossRef]
20. McNeil, A.J.; Frey, R.; Embrechts, P. *Quantitative Risk Management: Concepts, Techniques and Tools*; Princeton University Press: Princeton, NY, USA, 2005. [CrossRef]
21. Lewis, P.A.W.; Shedler, G.S. *Statistical Analysis of Non-Stationary Series of Events in a Data Base System*; Naval Postgraduate School: Monterey, CA, USA, 1976. [CrossRef]

22. Liesenfeld, R.; Nolte, I.; Pohlmeier, W. Modelling Financial Transaction Price Movements: A Dynamic Integer Count Data Model. *Appl. Econ.* **2006**, *30*, 795–825.
23. Bień, K.; Nolte, I.; Pohlmeier, W. An inflated multivariate integer count hurdle model: an application to bid and ask quote dynamics. *J. Appl. Economet.* **2011**, *26*, 669–707. [CrossRef]
24. Yamai, Y.; Yoshiba, T. Comparative analyses of expected shortfall and value-at-risk: their validity under market stress. *Monet. Econ. Stud.* **2002**, *20*, 181–237. [CrossRef]
25. Kupiec, P.H. Techniques for Verifying the Accuracy of Risk Measurement Models. *J. Deriv.* **1995**, *3*, 73–84. [CrossRef]
26. Christoffersen, P.F. Evaluating Interval Rorecasts. *Int. Econ. Rev.* **1998**, *39*, 841–862. [CrossRef]
27. Engle, R.F.; Manganelli, S. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *J. Bus. Econ. Stat.* **2004**, *22*, 367–381. [CrossRef]
28. Dumitrescu, E.; Hurlin, C.; Pham, V. Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests. *Finance* **2012**, *33*, 79–112. [CrossRef]
29. Bank of International Settlements. *Triennial Central Bank Survey. Foreign Exchange Turnover in April 2016*; Bank for International Settlements: Basel, Switzerland, 2016. [CrossRef]
30. Fernandes, M.; Grammig, J. Non-parametric specification tests for conditional duration models. *J. Economet.* **2005**, *127*, 35–68.
31. Basel Committee on Banking Supervision. *Minimum Capital Requirements for Market Risk*; Bank for International Settlements: Basel, Switzerland, 2015.
32. Roncalli, T. *Handbook of Financial Risk Management*; CRC Press: Boca Raton, FL, USA, 2020.

entropy

MDPI

Article

Systemic Importance of China's Financial Institutions: A Jump Volatility Spillover Network Review

Xin Yang [1,2], Xian Zhao [1,2], Xu Gong [3], Xiaoguang Yang [4] and Chuangxia Huang [1,2,*]

1 School of Mathematics and Statistics, Changsha University of Science and Technology, Changsha 410114, China; yangxintaoyuan@csust.edu.cn (X.Y.); zhaoxian1714@163.com (X.Z.)
2 Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering, Changsha University of Science and Technology, Changsha 410114, China
3 School of Management, China Institute for Studies in Energy Policy, Xiamen University, Xiamen 361005, China; xugong@xmu.edu.cn
4 Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; xgyang@iss.ac.cn
* Correspondence: cxiahuang@csust.edu.cn ; Tel.: +86-731-85258293 or +86-731-85258787

Received: 18 April 2020; Accepted: 19 May 2020; Published: 24 May 2020

Abstract: The investigation of the systemic importance of financial institutions (SIFIs) has become a hot topic in the field of financial risk management. By making full use of 5-min high-frequency data, and with the help of the method of entropy weight technique for order preference by similarities to ideal solution (TOPSIS), this paper builds jump volatility spillover network of China's financial institutions to measure the SIFIs. We find that: (i) state-owned depositories and large insurers display SIFIs according to the score of entropy weight TOPSIS; (ii) total connectedness of financial institution networks reveal that Industrial Bank, Ping An Bank and Pacific Securities play an important role when financial market is under pressure, especially during the subprime crisis, the European sovereign debt crisis and China's stock market disaster; (iii) an interesting finding shows that some small financial institutions are also SIFIs during the financial crisis and cannot be ignored.

Keywords: financial institution; complex network; jump volatility; entropy weight TOPSIS

1. Introduction

With the development of economic globalization, the financial system has become more and more closely interconnected by investment networks, debtor–creditor and trade contacts [1–4]. Financial institutions such as depositories, broker-dealers and insurance companies permeate each other by related business and display significant complex network properties [5–7]. The failure of several financial institutions may lead to a severe economic crisis [8–10]. One of the typical examples is the global financial crisis triggered by the collapse of Lehman Brothers in 2008 [11–13]. Therefore, how to accurately evaluate the systemic importance of financial institutions (SIFIs) so as to provide early warning and deal with the crisis effectively has become an emergent work [14–16].

Usually, there are three ways to measure the SIFIs. The first way is to employ Pearson correlation coefficient to calculate the financial institutions' default probabilities [17–19]. Pearson correlation coefficient ignores the heterogeneity of financial data at different times [20]. Adopting a tail-dependence method to measure the systemic risk contributions between financial institutions is the second method. Girardi and Ergün (2013) [21] used the conditional value-at-risk (CoVaR) method to estimate systemic risk of each financial institution. Acharya et al. (2017) [22] employed the systemic expected shortfall (SES) model to calculate financial institution's losses by considering its leverage. Wang et al. (2019) [23] proposed CSRISK model to investigate financial institutions' capital shortfall under the market crash. The above two methods are based on the local correlation and disregard the interlinked

among the financial institutions, which may underrate systemic risk contribution [24]. The latest financial crisis manifests that intricate connections among financial markets can spread risk [25–29]. Using the complex network theory to research SIFIs comes up to the third method. Billio et al. (2012) and Gong et al. (2019) [30,31] applied Granger causality model to build financial institution network and utilized the out degree to compute the total connectedness. Hautsch et al. (2014) [24] proposed VaR model to set up financial institution network and adopted systemic risk betas to investigate the systemic risk contribution. Härdle et al. (2016) and Wang et al. (2018) [32,33] adopted CoVaR model to construct financial institution network and used the index of out degree to measure the system risk contribution.

Most of the literature evaluates the SIFIs by out degree [31–33], which can reveal the range of risk contagion but restrict to the local information of the network [34,35]. Recently, considering financial institutions have the characteristics of deeper risk contagion extent, higher risk contagion efficiency and greater risk contagion degree after the outbreak of financial risk, some other indicators such as, clustering coefficient [36], closeness centrality [37] and Leaderrank value [38] are also applied to measure the SIFIs. Although most of the above indicators have been investigated extensively and many findings on SIFIs have been reported recently, which only reflect one characteristic of the network [36–38]. A comprehensive evaluation with respect to the entire network appears to be very few. Such studies are however essential to accurately evaluate SIFIs in practice. To deal with this issue, combining four indicators (out degree [32], clustering coefficient [36], closeness centrality [37] and Leaderrank value [38]) and assigning different weights to each indicator may give a better evaluation. However, the selection of weight is often based on the subjective experience of researchers, rather than sufficient scientific support, which may lead to inaccurate evaluation results. As we all known, entropy weight technique for ioder preference by similarities to ideal solution (TOPSIS) is a multiple criteria decision making method, and it bases on the conception that the selected alternative should have the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution. Entropy weight TOPSIS has been proved to be a good method in strategic decision making and successfully applied in some fields, such as coal mine safety [39], multinational consumer electronics company [40] and transport [41]. Therefore, it seems that adopting entropy weight TOPSIS to comprehensively assess the SIFIs might be a better choice.

It should be pointed out that the all above mentioned literatures on measuring SIFIs have been greatly limited to low-frequency data. The low-frequency data with daily, weekly, monthly, quarterly or annual sampling frequency can not accurately measure the whole-day volatility information [42]. Nowadays, more and more scholars have realized that the high-frequency data with the frequency of hours, minutes or even shorter includes the rich information of asset price, and it has been intensively studied in applied finance risk management [43–46]. On the other hand, with the unexpected changes of macroeconomic conditions, international events and economic policy in recent years, financial markets are increasingly volatile [47]. Some researches detected jump volatility in the volatile process of financial assets based on high-frequency data [48]. For example, Wright and Zhou (2007) [49] found that jump volatility can explain much of the countercyclical movements in bond risk premium. Zhang et al. (2016) [50] found that jump volatility is an important component of Dow Jones Industrial Average stocks' volatility. Audrino and Hu (2016) [51] found that jump volatility can improve the forecast of S&P 500's volatility.

The jump volatility depicts an infrequent but a sharp change of asset price, and it can better describe violent volatility of financial market than continuous volatility [52]. Measuring SIFIs associated with jump volatility spillover network and high-frequency data has not been reported yet and it still remains a challenging problem. Motivated by the above discussions, in this paper, we aim to employ high-frequency data of China's financial institutions to construct jump volatility spillover network, and then utilize entropy weight TOPSIS to comprehensively assess the SIFIs. The innovations of this paper are as follows:

(1) Many scholars investigated the jump volatility of a single financial asset on its price fluctuation from the perspective of prediction. We first propose Granger-causality test to identify the jump volatility spillover among financial institutions.
(2) Financial markets are extremely volatile, and the low-frequency data might lose a lot of important information. By employing 5-min high-frequency data, we establish the jump volatility spillover network, which can capture the jump volatility spillover among financial institutions.
(3) We use entropy weight TOPSIS rather than a single indicator to comprehensively assess the SIFIs.

The reminder of this paper is arranged as follows. In Section 2, we introduce the methodology. In Section 3, we present the data. In Section 4, we give an empirical analysis. Finally, we make conclusions and discuss our findings in Section 5.

2. Methodology

In this section, we introduce the method of network construction and the indicator for assessing the SIFIs. Specifically, in Section 2.1, we use Granger causality test to build the network, which reflects statistically significant relations between jump volatility spillover of financial institutions. In Section 2.2, out degree, clustering coefficient, closeness centrality and leaderrank algorithm are employed to evaluate the SIFIs, respectively. In Section 2.3, by the method of entropy weight TOPSIS, we integrate the above four indicators into a comprehensive indicator to measure the SIFIs.

2.1. Network Construction

We establish jump volatility spillover network of financial institution according to the following three steps, where each financial institution represents a network node, and each pair of the financial institution is connected with an edge calculated by Granger-causality test.

In the first step, we employ Andersen et al. (2007, 2012) tests to extract jump volatility of financial institutions. We suppose that the logarithmic price of a financial institution ($p_t = \ln P_t$) within the trading day obeys a standard jump diffusion process:

$$dp_t = \mu_t dt + \sigma_t dW_t + \kappa_t dq_t, \quad 0 \leqslant t \leqslant T, \tag{1}$$

where μ_t denotes the drift term, which includes a continuous volatility sample path; σ_t represents a strictly positive stochastic volatility process; W_t stands for a standard Brownian motion; $\kappa_t dq_t$ is the pure jump component.

Meanwhile, the logarithmic return volatility can be expressed as quadratic volatility, which contains jump volatility rather than unbiased estimator of integrated volatility:

$$QV_t = \int_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1<s\leq t} k_s^2 \tag{2}$$

where $\int_{t-1}^{t} \sigma_s^2$ represents the continuous volatility, and $\sum_{t-1<s\leq t} k_s^2$ stands for intra-day jump volatility.

Since the quadratic volatility can not be gained directly, this paper employs the estimated realized volatility RV_t to replace it based on Andersen et al. (2012) [53]:

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2, \tag{3}$$

where $r_{t,i} = (\ln P_{t,i} - \ln P_{t-1,i}) \times 100$, $P_{t,i}$ denotes the closing price of financial institution i at time t, and $M = 48$ represents the daily trading frequency.

In addition, when $M \to \infty$, $\int_{t-1}^{t} \sigma_s^2 ds$ can be calculated by the realized bipower volatility $MedRV_t$ based on Barndorff-Nielse et al. (2004) [54,55].

$$MedRV_t = \frac{\pi}{6 - 4\sqrt{3} - \pi} \left(\frac{M}{M-2} \right) \sum_{i=2}^{M-1} \text{Med} \left(\left| r_{t,i-1} \right| \left| r_{t,i} \right| \left| r_{r,i+1} \right| \right)^2, \tag{4}$$

If there is no jump in the price of financial institutions, the difference between realized volatility and bipower volatility is 0. Otherwise, Z-statistic is adopted to identify jump volatility [54]:

$$Z_t = \frac{\left(RV_t - MedRV_t \right) RV_t^{-1}}{\sqrt{\left(\mu_1^{-4} + 2\mu_1^{-2} - 5 \right) \frac{1}{M} \max \left(1, \frac{medRTQ_t}{medRV_t^2} \right)}} \to N(0,1), \tag{5}$$

where $\mu_1 = \sqrt{2/\pi}$, and $\text{med}\, RTQ_t = \frac{3\pi M}{9\pi + 72 - 52\sqrt{3}} \left(\frac{M}{M-2} \right) \sum_{i=2}^{M-1} \text{Med} \left(\left| r_{t,i-1} \right| \left| r_{t,i} \right| \left| r_{r,i+1} \right| \right)^4$, which stands for realized tri-power quarticity.

Based on Z statistics, we can obtain realized jump volatility:

$$J_t^d = I \left\{ Z_t > \Phi_t \right\} \left(RV_t - MedRV_t \right), \tag{6}$$

where $I(\cdot)$ is an indicator function, and α chooses as 0.95 (see Andersen et al., 2007 [56]).

In the second step, after extracting the jump volatility of a single financial institution, we investigate whether there is jump volatility spillover between financial institutions according to Granger-causality test [57]. If the p values of Granger-causality test are smaller than the critical values under the 5% significance level [58], there exists causality relationships between financial institutions.

In the last step, we construct a Granger-causality jump volatility spillover network of financial institutions. And the network can be represented by an adjacency matrix AD:

$$AD = (V, E) = \begin{bmatrix} 0 & \cdots & \cdots & AD_{1j} & \cdots & AD_{1n} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ AD_{i1} & \ddots & \ddots & AD_{ij} & \ddots & AD_{in} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ AD_{n1} & \cdots & \cdots & AD_{nj} & \cdots & 0 \end{bmatrix} \tag{7}$$

where V is nodes set and E is the edge set. n is the number of financial institutions. AD_{ij} is defined as follow:

$$AD_{ij} = \begin{cases} 1 & i\ Granger\ causes\ j\ significantly \\ 0 & i\ doesn't\ Granger\ causes\ j \end{cases} \tag{8}$$

2.2. Indicator for Assessing the Systemic Importance of Financial Institutions

There are a growing number of indicators to evaluate SIFIs. Taking into account financial institutions have the characteristics of wider risk contagion range, deeper risk contagion extent, higher risk contagion efficiency and greater risk contagion degree, we choose out degree, clustering coefficient, closeness centrality and leaderrank value to assess the SIFIs, respectively. And more and more scholars use these four indicators to study the SIFIs [33,36–38].

2.2.1. Out Degree

Out degree (OD) calculates the number of edges that node i point to other nodes. It is used to measure the risk contagion range [36]. When the risk occurs, it will directly transfer the risk to the connected nodes. The higher the out degree of nodes, the wider the range of risk transmission. The expression of out degree is as follows:

$$OD_{out}(i) = \sum_{j=1}^{n} AD_{ij}, \tag{9}$$

where AD_{ij} stands for the adjacency matrix of financial institution network.

2.2.2. Clustering Coefficient

The clustering coefficient (C) measures the degree of interconnection between the neighbors of a node in the graph. If one node owes high clustering coefficient, the risk may spread to their neighbor nodes when one financial institution fluctuates. Furthermore, the interconnectedness of neighbor nodes will cause risk contagion again and aggravate the risk contagion extent of the whole financial institutions [36]. Therefore, we employ clustering coefficient as the risk contagion extent of each financial institution, and it is computed as follows:

$$C_i = m_i / \alpha_i(\alpha_i - 1), \tag{10}$$

where $\alpha_i(\alpha_i - 1)$ represents the maximum number of possible edges of financial institution i, and m_i stands for the actual number of existing edges.

2.2.3. Closeness Centrality

Closeness centrality (CC) quantifies how close a node is to all other nodes in the financial institution network. The closeness centrality of a node is inversely proportional to the average shortest path distance from one node to any other nodes in the network. The larger value of the closeness centrality of a node, the faster the risk will be transferred from one node to any other nodes. Hence, the closeness centrality can depict how efficiently each node transmits risk to all other nodes [37], and it is expressed as follows:

$$CC_{out}(i) = \sum_{j=1, j\neq i}^{N} 2^{-d_{ij}}, \tag{11}$$

where d_{ij} is the shortest distance i to j.

2.2.4. Leaderrank Algorithm

LeaderRank (LR) algorithm is a method to identify key nodes in a complex network. The basic idea of the algorithm is as follows. We add a new node (called ground node) and connect it to all others by bidirectional edges for a directed network with M nodes and N edges. The new network is strongly linked, which owes $M + 1$ nodes and $N + 2M$ edges. Matrix $A = (a_{ij})$ depicts the connectivity of the network. If $AD_{ij} = 1$, which means that node i can pass financial risk to node j. The LR gives a score to each node, where score denotes the SIFIs. Scores are assign by $LR_g(0) = 0$ for ground node and $LR_i(0) = 1$ for other nodes. Thus, scores are updated by

$$LR_i(t) = \sum_{j=1}^{M+1} \frac{a_{ji}}{OD_{out}(j)} LR_i(t-1), \tag{12}$$

where $OD_{out}(j)$ is out degree.

After t iterations, the LR values of all nodes are stable. At this time, the ground node score is averagely distributed to each network node. Consequently, the final score of the network node reflects its cumulative risk ability. The higher the score, the stronger the cumulative risk degree of the node [38].

2.3. Entropy Weight TOPSIS

Measuring risk contagion of financial institutions from different indicators may lead to inconsistent results. Therefore, the construction of risk contagion composite index of financial institutions is an essential step in this paper. We adopt entropy weight TOPSIS to evaluate SIFIs. It can avoid the subjectivity of weight selection and make full use of the sample data [39–41].

The entropy of each indicator is calculated as below:

$$e_j = -\frac{1}{\ln N}\sum_{i=1}^{N} p_{ij}\ln p_{ij}, p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{N} x_{ij}^2}, \tag{13}$$

where $j = 1, \ldots N$; $i = 1, \ldots n$; $N = 24$; $n = 4$, x_{ij} denotes the jth indicator value of the ith financial institution of the initial matrix X; p_{ij} stands for the jth normalized indicator value of the ith normalized matrix P.

The weight (w_j) can be calculated as follows:

$$w_j = \frac{1 - e_j}{N - \sum_{i=1}^{n} e_j}, j = 1, 2, \cdots, N. \tag{14}$$

Then, the TOPSIS method ranks financial institutions based on their relative proximities, and the positive ideal solution and the negative ideal solution. The distance of each indicator from D_i^+ and D_i^- can be calculated as the following:

$$D_i^+ = \sqrt{\sum_{j=1}^{N}\left(S_{ij} - S_i^+\right)^2}, j = 1, 2, \cdots, N, \tag{15}$$

$$D_i^- = \sqrt{\sum_{j=1}^{N}\left(S_{ij} - S_j^-\right)^2}, j = 1, 2, \cdots, N, \tag{16}$$

where $S_{ij} = w_j p_{ij}$, $S_i^+ = \max\limits_{1\le j\le n}\left(S_{ij}\right)$, and $S_i^- = \min\limits_{1\le j\le n}\left(S_{ij}\right)$.

The relative proximity K_i, which is regarded as score of each financial institution, can be computed as follows:

$$K_i = D_i^- / \left(D_i^- + D_i^+\right), \quad j = 1, 2, \cdots, N. \tag{17}$$

Finally, in order to measure ability of risk contagion, we use K_i to rank the SIFIs.

3. Data

We select 24 listed financial institutions from 2008 to 2018 in China, similar sample selection can be found in Wang et al. (2018) [33]. We choose 2008 as the starting date due to several important financial institutions do not go public until 2007, such as the China Construction Bank. We divide listed financial institutions into three sectors: (1) depositories, (2) insurance companies and (3) broker dealers. The data are available from Wind Financial dataset, and the descriptive statistics of 5-min high-frequency closing price data of financial institutions are shown in Table 1.

Table 1. The descriptive statistics of 5-min high-frequency closing price data of 24 financial institutions.

Code	Financial Institution	Abbreviation	Mean	SD	MAX	MIN
Panel A: Depositories						
S000001.SZ	Ping An Bank	PAB	14.77	5.38	44.45	8.03
S002142.SZ	Bank of Ningbo	NBCB	13.00	3.58	24.27	5.87
S600000.SH	Shanghai Pudong Development Bank	SPDB	14.73	6.91	61.80	7.11
S600015.SH	Huaxia Bank	HXB	10.37	2.14	23.52	6.14
S600016.SH	China Minsheng Banking Corp., Ltd.	CMBC	7.56	1.85	16.38	3.89
S600036.SH	China Merchants Bank	CMB	17.21	6.57	43.58	9.42
S601009.SH	Bank of Nanjing	NJBK	11.24	3.55	23.49	6.41
S601166.SH	Industrial Bank	CIB	18.94	8.47	61.79	8.62
S601169.SH	Bank of Beijing	BOB	10.24	3.19	22.57	5.53
S601328.SH	Bank of Communications	BOCOM	6.01	1.76	16.15	3.60
S601398.SH	Industrial and Commercial Bank of China Ltd.	ICBC	4.61	0.83	8.35	3.15
S601939.SH	China Construction Bank	CCB	5.39	1.17	10.18	3.50
S601988.SH	Bank of China	BOC	3.53	0.69	6.93	2.44
S601998.SH	China CITIC Bank	CNCB	5.60	1.29	10.93	3.39
Panel B: Broker-dealers						
S000686.SZ	Northeast Securities	NESC	17.82	8.78	56.33	5.10
S000728.SZ	Guoyuan Securities	GYSC	15.36	7.00	45.70	5.55
S000783.SZ	Changjiang Securities	CJSC	11.50	4.78	40.15	4.07
S600030.SH	CITIC Securities	CITICS	19.05	10.57	98.07	9.10
S600109.SH	Sinolink Securities	SLSC	17.83	9.28	71.66	5.77
S600837.SH	Haitong Securities	HTSEC	14.17	7.25	61.58	7.06
S601099.SH	Pacific Securities	PSC	9.66	6.30	47.45	1.94
Panel C: Insurances						
S601318.SH	Ping An Insurance (Group) Co. of China, Ltd.	PAI	47.47	14.37	112.55	19.90
S601601.SH	China Pacific Insurance (Group) Co., Ltd.	CPIC	24.72	7.09	50.25	10.36
S601628.SH	China Life Insurance (Group) Co., Ltd.	CLI	23.24	6.65	60.60	12.89

Note: SZ denotes that financial institution is transacted by the Shenzhen stock exchange market, and SH means that financial institution is transacted by the Shanghai stock exchange market.

4. Empirical Analysis

4.1. Jump Volatility Spillover Network Construction of Financial Institution

The sampling frequency selection of intraday high-frequency data is very important for jump volatility measurement. The low sampling frequency may not fully express the jump volatility information, while high sampling frequency can cause micro structural noise. According to Haugom et al. (2014) [59], Gong and Lin (2018) [46] and Wen et al. (2019) [55], this paper first adopts 5-min high-frequency closing price data of financial institutions to compute jump volatility based on Equations (1)–(9). Then, using Granger causality test to investigate jump volatility spillover relations among financial institutions, we can obtain the p value of a 24×24 matrix. This paper chooses a threshold of 0.05 according to Jiang et al. (2017) [58]. Finally, the financial institution network gets a total of 24 nodes and 137 edges, and the results are illustrated by Figure 1.

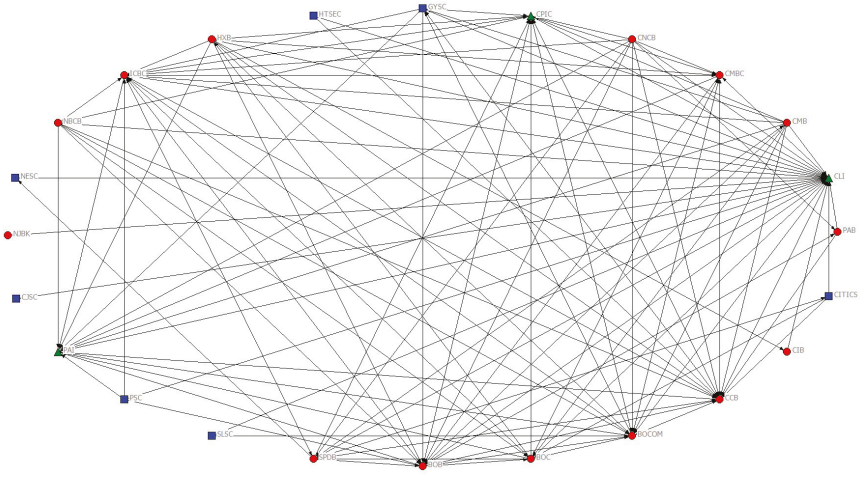


Figure 1. Jump volatility spillover network of financial institution. Note: Nodes (financial institutions) from the same sector are signed as the same shape and color. Depositories, broker dealer, and insurances are labelled as red circle, blue square, and green triangle, respectively.

4.2. Assessing the Systemic Importance of Financial Institutions

The research on the systemic importance of financial institutions (SIFIs) has become a hot topic in financial risk management. In this section, we choose out degree, clustering coefficient, closeness centrality and leaderrank value to measure the SIFIs. The results are shown in Table 2.

Table 2 displays four dimensions of risk contagion measurement. (1) In terms of risk contagion range, the larger the out degree value, the wider risk contagion range of financial institution. We can find that CMB, BOC and CNCB are all from depository sector, with the highest out degree value of 10. One possible reason is that China's financial system is a depository-led system. As claimed by the annual reports of CBRC, CSRC, and CIRC in 2019, the total assets of depository, broke-dealer and insurance sectors were 261.4 trillion Yuan, 6.2 trillion Yuan and 18.3 trillion Yuan, respectively. This indicates that the depository sector size is 42 times larger than the broke-dealer sector, or 14 times larger than the insurance sector. (2) In terms of risk contagion extent, the larger the clustering coefficient value, the greater risk contagion extent of financial institution. We can see that SLSC and HTSEC, which are all from broke-dealer sector, have the highest clustering coefficient value. The results show that broke-dealer sector's risk contagion ability can not be neglected in China's financial system. (3) In terms of risk contagion efficiency, the larger the closeness centrality value, the faster the risk will be transferred from one financial institution to any other financial institutions.

We can discover that CJSC and PAI have the highest closeness centrality value, indicating that some broke-dealers and insurances have gradually become important departments in China's banking-led financial system. (4) In terms of risk contagion degree, the larger the leaderrank value, the higher the risk contagion degree. The risk contagion degree of financial institutions manifests a hierarchical feature, i.e., the greatest degree of risk contagion is insurance sector, followed by broke-dealer sector, and depository sector have the lowest risk contagion degree.

Table 2. Indicators score of 24 financial institutions.

Symbol	OD	C	CC	LR
PAB	2	1.0000	0.2602	0.0200
NBCB	8	0.8393	0.1667	0.0120
SPDB	8	0.5600	0.1684	0.0389
HXB	9	0.6889	0.1633	0.0151
CMBC	8	0.7632	0.1739	0.0645
CMB	10	0.6574	0.1509	0.0157
NJBK	1	0.0000	0.2883	0.0120
CIB	1	1.0000	0.2909	0.0129
BOB	9	0.5688	0.1739	0.0806
BOCOM	8	0.6081	0.1928	0.0747
ICBC	9	0.5854	0.1633	0.0786
CCB	8	0.5112	0.1882	0.0930
BOC	10	0.6136	0.1538	0.0403
CNCB	10	0.6667	0.1495	0.0120
NESC	1	1.5000	0.2991	0.0157
GYSC	6	0.8429	0.1975	0.0302
CJSC	1	0.0000	0.3107	0.0271
CITICS	3	0.8333	0.2222	0.0157
SLSC	2	2.0000	0.2520	0.0120
HTSEC	2	2.0000	0.2520	0.0120
PSC	4	1.0833	0.2133	0.0120
PAI	1	0.5278	0.3107	0.0764
CPIC	7	0.6685	0.1975	0.0720
CLI	9	0.3382	0.1684	0.1563

The results of the above four dimensions are inconsistent in measuring the SIFIs. In order to comprehensively assess the SIFIs, this paper proposes entropy weight TOPSIS (EWTOPSIS) to obtain the weight of each indicator. As a result, we gain the weight of out degree, clustering coefficient, closeness centrality and leaderrank algorithm by 0.2807, 0.2499, 0.0401 and 0.4293, respectively.

Table 3 shows the score of SIFIs computed by EWTOPSIS in the whole period. We can find that: (1) CLI is the most SIFIs. This may be related to the deregulation reform of the China's insurance sector in 2014. China's state council issue that the insurance depth will reach 5% and the insurance density will reach 3500 yuan/person by 2020. (2) CCB, BOB, ICBC and BOCOM are deem as more systemically important financial institutions. Because most of them come from state-owned depositories, which dominate China's depository sector about 45% of the lending business in 2018. (3) CJSC and NJBK have a relatively low score, which imply that the impact of some small financial institutions on the financial system can be neglected.

Table 3. The score of systematic importance of financial institutions.

Rank	Symbol	Score	Rank	Symbol	Score
1	CLI	0.7489	13	GYSC	0.2846
2	CCB	0.5160	14	CMB	0.2829
3	BOB	0.4962	15	CNCB	0.2703
4	ICBC	0.4898	16	HXB	0.2658
5	BOCOM	0.4605	17	NBCB	0.2539
6	CPIC	0.4385	18	NESC	0.2239
7	CMBC	0.4409	19	PSC	0.2088
8	BOC	0.3697	20	PAB	0.1914
9	PAI	0.3244	21	CITICS	0.1701
10	SPDB	0.3179	22	CIB	0.1472
11	SLSC	0.2920	23	CJSC	0.0705
11	HTSEC	0.2920	24	NJBK	0.0123

Compared with the traditional evaluation SIFIs methods such as Equal weight [38], principal component analysis (PCA) and TOPSIS method [60], we will show our proposed EWTOPSIS method has obvious advantages. Just as reported by Sandoval (2014) [61], Wang et al. (2018) [33] and Wang et al. (2019) [62], market capitalization as a financial indicator could reflect the market influence of financial institutions. We calculate the correlation between each index (depending on different methods) and market capitalization, and the results can be presented as Table 4. It is easy to see that the EWTOPSIS is most effective method to measure the SIFIs because of the correlation between EWTOPSIS and market capitalization is the largest at a significant level of 1%.

Table 4. The correlation between the market capitalization and four indicators.

Index	Equal Weight	PCA	TOPSIS	EWTOPSIS
Correlation coefficients	0.3355	0.5193 ***	0.4871 ***	0.5647 ***

Note: *** denotes significant at 1%.

4.3. Assessing the Dynamic Systemic Importance of Financial Institutions

As we all know, the financial markets are complex dynamic systems. The jump volatility spillover between financial institutions is time-varying. Thus, we employ time-varying Granger causality test to build dynamic jump volatility spillover of financial institution networks [63]. 2677 financial institution networks are obtained.

Figure 2 exhibits the evolution of the number of total linkages as a percentage of all possible linkages (TP). We can discover that it has three prominent cycles based on high TP values. The first cycle started at January 2008 and ended until April 2008, which was in the period of the subprime crisis. During this time, the CSI 300 Index decreased by 661 points (almost 17%) from 4620 to 3959. Then, TP value followed by a quickly descending trend until June 2008. The second cycle started at June 2008 and ended at October 2008, which was in the later period of the subprime crisis, and in the initial period of the European sovereign debt crisis. During this period, the CSI 300 Index dropped to 1948 points (almost 54%) from 3611 to 1663. Thereafter, the TP value entered a stationary period from October 2009 to June 2015. The Third cycle began at May 2015 and ended at August 2015, which was in the period of China's stock market disaster. At this stage, the CSI 300 Index declined by 1474 points (nearly 31%) from 4840 to 3366. Henceforth, the TP value entered another stationary period.

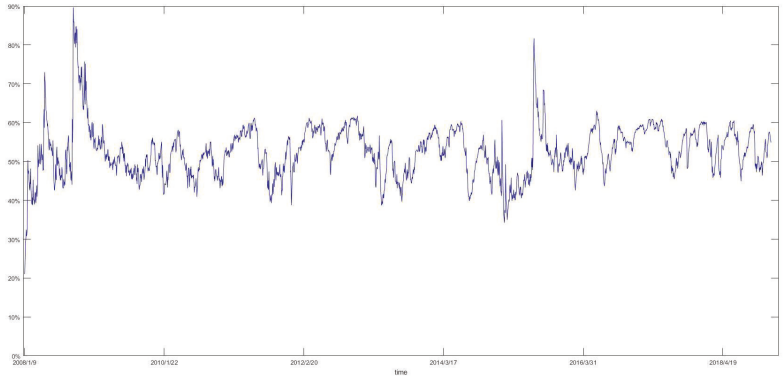


Figure 2. The number of total linkages of the jump volatility spillover network as a percentage of all possible linkages over time.

Furthermore, we employ the index of out degree, clustering coefficient, closeness centrality and leaderrank algorithm to measure the risk contagion range, risk contagion extent, risk contagion efficiency and risk contagion degree in each period, respectively. Then, entropy weight TOPSIS method is adopted to compute the score of SIFIs. We list the top 10 of financial institutions with systematically important score and market capitalization (MC) corresponding to the highest TP in three cycles, and the results are shown in Tables 5–7.

Table 5. Top 10 financial institutions ranked by the score of SIFIs on 25 April 2008.

Rank	Symbol	Score	Rank of MC
1	PSC	0.9087	18 (46,722,978,887)
2	CIB	0.8685	11 (199,800,000,000)
3	NBCB	0.8603	20 (36,175,000,000)
4	NESC	0.8553	24 (18,917,836,544)
5	HTSEC	0.8381	12 (176,610,181,629)
6	ICBC	0.8361	1 (2,167,782,336,669)
7	BOCOM	0.8336	5 (522,280,130,274)
8	PAB	0.8219	16 (64,100,729,703)
9	GYSC	0.7977	19 (42,005,029,000)
10	SLSC	0.7787	23 (22,425,428,420)

Note: This table lists the market capitalization (MC) and its corresponding rank of top 10 financial institutions.

Table 6. Top 10 financial institutions ranked by the score of SIFIs on 23 September 2008.

Rank	Symbol	Score	Rank of MC
1	CIB	0.9579	14 (79,900,000,000)
2	CMB	0.9230	7 (251,933,891,562)
3	PSC	0.9229	18 (27,435,468,619)
4	NBCB	0.9100	21 (18,750,000,000)
5	PAB	0.8718	16 (36,046,919,598)
6	ICBC	0.8123	1 (1,452,981,997,613)
7	CCB	0.8122	2 (1,088,991,131,440)
8	CITICS	0.8014	10 (133,603,922,140)
9	PAI	0.7996	6 (261,557,349,223)
10	BOCOM	0.7904	5 (292,496,470,707)

Table 7. Top 10 financial institutions ranked by the score of SIFIs on 10 July 2015.

Rank	Symbol	Score	Rank of MC
1	HTSEC	0.9081	14 (256,372,893,000)
2	PAB	0.8741	15 (212,626,927,426)
3	CJSC	0.8731	19 (62,553,148,673)
4	CIB	0.8451	10 (336,845,313,758)
5	PSC	0.8150	23 (40,635,675,469)
6	HXB	0.7961	16 (143,827,801,960)
7	CMB	0.7886	7 (478,924,867,963)
8	SPDB	0.7878	12 (322,331,986,051)
9	CITICS	0.7845	11 (335,880,700,848)
10	CLI	0.7522	4 (1,019,790,556,400)

Tables 5–7 display the top 10 financial institutions ranked by the systematically important score on 25 April 2008, 23 September 2008 and 10 July 2015. We can find that CIB, PAB and PSC are included in the three periods, indicating that large commercial banks and insurances play an important role during the financial crisis. Moreover, the top five largest financial institutions ranked by market capitalization are not all included in the table, and some small financial institutions are also systemically important. It means that the SIFIs in network may be "too big to fail" or "too interconnected to fail".

5. Conclusions and Discussion

This paper adopts 5-min high-frequency data of China's financial institutions to extract realized jump volatility. Then, we employ Granger-casuality test to construct the jump volatility spillover network. Furthermore, out degree, clustering coefficient, closeness centrality and leaderrank value are used to evaluate the SIFIs, respectively. In addition, we utilize entropy weight TOPSIS to comprehensively evaluate the SIFIs.

Some basic results of our research can be summed up as follows: (1) The highest frequency of jump volatility is 44.30% in 2008. This may be related to the outbreak of the subprime mortgage crisis in 2008. (2) We measure the SIFIs from four dimensions. In terms of risk contagion range, we can find that CMB, BOC and CNCB, which are all from depository sector, possess the highest out degree value of 10. One possible explanation is that China's financial system is a depository-led system. In terms of risk contagion extent, one can see that SLSC and HTSEC, which are all from broke-dealer sector, have the highest clustering coefficient value. This indicates that broke-dealer sector's risk contagion ability can not be neglected in China's financial system. In terms of risk contagion efficiency, we discover that CJSC and PAI have the highest closeness centrality value, which means that some insurances have gradually become important departments. In terms of risk contagion degree, the results reveal that the greatest degree of risk contagion is insurance sector, followed by broke-dealer sector, and depository sector have the worst risk contagion degree. (3) Based on the comprehensive evaluation of the SIFIs, by the method of entropy weight TOPSIS, the obtained results show that CLI, CCB, BOB, ICBC and BOCOM are identified as the influential nodes. (4) According to highest values of total linkages in each period, we can find three prominent cycles. The first cycle started at January 2008 and ended at April 2008, which was in the period of the subprime crisis. The second cycle started at June 2008 and ended at October 2008, which was in the later period of the subprime crisis, and in the initial period of the European sovereign debt crisis. The third cycle began at May 2015 and ended at August 2015, which was in the period of China's stock market disaster. (5) Total connectedness of financial institution networks reveal that large commercial banks and insurances play an important role when financial market is under pressure, especially during the subprime crisis, the European sovereign debt crisis and China's stock market disaster. Meanwhile, some small financial institutions are also systemic importance, which may be related to their too much interconnection with other financial institutions.

By the way, the work presented in this article does not consider the following points: (1) The data do not contain all publicly listed financial institutions in China because we have deleted those for

which we have experienced long suspension periods. Therefore, developing new tools that have limited sample for investigating SIFIs is a worthy target. (2) We just select 24 top financial institutions. Non-financial institutions may also play an important role as a result of their interactions with these financial ones. It would be interesting to research some financial institutions and non-financial institutions at the same time. We will leave this challenging yet interesting problem as future research. (3) It is important to measure SIFIs based on the network. This paper employs linear Granger-casuality test to construct the jump volatility spillover network. There is a nonlinear relationship between financial markets. The method of network construction could be extended by employing the Granger-casuality test. We have to leave these challenging yet interesting problems as future research.

Author Contributions: Conceptualization, C.H.; Data curation, C.H.; Formal analysis, X.Z., X.G. and X.Y. (Xiaoguang Yang); Funding acquisition, X.G.; Investigation, X.Y. (Xin Yang), X.Y. (Xiaoguang Yang) and C.H.; Methodology, X.Y. (Xin Yang) and X.Z.; Project administration, X.Y. (Xin Yang) and C.H.; Resources, X.G.; Software, X.Y. (Xin Yang) and X.G.; Supervision, X.Y. (Xiaoguang Yang); Validation, X.Z.; Visualization, X.Z. and X.G.; Writing—original draft, X.Y. (Xin Yang); Writing—review & editing, X.Y. (Xin Yang) and C.H. All authors have read and agreed to the published version of the manuscript.

Funding: Partially supported by the National Natural Science Foundation of P. R. China (Nos. 11971076, 51839002, 71701176), Hunan Provincial Natural Science Foundation (No. 2019JJ50650), Scientific Research Fund of Hunan Provincial Education Department (No. 18C0221),Research Promotion Program of Changsha University of Science and Technology (No. 2019QJCZ050).

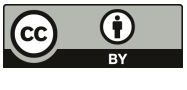
Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lentz, H.H.K.; Konschake, M.; Teske, K.; Kasper, M.; Rother, B.; Carmanns, R.; Petersen, B.; Conraths, F.J.; Selhorst, T. Trade communities and their spatial patterns in the German pork production network. *Prev. Vet. Med.* **2011**, *98*, 176–181. [CrossRef]
2. Bowden, R.J. Directional entropy and tail uncertainty, with applications to financial hazard. *Quant. Financ.* **2011**, *11*, 437–446. [CrossRef]
3. Dimpfl, T.; Peter, F.J. Using transfer entropy to measure information flows between financial markets. *Stud. Nonlinear Dyn. Econom.* **2013**, *17*, 85–102. [CrossRef]
4. Stutzer, M.J. The Role of Entropy in Estimating Financial Network Default Impact. *Entropy* **2018**, *20*, 369. [CrossRef]
5. Gencay, R.; Signori, D.; Xue, Y.; Yu, X.; Zhang, K.Y. Economic links and credit spreads. *J. Bank. Financ.* **2015**, *55*, 157–169. [CrossRef]
6. Li, C.; Wang, L.; Sun, S.W.; Xia, C.Y. Identification of influential spreaders based on classified neighbors in real-world complex networks. *Appl. Math. Comput.* **2018**, *320*, 512–523. [CrossRef]
7. Wen, S.; Tan, Y.; Li, M.; Deng, Y.; Huang, C. Analysis of global remittance based on complex networks. *Front. Phys.* **2020**, *8*, 85. [CrossRef]
8. Acharya, V.V.; Richardson, M. Causes of the financial crisis. *Crit. Rev.* **2009**, *21*, 195–210. [CrossRef]
9. Braouezec, Y.; Wagalath, L. Risk-Based Capital Requirements and Optimal Liquidation in a Stress Scenario. *Rev. Financ.* **2018**, *22*, 747–782. [CrossRef]
10. Braouezec, Y.; Wagalath, L. Strategic fire-sales and price-mediated contagion in the banking system. *Eur. J. Oper. Res.* **2019**, *274*, 1180–1197. [CrossRef]
11. De Haas, R.; Van Horen, N. International shock transmission after the Lehman Brothers collapse: Evidence from syndicated lending. *Am. Econ. Rev.* **2012**, *102*, 231–237. [CrossRef]
12. Gradojevic, N.; Caric, M. Predicting Systemic Risk with Entropic Indicators. *J. Forecast.* **2017**, *36*, 16–25. [CrossRef]
13. Erdemlioglu, D.; Gradojevic, N. Heterogeneous investment horizons, risk regimes, and realized jumps. *Int. J. Financ. Econ.* **2020**, in press. [CrossRef]
14. Yang, X.; Wen, S.G.; Liu, Z.F.; Li, C.; Huang, C.X. Dynamic Properties of Foreign Exchange Complex Network. *Mathematics* **2019**, *7*, 832. [CrossRef]

15. Yang, X.; Wen, S.G.; Zhao, X.; Huang, C.X. Systemic importance of financial institutions: A complex network perspective. *Phys. A Stat. Mech. Appl.* **2020**, *545*, 123448. [CrossRef]
16. Huang, C.X.; Wen, S.G.; Li, M.G.; Wen, F.H.; Yang, X. An empirical evaluation of the influential nodes for stock market network: Chinese A-shares case. *Financ. Res. Lett.* **2020**, 101517. [CrossRef]
17. Huang, X.; Zhou, H.; Zhu, H.B. A framework for assessing the systemic risk of major financial institutions. *J. Bank. Financ.* **2009**, *33*, 2036–2049. [CrossRef]
18. Valahzaghard, M.; Bahrami, M. Prediction of default probability in banking industry using CAMELS index: A case study of Iranian banks. *Manag. Sci. Lett.* **2013**, *3*, 1113–1118. [CrossRef]
19. Derbali, A.; Jamel, L. Dependence of Default Probability and Recovery Rate in Structural Credit Risk Models: Case of Greek Banks. *J. Knowl. Econ.* **2018**, *10*, 711–733. [CrossRef]
20. Wen, F.H.; Yang, X.; Zhou, W.X. Tail dependence networks of global stock markets. *Int. J. Financ. Econ.* **2019**, *24*, 558–567. [CrossRef]
21. Girardi, G.; Ergün, A.T. Systemic risk measurement: Multivariate GARCH estimation of CoVaR. *J. Bank. Financ.* **2013**, *37*, 3169–3180. [CrossRef]
22. Acharya, V.V.; Pedersen, L.H.; Philippon, T.; Richardson, M. Measuring systemic risk. *Rev. Financ. Stud.* **2017**, *30*, 2–47. [CrossRef]
23. Wang, J.N.; Hsu, Y.T.; Lee, J.M.; Chen, C.C. Measuring Systemic Risk: Capital Shortfall and CSRISK. *Int. Rev. Financ.* **2019**, 1–11. [CrossRef]
24. Hautsch, N.; Schaumburg, J.; Schienle, M. Financial network systemic risk contributions. *Rev. Financ.* **2014**, *19*, 685–738. [CrossRef]
25. Diebold, F.X.; Yilmaz, K. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *J. Econom.* **2014**, *182*, 119–134. [CrossRef]
26. Levy-Carciente, S.; Kenett, D.Y.; Avakian, A.; Stanley, H.E.; Havlin, S. Dynamical macroprudential stress testing using network theory. *J. Bank. Financ.* **2015**, *59*, 164–181. [CrossRef]
27. Dai, Z.F.; Zhou, H.T.; Wen, F.H.; He, S.Y. Efficient predictability of stock return volatility: The role of stock market implied volatility. *N. Am. J. Econ. Financ.* **2020**,*52*, 101174. [CrossRef]
28. Dai, Z.F.; Zhu, H. Stock return predictability from a mixed model perspective. *Pac. Basin Financ. J.* **2020**, *60*, 101267. [CrossRef]
29. Dai, Z.F.; Dong, X.D.; Kang, J.; Hong, L.Y. Forecasting stock market returns: New technical indicators and two-step economic constraint method. *N. Am. J. Econ. Financ.* **2020**, *53*, 101216. [CrossRef]
30. Billio, M.; Getmansky, M.; Lo, A.W.; Pelizzon, L. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *J. Financ. Econ.* **2012**, *104*, 535–559. [CrossRef]
31. Gong, X.L.; Liu, X.H.; Xiong, X.; Zhang, W. Financial systemic risk measurement based on causal network connectedness analysis. *Int. Rev. Econ. Financ.* **2019**, *64*, 290–307. [CrossRef]
32. Härdle, W.K.; Wang, W.N.; Yu, L.N. TENET: Tail-Event driven NETwork risk. *J. Econom.* **2016**, *192*, 499–513. [CrossRef]
33. Wang, G.J.; Jiang, Z.Q.; Lin, M.; Xie, C.; Stanley, H.E. Interconnectedness and systemic risk of China's financial institutions. *Emerg. Mark. Rev.* **2018**, *35*, 1–18. [CrossRef]
34. Li, P.; Wang, B.H. Extracting hidden fluctuation patterns of Hang Seng stock index from network topologies. *Phys. A Stat. Mech. Appl.* **2007**, *378*, 519–526. [CrossRef]
35. Zhang, Q.; Li, M.Z.; Deng, Y. Measure the structure similarity of nodes in complex networks based on relative entropy. *Phys. A Stat. Mech. Appl.* **2018**, *491*, 749–763. [CrossRef]
36. Sun, Q.R.; Gao, X.Y.; Wen, S.B.; Chen, Z.H.; Hao, X.Q. The transmission of fluctuation among price indices based on Granger causality network. *Phys. A Stat. Mech. Appl.* **2018**, *506*, 36–49. [CrossRef]
37. Ji, Q.; Bouri, E.; Roubaud, D. Dynamic network of implied volatility transmission among US equities, strategic commodities, and BRICS equities. *Int. Rev. Financ. Anal.* **2018**, *57*, 1–12. [CrossRef]
38. Deng, X.R.; Cao, H. Systematic Risk, network contagion and sssessment of SIFIs. *J. Cent. Univ. Financ. Econ.* **2016**, *3*, 52–60. (In Chinese)
39. Li, X.X.; Wang, K.S.; Liu, L.W.; Xin, J.; Yang, H.R.; Gao, C.Y. Application of the entropy weight and TOPSIS method in safety evaluation of coal mines. *Procedia Eng.* **2011**, *26*, 2085–2091. [CrossRef]
40. Onar, S.C.; Oztaysi, B.; Kahraman, C. Strategic Decision selection using hesitant fuzzy TOPSIS and interval type-2 fuzzy AHP: A case study. *Int. J. Comput. Intell. Syst.* **2014**, *7*, 1002–1021. [CrossRef]

41. Kim, A.R. A Study on competitiveness analysis of ports in Korea and China by entropy weight TOPSIS. *Asian J. Shipp. Logist.* **2016**, *32*, 187–194. [CrossRef]
42. Yang, R.J.; Yu, L.; Zhao, Y.J.; Yu, H.X.; Xu, G.P.; Wu, Y.T.; Liu, Z.K. Big data analytics for financial Market volatility forecast based on support vector machine. *Int. J. Inf. Manag.* **2020**, *50*, 452–462. [CrossRef]
43. Kitamura, Y.; Stutzer, M.J. Connections Between Entropic and Linear Projections in Asset Pricing Estimation. *J. Econom.* **2002**, *107*, 159–174. [CrossRef]
44. Liu, A.Q.; Paddrik, M.E.; Yang, S.Y.; Zhang, X.J. Interbank Contagion: An Agent-Based Model Approach to Endogenously Formed Networks. *J. Bank. Financ.* **2017**, *112*, 105191. [CrossRef]
45. Gong, X.; Wen, F.H.; Xia, X.H.; Huang, J.B.; Pan, B. Investigating the risk-return trade-off for crude oil futures using high-frequency data. *Appl. Energy* **2017**, *196*, 152–161. [CrossRef]
46. Gong, X.; Lin, B.Q. The incremental information content of investor fear gauge for volatility forecasting in the crude oil futures market. *Energy Econ.* **2018**, *74*, 370–386. [CrossRef]
47. Zhao, S. Spatial restructuring of financial centers in mainland China and Hong Kong: A geography of finance perspective. *Urban Aff. Rev.* **2003**, *38*, 535–571. [CrossRef]
48. Bajgrowicz, P.; Scaillet, O.; Treccani, A. Jumps in high-frequency data: Spurious detections, dynamics, and news. *Manag. Sci.* **2016**, *62*, 2198–2217. [CrossRef]
49. Wright, J.H.; Zhou, H. Bond Risk Premia and Realized Jump Volatility. *Soc. Sci. Res. Netw.* **2007**, *2007*, 1–37. [CrossRef]
50. Zhang, X.; Kim, D.; Wang, Y.Z. Jump Variation Estimation with Noisy High Frequency Financial Data via Wavelets. *Econometrics* **2016**, *4*, 34. [CrossRef]
51. Audrino, F.; Hu, Y.J. Volatility Forecasting: Downside Risk, Jumps and Leverage Effect. *Econometrics* **2016**, *4*, 8. [CrossRef]
52. Barndorff-Nielsen, O.E.; Shephard, N. Realized power variation and stochastic volatility models. *Bernoulli* **2003**, *9*, 243–265. [CrossRef]
53. Andersen, T.G.; Dobrev, D.; Schaumburg, E. Jump-robust volatility estimation using nearest neighbor truncation. *J. Econom.* **2012**, *169*, 75–93. [CrossRef]
54. Barndorff-Nielsen, O.E.; Shephard, N. Power and bipower variation with stochastic volatility and jumps. *J. Financ. Econom.* **2004**, *2*, 1–37. [CrossRef]
55. Wen, F.H.; Zhao, Y.P.; Zhang, M.Z.; Hu, C.Y. Forecasting realized volatility of crude oil futures with equity market uncertainty. *Appl. Econ.* **2019**, *51*, 6411–6427. [CrossRef]
56. Andersen, T.G.; Bollerslev, T.; Diebold, F.X. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Rev. Econ. Stat.* **2007**, *89*, 701–720. [CrossRef]
57. Zheng, Q.H.; Song, L.R. Dynamic Contagion of Systemic Risks on Global Main Equity Markets Based on Granger Causality Networks. *Discret. Dyn. Nat. Soc.* **2018**, *2018*, 1–13. [CrossRef]
58. Jiang, M.; Gao, X.; An, H.; Li, H.; Sun, B. Reconstructing complex network for characterizing the time-varying causality evolution behavior of multivariate time series. *Sci. Rep.* **2017**, *7*, 10486. [CrossRef]
59. Haugom, E.; Langeland, H.; Molnár, P.; Westgaard, S. Forecasting volatility of the US oil market. *J. Bank. Financ.* **2014**, *47*, 1–14. [CrossRef]
60. Du, Y.X.; Gao, C.; Hu, Y.; Mahadevan, S.; Deng, Y. A new method of identifying influential nodes in complex networks based on TOPSIS. *Phys. A Stat. Mech. Appl.* **2014**, *399*, 57–69. [CrossRef]
61. Sandoval, L. Structure of a global network of financial companies based on transfer entropy. *Entropy* **2014**, *16*, 4443–4482. [CrossRef]
62. Wang, Z.; Gao, X.Y.; An, H.Z.; Tang, R.W.; Sun, Q.R. Identifying influential energy stocks based on spillover network. *Int. Rev. Financ. Anal.* **2019**. [CrossRef]
63. Lu, F.B.; Hong, Y.M.; Wang, S.Y.; Lai, K.K.; Liu, J. Time-varying Granger causality tests for applications in global crude oil markets. *Energy Econ.* **2014**, *42*, 289–298. [CrossRef]



entropy

MDPI

Article

A Simple Mechanism Causing Wealth Concentration

Michał Cieśla [1] and Małgorzata Snarska [2,*]

[1] M. Smoluchowski Institute of Physics, Jagiellonian University, Łojasiewicza 11, 30-348 Kraków, Poland; michal.ciesla@uj.edu.pl

[2] Department of Financial Markets, Cracow University of Economics, Rakowicka 27, 31-510 Kraków, Poland

* Correspondence: malgorzata.snarska@uek.krakow.pl

Received: 1 September 2020; Accepted: 8 October 2020; Published: 13 October 2020

Abstract: We study mechanisms leading to wealth condensation. As a natural starting point, our model adopts a neoclassical point of view, i.e., we completely ignore work, production, and productive relations, and focus only on bilateral link between two randomly selected agents. We propose a simple matching process with deterministic trading rules and random selection of trading agents. Furthermore, we also neglect the internal characteristic of traded goods and analyse only the relative wealth changes of each agent. This is often the case in financial markets, where a traded good is money itself in various forms and various maturities. We assume that agents trade according to the rules of utility and decision theories. Agents possess incomplete knowledge about market conditions, but the market is in equilibrium. We show that these relatively frugal assumptions naturally lead to a wealth condensation. Moreover, we discuss the role of wealth redistribution in such a model.

Keywords: wealth condensation; agent-based computational economics; bargaining; gain function

1. Introduction

Study of wealth distribution among the population has been labelled as one of the key problems in modern economic theory and is often described by a power-law function known as Pareto distribution [1]. In this sense, research related to wealth distribution and wealth inequalities is two-fold. The well-studied macro-perspective focuses on the issue of poverty arising from wealth inequalities, its social and economic consequences, where it is typical that a small fraction of the population owns most of the total wealth. This approach stems from macroeconomic theory and general equilibrium (c.f. [2]) like the infinite-lived dynasty model [3] and overlapping generations model [4]. Other concepts refer to asymmetric knowledge [5], a different number of connections or opportunities to exchange or increase wealth [6] or only to luck or different competencies [7]. Most of these models usually rely on representative agent paradigm, while completely ignoring immanent aspects of human nature and psychological biases or even microstructural characteristics of trade mechanism [8]. On the other hand, micro-foundations of wealth concentration arise from bilateral trade or exchange of goods among two economic agents, where wealth typically is highly related to the individual investment decision process. This observation led to several mathematical models attempting to explain this phenomenon, i.e., so-called kinetic wealth exchange models that are based on microeconomic interactions between economic actors who exchange wealth between them over the trade cycle [9]. These include models introduced by Angle [10], Bennati [11], Chakraborti and Chakrabarti [12], Dragulescu and Yakovenko [13] and recently also the approaches by Vallejos et al. [14] and Lim and Min [15], which share some common features with our approach.

1.1. Some Stylised Facts Related to Empirical Wealth Distribution

A well-known fact about wealth distribution in developed-economy states is that wealth is highly concentrated and very unequally distributed. Data sets gathered over 30 years by Census Survey of Consumer Finances confirm via, e.g., observation of historical trends that a degree of wealth concentration in the United States is high, i.e., almost one-third of total wealth is kept by only 1% of households, while the top 5% of the population holds more than one half of total wealth. At the other edge, there is a large fraction of the community, who has pretty little wealth or no wealth at all. These results are quite persistent over time, and substantial changes in net wealth are subject to a boom-bust cycle of financial and economic crises [16]. Little work has, however, been done in the area of understanding mechanisms leading to wealth concentration during economic upturns and equalisation effects during recessions. We contribute to this field by introducing new kinetic wealth exchange model with simplified assumptions, that can reproduce stylised facts observed in the empirical distribution of wealth in the crisis and post-crisis times. As a starting point, we use a real-life example of wealth distributions in the U.S. in the years 2010 (crisis) and 2018 (post-crisis). We believe that our model can shed light into better recognition of patterns leading to changes in wealth inequalities during the real business cycle.

Although in economics income is typically defined as the amount of money an economic agent or household receives on regular basis and wealth is related to the length of time that a family could maintain their current lifestyle without receiving compensation for performing additional work, we treat these two categories as a whole. The primary source of our data on wealth in the U.S. for our empirical examples is the U.S. Bureau of Census and Bureau of Labour Statistics Current Population Survey for Household Income from the years 2010 and 2018. The survey has been conducted monthly for over 50 years, with over 54,000 households selected based on an area of residence to represent the nation as a whole, individual states, and other specified areas. Each family is interviewed once a month for 4 consecutive months one year, and again for the corresponding period a year later. These data are however available only in a binned or aggregated form, so the only available data include the number of households in each bin, mean income, standard error and income limits assigned to each bin. To estimate income or wealth probability density function, we use the entropy-based divergence method and seek a probability density function, that is as close to the uniform distribution as the data sample will permit [17,18] (see the Appendix A). Results of estimation are presented in Figure 1. In Table 1, we have also gathered Gini coefficients and information criteria, for various distributions from the Creesse and Read (C.R) family [19].

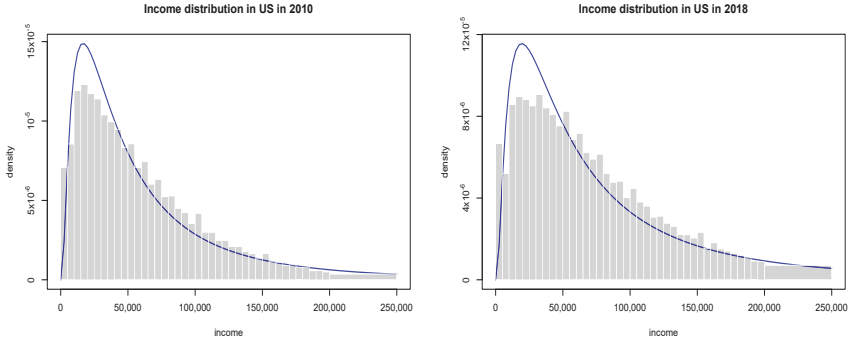


Figure 1. The empirical distribution of wealth in the U.S. based on Census Data. The blue line represents one of fitted generalised gamma distribution of the second kind.

Table 1. Information criteria for fitted distributions.

Distribution	2010			2018		
	Gini	Aic	Bic	Gini	Aic	Bic
Generalised Beta of the Second Kind	0.4504	825,368.9827	825,407.7613	0.4656	914,590.8355	914,629.8928
Generalised Gamma	0.4485	825,598.8432	825,627.9271	0.4526	915,341.0957	915,370.3886
Beta of the Second Kind	0.4545	825,501.8130	825,530.8969	0.4636	915,233.3365	915,262.6295
Dagum		1,258,915.0143	1,258,944.0982	0.4693	914,642.1348	914,671.4277
Singmad	0.4531	827,239.3961	827,268.4800	0.4600	914,833.1633	914,862.4562
Lognormal	0.5013	832,408.2444	832,427.6337	0.5206	924,094.3485	924,113.8772
Weibull	0.4432	827,065.1604	827,084.5496	0.4462	916,179.0877	916,198.6163
Gamma	0.4409	826,112.8345	826,132.2238	0.4467	915,559.4152	915,578.9439
Doubly lognormal		1,375,275.6949	1,375,295.0841		920,090.4281	920,109.9568
Pareto	0.5047	832,191.8408	832,211.2301	0.5061	920,845.1144	920,864.6431

Results presented in Tables 1 and 2 confirm that empirical distributions of wealth exhibit a Pareto power-law tail

$$f(x) \sim \frac{1}{x^{1+\alpha}} \qquad 1 < \alpha < 2, \tag{1}$$

and the actual shape of distribution at intermediate values of wealth is well fitted by a generalised gamma distribution of the second kind. So they can be reproduced by a simple kinetic wealth exchange model with either homogeneous or heterogeneous agents [9]. Furthermore, as seen in Figure 1, the post-crisis inequalities are larger than the crisis ones.

Table 2. Estimated distribution parameters: location μ, scale σ, skewness ν and kurtosis τ.

Distribution	2010				2018			
	μ	σ	ν	τ	μ	σ	ν	τ
Generalised Beta of the Second Kind	108,564.1708	1.7786	0.7034	2.0083	113,253.3847	2.1917	0.5323	1.2229
Generalised Gamma	60,663.6500	0.9001	0.7612		81,899.1321	0.9057	0.8421	
Beta of the Second Kind	283,044.3149	1.5608	7.5992		372,137.5280	1.4877	7.2873	
Dagum	1,012,451,669.9591	0.9721	0.1021		105,486.1708	2.4436	0.4689	
Singmad	1,012,451,669.9591	1.1348	53,697.3940		190,335.2758	1.3413	3.3841	
Lognormal	10.6958	0.9900			10.9507	1.0373		
Weibull	69,527.7650	1.1699			89,843.3170	1.1589		
Gamma	6,5806.3882	0.8652			85,510.9515	0.8789		
Doubly lognormal	1,012,451,669.9591	0.1514			60,235.5085	1.7064		
Pareto		1,549,526.7678	24.2996					

1.2. The General Structure of Kinetic Exchange Models

In this section, we will briefly review the basic features of kinetic wealth exchange models following [9,14,15]. As usual, the economy is assumed to consist of N agents with wealth $\{a_k \geq 0\}$ $(k = 1, 2, \ldots, N)$. At each cycle, an agent i exchanges a quantity Δa of wealth with another agent j. Both agents are chosen randomly. The total wealth $X = \sum_i a_i$ and the average wealth $\langle a \rangle = X/N$ are constant. After the wealth exchange, a_i and a_j are updated according to the rule:

$$\begin{aligned} a_i' &= a_i - \Delta a\,, \\ a_j' &= a_j + \Delta a\,, \end{aligned} \tag{2}$$

under the condition $(a_i', a_j' \geq 0)$. The signs have been chosen without the loss of generality and the function $\Delta a = \Delta a(a_i, a_j)$ is responsible for the dynamics of the underlying wealth concentration mechanism. Furthermore, agents can be parametrised by a maximum fraction of wealth $\omega \in (0, 1]$ that enters each cycle or exchange process, which determine the time scale of the relaxation process and the mean value $\langle a \rangle$ at equilibrium [9]. If the value of ω is identical for all agents, then models belong to a homogeneous class that can reproduce the shape of the gamma wealth distribution. For $\omega < 1$, models converge toward a stable state with a wealth distribution with non-zero median, and for diversified agents, a power-law tail behaviour can be recovered. If ω_k is different for every agent, then models are called heterogeneous.

In the Angle model [10], changes of wealth between agents are determined by

$$\Delta a = r\,\omega\,[\eta_{ij}\,a_i - (1-\eta_{ij})\,a_j]\,, \tag{3}$$

where random variable $r \in (0,1)$ is distributed uniformly or with a certain probability distribution $g(r)$, and η_{ij} is a random dichotomous variable responsible for the direction of the changes. The value $\eta_{ij} = 1$ produces a wealth transfer $|\Delta a| = r\,\omega\,a_i$ from agent i to j, while the value $\eta_{ij} = 0$ corresponds to a wealth transfer $|\Delta a| = r\,\omega\,a_j$ from j to i.

Another model is a One-Parameter Inequality Process model [20]

$$\Delta a = -\eta_{ij}\omega a_i + (1-\eta_{ij})\omega a_j\,, \tag{4}$$

where $\eta_{ij} = 0$ or $\eta_{ij} = 1$ are chosen randomly in each cycle. In these models, wealth distribution is best described by a gamma distribution.

Bennati [11] proposed a model, where agent exchange constant amount of wealth Δa_0 and transaction between agent i and j take place if and only if $a_i', a_j' \geq 0$, which leads to an exponential distribution of wealth.

Chakraborti and Chakrabarti [12] introduced a model, where new wealth a_i' (a_j') is expressed as a sum of the saved fraction $\lambda a_i'$ ($\lambda a_j'$) of the initial wealth and a random fraction r ($\bar{r}$) of the total remaining wealth, obtained summing the respective contributions of agents i and j.

$$\Delta a = \omega(\bar{r}\,a_i - r\,a_j) = (1-\lambda)(\bar{r}\,a_i - r\,a_j)\,. \tag{5}$$

Dragulescu and Yakovenko introduced another model [13] with dynamics described as:

$$\Delta a = \bar{r}\,a_i - r\,a_j\,, \tag{6}$$

leading to an exponential model for wealth distribution.

Lim and Min [15] consider various kinetic exchange models in search for solidarity effects consider multiple variants of the model introducing heterogeneous savings parameter λ_k and wealth dependent trading rules

$$\Delta a = r\min(a_i, a_j) \tag{7}$$

with r being random variable uniformly distributed over $[0,1]$ and updated every transaction.

The common factor of all these models is the wealth conservation: $a_i + a_j = a_i' + a_j'$, which means that while one agent gains money from a transaction, the other one has to lose some wealth. Therefore, without any preference of richer agents over the poorer ones and due to random character of interactions between agents, these kinetic exchange models can be characterised by a stationary wealth distributions, which exhibit exponential tails. After reaching this distribution wealth, inequalities do not increase anymore. It should be noted that there are also kinetics exchange models that exhibit power-law wealth distribution as a stationary state [21,22]. These models often assume some preferences of richer individuals; for example, individuals' wealth is repeatedly multiplied by a random factor, different for each individual.

Another type of agent model is one with a growing economy, where wealth is continuously added to the system and divided among the agents. One such model has been used recently by Vallejos et al. to study the growth of wealth inequalities in the U.S. [14]. There the assumption is, however, that individuals with greater wealth get significantly more of this added wealth than poorer agents. Namely, this wealth is divided into several equal parts, and each part is given to the agent i with probability proportional to a_i^β. In this setup, Authors studied how the initial Pareto like wealth distribution changes depending on β. For $0 < \beta < 1$, the power of wealthier individuals is diminished much more than the power of poorer individuals and wealth inequalities lowered over time. For $\beta = 1$,

the model gives all individuals a proportional amount of power, and thus initial wealth distribution does not change. For $\beta > 1$, the model provides a disproportionate amount of power to the wealthier individuals in the market, and the wealth inequalities grew.

To sum up, the discussed agent models reaches a stationary state when wealth inequalities do not increase any more or exhibit constant grow of wealth concentration under the assumption that richer agents are disproportionately better treated than the poorer ones. In this study, we propose a model of trading agents where inequalities grow incessantly, but the gain of each agent is at average proportional to its wealth, which corresponds to the case of $\beta = 1$ in Vallejos et al. model [14], where inequalities do not change.

2. Model

In our model, an agent is a participant of a market game. An agent can be a company, an institution as well as a single person. Agents can trade with themselves using their assets. It is worth noting that for our study "assets" are not only goods or money that an agent has, but also widely understood services he can make or even its skills that allows him to be more effective in a market. Agents can interact with themselves, which affects their assets. In general, such interaction covers all possible activities like, for example, buying or selling products, services, financial market instruments, etc., as well as making money at work or employee hiring. From our perspective, all activities mentioned above are indistinguishable, so we will call each of them using the same word—a trade. As mentioned, a simple example of trade is the buying of a product in a shop. Another example is hiring an employee by a company. Trade is also when one agent exchanges his knowledge with another one.

In a further study, we have considered two assumptions:

(i) agents are equal in the sense that each of them has the same access to the market and the same knowledge about it.
(ii) agent trade only when it is profitable from their perspective.

These assumptions are quite general. The first one reflects capabilities given by a modern technology where at least virtual access to goods, financial markets and stock-exchanges is common. Therefore the number of transaction that agent can make is limited only by the number of his assets. The second assumption corresponds to decision and utility theories, which tell us that action will be undertaken by the individual only when it causes maximisation of the individual utility [23,24]. It is worth noting that typically utility and assets are not the same quantities. They are not even measured in the same units. However, we think that they can be somehow compared to the money, which measures assets, spent to increase utility. Therefore there is a relation between these two concepts, and in further considerations, we will treat the utility as an asset.

The society consists of $N = 10^5$ individuals (agents). Each of them possesses some assets. Let a_i denotes a share of i-th agent assets in the whole population wealth. Thus

$$\sum_{i=1}^{N} a_i = 1. \tag{8}$$

Agents interact with themselves, which affects their wealth. As stated before a trade is a win-win situation, i.e., both trading agents gain a profit from it. Because a market is in equilibrium and all agents have comparable information about it, their profits from a single trade should also be comparable. Here, we assume that profits from trade are equal for both agents and are given by a deterministic gain function $g(i,j) = g(j,i)$, where i and j are the trading agents. Thus, the trade changes the trading agent's assets as follows:

$$\begin{aligned} a_i &\to a_i + g(i,j) \\ a_j &\to a_j + g(j,i) \end{aligned} \tag{9}$$

The above, fully deterministic, rules reflect the second condition made in the Introduction section. The first condition about equal access to the market is fulfilled by a specific matching process between

two bargaining agents. They are selected according to their wealth. Thus the probability that i-th agent will be chosen for trading is equal to a_i. It reflects the fact that wealthier agents have more opportunities to trade, but on the other hand, it does not exclude poorer ones from the market.

Trades are grouped into cycles. A single cycle contains $N/2$ trades. The protocol used for choosing agents involving in trade is as follows

1. The first agent i is chosen randomly with the probability equal to its wealth a_i.
2. The second agent j is chosen randomly with the probability equal to its wealth a_j.
3. If $i = j$ or agent i has traded with agent j in this cycle, go to point 1. Otherwise, make the trade.

Each pair of agents can trade at most once during a single cycle. This prevents a situation that all the trades are only between the richest individuals, and thus, increase the chance of gaining profits by poorer agents. (Without this assumption, two richest agents may perform all the trades. However even with this restriction, it may occur that $\sqrt{N}$ number of agents will concentrate all the wealth and they will trade with themselves only, but it is not possible to limit further the number of trading agents.) After each cycle, individuals portfolios of assets are normalised to fulfil condition (8)

$$a_i \to \frac{a_i + \sum_j g(i,j)}{1 + \sum_{i,j} g(i,j)}, \tag{10}$$

where $\sum_j$ sums all profits made by the i-th agent, and $\sum_{i,j}$ sums all profits from all transactions within a cycle. To fully specify the model, a particular gain function $g(i,j)$ is needed. The simplest, symmetrical functions of two variables is a constant function: $g(i,j) = r/N$, i.e., each agent receives a lump sum of money during a single trade. Note that the number of trades depends on agents wealth—richer individuals trades more because they split their wealth into a larger number of transaction. Thus, each transaction in the model involves the same amount of assets, and therefore, a constant payoff is justified. Although it might seem that such a mechanism is similar to preferential treatment of some agents as in other models e.g., [25], we will show in the following section that it is not in the case of this model. The payoff r was typically set to 0.1. Note that r is equal to the global income from all trades within a cycle when the global wealth is equal to 1. Therefore the specific value of r determines the speed of wealth distribution changes between cycles.

3. Results and Discussion

3.1. Wealth Condensation

The model was tested numerically. The evolution of wealth distribution is presented in Figure 2 (See the software used for simulations and data analysis in Supplementary).

The plots differ in initial wealth distribution among agents. Here, we used the following distributions.

(a) delta distribution—all agents started with the same amount of money;
(b) uniform distribution—the initial wealth of each agent was uniformly distributed on the interval $[0, 2/N)$;
(c) exponential distribution—the initial wealth was drawn according to the exponential distribution of the unit mean and variance;
(d) Gaussian distribution—the initial wealth of each agent was an absolute value of a number drawn according to the normal distribution of the zero mean value and unit variance;
(e) Cauchy distribution—the initial wealth of each agent was an absolute value of a number drawn according to the following probability distribution function

$$p(x) = \frac{1}{\pi(x^2 + 1)}; \tag{11}$$

(f) 1% of richest agents possessed 100 times more money than the remaining 99% of poorer agents.

After choosing initial wealth distribution, the assets of each agent were normalised to fulfil condition (8).

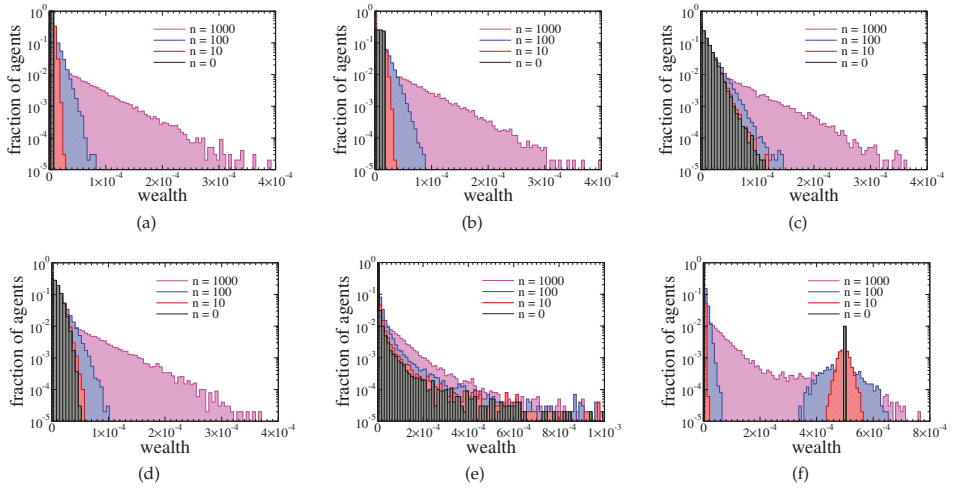


Figure 2. Histograms of agents assets after 0, 10, 100, and 1000 cycles. Different plots correspond to the different initial distribution of wealth among the agents: a_i is (**a**) equal to $1/N$, (**b**) uniformly distributed in the interval [0, 2/N), (**c**) exponentially distributed, (**d**) normally distributed, (**e**) Cauchy distributed, (**f**) 1% of richest agents possess 100 times more than the rest 99% of agents.

In most of these cases, the final wealth distribution (after 1000 cycles) was almost the same. Only in the last case we did not obtain purely exponential distribution after 1000 cycles, but here, the wealth distribution also should converge to an exponential distribution with a growing number of cycles. These results suggest that the wealth condensation was a feature of the model and not of the specific initial wealth distribution among agents. The effect of wealth condensation was confirmed by analysis of the Gini coefficient:

$$G = \frac{2\sum_{i=1}^{N} i a_i}{N\sum_{i=1}^{N} a_i} - \frac{N+1}{N}, \tag{12}$$

as well as the wealth of the richest and the middle agent (see Figure 3).

For most of the studied cases, the Gini index grew monotonically with the evolution of the system. The only exception was when initial wealth was drawn according to Cauchy distribution, which is an example of power-law, long-tail distribution. Here, the Gini index started from a relatively high value, as the condensation was an intrinsic property of power-laws. However, after the initial decrease corresponding to recombination to exponential distribution preferred by the model, it started to grow again. It is worth noting that the richest agent in this scenario lost most of its initial assets, but a rapid decrease of the median asset in the population (see Figure 3 inset) shows that wealth condensation occurred anyway. It should be stressed that the final state of these simulations was not stable in either case. The wealth inequalities seemed to grow endlessly. After 1000 cycles the Gini coefficient was above 0.89, and the richest agent owned approximately 5.3×10^{-4} of the total wealth and the median wealth several orders of magnitude smaller.

Next, we check if the population size affected condensation. It was done by studying systems consisting of $N = 10^4$ up to 10^7 agents—see Figure 4.

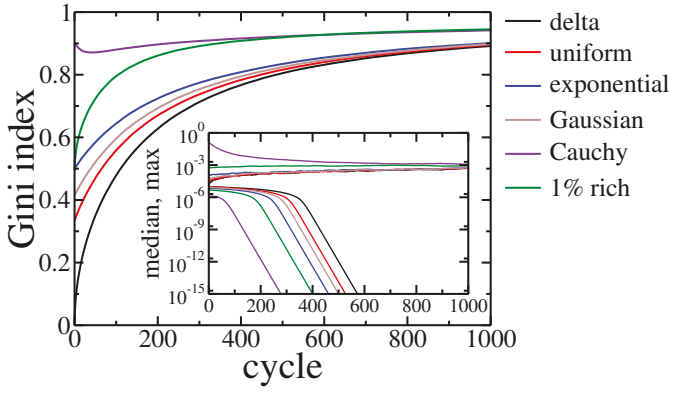


Figure 3. Gini index evolution for all studied agents initial wealth distributions. Inset shows the evolution of the maximal and median wealth in the population of agents. Black line corresponds to equal initial wealth of all agents, red—the uniform distribution of wealth on the interval $[0, 2/N)$, blue—the exponential wealth distribution, brown—Gaussian distribution, violet—Cauchy distribution, green—1% of richest agents have 100 bigger assets than the rest 99% of the population.

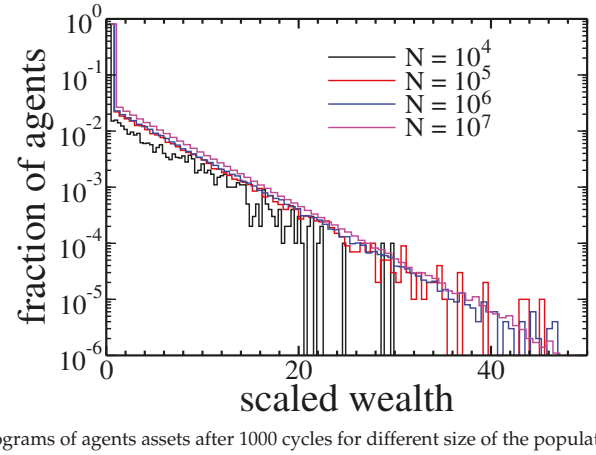


Figure 4. Histograms of agents assets after 1000 cycles for different size of the population. Because the total wealth changes with population size, we rescaled agents wealth by multiplying it by the number of agents N.

After 1000 cycles all populations exhibited a similar, exponential distribution of wealth.

At last we checked if the appearance of wealth condensation does not depend on model details. Therefore, instead the $g(i,j) = r/N$ we studied numerically other examples of payoff functions. In particular we performed independent analysis of the model using the following different utility functions, namely:

(i) linear preferences utility function: $g(i,j) = r\frac{a_i + a_j}{2}$—gains from individual transaction depend on assets of both sides of trade process.
(ii) Cobb–Douglas utility function: $g(i,j) = r\sqrt{a_i a_j}$—similar as in the above case but gains were much lower when agents assets differed significantly.
(iii) Koopmans and Leontieff utility function: $g(i,j) = r\min(a_i, a_j)$—gains are determined by a poorer trader.

Results are presented in Figure 5.

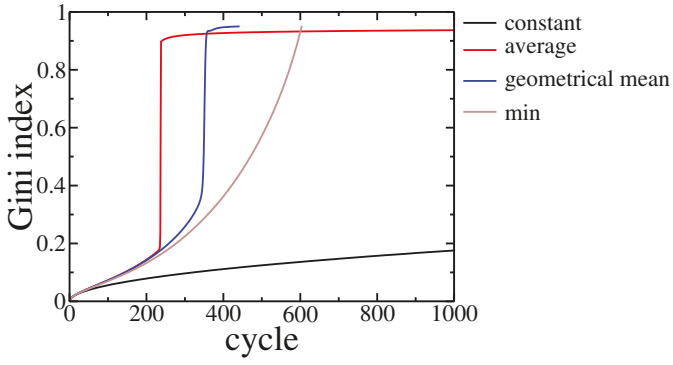


Figure 5. Gini index evolution for different payoff function used in the model. The black line corresponds to constant payoffs while the other ones correspond to payoffs depending on trading agents assets: red—payoff proportional to average assets of trading agents, blue—payoff proportional to geometrical mean of trading agents assets, brown—payoff proportional to poorer agent asset. The parameter $r = 0.01$ and population size $N = 10^5$.

In all studied cases, we observed wealth condensation. It occurred even faster than for the constant gain, as the above functions give additional profits from transactions between richer agents, which are more probable within the studied model. Moreover, for Cobb–Douglas and Koopmans–Leontieff functions, a small group of agents took all the assets, and thus, further evolution, according to model rules, became impossible. The model rules assume that in a single cycle there were $N/2$ trades between different pairs of agents. However, these agents were chosen with probability given by their assets, and thus there was practically no opportunity to draw randomly one of the poor agents as their wealth was negligibly small. Even if such trade occurred the assets of the poorer did not change significantly due to properties of these utility functions.

Because the wealth condensation occurred for all studied cases of the initial wealth distribution, we further focused on the case where all agents had an equal lump of money initially, and the population size was 10^5. To get some insight into obtained results, we analysed a simpler model, where an agent could trade with himself, and there was no restriction that each pair could trade at most once during a cycle. Note that these rules were more generous for a richer agent than the ones used in numerical simulations. We checked numerically that these restrictions had no qualitative influence on the phenomenon of wealth concentration as well as the type of wealth distribution. In such a case, each agent had N independent opportunities to trade, and each option was used with probability equal to the agent's wealth. Thus, the number k of i-th agent transactions during the cycle is given by the binomial distribution:

$$p_i(k) = \begin{pmatrix} N \\ k \end{pmatrix} a_i^k (1 - a_i)^{N-k}. \tag{13}$$

The gain of i-th agent after the cycle is kr/N, and the total gain of all agents is r. Thus, after normalisation (see Equation (10)) the wealth of the i-th agent will be

$$a_i \to a_i + \Delta a_i, \tag{14}$$

where

$$\Delta a_i(k) = \frac{r}{1+r} \left(\frac{k}{N} - a_i \right). \tag{15}$$

The mean value of such binomial distribution is $\langle k \rangle = N a_i$, and the variance $var(k) = N a_i (1 - a_i)$. Because Δa_i is a linear function of k, its mean value equals to

$$\langle \Delta a_i \rangle = \frac{r}{1+r}\left(\frac{N a_i}{N} - a_i\right) \equiv 0. \tag{16}$$

This result depended neither on a particular value of a_i or distribution of wealth among agents. It means that in this model, share of wealth of each agent, at average, remained constant. It corresponds to the case $\beta = 1$ in the Vallejos et al. model [14], where the initial wealth distribution was stable. Remembering that the presented here model had additional restrictions limiting the number of transactions mainly for the richest individuals, the observed concentration of wealth is therefore highly unexpected. To give some explanation of this phenomenon, we studied the variance of Δa_i, which is equal to

$$var(\Delta a_i) = \langle (\Delta a_i)^2 \rangle = \left(\frac{r}{1+r}\right)^2 \frac{a_i(1-a_i)}{N}. \tag{17}$$

The variance is a square function of the agent's wealth a_i and has a maximum for $a_i = 0.5$. In our model, the initial value of a_i was typically much smaller than 0.5 due to a large number of agents, so in that case, it was safe to assume that for poorer individuals the wealth would change slower than for richer ones. In other words, if someone became poor, it would be tough for him to regain his wealth. However, if a relatively small group of agents accumulated wealth that made the probability significantly higher (for the whole group) than 0.5, that the wealth of this group, as a whole, would be conserved due to decreasing variance for $a_i > 0.5$. This may be the mechanism which stabilises inequalities caused initially by random fluctuations.

3.2. Income and Wealth Tax Influence on the Model

Other important aspects of wealth concentration are the redistribution effect and optimal taxation problem [26,27]. To check how taxes influence the results of our exchange game model we analysed two different approaches to trade taxation. One was based on a linear income tax, i.e., where tax was collected within a single trade cycle, and the other on wealth tax, and affected total holdings of each agent collected over multiple trading cycles. In both cases, the tax was collected from all the agents and then was distributed equally among them. Thus the Equation (10) becomes:

$$a_i \to \frac{(1-t_W)a_i + (1-t_I)\sum_j \frac{r}{N} + \frac{t_W + t_I r}{N}}{1+r}, \tag{18}$$

where $\Delta a_i = \sum_j \frac{r}{N}$ is an income of i-th agent during one cycle, and t_I and t_W are income and wealth tax rates, respectively. Note that in the above relation we took into account that the global income was equal to r while the total wealth was normalised to 1. Due to the latest opinions that only wealth tax can lower wealth inequalities [27–29], we are particularly interested in comparing these taxes within our model. Therefore we studied two different situations—pure income tax with tax rate set at 10% ($t_I = 0.1$ and $t_W = 0$), and pure wealth tax with rate set at 1% ($t_I = 0$ and $t_W = 0.01$). For $r = 0.1$ such choice of rates causes redistribution of 1% of the global wealth per cycle in both cases. In general, to get the same global income from the wealth tax and the income tax, the ratio of their rates should be r. The comparison of both cases is presented in Figure 6.

In fact there was no significant difference between these two kinds of taxation. The detailed analysis of the Gini coefficient (see Figure 7) suggests that the wealth tax followed to slightly larger inequalities than the income tax. It is in contradiction with above-mentioned, well-established opinions [27–29]. On the other hand, this effect is quite easy to explain. Existence of wealth condensation meant that the relation of income to the wealth was effectively higher for richer individuals. Thus, the linear income tax hit the rich more than the wealth tax. However, in a real economy, it is easier to hide or reduce declared income than wealth. Therefore, in general, the wealth tax can be more effective as easier to enforce.

Until now we have shown that within our model there were no significant differences between income and wealth based redistribution when initially wealth was equally distributed among agents.

However, in a real society, we never have equally distributed goods. Therefore, it is particularly interesting how these two taxes affects the evolution of wealth in the case of high inequalities, such as, for example, one presented in Figure 2 obtained after 1000 cycles. The evolution of wealth in this case is presented in Figure 8.

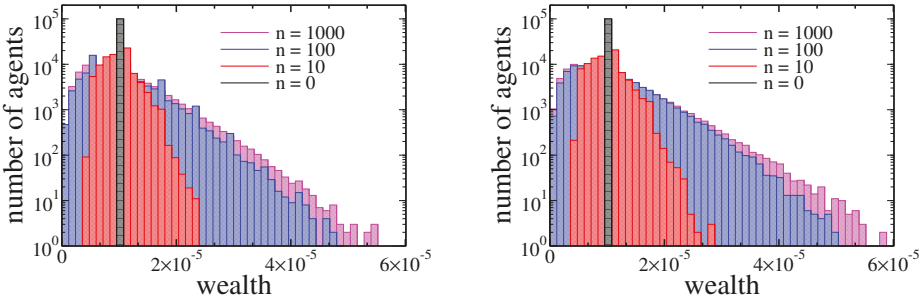


Figure 6. Histograms of agents assets after 0, 10, 100, and 1000 cycles. In both plots, the wealth of agents was initially equal ($a_i = 10^{-5}$). The left plot corresponds to income tax ($t_I = 0.1$ and $t_W = 0$), and the right one corresponds to wealth tax ($t_I = 0$ and $t_W = 0.01$).

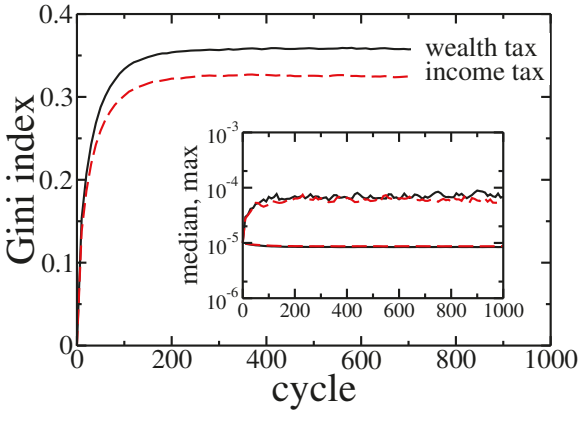


Figure 7. Gini index evolution for pure wealth (solid black line) and income (dashed red line) tax applied to the model. Initially, the wealth was distributed equally among agents ($a_i = 10^{-5}$). Inset shows the evolution of the maximal and median wealth in the population of agents.

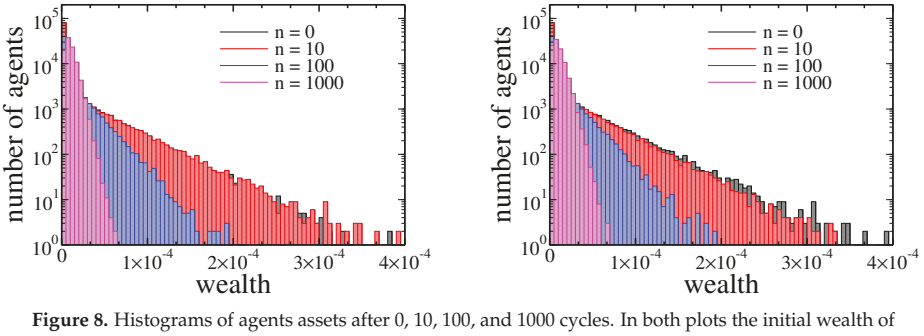


Figure 8. Histograms of agents assets after 0, 10, 100, and 1000 cycles. In both plots the initial wealth of agents was taken from the final state of simulation (1000 cycles) presented in a Figure 2. The left plot corresponds to income tax ($t_I = 0.1$ and $t_W = 0$), and the right one corresponds to wealth tax ($t_I = 0$ and $t_W = 0.01$).

Note that, in contrast to previous cases the wealth distribution shrank with time (growing number of cycles). Again however, there was no significant difference between income and wealth based redistribution. All these observations were confirmed by Gini coefficient evolution presented in Figure 9.

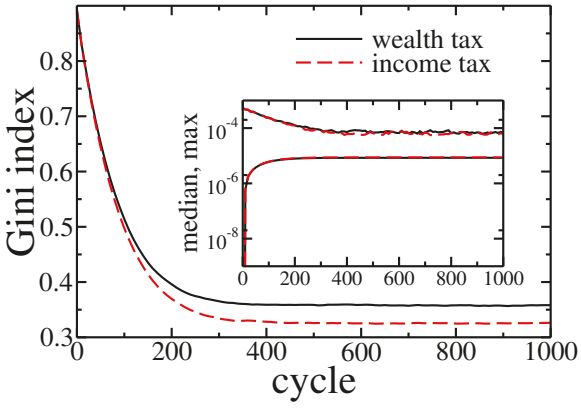


Figure 9. Gini index evolution for pure wealth (solid black line) and income (dashed red line) tax applied to the model. The initial wealth of agents was taken from the final state of simulation (1000 cycles) presented in Figure 2a. Inset shows the evolution of the maximal and median wealth in the population of agents.

In the beginning, wealth inequalities and the Gini coefficient was high, but wealth redistribution quickly tamed them. As previously, the income tax was slightly more effective. What is also important, the final state reached after 1000 cycles was similar to one with obtained for equal wealth distribution at the beginning of simulations—note the horizontal scale difference between Figure 2 and Figure 8. It allows trusting, that this final state was universal and did not depend on initial conditions. In both cases, the median wealth increased by several orders of magnitude. For example, for income tax based redistribution, the median raised from 2.1×10^{-33} to 8.7×10^{-6} at the cost of richest agent wealth, which decreased from 5.9×10^{-4} down to 6.2×10^{-5}, and the 90% of total wealth was owned by 74% of richest agents. What is even more spectacular, only 10 cycles were needed to raise the median to 6.1×10^{-7} at a cost of the richest agents wealth decreasing down to 4.7×10^{-4}.

To ultimately prove that within presented model redistribution based on the income and wealth taxes gives similar effects we analysed the dependence of Gini coefficient after 1000 cycles on redistributed amount of wealth, for high wealth inequalities at the beginning of simulations. Results are presented in Figure 10.

As expected, inequalities dropped down with growing redistribution. Moreover, the final value of Gini coefficient, the richest agent wealth and median of wealth in the society depended on the amount of redistributed wealth but they almost did not depend on a type of applied tax.

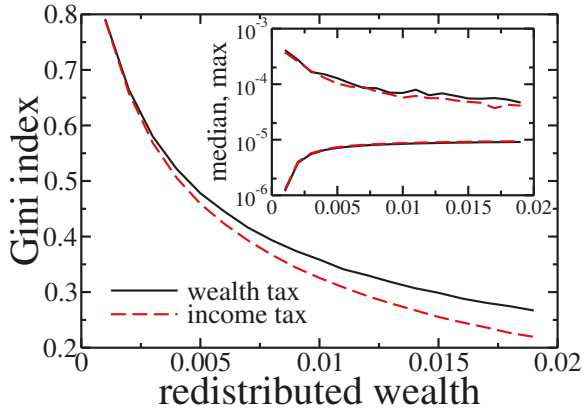


Figure 10. Gini index after 1000 cycles for wealth (solid black line) and income (dashed red line) tax based redistribution. The initial wealth of agents was taken from the final state of simulation (1000 cycles) presented in Figure 2. Inset shows final maximal and median wealth in the population of agents.

4. Conclusions

We proposed a simple mechanism owing utility theory, in which an individual with more assets has more opportunities to interact with others, but, at average, the gain from these interactions is proportional to individual's wealth. Despite this, using agent-based modelling, we showed that this mechanism causes wealth condensation independently on details of the studied model as well as its initial condition. The observation has been also supported by analytic arguments. In contrast to kinetic exchange models e.g., [10–13,15], here, the wealth inequalities grow for a large range of initial conditions and this growth is not limited by a specific distribution with exponential tails. In contrast to other models with growing economy [14], no disproportionately better treatment of wealthier agents is required to fuel the growth of wealth inequalities.

It suggests that the phenomenon of wealth condensation can be much more fundamental than expected, as it may appear even in the absence of any form of disproportionately preferential treatment of some groups of individuals.

We also studied the influence of wealth redistribution based on income and wealth taxes within the model. It occurs that while the level of inequalities depends on the amount of redistributed wealth, it almost does not depend on the type of applied tax.

Supplementary Materials: The following are available online at http://www.mdpi.com/1099-4300/22/10/1148/s1. The software used for simulations and data analysis is attached to the article.

Author Contributions: Conceptualization, M.C. and M.S.; methodology, M.C. and M.S.; software, M.C. and M.S.; formal analysis M.S.; validation; M.C.; investigation; M.C.; resources, M.C. and M.S.; visualization, M.C. and M.S.; writing–original draft preparation, M.C. and M.S.; writing–review and editing, M.C. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the Ministry of Science and Higher Education grant within the "Regional 158 Initiative of Excellence" Programme for 2019-2022. Project no.: 021/RID/2018/19.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

To find estimation and inference measures that will enable linking the model to a family of possible likelihood functions related to income data, we use a single parameter family of entropic function-power divergence measures given by [17]:

$$I(p,q,\gamma) = \frac{1}{\gamma(\gamma+1)} \sum_{i=1}^{n} p_i \left[\left(\frac{p_i}{q_i} \right)^{\gamma} - 1 \right], \tag{A1}$$

with γ being a parameter indexing Creesse and Read (C.R.)-entropy family of divergence measures-distributions, p_i are probabilities that need to be estimated, and q_i are reference probabilities from a uniform distribution.

Estimation of income distribution from a sample of real data is then equivalent to the solution of the optimisation problem [18]:

$$\hat{p} = \arg\min_{p} \left[I(p,q,\gamma) \sum_{j=1}^{n} p_j d_j \right]; \quad \sum_{j=1}^{n} p_j = 1; \quad p_j \geq 0, \tag{A2}$$

where d_j is a discrete random income variable, representing mean income with j-th bin. In the limit criterion $\gamma \to 0$, the problem further converges to:

$$\max_{p} \left[-\sum_{j=1}^{K} p_j \log p_j \sum_{j=1}^{K} p_j d_j \right], \tag{A3}$$

which can be solved by via Lagrange multiplier $\hat{\lambda}$ leading to

$$\hat{p}_j = \frac{\exp(-d_j \hat{\lambda})}{\sum_{j=1}^{K} \exp(-d_j \hat{\lambda})}. \tag{A4}$$

This procedure is part of a statistical generalised additive model for location, scale and shape. In principle, a parametric distribution, which might be heavy-tailed and positively skewed, is assumed for target variable and distribution can vary according to explanatory variables using smooth functions. This distribution is characterised by location μ, scale σ and remaining parameters are shape parameters, i.e., skewness ν and kurtosis τ [19]. To select a proper parametrisation M for wealth distribution, we apply Akaike (aic) and Bayes (bic) information criteria.

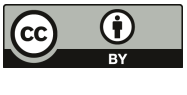
$$bic(M) = k\log(n) - 2logL \qquad aic(M) = 2k - 2logL \tag{A5}$$

with k equal to the number of model parameters, n being a sample size, and L being the maximised likelihood function and look for a distribution model M with minimal values of either bic or aic.

References

1. Pareto, V. *Cours d'économie Politique*; Librairie Droz: Geneva, Switzerland, 1964; Volume 1.
2. Heathcote, J.; Perri, F.; Violante, G.L. Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Rev. Econ. Dyn.* **2010**, *13*, 15–51. [CrossRef]
3. Ljungqvist, L.; Sargent, T.J. *Recursive Macroeconomic Theory*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2004.
4. Huggett, M. Wealth distribution in life-cycle economies. *J. Monet. Econ.* **1996**, *38*, 469–494. [CrossRef]
5. Golosov, M.; Lorenzoni, G.; Tsyvinski, A. Decentralized trading with private information. *Econometrica* **2014**, *82*, 1055–1091.
6. Bentley, R.A.; Lake, M.W.; Shennan, S.J. Specialisation and wealth inequality in a model of a clustered economic network. *J. Archaeol. Sci.* **2005**, *32*, 1346–1356. [CrossRef]

7. Mohanty, P. Why only few are so successful? *Physica A* **2007**, *384*, 75–79. [CrossRef]
8. De Grauwe, P. Animal spirits and monetary policy. *Econ. Theory* **2011**, *47*, 423–457. [CrossRef]
9. Patriarca, M.; Heinsalu, E.; Chakraborti, A. Basic kinetic wealth-exchange models: Common features and open problems. *Eur. Phys. J. B* **2010**, *73*, 145–153. [CrossRef]
10. Angle, J. The surplus theory of social stratification and the size distribution of personal wealth. *Soc. Forces* **1986**, *65*, 293–326. [CrossRef]
11. Bennati, E. Un metodo di simulazione statistica nell'analisi della distribuzione del reddito. *Riv. Internazionale Sci. Econ. Commer.* **1988**, *35*, 735–756.
12. Chakraborti, A.; Chakrabarti, B.K. Statistical mechanics of money: How saving propensity affects its distribution. *Eur. Phys. J. Condens. Matter Complex Syst.* **2000**, *17*, 167–170. [CrossRef]
13. Dragulescu, A.; Yakovenko, V.M. Statistical mechanics of money. *EPJ B* **2000**, *17*, 723–729. [CrossRef]
14. Vallejos, H.A.; Nutaro, J.J.; Perumalla, K.S. An agent-based model of the observed distribution of wealth in the United States. *J. Econ. Interact. Coord.* **2018**, *13*, 641–656. [CrossRef]
15. Lim, G.; Min, S. Analysis of solidarity effect for entropy, Pareto, and Gini indices on two-class society using kinetic wealth exchange model. *Entropy* **2020**, *22*, 386. [CrossRef]
16. Cowell, F.; Karagiannaki, E.; McKnight, A. The changing distribution of wealth in the pre-crisis US and UK: the role of socio-economic factors. *Oxf. Econ. Pap.* **2019**, *71*, 1–24. [CrossRef]
17. Cressie, N.; Read, T.R. Multinomial goodness-of-fit tests. *J. R. Stat. Soc. Ser. Methodol.* **1984**, *46*, 440–464. [CrossRef]
18. Judge, G.G.; Mittelhammer, R.C. Implications of the Cressie-Read family of additive divergences for information recovery. *Entropy* **2012**, *14*, 2427–2438. [CrossRef]
19. Rigby, R.A.; Stasinopoulos, D.M. Generalized additive models for location, scale and shape. *J. R. Stat. Soc. Ser. Appl. Stat.* **2005**, *54*, 507–554. [CrossRef]
20. Angle, J. The inequality process as a wealth maximizing process. *Phys. A Stat. Mech. Its Appl.* **2006**, *367*, 388–414. [CrossRef]
21. Bouchaud, J.P.; Mézard, M. Wealth condensation in a simple model of economy. *Physica A* **2000**, *282*, 536–545. [CrossRef]
22. Burda, Z.; Johnston, D.; Jurkiewicz, J.; Kamiński, M.; Nowak, M.A.; Papp, G.; Zahed, I. Wealth condensation in pareto macroeconomies. *Phys. Rev. E* **2002**, *65*, 026102. [CrossRef]
23. Vatn, A. Cooperative behavior and institutions. *J. Socio-Econ.* **2009**, *38*, 188–196. [CrossRef]
24. Silberberg, E.; Suen, W.C. *The Structure of Economics: A Mathematical Analysis*, 3rd ed.; McGraw-Hill: New York, NY, USA, 2000.
25. Goswami, S.; Sen, P. Agent based models for wealth distribution with preference in interaction. *Physica A* **2014**, *415*, 514–524. [CrossRef]
26. Bertotti, M.L.; Modanese, G. From microscopic taxation and redistribution models to macroscopic income distributions. *Phys. A Stat. Mech. Its Appl.* **2011**, *390*, 3782–3793. [CrossRef]
27. Kulp, C.W.; Kurtz, M.; Wilston, N.; Quigley, L. The effect of various tax and redistribution models on the Gini coefficient of simple exchange games. *Chaos Interdiscip. J. Nonlinear Sci.* **2019**, *29*, 083118. [CrossRef] [PubMed]
28. Diniz, M.; Mendes, F. Effects of taxation on money distribution. *Int. Rev. Financ. Anal.* **2012**, *23*, 81–85. [CrossRef]
29. Piketty, T.; Saez, E. A theory of optimal inheritance taxation. *Econometrica* **2013**, *81*, 1851–1886.



entropy

MDPI

Article

Macroprudential Policy in a Heterogeneous Environment—An Application of Agent-Based Approach in Systemic Risk Modelling

Jagoda Kaszowska-Mojsa [1,†] and Mateusz Pipień [2,*,†]

[1] Institute of Economics, Polish Academy of Sciences, Nowy Swiat St. 72, 00-330 Warsaw, Poland; jagoda.kaszowska@inepan.waw.pl
[2] Department of Empirical Analyses of Economic Stability, Cracow University of Economics, Rakowicka St. 27, 31-510 Cracow, Poland
* Correspondence: eepipien@cyf-kr.edu.pl
† These authors contributed equally to this work.

Received: 29 December 2019; Accepted: 17 January 2020; Published: 21 January 2020

Abstract: Assessment of welfare effects of macroprudential policy seems the most important application of the Dynamic Stochastic General Equilibrium (DSGE) framework of macro-modelling. In particular, the DSGE-3D model, with three layers of default (3D), was developed and used by the European Systemic Risk Board and European Central Bank as a reference tool to formally model the financial cycle as well as to analyze effects of macroprudential policies. Despite the extreme importance of incorporating financial constraints in Real Business Cycle (RBC) models, the resulting DSGE-3D construct still embraces the *representative agent* idea, making serious analyses of diversity of economic entities impossible. In this paper, we present an alternative to DSGE modelling that seriously departs from the assumption of the representativeness of agents. Within an Agent Based Modelling (ABM) framework, we build an environment suitable for performing counterfactual simulations of the impact of macroprudential policy on the economy, financial system and society. We contribute to the existing literature by presenting an ABM model with broad insight into heterogeneity of agents. We show the stabilizing effects of macroprudential policies in the case of economic or financial distress.

Keywords: systemic risk; macroprudential policy; agent-based modelling; inequality; central-banking

1. Introduction

The new setting of financial supervision tailored after the global financial crisis of 2008–2009 has highlighted the need for research on the nature and measurement of risk in the financial system, also called systemic risk [1–4]. In response to the problems that occurred during the global financial crisis, the Basel III regulatory framework for financial institutions was adopted in 2010–2011. In the updated version of the Basel document the capital and liquidity requirements were established. In addition, the methods of conducting stress tests in the financial system have been subject to revision. Basel III was designed to strengthen the effects of banks' capital requirements by increasing the liquidity of the banking sector and reducing leverage undertaken by banks. In the European Union (EU), the implementing act of the Basel Agreements has been issued in the form of a new legislative package covering CRD IV/CRR (i.e., CRD IV Directive No. 2013/36/EU on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms and CRR Regulation No. 575/2013 on prudential requirements for credit institutions and investment firms).

The literature on selected macroprudential policy tools [5,6] presented in Basel II and III has been mainly focused on theoretical and empirical research on linkages between real sector and the financial

system. According to International Monetary Fund, Financial Stability Board and Bank for International Settlements (IMF-FSB-BIS) [7], the assessment of the *effectiveness* of macroprudential policies includes: the assessment of the extent to which the macroprudential instrument increases the resilience of the macro-financial system; and the assessment of the extent to which the macroprudential instrument has impact on credit dynamics and asset prices (the impact on the cycle). The adoption of the new institutional framework for macroprudential supervision in the EU Member States took place in most countries during last three years. Therefore, the results of empirical studies on the effectiveness of macroprudential instruments are biased by substantial uncertainty. Alternatively one may carry out counterfactual analyses on the impact of a *combination* of macroprudential instruments on a stylised economy using the following simulation methods: dynamic stochastic general equilibrium models (DSGE) or non-equilibrium models (e.g., ABMs) [8]. Although from the presentation of ABM models in opposition to DSGE models, the conclusion can be drawn that the DSGE models are always less useful, it is important to remember that DSGE models are an extremely important tool used mainly in central banking. A defence of the legitimacy of using DSGE models even after the financial crisis can be found in Christiano et al. [9].

Analyses of the impact of macroprudential policies on the financial system and the real part of the economy have been primarily focused on capital requirements [10–13], countercyclical capital buffer [14–17] and leverage [18]. In the literature, we can also observe successive attempts to incorporate stylised macroeconomic and macroprudential policies in the form of financial frictions into Dynamic Stochastic General Equilibrium (DSGE) models [19–25]. The main goal of these attempts was to examine the effects of monetary policy or the general equilibrium welfare effects of capital requirements and leverage. Despite the role these studies played in formulating theoretical background to design of macroprudential policy, the assumptions of DSGE models are subject to critique [26,27]. DSGE models share the assumption of a perfectly rational representative agent that dynamically optimizes the use of resources. Failure to take into account the heterogeneity of agents in most DSGE models is particularly acute from the perspective of social welfare analysis performed within an equilibrium environment [28].

Both empirical and theoretical studies on the effects of the Basel III have led to the formulation of criticisms of the adopted regulatory framework. In particular, some researchers highlighted insufficient risk and uncertainty sensitivity of macroprudential tools, over-reliance on external rating regulations, improper tool calibration and lack of synchronization of adopted rules at institutional and national level. EU expert groups are still working to incorporate changes within these areas into the Basel IV framework. New research tools are required to examine the impact of regulatory changes on the economy, financial sector and society. The new setting should allow greater flexibility in modelling of risk-taking, risk aversion and decision-making under uncertainty. Consequently assumptions of macromodels should be more realistic to allow for a study of changes in welfare in heterogeneous economy beyond the social planner framework.

The aim of this paper is to analyse the impact of selected macroprudential policy tools on the economic and financial system using agent-based modelling (ABM). Modelling of interactions between agents within the ABM approach was confronted with the DSGE model with three layers of default ('3D') [21], which is currently used by experts within the European Systemic Risk Board (ESRB) and European Central Bank (ECB).

This paper contributes to the existing literature of agent-based modelling through detailed and relatively broad insight into heterogeneity of agents. In the approach taken, decision-making rules, preferences and behaviours may vary across units. In our model, all agents, ie banks, individuals, households, consumers, firms, establishments, industries, suppliers, properties, are heterogenous.

We conducted extended simulation experiments that were based on an ABM model that had been calibrated to reflect the features of a small open economy. Our choice was Poland as an exemplar case. The reason for calibrating the model relying on Polish data is that among the EU countries, the Polish economy is relatively small, open and strongly connected to the rest of the European Union countries.

Moreover, the Polish banking system still remains strongly influenced by investors from the European Union, who treat Poland as a host country. Generally, the smaller Central and Eastern European (CEE) countries that host foreign financial institutions are exposed to various dimensions of systemic risks more strongly. At the same time, the degree of the development of financial intermediation is relatively low, which results in a rather weak credit channel, especially in the case of investments. Although the financial system in Poland generates limited systemic risk, it is more vulnerable to regulatory arbitrage and the propagation of the shocks that are caused by the activity of international financial groups.

Consequently, the CEE economies and other emerging economies may need to conduct a more active macroprudential policy because of the higher risks that stem from volatile capital flows or credit booms and so forth. These issues also relate to the Polish economy and its financial system. Hence, both the ABM model and the simulations presented in our paper are valuable for gaining a detailed insight into the effects of macroprudential policy, especially in the case of small emerging open economies.

In order to study the macroeconomic effects of macroprudential instruments and their interaction with monetary policy in the case of a hypothetical small open economy, Aoki et al. [29] applied a DSGE framework. The analysed model captured some critical features of the emerging market economies with macroprudential instruments that were defined as the capital requirements that were imposed on banks and a tax on foreign currency (FX) lending. However, there are some relevant aspects that were not taken into consideration in the Aoki et al. [29] model. For example, the possibility of the government or central bank intervening in the foreign exchange markets through the use of official foreign reserves is not discussed. Moreover, what is missing in the model is a more flexible specification of international capital flows (no equity flows or foreign direct investment) and the role of cross border gross flows, which could play a destabilizing role for financial stability. The ABM construct that is presented in our paper and the simulation study seem to be a step forward in addressing some of these issues but in particular in relaxing the assumption of the homogeneity of the economic units that interact in a system.

2. Comparison of the ABM and DSGE-3D Model

The use of DSGE models historically has been primarily focused on the analysis of technological changes and their impact on the real economy [30,31] as well as the impact of monetary policy on the business cycle [32,33]. The first DSGE models with financial frictions were not used for impact studies of macroprudential policies on the macro-financial system and social welfare. The research has been mainly focused on the formal explanation of the financial accelerator hypothesis [34,35] and the role of the net worth channel in credit supply [23,36]. In a few models, the impact of LTV changes and capital requirements on the economy was analysed explicitly [25].

After the global financial crisis, interest in systemic risk assessment and macroprudential policies increased [37–39]. Currently, one of the most important examples showing the use of DSGE models in research on macroprudential policies is the DSGE model with 3 layers of default ('3D' model) [21] developed in ECB. The main goal of the '3D' model was to create a framework for analysing positive and normative macroprudential policies. This model enables to set the optimal levels of capital requirements as well as to analyse welfare within the social planner framework. However, heterogeneity of agents within the model which seems more realistic would change the optimal values for capital requirements and would make the welfare analyses more meaningful.

The paper includes a comment on the '3D' model, due to its similarity to the prepared heterogeneous agent-based simulation. In both cases, the behaviour and decisions of major market players are taken into account. The insolvency of individuals (households), companies and banks describe formally sources of systemic risk. In the '3D' model and in the ABM model, stability of the financial sector is related to the default of agents. Nonetheless, the method of modelling agent's decisions differs between both approaches. More importantly definition of the insolvency also differs.

In the '3D' model, individuals can deposit funds in banks and take loans for the purchase of real estate; entrepreneurs borrow from banks to accumulate capital. Business insolvency is associated with the occurrence of idiosyncratic and aggregate shocks. In the ABM model, the insolvency of agents is tied to internal market dynamics, driven by business and financial factors mainly the supply and investment finance policies of the banks and the demographic factors. The external shocks can be taken into account in the analyses but their significance is smaller than in the DSGE approach.

The '3D' model has been a novelty in the DSGE literature. Traditionally, in the DSGE models, due to appropriately formulated contracts, insolvency at steady state was impossible. Risk of insolvency itself was fully hedged in the model [35]. In a '3D' model, a borrower's insolvency implies changes in the lender's balance, which in turn affects his or her optimal behaviour in the market. Moreover, the bank's insolvency also entails costs to individuals and businesses, in spite of deposit insurance, which in turn strengthens the impact of bank insolvency.

In both approaches, the entire chain of interconnections between agents is formally modelled. Households save and deposit their funds in banks, while other households and companies borrow funds from the same banks. The ABM model departs from the stylised division into two dynasties: savers and borrowers. In an agent-based simulation, each household is made up of individuals with their own heterogeneous profiles in terms of savings, income, spending or additional financing. The way households make decisions depends on the interaction of an individual with family members and other agents in the model. In this way it is possible to trace the way of transferring the risk of insolvency between sectors. However, the transmission of default risk between sectors in DSGE models is accomplished by further optimisation of resources by a representative agent assuming appropriate restrictions and rational expectations. In the ABM model, insolvency transmission is accomplished not only through actual economic and balance-sheet relations and constraints but also the perception of risk and uncertainty of heterogeneous agents in the system [40–42]. DelliGatti [41] binds DSGE models with the financial accelerator hypothesis, while ABM models utilise the instability hypothesis of H. Minsky, which take into account not only the financial accelerator hypothesis but also Knightian uncertainty [43]. Additionally, the ABM model presented includes not only the transmission of risk of insolvency between the agents but also between industries.

In the '3D' model the attention is drawn to two types of distortions, which drive banks to excessive use of leverage and significant exposure to risk. They also explain the need for macroprudential policy. The first is related to the existence of deposit insurance agency. Banks run increased risk at the expense of an external agency, leading to higher credit supply and increased demand for deposits. The second distortion is related to the fact that the insolvency is expensive, not only for the lender but also for the borrowers. Occurrence of costly state verification [19] leads to lower demand for credit. The net effect of the two market distortions may be different for each sector. Consequently, the supply of credit for each sector may be lower or greater than the level that maximises the welfare of society.

According to the logic of the '3D' model, in a steady state, when the probability of bank insolvency is high, the risk premium is raised. We assume in the '3D' model that the risk premium is for the whole system and not for a given bank; therefore, banks are willing to take a higher risk because their funding cost is the result of decisions made by all participants within the market. These results seem obvious and a natural conclusion of homogeneity of individuals particularly banks. In the '3D' model, we assume a certain probability of bank default, which is characteristic of the state of equilibrium that we analyse. In financial markets, the decisions of a bank depend on its perception of counterparty risk and estimation of the way a bank is assessed by other units (the »perception of perception« of uncertainty). One of the dimensions of the heterogeneity of agents is related to differences in perception of reality, inter alia the perception of counterparty risk, overall uncertainty in the market and the state of the economy. Those elements are clearly omitted in the aforementioned DSGE '3D' model. General frameworks built upon an idea of equilibrium make it impossible to generalise existing DSGE models regarding the aforementioned issues. Also, assumption that a particular individual's decision may or may not lead to achieving equilibrium seems more realistic. The ABM approach

helps overcome these drawbacks. Analyses conducted only within the state of general equilibrium are departed from and the field of interest of non-equilibrium theories and the instability hypothesis is entered [44,45]. The ABM approach is therefore a step towards overcoming the DSGE modelling imperfections indicated by many authors [46,47]. According to empirical literature, the response of systems to the occurrence of shocks may be nonlinear and the financial system itself may be unstable. In addition, shock effects exhibit asymmetric nature. By design, mainly due to the use of log-linearisation, DSGE models are not able to accommodate non-linearity.

DSGE models with financial frictions [19,48], including the '3D' model, refer to the hypothesis of the financial accelerator system. At the same time, the system is characterised by market failures, including the asymmetric information and externalities. A number of researchers highlight the presence of pecuniary externalities [49–51]. Pecuniary externalities complement technological externalities and aggregate demand externalities [52]. After the recent financial crisis, one may observe an increased interest in explaining the impact of pecuniary externalities on the system and social welfare. Pecuniary externalities are incorporated mainly into the models by means of credit restrictions. An example of pecuniary externality may be the lack of internalisation of effects of investment decisions in housing and capital prices, which in turn affects the required collateral. In the '3D' model the level of leverage of households and firms is affected endogenously by prices, including real estate prices. At the same time, the direction and size of the impact of pecuniary externalities on allocation of resources are difficult to estimate using the '3D' model.

In agent-based models, it is possible to go one step further to include the premises of instability hypothesis in the analysis. The instability hypothesis is closely tied to financial accelerator hypothesis and it can assume existence of pecuniary externalities. Nonetheless, it goes far beyond it. DelliGatti [41] distinguishes two ways of presenting Minsky's hypothesis. The first does not refer explicitly to the heterogeneity of agents, and the second assumes the existence of three types of agents; hedging, speculative and Ponzi agents. According to the first explanation, the level of investment in the economy depends on the volume of internal finance and the difference between the market price of assets and the price of the final good. The market price of assets depends on long-term profit expectations. Final good price depends on the expected demand for that good. In the absence of heterogeneity of agents, the representative agent's investment level decisions are a function of internal financing, which in principle is consistent with the financial accelerator hypothesis. In practice, investment decisions are also made according to how the agents perceive risk. Hence, according to Delli Gatti, in order to fully understand the Minsky's hypothesis, we need to distinguish between three types of agents that have different attitudes towards external financing.

During the economic boom period, both the borrower and lender expect future cash flows to increase at a pace that will allow the borrower to repay their obligations. As the expectations develop, asset and product prices increase, stimulating investment growth. As a result, production, profits and employment in the economy increase. Banks increase the supply of credit, often requiring lower collateral. Companies are less cautious when they borrow money. Consequently, the proportion of speculative and Ponzi agents increase and the financial system resilience decreases. If the level of debt in relation to its service is perceived as too high on aggregate, the number of insolvency announcements in the system increases, leading to an eventual financial crisis.

Both the '3D' model and the ABM simulation are examples of stochastic dynamic systems that describe the evolution of basic components of the economy. However, while in the '3D' model the economic agents are homogeneous, fully rational and dynamically optimising, in the ABM simulation model, the agents are fully heterogeneous, bounded rational and they perform heuristic optimisation [53–56].

In order to include the conclusions of the Minsky's hypothesis in analyses, heterogeneity of agents was included in the agent-based simulation. The heterogeneity of the economy is understood here, however, more broadly than the differentiation of attitudes towards external financing. It is understood

as a differentiation of states, behaviour rules and expectations; this implies heterogeneous distributions of variables *ex ante* and *ex post*.

Both groups of researchers, working on the DSGE models and ABM models respectively have recognised the need to consider heterogeneity of the economy in order to analyse the optimality of policies as well as welfare implications. The heterogeneity of agents in the ABM approach is, however, understood differently than in DSGE models with heterogeneous agents [57–59]. In the DSGE models with heterogeneous agents, the discontinuation of the assumption of a representative agent is made primarily by allowing for idiosyncratic shocks and by removing the assumption of completeness of asset markets. In particular, such a definition of heterogeneity requires the redefinition of basic model and analyses elements, including the definition of steady state and equilibrium [58]. States of the economy are generally considered to be the realisation of a complex stochastic process with approximate properties to the Markov processes. Therefore, in such models, stationary equilibrium is considered, within which the stationary (ergodic) distribution exists. Within these model types, decision functions and price process realisations are approximated numerically. Some of these techniques were also adopted in the ABM approach [60–62].

The heterogeneity of agents in ABM is understood differently. This could be due to differentiation of attributes and states, differentiation of decision rules [63,64], attitudes towards risks or expectations [65–67]. In most ABM models with heterogeneous expectations, agents typically have adaptive expectations, as opposed to the rational expectations of representative agent within the DSGE approach. All these dimensions of heterogeneity appear in the simulation presented in the next section. Among others, the empirical distributions of basic economic categories such as income, expenses or credits were used to calibrate the states and the decision functions and procedures were selected after conducting the empirical research on the patterns of consumption and production on the market. The adaptive expectations were imposed as well on simulation design. A key difference between the presented simulation and other agent-based models is inclusion of varied attitudes towards risk and uncertainty in Knight's sense and the risk sensitivities.

The final distinguishing element of the ABM approach is the possibility to introduce more realistic assumptions in the model than in the DSGE approach. Good examples are the assumption of visibility and satiation. This visibility means that agent decisions take into account not only purely economic conditions and factors, such as the price of the product but also the proximity of the supplier in a spatial sense. In the case of analysis of financial system, the idea of visibility has an additional dimension. It does not come down to visibility in a geographic or spatial sense but rather to the perception of banks as relatively safe institutions. The perception of banks does not have to be reflected in economic foundations or stances [42]. Adopting the saturation principle leads to a departure from the global optimisation of underlying criteria with restrictions on the choice of local maxima.

The presented ABM model is based on the traditions of the EURACE [68,69], FP7 MOSIPS and the population dynamics model in the EU regions models [70] as well as FP7 CRISIS models. Our ABM model is also consistent with the stock-flow approach [71,72]. Impact studies of macroprudential policies on the economy within the ABM approach is relatively new. However, the topic refers to the tradition of agent-based models within financial markets [73–77] as well as literature on credit and financial markets from the agent-based perspective [72,78–97]. In the broader sense, the study also refers to the coevolution models successfully applied in [98,99] to explain the stylized fact of persistency in a time series. For more general reviews on complex network theory refer to References [100,101], while spatial interactions in agent-based modeling were discussed in Reference [102,103].

The purpose of the model is to bridge the gap in the literature on the role of macroprudential policies in systemic risk mitigation. In the following section, details of the ABM model, simulation results and an explanation of the logic behind robustness checks is provided. Comments on welfare analysis within the DSGE and ABM approach are also provided; hence a critical perspective on '3D' modelling results is presented.

3. Model Description

Presented within this section is an ABM model suitable for performing simulations that provide detailed insight into the nature of the relationship between the financial system and the real part of the economy. Due to the specifics of agent-based modelling, initially presented is the software environment in which simulations were developed. Next, attributes and activities of agents are presented. Also discussed is the method of sequential updating of the states within simulation modules. The form used to present the model and simulation is consistent with 'A Common Protocol for Agent-Based Social Simulations' [104].

3.1. The Software

The simulation was developed using object-oriented programming in Java-NetBeans and Eclipse environments. Statistical data to determine the attributes of system agents were grouped in a relational database (PostGreSQL-pgAdmin III). The simulation was linked to the database using Hibernate and SQL queries. According to the logic of object-oriented programming, initially agents and their attributes are described, and then the simulation workings from the perspective of individual agents and their activities are discussed. In the next subsection, a sequential update of agents' attributes in simulation modules is presented.

3.2. Agents and Attributes

In the macro-finance model, 9 agents are distinguished: *Banks*, *Individuals*, *Households*, *Consumers*, *Firms*, *Establishments*, *Suppliers*, *Properties* and *Industries*. All *Parameters* in a separate object are also defined. The relations between agents in the model are presented in the Figure 1. The attributes of the agents can be found in the Tables A1–A10 in the Appendix A.

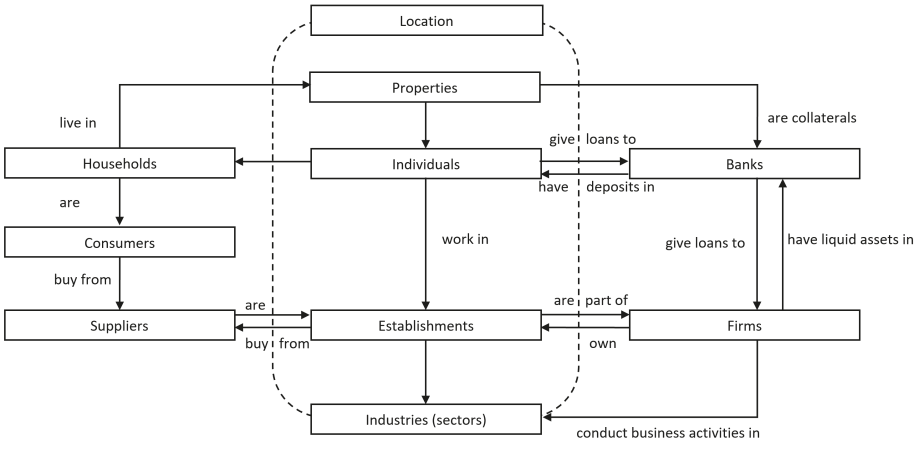


Figure 1. Relations between agents in the model.

3.2.1. Individuals, Households, Consumers & Properties

Individuals do not determine the behaviour of the system in the initial *Initialisation*, *Production*, *Supply chain* and *Public contracts* modules. Their significance is enhanced only in the *Households consumption*, *Households mobility* and *Individuals' records updating* modules. Nonetheless, individuals play a special role in the model because they determine the functioning of the program and the way data that describe agents' attributes is mapped in the relational database.

In the *Households consumption* module, the sum of the income of individual entities forming the given household is initially calculated. Total household income is not only equal to the sum

of family member income but also includes additional income from rental property and alimony payments. After calculating total household income, resources are divided between consumption and savings. The program counts the number of household consumers and their total disposable income. Depending on the level of disposable income, household savings are determined to update household members' deposits. If a household consists of a couple (with or without children), then savings are distributed between them. Otherwise, if the household consists of a single adult or single parent, the savings are given to that adult. If the disposable income is exceptionally low, the scheme works similarly, with the common deposits used primarily to cover the basic needs. Decisions on the distribution of income between consumption and savings and the distribution of savings between household members are dynamic. For example, if the household changes as a result of divorce or death of the spouse, the states are updated.

Households represent different types of consumers in the model that determine the purchase of goods and services from a given industry. Households purchase products from suppliers according to the price offered by the supplier relative to the average industry price, quality of goods or services relative to average industry quality and depending on supplier's spatial location. Households of a certain consumer type seek suppliers sequentially further from their location. Ultimately, they seek suppliers globally, taking into account only the price and quality of the product or service.

Purchases of goods and services do not need to be financed solely by funds deposited in a bank account. Households with creditworthiness are eligible for consumer credit. In the model, loan maturity depends on the amount of the loan.

In the *Households mobility* module, households decide the place of residence and purchase or lease the property. If the household already owns the property, the model verifies whether the cost of repaying the mortgaged loan relative to the household income is too high. If repayment cost is too high, the property is designated for sale. If the property in the previous iteration has already been marked for sale, the price is reduced.

If a household leases a property, rent is equal to the sum of the property owners bank loan repayment obligations; rent is calculated as the ratio of property price to the number of families renting the property. If the rental cost is too high, the household seeks a new rental property. Preference is given to properties near the current place of residence. If household members are not working, further conditions are laid down for disposable income and burdens on household loans. If income after deductions is relatively low and the person is over 25 and not working, then the individual defaults. The banks which have granted the loans to this individual update the non-performing loans value. In this module, income from renting real estate property to other households is also updated. If the property is not leased, it is marked for sale. If it is not purchased, subsequent iterations reduce the desired property sale value.

One of the most important elements of this module is the ability to purchase real estate. Households with high savings buy property in cash with a given probability. Nonetheless, some households, despite their resources, decide to apply for a mortgage loan. Households that do not have sufficient savings, but meet the requirements, also take a loan. If the household is already the owner of one of the properties, it may additionally take a non-residential loan on the pledge of the first property. In practice, either the household takes a residential or non-residential loan, taking into account whether it is already a property owner. Probabilities of taking a residential and non-residential loan were estimated based on empirical data.

Attributes of individual entities change their values in the *Individuals' records updating* module. Depending on age and sex, the program determines the probability of death of an individual. If the probability is high, the person dies and the assets capital and deposits go to the heir. If the deceased person was the owner of a business, the heir can continue the business or forgo, depending on their previous income as an employee in one of the firms. Simultaneously, the division of capital in the economy changes. If the probability of individual's death is low, the program directs

the individual to the *Education* and *Pairing* modules. In the selected modules the consumer type of households is updated and in the final modules the remaining individuals' records are updated.

3.2.2. Firms & Establishments

The role of *Firms* is highlighted primarily in the modules: *Firm demography, Mergers* & *Acquisitions* and *Firm growth*. An entrepreneur can open a new business according to a certain probability that depends on the experience of running the company, the age and level of education completed by the individual. When deciding to set up a new business, the entrepreneur takes into account the average profitability of individual industries in the economy, the ease of obtaining licenses for running a business in a given industry and the ability to raise additional funds for opening and running a business. In the case of small and medium-sized enterprises, these funds can be obtained from banks, while in the case of large companies, capital is obtained from many individuals, who are henceforth shareholders of this company. Obtaining funding from a bank requires a number of formal requirements to be met, including leverage, investment and industry risk, investment risk mitigation and a good credit history of the applicant. Implicitly, banks also take into account the cost of labour, equity and size of the enterprise relative to the average enterprise size in the industry.

As a result of the acquisition of another company or as a result of the company's strategic development, firms can increase the number of establishments they own. In the model, it is possible to obtain additional funds from banks for expansion. Firms can also cease their business activity.

In the simulation, companies announce insolvency when the level of indebtedness and the risk of business exceed the acceptable level. A low percentage of business that has been run in the low-profit sector at high operating costs defaults as well. In the event of a company's inactivity on the market for six quarters, the program automatically removes the company from the database and program. The adoption of such assumptions ensures adequate market dynamics in the simulation. Firms and establishments are auxiliary objects for the remaining agents in the model. In the final modules, other attributes of firms and establishments are updated.

3.2.3. Establishments & Suppliers

Establishments in the simulation allow for the spatial location of businesses. In the *Initialisation* module, the maximum potential production of goods and services of the establishment is computed. The price of the goods produced by the establishments changes depending on the demand. In each period, the optimal number of products to be stored for future sales is calculated. In the next period, the facility will only produce the number of goods equal to the number of goods demanded minus the number of goods stored. In addition, the production process takes into account the manufacturing risk and the overall level of corporate debt, which should not exceed the level specified by prudential regulation and policies. For the production of final goods, the establishments purchase inputs from others acting as suppliers. The choice of supplier is designed in such a way as to take into account economic categories such as price or product quality but also supplier availability in a geographical or spatial sense.

The demand from the private sector is supplemented by the need for goods from the public sector. It was assumed that the ability to sign a contract in a public sector depends on the size of the establishments producing the given good and the price and quality of the product offered compared to the average values in the industry. The establishments then decide on the destination of the final goods. Establishments may allocate all products for sale in a given sector or export some of their products to another sector and other spatial units.

In addition, the establishments play an auxiliary role in other modules. Individuals seek buildings near the workplace, that is, in the territorial unit within which the establishments are located. At the time of setting up a new company, new establishments are also created. Similarly, when a new company is created as a result of an acquisition or merger, the affiliation of the company and the owner of the establishments change. The owners of establishments decide to increase or

decrease the workforce and firms with a strong market position increase the number of establishments. In addition, depending on the demand for work, the number of employees to be hired and fired in each establishment is calculated. In the final modules, firms pay salaries to establishments' employees and the remaining attributes of establishments are updated.

3.2.4. Industries

The existence of major branches of the economy is assumed. The role of industries is crucial when calculating large exposures (LE) of banks to particular industries. *Industries* is an auxiliary object as well. Firms and establishments operate within the industries. The establishments may import and export goods between industries as part of the purchase of inputs for production. When firms apply for a loan, the total exposure to the industry is checked, which should not exceed the regulatory thresholds. The average values of variables for the industries are treated as a benchmark for business operations of firms and establishments. The main values calculated by the program are the number of units operating in the sector, average product price in the sector, average good quality, average firm size in the sector, average import and export, average industry earnings and average industry workforce. When entrepreneurs decide to run a business, first they attempt it in the relatively most profitable sector, and then in the next sectors.

3.2.5. Banks & Macroprudential Policies

Banks provide loans to individuals, households and companies. In the case of individuals and households, we distinguish between consumer loans and mortgages for residential and non-residential purposes. Companies can apply for a loan to purchase inputs and increase sales and business development (investment loans). Banks analyse loans granted to each of the industries and examine the risks associated with high exposures to the industry. According to regulatory requirements, banks are not allowed to lend to specific industries above certain thresholds.

Individuals and establishments accumulate funds in a bank account. To each individual and establishment at least one bank is matched based on survey data. The model assumes the existence of network and reputation effects. According to the results of empirical research, individuals and households are not guided solely by price in the decision to allocate funds. With relatively high probability, the entity will decide to remain with the bank assigned to them. On the other hand, if they decide to change bank, they will take into account interest rates and additional transaction costs. In the case of consumer credit and the purchase of inputs, agents are more driven by the reputation of the bank than the interest rate. On the other hand, in the case of residential and non-residential loans as well as investment loans, households and firms are primarily guided by the long-term interest rates of banks.

In the module *Banks Supply side & regulatory requirements*, banks set the supply of different types of loans. Banks compete with each other in terms of price, offering different interest rates on deposits and loans, as well as in terms of creditworthiness criteria. Risk-taking banks provide loans to individuals, households or companies with a lower credit history or lower creditworthiness (income or equity respectively). In the absence of macroprudential policy solutions, banks could be willing to lend to more risk-prone entities, which would jeopardise the stability of the financial sector. Thus, the model takes into account the examples of macroprudential policy tools. Firstly, the existence of capital requirements (CAR) and recommendations of the financial supervision authority regarding the capital maintained by banks were taken into account. The model then included the large exposures (LE) and exposures of risk to the industries. Banks cannot lend to a given industry over a specified amount and will not choose to lend to a company that runs risk-prone business without sufficient collateral. By providing housing and non-residential loans, banks pay attention to the following indicators: debt to assets (DTA), debt service to income (DSTI) and loan to value ratios (LTV). Moreover, the liquidity ratio (LCR) and leverage ratio (LR) have been taken into account as well.

From the supervisor's point of view, it is extremely important to analyse the value of these indicators on aggregate for the economy. The survey data allows only for a static description of the level of these ratios for a given period. The simulation allows for the investigation of changes that occur in the indicator values as a result of dynamic interactions between fully heterogeneous agents. Similarly, for companies and premises leverage requirements (LR) are analysed.

Individuals, households, firms and establishments make decisions about depositing funds and obtaining a loan based on interest rates. In the system it is possible to introduce interest rates offered by banks. In this case, after taking into account the network and reputation effects, banks compete on interest rates. It is possible to take into account counterparty risk indicators and indicators of »perception of perception« in the analysis. It is then possible to integrate a macro-financial model into a financial model that simulates the role of risk perception and uncertainty in generating systemic risk in the interbank market [42].

4. Sequential Updating of States in the Model

The graphical representation of sequential updating of states in the model is presented in Appendix B.

Module 0: Initialisation Sector Profitability (M.1)

In the *Initialisation* module, we calculate the average profit (that is $\overline{\pi}_t^s$) of the $N_t^{firm_s}$ firms doing business in the S industries (s, where $s \in \{1, 2 \dots S\}$, $S \in \mathbb{N}$) at time t. The procedure classifies sectors according to their average profitability. Information on average profitability is used by individuals when they decide to establish a new firm in a given industry. Each firm in the sector s generates profits ($\Pi_t^{firm_s}$).

$$\overline{\pi}_t^s = \sum_{firm_s=1}^{N} \frac{\Pi_t^{firm_s}}{N_t^{firm_s}}. \tag{1}$$

In this module, the model also stores the initial supply of different credit types (for each bank b): consumer loans $S_t^{l.C^{ind}}$, residential $S_t^{l.H^{ind}}$ and non-residential loans ($S_t^{NH^{ind}}$), firm investment loans ($S_t^{l.I^{est}}$) and short-term loans for firms ($S_t^{l.SH^{est}}$) as temporal variables. This information is used later in module (M.55) to determine how much supply was used during this iteration by different agents.

Module 1: Production Price Updating (M.2)

In the *Price updating* submodule, the price (P_t^{est}) is updated according to the demand $\left(Q_t^{d_{est}}\right)$ for a given good or service in relation to the expected demand ($E\left(Q_t^{d_{est}}\right)$) and the level of production $\left(Y_t^{est}\right)$ relative to the maximum potential production of each establishment ($Y_{t_{max}}^s$). In this sub-module, the values of variables determining the number of employees to be hired $\left(L_t^{HI^{est}}\right)$ and fired $\left(L_t^{FI^{est}}\right)$ in the current period in each establishment are set to zero.

The maximum production of premises in a given sector is then calculated according to the Cobb and Douglas production function:

$$Y_{t_{max}}^s = \alpha_{1.s}(L_t^{est} + 1)^{\alpha_{2.s}}(K_t^{est})^{\alpha_{3.s}}\left(\frac{(Ql)_t^{est}}{(Ql)_t^s}\right)^{\alpha_{4.s}}(A_t^{est})^{\alpha_{5.s}}, \tag{2}$$

where L_t^{est} is the labour force, K_t^{est} is the capital, A_t^{est} is technology and $(Ql)_t^{est}/(Ql)_t^s$ represents the relative quality of the establishment's product (or service) with respect to the average quality of product and services in the sector (industry). The value of parameters $\alpha_{1.s}$ $\alpha_{2.s}$ $\alpha_{3.s}$ $\alpha_{4.s}$ $\alpha_{5.s}$ are specific to each industry. When initialising the system, the price is defined based on the initial

conditions in the database. In subsequent iterations, the price of the previous period is assumed to be the initial value of the good ($P_t^{est} = P_{t-1}^{est}$). This price may change depending on the demand relative to expected demand and production in comparison to the maximum capacity, that is to say the maximum potential production. If the demand for good produced by a given facility is greater than the expected demand for that good and production is greater than the specified part α_6 of maximum production, then the price increases in proportion to the given parameter $\alpha_{7.s}$. The $\alpha_{7.s}$. parameter is industry-specific. If the demand for a good is less than the part α_8 of expected demand for this good and output is less than the specified α_9 part of maximum output, the price drops by the percentage given by the parameter $\alpha_{10.s}$. The $\alpha_{10.s}$ parameter is industry-specific. This procedure is consistent with the adoption of adaptive expectations in the model.

Module 1: Production Expected Demand Updating (M.3)

In this sub-module, we update the expected demand for the next iteration ($E(Q_t^{d_{est}})$). The formula of the expected demand for a good depends on the production experience of the establishment. If the premises have been operating on the market for at least a quarter, the expected demand for its good is calculated according to the following formula:

$$E\left(Q_t^{d_{est}}\right) = \left[\alpha_{11.s}\left(\frac{P_t^{est}}{P_t^s}\right) + \alpha_{12.s}\left(\frac{P_{t-1}^{est}}{P_{t-1}^s}\right) + \alpha_{13.s}\left(\frac{(Ql)_t^{est}}{(Ql)_t^s}\right) + \alpha_{14.s}\left(\frac{(Ql)_{t-1}^{est}}{(Ql)_{t-1}^s}\right)\right] \times Q_{t-1}^{d_{est}}. \quad (3)$$

The expected demand depends on the price of the product relative to the average price in the industry within the periods t and $t-1$, product quality relative to the average quality in the industry within the periods t and $t-1$, given parameter values, and demand for the goods up to date. The parameters $\alpha_{11.s}$, $\alpha_{12.s}$ $\alpha_{13.s}$ and $\alpha_{14.s}$ are industry-specific.

However, if the establishment is new, then the expected demand is calculated to take into account the workforce in the newly-created establishment and the average sales per worker in the industry $(Sl)_t^s$ within which it operates.

$$E\left(Q_t^{d_{est}}\right) = \left[\alpha_{11.s}\left(\frac{P_t^{est}}{P_t^s}\right) + \alpha_{12.s}\left(\frac{P_{t-1}^{est}}{P_{t-1}^s}\right) + \alpha_{13.s}\left(\frac{(Ql)_t^{est}}{(Ql)_t^s}\right) + \alpha_{14.s}\left(\frac{(Ql)_{t-1}^{est}}{(Ql)_{t-1}^s}\right)\right] \times (L_t^{est}+1) \times (Sl)_t^s. \quad (4)$$

Module 1: Production Expected Stock Updating (M.4)

If the establishment is new ($(New)_t^{est} = 1$), we calculate the optimal level of stock ($Inv_{opt_t}^{est}$) as part of the expected demand, specified by the parameter α_{15}. If the establishment is already established, the optimum stock level is calculated to take into account the expected demand for the good, the ratio of expected revenue from the sale of goods ($P_t^{est} \times E\left(Q_t^{d_{est}}\right)$) to the costs of producing that good in the current period (TC_t^{est}), and the ratio of sales revenue in the current period ($P_t^{est} \times (Sl)_t^{est}$) to the total costs incurred in the previous period (TC_{t-1}^{est}).

$$Inv_{opt_t}^{est} = \alpha_{15} \times E\left(Q_t^{d_{est}}\right) + \alpha_{16} \times \left(\frac{P_t^{est} \times E\left(Q_t^{d_{est}}\right)}{TC_t^{est}}\right) + \alpha_{17} \times \left(\frac{P_t^{est} \times (Sl)_t^{est}}{TC_{t-1}^{est}}\right). \quad (5)$$

Module 1: Production Production Decision Making (M.5)

If the optimum stock of products ($Inv_{opt_t}^{est}$) is less than or equal to the actual stock (Inv_t^{est}), the establishment will not produce goods in the current period. However, if the optimal stock is greater than the stored number of goods, the establishment should produce the difference between the optimum level and the current stock.

$$Y_t^{est} = \min(Y_{t_{max}}^{est};\ Inv_{opt_t}^{est} - Inv_t^{est}). \quad (6)$$

Nonetheless, these establishments will produce goods only when the level of leverage and financial risk associated with the debt of the establishments does not exceed the levels specified by the parameters α_{18} and α_{19}. If the establishment meets the conditions to produce goods, the production is equal to the lower value of either maximum production of the establishment or the difference between the optimum stock and the actual stock (inventory) of goods.

Module 2: Supply Chain Quantity of Inputs, Import & Export (M.6)

In this module, establishments buy inputs and decide on the import and export of goods between industries. In order to minimise costs, they choose a supplier from the nearest spatially located area, thus limiting the cost of transport. In addition, the adopted mechanism allows the modelling of continuation of transaction relationships between suppliers and recipients of goods. Each company is located spatially in the form of establishments. Each establishment is a supplier for another establishment. For each establishment in all sectors, the initial value of inputs ($(Inp)_t^{est}$), profits from sales ($(Sl)_t^{est}$) and demand for goods ($Q_t^{d_{est}}$) are set to zero. Next, the amount of inputs ($q_t^{est.buy.(sup)(.s)}$) (provided by suppliers from sectors s) necessary to ensure continuity of production, taking into account import and export of inputs between industries is calculated. If the facility imports or exports semi-finished products, the amount of inputs that the establishment is going to purchase is obtained using the following formula:

$$q_t^{est.buy.(sup)(.s)} = \alpha_{20.s-buy.s} \times Y_t^{est} \times \alpha_{21.s-buy.s}. \tag{7}$$

In the model there are $2 \times s$ values for parameters $\alpha_{20.s-buy.s}$ and $\alpha_{21.s-buy.s}$ (*Cf.* Calibration for the explanation how the values of parameters were obtained). For each establishment in the industry, the value of the parameter is the same but the values vary between sectors (industries). If the establishment does not import goods then the quantity of purchased goods is equal to the part of production specified by parameter $\alpha_{20.s-buy.s}$:

$$q_t^{est.buy.(sup)(.s)} = \alpha_{20.s-buy.s} \times Y_t^{est}. \tag{8}$$

The total quantity of inputs is the sum of inputs purchased from all suppliers in all sectors ($q_t^{est.buy}$).

Module 2: Supply Chain Supplier Selection (M.7)

When searching for a supplier, the establishment takes into account the amount of goods stored by the supplier ($Inv_t^{est(.\mathrm{sup})} > q_t^{est.buy(.s)}$), and compares the ratio of quality to price of a supplier (establishment in the sector s) with the average ratio within the industry (sector):

$$\frac{\alpha_{22.s-est} \times (Ql)_t^{est(.\mathrm{s})}}{\alpha_{23.s-est} \times P_t^{est(.\mathrm{s})}} > \frac{\alpha_{24.s-est} \times (Ql)_t^{s}}{\alpha_{25.s-est} \times P_t^{s}}. \tag{9}$$

In addition, it also takes into account supplier location (i.e., compares the spatial codes at NUTS 1-4 levels: $\vartheta_t^{est^1}$, $\vartheta_t^{est^2}$, $\vartheta_t^{est^3}$, $\vartheta_t^{est^4}$). If the current supplier has a sufficient number of inputs for sale, and the quality and price of the good are acceptable in relation to the average price and quality in the sector, then the establishment can buy inputs from the supplier. The model consolidates the network effects developed during the cooperation of businesses. If the supplier does not meet the requirements, the establishments seek a new supplier locally in increasingly distant locations and then globally. The parameter values from $\alpha_{22.s-est}$ at time t to $\alpha_{25.s-est}$ are specific to the supplier's sector.

Module 2: Supply Chain Inputs Purchase (M.8)

After selecting a supplier, the establishment purchases inputs. To purchase inputs, the establishment must have sufficient liquid assets to cover the wages and the cost of buying the inputs:

$(LA)_t^{est.buy} - W_t^{est.buy} \geq (1 + \alpha_{26.s-sup} \times \varpi) \times P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)}$, where ϖ is the binary variable expressing whether the cost of transportation should be added. If it has sufficient liquid assets, it can finance the purchase of inputs from accumulated funds. Therefore, in the model, with the probability pr_1 the establishment will not apply for a loan. In that case, for the establishment-buyer inputs $((Inp)_t^{est.buy})$ and liquid assets $((LA)_t^{est.buy})$ are updated to. The signs "+=" shall be interpreted as the incrementation of the value of the variable by the amount quantified by the formula given on the right-hand side. Respectively, "-=", shall be interpreted as a decrease in the value.

$$(Inp)_t^{est.buy} += (1 + \alpha_{26.s-sup} \times \varpi) \times P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)} \tag{10}$$

$$(LA)_t^{est.buy} -= (1 + \alpha_{26.s-sup} \times \varpi) \times P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)}. \tag{11}$$

While for all establishments-suppliers from each sector, the sales expressed in monetary terms $((Sl)_t^{est.(sup)(.s)})$, demand for goods $(Q_t^{d_{est(.sup).(s)}})$, liquid assets $((LA)_t^{est.(sup)(.s)})$ and stock $((Inv)_t^{est.(sup)(.s)})$ are updated.

$$(Sl)_t^{est.(sup)(.s)} += P_t^{est(.sup).(s)} \times q_t^{est.buy(.s)} \tag{12}$$

$$(LA)_t^{est.(sup)(.s)} += P_t^{est(.sup).(s)} \times q_t^{est.buy(.s)} \tag{13}$$

$$Q_t^{d_{est.(sup).(s)}} += q_t^{est.buy(.s)} \tag{14}$$

$$(Inv)_t^{est.(sup)(.s)} -= q_t^{est.buy(.s)}. \tag{15}$$

In particular cases, with the probability of $1 - pr_1$, despite sufficient liquid assets, the establishment may apply for a loan to purchase additional inputs that will allow the facility to increase its production capacity and sales. If the establishment applies for a loan, the applicant's creditworthiness is checked even if its accumulated funds are sufficient to cover the purchase. In the submodule *Bank credit admissibility 1* (M.9), conditions in addition to liquidity funds are checked. In accordance to the market dynamics of short-term loans, some of the applicants will not obtain a loan from the bank due to lack of creditworthiness. The possibility of establishments applying for a short-term loan in the case of temporary liquidity problems has also been included in the model (i.e., $(LA)_t^{est.buy} < (1 + \alpha_{26.s-sup} \times \varpi) \times P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)}$). If, in spite of short-term liquidity problems, the establishment has not completely lost its creditworthiness, the bank may grant him credit for the purchase of inputs in the submodule *Bank credit admissibility 2* (M.10). If the establishment has no creditworthiness, it has to *adjust* the quantity to buy $(\tilde{q}_t^{est.buy})$:

$$\tilde{q}_t^{est.buy(.s)} = round\left(\frac{((LA)_t^{est.buy} - W_t^{est.buy}) \times \alpha_{20.s-buy.s}}{P_t^{est.(sup)(.s)} \times (1 + \alpha_{26.s-sup} \times \varpi)}\right). \tag{16}$$

After the purchase, the value of inputs $((Inp)_t^{est})$ and liquid assets of establishment-buyer $((LA)_t^{est.buy})$ are updated.

$$(Inp)_t^{est.buy} += \tilde{q}_t^{est.buy.(s)} \times P_t^{est.sup(.s)} \times (1 + \alpha_{26.s-sup} \times \varpi) \tag{17}$$

$$(LA)_t^{est.buy} -= \tilde{q}_t^{est.buy.(s)} \times P_t^{est.sup(.s)} \times (1 + \alpha_{26.s-sup} \times \varpi) \tag{18}$$

At the same time, we update the values of sales ($(Sl)_t^{est.(sup)}$), demand for a good ($Q_t^{d_{est.(sup)}}$), liquid assets ($(LA)_t^{est.buy}$) and stock of suppliers from all sectors that provided inputs to establishments ($Inv_t^{est.sup}$) are updated:

$$(Sl)_t^{est1.(sup)} += P_t^{est.sup(.s)} \times Inv_t^{est1.sup} \quad (19)$$

$$(Sl)_t^{est2.(sup)} += P_t^{est.sup(.s)} \times (\tilde{q}_t^{est.buy.(s)} - Inv_t^{est1.sup}) \quad (20)$$

$$Q_t^{d_{est1.(sup)}} += Inv_t^{est1.sup} \quad (21)$$

$$Q_t^{d_{est2.(sup)}} += \tilde{q}_t^{est.buy.(s)} - Inv_t^{est1.sup} \quad (22)$$

$$(LA)_t^{est1.buy} += P_t^{est.sup.(s)} \times Inv_t^{est1.sup} \quad (23)$$

$$(LA)_t^{est2.(sup)} += P_t^{est.sup.(s)} \times (\tilde{q}_t^{est.buy.(s)} - Inv_t^{est1.sup}) \quad (24)$$

$$Inv_t^{est2.sup} -= (\tilde{q}_t^{est.buy.(s)} - Inv_t^{est1.sup}) \quad (25)$$

$$Inv_t^{est1.sup} = 0 \quad (26)$$

where $\tilde{q}_t^{est.buy.(s)}$ is the quantity of inputs that has been selected according to the adaptive algorithm.

Module 2: Supply Chain Short Term Credit Admissibility 1 (M.9)

In this submodule we analyse the case of an establishment without liquidity problems. The requested amount is given by the formula:

$$l_t^{SH^{est}} += \alpha_{27} \times ((LA)_t^{est.buy} - W_t^{est.buy} - P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)}) \quad (27)$$

In the future, the model could also recognize different business types, similarly to the consumer types in the model, however at this stage, access to such disaggregated data was unavailable. Firstly, it is checked whether the matched bank in the database is able to loan this quantity ($S_t^{l.SH^{est}} \geq l_t^{SH^{est}}$), as is the creditworthiness of the applicant. The values of ROA, ROE and leverage ratios as well as the value of average financial risk associated with the establishment operating in a given sector and its default history are checked. If the loan is granted, then the values of loans ($l_t^{SH^{est}}$), quarterly payments ($l_t^{SH_q^{est}}$), interest to be paid (in total) ($R_t^{SH^{ind}}$) and quarterly ($R_t^{SH_q^{ind}}$), inputs ($(Inp)_t^{est}$) and liquidity assets ($(LA)_t^{est}$) are updated (for the establishment-buyer).

$$R_t^{SH^{ind}} += \frac{1}{\kappa} \times [\alpha_{27} \times ((LA)_t^{est.buy} - W_t^{est.buy} - P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)}) \times (1 + 0.25 \times i_{lSH})^{\mathrm{M}^{SH^{est}}} + - \alpha_{27} \times ((LA)_t^{est.buy} - W_t^{est.buy} - P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)}))] \quad (28)$$

$$R_t^{SH_q^{est}} = \frac{R_t^{SH^{est}}}{\mathrm{M}^{SH^{est}}} \quad (29)$$

$$l_t^{SH_q^{est}} = \frac{l_t^{SH^{est}}}{\mathrm{M}^{SH^{est}}} \quad (30)$$

$$(Inp)_t^{est.buy} += P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)} \quad (31)$$

$$(LA)_t^{est.buy} -= P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)} - \alpha_{27} \times ((LA)_t^{est.buy} - W_t^{est.buy} - P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)}) \quad (32)$$

At the same time, the revenue of banks ($(Rev_{lSH})_t^b$) and supply of short-term credit for firms ($S_t^{l.SH^{est}}$) are updates as well as sales ($(Sl)_t^{est}$), demand ($Q_t^{d_{est}}$), liquidity assets ($(LA)_t^{est}$) and stock of all establishments from sectors ($(Inv)_t^{est}$) that provided inputs to establishments (buyers).

$$(Rev_{lSH})_t^b += R_t^{SH_q^{est}} \tag{33}$$

$$S_t^{l.SH^{est}} -= \alpha_{27} \times ((LA)_t^{est.buy} - W_t^{est.buy} - P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)}) \tag{34}$$

$$(Sl)_t^{est(.sup)} += P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)} \tag{35}$$

$$Q_t^{d_{est}} += q_t^{est.buy(.s)} \tag{36}$$

$$(LA)_t^{est(.sup)} += P_t^{est(.sup)(.s)} \times q_t^{est.buy(.s)} \tag{37}$$

$$(Inv)_t^{est(.sup)} -= q_t^{est.buy(.s)}. \tag{38}$$

Short-term loans make it possible to guarantee the solvency of establishments in everyday business transactions. Restrictions to funding provision could result in an establishment's loss of liquidity and production capacity. If the matched bank does not agree to grant credit, the same conditions are checked with other banks in the market. Firstly, the conditions are checked in the bank that offers the lowest interest rate. If none will grant the loan, the establishment needs to adjust the quantity of inputs to be purchased ($\tilde{q}_t^{est.buy}$).

Module 2: Supply Chain Short Term Credit Admissibility 2 (M.10)

The requested amount is given by the formula:

$$l_t^{SH^{est}} = q_t^{est.buy(.s)} \times P_t^{est(.sup)(.s)} \times (1 + \alpha_{26.s-sup} \times \varpi) - ((LA)_t^{est.buy} - W_t^{est.buy}). \tag{39}$$

Similar to the submodule 9, the supply conditions and creditworthiness are checked. In this case, the conditions for granting credit are also tightened, hence the differences in the parameters in the sub-modules *Bank credit admissibility 1* (M.9) and *Bank credit admissibility 2* (M.10). If a loan is not granted by a given bank, the establishment tries to obtain a loan from another bank. If there is no bank that is willing to supply a loan, the establishment is only able to purchase a portion of the planned amount of inputs. The logic of adaptive algorithms is used here. The values of variables are updated in the similar way as in the previous module (M.9). Purchases by establishments from the suppliers are supplemented by the purchases of consumers and the governmental sector.

Module 3: Household consumption - Individuals & households income (M.11) & (M.12)

Firstly, in the module, the individuals' and households' incomes (respectively y_t^{ind} and y_t^{HH}) are computed. Individual income includes income from various sources: wages, business activity, dividends, public pensions, pension benefits, pre-retirement benefits and training allowances. The model distinguishes three main categories: wage (w_t^{ind}), subsidy ($(sub)_t^{ind}$) and interest from bank savings (deposits) ($i_d \times d_t^{ind}$). An individual's income is expressed by the following formula:

$$y_t^{ind} = w_t^{ind} + (sub)_t^{ind} + i_d \times d_t^{ind}. \tag{40}$$

Household income includes the sum of individual incomes, supplemental security income, alimony, donations, property rental income, interest and dividends from savings accounts, bonds, investment funds and income earned from participation in companies in which family members were investors or

inactive partners. All items have been grouped into categories at the database level. The total income of the household (composed of $N^{ind.HH}$ individuals) is given by:

$$y_t^{HH} = \sum_{i=1}^{N^{ind.HH}} y_{it}^{ind} + (Don)_t^{HH} + (Rent)_{RE_t}^{HH}, \tag{41}$$

where $(Don)_t^{HH}$ are donations and $(Rent)_{RE_t}^{HH}$ is an additional income from renting apartments.

At the beginning of the cycle, at least two banks are assigned to each individual; the bank to which the individual entrusted their savings ($Id_{d_t}^{bank.(ind)}$) and the bank that may grant consumer loans to the individual ($Id_{l_t^C}^{bank.(ind)}$). In the model, it is assumed that individuals are less prone to change the bank to which they entrusted their saving and current account funds than to change the 'bank-lender.' If they decide to change bank, when looking for a new bank, they take into account the offered interest rates on deposits and transaction costs, that is, the costs of changing the bank and opening a new bank account. The likelihood of a bank changing in the case of a consumer loan is higher than in the case of deposits. In the case of deposits, psychological factors such as habit formation or the perception of a bank as a reputable institution, which reinforce network effects, play a greater role. In the case of mortgages, households primarily rely on the interest rate on the loan.

Module 3: Households Consumption Net Savings (M.13)

After calculating household income, savings (s_t^{HH}) are calculated. Disposable income is the household income after deducting the cost of living (h_t^{HH}), which includes the cost of renting or repaying the mortgage. If the disposable income is higher than the specified income per person in the household, where the per person income is given by the parameter α_{41}, then the savings are computed according to the following formula:

$$s_t^{HH} = max\left\{0;\ \frac{\sqrt{(Age)_t^{eldest.ind}}}{\alpha_{42}} \times \left(y_t^{HH} - h_t^{HH} - \alpha_{41} \times N_t^{ind.HH}\right)\right\}, \tag{42}$$

where α_{41} is the minimum cost of food per person in the household according to Central Statistical Office statistics, $N_t^{ind.HH}$ counts the number of individuals in the household.

However, if disposable income is lower than the specified parameter α_{38} (income per capita) multiplied by the number of consumers in the household ($N_t^{ind.HH}$), two possibilities are considered. First, when the disposable income is positive, the savings are calculated according to the following formula:

$$s_t^{HH} = max\left\{0;\ \frac{\sqrt{(Age)_t^{eldest.ind}}}{3 \times \alpha_{42}} \times \left(y_t^{HH} - h_t^{HH} - \alpha_{41} \times N_t^{ind.HH}\right)\right\}. \tag{43}$$

Household savings are redistributed between the accounts of adult household members. If the household is a single person, then all savings are transferred to his or her account. If the household is a couple or an extended one (more than two adults in the family), then the corresponding parts of the savings are transferred to the adults' bank accounts.

$$d_t^{HH} += \frac{1}{N_t^{ind.HH}} \times s_t^{HH}. \tag{44}$$

Second, if the disposable income is negative, household members spend their savings and current deposits on consumption and try to sell the properties on the real estate market. The algorithm looks for households which have sufficient funds after deducting the debt burden to pay for the property:

$$\sum_{i=1}^{N^{ind}} d_{it}^{ind} - \sum_{i=1}^{N^{ind}} \left(l_{it}^{C_q^{ind}} + R_{it}^{C_q^{ind}} + l_{it}^{H_q^{ind}} + R_{it}^{H_q^{ind}} + l_{it}^{NH_q^{ind}} + R_{it}^{NH_q^{ind}} \right) \geq P_t^{prop}. \tag{45}$$

If this kind of household is found, then the values of deposits of buyers and sellers are updated as well as the status of the ownership of buyers ($(Own)_t^{HH} = 1$).

$$d_{it}^{ind.seller} += \frac{1}{N^{ind}} P_t^{prop} \tag{46}$$

$$d_{it}^{ind.buy} -= \frac{1}{N^{ind}} P_t^{prop}. \tag{47}$$

If the household does not receive additional resources, this can eventually lead to default ($PD_t^{ind} + +$) and the household is removed from the database. The program updates the value of non-performing loans ($(NPL_{l^I})_t^b$, $(NPL_{l^C})_t^b$, $(NPL_{l^H})_t^b$, $(NPL_{l^{NH}})_t^b$). The probability of default of the bank increases.

$$(NPL_{l^I})_t^b += (pl)_t^{I^{ind}} \tag{48}$$

$$(NPL_{l^C})_t^b += (pl)_t^{C^{ind}} \tag{49}$$

$$(NPL_{l^H})_t^b += (pl)_t^{H^{ind}} \tag{50}$$

$$(NPL_{l^{NH}})_t^b += (pl)_t^{NH^{ind}}, \tag{51}$$

where $(pl)_t^{I^{ind}}$ is the sum of liabilities to the bank for outstanding investment loans that have to be repaid in the given iteration, $(pl)_t^{C^{ind}}$ is the sum of liabilities to the bank for outstanding consumer loans, $(pl)_t^{H^{ind}}$ is the sum of liabilities to the bank for outstanding housing loans and $(pl)_t^{NH^{ind}}$ is the sum of liabilities to the bank for outstanding non-housing loans.

The bank may become a new temporary owner of the property if the household was removed from the database after the default. The property is marked for sale. After selling, we update the value of bank's revenues:

$$(Rev_{l^H})_t^b += P_t^{prop}. \tag{52}$$

Module 3: Households Consumption Consumer Loans Update (M.14)

In this submodule, the desired consumption of goods in the current period in a particular industry by the household is determined. Households can finance their consumption entirely by their own means or apply for consumer credit. The basic amount of good purchased from the industry ($Q_{basic}^{HH.(cons)}$) is given by:

$$Q_{basic}^{HH.(cons)} = \alpha_{43.tc.s} \times \frac{\sum_{i=1}^{N^{ind.HH}} d_{it}^{ind} - \sum_{i=1}^{N^{ind.HH}} (l_{it}^{C_q^{ind}} + R_{it}^{C_q^{ind}} + l_{it}^{H_q^{ind}} + R_{it}^{H_q^{ind}} + l_{it}^{NH_q^{ind}} + R_{it}^{NH_q^{ind}})}{P_t^s \times (1 + t_{VAT}^s)}, \tag{53}$$

where parameter value $\alpha_{43.tc.s}$ is specific to industry and customer type. The parameter expresses the percentage of total consumption that is, household purchases from all industries. When buying from several industries, we assume $\sum_{s=1}^{S} \alpha_{43.tc.s} = 1$, where S is the number of industries. If the household consume only the quantity $Q_t^{HH.(cons)} = Q^{HH.(cons)_{basic}}$ we proceed to the *Supplier searching module*.

In the case of taking a loan, consumption funds are increased by the amount of the loan less debt burden and service. The loan can only be granted if the following basic condition is met:

$$\left\{ y_t^{HH} - h_t^{HH} - \sum_{i=1}^{N^{ind.HH}} (l_{it}^{C_q^{ind}} + R_{it}^{C_q^{ind}} + l_{it}^{H_q^{ind}} + R_{it}^{H_q^{ind}} + l_{it}^{NH_q^{ind}} + R_{it}^{NH_q^{ind}}) \right\} \geq \alpha_{44} \times N_t^{ind.HH}. \quad (54)$$

In this case the quantity of loans is given by the formula:

$$l_{temp}^{C^{ind}} = \alpha_{45.tc.s} \times \left\{ y_t^{HH} - h_t^{HH} - \sum_{i=1}^{N^{ind.HH}} (l_{it}^{C_q^{ind}} + R_{it}^{C_q^{ind}} + l_{it}^{H_q^{ind}} + R_{it}^{H_q^{ind}} + l_{it}^{NH_q^{ind}} + R_{it}^{NH_q^{ind}}) \right\} + \alpha_{46.s}. \quad (55)$$

The parameter value $\alpha_{45.tc.s}$ is specific to industry and customer type, while $\alpha_{46.s}$ is industry-specific. With a certain probability (pr_2), an individual tries to obtain the quantity in the matched bank in the database ($Id_{l_t^C}^{bank.(ind)}$). In this case, the supply from the bank and creditworthiness are checked in the *Supply side checking 1* (M.15) and *Consumer credit admissibility 1* (M.17). The individual can also try to obtain the loan from other banks. In such cases, interest rates are compared, and *Supply side checking 2* (M.16) and *Consumer credit admissibility 2* (M.18) are proceeded to.

Module 3: Households Consumption Supply Side Checking 1 & 2 (M.15) & (M.16)

In the *Supply side checking 1* (M.15) submodule it is checked whether the bank assigned to the household has sufficient funds to grant the loan ($S_t^{l.C^{ind}} \geq l_t^{C^{ind}}$), whether the bank's policy will allow another loan to be granted, and whether the regulatory requirement for sectoral exposures is met ($l_t^{C^{ind}} < \alpha_{47}$). In the *Supply side checking 2* (M.16) submodule, we check whether any bank selected from the list of banks offering consumer loans at a specified interest rate has sufficient credit supply and that the regulatory requirements for sectoral exposures are met. If none is able to give this amount, the amount of loan is adjusted using adaptive algorithm ($\tilde{l}_t^{C^{ind}}$).

Module 3: Households Consumption Consumer Credit Admissibility 1 & 2 (M.17) & (M.18)

In submodules M.17 & M.18, household creditworthiness is checked with the bank assigned to the household or other bank (from the list of banks) selected in the *Supply side checking 2* (M.18) submodule. The first condition relates to the level of income per person after deduction of repayments of other loans. This level is specific to each bank:

$$\left\{ y_t^{HH} - h_t^{HH}{}_t - (\sum_{i=1}^{N^{ind.HH}} \left(l_{it}^{C_q^{ind}} + R_{it}^{C_q^{ind}} + l_{it}^{H_q^{ind}} + R_{it}^{H_q^{ind}} + l_{it}^{NH_q^{ind}} + R_{it}^{NH_q^{ind}} \right)) \right\} \geq \alpha_{50.b} \times N_t^{ind.HH}. \quad (56)$$

The next conditions relate to credit history of the household (the probability of default of members of the family) ($(PD)_t^{ind.HH} \leq \alpha_{49.b}$), and the maximum number of loans that can be granted to one household. If all conditions are met, the loan can be granted. For all loans granted to individuals and households, the applicant age ($(Age)_t^{ind} \geq 18$) and status of the labour market ($\Xi_t^{ind} = \{3||5\}$) are also checked. It is also possible to include a variable which counts the elapsed time since the last change in status on the labour market. If the individual works less than 4 quarters and is under 30, it is assumed that they have a fixed-term contract and no creditworthiness.

Module 3: Households Consumption Consumer Credit and Purchase After Passing Supply Side Conditions and Credit Admissibility (M.19)

In this submodule, maturity is assigned to the loan ($M^{C^{ind}}$) depending on the amount of loan granted. Next, the value of debt service is updated, taking into account the civil status of the household members.

$$l_t^{C^{ind}} = \frac{1}{N_t^{adults.HH}} \times l_{temp}^{C^{ind}} \tag{57}$$

$$l_t^{C_q^{ind}} = \frac{1}{N_t^{adults.HH}} \times \frac{l_{temp}^{C^{ind}}}{M^{C^{ind}}} \tag{58}$$

$$R_t^{C^{ind}} += \frac{1}{N_t^{adults.HH}} \times l_{temp}^{C^{ind}} \times (1 + i_{l^C})^{M^{C^{ind}}} - l_{temp}^{C^{ind}} \tag{59}$$

$$R_t^{C_q^{ind}} = \frac{R_t^{C^{ind}}}{M^{C^{ind}}}. \tag{60}$$

The amount of credit granted by the bank, the supply of the credit and the revenues of the bank in the given period are also increased.

$$S_t^{l.C^{ind}} -= l_{temp}^{C^{ind}} \tag{61}$$

$$(Rev_{l^C})_t^b += R_t^{C_q^{ind}}. \tag{62}$$

Finally, the amount of consumer goods to be purchased from different industries is computed.

$$Q_t^{HH.(cons)} = Q_{basic}^{HH.(cons)} + \frac{\alpha_{43.tc.s} \times l_t^{C^{ind}}}{M^{C^{ind}} \times \{P_t^s \times (1 + t_{VAT})\}}. \tag{63}$$

Module 3: Households Consumption Supplier Searching & Purchase.hh (M.20)

The household next chooses the supplier. If the current supplier has sufficient stock of good ($Inv_t^{est.(sup)}$) and the ratio of quality to price is higher than average ratio in the sector, then the household purchases the goods ($Q_t^{HH.(cons)}$) from this supplier.

$$Inv_t^{est.(sup)} \geq Q_t^{HH.(cons)} \;\&\&\; \frac{\alpha_{22.s-sup} \times (Ql)_t^{est.(sup)}}{\alpha_{23.s-sup} \times P_t^{est(.sup)}} > \frac{\alpha_{24.s-sup} \times (Ql)_t^s}{\alpha_{25.s-sup} \times P_t^s}. \tag{64}$$

Otherwise, the household seeks a new supplier from incrementally more distant spatial locations. The requirements for the same spatial codes are loosened sequentially. Later, in the Purchase.hhs submodule, the profits from sales ($(Sl)_t^{est.(sup)}$), demand ($Q_t^{d_{est.(sup)}}$), stock ($Inv_t^{est.(sup)}$), liquid assets ($(LA)_t^{est.(sup)}$) of suppliers are updated. In addition, deposits of consumers are updated.

$$(Sl)_t^{est.(sup)} += P_t^{est(.\mathrm{sup})} \times Q_t^{HH.(cons)} \tag{65}$$

$$Q_t^{d_{est.(sup)}} += Q_t^{HH.(cons)} \tag{66}$$

$$Inv_t^{est.(sup)} -= Q_t^{HH.(cons)} \tag{67}$$

$$(LA)_t^{est.(sup)} += P_t^{est(.sup)} \times Q_t^{HH.(cons)} \tag{68}$$

$$d_{it}^{ind} -= \frac{1}{N_t^{adults.HH}} P_t^{est(.sup)} \times Q_t^{HH.(cons)}. \tag{69}$$

Module 4: Public Contracts (M.21)

In this module, we complement the demand of the private sector with demand from the public sector. Public contracts are usually signed by large companies. A public contract is awarded when the product price is lower than the average price of the product in the sector ($P_t^{est} \leq P_t^s$). In addition, the probability of signing contracts increases with the size of the business and the quality of the product. If the terms of the contract are met, the value of the stored goods (Inv_t^{est}), the demand for good ($Q_t^{d_{est.(sup)}}$), liquid assets of the supplier $(LA)_t^{est}$ and the value of sales $(Sl)_t^{est}$ are updated. The stock cannot be lower than the minimum fraction of production (Y_t^{est}) given by the parameter α_{57}.

$$Inv_t^{est.(sup)} - = max\{0; min\left(Inv_t^{est.(sup)},\ \alpha_{57} \times Y_t^{est.(sup)}\right)\} \tag{70}$$

$$Q_t^{d_{est.(sup)}} + = min\left(Inv_t^{est.(sup)},\ \alpha_{57} \times Y_t^{est.(sup)}\right) \tag{71}$$

$$(LA)_t^{est.(sup)} + = \alpha_{58} \times min\left(Inv_t^{est.(sup)},\ \alpha_{57} \times Y_t^{est.(sup)}\right) \tag{72}$$

$$(Sl)_t^{est.(sup)} + = \alpha_{58} \times min\left(Inv_t^{est.(sup)},\ \alpha_{57} \times Y_t^{est.(sup)}\right). \tag{73}$$

Module 5: Households Mobility Accommodation Cost and Housing Stress (M.22)

The term property refers to the value of apartment or houses, with or without land. Households live in properties they own ($(Own)_t^{HH} = 1$) or rent property from other households ($(Own)_t^{HH} = 0$). The household may own more than one property. Real estate may be subject to residential $\left(l_t^{H^{ind}}\right)$ and non-residential loans ($l_t^{NH^{ind}}$). If the household lives in their own property, then the cost of living is equal to the sum of the financial obligations of the owners of the building, that is, the adult individuals forming the household who bought the property:

$$h_t^{HH} = \beta_0 + \beta_1 \times \{\sum_{i=1}^{N^{ind.HH}} (l_t^{H_q^{ind.HH}} + R_t^{H_q^{ind}})\}, \tag{74}$$

where parameters β_0 and β_1 adjust the fixed and the variable parts of accommodation cost.

The cost of renting is calculated as a part β_1 of the sum of liabilities to the bank (loans) and the part β_4 of the cost of rent, calculated as the ratio of the price of the property to the number of households that rent this property.

$$h_t^{HH} = \beta_1 \times \{\sum_{i=1}^{N^{ind.HH}} (l_t^{H_q^{ind.HH}} + R_t^{H_q^{ind}})\} + \beta_4 \times \frac{P_t^{prop}}{(\#Rent)_t^{prop}} \tag{75}$$

In the first case of ownership, if the rental cost is greater than the specified parameter β_2 (part of the income y_t^{HH}), the household decides to sell the property. If the property was already marked for sale ($(ForSale)_t^{prop} = 1$), then the price (P_t^{prop}) has to be decreased by a percentage β_3. In practice, the parameter β_3 reflects how much the price has to be lowered in order to sell the property in next iteration. In the second case, if a household rents a property and the cost of rent is too high, it will start looking for a new home. The household looks for another building considering the status on the labour market ($\Xi_t^{int} = \{3||5\}$) and the age of the household members ($(Age)_t^{ind} \geq 18$). If two adults in the household are working, the program randomly selects one of them and searches for a building near the person's workplace (the algorithm checks and compares the spatial codes: $\vartheta_t^{prop^1}$, $\vartheta_t^{prop^2}$, $\vartheta_t^{prop^3}$, $\vartheta_t^{prop^4}$, $\vartheta_t^{est^1}$, $\vartheta_t^{est^2}$, $\vartheta_t^{est^3}$, $\vartheta_t^{est^4}$). In addition to property locations, households take into account the price of the property from the appropriate price range and whether the building is for sale. If more than one property meets the criteria, one is selected at random and the spatial codes attributing the individual to the respective spatial units are updated.

If two adults are not working ($\Xi_t^{int} = \{1||2\}$) and the difference between the sum of individual deposits and the sum of the individual liabilities of family members is less than the subsistence level expressed by the parameter β_7, then the individuals are removed from the database, the household is removed from the database, and the corresponding records from the object *Consumers* are deleted as well.

$$\sum_{i=1}^{N^{ind.HH}} d_{it}^{ind.HH} - \sum_{i=1}^{N^{ind.HH}} (l_t^{H_q^{ind}} + R_t^{H_q^{ind}})\} + \beta_4 \times \frac{P_t^{prop}}{(\#Rent)_t^{prop}} < \beta_7 \tag{76}$$

The exception is a situation in which an adult is under 25. Then the assumption applies that they are still living with their parents. If individuals and households are removed from the database, all records related to insolvency are updated. Consequently, the non-performing loans for a given bank and sectors are increased. The probability of bank's default increases as well. Especially,

$$(NPL_{l^I})_t^b += (pl)_t^{I^{ind}} \tag{77}$$

$$(NPL_{l^C})_t^b += (pl)_t^{C^{ind}} \tag{78}$$

$$(NPL_{l^H})_t^b += (pl)_t^{H^{ind}} \tag{79}$$

$$(NPL_{l^{NH}})_t^b += (pl)_t^{NH^{ind}} \tag{80}$$

where $(pl)_t$ is the sum of liabilities to the bank for outstanding (respectively investment, consumer, housing and non-housing) loans that have to be repaid in the given iteration.

Module 5: Households Mobility Profits from Rent (Accommodation & Housing Stress) (M.23)

In this module, property attributes are updated if the household obtains profits from renting the property. The algorithm checks all household properties. The primary property ($(PH)_t^{prop} = 1$) cannot be sublet to another household. If the household has a second property ($(PH)_t^{prop} = 0$), the number of households that live in the house is checked. If none live there, it is marked for sale ($(ForSale)_t^{prop} = 1$). If it was previously marked for sale ($(ForSale)_{t-1}^{prop} = 1$), its price is reduced by a certain percentage of the value that is specified by the parameter β_3.

$$P_t^{prop} = \beta_3 \times P_{t-1}^{prop}. \tag{81}$$

Revenues from renting second property ($(Rent)_{RE_t}^{HH}$) are updated according to:

$$(Rent)_{RE_t}^{HH} += \beta_4 \times P_t^{prop}. \tag{82}$$

Module 5: Households Mobility - Decisions About Funding Housing and Non-Housing Purchase (M.24)

In this sub-module the household decides how to finance the purchase of houses and other non-housing purchases. If the current funds and savings in the bank account are greater than the price of the cheapest property on the market that has already been marked for sale ($(ForSale)_{t-1}^{prop} = 1$), it can potentially buy a property in cash.

$$\sum_{i=1}^{N^{ind.HH}} d_{it}^{ind.HH} - \sum_{i=1}^{N^{ind.HH}} (l_t^{H_q^{ind}} + R_t^{H_q^{ind}})\} > P_{min_t}^{prop}. \tag{83}$$

In practice, we assume that a household buys a property in cash, only with a given probability (pr_3) in the model. If a member of the household is an individual who owns the firm ($(Entr)_t^{ind} = 1$) or is unemployed ($\Xi_t^{ind} = 2$) with a high entrepreneurial spirit ($(EntrS)_t^{ind} > 0.5$), then the funds will be first invested in the firm rather than housing or non-housing purchases. We update the status of

individual as an entrepreneur ($(Entr)_t^{ind} = 1$). If a household does not have enough resources to buy a property in cash, it will apply for a loan and the sub-module *Macroprudential ratios* is moved to.

Module 5: Households Mobility Macroprudential Ratios (M.25)

In this submodule, macroprudential ratios are computed. In addition, the total indebtedness, debt servicing and total assets held by the household are computed as a temporal variables. The following macroprudential ratios were included in the system for residential and non-residential loans: total debt to assets ratio, debt to income ratio, debt service to income ratio, loan-to-value ratio.

$$(IndebtQ)_{temp}^{HH} = \sum_{i=1}^{N^{ind.HH}} (l_{it}^{C_q^{ind}} + l_{it}^{H_q^{ind}} + l_{it}^{NH_q^{ind}} + l_{it}^{I_q^{ind}}) \tag{84}$$

$$(DServiceQ)_{temp}^{HH} = \sum_{i=1}^{N^{ind.HH}} (R_{it}^{C_q^{ind}} + R_{it}^{H_q^{ind}} + R_{it}^{NH_q^{ind}} + R_{it}^{I_q^{ind}}) \tag{85}$$

$$(AssetsQ)_{temp}^{HH} = \sum_{i=1}^{N^{ind.HH}} d_{it}^{ind.HH} \tag{86}$$

$$(DTA)_t^{HH} = \frac{(IndebtQ)_{temp}^{HH}}{(AssetsQ)_{temp}^{HH}} \tag{87}$$

$$(DTI)_t^{HH} = \frac{(IndebtQ)_{temp}^{HH}}{y_t^{HH}} \tag{88}$$

$$(DSTI)_t^{HH} = \frac{(DServiceQ)_{temp}^{HH}}{y_t^{HH}} \tag{89}$$

$$(LTV)_t^{HH} = \frac{\sum_{i=1}^{N^{ind.HH}} l_{it}^{H^{ind}}}{P_t^{prop}} \ . \tag{90}$$

Module 5: Households Mobility Housing and Non-Housing Loans (M.26)

In this sub-module, the household's creditworthiness and supply side conditions are checked. Firstly, basic requirements are checked. If the household debt ratios exceed regulatory requirements, the household lacks creditworthiness ($(DSTI)_t^{HH} > \beta_{8.b} || \ (DTI)_t^{HH} > \beta_{9.b} || \ (DTA)_t^{HH} > \beta_{10.b}$). Likewise, if all household members study, they are inactive on the labour market or are living on social benefits ($\Xi_t^{ind} = \{1||2||4\}$). The model also takes into account the situation of people under the age of 30 who work less than 1 year in a company that also do not have creditworthiness ($(\Xi_t^{ind} = 3 \ \&\& \ (Age)_t^{ind} < 30 \ \&\& \ \Psi_t^{ind} \leq 4) \ || \ (\Lambda_t^{ind} = 2 \ \&\& \ \Xi_t^{ind} = (1||2||4) \ \&\& \ \Xi_t^{ind} = 3 \ \&\& \ (Age)_t^{ind} < 30 \ \&\& \ \Psi_t^{ind} \leq 4)$). Then the value of collateral is estimated. For the all properties that the household owns, the property prices are summed up ($(Collat)_{temp}^{max} += P_t^{prop}$). Households may apply for a residential or non-residential loan under a pledge of real estate. If the household owns at least one property, it will apply for a residential loan with a given probability, and for a non-residential loan with one minus that probability. On the other hand, if a household has rented a property (i.e., is not an owner), it will first apply for a residential loan. In most cases, the first property is bought in cash. The value of the mortgage loan is equal to the difference between the price of the cheapest real estate on the market, and the savings (less the charges for other loans); this is according to the following equation:

$$l_{temp}^{H^{HH}} = P_{min_t}^{prop} - \beta_{15} \times (\sum_{i=1}^{N^{ind.HH}} d_{it}^{ind.HH} - (Debt)_{temp}^{HH} - (DebtServ)_{temp}^{HH}) \tag{91}$$

where $(Debt)_{temp}^{HH}$, $(DebtServ)_{temp}^{HH}$ are given by the expressions:

$$(Debt)_{temp}^{HH} = \sum_{i=1}^{N^{ind.HH}} \left((pl)_{it}^{C_q^{ind}} + (pl)_{it}^{H_q^{ind}} + (pl)_{it}^{NH_q^{ind}} + (pl)_{it}^{I_q^{ind}} \right) \quad (92)$$

$$(DebtServ)_{temp}^{HH} = \sum_{i=1}^{N^{ind.HH}} \left(R_{it}^{C_q^{ind}} + R_{it}^{H_q^{ind}} + R_{it}^{NH_q^{ind}} + R_{it}^{I_q^{ind}} \right) \quad (93)$$

Households choose a bank with which to apply for a loan, taking into account the interest rate. At the same time, the bank must have sufficient funds and be willing to accept the LTV ratio of applicant ($LTV_{temp}^{HH} = ((l_t^{H^{ind}} + l_t^{NH^{ind}})/P_t^{prop})$). If the household does not obtain credit in the bank that offers the lowest interest rate, it resubmits the request to another bank on the list that requires a higher interest rate. After receiving the loan, the household buys the property. The value of housing loans granted by the bank is also updated. The value of the household loans obtained is then updated. If the household consists of a marriage, the value of granted loans is halved. If the property is the main residence, the cost of accommodation is also updated. In the case of non-residential loans, the scheme works in an analogous way, with the value of the residential loan being equal to:

$$l_{temp}^{NH^{HH}} = \beta_{13.tc} \times \{P_{min_t}^{prop} - \beta_{15} \times (\sum_{i=1}^{N^{ind.HH}} d_{it}^{ind.HH} - (Debt)_{temp}^{HH} - (DebtServ)_{temp}^{HH})\}. \quad (94)$$

After checking the creditworthiness and credit supply of the selected bank, the value of non-residential loans obtained are updated, taking into account the civil status of household members.

Module 6: Individuals' Records Updating

Individuals' records updating module ensures the maintenance of demographic trends observed empirically in the simulation. The heterogeneity of individuals has been taken into account, as has population dynamics.

Module 6: Individuals' Records Updating Inheritor or Life (M.27)

In the first sub-module the individuals may die with probabilities ($\rho_{age.gender}$) depending on age ($(Age)_t^{ind}$) and gender ($G^{ind} = \{0||1\}$). If a person dies, after the submodules *Inheritor* (M.28) and *Inheritance* (M.29 & M.30) have been applied, the individual is deleted from the database. This insolvency has a direct impact on the banks in the form of an increase in non-performing loans. In the case of survival of an individual, the program increases the age of the person ($(Age)_t^{ind}++$) and the period since the last change of status on the labour market ($\Psi_t^{ind}++$) and continues in the sub-module *Updating consumer type* (M.31).

Module 6: Individuals' Records Updating Inheritor (M.28)

In this submodule, the heir in the event of death of an individual is determined. First, the age of the deceased person is checked. If the deceased person was an adult ($(Age)_t^{ind} \geq 18$), inheritance may be considered. Then, all members of the households and adults in the households are counted. If an adult over the age of 18 has died, who does not have family, the labor status of this individual is checked (Ξ_t^{ind}). If this individual was an entrepreneur ($\Xi_t^{ind} = 5$), the number of employees (L_t^{firm}), in the company is checked. If it was a sole proprietorship, then the firm is for sale ($(ForSale)_t^{firm} = 1$). Otherwise, one of workers in their company who earned the highest wage in the previous period is selected. If the deceased person was not an entrepreneur, then the algorithm selects an inheritor at random from the group of working adults ($\Xi_t^{ind} \neq 1$ && $(Age)_t^{ind} \geq 18$). If a child dies ($(Age)_t^{ind} \leq 18$), they do not leave material property, rather the consumer type of the child's parents changes. If the parents have no more children and if all household members are over 67 years old then we update

the consumer type ($(ConsT)_t^{cons} = 3$). If both parents were under 67, they are also updated to another consumer type ($(ConsT)_t^{cons} = 4$). If the family consists of more than two adults then the consumer type is also updated ($(ConsT)_t^{cons} = 6$). If the deceased adult person has a family, then two cases must be differentiated. If they had children, the eldest person in the family is the inheritor and the consumer type does not change ($(ConsT)_t^{cons} = \{5||6\}$). If there were no children, then the spouse is the inheritor and the consumer type changes depending on the age of the spouse ($(ConsT)_t^{cons} = \{1||2\}$). If the deceased adult was a single parent, then the adoption of a child is considered in the model in the module *Adoption*. The algorithm looks for a new couple or a single individual to be a parent and sum up deposits, ownership of properties and firms. If they own more than two properties, they are designated for sale ($(ForSale)_t^{firm} = 1$). We remove the deceased person from the database. We continue to the *Inheritance* modules (M.29 & M.30).

Module 6: Individuals' Records Updating Inheritance: Deposits & Firms (M.29)

In this submodule, we pass the deposits and firms to the inheritor. Firstly, savings after deductions of housing loans pending to be paid are given to the inheritor. If these savings are greater than zero ($d_t^{dec.ind} - \left((pl)_t^{H^{ind}} + (pl)_t^{NH^{ind}}\right) > 0$), then the deposits of the heir are updated.

$$d_t^{heir.ind} += (1 - t_{inh}) \times \left(d_t^{dec.ind} - (pl)_t^{H^{ind}} - (pl)_t^{NH^{ind}}\right) \tag{95}$$

The parameter t_{inh} stands for taxes that must be paid. In the model we assumed that consumer loans are not inherited and hence the inflow of non-performing loans of banks ($(NPL_{lC})_t^b$) and the probability of default of bank ($(PD)_t^b$) are updated.

$$(NPL_{lC})_t^b += (pl)_t^{C^{ind}} \tag{96}$$

If individual debts are greater than savings, then the variables are updated separately. In this way it is possible for the heir to inherit the debt from the loan taken in the past.

$$d_t^{heir.ind} += (1 - t_{inh}) \times d_t^{dec.ind} \tag{97}$$
$$l_t^{H^{heir.ind}} += (pl)_t^{H^{dec.ind}} \tag{98}$$
$$(pl)_t^{H^{heir.ind}} += (pl)_t^{H^{dec.ind}} \tag{99}$$
$$l_t^{H_q^{heir.ind}} += l_t^{H_q^{dec.ind}} \tag{100}$$
$$R_t^{H^{heir.ind}} += R_t^{H^{dec.ind}} \tag{101}$$
$$R_t^{H_q^{heir.ind}} += R_t^{H_q^{dec.ind}} \tag{102}$$
$$l_t^{NH^{heir.ind}} += (pl)_t^{NH^{dec.ind}} \tag{103}$$
$$(pl)_t^{NH^{heir.ind}} += (pl)_t^{NH^{dec.ind}} \tag{104}$$
$$l_t^{NH_q^{heir.ind}} += l_t^{NH_q^{dec.ind}} \tag{105}$$
$$R_t^{NH^{heir.ind}} += R_t^{NH^{dec.ind}} \tag{106}$$
$$R_t^{NH_q^{heir.ind}} += R_t^{NH_q^{dec.ind}} \tag{107}$$
$$(NPL_{lC})_t^b += (pl)_t^{C^{ind}}. \tag{108}$$

As the non-performing consumer loans increase, the probability of default of bank ($(PD)_t^b$) increases as well. If the deceased person was not the owner of the firm, then we update savings (deposits) only.

Otherwise, the inheritor may or may not run the business in the future. The decision on running a business depends on their previous earnings relative to the potential income from business activity.

$$w_t^{inh.ind} + (sub)_t^{inh.ind} < rndm(0,1) \times w_t^{dec.ind} \tag{109}$$

If the inheritor decides to run the business, the status of entrepreneur ($(Entr)_t^{ind} = 1$), the labor status ($\Xi_t^{ind} = 5$) and the number of periods since the last change in labor status ($\Psi_t^{ind} = 0$) are updated. The public assistance is ceased ($(sub)_t^{ind} = 0$). The inheritor is responsible for paying back the investment loans.

$$l_t^{heir.ind} += (pl)_t^{dec.ind} \tag{110}$$

$$(pl)_t^{heir.ind} += (pl)_t^{dec.ind} \tag{111}$$

$$l_t^{heir.ind}{}_q += l_t^{dec.ind}{}_q \tag{112}$$

$$R_t^{heir.ind} += R_t^{dec.ind} \tag{113}$$

$$R_t^{heir.ind}{}_q += R_t^{dec.ind}{}_q \tag{114}$$

If the inheritor decides to sell, the company is marked for sale ($(ForSale)_t^{firm} = 1$). The inflow of non-performing loans is increased by the amount of investment loans that will not be paid. The probability of default of bank ($(PD)_t^b$) is also increased in that case.

$$(NPL_{l^I})_t^b += (pl)_t^{I^{ind}}. \tag{115}$$

Module 6: Individuals' Records Updating Inheritance: Properties (M.30)

In this module, the number of adults who were in the household of deceased person is checked, as is whether the deceased person rented ($(Own)_t^{HH} = 0$) or owned their residence ($(Own)_t^{HH} = 1$). If the the deceased individual rented his residence, then the number of households that rents this property needs to decrease ($(\#Rent)_t^{prop} --$). If the deceased person owned their residence and had no mortgage, then the property is marked for sale ($(ForSale)_t^{firm} = 1$). However, if the deceased person had a housing loan to buy the property, the bank receives the property and attempts to sell it ($(ForSale)_t^{firm} = 1$). The bank looks for any household which sum of deposits ($\sum_{i=1}^{N^{ind.HH}} d_{it}^{ind.HH}$) is greater than the property price (P_t^{prop}). If it is found, then the bank updates the inflow of non-performing loans ($(NPL_{l^H})_t^b$, $(NPL_{l^{NH}})_t^b$) and the revenue from selling the property ($(Rev_{l^H})_t^b$).

$$(NPL_{l^H})_t^b += (pl)_t^{H^{dec.ind}} \tag{116}$$

$$(NPL_{l^{NH}})_t^b += (pl)_t^{NH^{dec.ind}} \tag{117}$$

$$(Rev_{l^H})_t^b += (1 - \beta_{22}) \times P_t^{prop} \tag{118}$$

where β_{22} expresses the transaction costs.

The new owners of the property proportionally update the deposits ($d_t^{ind.HH}$) after the purchase.

$$d_t^{ind.HH} -= \frac{1}{N^{ind.HH}} \times (1 - \beta_{22}) \times P_t^{prop} \tag{119}$$

If the bank is not able to sell the property, the price (P_t^{prop}) is lowered gradually:

$$P_t^{prop} = 0.95 \times P_{t-1}^{prop}. \tag{120}$$

Module 6: Individuals' Records Updating Updating Consumer Type (M.31)

In this module the consumer type of the household is updated. If the individual reaches the age of 67, then the consumer type of the household the individual belongs to is updated (ie. $(ConsT)_t^{cons}$=1 if the individual was single, while $(ConsT)_t^{cons} = 3$ if was married). If the individual become an adult ($(Age)_t^{ind} = 18$), then a new household is created. The individual continues to live in the same property. The accommodation cost and the income are computed according to the formulas expressed in the previous modules (M.11, M.12 & M.22).

Module 6: Individuals' Records Updating Education Level (M.32)

In this submodule the level of education is updated. If an individual continues to study ($(EducP)_t^{ind} + +$) and exceeds the number of periods that is needed to complete a particular level of education (i.e., primary school, secondary school, college, university degree, doctoral school), then a variable describing their education level ($(Educ)_t^{ind} + +$) is updated.

Module 6: Individuals' Records Updating Continue Education (M.33)

The individuals continue education with a probability depending on age and gender ($pr_{age.gender}$). According to this probability, the completed level of education ($(Educ)_t^{ind}$), labour market status (Ξ_t^{ind}) of an individual and the number of periods that have passed since the last change of status in the labour market ($\Psi_t^{ind} + +$) are updated. In this module the labor status of individuals that continue education is set to one (Ξ_t^{ind}=1).

Module 6: Individuals' Records Updating Divorces (M.34)

In this submodule, individuals may divorce. The algorithm checks the civil status of individuals (Λ_t^{ind}). The probability of divorce depends on the age of the spouses ($pr_{40.age}$). After the divorce, a new household is created for ex-husband and the consumer types ($(ConsT)_t^{cons}$) are adjusted. In the case of divorce, children always stay with the mother. The ex-husband pays the alimony ($(Don)_t^{(husb).HH}$) which depends on his earnings (w_t^{ind}) and the number of children ($(\#child)_{temp}$).

$$(Don)_t^{(wife).HH} + = \beta_{39} \times w_t^{ind} \times (\#child)_{temp} \quad (121)$$

$$(Don)_t^{(husb).HH} - = \beta_{39} \times w_t^{ind} \times (\#child)_{temp} \quad (122)$$

Both adults remain in the same property hence they both contribute to rent to cover the accommodation cost. The ex-wife is the owner, while the ex-husband lives there temporary. If the marriage had more than two properties, then each of them stays in one of them and the rest is for sale ($(ForSale)_t^{prop} = 1$). Their status of the owner ($(Own)_t^{HH} = 1$), the variable principal housing ($(PH)_t^{prop} = 1$) and the number of families that rent the property ($(\#Rent)_t^{prop} - -$) are updated. If the property was purchased by credit, then the amount of loans pending to be paid by ex-wife has to be updated as well.

$$l_t^{H^{(wife).ind}} += l_t^{H^{(husb).ind}},\ l_t^{H^{(husb).ind}} = 0 \tag{123}$$

$$(pl)_t^{H^{(wife).ind}} += (pl)_t^{H^{(husb).ind}},\ (pl)_t^{H^{(husb).ind}} = 0 \tag{124}$$

$$l_t^{H_q^{(wife).ind}} += l_t^{H_q^{(husb).ind}},\ l_t^{H_q^{(husb).ind}} = 0 \tag{125}$$

$$R_t^{H^{(wife).ind}} += R_t^{H^{(husb).ind}},\ R_t^{H^{(husb).ind}} = 0 \tag{126}$$

$$R_t^{H_q^{(wife).ind}} += R_t^{H_q^{(husb).ind}},\ R_t^{H_q^{(husb).ind}} = 0 \tag{127}$$

$$l_t^{NH^{(wife).ind}} += l_t^{NH^{(husb).ind}},\ l_t^{NH^{(husb).ind}} = 0 \tag{128}$$

$$l_t^{NH^{(wife).ind}} += l_t^{NH^{(husb).ind}},\ l_t^{NH^{(husb).ind}} = 0 \tag{129}$$

$$(pl)_t^{NH^{(wife).ind}} += (pl)_t^{NH^{(husb).ind}},\ (pl)_t^{NH^{(husb).ind}} = 0 \tag{130}$$

$$l_t^{NH_q^{(wife).ind}} += l_t^{NH_q^{(husb).ind}},\ l_t^{NH_q^{(husb).ind}} = 0 \tag{131}$$

$$R_t^{NH^{(wife).ind}} += R_t^{NH^{(husb).ind}},\ R_t^{NH^{(husb).ind}} = 0 \tag{132}$$

$$R_t^{NH_q^{(wife).ind}} += R_t^{NH_q^{(husb).ind}},\ R_t^{NH_q^{(husb).ind}} = 0. \tag{133}$$

Module 6: Individuals' Records Updating Marriages (M.35) & Births (M.36)

Each single adult can get married in the model. The probability of getting married depends on civil status, age and gender ($pr_{civ.st.age.gend}$). In the model, the age difference between partners should not exceed 10 years ($|(Age)_t^{ind1} - (Age)_t^{ind2}| < 10$ && $G^{ind1} \neq G^{ind2}$ && $(Age)_t^{ind1}, (Age)_t^{ind2} > 18$, $\Lambda_t^{ind} \neq 2$). In the case of marriage, the marital status of individuals ($\Lambda_t^{ind} = 2$) as well as the codes identifying the households ($(Id)^{HH}$) are updated. In addition, the total number of properties and their ownership ($(Own)_t^{HH}$) are updated. One of properties is marked as the principal residence ($(PH)_t^{prop} = 1$). If the marriage owns more than two properties, they are marked for sale ($(ForSale)_t^{prop} = 1$). The algorithm looks for a household-buyer who has sufficient deposits to pay for the property.

$$\sum_{i=1}^{N^{ind.HH}} d_{it}^{ind.HH} - \sum_{i=1}^{N^{ind.HH}} (l_t^{H_q^{ind.HH}} + R_t^{H_q^{ind}}) > P_t^{prop}. \tag{134}$$

When the buyer is found, then deposits are reduced by the amount equal to the sum of prices of J properties.

$$d_{it}^{buy.ind.HH} -= \frac{1}{N^{ind.HH}} \times \sum_{j=1}^{\mathrm{J}} P_{jt}^{prop} \tag{135}$$

If the buyer is not found, the price is decreased by 10%. After selling the properties, the marriage can pay back the housing and non-housing loans. If housing loans pending to be paid are greater than the revenue from selling the properties, then the algorithm decrease the amount of housing loans to be paid.

$$(pl)_t^{H^{seller.ind}} -= \frac{1}{N^{ind.HH}} \times \sum_{j=1}^{\mathrm{J}} P_{jt}^{prop}. \tag{136}$$

Otherwise, we check whether the individual could also pay back the non-housing loans. In that case, if the non-housing loans pending to be paid are greater than the difference between the revenue from selling the properties and housing loans pending to be paid, then the non-housing loans are decreased.

$$(pl)_t^{NH^{seller.ind}} - = \frac{1}{N^{ind.HH}} \times \sum_{j=1}^{J} P_{jt}^{prop} - (pl)_t^{H^{seller.ind}}. \tag{137}$$

Then, housing loans pending to be paid are set to zero. If the price for which the properties were sold was greater than the sum of all liabilities then, deposits are updated.

$$d_{it}^{seller.ind.HH} + = \frac{1}{N^{ind.HH}} \times \sum_{j=1}^{J} P_{jt}^{prop} - (pl)_t^{H^{seller.ind}} - (pl)_t^{NH^{seller.ind}}. \tag{138}$$

If the wife had children from a previous marriage, then the consumer type does not change. Otherwise, the model checks the age of the individuals. If at least one member of the household is over 67, the consumer type is updated ($(ConsT)_t^{cons} = 3$). In the *Births* sub-module, children are born with the probability depending on age and civil status of the mother ($pr_{age.st.civil}$). All necessary records are created for a newborn in submodule (M.36).

Module 6: Individuals' Records Updating Updating Entrepreneurs (M.37)

In this module, decisions are made on whether an adult individual ($(Age)_t^{ind} > 18$) becomes an entrepreneur. If the individual is already an entrepreneur ($(Entr)_t^{ind} = 1$), then the program skips this module and continues in the *Probability of opening a new firm* module (M.38). For the individuals unemployed ($\Xi_t^{ind} = 2$), employed in private sector ($\Xi_t^{ind} = 3$), inactive ($\Xi_t^{ind} = 4$), or employed in the public sector ($\Xi_t^{ind} = 5$), the probability of becoming an entrepreneur is computed. The probability depends on the experience in running a business ($(EntrP)_t^{ind} = 1$), gender (G^{ind}), age ($(Age)_t^{ind}$), level of education ($(Educ)_t^{ind}$) and the period that has passed since the last change of status on the labour market (Ψ_t^{ind})

$$\begin{aligned}(EntrS)_t^{ind} = &[(\max\{\min(1, (EntrS)_{t-1}^{ind} \times \\ &\times \left(\beta_G G^{ind} + \beta_{Educ}(Educ)_t^{ind} + \beta_{EntrP}\left(EntrP\right)_t^{ind} + \beta_{Age}(Age)_t^{ind} + \beta_\Psi \Psi_t^{ind}\right), \beta_{EntrS_{max}}\}].\end{aligned} \tag{139}$$

If the entrepreneur spirit is high enough, the individual becomes an entrepreneur ($(Entr)_t^{ind} = 1$). Otherwise, the program moves to module *Updating individual labour status* (M.47).

Module 7: Firm Demography Probability of Opening a New Firm (M.38)

In the *Firm demography* module, we analyse the possibility of entrepreneurs opening new companies in the system. As a part of the first *Probability of opening a new firm* submodule, the probability of opening a firm in a given industry is determined. This probability depends on previous experience in the industry ($(EntrP)_t^{ind}$), level of education ($(Educ)_t^{ind}$) and age of the entrepreneur ($(Age)_t^{ind}$). Firstly, temporary variables are calculated that are required to calculate the probability of creating a new company in a given sector.

$$temp_1 = \gamma_2 + \gamma_3 \times (EntrP)_t^{ind} + \gamma_4 \times (Educ)_t^{ind} + \gamma_5 \times ((Age)_t^{ind} + \gamma_6) \tag{140}$$

$$temp_2 = \gamma_7 + \gamma_8 \times (EntrP)_t^{ind} + \gamma_9 \times (Educ)_t^{ind} + \gamma_{10} \times ((Age)_t^{ind} + \gamma_{11}) \tag{141}$$

According to a given probability, the entrepreneur will open a new company in the most profitable sector (if $pr < temp_1$) and then in the second most profitable sector (if $pr < temp_2$). In this module we use information from the *Initialisation* module (M.1). If the entrepreneur cannot open up activity in these two sectors, they will start operating in a randomly selected industry.

Module 7: Firm Demography New License (M.39)

A number of criteria must be met to create a new business. First and foremost, it is necessary to obtain a license. The likelihood of obtaining a license depends on the size of the company compared to the average firm size in the industry and industry-specific parameters. The company's initial size $(Size)_{temp}$ is drawn from the normal distribution with the expected value $((AL)_t^s)$ and the standard deviation of the size of firms (empirically determined) $((SD)_t^s)$.

$$(Size)_{temp} \text{ was drawn from } N((AL)_t^s, (SD)_t^s) \tag{142}$$

$$(InitSize)_{temp}^{firm} = \max\{\min\left(\gamma_{12} \times (AL)_t^s,\ (Size)_{temp}\right), 0\}. \tag{143}$$

Module 7: Firm Demography Funding New Firm Creation (M.40)

This module analyses the possibility of opening a company from funds held by the entrepreneur or co-financing the opening of the company from external sources. Large companies can be co-financed by various individuals. In this case, the initial capital of the company is equal to the sum of the contributions of the individuals. The amount depends on the type of household to which the individual belongs. These individuals hold shares in the company. According to the algorithm, the cost of setting up a company depends on the size of the company measured by the number of employees $((InitSize)_{temp}^{firm})$, average salaries in the industry $\overline{W}_t^s$, and industry-specific fixed costs specified by the parameter $\gamma_{17.s}$:

$$(CostNew)_{temp}^{firm} = (InitSize)_{temp}^{firm} \times \overline{W}_t^s \times \gamma_{16} + \gamma_{17.s}. \tag{144}$$

The value of the individual's savings after deducting the charges for other repayments of loans is then calculated. If these funds are greater than the cost of establishing a business, the entrepreneur will allocate funds to set up a business. Consequently, the amount of savings in the entrepreneur's account is updated. Otherwise, depending on the size of the business, the entrepreneur can apply for bank credit (*Bank firm creation funding* (M.41.1.) & *New firm creation* (M.41.1.1)) or look for co-shareholders (*Shares* (M.41.2)).

Module 7: Firm Demography Bank Firm Creation Funding (M.41.1)

The entrepreneur seeks a bank which is both willing to lend funds and that offers a lower interest rate than other banks (i_{l^l}). In addition, the sectoral exposure requirements are checked $(CostNew)_{temp}^{firm} \leq \gamma_{sect.exp}$). If the selected bank has the requested funds ($S_t^{l.I^{est}} \geq (CostNew)_{temp}^{firm}$), creditworthiness is checked. In order to obtain the loan, the following criteria are to be fulfilled:

(1) the leverage ratio needs to be acceptable:

$$\frac{(CostNew)_{temp}^{firm}}{K_{init}^{firm}} < \gamma_{LR.b}, \tag{145}$$

where K_{init}^{firm} is a sum of funds (deposits) that the owners or shareholders provided to fund the capital of the company.

(2) the financial risk of the business has to be relatively low in comparison to the sectorial risk

$$((Risk)_t^{firm} < \gamma_{Risk.b} \ \&\& \ (Risk)_t^{firm} < \overline{(Risk)}_t^s), \tag{146}$$

(3) the size of the company has to be at least average for the industry ($L_t^{firm}/\overline{L}_t^s$) and credit history has to be good ($(PD)_t^{firm} < \gamma_{cred.h.b}$).

Banks may set their own thresholds for the proposed criteria. In this way banks compete not only on interest rates but also creditworthiness criteria. Thresholds are an expression of their degree of risk aversion and the adopted strategy of a bank's activity. If the trader does not get the requested amount in a given bank, they try again to apply for a loan in another bank, starting with the bank that offers the lowest interest rate. In extreme cases, the amount of funding obtained from the bank is reduced, gradually up to the limit below which it would not be possible to open a business in the sector. In the *New firm creation* (41.1.1) sub-module, records are created for the new firm, and the firm loans of the individual are updated. Maturity of the loan ($\mathrm{M}^{I^{ind}}$) is assigned depending on the amount of loan taken.

$$l_t^{I^{ind}} += (CostNew)_{temp}^{firm} \tag{147}$$

$$(pl)_t^{I^{ind}} += (CostNew)_{temp}^{firm} \tag{148}$$

$$l_t^{I_q^{ind}} += \frac{(CostNew)_{temp}^{firm}}{\mathrm{M}^{I^{ind}}} \tag{149}$$

$$R_t^{I^{ind}} += \frac{1}{\kappa_4} \times ((CostNew)_{temp}^{firm} \times (1 + 0.25 \times i_{l^I})^{\mathrm{M}^{I^{ind}}} - (CostNew)_{temp}^{firm}) \tag{150}$$

$$R_t^{I_q^{ind}} += \frac{R_t^{I^{ind}}}{\mathrm{M}^{I^{ind}}} \tag{151}$$

The bank's revenues are updated as well.

$$(Rev_{l^I})_t^b += R_t^{I_q^{ind}}. \tag{152}$$

Module 7: Firm Demography Shares (M.41.2)

For large companies, their initial capital is the sum of the contributions of the individuals. The share of each individual is determined by the consumer type of the household to which this individual belongs and the deposits. The algorithm looks for M individuals which shares will fund the cost of setting up the firm ($(CostNew)_{temp}^{firm}$). For all individuals we update the values of shares (ψ_t^{ind}) and deposits (d_t^{ind}). Large companies can apply for short-term loans but are funded by their shares. In the case of short-term loans, mechanisms are analogous to the one presented for small and medium-sized businesses but different credit ratings are adopted.

Module 8: Mergers & Acquisitions Updating Firm Market Value (M.42)

The *Mergers & Acquisitions* module analyses the possibility of one company buying another. First, the value of the company on the market ($(MV)_t^{firm}$) is updated, depending on the company's generated profits relative to average firm profits in the sector ($(RelProf)_{temp}^{firm} = \Pi_t^{firm}/(\overline{\pi})_t^s$) and the company's financial risk relative to the average risk in the industry ($(RelRisk)_{tem}^{firm} = (Risk)_t^{firm}/\overline{(Risk)}_t^s$). The market value of a company is correlated with the likelihood of bankruptcy of a particular company. By default, the maximum value of this variable is 1. The closer this value is to 1, the higher the probability of bankruptcy. If the variable is close to 1, then the probability of the business declaring insolvency increases.

Module 8: Mergers & Acquisitions Mergers (M.43)

If the business debt ($(Debt)_t^{firm}$) is lower than that specified by the parameter γ_{π} part of the firm's profits (Π_t^{firm}), the firm was not marked as for sale ($(ForSale)_t^{firm} = 0$), the firm debt is lower than that determined by the parameter γ_{LA_K} part of liquid assets ($(LA)_t^{firm}$) and capital (K_t^{firm}), the firm value ($(MV)_t^{firm}$) is greater than that specified by the parameter γ_{MV}, a percentage γ_{merg1} of firms

in the sector is searched for. The company will merge with one of these firms. The company to be merged with must meet several requirements. First, the sum of liquid assets of the firm and the fixed capital of this firm need to be lower than that of the second company.

$$((LA)_t^{merged.firm} + K_t^{merged.firm}) < ((LA)_t^{firm} + K_t^{firm}). \tag{153}$$

Secondly, it should be a smaller company, in a sense that the work force of this firm is smaller than the work force of the second company ($L_t^{merged.firm} < L_t^{firm}$). Next, the market value of the firm should be acceptable ($(MV)_t^{merged.firm} > \gamma_{MV} \times (MV)_t^{firm}$) and this firm should not be marked for sale ($(ForSale)_t^{firm} = 0$). If the merger occurs, then the attributes of the firm and the entrepreneur are updated.

$$(Debt)_t^{firm} += (Debt)_t^{merged.firm} + (MV)_t^{merged.firm} \times ((LA)_t^{merged.firm} + K_t^{merged.firm}) \tag{154}$$

$$(N_{est})_t^{firm} += (N_{est})_t^{merged.firm} \tag{155}$$

$$((LA)_t^{firm} += ((LA)_t^{firm} \tag{156}$$

$$K_t^{firm} += K_t^{firm}. \tag{157}$$

The owner of the firm changes and the corresponding variables such as labor status ($\Xi_t^{ind} = 5$), entrepreneurship ($(Entr)_t^{ind} = 1$), entrepreneurship in the past ($(EducP)_t^{ind} = 1$), periods since the last change on the labor market ($\Psi_t^{ind} = 0$) and the deposits are updated as well.

$$d_t^{owner.ind} = \gamma_{53} \times (MV)_t^{merged.firm} \times ((LA)_t^{merged.firm} + K_t^{merged.firm}). \tag{158}$$

If the firm wants to merge but there is no company to merge within the set γ_{merg1} of firms, the procedure is repeated in the broader set specified by γ_{merg2}.

Module 8: Mergers & Acquisitions Acquisitions (M.44)

If a company is successful and its financial risk is below average risk in the industry ($(Risk)_t^{firm} \leq \overline{(Risk)_t}^{s}$) and the firm is not marked for sale ($(ForSale)_t^{firm} = 0$), then it seeks to acquire another firm. In practice, this company should be relatively small ($L_t^{est} < \gamma_{55} \times L_t^{est}$), debt should be lower than the one expressed by debt percentage of the firm before acquisition ($(Debt)_t^{firm} < \gamma_{56} \times (Debt)_t^{firm}$), and this company should be marked for sale ($(ForSale)_t^{firm} = 1$). If a company to acquire is found, then the attributes of the firms and the entrepreneur are updated. The entrepreneur is now the co-owner of the firm.

$$\begin{aligned}(Debt)_t^{firm} += (Debt)_t^{acq.firm} + (MV)_t^{acq.firm} \times ((LA)_t^{acq.firm} + K_t^{acq.firm} + \\ - (Debt)_t^{acq.firm}) \times \sqrt{rndm(0,1)}\end{aligned} \tag{159}$$

$$K_t^{firm} += K_t^{acq.firm} \tag{160}$$

$$(N_{est})_t^{firm} += (N_{est})_t^{acq.firm} \tag{161}$$

$$(LA)_t^{firm} += (LA)_t^{acq.firm} \tag{162}$$

$$L_t^{firm} += L_t^{acq.firm}. \tag{163}$$

Module 9: Firm Growth Individuals Records Updating When Firm Closes (M.45)

This module deals with the opening and closure of companies. If the company was registered more than 1.5 years ago ($(\#Oper)_t^{firm} > 6$) and the number of employees is 0, then the company is removed from the system. If the entrepreneur does not own other firms, the value of variables which describe the experience of being an entrepreneur ($(Entr)_t^{ind} = 0$, $(EntrP)_t^{ind} = 1$), status on

the labour market ($\Xi_t^{ind} \neq 5$) and the period since the last change of status on the labour market ($\Psi_t^{ind} = 0$) are updated. If the company qualifies for closure, all employees will be first fired, and then the establishments will be closed.

Module 9: Firm Growth New Establishment Creation (M.46)

In this sub-module, firms with a strong position can also increase the number of establishments they own. If the company's financial risk is lower than the average risk in the sector ($(Risk)_t^{firm} < \overline{(Risk)}_t^{s}$) and the profit per employee $\frac{\Pi_t^{firm}}{(L_t^{firm}+1)}$ is higher than the percentage of the average profit per employee in the sector $v_2 \times \frac{\overline{\pi}_t^{s}}{(N_t^{firms} \times (\overline{L}_t^{s}+1))}$, then the company can create new establishments.

Module 10: Labour Market Updating Individual Labour Status (M.47)

Within the framework of the model, a stylized labour market has been designed. In the model the parameters $\alpha_{57.age.gender-62.age.gender}$ depend on age and gender. In the module the status on the labour market of all individuals (Ξ_t^{ind}) and the time since the last change of status on the labor market (Ψ_t^{ind}) are updated. If the individual is a woman over the age of 55, employed in the private or public sector, she is likely to be unemployed with the probability $\alpha_{57.age}$. If the individual is a man over the age of 55, employed in the private or public sector, he is likely to be unemployed with the probability $\alpha_{58.age}$. If the individual is a woman, who is unemployed for more than 2 years, then she becomes inactive with probability $\alpha_{59.age}$. If the individual is a man who is unemployed for more than 2 years, then he becomes inactive with probability $\alpha_{62.age}$. If a person is inactive on the labour market and is under the age of 55 and their total family income is less than the established threshold, then this person will be registered as an active job seeker, even if only for the purpose of receiving social benefits.

Module 10: Labour Market Workers Skills (M.48)

This sub-module updates the variable describing employee skills. These skills depend on age, gender, time elapsed since the last change of status on the labour market and completed education.

$$\Theta_t^{ind} = max\{0; \gamma_{\Psi} \times \Psi_t^{ind} + \gamma_{Educ} \times (Educ)_t^{ind} + \gamma_{G} \times G^{ind} + \gamma_{\Xi} \times \Xi_t^{ind} + \gamma_{Age} \times (Age)_t^{ind}\}. \quad (164)$$

Module 10: Labour Market Hiring & Firing in Establishments (M.49)

This sub-module carries one from the *Firm growth* module. In this sub-module a decision is made to reduce or increase the number of employees. First, the number of employees to be fired ($L_t^{FI^{est}}$) and hired in the establishment ($L_t^{HI^{est}}$) is compared. Employees with the lowest working skills are fired. If the number of employees to be hired is greater than the number of employees to be fired, then employees are searched for according to their location (using NUTS1-4 codes: $\vartheta_t^{ind^1}$, $\vartheta_t^{ind^2}$, $\vartheta_t^{ind^3}$, $\vartheta_t^{ind^4}$). The model assumes that employees living closer to the firm are beneficial to the company.

Module 11: Cycle Ends

The final module updates the remaining agent records that have not been updated in previous modules. Loans and deposits of individuals, households, companies and, above all, banks are settled. Non-performing loans in the period are also counted.

Module 11: Cycle Ends Labor Status & Wage & Subsidy Updating (M.50)

Individuals from the *Labor market* modules (M.47–M.49) are carried on into this sub-module. If individuals have just become unemployed ($\Xi_t^{ind} = 2$ && $\Psi_t^{ind} = 0$), they receive unemployment benefits ($(sub)_t^{ind} = \delta_1$) and their earnings from previous employment ($w_t^{ind} = 0$) are equal to zero. If individuals have become unemployed more than one quarter ago ($\Xi_t^{ind} = 2$ && $\Psi_t^{ind} > 1$), then they

receive lower benefits ($(sub)_t^{ind} = \delta_2$, $\delta_2 < \delta_1$). If a person is employed in the private sector ($\Xi_t^{ind} = 3$), then no social benefits are to be paid ($(sub)_t^{ind} = 0$) and the earnings are calculated according to the formula that depends on age ($(Age)_t^{ind}$), gender (G^{ind}), average wage per employee in the sector $\frac{\overline{W_t^s}}{\overline{L_t^s}}$ and a random factor.

$$w_t^{ind} = \frac{\overline{W_t^s}}{\overline{L_t^s}} \times (\delta_6 + \delta_7 \times G^{ind}) \times (\delta_8 + \delta_9 \times (Age)_t^{ind}) \times (rndm(0.04, 0.06)). \tag{165}$$

If the person is 67 years old and the subsidy was equal to 0, then the subsidy depending on average wage in the sector, gender and values of parameters is computed. If the person is inactive, then the gender and age of the individual are checked. If individuals are older than 55 ($G^{ind} = 1$ && $(Age)_t^{ind} > 55$), and had positive wage ($w_t^{ind} > 0$), they may retire early. The subsidy depending on gender and previous wage is computed, and the wage is set to zero ($w_t^{ind} = 0$, $(sub)_t^{ind} = \delta_2 \times w_t^{ind}$ for women and $(sub)_t^{ind} = \delta_2 \times w_t^{ind}$ for man). If the person in question runs a business and the profits of the firm are positive ($\Pi_t^{firm} > 0$), then for each of the companies they run, the portion of the profit paid to dividends is calculated; this is equal to the remuneration of the owner(s).

$$(Divid)_{temp} += \Pi_t^{firm} \times \delta_{14} \times rndm(0, 0.02) \tag{166}$$

$$w_t^{ind} += (Divid)_{temp} \tag{167}$$

$$(sub)_t^{ind} = 0 \tag{168}$$

$$(Debt)_t^{firm} += (Divid)_{temp}. \tag{169}$$

If the total family income is below the subsistence level (below δ_3), the public assistance is provided ($(sub)_t^{ind} = \delta_4$)

Module 11: Cycle Ends Establishments & Firms Records Updating (M.51) & (M.52)

The submodules *Updating establishment & firm records* first change the status of new establishments to experienced ones ($(New)_t^{est}--$) and wages of these establishments are rested from the liquidity assets ($(LA)_t^{est} -= W_t^{est}$) and then set to zero ($W_t^{est} = 0$). Then we compute the cost of wages ($W_t^{est} += w_t^{ind}$) and the total costs of the establishments.

$$(TC)_t^{est} += \delta_{16} \times (Inp)_t^{est} + \delta_{16} \times ((Debt)_t^{firm}/(\#Oper)_t^{firm}) + t_{pr} + \delta_{18.s} \times Y_t^{est} + \delta_{19.s} \times L_t^{est} + \delta_{20} \times W_t^{est} \tag{170}$$

Then the costs are paid from the liquid assets. If the liquid assets of the firm are negative ($(LA)_t^{est} < 0$), the establishment has liquidity shortage. We first check whether the firm is able to move sources from one establishment to the other one. If the firm owns more than one establishment and the liquidity of the firm and establishment are positive ($(N_{est})_t^{firm} > 1$ && $((LA)_t^{est} + (LA)_t^{firm}) > 0$), one of the establishments is closed ($(N_{est})_t^{firm} --$) and liquidity debts of this establishment are paid ($(LA)_t^{firm} -= (LA)_t^{est}$), and the value of default is updated ($(PD)_t^{est} ++$). The firm is at risk of default, but still has not defaulted ($(PD)_t^{firm}$ increases). Otherwise, if the firm consists of only one establishment that has serious liquidity shortage, the firm defaults ($(PD)_t^{firm} = 1$). The non-performing loans are updated ($(NPL_{l^I})_t^b += (pl)_t^{I^{ind}}$). If the owner of this firm does not own any other firm, then the status on the labor market ($\Xi_t^{ind} = 2$), the periods since the last change of the status on the labor market ($\Psi_t^{ind} = 0$) and the variable expressing experience in being an entrepreneur ($(Entr)_t^{ind} = 0$, $(EntrP)_t^{ind} = 1$) are updated. In this sub-module, the sales per employee are computed ($(SE)_t^{est} = \frac{(SE)_t^{est}}{L_t^{est}}$) and the liquidity assets and fixed capital are updated according to generated profits

(or loses). If the profits per employee are sufficiently high (i.e., if $\frac{\Pi_t^{firm}}{(L_t^{firm}+1)} > \delta_{21}$), then the liquid assets and the fixed capital are updated.

$$(LA)_t^{firm} += (\delta_{22} + \delta_{23}) \times K_t^{firm} \tag{171}$$

$$K_t^{firm} = (1 + \delta_{22} - \delta_{23}) \times K_t^{firm} \tag{172}$$

Otherwise, only fixed capital is updated.

$$K_t^{firm} = (1 - \delta_{23}) \times K_t^{firm} \tag{173}$$

If the liquid assets per employee are sufficiently high (i.e., if $\frac{(LA)_t^{firm}}{(L_t^{firm}+1) > \delta_{25}}$), then part of liquid assets can be used to pay back the debts. If the firm's debt is higher than the liquid assets, then the firm's debt and liquid assets are updated.

$$(Debt)_t^{firm} -= (LA)_t^{firm} - (L_t^{firm} + 1) \times \delta_{25} \tag{174}$$

$$(LA)_t^{firm} = (L_t^{firm} + 1) \times \delta_{25} \tag{175}$$

Otherwise, the debt is paid in full.

$$(LA)_t^{firm} -= (Debt)_t^{firm} \tag{176}$$

$$(Debt)_t^{firm} = 0. \tag{177}$$

In the worst case scenario, the firm is accumulating debt and the corresponding variables are updated.

$$(Debt)_t^{firm} += \left(L_t^{firm} + 1\right) \times \delta_{26} - (LA)_t^{firm} \tag{178}$$

$$(LA)_t^{firm} = (L_t^{firm} + 1) \times \delta_{26}. \tag{179}$$

At the end of this sub-module, lump sum property tax is updated (t_{pr}). In the *Updating firm records* sub-module, the age of the firm ($(\#Oper)_t^{firm} ++$) and the financial risk are updated.

$$(Risk)_t^{firm} = \delta_{29.s} + \delta_{30} \times \left(\frac{1}{L_t^{firm} + 1}\right) + \delta_{30} \times \left(\frac{L_t^{firm}}{(Debt)_t^{firm}}\right). \tag{180}$$

The values of a firm's fixed capital, liquidity assets, work force, firm profit and tax to be paid are the sum of respective values of variables of firms' establishments.

Module 11: Cycle Ends Updating Industries (Sectors) (M.53)

This module recalculates average values for sectors that become benchmarks for agents' decisions. The values of the quality and average price of each sector are updated ($(Ql)_{t-1}^s = (Ql)_t^s$, $P_{t-1}^s = P_t^s$). Then, the profit in the sector and number of firms in the sector are computed as a sum of firms' profits and the sum of firms. The average financial risk is computed as well. Then, the values of average sales in the sector, price in the sector, quality, average work force of the establishments in the sector, percentage of establishments that import and export as well as average wage in the sector are updated.

Module 11: Cycle Ends Paying Back of Loans (M.54)

This submodule computes how much has already been paid back from the loan, how much is due and how much interest was paid and is due. This is performed for all types of credit: consumer loans ($l_t^{C^{ind}} -= l_t^{C_q^{ind}}$, $(pl)_t^{C^{ind}} -= l_t^{C_q^{ind}}$, $R_t^{C^{ind}} -= R_t^{C_q^{ind}}$), real estate housing ($l_t^{H^{ind}} -= l_t^{H_q^{ind}}$,

$(pl)_t^{H^{ind}} -= l_t^{H_q^{ind}}$, $R_t^{H^{ind}} -= R_t^{H_q^{ind}}$) and non-housing loans ($l_t^{NH^{ind}} -= l_t^{NH_q^{ind}}$, $(pl)_t^{NH^{ind}} -= l_t^{NH_q^{ind}}$, $R_t^{NH^{ind}} -= R_t^{NH_q^{ind}}$), and firm investment loans ($l_t^{I^{ind}} -= l_t^{I_q^{ind}}$, $(pl)_t^{I^{ind}} -= l_t^{I_q^{ind}}$, $R_t^{I^{ind}} -= R_t^{I_q^{ind}}$). The same is computed for establishments paying back short-term loans ($l_t^{SH^{est}} -= l_t^{SH_q^{est}}$, $(pl)_t^{SH^{est}} -= l_t^{SH_q^{est}}$, $R_t^{SH^{est}} -= R_t^{SH_q^{est}}$).

Module 11: Cycle Ends Bank's Balance Sheet Positions With Non-Financial Sector Updating (M.55)

This submodule computes how much supply was used to give new credits in this iteration. For this purpose, the value of supply from the *Initialisation module* (M.1) and the value of supply after all updates (module M.55) is used. The amount of loans of each type which have been granted is computed.

$$l_{temp}^{C.granted.b} = S_{temp}^{l.C^{ind}} - S_t^{l.C^{ind}} \tag{181}$$

$$l_{temp}^{H.granted.b} = S_{temp}^{l.H^{ind}} - S_t^{l.H^{ind}} \tag{182}$$

$$l_{temp}^{NH.granted.b} = S_{temp}^{l.NH^{ind}} - S_t^{NH^{ind}} \tag{183}$$

$$l_{temp}^{I.granted.b} = S_{temp}^{l.I^{ind}} - S_t^{l.I^{est}} \tag{184}$$

$$l_{temp}^{SH.granted.b} = S_{temp}^{l.SH^{ind}} - S_t^{l.SH^{est}} \tag{185}$$

$$l_{temp}^{sum.granted.b} = l_{temp}^{C.granted.b} + l_{temp}^{H.granted.b} + l_{temp}^{NH.granted.b} + l_{temp}^{I.granted.b} + l_{temp}^{SH.granted.b} \tag{186}$$

$$(pl)_{temp}^{C.granted} += (pl)_t^{C^{ind}} \tag{187}$$

$$(pl)_{temp}^{H.granted} += (pl)_t^{H^{ind}} \tag{188}$$

$$(pl)_{temp}^{NH.granted} += (pl)_t^{NH^{ind}} \tag{189}$$

$$(pl)_{temp}^{I.granted} += (pl)_t^{I^{ind}} \tag{190}$$

$$(pl)_{temp}^{I.granted} += (pl)_t^{SH^{est}} \tag{191}$$

$$(pl)_{temp}^{sum.granted} = (pl)_{temp}^{C.granted} + (pl)_{temp}^{H.granted} + (pl)_{temp}^{NH.granted} + (pl)_{temp}^{I.granted} + (pl)_{temp}^{SH.granted} \tag{192}$$

$$(NPL)_{temp}^{sum} = (NPL_{l^C})_t^b + (NPL_{l^H})_t^b + (NPL_{l^{NH}})_t^b + (NPL_{l^I})_t^b + (NPL_{l^{SH}})_t^b \tag{193}$$

$$(NPL)_{temp}^{ratio} = \frac{(NPL)_{temp}^{sum}}{(pl)_{temp}^{sum.granted}} \tag{194}$$

$$(NPL)_{temp}^{ratio.l^C} = \frac{(NPL_{l^C})_t^b}{(pl)_{temp}^{sum.granted}} \tag{195}$$

$$(NPL)_{temp}^{ratio.l^{NH}} = \frac{((NPL_{l^H})_t^b}{(pl)_{temp}^{sum.granted}} \tag{196}$$

$$(NPL)_{temp}^{ratio.l^{NH}} = \frac{(NPL_{l^{NH}})_t^b}{(pl)_{temp}^{sum.granted}} \tag{197}$$

$$(NPL)_{temp}^{ratio.l^I} = \frac{(NPL_{l^I})_t^b}{(pl)_{temp}^{sum.granted}} \tag{198}$$

$$(NPL)_{temp}^{ratio.l^{SH}} = \frac{(NPL_{l^{SH}})_t^b}{(pl)_{temp}^{sum.granted}} \tag{199}$$

$$d_{temp}^{ind} += d_t^{ind} \tag{200}$$

$$d^{firm}_{temp} += (LA)^{est}_t \tag{201}$$

$$d^b_t = d^{ind}_{temp} + d^{firm}_{temp}. \tag{202}$$

Equity is updated in the next module, after computing the profits and costs of banks' activity.

Module 11: Cycle Ends Updating Banks' Profits & Costs & Equity (M.56)

This submodule updates the equity by adding the profits generated in this period and subtracting costs associated with the banks' activity.

$$(Rev^b)^{sum}_{temp} = (Rev_{lC})^b_t + (Rev_{lH})^b_t + (Rev_{lNH})^b_t + (Rev_{lI})^b_t + (Rev_{lSH})^b_t \tag{203}$$

$$(Cost)^b_{temp}\ draw\ from\ N((AvC)^b_t, (SdC)^b_t) \tag{204}$$

$$E^b_t += (Rev^b)^{sum}_{temp} - (Cost)^b_{temp} - i_d \times d^b_t \tag{205}$$

Also considered is the possibility of the banks' default. If the following value of temporal variable is below zero, the bank defaults. In practice, the bank can default when it has not enough equity or in the case of withdrawal of deposits or accumulation of non-performing loans.

$$(Default)^b_{temp} = E^b_t + d^b_t + (Rev^b)^{sum}_{temp} - i_d \times d^b_t - (NPL)^{sum}_{temp}. \tag{206}$$

Module 11: Cycle Ends Supply Side Decisions for T+1 and Regulatory Requirements (M.57)

This submodule analyses the regulatory requirements and the supply side decisions. Definitions of macroprudential ratios have been adapted from Popoyan et al. (2017). Firstly, the value of most liquid assets held by the bank (cash) as a percentage of all loans granted to the entities in the model is approximated.

$$(Appr.Cash)^b_{temp} = \varphi_1 \times l^{sum.granted.b}_{temp} \tag{207}$$

We also compute the quantity of loans less the most liquid assets held.

$$\Theta^{\ b}_{temp} = l^{sum.granted.b}_{temp} - (Appr.Cash)^b_{temp} \tag{208}$$

Next, the quantity of liquid assets demanded from the central bank is approximated.

$$(LiqDemCB)_{temp} = \varphi_{LCR.min} \times d^b_t - B^b_t - (Res)^b_t - \varphi_3 \times (Appr.Cash)^b_{temp} \tag{209}$$

The liquid assets from the central bank are only demanded when LCR is lower than $\varphi_{LCR.min}$. The computation of risk weighted assets and the approximation of TIER 1 equity allows the value of capital adequacy ratio according to Basel III to be obtained.

$$(RWA)_{temp} = (pl)^{sum.granted}_{temp} \times \delta_{\ RW} \tag{210}$$

$$(E^b_{T1})_{temp} = E^b_t \times \delta_{\ T1} \tag{211}$$

$$(CAR^b)_{temp} = \frac{(E^b_{T1})_{temp}}{(RWA)_{temp}} \tag{212}$$

The model also computes the value of bank leverage ratio, as a ratio of equity TIER 1 and total assets, as well as high quality liquid assets, expected cash inflow and outflow to be able to compute the liquidity coverage ratio.

$$(LR^b)_{temp} = \frac{(E^b_{T1})_{temp}}{(I^{sum.granted.b}_{temp} + B^b_t + (Res)^b_t + \varphi_3 \times (Appr.Cash)^b_{temp})} \quad (213)$$

$$(HQLA^b)_{temp} = (Res)^b_t + (Appr.Cash)^b_{temp} + \min\{0.85; 0.75 \times ((Res)^b_t + (Appr.Cash)^b_{temp})\} \quad (214)$$

$$(E(CashOutflow)^b)_{temp} = \varphi_4 + \varphi_5 \times d^b_t + \varphi_6 \times (LiqDemCB)_{temp} \quad (215)$$

$$(E(CashInflow)^b)_{temp} = \varphi_7 - \varphi_8 + \varphi_9 \times (Appr.Cash)^b_{temp} + \varphi_{10} \times B^b_t + \varphi_{11} \times (Res)^b_t \quad (216)$$

$$(NCOF^b)_{temp} = (E(CashOutflow)^b)_{temp} - (E(CashInflow)^b)_{temp} \quad (217)$$

$$(LCR^b)_{temp} = \frac{(HQLA^b)_{temp}}{(NCOF^b)_{temp}} \quad (218)$$

Then, the decisions about the supply of credit are made:

$$(S^b)_{temp} = \frac{E^{b_t}}{\varphi_{min.c.req}} - (pl)^{sum.granted}_{temp} \quad (219)$$

If the capital requirements are not met, the supply is set to zero ($(S^b)_{temp} = 0$) and the value of variable ($(PD)^b_t = 0.5$) is updated to emphasise that this bank is at threat of default and the riskiness of the bank's activity increases. Next, it is checked if the liquidity requirement is met. If so, the bank does not have to recur to the central bank. Otherwise $\left(\left(LCR^b\right)_{temp} < \varphi_{min.LCR} \times d^b_t\right)$, bank deposits are increased by the resources obtained from the central bank.

$$d^b_t + = (LiqDemCB)_{temp} \quad (220)$$

Finally, the leverage ratio requirement is checked. If the leverage ratio is lower than the value specified by the parameter ($(LR^b)_{temp} \leq \varphi_{13}$), the supply is set to zero ($(S^b)_{temp} = 0$) and the value of the variable expressing the probability of default ($(PD)^b_t = 0.5$) is updated to emphasise that this bank is at threat of default and the riskiness of the bank's activity increases.

Next, each bank needs to compute the supply of each type of loan to the market, taking into account the profitability of each and the non-performing loans ratios. There are two procedures that can be initiated in the system. The first one is initiated if the values of parameters μ_1, μ_2, μ_3, μ_4, μ_5, are not equal to zero. It is implicitly assumed that the banks are committed to providing a supply of credit of each type and all banks are universal. In the consequence of application of the first strategy, groups of banks that pursue similar strategies on the market are obtained.
At the beginning, the following auxiliary variables are computed:

$$(SumForRatio)^{l.C^{ind}}_{temp} = S^{l.C^{ind}}_{temp} + (Rev_{lC})^b_t - (NPL_{lC})^b_t \quad (221)$$

$$(SumForRatio)^{l.H^{ind}}_{temp} = S^{l.H^{ind}}_{temp} + (Rev_{lH})^b_t - (NPL_{lH})^b_t \quad (222)$$

$$(SumForRatio)^{l.NH^{ind}}_{temp} = S^{l.NH^{ind}}_{temp} + (Rev_{lNH})^b_t - (NPL_{lNH})^b_t \quad (223)$$

$$(SumForRatio)^{l.I^{ind}}_{temp} = S^{l.I^{ind}}_{temp} + (Rev_{lI})^b_t - (NPL_{lI})^b_t \quad (224)$$

$$(SumForRatio)^{l.SH^{est}}_{temp} = S^{l.SH^{est}}_{temp} + (Rev_{lSH})^b_t - (NPL_{lSH})^b_t \quad (225)$$

These auxiliary variables help us to understand how much has been earned and lost from the granted loans of each type. Then, the following ratios are computed:

$$(Sum1ForRatio)_{temp}^{l.C^{ind}} = \frac{(SumForRatio)_{temp}^{l.C^{ind}}}{S_t^{l.C^{ind}}} \tag{226}$$

$$(Sum1ForRatio)_{temp}^{l.H^{ind}} = \frac{(SumForRatio)_{temp}^{l.H^{ind}}}{S_t^{l.H^{ind}}} \tag{227}$$

$$(Sum1ForRatio)_{temp}^{l.NH^{ind}} = \frac{(SumForRatio)_{temp}^{l.NH^{ind}}}{S_t^{NH^{ind}}} \tag{228}$$

$$(Sum1ForRatio)_{temp}^{l.I^{ind}} = \frac{(SumForRatio)_{temp}^{l.I^{ind}}}{S_t^{l.I^{est}}} \tag{229}$$

$$(Sum1ForRatio)_{temp}^{l.SH^{est}} = \frac{(SumForRatio)_{temp}^{l.SH^{est}}}{S_t^{l.SH^{est}}} \tag{230}$$

The ratios are ordered from highest to the lowest. According to the ratios, it is possible to indicate the most profitable and less profitable types of loans. In the first procedure, it is assumed that the regulator introduced the maximum sectorial exposures, that is, μ_1, μ_2, μ_3, μ_4, μ_5, are fixed by the regulator. In practice, banks invest the maximum allowed by the regulator in the most profitable type of loans according to them. In the second procedure, it is assumed that the regulator introduced only one recommendation; that the exposure to any loan cannot be higher than 35%, that is $\mu_i \leq 0.35$ for $i \in \{1, 2, 3, 4, 5\}$. The auxiliary assumption needed states that all weights have to sum up to 1: $\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5 = 1$.

In this method, μ_1 corresponds always to the first ratio, μ_2 corresponds always to the second ratio, μ_3 corresponds always to the third ratio, μ_4 corresponds always to the fourth ratio and, μ_5 corresponds always to the fifth ratio. The highest ratio is checked, and then the value of μ_i, is drawn from the corresponding interval of the highest allowed values, $\mu_i \in (0{:}30; 0{:}35]$. Then, the second best, μ_j is drawn from the interval $\mu_j \in (\mu_i - 0.06; \mu_i)$. Analogically, the third and the fourth best, ie. μ_k is drawn from the interval $\mu_k \in (\mu_j - 0.06; \mu_j)$, $\mu_l \in (\mu_k - 0.06; \mu_k)$. The fifth one is computed from the restriction that states that all weights have to sum up to one. In this procedure, each bank has a different ordering of ratios and in addition weights are obtained in a stochastic procedure.

Module 11: Cycle Ends Interest Rates (M.58)

This sub-module computes the value of interest rates for next iteration. If the values of perception indicators are equal to zero then the same empirical interest rates from the database in the cycle can be used. Otherwise, the indicators of perception to change the values of interest rates in the first iteration are used and then random values of perception indicators from the distribution are used. The perception of risk is different on the ON ($\varsigma_t^{ind^1}$) and longer-maturity markets (SW, 1M, 3M) (respectively $\varsigma_t^{ind^2}$, $\varsigma_t^{ind^3}$,$\varsigma_t^{ind^4}$). The parameter ζ_t^{ind} expresses the value of perception of uncertainty indicator.

$$\varsigma_t^{ind^1} = 0 \;\&\&\; \varsigma_t^{ind^2} = 0 \;\&\&\; \varsigma_t^{ind^3} = 0 \;\&\&\; \varsigma_t^{ind^4} = 0 \;\&\&\; \zeta_t^{ind} = 0. \tag{231}$$

5. Inequality Measures and Distributional Effects

5.1. Income Distribution, Inequality and Concentration Measures

At $t = 0$, the simulation uses empirical data. In Figure 2 the household income according to the percentile has been presented as well as the approximation of wealth has been shown.

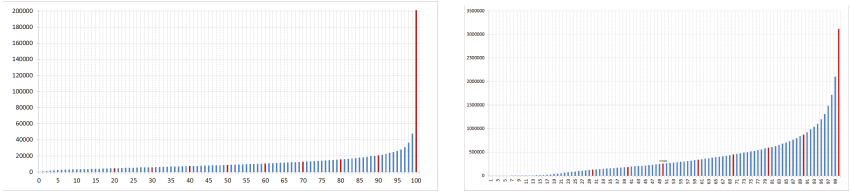


Figure 2. Household income by percentile (on the left) & Household wealth by percentile (on the right).

Based on empirical data, it is possible to calculate the approximation of the degree of income and wealth inequality in the sample at $t = 0$, which has been compared with results of simulations of wealth and income distributions in counterfactual scenarios (at $t > 0$).

5.2. Gini Coefficient and Measures of Asymmetry

Typically, the degree of inequality is calculated based on the Gini coefficient [105,106], and the Lorenz curve [107]. This section presents the results of the calculation of the basic Gini coefficient and the further studies on income and wealth inequality. Based on the sample, both the Lorenz curve and the income Gini coefficient were determined from the data on a quarterly and annual basis. The Gini coefficient for income distribution calculated on the basis of the quarterly sample data is equal to 0.3943179 (39.43%). In the case of calculations on annual figures, the coefficient value is slightly different, 0.3945987 (39.46%). If we include the corrections for unbiasedness, the value is 0.3947129 (39.37%).

These values differ slightly from the Gini coefficient for income distributions reported in the 2008 Report on Household Wealth and Debt, 39.4% versus 38.4%. This is due to the adoption of a different procedure of over-sampling of most affluent households. In addition, Grejcz and Żołkiewski [108] present the coefficient as 39.2%. At the same time, the authors emphasise that the procedure of oversampling of the most affluent households makes the value of Gini coefficient higher in comparison to the ones published by OECD, 35.5% or World Bank in 2014, 32.08%. In literature, other values of Gini coefficients for income distributions, computed using different samples, have been presented: 32.6% [109], 30.7% [110], 28.5% [111]. Accordingly, the Gini coefficient for wealth is equal to 0.6004731 (60.04%). The obtained results are relatively similar to the ones presented in the NBP [112] report 57.9%, and Grejcz and Żołkiewski [108] 58.7%.

According to Eurostat, Gini coefficient for income distribution in Poland in 2014 and 2015, were respectively 30.8% and 30.6%. The average Gini coefficient in the European Union for 28 countries was 30.9% in 2014 and 31% in 2015. A similar Gini coefficient for income distribution, suggesting a similar level of inequality, was reported in countries such as Croatia, 30.2% in 2014 and 2015, Ireland, 31.1% in 2014 and 29.8% in 2015, Germany, 30.7% in 2014 and 30.1% in 2015, and Great Britain, 31.6% in 2014 and 32.4% in 2015. The reported value of Gini coefficient of wealth for euro area by Grejcz and Żołkiewski [108] was 68.6%. In the case of the Gini coefficient for wealth distribution, similar results to the one for Poland were obtained in countries such as Malta, Belgium and Italy [113,114]. Interestingly, Gini coefficients for income and wealth of countries geographically and socially closer to Poland differ significantly from values reported for Poland.

However, as Chen [105] notes, "not all inequality curves yielding the same Gini coefficient are unequal in the same way". This observation is important not only for comparisons between countries but also for monitoring changes in inequalities in a given country when applying or analysing policies in a particular counterfactual scenario. Using agent-based simulation, we can analyse how the Gini coefficients change between iterations as a consequence of the introduction of specific policies. Furthermore, we can also describe changes in inequalities using more accurate measures of asymmetry and measures of spatial inequality.

Among the most useful measures in identifying patterns of inequality are the Lorenz asymmetry coefficient [115–117], and radial triangular measures derived from the Lorenz asymmetry coefficient expressed in polar coordinates. The Lorenz asymmetry coefficient (LAC; abbrev. S) is defined as:

$$S = F(\mu) + L(\mu) \tag{232}$$

where the functions F and L are defined as for the Lorenz curve, and μ is the mean. Based on an ordered set of data $(x_1, x_2, ...x_m, x_{m+1}, ..., x_n)$, we can compute statistics S in the following way

$$\delta = \frac{\mu - x_m}{x_{m+1} - x_m} \tag{233}$$

$$F(\mu) = \frac{m + \delta}{n} \tag{234}$$

$$L(\mu) = \frac{L_m + \delta\, x_{m+1}}{L_n}. \tag{235}$$

Damgaard and Weiner [115] show that Lorenz's asymmetry coefficient allows us to notice that although two Lorenz curves allow in practice the same Gini coefficient to be obtained, their shapes are different, which translates into other patterns of inequality. If the LAC is less than one, the inequality is related to the presence of relatively much poorer units. If the LAC is greater than one then the inequality is related to the existence of extremely rich individual entities. For the purposes of further discussion, the point at which the first derivative of the Lorenz curve is equal to 1 in the interval $(0, 1)$, it can be denoted, similarly to Chen [105], $[\mu, f(\mu)]$.

When calculating Lorenz's asymmetry coefficient for income distribution based on quarterly data, a result of 0.9790896 is obtained. For annual data, this is 0.9788734. Similarly, for wealth distribution it is 0.942073944. Intuitively, if the value of the Lorenz coefficient is less than 1, the point $[\mu, f(\mu)]$ is below and to the left of the symmetry axis. In fact, this case describes the situation in Poland, although the value does not differ significantly from the value of 1 corresponding to the point on the axis of symmetry. For the income distribution, the point $[\mu, f(\mu)] = (0.63; 0.35)$. This deals with the first of the two discussed situations. The income inequality in Poland is related to the relatively large number of poor agents. For wealth distribution, the point is equal $(0.68; 0.26)$ which is for the first case. Indeed, the wealth inequality is Poland is related to the relatively much poorer units.

The lack of access to disaggregated income and wealth data for countries such as Croatia, Ireland, Germany or Great Britain makes it impossible to compare patterns of inequality in these countries. However, it is still possible to compute Gini coefficients for Poland in different counterfactual scenarios. In the case of subtle changes in inequality in one country, it may be helpful not only to calculate the LAC but also to describe Lorenz's asymmetry with polar coordinates and to calculate the additional measure of adjusted azimuthal asymmetry (AAA) (*cf.* Chen [105]).

In this new approach the location of a point $[\mu, f(\mu)]$ relative to the point $(\frac{1}{2}, \frac{1}{2})$, a point dividing the line of equality in half can be described. In fact, $[\mu, f(\mu)]$ can be designated according to its radial distance from $(\frac{1}{2}, \frac{1}{2})$ and according to the azimuthal angle formed by that radius, relative to the line of symmetry. According to convention, they can be denoted accordingly ρ and θ. For the graphical explanation of ρ and θ based on Chen [105], see: Figure 3. Both enhance the interpretation of Lorenz asymmetry coefficient. The maximum radial distance of $[\mu, f(\mu)]$ from $(\frac{1}{2}, \frac{1}{2})$ is $\frac{1}{\sqrt{2}}$ and the azimuthal angle formed by the axis of symmetry and the line segment connecting $(\frac{1}{2}, \frac{1}{2})$ to $[\mu, f(\mu)]$ falls within the range of $+/-\frac{\pi}{2}$.

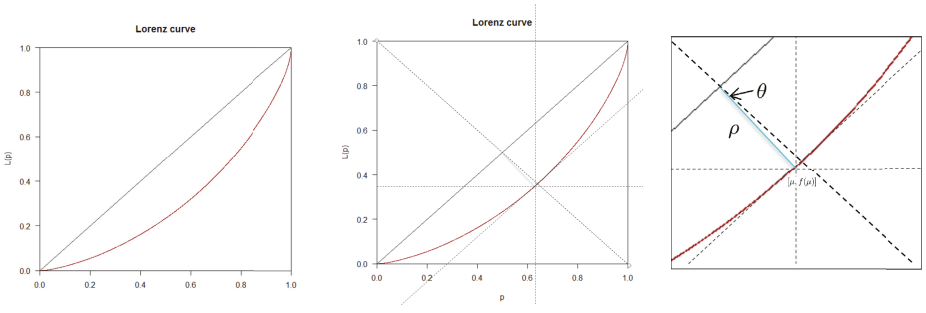


Figure 3. Lorenz curve (on the left) and the radial interpretation of ρ and θ (in the middle and on the right).

The radial distance ρ measures the concavity of the Lorenz curve and it is expressed by the formula:

$$\rho = \sqrt{(\mu - \frac{1}{2})^2 + (f(\mu) - \frac{1}{2})^2}. \tag{236}$$

Azimuthal angle θ measures the distance of $[\mu, f(\mu)]$ from the axis of symmetry in precise angular terms. It is expressed by the formula:

$$\theta \ = \arcsin(\frac{S-1}{\rho\sqrt{2}}). \tag{237}$$

It adds precision and intuitive geometric appreciation to the Lorenz asymmetry coefficient. "Whereas the primary use of S has been binary negative values lie to the left of the axis of symmetry, while positive uses lie to the right θ locates $[\mu, f(\mu)]$ with precision within the unit triangle" [105] Note that S, ρ and θ are constrained by the Gini coefficient. In practice, for a given value G, we have specified boundaries on the values S, ρ and θ . For any Gini coefficient G, the minimum value of S is G itself. The maximum value is $2 - G$.

For $t = 0$ during the simulation, the following results are obtained for income distribution:

$$\rho = 0.19908, \theta \ = -0.07434, t = \frac{\theta}{\pi/2} = -0.04733. \tag{238}$$

The distance from $(\frac{1}{2}, \frac{1}{2})$ to the orthogonal intercept along the line of symmetry is 0.19886. The maximum θ given the value of Gini coefficient for quarterly data is equal to 0.99369. The adjusted azimuthal asymmetry coefficient for $t = 0$ is equal to:

$$AAA = \frac{\theta}{\theta_{max}} = -0.07481. \tag{239}$$

Accordingly, for the wealth distribution, the following is obtained:

$$\rho = 0.30767, \theta \ = -0.13353, t = \frac{\theta}{\pi/2} = -0.08500. \tag{240}$$

The distance from $(\frac{1}{2}, \frac{1}{2})$ to the orthogonal intercept along the line of symmetry is 0.3049. The maximum θ given the value of Gini coefficient for quarterly data is equal to 0.5880. The adjusted azimuthal asymmetry coefficient for $t = 0$ is equal to:

$$AAA = \frac{\theta}{\theta_{max}} = -0.22783. \tag{241}$$

As Chen [105] states, "adjusted azimuthal asymmetry promises a two-fold advantage over the unadorned Lorenz asymmetry coefficient. First, the asymmetry is more intuitively and more accurately expressed in angular terms than as the sum of two Cartesian coordinates. Second, because adjusted azimuthal asymmetry accounts for the maximum angular distance from the axis of symmetry for a particular value of G, it expresses a sense of proportionality that raw θ, to say nothing of S, cannot".

5.3. Spatial Dimensions of Inequality and Inequality Aversion

Although the spatial analysis is based on basic Theil index (TI), the generalised entropy index (that the Theil index is a special case of) can be started at. The generalised entropy index is defined as:

$$GE(\alpha) = \begin{cases} \frac{1}{N\alpha(\alpha-1)} \sum_{i=1}^{N} [(\frac{y_i}{\overline{y}})^{\alpha} - 1] & \alpha \neq 0, 1 \\ \frac{1}{N} \sum_{i=1}^{N} \frac{y_i}{\overline{y}} \ln \frac{y_i}{\overline{y}} & \alpha = 1 \\ -\frac{1}{N} \sum_{i=1}^{N} \ln \frac{y_i}{\overline{y}} & \alpha = 0 \end{cases} \tag{242}$$

where N is the number of individuals or households, y_i is the income of entity i, α is the weight given to distances between incomes at different parts of the income distribution. The generalised entropy index has been proposed as a better measure of income inequality with regard to Gini coefficient due to the fact that it is an additively decomposable inequality measure [118].

If it is assumed that $\alpha = 1$, that is $GE(1)$, the TI is obtained. One of the advantages of the TI is that it is a weighted average of inequality within subgroups, plus inequality between those subgroups. For example, inequality within Poland is the average inequality within each region, weighted by region's income, plus inequality between regions. We can express the decomposition formally. If for the TI the population is divided into k subgroups and s_i is the income share of group i, T_{T_i} is the Theil index for that subgroup, and $\overline{x}_i$ is the average income in group i, then then the TI is expressed by the formula:

$$T_T = \sum_{i=1}^{k} s_i T_{T_i} + \sum_{i=1}^{k} s_i \ln \frac{\overline{x}_i}{\overline{x}}. \tag{243}$$

The TI allows us to analyse the relative importance of spatial dimension of inequality, that is, the regional inequality [119].

For time $t = 0$ during the simulation, based on empirical data of income and wealth, the entropy measure (EM) and TI for Poland is computed. We obtain respectively $EM = 0.2721004$ and $TI = 0.2836961$ for income, and $EM = 0.7599967$ and $TI = 0.6753153$ for wealth. The decomposition of TI allows us to analyse the spatial inequality in Poland. The results of measures of spatial inequality are not reported for confidentiality reasons. The authors cannot present which spatial codes 1-4 represent which region and districts in Central Statistical Office but it is feasible to present the data and visualise changes in spatial inequality in each counterfactual simulation using maps.

The generalised entropy index can also be transformed into a subclass of the Atkinson index (AI) with $\epsilon = 1 - \alpha$ for $0 \leq \alpha < 1$, defined as:

$$A_\varepsilon(y_1, ..., y_N) = \begin{cases} 1 - \frac{1}{\mu}(\frac{1}{N}\sum_{i=1}^{N} y_i^{1-\varepsilon})^{\frac{1}{1-\varepsilon}} & \text{for } 0 \leq \varepsilon \neq 1 \\ 1 - \frac{1}{\mu}(\prod_{i=1}^{N} y_i)^{\frac{1}{N}} & \text{for } \varepsilon = 1 \end{cases}, \quad (244)$$

where ε is an inequality aversion parameter. When ε approaches 0, it becomes more sensitive to changes in the upper end of the income distribution. Then, there is no aversion to inequality; no social utility is gained by complete redistribution. The higher ε, the more AI becomes sensitive to changes at the lower end of the income distribution. At the same time, more social utility can be gained from the redistribution as aversion to inequality is higher. Relatively low values of AI indicate a more equal distribution than higher values, given a particular degree of risk aversion. The computation of AI for the empirical distribution of income for Poland allows us to obtain the result of 0.1314228. The AI for the empirical distribution of wealth is equal to 0.3438987. A number of other measures of concentration and inequality have also been calculated to investigate the sources of income inequality in Poland. The Herfindahl concentration index (HHI) is equal 0.0005451883, while the Rosenbluth concentration index (RI) is 0.0004778672. The values of additional measures of wealth inequality are equal to 0.001255462 in the case of HHI and 0.4313275 in the case of RI.

5.4. Indebtedness of Households & Macroprudential Ratios

The distributions of debt service to income (DSTI), debt to income (DTI), loan to value for housing loans (LTV_H) and debt to income (DTA) at $t = 0$ are analysed, and then the changes in the empirical distribution of these ratios at $t > 0$ are presented in Appendix D. The percentage of households that have DSTI $\geq$ 30% (6%) and 40% (3.5%), as well as DTI $\geq$ 3 (3.6%) and 4.5 (1.7%) is computed. In the sample, 101 504 individuals have a total of 137 216 consumer loans. None of the individuals have more than 5 consumer loans. On average, individual customers have been charged statistically with 1.4 of consumer loans. In the sample, 43 008 housing loans were still to be repaid.

6. Results of Simulations

Using this agent-based simulation, it is possible to analyse the behaviour of a heterogeneous economy within certain counterfactual scenarios and the effects of public policies, especially stabilising and distributional effects of macroprudential policies. Firstly, the scenario is analysed in which the main mechanisms described by H. Minsky in 'Stabilizing unstable economy' (1986) [44] and in 'Financial Instability Hypothesis' (1992) [45] are simulated. Under the second scenario, the behaviour of the system is investigated, assuming the use of macroprudential policies to mitigate financial risk and stabilise an unstable economy. Then, the macroprudential policy stance which would best allow for the stabilization of heterogeneous economy is analysed, as is the distributional effects of these policies.

6.1. Minsky Moment

According to Reference [44], 'capitalist economies exhibit inflations and debt deflations which seem to have the potential to spin out of control'. The economies are unstable by nature (this instability, in the most general terms, was mostly related to industrial structure and finance). In this statement Minsky refereed to the Kindleberger's definition of self-sustaining disequilibrium from the 1970s. The capitalist economy is conditionally coherent. Instead of accepting the state of equilibrium, the focus should rather be on periods of tranquillity, apparent stability, which, in essence, are destabilising.

The neoclassical synthesis tries to explain how a decentralized economy achieves *coherence* and *coordination* in production and distribution (in other words, how market mechanisms lead to a sustained, stable-price, full-employment equilibrium). In opposition, the Minsky theory focuses on capital development of an economy and *the impact of financial institutions' activity on production* and

distribution. The optimum that is derived from the neoclassical decentralized market processes 'rules out interpersonal comparisons of well-being and ignores the inequity of the initial distribution of resources and thus of income'. As Minsky stated, 'inasmuch as our aim is to indicate how we can do better than we have, and as the best is often the enemy of the good, we can forget about the optimum [in neoclassical terms]. Even though a tendency toward coherence exists because of the processes that determine production and consumption in market economics, the processes of a market economy can set off interactions that disrupt coherence' [44].

Minsky's theory can be interpreted in two ways. In Keynesian terms, it should be understood in terms of accumulation of capital in the economy. In Knightian terms, it should rather be interpreted as a problem of allocating resources under risk and uncertainty. The paper refers to the first interpretation. In *Extensions*, the second interpretation is referred to. Hence, the role of risk perception and uncertainty in generation and amplification of risk within the system is analysed [42].

We begin with interpretations of the instability hypothesis in Keynesian terms. Minsky's theory refers to the general theory of Keynes from the 1930s. Nonetheless, in 'Stabilizing an unstable economy', the process of capital accumulation described by Keynes is accompanied by an exchange of current money for the future. At the heart of this theory are not only capital stock and investment but primarily cash flows. Any attempt to model the instability hypothesis must therefore be stock-flow consistent. In reference to the Minsky theory, in the simulation, three types of stylized cash flows are developed: income, balance sheet and portfolio. The income cash flow refers to all payments in the production and sale of inputs and final goods, thus also the one in the supplier searching and purchase modules. The balance sheet cash flow refers to repayment of debts. The third type of cash flow occurs due to transactions in which capital and financial assets are acquired by a new agent. Money in cash form is not analysed explicitly in the model, but the general idea of money and settlements has been taken into account in the model. As Minsky noted, 'money is created in the process of financing investment and positions in capital assets. An increase in the quantity of money first finances either an increase in the demand for investment output or an increase in the demand for items in the stock of capital or financial assets. As money is created, borrowers from banks enter upon commitments to repay money to the lending banks. In its origins in the banking process, money is part of a network of cash-flow commitments, a network that for the business side of the economy ultimately rests upon the gross profits, appropriately defined, that firms earn' [44].

The current money is used to purchase inputs that are used in the production of investment and consumption goods. The purchase of inputs can be funded by the firm from the liquid assets and capital. Otherwise, external funding can be obtained. Financial institutions on the market generate profits from lending funds. 'The financing terms affect the prices of capital assets, the effective demand for investment, and the supply price of investment outputs' [44]. Firms, establishments and individuals are obliged to repay debt, that is, principal and interest rates, within designated deadlines. In Minsky's basic theory, the focus was on investment. The government budget, the behaviour of consumption and the path of wages were secondary. To make the model more realistic and ensure it can be applied in modern policy making, all the above-mentioned elements were included. Similarly, the flow of money described by Minsky included only loans to firms, while in the simulation, loans to individuals and households were also added. The flow of money in Minsky's theory, similarly to the simulation, has the following two directions. Individuals make deposits to banks and banks lend money to firms and individuals. In a later period, firms and individuals return funds to banks and banks to depositories.

In this theory the cash flows are influenced by the *expected* future profits; the flow of money from companies to banks takes place after the realisation of real profits. These expectations need not be always consistent with the final realisation of profits.

Funds can be obtained through negotiation, in which risk perception and uncertainty as well as expectations play an important role. Firms willing to obtain money for carrying out their activity interpret financial results and the economic situation in an enthusiastic way, and in fact often create overly optimistic expectations. While bankers are inherently conservative, though profit-seeking, they

are more restrictive in assessing potential gains from a deal. Despite being more restrictive, bankers are also aware that investment in innovation and new products, services and industries is the most profitable. Consequently, their propensity to finance projects from the most dynamic and profitable sectors is usually greater than the projects from other industries, even if the riskiness of the project is higher than average.

Minsky's theory incorporates elements of Kalecki-Levy theory, according to which the structure of aggregate demand determines profits. The profit expectations depend on the expected level of investment in the future and actual returns on investment. Investments continue on the assumption of both entrepreneurs and bankers that investment will also take place in the future. In order to increase investment, the agents borrow additional funds, frequently assuming unrealistic returns in the future.

According to Minsky, there are three types of agents that are characterised by their relation between income and debt. 'Hedge financing units are those which can fulfil all of their contractual payment obligations by their cash flows. (...) Speculative finance units need to roll over their liabilities, [that is,] issue new debt to meet commitments on maturing debt. (...) For Ponzi units, the cash flows from operations are not sufficient to fulfil either the repayment of principle or the interest due on outstanding debts by their cash flows from operations. Such units can sell assets or borrow. Borrowing to pay interest or selling assets to pay interest on common stock lowers the equity of a unit, even as it increases liabilities and the prior commitment of future incomes. A unit that Ponzi finances lowers the margin of safety that it offers the holders of its debts' [44]. The higher the leverage ratio of firms, the higher the probability of defaulting on debt. Therefore, the reduction of collateral required and credit rating requirements in this scenario leads to an increase in the percentage of speculative and Ponzi agents. Changes in creditworthiness requirements also affect the percentage of hedging, speculative and Ponzi-type agents in the economy. The higher the level of speculators and Ponzi-type agents, the greater the probability of a crisis.

Investment boom increases aggregate demand and spending through a multiplier effect and sales increase. Profits increase with increasing investment, encouraging further investment. In this way, the instability of the system is strengthened until the percentage of speculative investors increases significantly. In the case of prolonged prosperity, assuming that no prudential policies have been applied, the economy becomes unstable, due to the increasing number of speculative and Ponzi-type agents.

Minsky's instability hypothesis is a study about the extent to which debt affects the behaviour of the system and how the level of debt is considered to be adequate in terms of dynamics of the system. According to this theory, there is no need for external shock for the crisis to occur. The crisis is generated endogenously by agents taking too much risk and by the desire by entrepreneurs and bankers to obtain ever-increasing profits. At the same time, this situation also has consequences for individuals and households. In the event of an increase in insolvencies, banks may default and thus depositors would lose the funds (the case in which by increasing uncertainty individuals decide to withdraw deposits is analysed in *Extensions*, see: [8]). In addition, the situation on the labour market determines changes in the income and wealth of households.

According to Minsky's theory, increase in investment 'would never trickle down to the poor and would tend to increase inequality by favouring the workers with the highest skills working in industries with the greatest pricing power' [44]. In the simulation, this mechanism was modelled. Firms hire individuals according to their work skills and in industries with the greatest pricing power. Consequently, it leads to inequality. The changes in inequality can be measured using the Gini coefficient as well as other measures of asymmetry and spatial inequalities.

The dynamics of the real estate market is also changing. The increased percentage of job seekers is influencing their shift to cheaper properties and partly ineligible mortgage loan repayment, which again puts banks in a difficult position. Banks are in possession of properties for which there is no demand, and their prices are gradually lowered in order to find a buyer.

Two main components of the theory are the two-price system and the lender's and borrower's risk, both derived from theories of Kalecki and Keynes.

The first price system concerns current output prices, that is, costs and mark-ups, that need to be set at a level that will generate profits for the firm from the sale of consumer goods, investment goods, and goods and services purchased by the government in public contracts. If, in the analysis, the external fund increase is also taken into account, the supply price of capital plus the interest rate and lender's risk must also be considered.

The second pricing system refers to assets. Assets are expected to generate cash flow in the future. These flows are not known and their estimation depends on subjective expectations. How much is able to be paid for such a financial asset depends on the amount of external finance required. The more the borrower becomes indebted, the greater the risk of insolvency; in this sense the price of the asset includes the borrower's risk. Investments occur when demand price exceeds the supply price of assets. Prices include collateral. After the crisis, usually larger collateral are required, in the expansion period, they are lowered significantly.

As Minsky [44] states, 'the costs of financing the production of investment is a cost that enters the supply price of output like the costs of labour and purchased inputs. The fact that a firm has to borrow to pay wages raises the effective costs by the interest payments on the borrowings'. 'The decision to invest therefore involves a supply function of investment, which depends upon labour costs and short-term interest rates, a demand function for investment, which is derived from the price of capital assets, and the anticipated structure and conditions of financing. Whereas the structure of balance sheets reflects the mix of internal funds (gross retained earnings) and the external funds (bonds and equity issues) actually used, the investment decision is based upon expected flows of internal and external funds' [44].

6.2. Simulating an Unstable Economy

In the first scenario, the role and impact of financial institutions on production and distribution of income and wealth in the economy is simulated. As Minsky pointed out, according to neoclassical assumptions, the initial distribution of income and wealth did not matter. Neoclassical models ignore the impact of policies on income and wealth distributions *ex post*. The inclusion of *ex ante* heterogeneity was not relevant to the analysis of outcome *ex post*. In contrast, in the simulated scenario, we show that the processes of a market economy do 'set off interactions that disrupt coherence', and that incorporation of heterogeneity *ex ante* into the model makes it possible to observe changes in distributions *ex post*, including income and wealth distributions.

In accordance with Minsky's hypothesis, in the database and model design, the focus was put not solely on heterogeneity in capital stock and on different investment decisions but also on balance sheet changes and cash flows.

In the scenario, firms operate in eight different sectors through establishments. In the first iteration, the most profitable sector is the first sector, thus it attracts the highest percentage of new investors. Other sectors attract a smaller percentage of new businesses. During first iterations we observe fluctuations in the relative profitability of sectors, related to business fluctuations, that is, the effect of the business cycle. In the initial iterations, only stock fluctuations related to intrinsic dynamics of the economy are observed, while in further iterations, it is observed that increased stock fluctuations and a cease in production is experienced by a higher percentage of establishments. The number of firms in each of sectors decreased over a year due to both higher rate of bankruptcies and higher concentration on the market, *see:* Figure A6 (on the right) in Appendix D. At $t = 0$ there are 18727 establishments in the database. At $t = 1$, 14112 establishments from the database at $t = 0$ continue operating on the market. 4615 establishments ceased to produce, while 9887 new establishments were attracted by the market due to higher relative profitability of selected sectors. At $t = 2$, 15727 establishments from $t = 1$ operate on the market, while 2773 new establishments are created. In total, 18500 establishments are in operation. At $t = 3$, 15357 establishments from $t = 2$ operate on the market

but only 1546 new establishments are established due to deterioration of market conditions. In total, a lower number of establishments (16,903) with respect to $t = 0$ operate on the market. Some firms go bankrupt, while others cease their activity in selected establishments that are operating in less profitable sectors. At $t = 4$, 15,357 establishments stay on the market and 1491 new establishments are created. In total, 16848 establishments continue operation.

In next iterations, there are higher levels of debt and leverage and increased financial risk in the sectors that were the most profitable in the previous periods. The risk spreading between sectors was visualised in Figure A5 (on the left and on the right side) and Figure A6 (on the left). Over a year, the higher number of firms' bankruptcy with respect to the previous periods was observed, indicating the gradual *endogenous* generation of the crisis. Occasionally, though gradually more frequently, firms had problems finding a supplier, which is related to the cessation of business activity by selected establishments, and thus the interruption of network of transactions established between establishments and suppliers. The data could be analysed for further iterations than four (one year) but the goal was to show that a crisis can emerge endogenously in a relatively short period of time, such as 1–4 quarters, as it was the case during the financial crisis (Q2 2008–Q2 2009).

The problem of searching and matching the suppliers in the supply chain module gains importance. The time required to searching for supplier is increased due to the fact that new transaction relations are established with suppliers in further territorial units than previously.

When the uncertainty on the market increases and indicators of profitability and riskiness deteriorate, a lower percentage of firms is eligible for a loan. Establishments lose creditworthiness and are forced to restrict production, which in turn worsens economic conditions. At $t = 1$, 62.5% of establishments were checked for creditworthiness in the module *Consumer credit admissibility* 1, while 37.5% in *Consumer credit admissibility* 2. In the following iterations, these values were respectively: 48.61% and 51.39% at $t = 2$, 55% and 45% at $t = 3$, and 57.5% and 42.5% at $t = 4$. The firms and establishments' conditions are also closely related to the problems of specific industries and fluctuations in the economy.

Expectations of lower profits lead to a reduction in the flow of money from companies to banks and from banks to companies, worsening the state of the economy. Banks are on the one hand aware of the higher risk related to granting loans to companies that qualify for the liquidity problem procedure, on the other hand it allows them to earn a premium. Overall, however, they provide loans for shorter periods, which is related to the general uncertainty in the market.

In the simulation cycle under first scenario, the number of companies with higher debt and higher leverage gradually increases. Using the simulation, changes can be observed in the distribution of the firms' debt and leverage, rather than just analysing growth of the overall average debt and leverage. However, as the crisis becomes more severe, more firms with excessive debt and leverage go bankrupt or cease their activities in selected establishments; see: Table A11 in Appendix D.

The initial boom also encourages households to increase consumption, which in turn increases their overall debt, also affecting the value of debt service to income (DSTI), debt to income (DTI), debt to assets (DTA) and loan to value (LTV) ratios, see: Figures A12–A16. After the first iteration, the percentages of households that have DSTI $\geq$ 30% equal to 4.7%, and DSTI $\geq$ 40% equal to 2.7%, while the percentage of households with DTI $\geq$ 3 equal to 3% and DTI $\geq$ 4.5 equal to 1.3%. After the second iteration, the percentage of households with DSTI $\geq$ 30% was equal to 3.6%, and DSTI $\geq$ 40% equal to 2.2%, while the percentage of households with DTI $\geq$ 3 equal to 2.1% and with DTI $\geq$ 4.5 equal to 0.9%. After the third iteration, the percentage of households with DSTI $\geq$ 30% is equal to 4.6%, and with DSTI 40% $\geq$ 3.2%, while those with DTI $\geq$ 3 is equal to 1.9% and with DTI $\geq$ 4.5 equal to 0.9%. After one year, the percentage of households with DSTI $\geq$ 30% was equal to 5.04%, and with DSTI $\geq$ 40% was equal to 3.5%, while with DTI $\geq$ 3 this percentage was equal to 2.2% and with DTI $\geq$ 4.5 the percentage was equal to 1.1%. The percentage of households with respective values of DSTI and DTI at $t > 0$ are lower than at $t = 0$. There are two reasons that explain this pattern. Firstly, households with very excessive debt default are removed by the system. Secondly, highly indebted households

will not apply for any new loans and they pay back part of the debt. Nonetheless, as soon as their creditworthiness improves due to loan repayment, they are attracted by the market via mechanisms described by Minsky and the ratios start to deteriorate. Changes in investment and external financing directly affect the market imbalances. Individuals with lower qualifications and level of education experience a gradually deteriorating situation on average.

After a year the Gini coefficient for income distribution is equal to 39.9% and for wealth distribution is 60.8%. After two years, the values are 41.1% and 61.3%, respectively. The LAC is equal to 0.96901 and 0.93896 after four quarters. The point $[\mu, f(\mu)] = (0.63; 0.33)$ for income, while it is equal $(0.66; 0.26)$ the case of wealth. The adjusted azimuthal asymmetry (AAA) is equal to -0.07541 for income and -0.21341 for wealth.

Due to endogenous changes in the decisions of agents and their interactions, the dynamics of the markets and economy also changes. In particular, the dynamics of the labour and real estate markets change. Higher mortgages and a difficult situation on the labour market affect the dynamics of the real estate market. More properties are marked for sale and their price is reduced accordingly. On average, the prices of properties decreased by 2% in one year. It is possible to analyse a decrease in prices in each spatial unit, for example, region, however the results are not reported for confidentiality reasons.

With the deterioration of a favourable economic situation, the number of insolvencies and the NPLs ratios of banks increased; see: Appendix D (Figure A7 (on the right) and Figures A9 and A10 (on the left and right). Higher non-performing loans affect the supply of credit in subsequent iterations. In the model it is possible to use two methods of determining the supply. If the first one is used, banks adopt similar strategies in groups. If the second is used, there may be a greater degree in the heterogeneity of strategies. In both cases, these patterns were observed. The model allows the analysis of differences in non-performing loans according to the loan type which thus reflects another aspect of heterogeneity. Using simulation, we can also analyse the changes in profits (Figure A8 (on the left)), equity (Figure A8 (on the right)) or the fulfilment with the capital and liquidity requirements which were set on too low levels (Figure A11 (on the left) and Figure A11 (on the right)).

Stabilizing Unstable Economy via Macroprudential Tools

In the second scenario, the behaviour of the economy is simulated, assuming the implementation of macroprudential policy aimed at stabilising the unstable economy. All banks set the capital requirements set by the regulator as well as apply the recommendations with respect to debt to income (DTI), debt service to income (DSTI), debt to assets (DTA), loan to value (LTV) ratios and leverage ratio (LR). In the economy, the activity of companies is conducted in the eight sectors grouping NACE industries. Among the analysed sectors, the most attractive for potential entities is the third sector, while the relatively least profitable for investors is the eight sector. In the case of any industry, there is no significant decrease in relative profitability. In first scenario, it was observed that the role of some of the sectors, that is, real estate sector, increased sharply, and then was significantly reduced. In the second scenario, no sudden discontinuity of production or increasing number of bankruptcies were observed. There are were no significant changes in product prices.

The expectations are relatively much more consistent in time than in the case of the first scenario. In this scenario, these expectations are only slightly different than those formed in the previous period and adapt to the fluctuations of the economy. Stock fluctuations are consistent with the dynamics of the economy. There is no stock accumulation or stop to production and sales of stored goods. In this scenario, the leverage level is moderate. Only in the case of a small percentage of establishments was production stopped completely.

In this scenario there are no searching-matching problems between the establishments in the search for suppliers. Quality-to-price ratio of establishment in relation to quality-to-price ratio in a given sector is the appropriate determinant of the supplier's search decision. In most cases, it is possible to find a supplier in a given spatial unit that has a sufficient quantity of produced and stored goods. The search for a supplier in the unit that exceeded the basic spatial unit is moderate and corresponds

to normal time market dynamics. The additional costs generated by transportation of inputs for establishments' production of goods were also moderate.

Most firms have adequate liquid assets and adequate creditworthiness. Additional funds will be used for further investment. A very low percentage of establishments were eligible for the creditworthiness check in connection with transitional liquidity problems. Transient problems of establishments on the market are related to temporary economic fluctuations in industries, as well as market dynamics. In this scenario, banks have adopted similar credit requirements for less risk-prone companies. However, banks differ in credit requirements for more risk-prone businesses, especially firms with temporary liquidity or financial problems. The tightening of the criteria has countered the financial crisis. Of particular importance was the observation of return on assets (ROA), return on equity (ROE), financial risk and credit history of companies.

In the case of a low number of establishments, a readjustment of quantity was required, and the final purchase of the inputs was lower than the intended one. Banks providing loans for the purchase of inputs were characterised by different risk aversion and strategies on the market. In some cases, the network effects have been maintained. The establishments applied for a loan in a matched bank, thus maintaining a transaction link (edges in the network). For other transactions, there was a change of the loan-granting bank. When searching for a bank that will grant short-term credit for the purchase of inputs, the companies were mainly driven by the interest rate and the supply of credit.

In the model it is possible to use two methods of determining the supply. If the first one is used, banks adopt similar strategies in groups. If the second one is used, a greater degree of heterogeneity of strategies is observed. In the first case, network effects associated with transaction relationships in the market have been maintained. Individuals choose banks to apply for a loan according to the interest rates on consumer loans offered by banks. Granting a loan also depended on credit supply and sectoral exposure requirements. Some banks have exhausted the supply of consumer loans, which means that according to the strategy defined for a given period, the funds were spent on other activities such as granting short-term and long-term (investment) loans to companies, and residential and non-residential loans to individuals.

Within the second method, during the year the same number of individuals applied for a loan at the matched banks; a very high percentage of these applications were successful. In this case, network effects associated with transaction relationships in the market were maintained. The supply of consumer loans and sectoral requirements limited the number of loans granted.

In both cases there were also no searching-matching problems between consumers and suppliers of goods. Also, the quality-to-price ratio in relation to quality-to-price ratio in a given sector is the appropriate determinant of the supplier's search decision. In most cases, it has been possible to find a supplier in a given spatial unit that had a sufficient quantity of produced and stored goods. In some cases, the search for suppliers exceeded the basic spatial unit, generating a modest cost of goods transportation. It did not generate an excessive burden on households.

Demand of establishments and households was complemented by demand from the government, through public procurement. The establishments that signed the public contracts produce high-quality products. These establishments were also part of the largest companies on the market.

In the case of households within 'Housing stress!', the cost of accommodation, whether in the case of servicing a housing loan or renting, proved to be excessive. These households changed their residence to a cheaper property. The dynamics of the housing market was not excessive and was consistent with household income fluctuations. In the case of a very low percentage of households, their insolvency was observed, as was an increase of NPLs in the banks by the value of their loans. The prices of properties fluctuate on the market as a whole but the prices remain relatively stable within a determined region.

Using the simulation, it is also possible to compute which percentage of households purchased a new property in cash and which have applied for a residential or non-residential loan. Demographic

trends were fully preserved in the scenario. The probability of survival or death has been specified according to Central Statistical Office data. As a consequence of the death of individuals, inheritors gained additional deposits. Most companies were taken over by the heirs, while in cases of negligible value the inheritance was rejected. In some cases, the deceased was the owner of the property charged with the loan. These properties were taken over by the bank and were resold at a lower value. In the case of survivors, the tendencies of completed education by individuals were also maintained.

Some companies were not created despite starting the company opening process. Some did not obtain a business license, and in some cases, funds were not sufficient to run a business. Credit applications are a special case in which the entrepreneurs compared the interest charged by the banks. The banks also checked whether the supply side restriction was fulfilled and whether the sectoral exposure did not exceed the requirements. Sectoral exposures have made it possible to limit potentially excessive credit growth in the most profitable sectors at any given moment, including in the real estate sector, which normally expands dynamically during prosperity. The larger companies were funded with contributions from entrepreneurs. There was an increase in average goodwill points during an expansive phase of a business cycle. At the same time, the structure on the market changed, with some activities ceasing and other new activities being created. New establishments were mostly opened in sectors with high or moderate risk. Changes in the number and structure of firms on the market corresponded to the usual dynamics of the economy. There was no increased concentration of capital, business clustering, excessive bankruptcies or escalation of reductions of labour force in already operating establishments.

An analysis of the distributions describing the attributes of firms and establishments does not allow the identification of situations typical of financial or economic crises and symptoms of overheating of the economy. In this scenario, a significant increase in the risk of activity of firms and establishments on the market was not observed, nor was the strong growth of the economy and the increased financial risk of a particular industry. The average risk to the business activity in the market was moderate.

The lack of concentration of financial risk in a given industry is largely due to the introduction of regulations for maximum exposure to a given industry. The situation of banks is stable. For the first and the second procedure of supply determination, the NPL ratios were moderate. Liquidity requirements are met in the case of most banks. Capital requirements were fulfilled at the level of 8%, introduced by bank regulator. None of the banks declared insolvency. A very low percentage of individuals declared insolvency as a result of net savings at a negative level. The NPL ratios and the growth (inflow) of NPLs of banks was modest.

The stabilising effects of macroprudential policies for the economy and the financial system is significant, however the effect on reducing inequality is ambiguous. The richest agents on the market seem to remain unaffected by the introduction of the policies. In extreme cases, the rich get richer. After one year, the Gini coefficient for income distributions was equal to 39.7% and 60.6% for wealth. After two years, the values were 40.1% and 60.9%, respectively. The Lorenz Asymmetry Coefficient (LAC) was equal to 0.97672 and 94782 respectively. The point $[\mu, f(\mu)] = (0.62; 0.36)$ for income, while it is equal $(0.67; 0.26)$ for wealth. The adjusted azimuthal asymmetry (AAA) is equal to -0.07541 in the case of income and -0.21341 in the case of wealth. The results do not differ significantly in the next four iterations.

6.3. Optimality of Macroprudential Policy Stance

Thus far, the possibility of stabilising an unstable economy using macroprudential policy has been analysed. Simulation has allowed analysis of the degree to which macroprudential policies influence the distribution of variables, including the distribution of income or the distribution of wealth, allowing a fuller welfare analysis. However, the main question is which policy combination to choose from, utilising possible combinations of tools to stabilise the economy and financial system, as well as to ensure a more homogeneous distribution of wealth and income. Optimal choice is understood here to

mean the method of choosing the most advantageous combination of macroprudential tools (from the point of view of social well-being). In contrast to the social planner approach, the abstract concept of Bentham's social utility is not focused on; rather there is focus on the stability of the economic and financial system and the measures of welfare that quantify changes in income and wealth distributions.

Based on the scenarios, it can be concluded that the use of a combination of macroprudential tools (appropriately calibrated) may lead to stabilisation of the system. Nonetheless, this combination of policies was not optimal in a sense that the statistical equilibrium was achieved at a level which does not guarantee a sufficient reduction of inequality in the system. The optimal policies would be the ones which enable a decrease in inequality in the system (to an acceptable level). The results were shown to be robust by carrying out the Monte Carlo procedure. The exact results for the Polish system (including optimal combination of macroprudential tools) are not reported here due to confidentiality of individual banks' and firms' data. The simulation can, however, be performed using data available from the National Bank of Poland, which would allow plausible results to be obtained. The results in the scenarios are mostly illustrative to the new methodology developed.

7. Final Remarks

The paper has analysed the stabilising effects of macroprudential policies on a heterogeneous economy using an agent-based approach. The presented simulation is a novel application of agent-based approach in systemic risk modelling. The model constitutes an alternative to the '3D' model with three layers of default that is widely-used by experts of the European Systemic Risk Board and the European Central Bank. The main advantage of the ABM model with respect to the '3D' model is the possibility to carry out counterfactual simulations within the framework of fully heterogeneous agents.

The simulation results show the stabilising effects of binding macroprudential policies on the unstable heterogeneous economy. However, in opposition to mainstream literature, the use of macroprudential policies as an alternative to redistribution policies does not always lead to the same results if heterogeneity is assumed in the model. Even if the shape of the income and wealth distributions could be smoothed for most percentiles, the macroprudential policies do not affect the richest individuals in the desired way; the richest get even richer with respect to the rest of society. In general, the macroprudential policies and regulation should counteract the negative effects of the crisis, including the impact of the crisis on society (distributional effects). Nonetheless, this paper shows that suboptimal or non-binding macroprudential policies in the economy with heterogeneous agents *ex ante* and *ex post* do not remove the all negative distributional effects. Poor calibration of macroprudential policies would not only counteract the endogenous generation of the crisis in Minsky's sense but would also lead to higher inequality. This finding is consistent with the ECB studies of P. Hartmann [120] on the distibutional effects of macroprudential policies and the possible interactions between social policies and macroprudential policies. This view was also supported by van der Heuvel [13] and Claessens [121]. Moreover, the policy recommendations and results of welfare analysis obtained from the DSGE models with homogeneous agents may be in practice very misleading for central bankers.

Author Contributions: Conceptualization, J.K.-M. and M.P.; methodology, J.K.-M. and M.P.; software, J.K.-M.; validation, J.K.-M. and M.P.; formal analysis, J.K.-M. and M.P.; investigation, J.K.-M. and M.P.; resources, J.K.-M. and M.P.; data curation, J.K.-M. and M.P.; writing—original draft preparation, J.K.-M. and M.P.; writing—review and editing, J.K.-M. and M.P.; visualization, J.K.-M.; supervision, M.P.; project administration, J.K.-M. and M.P.; funding acquisition, J.K.-M. and M.P. All authors have read and agreed to the published version of the manuscript.

Funding: The research of J. Kaszowska-Mojsa was partially funded by the National Science Centre (2013/09/N/HS4/03740) and the Polish-U.S. Fulbright Commission. The research of M. Pipień was supported by the research grant no. 2017/25/B/HS4/02529, financed by the National Science Centre, Poland.

Acknowledgments: We would like to thank Alberto Russo and Silvano Cincotti for valuable comments during the seminars "Agent Based Modelling in Macroeconomics" on 8 May 2015, and "Agent-based Macroeconomics: Endogenous credit-driven business cycles in Eurace" on 4 October 2017 at the National Bank of Poland. We also would like to thank James M. Chen for valuable comments on inequality measures. The model code and pseudocodes can be found on: https://www.comses.net/codebase-release/9b39dc1c-506e-4a02-b2ca-56b9c50738e4/.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AAA	Adjusted Azimuthal Asymmetry
ABM	Agent-Based Modelling
AI	Atkinson Index
BZGD	Household Wealth and Debt Survey
CAR	Capital Adequacy Ratio
CEE	Central and Eastern Europe
CPA	Classification of Products by Activity
CRISIS	Complexity Research Initiative for Systemic Instabilities
DSGE(-3D)	Dynamic Stochastic General Equilibrium Model (with 3 layers of default)
DTA	Debt to Assets Ratio
DSTI	Debt Service to Income Ratio
DTI	Debt to Income Ratio
ECB	European Central Bank
EM	Entropy Measure
ESRB	European Systemic Risk Board
EU	European Union
EURACE	A Massively Parallel Agent-Based Model of the European Economy
FX	Foreign Currency
GE(α)	Generalised Entropy Index
GUS	Central Statistical Office
HHI	Herfindahl Concentration Index
IMF-FSB-BIS	International Monetary Funds-Financial Stability Board-Bank for International Settlements
LAC	Lorenz Asymmetry Coefficient
LCR	Liquidity Coverage Ratio
LE	Large Exposures
LR	Real Business Cycle
LTV	Loan to Value Ratio
LTV_H	Loan to Value Ratio (only for mortgages)
MOSIPS	Modeling and Simulation of the Impact of Public Policies on SMEs (FP7)
NBP	National Bank of Poland
NPLs	Non-performing loans
NUTS (1–4)	Nomenclature of Territorial Units for Statistics
PFSA	Polish Financial Stability Authority
PKD (NACE)	Statistical Classification of Economic Activities in the European Community
RBC	Real Business Cycle
RI	Rosenbluth Concentration Index
ROA	Return on Assets
ROE	Return on Equity
TI	Theil index

Appendix A. Notation and List of Variables (Agent's Attributes)

Table A1. Individuals ($-.ind$).

Variable	Explanation	Type	Restr.
$(Id)^{ind}$	Individual ID (code identifying individuals)	Int.	≥ 0
$(Id)^{HH.(ind)}$	Identification code of household to which the individual belongs to	Int.	≥ 0
$(Id)^{est.(ind)}$	Identification code of establishment in which the individual works	Int.	≥ 0
$(Age)_t^{ind}$	Age of individual at time t	Double	≥ 0
$(Educ)_t^{ind}$	Level of education completed by the individual at time t	Int	$\in \{1,2,3,4,5,6,7,8\}$
$(EducP)_t^{ind}$	Periods of education	Int.	
$(Entr)_t^{ind}$	The variable determining whether an individual is an entrepreneur at time t	Boolean	$\in \{0,1\}$
$(EntrP)_t^{ind}$	The variable determining whether an individual was an entrepreneur in the past	Boolean	$\in \{0,1\}$
$(EntrS)_t^{ind}$	The probability of becoming an entrepreneur in the future (expectations)	Double	$\in \langle 0,1 \rangle$
G^{ind}	Gender of an individual	Boolean	$\in \{0,1\}$
$Id_{d_t}^{bank.(ind)}$	Bank matched to the individual in which deposits (savings) are hold at time t	Int.	≥ 0
d_t^{ind}	Deposits (savings and current account funds) on bank account of individual at time t	Double	≥ 0
$Id_{l_t^C}^{bank.(ind)}$	Bank matched to the individual in which wants to take a consumer loan at first at time t	Int.	≥ 0
$l_t^{C^{ind}}$	The amount of consumer loan taken at time t	Double	≥ 0
$(pl)_t^{C^{ind}}$	The sum of liabilities to the bank for outstanding consumer loans that have to be repaid in the given iteration	Double	≥ 0
$\mathrm{M}^{C^{ind}}$	The maturity of consumer loan in quarters	Double	≥ 0
$l_t^{C_q^{ind}}$	Consumer loans taken by an individual (quarterly payments at time t)	Double	≥ 0
$R_t^{C^{ind}}$	Interest to be paid from the total amount of consumer loans	Double	≥ 0
$R_t^{C_q^{ind}}$	Interest to be paid from the total amount of consumer loans in a quarter	Double	≥ 0
$Id_{l_t^H}^{bank.(ind)}$	Bank matched to the individual in which wants to take a housing loan at first at time t	Int.	≥ 0
$l_t^{H^{ind}}$	The amount of housing loan taken at time t	Double	≥ 0
$(pl)_t^{H^{ind}}$	The sum of liabilities to the bank for outstanding housing loans that have to be repaid in the given iteration	Double	≥ 0
$\mathrm{M}^{H^{ind}}$	The maturity of housing loan in quarters	Double	≥ 0
$l_t^{H_q^{ind}}$	Housing loans taken by an individual (quarterly payments at time t)	Double	≥ 0

Table A2. Individuals ($-.ind$) (*continuation*).

Variable	Explanation	Type	Restr.
$R_t^{H^{ind}}$	Interest to be paid from the total amount of housing loans	Double	≥ 0
$R_t^{H_q^{ind}}$	Interest to be paid from the total amount of housing loans in a quarter	Double	≥ 0
$Id_{l_t^{NH}}^{bank.(ind)}$	Bank matched to the individual in which wants to take a non-housing loan at first at time t	Int.	≥ 0
$l_t^{NH^{ind}}$	The amount of non-housing loan taken at time t	Double	≥ 0
$(pl)_t^{NH^{ind}}$	The sum of liabilities to the bank for outstanding non-housing loans that have to be repaid in the given iteration	Double	≥ 0
$\mathrm{M}^{NH^{ind}}$	The maturity of non-housing loan in quarters	Double	≥ 0
$l_t^{NH_q^{ind}}$	Non-housing loans taken by an individual (quarterly payments at time t)	Double	≥ 0
$R_t^{NH^{ind}}$	Interest to be paid from the total amount of non-housing loans (at time t)	Double	≥ 0
$R_t^{NH_q^{ind}}$	Interest to be paid from the total amount of non-housing loans in a quarter (at time t)	Double	≥ 0
$Id_{l_t^I}^{bank.(ind)}$	Bank matched to the individual in which wants to take an investment loan at first at time t	Int.	≥ 0
$l_t^{I^{ind}}$	Investment loans taken by an individual to invest in a firm that he owns or co-owns (at time t)	Double	≥ 0
$(pl)_t^{I^{ind}}$	The sum of liabilities to the bank for outstanding investment loans that have to be repaid in the given iteration	Double	≥ 0
$\mathrm{M}^{I^{ind}}$	Maturity of investment loan	Int.	≥ 0
$l_t^{I_q^{ind}}$	Investment loans taken by an individual (quarterly payments at time t)	Double	≥ 0
$R_t^{I^{ind}}$	Interest to be paid from the total amount of investment loans (at time t)	Double	≥ 0
$R_t^{I_q^{ind}}$	Interest to be paid from the total amount of investment loans in a quarter (at time t)	Double	≥ 0
PD_t^{ind}	Probability of default of individual at time t	Double	$\in \langle 0, 1 \rangle$
y_t^{ind}	Income of an individual at time t	Double	≥ 0
w_t^{ind}	Wage of an individual at time t	Double	≥ 0
$(sub)_t^{ind}$	Public assistance for an individual at time t	Double	≥ 0
ψ_t^{ind}	Shares of an individual in large and medium-sized companies at time t	Double	≥ 0
Ξ_t^{ind}	Status on the labour market of individual at time t	Int.	$\in \{1, 2, 3, 4, 5\}$
Ψ_t^{ind}	Number of periods since the last change of status on the labor market in quarters (at time t)	Int.	≥ 0
Θ_t^{ind}	Variable determining the skills of workers	Double	$\in \langle 0, 1 \rangle$
Λ_t^{ind}	Status civil	Int.	$\in \{1, 2, 3, 4\}$
$\vartheta_t^{ind^1}$	Spatial code 1 (NUTS1)	Int.	≥ 0
$\vartheta_t^{ind^2}$	Spatial code 2 (NUTS2)	Int.	≥ 0
$\vartheta_t^{ind^3}$	Spatial code 3 (NUTS3)	Int.	≥ 0
$\vartheta_t^{ind^4}$	Spatial code 4 (NUTS4)	Int.	≥ 0
$\varsigma_t^{ind^1}$	Perception of risk indicator (ON market)	Double	$\in \langle 0, 1 \rangle$
$\varsigma_t^{ind^2}$	Perception of risk indicator (SW)	Double	$\in \langle 0, 1 \rangle$
$\varsigma_t^{ind^3}$	Perception of risk indicator (1M)	Double	$\in \langle 0, 1 \rangle$
$\varsigma_t^{ind^4}$	Perception of risk indicator (3M)	Double	$\in \langle 0, 1 \rangle$
ζ_t^{ind}	Perception of uncertainty indicator	Double	$\in \langle 0, 1 \rangle$

Table A3. Households $(-.HH)$.

Variable	Explanation	Type	Restr.
$(Id)^{HH}$	Household ID	Int.	≥ 0
h_t^{HH}	The cost of living in own property or renting a property by the household at time t	Double	≥ 0
s_t^{HH}	The net savings of households (temporal variable) at time t	Double	≥ 0
$(Don)_t^{HH}$	Donations at time t	Double	≥ 0
$(DSTI)_t^{HH}$	Debt service to income at time t	Double	≥ 0
$(DTA)_t^{HH}$	Debt to assets at time t	Double	≥ 0
$(DTI)_t^{HH}$	Debt to income at time t	Double	≥ 0
$(LTV)_t^{HH}$	Loan to value (for housing loans) at time t	Double	≥ 0
$ID_t^{prop1.(HH)}$	Property id (the property in which the household lives) at time t	Double	≥ 0
$ID_t^{prop2.(HH)}$	Property id (this second property can be rented by other households or can be for sale)	Double	≥ 0
$(Rent)_{RE_t}^{HH}$	Household income from renting property at time t	Double	≥ 0
$(Own)_t^{HH}$	Status of the owner at time t	Boolean	$\in \{0,1\}$
y_t^{HH}	Total household income at time t	Double	≥ 0

Table A4. Firm $(-.firm)$.

Variable	Explanation	Type	Restr.
$(Id)^{firm}$	Firm ID	Int.	≥ 0
$(Id)^{s}$	Sector ID in which the establishments of the firm operate	Int.	≥ 0
$(PD)_t^{firm}$	The probability of default of a firm based on credit history	Double	$\in \langle 0,1 \rangle$
$(Risk)_t^{firm}$	The financial risk of the company	Double	$\in \langle 0,1 \rangle$
$(\#Oper)_t^{firm}$	Number of periods that a firm operates on the market	Int.	≥ 0
$(Debt)_t^{firm}$	Debt of the company	Double	≥ 0
$(MV)_t^{firm}$	The value of the firm in monetary terms	Double	≥ 0
Π_t^{firm}	Profits generated by the firm	Double	≥ 0
$(ForSale)_t^{firm}$	Variable determining whether a firm has been marked for sale	Boolean	$\in \{0,1\}$
K_t^{firm}	Fixed capital of a firm	Double	≥ 0
$(Id)^{ind.(firm)}$	Codes identifying business owner	Int.	≥ 0
$(LA)_t^{firm}$	Liquid assets of a firm	Double	≥ 0
$(N_{est})_t^{firm}$	Number of establishments of a firm in all sectors	Int.	≥ 0
t_{profit}	Taxes on profits	Double	≥ 0
L_t^{firm}	Number of employees of a firm	Int.	≥ 0

Table A5. Establishments ($-.est$, $-.est.buy$).

Variable	Explanation	Type	Restr.
$(Id)^{est}$	Establishment ID	Int.	≥ 0
$(Id_{ISH})^{b}$	ID of bank that granted short-term loan to establishment	Int.	≥ 0
$(TC)_t^{est}$	Costs of the establishment at time t	Double	≥ 0
$(TC)_{t-1}^{est}$	Cost of the establishment at time $t-1$	Double	≥ 0
$(PD)_t^{est}$	Probability of default of establishment at time t	Double	$\in \langle 0,1 \rangle$
$Q_t^{d_{est}}$	Demand for a good or service provided by the establishment at time t	Double	≥ 0
$(New)_t^{est}$	Variable determining whether the establishment has been recently created	Boolean	$\in \{0,1\}$
$E\left(Q_t^{d_{est}}\right)$	Expected demand for a good or service	Double	≥ 0
$(ExpImp)_t^{est}$	Variable determining whether the establishment exports, imports or both	Double	$\in \{0,1,2,3\}$
$(Id)^{firm(.est)}$	The code identifying the business which premises belongs to	Int.	≥ 0
K_t^{est}	Fixed capital	Double	≥ 0
$(Inp)_t^{est}$	Quantity of inputs required in the production of final goods expressed in monetary terms	Double	≥ 0
$R_t^{SH^{est}}$	Interest on short-term firm loans to be paid	Double	≥ 0
$R_t^{SH_q^{est}}$	Interest on short-term firm loans to be paid quarterly	Double	≥ 0
A_t^{est}	Know-how (technology)	Double	≥ 0
$(LA)_t^{est}$	Liquid assets of an establishment	Double	
ς_t^{est}	Perception indicator of establishment	Double	$\in \langle 0,1 \rangle$
ζ_t^{est}	Indicator of perception of uncertainty on the market	Double	$\in \langle 0,1 \rangle$
P_t^{est}	Price of goods produced by the establishment at time t	Double	≥ 0
P_{t-1}^{est}	Price of goods produced by the establishment in the previous period $(t-1)$	Double	≥ 0
Y_t^{est}	Production	Double	≥ 0
$q_t^{est.buy.(s)}$	The quantity of goods purchased from the supplier from sector s	Double	≥ 0
$l_t^{SH^{est}}$	Loans taken by establishment to purchase inputs	Double	≥ 0
$\mathrm{M}^{SH^{est}}$	Maturity of short-term loans	Int.	≥ 0
$l_t^{SH_q^{est}}$	Loans taken by establishments to purchase inputs (quarterly)	Double	≥ 0
$(Ql)_t^{est}$	Quality of products at time t	Double	$\in \langle 0,1 \rangle$
$(Ql)_{t-1}^{est}$	Quality of products at time $t-1$	Double	$\in \langle 0,1 \rangle$
$(SE)_t^{est}$	Sales per employee (expressed in monetary terms)	Double	≥ 0
$(Sl)_t^{est}$	Sales expressed in monetary terms	Double	≥ 0
$(Id)^{s.(est)}$	Sector ID	Int.	≥ 0
$\vartheta_t^{est^1}$	Spatial code – level 1 (NUTS1)	Int.	≥ 0
$\vartheta_t^{est^2}$	Spatial code – level 2 (NUTS2)	Int.	≥ 0
$\vartheta_t^{est^3}$	Spatial code – level 3 (NUTS3)	Int.	≥ 0
$\vartheta_t^{est^4}$	Spatial code – level 4 (NUTS4)	Int.	≥ 0
Inv_t^{est}	Inventories (stock)	Double	≥ 0
$Inv_{opt_t}^{est}$	Optimal stock of goods stored in the current period	Double	≥ 0
t_{pr}	Property tax	Double	≥ 0
W_t^{est}	The cost associated with the payment of wages by the premises	Double	≥ 0
L_t^{est}	Total labor force (number of employees) of the premises	Double	≥ 0
$L_t^{FI^{est}}$	Number of employees (labor force) to be fired in the current period	Double	≥ 0
$L_t^{HI^{est}}$	Number of employees (labor force) to be hired in the current period	Double	≥ 0

Table A6. Bank $(-.bank,\ -.b)$.

Variable	Explanation	Type	Restr.
$(Id)^b$	Bank ID	Int.	≥ 0
B_t^b	Bonds	Double	≥ 0
$(CT)_t^b$	Operating costs	Double	≥ 0
$(PD)_t^b$	The probability of default of the bank	Double	$\in \{0, 0.5,\ 1\}$
d_t^b	Deposits of the bank	Double	≥ 0
$(Rev)_t^b$	Revenues of the bank	Double	≥ 0
$(AvC)_t^b$	Average costs of banking operating in the sample	Double	≥ 0
$(SdC)_t^b$	Standard deviation of costs of banks operating in the sample	Double	≥ 0
E_t^b	Capital	Double	≥ 0
i_{l^C}	Interest rate on consumer loans	Double	≥ 0
i_d	Interest rate on deposits	Double	≥ 0
i_{l^I}	Interest rate on loans for investment purposes	Double	≥ 0
$i_{l^{SH}}$	Interest rate on loans for the purchase of inputs	Double	≥ 0
i_{l^H}	Interest rate on mortgages	Double	≥ 0
$i_{l^{NH}}$	Interest rate on non-housing loans (on the pledge of real estate)	Double	≥ 0
$(\#pd)_t^b$	Number of insolvencies of agents that had loans in this bank	Double	≥ 0
$(NPL_{l^{SH}})_t^b$	Net inflow of non-performing loans (for short term loans)	Double	$\in \langle 0, 1\rangle$
$(NPL_{l^I})_t^b$	Net inflow of non-performing loans (for investment loans)	Double	$\in \langle 0, 1\rangle$
$(NPL_{l^C})_t^b$	Net inflow of non-performing loans (for consumer loans)	Double	$\in \langle 0, 1\rangle$
$(NPL_{l^H})_t^b$	Net inflow of non-performing loans (for mortgages)	Double	$\in \langle 0, 1\rangle$
$(NPL_{l^{NH}})_t^b$	Net inflow of non-performing loans (for non-housing loans)	Double	$\in \langle 0, 1\rangle$
$(Rev_{l^C})_t^b$	Revenues from granting consumer loans	Double	≥ 0
$(Rev_{l^{SH}})_t^b$	Revenues from granting short term loans for firms	Double	≥ 0
$(Rev_{l^I})_t^b$	Revenues from granting investment loans	Double	≥ 0
$(Rev_{l^H})_t^b$	Revenues from granting mortgages	Double	≥ 0
$(Rev_{l^{NH}})_t^b$	Revenues from granting non-housing loans	Double	≥ 0
$S_t^{l.C^{ind}}$	Supply for consumer loans	Double	≥ 0
$S_t^{l.I^{est}}$	Supply for firm investment loans	Double	≥ 0
$S_t^{l.SH^{est}}$	Supply for short term loans for firms	Double	≥ 0
$S_t^{l.H^{ind}}$	Supply for mortgages	Double	≥ 0
$S_t^{NH^{ind}}$	Supply for non-housing loans	Double	≥ 0
i_{Bid3M}	Interest rate WIBID3M	Double	≥ 0
i_{OR3M}	Interest rate WIBOR3M	Double	≥ 0

Table A7. Sector (industry) $(-.s)$.

Variable	Explanation	Type	Restr.
$(Id)^s$	Sector ID	Int.	≥ 0
$\overline{\pi}_t^s$	Average profitability of firms in the sector s	Double	≥ 0
$\overline{(Risk)}_t^s$	Average financial risk of the industry (sector)	Double	$\in \langle 0, 1 \rangle$
$(Sl)_t^s$	The average sales of goods in firms in a given industry	Double	≥ 0
$\overline{L}_t^s$	Average work force of establishments in the industry	Double	≥ 0
$\overline{W}_t^s$	Average wage in industry (sector)	Double	≥ 0
$(\%Exp)_t^s$	Percentage of establishments that import to other industries	Double	≥ 0
$(\%Imp)_t^s$	Percentage of establishments that exports from other industries	Double	≥ 0
P_t^s	Average price of good or service in a given industry at time t	Double	≥ 0
P_{t-1}^s	Average price of goods or services in a given industry at time $t-1$	Double	≥ 0
Π_t^s	Total profit from the business of establishments in a given industry	Double	≥ 0
$(Ql)_t^s$	The average quality of a good or service in a sector in the current period	Double	≥ 0
$(Ql)_{t-1}^s$	The average quality of a good or service in a sector in the previous period	Double	≥ 0
$(AL)_t^s$	Average size of a firm in a given industry (sector)	Double	≥ 0
$(SD)_t^s$	Standard deviation of the size of firms in the sector	Double	≥ 0
$N_t^{firm_s}$	Number of firms operating in the industry	Int.	≥ 0
t_{VAT}	Value added tax	Double	≥ 0

Table A8. Properties $(-.prop)$.

Variable	Explanation	Type	Restr.
$(Id)^{prop}$	Property ID	Int.	≥ 0
$(Id)^{HH.(prop)}$	Household ID	Int.	≥ 0
$(\#Rent)_t^{prop}$	Number of households that rent the property at time t	Int.	≥ 0
$(PH)_t^{prop}$	Principal housing	Boolean	$\in \{0, 1\}$
P_t^{prop}	Property price	Double	≥ 0
$(ForSale)_t^{prop}$	Property for sale	Boolean	$\in \{0, 1\}$
$\vartheta_t^{prop^1}$	Spatial code level 1 (NUTS1)	Int.	≥ 0
$\vartheta_t^{prop^2}$	Spatial code level 2 (NUTS2)	Int.	≥ 0
$\vartheta_t^{prop^3}$	Spatial code level 3 (NUTS3)	Int.	≥ 0
$\vartheta_t^{prop^4}$	Spatial code level 4 (NUTS4)	Int.	≥ 0

Table A9. Consumers $(-.cons)$.

Variable	Explanation	Type	Restr.
$(Id)^{HH.(cons)}$	Household ID	Int.	≥ 0
$(ConsT)_t^{(cons)}$	Consumer type	Int.	$\in \{1,2,3,4,5,6\}$
$(Id)^{(cons)}$	Consumer ID	Int.	≥ 0
$(Id)^{s.(cons)}$	Sector ID	Int.	≥ 0
$(Id)^{sup.(cons)}$	Supplier ID	Int.	≥ 0
$Q_t^{HH.(cons)}$	The quantity of goods that a household will buy from suppliers in all sectors at time t	Double	≥ 0

Table A10. Suppliers $(-.sup)$.

Variable	Explanation	Type	Restr.
$(Id)^{est.(sup)}$	Establishment ID	Int.	≥ 0
$Q_t^{buy.est}$	Quantity of inputs to be purchased by establishment at time t	Double	≥ 0
$(Id)^{s.(sup)}$	Sector ID	Int.	≥ 0
$(Id)^{sup}$	Supplier ID	Int.	≥ 0

Appendix B. Sequential Updating of States in the Model

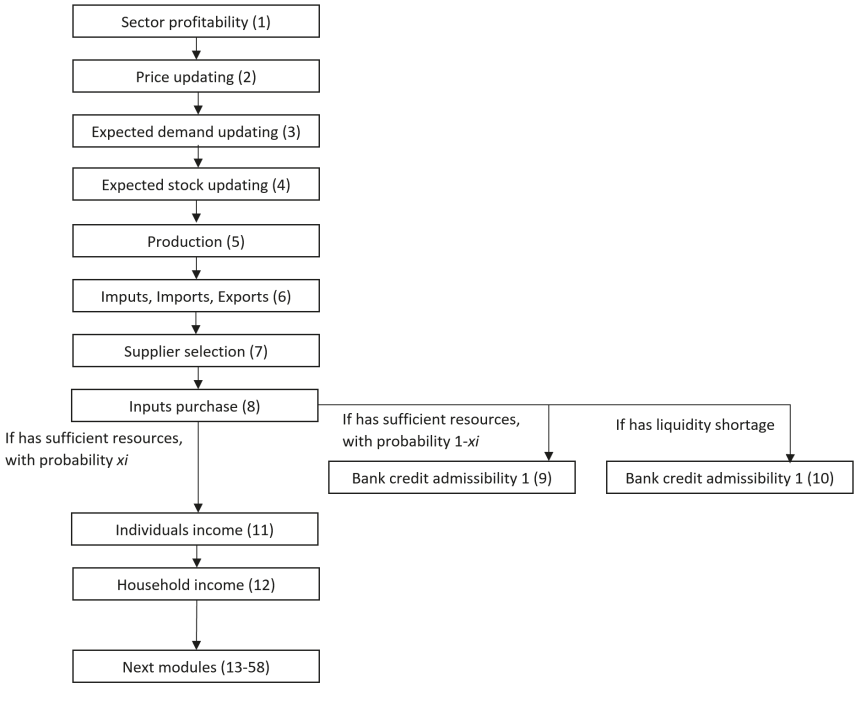


Figure A1. Sequential updating of states in the model (modules 1–12).

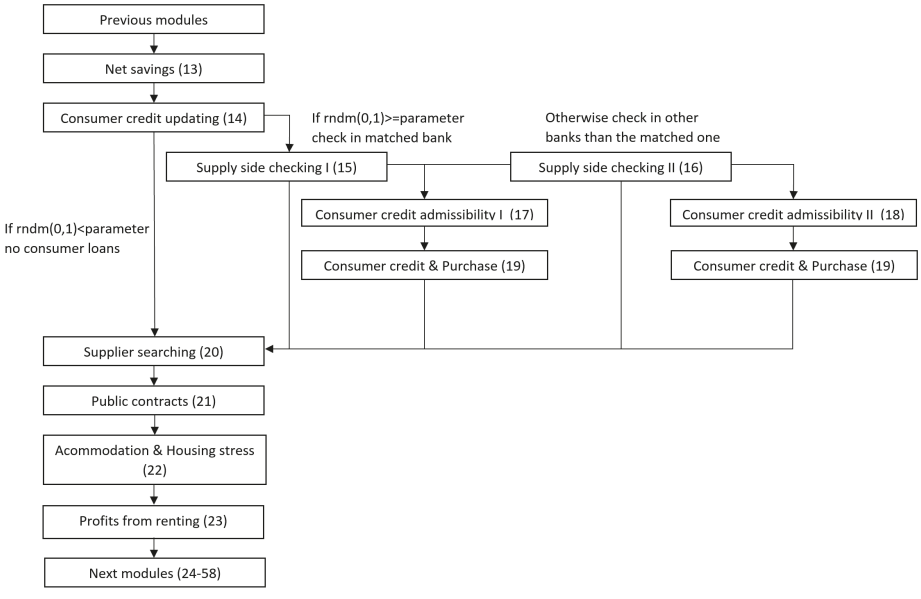


Figure A2. Sequential updating of states in the model (modules 13–23).

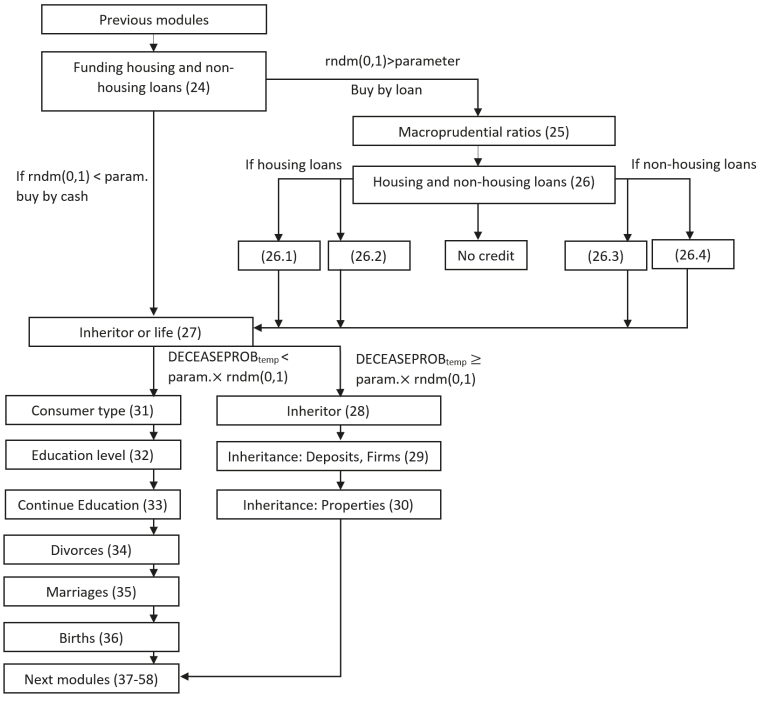


Figure A3. Sequential updating of states in the model (modules 24–36).

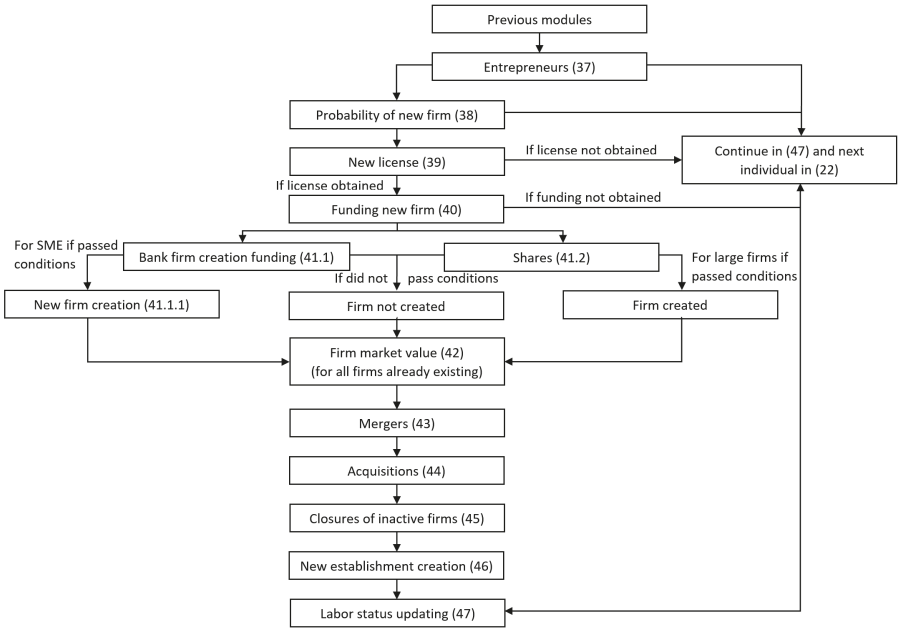


Figure A4. Sequential updating of states in the model (modules 37–47).

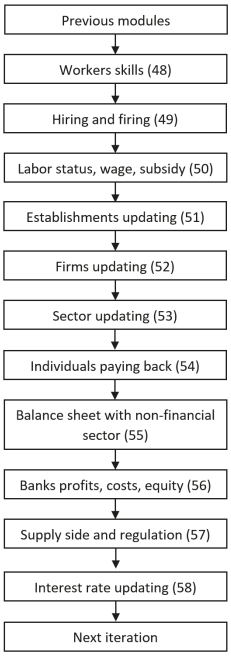


Figure A5. Sequential updating of states in the model (48–58).

Appendix C. Calibration

At the start of the simulation, the attributes of individuals, households, and consumers, as well as some attributes of properties were calibrated on the basis of data from the Household Wealth and Debt Survey from 2014 [112]. This first study on household wealth and indebtedness in Poland was conducted between January, 20 and February 28, 2014, on a sample of 7000 households residing in the country. The study was conducted through direct interviews. The test sample was generated by a two-stage random sampling scheme, where the first stage of sampling was stratified; the census units were randomly selected from the layers, and the second stage of household sampling was randomised. The stratification was based on the following criteria: province, size of the spatial unit (6 categories) and wealth (on a scale of 1–4). In the applied scheme, procedures of over-sampling of the most affluent households and imputation of missing observations were applied [122]. The completion ratio of the surveys was approximately 49.4%.

Appendix C.1. Individuals, Households & Consumers at t = 0

Based on the information from the Household Wealth and Debt Survey sample, the distributions of agent's attributes were approximated. In the model 971 520 individuals were distinguished at $t = 0$. There are 509 056 (0.52%) women and 462 464 (0.48%) men in the database. The distributions of sex and age were compared with the data from 'Population. Size and structure of population and vital statistics in Poland by territorial division as of 31 December 2014' [123], and 'Size and structure of population by age groups in 1989–2016' [124], provided by the Central Statistical Office (GUS). To ensure the proper dynamics of the model, data from 'Marriages contracted and dissolved in 1970–2015' [125], 'Births in 1970–2015' (GUS) [126] and 'Demographic Yearbook of Poland 2015' [127] were also used. In addition, data from life expectancy tables of 2014 provided by GUS were used [128]. Data was supplemented by 'Education in the school year 2014/2015' (GUS) [129]. Supplementary data on labour market and wages was used from 'Yearbook of Labour Statistics, 2015' (GUS) [130]. In addition, information from 'Social assistance. Child and Family services' (GUS) [131] was used. In the model, at $t = 0$, 442,240 households were distinguished. Each household is described by a joint distribution of attributes. BZDG data was supplemented by the data from 'Household budget survey in 2014' (GUS) [109] in order to obtain distributions. To each household a specific type of consumer has been assigned, which determines expenditures. There are therefore 442,240 $\times$ 8 (industries) $\times$ 6 (consumer types) = 21,227,520 data-points (records) in the Table *Consumers* database.

Appendix C.2. Firms & Establishments at t = 0

The values of variables of firms and establishments were approximated on the basis of GUS reporting forms of enterprises for 2014 [132]. In order to approximate inputs from different sectors and demand for final goods as well as direct product input coefficients and direct import input coefficients, information from input-output tables for 2010, provided by the Central Statistical Office [133], was used. The distributions of variables have not been reported for confidentiality reasons.

Appendix C.3. Industries & Suppliers at t = 0

The sectors in the model correspond to the branches of economy used in the input-output table at basic prices for domestic output for 2010, provided by GUS (2014) [133]. The classification used in the table corresponds to products and intermediate consumption (CPA 2008). Data from input-output tables was used to approximate the flows between sectors in the model. Each industry is described by the average values of variables that were approximated using the data from Central Statistical Office forms. The classification of industries (PKD (NACE) 2007) was matched with the classification of sectors used in the input-output tables. PKD 2007 is more disaggregated than the classification used in the published input-output tables. PKD-2007 is fully methodologically, conceptually, in the scope and coding system (up to fourth digit) coherent and compatible with the classification NACE Revision 2.

In the input-output tables the Classification of Products by Activity (CPA) was used. Some of PKD sectors were aggregated into one group in input-output tables for confidentiality reasons. A total number of 73 sectors were analysed, excluding sectors 04 (Coal and lignite), 08 (Tobacco products), 66 (Public administration and defence services; compulsory social security services) and 77 (Services of households as employers of domestic personnel & Undifferentiated goods and services produced by private households for own use). The suppliers for establishments have been matched based on input-output tables. The model expresses the network structure of relations between companies in the economy. For confidentiality reasons, the system was simulated with eight sectors, grouping certain industries.

Appendix C.4. Banks at t = 0

The model included 14 major banks in the market. The interest rates for these representative banks came from new MIR interest rate statistics [134]. Different interest rates were used for deposits, consumer credit, housing loans and non-residential loans. The distinction between flows and stocks was also taken into account. The interest rates for each bank were not reported for confidentiality reasons. The behavioural aspects of banking decisions while checking the creditworthiness, such as levels of LTV or DTI, and sectoral exposures and concentration limits, have been approximated using the PFSA Recommendations S, T and C, as well the Banking Code. In the future, it will be possible to use the PFSA surveys. The data from the surveys has not yet been accessible for confidentiality reasons.

Appendix C.5. Parameters and Decision Rules

All decision rules are obtained as a result of prior econometric analyses. In addition, expert knowledge about the functioning of markets and economic agents on the Polish markets were used. Parameter values were obtained after a thorough analysis of statistical data obtained from the National Bank of Poland and the Central Statistical Office and partly correspond to the values adopted in macro-models in the literature (e.g., labour elasticity, capital elasticity, relative quality elasticity and, know-how elasticity in the production function). The most important added value of the analysis is the linking of information obtained from the Central Statistical Office database on firms and the publicly available input-output table at basic prices for domestic output. All parameters assumed in modules 6–8, selected parameters in 14 & 19–20 were obtained directly using the procedure developed by Santos & Kaszowska (in the MOSIPS FP7 project), based on input-output table data. In order to obtain the parameters in demographic, education and labour market modules, Central Statistical Office and Household Wealth and Debt Survey data were used (especially, modules 27–36, 47–50). The parameters in modules 38–40, 42–46, 50–52 were calibrated based on Central Statistical Office data on enterprises.

The group of 'parameters' in the model includes not only the values of estimated or calibrated parameters of a functional form of decision rules, but also parameters that can be adjusted by the user to develop scenarios and to carry out a counterfactual simulation on the state of the economy and financial sector as well as macroprudential policies. For instance, in module 3 it was assumed that the level of leverage and financial risk that is acceptable by the companies in each of the industries to produce goods and services. Similarly, the levels of screening ratios of banks (e.g., ROA, ROE, acceptable financial risk of the applicant or the industry etc.) and macroprudential ratios can be assumed. Alternatively, the experts of the National Bank of Poland can use confidential data from the surveys about the banks' policies and stance, and/or the calibration based on the ESRB recommendations.

Appendix D. Figures & Tables

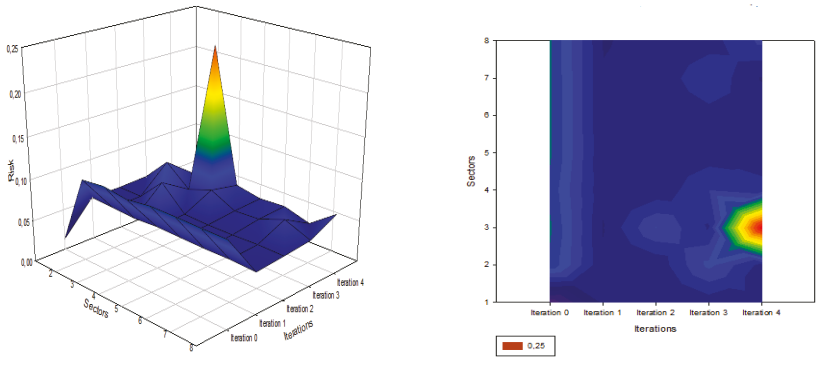


Figure A6. Risk spreading between industries over 1 year horizon (two visualisations on the left and on the right).

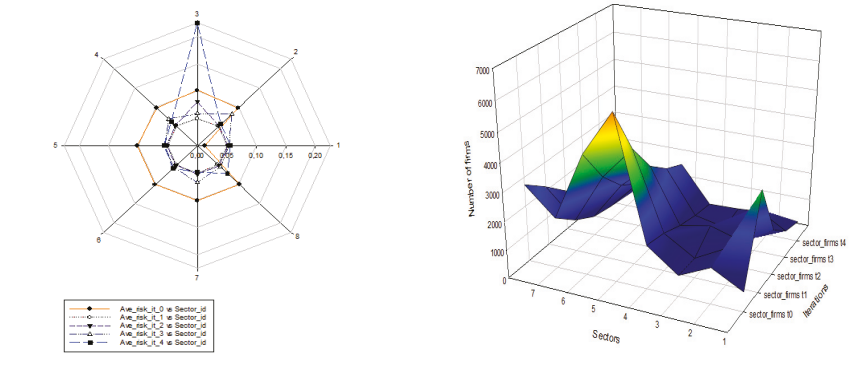


Figure A7. Risk spreading between industries (on the left) & Number of firms in the sectors over 1 year horizon (on the right).

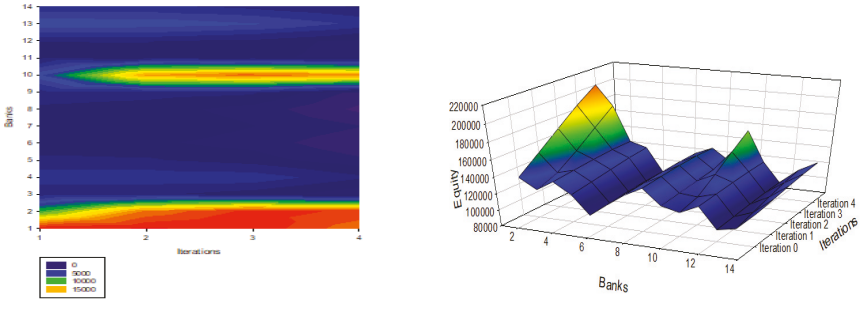


Figure A8. Profits of banks (on the left) & (**b**) Equity of banks over 1 year horizon (on the right).

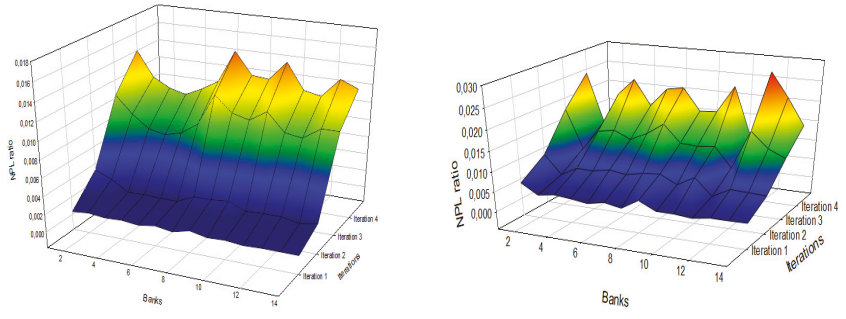


Figure A9. Total NPL ratio of banks (on the left) & NPL ratio of consumers loans of banks over 1 year horizon (on the right).

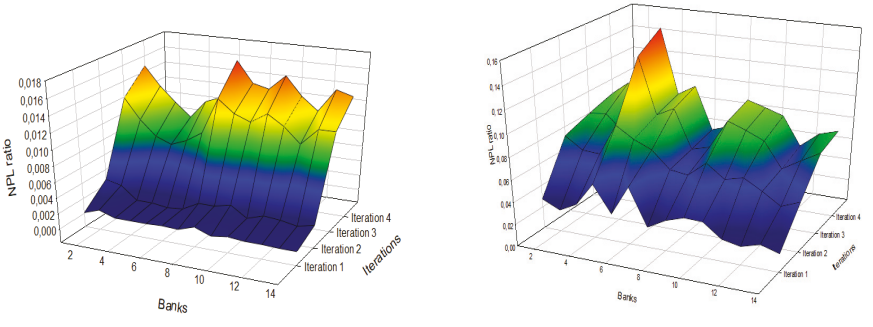


Figure A10. NPL ratio of housing loans of banks (on the left) & (**b**) NPL ratio of firm loans of banks over 1 year horizon (on the right).

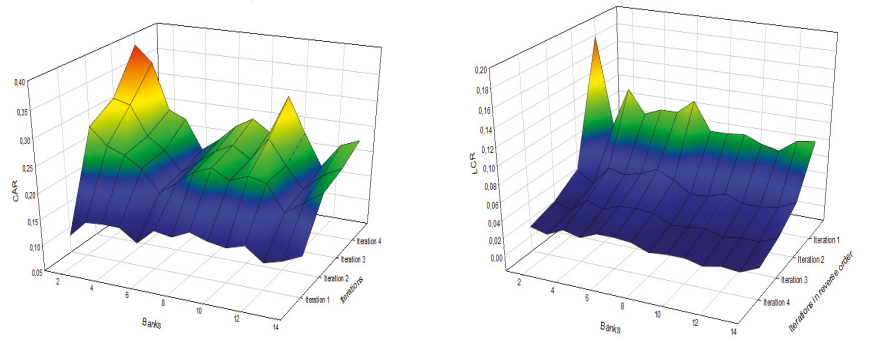


Figure A11. CAR of banks (on the left) & (**b**) LCR of banks over 1 year horizon (on the right).

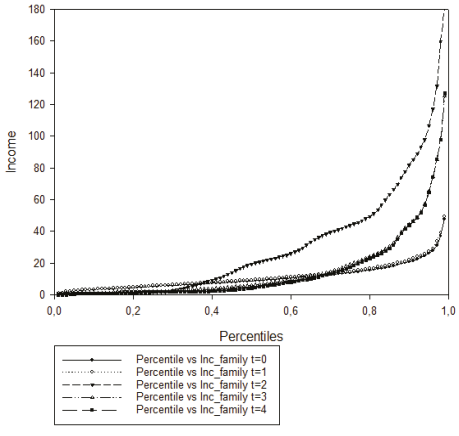


Figure A12. Households' income over 1 year horizon. (*) t = 0 BZGD. From t = 1 to t = 4 simulation results.

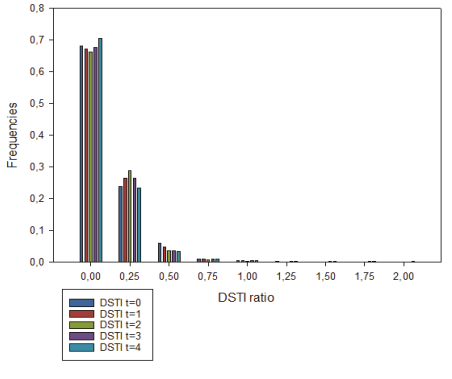


Figure A13. DSTI over 1 year horizon *. (*) t = 0 BZGD. From t = 1 to t = 4 simulation results.

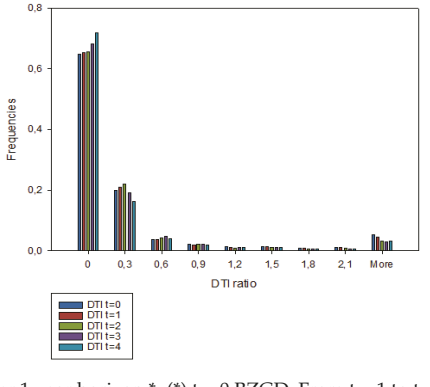


Figure A14. DTI over 1 year horizon *. (*) t = 0 BZGD. From t = 1 to t = 4 simulation results.

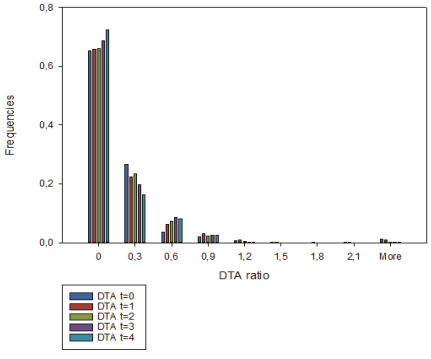


Figure A15. DTA over 1 year horizon *. (*) t = 0 BZGD. From t = 1 to t = 4 simulation results.

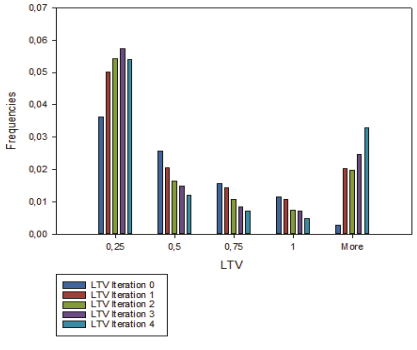


Figure A16. LTV over 1 year horizon *. (*) t = 0 BZGD. From t = 1 to t = 4 simulation results.

Table A11. Leverage of companies.

Value	$t = 0$	$t = 3$
Less than 0.1	0.1099	0.0081
0.1–0.2	0.0994	0.7886
0.2–0.4	0.3740	0.0946
0.4–0.6	0.2149	0.0469
0.6–0.8	0.0713	0.0174
More than 0.8	0.1305	0.0444
	1	1

References

1. Gai, P. *Systemic Risk: The Dynamics of Modern Financial Systems*; Oxford University Press: Oxford, UK, 2013; pp. 1–123, ISBN 9780199544493.
2. Kaszowska, J.; Santos, J.L. The role of risk perception in systemic risk generation and amplification: An agent-based approach. *ACRN J. Financ. Risk Perspect.* **2014**, *3*, 146–170.
3. Barkauskaite, A.; Lakstutiene, A.; Witkowska, J. Measurement of Systemic Risk in a Common European Union Risk-Based Deposit Insurance System: Formal Necessity or Value-Adding Process? *Risks* **2018**, *6*, 137. [CrossRef]
4. Kaszowska, J.; Santos, J.L. Systemic Risk Modeling: Agent-based Modeling Foundations and Validation. In *Systemic Actions in Complex Scenario*; Fabó, E., Chen, J., Eds.; Cambridge Scholars Publishing: Cambridge, UK, 2017; pp. 218–243, ISBN 10:1-5275-0006-3.

5. Galati, G.; Moessner, R. What do we know about the effects of macroprudential policies? *Economica* **2018**, *85*, 735–770. [CrossRef]
6. Zhang, X.; Li, F.; Li, Z.; Xu, Y. Macroprudential Policy, Credit Cycle, and Bank Risk-Taking. *Sustainability* **2018**, *10*, 3620. [CrossRef]
7. Elements of Effective Macroprudential Policies. *Lessons from International Experience*; Jointly publised by IMF-FSB-BIS; 2016; pp. 1–16. Available online: https://www.bis.org/publ/othp26.pdf (accessed on 5 January 2020).
8. Kaszowska, J. Systemic Risk in Space and Time. Forthcoming.
9. Christiano, L.; Eichenbaum, M.; Trabandt, M. On DSGE models. *J. Econ. Perspect* **2018**, *32*, 113–140. [CrossRef]
10. Hitaj, A.; Mateus, C.; Peri, I. Lambda Value at Risk and Regulatory Capital: A Dynamic Approach to Tail Risk. *Risks* **2018**, *6*, 17. [CrossRef]
11. Repullo, R. Capital Requirements, Market Power, and Risk-Taking in Banking. *J. Financ. Intermed.* **2004**, *13*, 156–182. [CrossRef]
12. Repullo, R.; Suarez, J. The Procyclical Effects of Bank Capital Regulation. *Rev. Financ. Stud.* **2013**, *26*, 454–490. [CrossRef]
13. van den Heuvel, S. The welfare cost of bank capital requirements. *J. Monet. Econ.* **2008**, *55*, 298–320. [CrossRef]
14. Drehmann, M.; Borio, C.; Tsatsaronis, K. Anchoring countercyclical capital buffers: The role of credit aggregates. *BIS Work. Pap.* **2011**, *335*, 1–33.
15. Drehmann, M.; Borio, C.; Gambacorta, L.; Jimenez, G.; Trucharte, C. Countercyclical Capital Buffers: Exploring Options. *BIS Work. Pap.* **2010**, *317*, 1–57. [CrossRef]
16. Kowalik, M. Countercyclical capital regulation: should bank regulators use rules or discretion. *Econ. Rev.* **2011**, *96*, 59–80.
17. Repullo, R.; Saurina, S. The Countercyclical Capital Buffer of Basel III. A Critical Assessment. *CEMFI Work. Pap.* **2011**, *1102*, 11–29.
18. Kiema, I.; Jokivuolle, E. Does a leverage ratio requirement increase bank stability? *J. Bank. Finance* **2014**, *39*, 240–254. [CrossRef]
19. Bernanke, B.; Gertler, M.; Gilchrist, S. The financial accelerator in a quantitative business cycle framework. In *Handbook of Macroeconomics*; Taylor, J., Woodford, M., Eds; Elsevier: North-Holland, The Netherlands, 1999; pp. 1341–1393, ISBN 978-0-444-50158-5.
20. Choi, W.; Cook, D. Liability dollarization and the bank balance sheet channel. *J. Int. Econ.* **2004**, *64*, 247–275. [CrossRef]
21. Clerc, L.; Derviz, A.; Mendicino, C.; Moyen, S.; Nikolov, K.; Stracca, L.; Suarez, J.; Vardoulakis, A. Capital regulation in a macroeconomic model with three layers of default. *ECB Work. Pap.* **2015**, *1827*, 1–58. [CrossRef]
22. Gertler, M.; Karadi, P. A model of unconventional monetary policy. *J. Monet. Econ.* **2011**, *58*, 17–34. [CrossRef]
23. Gertler, M.; Kiyotaki, N. Financial Intermediation and Credit Policy in Business Cycle Analysis. In *Handbook of Monetary Economics*; Friedman, B., Woodford, M., Eds; Elsevier: North-Holland, The Netherlands, 2010; Volume 3, pp. 547–599, ISBN 978-0-444-53470-5.
24. Gertler, M; Kiyotaki, N.; Queralto, A. Financial Crises, Bank Risk Exposure, and Government Financial Policy. *J. Monet. Econ.* **2012**, *59*, S17–S34. [CrossRef]
25. Iacoviello, M. House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle. *Am. Econ. Rev.* **2005**, *95*, 739–764. [CrossRef]
26. Caballero, R. Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge Syndrome. *J. Econ. Perspect* **2010**, *24*, 85–102. [CrossRef]
27. Korinek, A. Thoughts on DSGE Macroeconomics: Matching the Moment, But Missing the Point? Paper for the 2015 Conference: 'A Just Society' Honoring J. Stiglitz's 50 Years of Teaching. Available online: http://dx.doi.org/10.2139/ssrn.3022009 (accessed on 7 January 2020).
28. Nuño, G. Constrained Social Optimum with Heterogeneous Agents in Continuous Time. *ECB Work. Pap.* **2014**, *1608*, 1–19.
29. Aoki, K.; Benigno, G.; Kiyotaki, N. Monetary and Financial Policies in Emerging Markets. Mimeo, Princeton University 2016. Available online: https://www.princeton.edu/~kiyotaki/papers/ABKBankModel10-16-2018_GB.pdf (accessed on 11 January 2020).

30. Smets, F.; Wouters, R. An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area. *J. Eur. Econ. Assoc.* **2003**, *1*, 1123–1175. [CrossRef]
31. Christiano, L.; Trabandt, M.; Walentin, K. Introducing financial frictions and unemployment into a small open economy model. *J. Econ. Dyn. Control* **2011**, *35*, 1999–2041. [CrossRef]
32. Adolfson, M.; Laséen, S.; Lindé, J.; Svensson, L. Monetary policy trade-off in an estimated open-economy DSGE model. *J. Econ. Dyn. Control* **2014**, *42*, 33–49. [CrossRef]
33. Christiano, L.; Trabandt, M.; Walentin, K. DSGE Models for Monetary Policy Analysis, In *Handbook of Monetary Economics*; Friedman, B., Woodford, M., Eds.; Elsevier: North-Holland, The Netherlands, 2010; Volume 3A, pp. 285–367, ISBN 978-0-444-53238-1.
34. Greenwald, B.; Stiglitz, J. Financial Market Imperfections and Business Cycles. *Q. J. Econ.* **1993**, *108*, 77–114. [CrossRef]
35. Kiyotaki, N.; Moore, J. Credit cycles. *J. Polit. Econ.* **1997**, *105*, 211–248. [CrossRef]
36. Meh, C.; Moran, K. The role of bank capital in the propagation of shocks. *J. Econ. Dyn. Control* **2010**, *34*, 555–576. [CrossRef]
37. Galati, G.; Moessner, R. Macroprudential policy—A literature review. *J. Econ. Surv.* **2013**, *27*, 846–878. [CrossRef]
38. Hanson, S.; Kashyap, A.; Stein, J. A Macroprudential Approach to Financial Regulation. *J. Econ. Perspect* **2011**, *25*, 3–28. [CrossRef]
39. Osiński, J.; Seal, K.; Hoogduin, L. *Macroprudential and Microprudential Policies: Toward Cohabitation*; IMF Staff Discussion Note; International Monetary Fund: Washington, DC, USA, 2013; pp. 4–28.
40. Assenza, T.; Delli Gatti, D.; Gallegati, M. Financial Instability and Agents' Heterogeneity: A Post Minskyan Research Agenda. In *The Elgar Companion to Hyman Minsky*; Papadimitriou, D., Wray, L., Eds.; Edward Elgar: Cheltenham, UK, 2010; pp. 182–205, ISBN 978-1-84720-849-1.
41. Delli Gatti, D. Financial Instability After Minsky: Heterogeneity, Agent-based Models, and Credit Networks. INET Paper 2012. Available online: https://www.ineteconomics.org/research/research-papers/financial-instability-after-minskyheterogeneity-agent-based-models-and-credit-networks (accessed on 8 January 2020).
42. Kaszowska, J.; Santos, J.L. Rola percepcji ryzyka i niepewności w ocenie ryzyka systemowego. Podejście symulacyjne ABM (The role of risk and uncertainty perception in systemic risk assessment. An Agent-based Simulation). *Ekonomista* **2017**, *4*, 369–399. Available online: http://www.ekonomista.info.pl/?rok=2017&nr=4&lang=0&menu=0&t=1 (accessed on 5 January 2020).
43. Knight, F. Risk. In *Uncertainty and Profit*; Houghton Miffin: Boston, MA, USA; New York, NY, USA, 1921; pp. 1–253.
44. Minsky, H. *Stabilizing an Unstable Economy*; Yale University Press: New Haven, USA, 1986; pp. 1–385.
45. Minsky, H. The financial instability hypothesis. *Levy Econ. Inst. Bard Coll. Work. Pap.* **1992**, *74*, 1–10. [CrossRef]
46. Brunnermeier, M.; Eisenbach, T.; Sannikov, Y. Macroeconomics with Financial Frictions: A Survey. In *Advances in Economics and Econometrics: Tenth World Congress. Econometric Society Monographs*; Acemoglu, D., Arellano, M., Dekel, E., Eds.; Cambridge University Press: Cambridge, UK, 2013; pp. 3–94, ISBN 9781139060028.
47. Brunnermeier, M.; Sannikov, Y. A Macroeconomic Model with a Financial Sector. *Am. Econ. Rev* **2014**, *104*, 379–421. [CrossRef]
48. Brzoza-Brzezina, M.; Kolasa, M.; Makarski, K. The anatomy of standard DSGE models with financial frictions. *J. Econ. Dyn. Control* **2013**, *37*, 32–51. [CrossRef]
49. Bianchi, J. Credit Externalities: Macroeconomic Effects and Policy Implications. *Am. Econ. Rev* **2010**, *100*, 398–402. [CrossRef]
50. Bianchi, J. Overborrowing and Systemic Externalities in the Business Cycle. *Am. Econ. Rev* **2011**, *101*, 3400–3426. [CrossRef]
51. Bianchi, J.; Mendoza, E. Overborrowing, Financial Crises and Macroprudential Taxes. *NBER Work. Pap.* **2010**, *16091*, 1–60.
52. Farhi, E.; Werning, I. A Theory of Macroprudential Policies in the Presence of Nominal Rigidities. *Econometrica* **2016**, *84*, 1645–1704. [CrossRef]

53. An, G.; Fitrzpatrick, B.; Christley, S.; Federico, P.; Kanarek, A.; Miller Neilan, R.; Oremland, M.; Salinas, R.; Laubenbacher, R.; Lenhart, S. Optimization and Control of Agent-Based Models in Biology: A Perspective. *Bull. Math. Biol.* **2017**, *79*, 63–87. [CrossRef]
54. Gilli, M.; Winker, P. A global optimization heuristics for estimating agent-based models. *Comput. Stat. Data Anal.* **2003**, *42*, 299–312. [CrossRef]
55. Oremland, M. Optimization and Optimal Control of Agent-Based Models. Master's Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VI, USA, 3 May 2011.
56. Oremland, M.; Laubenbacher, R. Optimization of Agent-Based Models: Scaling Methods and Heuristic Algorithms. *J. Artif. Soc. Soc. Simul.* **2014**, *17*, 1–16. [CrossRef]
57. Costa, C.; Werning, I. On the optimality of the Friedman Rule with Heterogeneous Agents and Nonlinear Income Tax. *J. Polit. Econ.* **2008**, *116*, 82–112. [CrossRef]
58. Koloch, G. Optymalny Poziom Inflacji W Gospodarce Heterogenicznej (The Optimal Level of Inflation in A Heterogeneous-Agent Economy). Ph.D. Thesis, Warsaw School of Economics, Warsaw, Poland, 2013.
59. Brzoza-Brzezina, M.; Kolasa, M.; Makarski, K. Monetary policy in a non-representative agent economy: A Survey. *J. Econ. Surv.* **2013**, *27*, 641–669. [CrossRef]
60. Fabretti, A. A Markov Chain approach to ABM calibration. In Proceedings of the Social Simulation Conference 2014, Barcelona, Spain, 1–5 September 2014; pp. 404–408.
61. Grazzini, J.; Richiardi, M. Estimation of Ergodic Agent-Based Models by Simulated Minimum Distance. *J. Econ. Dyn. Control* **2015**, *51*, 148–165. [CrossRef]
62. Grazzini, J.; Richiardi, M.; Tsionas, M. Bayesian Estimation of Agent-based Models. *J. Econ. Dyn. Control* **2017**, *77*, 26–47. [CrossRef]
63. Chen, S.-H. Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective. *J. Econ. Dyn. Control* **2012**, *36*, 1–25. [CrossRef]
64. Landini, S.; Gallegati, M.; Stiglitz, J. Economies with heterogeneous interacting learning agents. *J. Econ. Interact. Coord.* **2015**, *10*, 91–118. [CrossRef]
65. Hommes, C. Heterogeneous Agent Models in Economics and Finance. In *Handbook of Computational Economics*; Tesfatsion, L., Judd, K., Eds.; Elsevier: North-Holland, The Netherlands, 2006; pp. 1109–1186, ISBN 978-0-444-64131-1.
66. Hommes, C. The heterogeneous expectations hypothesis: Some evidence from the lab. *J. Econ. Dyn. Control* **2011**, *35*, 1–24. [CrossRef]
67. Hommes, C. *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*; Cambridge University Press: Cambridge, UK, 2013; pp. 1–251, ISBN 9781139094276.
68. Deissenberg, C.; van Der Hoog, S.; Dawid, H. EURACE: A Massively Parallel Agent-Based Model of the European Economy. *Appl. Math. Comput.* **2008**, *204*, 541–552. [CrossRef]
69. Gencer, M.; Ozel, B. EURACE Project Description. Available online: https://ecomod.net/sites/default/files/document-conference/ecomod2010/1316.pdf (accessed on 8 January 2020).
70. Pablo Martí, F.; Santos, J.L.; Kaszowska, J. An Agent-Based Model of Population Dynamics for the European Regions. *Emerg. Complex. Organ.* **2015**, *17*, 1–19. [CrossRef]
71. Caiani, A.; Godin, A.; Caverzasi, E; Gallegati, M.; Kinsella, S.; Stiglitz, J. Agent Based-Stock Flow Consistent Macroeconomics: Towards a Benchmark Model. *J. Econ. Dyn. Control* **2016**, *69*, 375–408. [CrossRef]
72. van der Hoog, S. The Limits to Credit Growth: Mitigation Policies and Macroprudential Regulations to Foster Macrofinancial Stability and Sustainable Debt. *Bielef. Work. Pap. Econ. Manag.*. **2015**, *8*, 1–32. [CrossRef]
73. Baptista, R.; Farmer, J.D.; Hinterschweiger, M.; Low, K.; Tang, D.; Uluc, A. Macroprudential policy in an agent-based model of the UK housing market. *Bank Engl. Staff Work. Pap.* **2016**, *619*, 3–50. [CrossRef]
74. Cincotti, S.; Raberto, M.; Teglio, A. Macroprudential Policies in an Agent-Based Artificial Economy. *Revue de l'OFCE* **2012**, *124*, 205–234. [CrossRef]
75. Fagiolo, G.; Roventini, A. Macroeconomic Policy in DSGE and Agent-Based Models Redux: New Developments and Challenges Ahead. *J. Artif. Soc. Soc. Simul.* **2017**, *20*, 1–37. [CrossRef]
76. Popoyan, L.; Napoletano, M.; Roventini, A. Taming macroeconomic instability: Monetary and macroprudential policy interactions in an agent-based model. *J. Econ. Behav. Organ.* **2017**, *134*, 117–140. [CrossRef]
77. Choi, Y. Masked Instability: Within-Sector Financial Risk in the Presence of Wealth Inequality. *Risks* **2018**, *6*, 65. [CrossRef]

78. Ashraf, Q.; Gershman, B.; Howitt, P. Banks, market organization, and macroeconomic performance: An agent-based computational analysis. *J. Econ. Behav. Organ* **2017**, *135*, 143–180. [CrossRef]
79. Aymanns, C.; Farmer, J.D. The dynamics of the leverage cycle. *J. Econ. Dyn. Control* **2015**, *50*, 155–179. [CrossRef]
80. Battiston, S.; Delli Gatti, D.; Gallegati, B.; Greenwald, B.; Stiglitz, J. Liaisons dangereuses: increasing connectivity, risk sharing, and systemic risk. *J. Econ. Dyn. Control* **2012**, *36*, 8, 1121–1141. [CrossRef]
81. Da Silva, M.; Lima, G. Combining monetary policy and prudential regulation: an agent-based modeling approach. *J. Econ. Interact. Coord.* **2017**. [CrossRef]
82. Delli Gatti, D.; Di Guilmi, C.; Gaffao, G.; Giulioni, G.; Gallegati, M.; Palestrini, A. A new approach to business fluctuations: Heterogeneous interacting agents scaling laws and financial fragility. *J. Econ. Behav. Organ.* **2005**, *54*, 489–512. [CrossRef]
83. Delli Gatti, D.; Gallegati, M.; Greenwald, B.; Russo, A.; Stiglitz, J. The financial accelerator in an evolving credit network. *J. Econ. Dyn. Control* **2010**, *34*, 1627–1650. [CrossRef]
84. Dosi, G.; Fagiolo, G.; Roventini, A. Schumpeter meeting Keynes: A policy-friendly model of endogenous growth and business cycles. *J. Econ. Dyn. Control* **2010**, *34*, 1748–1767. [CrossRef]
85. Dosi, G.; Fagiolo, G.; Napoletano, M.; Roventini, A. Income distribution, credit and fiscal policies in an agent-based keynesian model. *J. Econ. Dyn. Control* **2013**, *37*, 1598–1625. [CrossRef]
86. Dosi, G.; Fagiolo, G.; Napoletano, M.; Roventini, A.; Treibich, T. Fiscal and monetary policies in complex envolving economies. *J. Econ. Dyn. Control* **2015**, *52*, 166–189. [CrossRef]
87. Gai, P.; Haldane, A.; Kapadia, S. Complexity, concentration and contagion. *J. Monet. Econ.* **2011**, *58*, 453–470. [CrossRef]
88. Geanakoplos, J.; Axtell, R.; Farmer, J.D.; Howitt, P.; Conlee, B.; Goldstein, J.; Hendrey, M.; Palmer, N.; Yang, C. Getting at Systemic Risk via an Agent-Based Model of the Housing Market. *Am. Econ. Rev* **2012**, *102*, 53–58. [CrossRef]
89. Klimek, P.; Poledna, S.; Farmer, J.D.; Thurner, S. To bail-out or to bail-in? Answers from an agent-based model. *J. Econ. Dyn. Control* **2015**, *50*, 144–154. [CrossRef]
90. Krug, S. The interaction between monetary and macroprudential policy: Should central banks » lean against the wind« to foster macroprudential stability? *Econ. Open-Access Open-Assess. E-J.* **2018**, *12*, 1–69. Available online: http://dx.doi.org/10.5018/economics-ejournal.ja.2018-7 (accessed on 7 January 2020).
91. Lengnick, M.; Krug, S.; Wohltmann, H. Money creation and financial instability: An agent-based credit network approach. *Economics* **2013**, *7*, 1–44. Available online: http://dx.doi.org/10.5018/economics-ejournal.ja.2013-32 (accessed on 3 January 2020). [CrossRef]
92. Napoletano, M.; Roventini, A.; Gaffard, J.L. Time-varying fiscal multipliers in an agent-based model with credit rationing. *Economics* **2017**, *19*, 1–21.
93. Poledna, S.; Thurner, S.; Farmer, J.D.; Geanakoplos, J. Leverage-induced systemic risk under Basel II and other credit risk policies. *J. Bank. Finance* **2014**, *42*, 199–212. [CrossRef]
94. Raberto, M.; Teglio, A.; Cincotti, S. Debt, deleveraging and business cycles. *Economics*, **2012**, *6*, 1–49. Available online: http://dx.doi.org/10.5018/economics-ejournal.ja.2012-27 (accessed on 8 January 2020).
95. Riccetti, L.; Russo, A.; Gallegati, M. Leveraged network-based financial accelerator. *J. Econ. Dyn. Control* **2013**, *37*, 1626–1640. [CrossRef]
96. Seppecher, P.; Salle, I. Deleveraging crises and deep recessions: A behavioral approach. *Appl. Econ.* **2015**, *47*, 3771–3790. [CrossRef]
97. Mazzocchetti, A.; Raberto, M.; Teglio, A.; Cincotti, S. Securitisation and Business Cycle: An Agent-Based Perspective. *Ind. Corp. Chang.* **2018**, *27*, 1091–1121. [CrossRef]
98. Petroni, F.; Ausloos, M.; Rotundo, G. Generating synthetic time series from Bak-Sneppen coevolution models mixtures. *Physica A* **2007**, *384*, 359–367. [CrossRef]
99. Rotundo, G.; Ausloos, M. Microeconomic co-evolution model for financial technical analysis signals. *Physica A* **2007**, *373*, 569–585. [CrossRef]
100. Varela, L.M.; Rotundo, G.; Ausloos, M.; Carrete, J. Complex network analysis in socioeconomic models. In *Complexity and Geographical Economics*; Commendatore, P., Kayam, S., Kubin, I., Eds.; Springer:Berlin/Heidelberg, Germany, 2015; pp. 209–245. [CrossRef]

101. Varela, L.M.; Rotundo, G. Complex network analysis and nonlinear dynamics. In *Complex Networks and Dynamics*; Commendatore, P., Matilla-García, M., Varela, L., Cánovas, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 3–25. [CrossRef]
102. Ausloos, M.; Herbert, D.; Merlone, U. Spatial interactions in agent-based modeling. In *Complexity and Geographical Economics*; Commendatore, P., Kayam, S., Kubin, I., Eds.; Springer: Berlin/Heidelberg, Germany, 2015; pp. 353–377, ISBN 978-3-319-12804-7.
103. Boeing, G. The Effects of Inequality, Density, and Heterogeneous Residential Preferences on Urban Displacement and Metropolitan Structure: An Agent-Based Model. *Urban Sci.* **2018**, *2*, 76. [CrossRef]
104. Richiardi, M.; Leombruni, R.; Saam, N.; Sonnessa, M. A Common Protocol for Agent-Based Social Simulation. *J. Artif. Soc. Soc. Simul.* **2006**, *9*, 1–22.
105. Chen, J.M. Gini's Crossbow. Presented at the 81st International Atlantic Economic Society Conference, Lisbon, Portugal, 16–19 March 2016. Available online: http://dx.doi.org/10.2139/ssrn.2608850 (accessed on 8 January 2020).
106. Gini, C. Concentration and dependency ratios. *Rivista di Politica Economica* **1909**, *87*, 769–789.
107. Lorenz, M. Methods of measuring the concentration of wealth. *Am. Stat. Assoc. Publ.* **1905**, *9*, 209–219. [CrossRef]
108. Grejcz, K.; Żółkiewski, Z. Household wealth in Poland: the results of a new survey of household finance. *Bank i Kredyt* **2017**, *48*, 295–326.
109. Household Budget Survey in 2014. Available online: https://stat.gov.pl/en/topics/living-conditions/living-conditions/household-budget-survey-in-2014,2,9.html (accessed on 8 January 2020).
110. European Union Statistics on Income and Living Conditions (2013). Available online: https://stat.gov.pl/obszary-tematyczne/warunki-zycia/dochody-wydatki-i-warunki-zycia-ludnosci/europejskie-badanie-dochodow-i-warunkow-zycia-eu-silc-w-2013-r-,7,5.html (accessed on 8 January 2020).
111. Czapiński, J.; Panek, T. Social diagnosis 2015. Objective and Subjective Quality of Life in Poland. *Contemp. Econ.* **2015**, *9*, 13–526. Available online: http://www.diagnoza.com/pliki/raporty/Diagnoza_raport_2015.pdf (accessed on 8 January 2020).
112. Household Wealth and Debt in Poland. Pilot Survey Report 2014. National Bank of Poland Report. Available online: https://www.nbp.pl/en/publikacje/inne/bzgd/bzgd_2014_en.pdf (accessed on 8 January 2020).
113. Eurostat, Gini Coefficient of Equivalised Disposable Income—EU-SILC Survey 2013. Available online: https://data.europa.eu/euodp/en/data/dataset/rkeZ1htI2J2YycyRb2rBXg (accessed on 8 January 2020).
114. Sierminska, E.; Medgyesi, M. *The Distribution of Wealth Between Households*; European Commission Report; 2013; pp. 1–30. Available online: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=2ahUKEwjSy9rWiZDnAhWC-yoKHShKDvAQFjACegQIAhAB&url=https%3A%2F%2Fec.europa.eu%2Fsocial%2FBlobServlet%3FdocId%3D11782%26langId%3Den&usg=AOvVaw2GdUJRp0z8WdwunE-joZVp (accessed on 8 January 2020).
115. Damgaard, C.; Weiner, J. Describing inequality in plant size or fecundity. *Ecology* **2000**, *81*, 1139–1142. [CrossRef]
116. Kotz, S.; Johnson, N.L.; Read, C.B. *Encyclopedia of Statistical Sciences*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1983.
117. Shumway, D.; Koide, R. Size and Reproductive Inequality in Mycorrhizal and Nonmycorrhizal Populations in Abutilon Theophrasti. *J. Ecol.* **1995**, *83*, 613–620. [CrossRef]
118. Schorrocks, A.F. The Class of Additively Decomposable Inequality Measures. *Econometrica* **1980**, *48*, 613–625. [CrossRef]
119. Novotný, J. On the measurement of regional inequality: Does spatial dimension of income inequality matter? *Ann. Reg. Sci.* **2007**, *41*, 563–580. [CrossRef]
120. Hartmann, P. Interactions Between Macroprudential and Other Policies. Presented at the SUERF/National Bank of Poland Conference: Challenges of Interactions between Macroprudential and other Policies, Warsaw, Poland, 15 February 2019.
121. Claessens, S. Moving forward with macroprudential framework. Presented at the SUERF/National Bank of Poland Conference: Challenges of Interactions between Macroprudential and other Policies, Warsaw, Poland, 15 February 2019.

122. Household Wealth and Debt in Poland. Methodological Annex to the Pilot Survey 2014, National Bank of Poland Report. Available online: https://www.nbp.pl/en/publikacje/inne/bzgd/bzgd_2014_metodology_en.pdf (accessed on 8 January 2020).
123. Population. Size and Structure of Population and Vital Statistics in Poland by Territorial Division. As of December 31, 2014. Available online: https://stat.gov.pl/en/topics/population/population/population-size-and-structure-of-population-and-vital-statistics-in-poland-by-territorial-divison-as-of-december-31-2014,3,13.html (accessed on 8 January 2020).
124. Structure of the Population by 2016: Size and Structure of Population by Age Groups in 1989–2016. Available online: https://stat.gov.pl/en/topics/population/population/structure-of-the-population-by-2016,7,1.html (accessed on 8 January 2020).
125. Demographic Yearbook of Poland: Marriages Contracted and Dissolved, Central Statistical Office Report; pp. 189–256. Available online: https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5515/3/11/1/rocznik_demograficzny_2017.pdf (accessed on 8 January 2020).
126. Structure of the Population by 2016: Births in 1970–2016. Available online: https://stat.gov.pl/en/topics/population/population/structure-of-the-population-by-2016,7,1.html (accessed on 8 January 2020).
127. Demographic Yearbook of Poland 2015. Available online: https://stat.gov.pl/en/topics/statistical-yearbooks/statistical-yearbooks/demographic-yearbook-of-poland-2015,3,9.html (accessed on 8 January 2020).
128. Life Expectancy by Age in Poland 1950–2018. Available online: https://stat.gov.pl/en/topics/population/life-expectancy/life-expectancy-in-poland,1,3.html (accessed on 3 January 2020).
129. Education in the School Year 2014/2015. Available online: https://stat.gov.pl/en/topics/education/education/education-in-the-school-year-20142015,1,7.html (accessed on 8 January 2020).
130. Yearbook of Labour Statistics 2015. Available online: https://stat.gov.pl/en/topics/statistical-yearbooks/statistical-yearbooks/yearbook-of-labour-statistics-2017,10,6.html (accessed on 8 January 2020).
131. Social Assistance, Child and Family Services in 2014. Available online: https://stat.gov.pl/en/topics/living-conditions/social-assistance/social-assistance-child-and-family-services-in-2016,1,8.html (accessed on 8 January 2020).
132. Activity of Non-Financial Enterprises in 2014. Available Online: https://stat.gov.pl/en/topics/economic-activities-finances/activity-of-enterprises-activity-of-companies/activity-of-non-financial-enterprises-in-2014,1,11.html (accessed on 8 January 2020).
133. Input—Output Table at Basic Prices in 2010. Available online: https://stat.gov.pl/en/topics/national-accounts/annual-national-accounts/input-output-table-at-basic-prices-in-2010,5,2.html (accessed on 8 January 2020).
134. MIR Interest Rate Statistics. National Bank of Poland Page. Available online: https://www.nbp.pl/en/statystyka/oproc/mir_new/manual_mir.pdf or https://www.nbp.pl/homen.aspx?f=/en/statystyka/oproc/oproc.html (accessed on 5 January 2020).

entropy

MDPI

Article

Innovativeness of Industrial Processing Enterprises and Conjunctural Movement

Aleksander Jakimowicz [1,*] and Daniel Rzeczkowski [2]

[1] Department of World Economy, Institute of Economics, Polish Academy of Sciences, Palace of Culture and Science, 1 Defilad Sq., 00-901 Warsaw, Poland

[2] Department of Market and Consumption, Institute of Economics and Finance, Faculty of Economic Sciences, University of Warmia and Mazury in Olsztyn, 1/327 Cieszyński Sq., 10-720 Olsztyn, Poland; daniel.rzeczkowski@uwm.edu.pl

* Correspondence: ajakimowicz@inepan.waw.pl

Received: 10 August 2020; Accepted: 17 October 2020; Published: 19 October 2020

Abstract: Singulation of components determining the innovative activity of enterprises is a complex issue as it depends on both microeconomic and macroeconomic factors. The purpose of this article is to present the results of research on the impact of the mutual interactions between ownership and the size of companies on the achievement of the objectives of innovative activity by Polish industrial processing enterprises in changing cyclical conditions. The importance of innovation barriers was also assessed. Empirical data came from three periods that covered different phases of the business cycle: prosperity 2004–2006, global financial crisis 2008–2010, and recovery 2012–2014. The research used a cybernetic approach based on feedback loops presenting interactions between variables. In addition, two statistical methods were used: the Pearson's χ^2 independence test and correspondence analysis. The following discoveries were made during the research: (1) consideration of the combined impact of ownership and the size of companies on their innovation activities makes it possible to study phenomena that may be overlooked if the impact of these factors is considered separately; (2) public enterprises achieve significantly worse results in terms of innovation than companies from other ownership sectors; (3) the Red Queen effect, which assumes that the best innovative enterprises exert selection pressure on all other companies, applies to industrial processing companies, and in particular public enterprises; (4) the industrial processing section is more sensitive to secular trends than to cyclical fluctuations; (5) confirmation of occurrence of the Polish Green Island effect, which assumes that companies achieve good results in terms of innovation, irrespective of the phases of the business cycle; and (6) statistical evidence is provided that the global financial crisis may be associated with the turn of the Fifth and Sixth Kondratieff waves. Most likely, the role of the communication channel between the world economy and the Polish manufacturing section is fulfilled by foreign ownership, whose percentage of share capital of this section is estimated at 50%.

Keywords: macroeconomics; innovative activity; manufacturing industry; conjunctural movements; cybernetics; feedback loops; correspondence analysis; Polish Green Island effect; Red Queen effect; Kondratieff waves

1. Introduction

The explanation of the relationship between innovation and conjunctural phenomena is one of the most important problems of modern economics. By their very nature, the processes that should be taken into account during such studies are long-term. Therefore, they should be considered not only within the time frame appropriate for traditionally understood business cycles, but also from the perspective of secular changes. At the heart of conjuncture theory (as well as of business cycle theory) is the division of the stunning complexity of economic interactions into a number

of heterogeneous forms of movement according to the criterion of their duration. According to the guidelines of the Harvard Business School, we can distinguish the following forms of movements: (1) the fundamental course of movement (or secular trend), (2) seasonal fluctuations, (3) cyclical fluctuations (conjuncture in the narrower sense), and (4) miscellaneous random fluctuations. With this approach to the problem, business cycle theory tackles the cyclical movements in a narrower sense, and thus variations of economic phenomena that are recurring in free rhythm [1]. Conjunctural movements refer to the entire wavelike evolution of economic life, so by definition, they include secular changes, which is well documented in the literature [2–6]. This is the justification for the use of the mentioned term in the title of the article, as we believe that it is advisable to extend the time frame of studies on innovation beyond those that are appropriate for traditionally understood business cycles.

It is difficult to imagine an increase in the innovative activity of enterprises without prior changes in production techniques, which usually occur over long periods. This implies the need to study interdependencies between innovation and traditional business cycles and secular cycles. Technological revolutions belong to the secular factors of economic growth and development and are the cause of supercycles or Kondratieff waves (K-waves), which last from 48 to 60 years. Changes occurring over several decades, as opposed to business cycles, are usually caused by extra-economic circumstances and events. According to Kondratieff, secular factors can be divided into the following four groups [7]: (1) changes in technology, (2) wars and revolutions, (3) the assimilation of new countries into the world economy, and (4) fluctuations in gold production. In the case of the Polish industrial processing section, at least two factors should be taken into account, the first and third, and maybe even all of them.

The modification of the Kondratieff long-waves theory was made by Šmihula, according to whom the global financial crisis is a phenomenon typical of the breakthrough associated with the end of one and the beginning of the next K-wave [8]. In this case, it would be a transition from the Fifth to Sixth K-Wave, which would mean the end of the information and telecommunications revolution and the initiation of the biomedical-hydrogen revolution. According to Šmihula, this breakthrough dates to 2015. The technological innovations underlying the secular cycles are usually the result of earlier technological revolutions. In addition, all long-waves have common features that are the cause of certain patterns of economic development. Each of them consists of two phases: the innovation phase in which the inventions find practical applications, and the application phase in which the existing innovative solutions are improved and integrated into everyday economic life. The end of a given wave of innovations is determined by the decreasing rate of profit from a new innovation to the level appropriate for traditional branches of industry. Thereby, a given technology achieves its proper development limit. In order to cross the limit, a new innovative technology is required. Thus, the end of the application phase of each K-wave signifies a period of stagnation caused by the economic crisis, which can be overcome by increased demand for new inventions and revolutionary technological innovations. In addition, Šmihula made an interesting observation regarding the shortening of the length of successive K-waves [8]. This implies the constant acceleration of technological development, which—if this trend continues—may lead in the years 2080–2090 to blurring of the differences between K-waves and classical business cycles.

Schumpeter's theory of innovation, which has been elaborated in the first half of the 20th century is the milestone of the study of interdependencies between innovation and economic growth and development. Schumpeter perceived innovation as the driving force of the economy, which on one hand ensures its development, but on the other hand, is the source of the business cycle because it brings the economy out of balance. In his opinion, innovations are at the center of almost all socio-economic phenomena, and the length of the two basic phases of the business cycle, prosperity and recession, depend on the essential features of the innovation that underlie the cycle [9]. The main figure in his theory is the entrepreneur, whose basic task is to search for new combinations of productive means, therefore, such that were not created as a result of improving existing combinations. These include the following five cases [10]: (1) placement of new products on the market, (2) implementation of new

production methods, (3) opening of new markets, (4) acquisition of a new source of raw materials or semi-manufactured goods, and (5) introduction of a new organization of any industry.

Schumpeter's interests also included issues related to the impact of ownership and size of enterprises on their innovation activity. The essence of ownership is the ability to freely dispose of means of production, which can be used directly to create a new combination of forces and materials or can be exchanged for the necessary goods and services. In his opinion, during the capitalist process, dematerialization of property occurs, as a result of which property ceases to perform its basic functions in business. Elimination of the material substance of property, which is done by exchanging factory walls and machinery for a mere parcel of shares, deprives ownership of its most important feature, which is moral allegiance. In this way, the holders of the title cannot freely dispose of their property. With respect to the size of the company, it is very important for innovation. In a competitive economy, new enterprises are the carriers of innovation, and these companies are not necessarily large. The situation is changed by the emergence of huge concerns that reduce the competitiveness of the economy and gain an advantage in the field of innovation over smaller companies due to their size [10,11]. The theory of innovation presented here shows that ownership and size are the basic factors determining the innovativeness of enterprises, therefore their impact must be considered together. Thus, innovations and related economic phenomena should be viewed from the point of view of cybernetics, where the importance of feedback loops is emphasized.

The aim of the article is to complete the gap in research on innovation, which consists of the failure to define the interrelationships between the innovative activity of enterprises at the microeconomic level and long waves occurring at the macroeconomic level. Most studies have focused on the separate impact of variables such as the type of enterprise and ownership sector on innovation activity and barriers to innovation, however, it is very important to capture the combined impact of these variables. The article proves that consideration of the combined impact of the type and ownership sector of enterprises on their innovative activity allows for the discovery of previously unknown economic phenomena. It should be emphasized that the adopted research methodology is characteristic of the complexity economics and therefore the macroeconomic and microeconomic levels are not distinguished [12] (pp. 97, 161–185). This allowed for the following discoveries: (1) consideration of the combined impact of ownership and the size of companies on their innovation activities makes it possible to study phenomena that may be overlooked if the impact of these factors is examined separately; (2) public enterprises achieve significantly worse results in terms of innovation than companies from other ownership sectors; (3) the Red Queen effect, which assumes that the best innovative enterprises exert selection pressure on all other companies, applies to industrial processing companies, and in particular public enterprises; (4) the manufacturing sector is more sensitive to secular trends than to cyclical fluctuations; (5) confirmation of occurrence of the Polish Green Island effect, which assumes that companies achieve good results in terms of innovation, irrespective of the phases of the business cycle; and (6) statistical evidence is provided that the global financial crisis may be associated with the turn of the Fifth and Sixth Kondratieff waves.

As noted by W. Brian Arthur [13] (pp. 16–17), the source of the complexity of economic systems is the presence of both negative and positive feedbacks, the effects of which overlap. The feedbacks from the manufacturing sector are presented below in cybernetic diagrams (2)–(8), which are used to study the mutual interactions between: ownership sector and enterprise type, ownership sector + enterprise type and innovative activity, ownership sector + enterprise type and barriers to innovation, barriers to innovation and innovative activity, and innovative activities in different periods. In many cases, the strength of the feedbacks is examined on a four-point scale: high, medium, low, and irrelevant. It is observed that—depending on specific conditions—all these feedbacks can be both positive and negative, and their effects overlap, which creates a complex pattern of industrial processing in Poland. The innovative activity of enterprises also depends on many other factors such as the national and international environment [14–16], management system [17], business support organizations [18,19], intellectual assets [20], sectoral patterns of cooperation and technology level [21], for which the *ceteris*

paribus assumption was made. It can also affect other economic variables such as total factor productivity [22], the level of firm productivity [23], and anti-crisis reputational sustainability [24]. These dependencies create additional feedback loops that spread throughout the economy and increase its complexity.

In the manufacturing sector in Poland, there are also at least four other sources of complex dynamics, apart from the positive and negative feedbacks discussed above. First, the innovative activity of enterprises is more dependent on secular factors than on the phases of the business cycle. Another big surprise was the steady decline in the significance of innovation barriers in successive periods prosperity (2004–2006), global financial crisis (2008–2010), and recovery (2012–2014). This brings to mind the fractal market hypothesis, which applies to capital markets, and highlights the importance of time scales in which investors operate [25]. Second, the industrial processing sector can be considered as a complex adaptive system in the meaning of Gell-Mann [26], because the operation of its companies is based on the creation and improvement of schemas or models describing the regularities observed in the environment. These schemas are then used by companies to operate in the real world. Third, the sector under study can be viewed from the point of view of thermoeconomics and trends analyzed as a result of changes in thermodynamic entropy and money entropy [27]. Fourth, entropy and information are closely related, leading to the conclusion that entropy can be used in economics to measure ignorance or uncertainty. Information and ignorance are opposites, but the measurement of one quantity can determine the other [28]. Ignorance related to thermodynamic entropy applies to both innovative strategies of industrial processing enterprises and the government's pro-innovation policy, therefore it is both microeconomic and macroeconomic in nature. Ignorance related to money entropy concerns the disorder of monetary policy, so it is only macroeconomic in nature.

2. Materials and Methods

2.1. General Characteristics of the Cybernetic Research Approach

Contemporary studies on innovation are dominated by the concept that ownership and the type of enterprise are treated as one of the most important variables determining the innovative activity of enterprises [29–32]. However, the impact of these factors is relatively often considered separately. This reasoning can be represented using mathematical formalism in the form of the following logical dependence:

$$\begin{array}{c} \boxed{Ownerships\ sectors} \\ \neg\,(\uparrow\downarrow) \\ \boxed{Types\ of\ enterprises} \end{array} \begin{array}{c} \xrightarrow{Unidirectional\ impact} \\ \neg\,(\uparrow\downarrow) \\ \xrightarrow{Unidirectional\ impact} \end{array} \boxed{\begin{array}{c} Innovation \\ activities \\ of\ enterprises \end{array}} \tag{1}$$

Negation symbol preceding two opposite arrows in brackets $\neg\ (\uparrow\downarrow)$ signifies bypassing the interactions between the ownership and size of the company, so each of these variables affects the innovation of companies separately. Furthermore, relationships between independent variables and the dependent variable are unidirectional. Feedbacks between innovation activities and the cumulative interaction of ownership and company size are therefore not included. Elimination of interactions between the ownership sector and the type of enterprise may critically affect the obtained results. Some publications recognize this problem [33,34].

This study used a cybernetic approach to the problem, which emphasizes the importance of feedback loops. After considering them, the logical relationship (1) is transformed into the following form:

$$\begin{array}{c} \boxed{Ownerships\ sectors} \\ \uparrow\ \downarrow \\ \boxed{Types\ of\ enterprises} \end{array} \begin{array}{c} \longrightarrow \\ \longrightarrow \\ \longrightarrow \end{array} \boxed{\begin{array}{c} Innovation \\ activities \\ of\ enterprises \end{array}} \tag{2}$$

$$\xleftarrow{Impact\ of\ innovation\ activity\ on\ ownership\ and\ type}$$

Such formulation of the problem allows for the examination of the impact of mutual interactions between ownership and type of enterprises on the innovation activity of enterprises. It also recognizes another, no less important feedback loop between variables containing a response record in the form of the impact of innovation activity on ownership and size of the enterprise. Therefore, the cybernetic scheme (2) is consistent with Schumpeter's theory of innovation.

The cybernetic approach was aimed at simultaneous examination of the following phenomena: (1) determination of the mutual influence of ownership sectors and types of enterprises on the innovative activity of companies, (2) determination of the mutual impact of ownership sectors and types of enterprises on innovation barriers, and (3) registration of the impact of barriers to innovation on the innovative activity of enterprises. In addition, other feedback loops were included depending on the information contained in individual databases. Innovations implemented in a given period may contribute to the growth of innovations in subsequent years by affecting independent variables [35–37].

The research used three databases on Polish industrial processing enterprises, each of which covered one of the following periods: prosperity from 2004–2006, global financial crisis from 2008–2010, and recovery from 2012–2014. The first database contained 10,149 enterprises, the second included 20,655, and the third 10,244 (Table 1) [36]. The data were collected by the Statistical Office in Szczecin based on the PNT-02 questionnaires, which were subject to some modifications in the above-mentioned periods, but these changes were not significant enough to polarize observations in a way that could hinder the comparability of the research results. Nevertheless, there were some differences in the design of questionnaires in each of the periods, which as a consequence necessitated the development and adaptation of the relationship (2) to each of the three-year periods of analysis [35–37].

Three ownership sectors, public, private, and mixed (50% public, 50% private), and three types of enterprises distinguished on the basis of the size criterion (i.e., small, medium and large were considered in this study). The typology of enterprises is based on European Union standards, where certain thresholds are considered in the form of the number of employees and the volume of annual turnover or the annual balance sheet total (Table 2) [38].

The research adopted the following method for encoding variables. Small, medium, and large enterprises were indicated by symbols FR_1, FR_2, and FR_3, respectively, while ownership sectors—public, private, and mixed—were represented by the symbols S1, S2, and S3, respectively. The analyses concerning the combined impact of ownership sectors and types of enterprises on the objectives of innovative activity or barriers to innovation, which are presented in logical relationships (3)–(8), used two-part designations that first identified the ownership sector, and second, the type of enterprise. In this convention, medium-sized private sector enterprises were represented by the symbol S2FR_2.

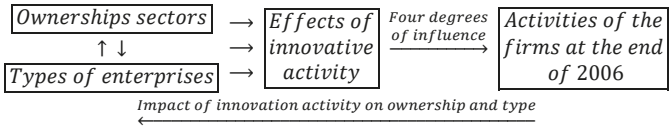
Tables 3–5 contain the percentage data referring to types and ownership sectors of enterprises in the three periods under examination (i.e., 2004–2006, 2008–2010, and 2012–2014) [37]. Each of these tables shows both the share of particular types of enterprises in the ownership sectors and the share of particular ownership sectors in the types of enterprises. Tables 3–5 are to be read as follows. The Type column provides the percentage share of a given type of enterprise in individual ownership sectors. Table 3 shows that in the first period, medium-sized enterprises (FR_2) accounted for 4.14% of the public sector (S1), 83.64% of the private sector (S2), and 12.22% of the mixed sector (S3). The Subtotal (FR) row contains the percentage shares of each type of enterprise in the total number of enterprises. It indicates that in the prosperity period (2004–2006), small enterprises accounted for 30.06%, medium enterprises for 55.66%, and large enterprises for 14.28% of the total number of enterprises. With regard to the Sector column, it represents the share of a given ownership sector in each type of enterprise, which is read horizontally, taking into account every second cell of a given row. To clarify this, the Mixed (S3) row can be examined here. It demonstrates that the mixed sector comprised 44.84% of small enterprises, 41.87% of medium-sized enterprises, and 13.29% of large enterprises. The Subtotal (S) column shows the percentage of enterprises from the given ownership sector in the total number of enterprises. As can be inferred from Table 3, in 2004–2006, the public sector (S1) included 4.37%, the private sector (S2)

was 79.39%, and the mixed sector (S3) was 16.24% of the total number of the investigated enterprises. Tables 4 and 5 should be read in the same way. The data provided in Tables 3–5 relate to the role and importance of individual types and ownership sectors of enterprises in the whole section of industrial processing. They enable precise interpretation of correspondence maps showing the co-occurrence of points representing the types and ownership sectors of enterprises, points indicating the effects (objectives) of innovative activity, and points responsible for innovation barriers.

In this study, ownership sectors and types of enterprises are grouping variables. This results from the adoption of the Schumpeterian point of view, according to which these variables and the interactions between them exert a crucial influence on the effects and objectives of innovative activities undertaken by companies. In this way, the data were sorted into categories and groups with clear economic sense.

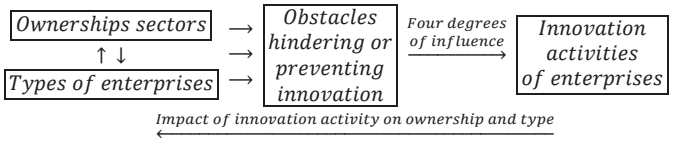
2.2. Feedback Loops in the Years of Prosperity

In the period of prosperity from 2004–2006, the impact of the mutual interactions of ownership sectors and types of enterprises on the effects of innovation activities and the degrees of influence of each of them on the activities of enterprises at the end of 2006 were determined. This problem can be illustrated by the following logical dependence:



(3)

Relationship (3) contains an additional feedback loop between the effects of innovation activities in 2004–2006 and the activities of companies at the end of 2006. Four degrees of impact are possible: high, medium, low, and irrelevant. In addition, during the prosperity period, it was necessary to recognize the impact of interactions between ownership sectors and business types on barriers to innovation, as illustrated by another relationship:



(4)

Innovation barriers may hinder enterprises from conducting innovative activities and even influence the decision not to conduct such activities. In this case, there are four degrees of the impact of innovation barriers on innovation activities: high, medium, low, and irrelevant. Undoubtedly, the considered feedbacks between innovation activities conducted at different times, as illustrated by relationship (3), and feedbacks between innovation activities and innovation barriers, as shown by relationship (4), are mediated by independent variables (i.e., ownership sectors and types of enterprises).

Nine effects of innovative activity and eleven barriers to innovation were taken into account during the period under study. Each of these variables can occur in four states, which, in conjunction with the nine states that ownership sectors and types of enterprises can collectively adopt, indicates the need to consider the simultaneous relationships between eighty-nine variable states.

Table 1. The number of enterprises depending on types and sectors of ownership in the three analyzed databases.

Type/ Ownership Sector (Code)	Database								
	2004–2006			2008–2010			2012–2014		
	Small (FR_1)	Medium (FR_2)	Large (FR_3)	Small (FR_1)	Medium (FR_2)	Large (FR_3)	Small (FR_1)	Medium (FR_2)	Large (FR_3)
Public (S1)	90	234	120	52	119	73	20	54	39
Private (S2)	2222	4725	1110	10,187	4327	1012	2052	1677	906
Mixed (S3)	739	690	219	3560	1039	286	1522	3467	507
Subtotal	3051	5649	1449	13,799	5485	1371	3594	5198	1452
Total		10,149			20,655			10,244	

Table 2. Typology of enterprises in the light of European Union standards and the method of coding.

Types of Enterprise/Code	Number of Employees (NE, in Persons)	Annual Turnover (AT, in EUR Million)	Annual Balance Sheet Total (ABS, in EUR Million)
Micro	NE < 10	AT ≤ 2	ABS ≤ 2
Small (FR_1)	10 ≤ NE < 50	2 < AT ≤ 10	2 < ABS ≤ 10
Medium (FR_2)	50 ≤ NE < 250	10 < AT ≤ 50	10 < ABS ≤ 43
Large (FR_3)	NE ≥ 250	AT > 50	ABS > 43

Table 3. Percentage share of enterprise types (FR) in ownership sectors (S) and percentage share of ownership sectors in enterprise types in the period 2004–2006.

Database 2004–2006 (%)							
Type/Ownership Sector (Codes)	**Small (FR_1)**		**Medium (FR_2)**		**Large (FR_3)**		**Subtotal (S)**
	Type	**Sector**	**Type**	**Sector**	**Type**	**Sector**	
Public (S1)	2.95	20.27	4.14	52.70	8.28	27.03	4.37
Private (S2)	72.83	27.58	83.64	58.64	76.61	13.78	79.39
Mixed (S3)	24.22	44.84	12.22	41.87	15.11	13.29	16.24
Subtotal (FR)	30.06		55.66		14.28		Total = 100

Table 4. Percentage share of enterprise types (FR) in ownership sectors (S) and percentage share of ownership sectors in enterprise types in the period 2008–2010.

Database 2008–2010 (%)							
Type/Ownership Sector (Codes)	**Small (FR_1)**		**Medium (FR_2)**		**Large (FR_3)**		**Subtotal (S)**
	Type	**Sector**	**Type**	**Sector**	**Type**	**Sector**	
Public (S1)	0.38	21.31	2.17	48.77	5.33	29.92	1.18
Private (S2)	73.82	65.61	78.89	27.87	73.81	6.52	75.17
Mixed (S3)	25.80	72.88	18.94	21.27	20.86	5.85	23.65
Subtotal (FR)	66.81		26.55		6.64		Total = 100

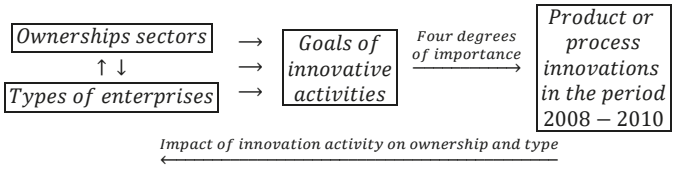
Table 5. Percentage share of enterprise types (FR) in ownership sectors (S) and percentage share of ownership sectors in enterprise types in the period 2012–2014.

Database 2012–2014 (%)							
Type/Ownership Sector (Codes)	**Small (FR_1)**		**Medium (FR_2)**		**Large (FR_3)**		**Subtotal (S)**
	Type	**Sector**	**Type**	**Sector**	**Type**	**Sector**	
Public (S1)	0.56	17.70	1.04	47.79	2.68	34.51	1.10
Private (S2)	57.09	44.27	32.26	36.18	62.40	19.55	45.25
Mixed (S3)	42.35	27.69	66.70	63.08	34.92	9.23	53.65
Subtotal (FR)	35.09		50.74		14.17		Total = 100

2.3. Feedback Loops in the Years of the Global Financial Crisis

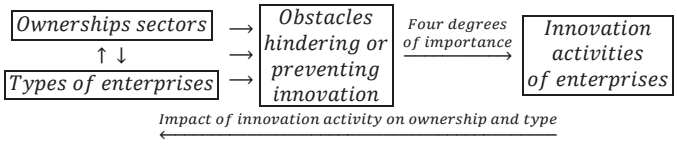
During the global financial crisis, there were slight changes in the statistical form PNT-02, which consisted of replacing effects with the goals of innovative activity and degrees of influence with degrees of importance. This time, the impact of the achieved goals on innovative activity in the field of product and process innovations in the years 2008–2010 was taken into account. The scale of impact included four degrees of importance: high, medium, low, and irrelevant. The overall cumulative impact of

ownership sectors and types of enterprises on the goals of innovative activities and the degrees of their impact on the activities in the field of product and process innovation is presented in the cause and effect loops in the form of:



(5)

During this period, the impact of mutual interactions of ownership and enterprise type on innovation barriers and the four-level significance of these barriers for the innovative activity of companies were examined. Barriers to innovation may have impeded the conducting of innovation activities or influenced the decision not to conduct such activities. It is important to consider the impact of innovative activity on innovation barriers through the cumulative impact of ownership and types of enterprises. Therefore, the following logical relationship is addressed below:

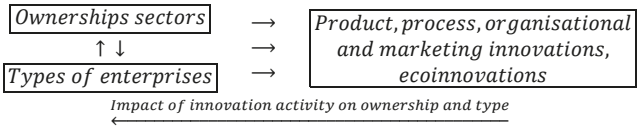


(6)

During the crisis years, the goals of innovative activities are described by ten variables and the barriers to innovation by eleven variables. All these variables can be in four states, which, together with nine states of ownership sectors and types of enterprises, indicates the necessity to examine interdependencies between ninety-three states of variables.

2.4. Feedback Loops during the Recovery Period 2012–2014

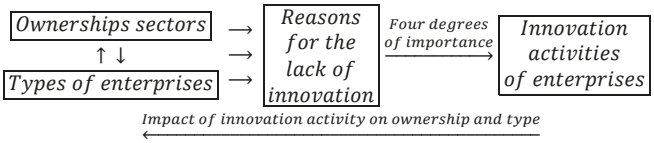
During the recovery period 2012–2014, the goals of the innovative activity of enterprises included four types of traditional innovations (product, process, organizational, and marketing innovations) and eco-innovations. During this period, the determination of the degrees of influence (importance) of these objectives on the further innovation activities of companies was abandoned. The feedback loops used in the research have the following form:



(7)

In the years of prosperity, as in the previous two periods, the joint impact of ownership sectors and types of enterprises on innovation barriers was also examined. The barriers to innovations included eleven variables, and their importance to the innovation activities of companies were on a four level scale. The barriers may have contributed to the lack of innovations to a high, medium, and low degree

or be irrelevant. The study of the importance of innovation barriers for the innovative activities of enterprises was based on the following feedback loops:



(8)

During this period, twenty-two goals of innovative activities were one-state variables, while eleven barriers to innovation were described by four-state variables. Considering nine ownership states and types of enterprises, it is needed to determine the interdependencies between seventy-five variable states.

2.5. Statistical Methods

Two statistical methods were used in the study: Pearson's χ^2 independence test and correspondence analysis. These methods were selected in such a way that the results obtained complement each other. The starting point for the calculations according to both methods is the summary of the data in the contingency tables. The first method tests the existence of significant relationships between variables, while the second provides information about the structure of the relationships between rows and columns of a contingency table.

The independence test is used to determine the relationship between two categorical variables [39]. The test relies on the comparison of the values resulting from (empirical) research with the expected values, which assume no relationship between variables. The options considered must be mutually exclusive and have a total probability of 1. The χ^2 statistic is used to evaluate the test value. The choice between the null hypothesis on the independence of variables and its opposite (i.e., the alternative hypothesis) is made on the basis of a comparison of the p-value with the significance level.

Correspondence analysis is a multidimensional statistical method for studying co-occurrence of phenomena [40,41]. It has an exploratory character, which differs from traditional methods of testing statistical hypotheses. Classic methods rely on a priori verification of hypotheses regarding relationships between variables, while correspondence analysis enables the discovery of systematic relations between variables without formulating expectations a priori about the nature of these relationships. Therefore, correspondence analysis is not a confirmation technique, but a method of discovering relationships and structures in empirical data. It is especially useful in economics because it allows the study of multidimensional phenomena such as irrationality [42]. The essence of correspondence analysis is to reduce the dimension of the studied problem, which consists of recreating the distance between points representing rows and columns of the contingency table in a space with fewer dimensions. Calculations are performed in such a way that the loss of information about the diversity of rows and columns of the contingency table is as small as possible. Contingency tables contain appropriate measures to describe the relationships between rows and columns. The final results of the correspondence analysis are two or three-dimensional charts called biplots, which graphically present the relations of co-occurrence between the studied variables. In this study, row and column profile standardization was chosen to simultaneously analyze points representing row profiles and column profiles. The evaluation of points representing the individual variables, the χ^2 metric is used, which is the weighted Euclidean distance. In the correspondence analysis, there is a total inertia that signifies the ratio of the χ^2 statistic to the grand total of quantity. This is a measure of the dispersion of row profiles and column profiles around average profiles. Inertia that is close to zero signifies a small dispersion of profiles around the average profile. For example, this situation occurs when all students have received the same exam grade.

On some correspondence maps (Figure 3, Figure 4 and Figure 5) [37], the analysis is conducted both from the perspective of enterprise types and ownership sectors, which indicates the need to introduce a method to increase the clarity of these two variables and their states. The easiest way was to combine the points representing the different types and sectors of business ownership so that they formed triangles. Ownership sectors are indicated by hatched triangles, while full-color triangles represent types of enterprises. The vertices of the triangles have two-part names, with the first segment representing the ownership sector and the second segment representing the type of enterprise. In this way, the mixed sector (S3) forms a hatched triangle with vertices S3FR_1, S3FR_2, and S3FR_3, since it includes all three types of enterprises (i.e., small FR_1, medium FR_2, and large FR_3). Since the same principle applies to the other two ownership sectors, in total, there are nine names for the vertices of these triangles. Following this convention, types of enterprises are identified by the second part of their name. Thus, medium-sized enterprises (FR_2) are represented by a triangle with vertices S1FR_2, S2FR_2, and S3FR_2, as this type includes enterprises belonging to public (S1), private (S2), and mixed (S3) sectors. The same principle applies to the other two types of enterprise. The small enterprise type FR_1 is represented by a yellow triangle, the medium-sized enterprise type FR_2 by a pink triangle, and the large enterprise type FR_3 is represented by a light blue triangle.

3. Results

Studies on the innovativeness of Polish industrial processing enterprises in three periods, prosperity (2004–2006), global financial crisis (2008–2010), and recovery (2012–2014) led to many interesting discoveries [35–37]. The most important of them include:

(1) demonstration of the significant impact of interactions between the ownership sectors and types of enterprises on research results;
(2) detection of a low level of innovativeness of public enterprises compared to other enterprises from different ownership sectors (i.e., private and mixed);
(3) exemplification of Schumpeter's creative destruction theory by showing that innovative enterprises are developing in accordance with the Red Queen dynamics;
(4) demonstration of the insensitivity of the effects and goals of innovative activity of companies to cyclical factors (business cycle phases);
(5) confirmation of the occurrence of the Polish Green Island effect as a fact and not government propaganda; and
(6) validation of the hypothesis that the global financial crisis is associated with the turn of the Fifth and Sixth Kondratieff waves.

3.1. Significant Impact of Interactions between Ownership Sectors and Types of Enterprises on Research Results

The importance of the interrelationships between ownership sectors and types of enterprises for the innovation activity of enterprises, and thus what cybernetic diagrams (3)–(8) show can be demonstrated empirically. During the 2012–2014 survey, it became apparent that the χ^2 independence test demonstrates a statistically significant relationship only between the types of enterprises and the twenty-two objectives of innovation, which include both the four basic types of innovation (product, process, organizational, and marketing) as well as eco-innovations (Tables 6–8 and Figures 1 and 2) [35,37]. When three ownership sectors are included (public, private, and mixed) in the analysis, this relationship disappears (Table 9 and Figure 3) [37]. In other words, during the recovery period, the goals of innovation activities are independent of the interaction between ownership sectors and types of enterprises. Compared to other periods, the correspondence maps (Figures 1–3) showed a significant decrease in average distances between points representing ownership sectors and types of enterprises, and points corresponding to the objectives of innovative activity (Table 10 and Figure 4, Table 11 and Figure 5) [35,37]. The total inertia of the entire industrial processing section has become close to zero (Table 12) [37]. This leads to the conclusion that, in principle, 98.9%

of companies achieve their goals of innovative activities, which is all but public sector enterprises (Table 5) [37]. In addition, the result is identical when the goals are limited to four basic types of innovations, and eco-innovations will be considered as supplementary points (Table 13) [37]. The χ^2 independence test confirms the cumulative impact of ownership sectors and enterprise types on eco-innovation alone (Table 14) [37].

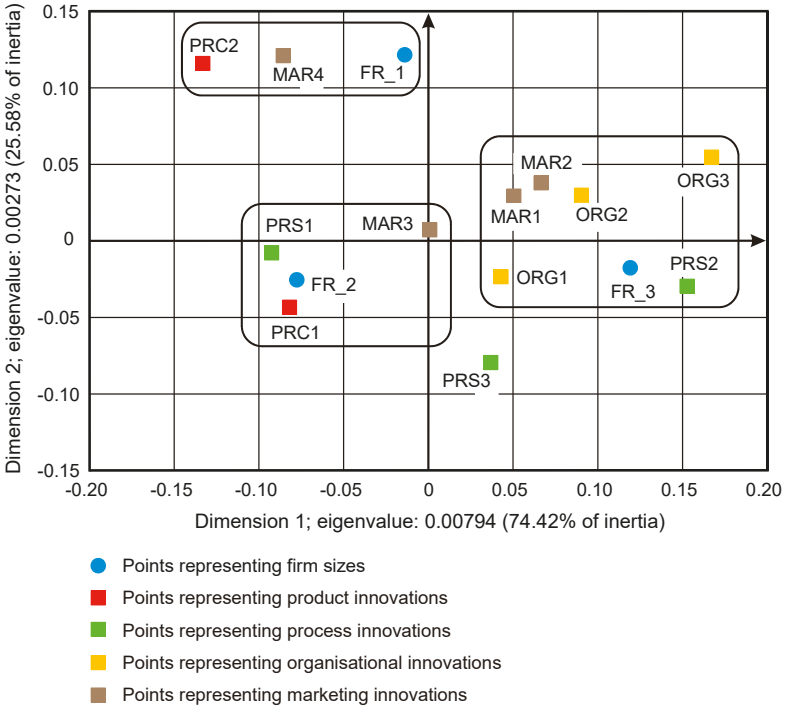


Figure 1. Correspondence map showing the co-occurrence of the types of innovation and the sizes of enterprises in the period 2012–2014 (dimensions 1–2; 100% of total inertia).

To draw binding conclusions regarding the calculations made for the years 2012–2014, it is required to compare them with the results obtained for the previous two periods. In the years of prosperity in 2004–2006 and during the global financial crisis in 2008–2010, there was a statistically significant relationship between the combined impact of ownership sectors and types of enterprises, and the effects or objectives of innovative activities (Tables 15 and 16) [36]. In addition, during the crisis, compared to the previous period of prosperity, the innovative activities of most enterprises increased (Figures 4 and 5) [37], which seems to be a peculiarity, but this can be explained by referring to the impact of secular factors. However, in the third period 2012–2014, this trend continued (Figure 3) [37]. Each company tried to be innovative. Only eco-innovations alone proved to be dependent on the combined impact of ownership and types of enterprises (Table 14) [37], but it should be noted that this was a relatively new type of activity for companies at the time. Generally, it should be noted that the last period was special and ground breaking. Almost all enterprises approached a certain development threshold, some more and some a little less [35–37]. The significance of these changes will be explained later.

Table 6. Variables describing the types of innovation, the goals of innovative activity in the years 2012–2014, and coding method.

Types of Innovation		Goals of Innovative Activity	Codes
Product innovations		New or significantly improved manufactured goods	PRC1
		New or significantly improved services	PRC2
Process innovations		New or significantly improved methods of producing goods and services	PRS1
		New logistic processes	PRS2
		New management processes	PRS3
Organizational innovations		New methods under the principles of operation adopted	ORG1
		New methods of distribution of tasks and decision-making powers among employees	ORG2
		New organizational methods in terms of relations with the environment	ORG3
Marketing innovations		Significant changes in the design/construction and/or packaging of goods and/or services	MAR1
		New media or product promotion methods	MAR2
		New methods in terms of product distribution or sales channels	MAR3
		New methods of pricing goods and services	MAR4
Eco-innovations	**Environmental benefits obtained during the production of products or services**	Reduction of material consumption or water consumption per unit of product	ECO1
		Reduction of energy intensity or carbon dioxide emissions	ECO2
		Reduction of soil, water, air or noise pollutions	ECO3
		Use of materials that are less polluting or less dangerous to the environment	ECO4
		Reduction of the fossil fuels, higher use of energy obtained from renewable sources	ECO5
		Re-use (recycling) of waste, water or materials for personal use or sale	ECO6
	Environmental benefits obtained during the period of use of the purchased product or use of the service by end users	Reducing energy consumption or carbon dioxide emissions	ECO7
		Reduction of air, water, soil or noise pollutions	ECO8
		Facilitating the re-use (recycling) of the product after use	ECO9
		Extending the life of products thanks to increased durability and strength	ECO10

Table 7. Results of the verification of the null hypothesis regarding the independence of innovation types from enterprise size (2012–2014).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	Types of innovation implemented do not depend on the enterprise size
Alternative hypothesis (H_1)	**Types of innovation implemented depend on the enterprise size**
χ^2 **statistics value**	117.36
Critical region	right-tailed
Level of Significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.0000$
Decision	Since $p < \alpha$, H_0 needs to be rejected in favour of H_1

Table 8. Results of the verification of the null hypothesis regarding the independence of eco-innovation form choice from enterprise types (2012–2014).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	Forms of eco-innovation do not depend on the type of enterprise
Alternative hypothesis (H_1)	**Forms of eco-innovation depend on the type of enterprise**
χ^2 **statistics value**	55.228
Critical region	right-tailed
Level of Significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.0001$
Decision	Since $p < \alpha$, H_0 needs to be rejected in favour of H_1

Table 9. List of assumptions and calculations necessary to verify the hypothesis regarding the relationship between the type and ownership sector of an enterprise and the goals of its innovative activity (2012–2014).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	**The type and ownership sector of enterprises have no effect on the goals of innovative activity**
Alternative hypothesis (H_1)	The type and ownership sector of enterprises have an effect on the goals of innovative activity
χ^2 **statistics value**	120.85
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.99759$
Decision	Since $p > \alpha$, there are no grounds for rejecting H_0

Table 10. Variables describing the effects of innovative activity of enterprises in 2004–2006, the degrees of influence of innovations introduced by enterprises in 2004–2006 on the activity of enterprises at the end of 2006, and the method of coding.

Effect Type	**Effects of Innovative Activity Scale: 1—High; 2—Medium; 3—Low; 4—Irrelevant**	**Codes**	**Degree of Influence**
Product effects	Increase of the product assortment	E1	1, 2, 3, 4
	Entering into new markets or increasing the existing market share	E2	1, 2, 3, 4
	Product quality increase	E3	1, 2, 3, 4
Process effects	Improvement in production flexibility	E4	1, 2, 3, 4
	Increase of production capacity	E5	1, 2, 3, 4
	Reduction of labor costs per unit of product	E6	1, 2, 3, 4
	Reduction of consumption of materials and energy per unit of product	E7	1, 2, 3, 4
Other effects	Reduction of harmfulness to the environment and improvement of work safety	E8	1, 2, 3, 4
	Compliance with regulations, norms or standards	E9	1, 2, 3, 4

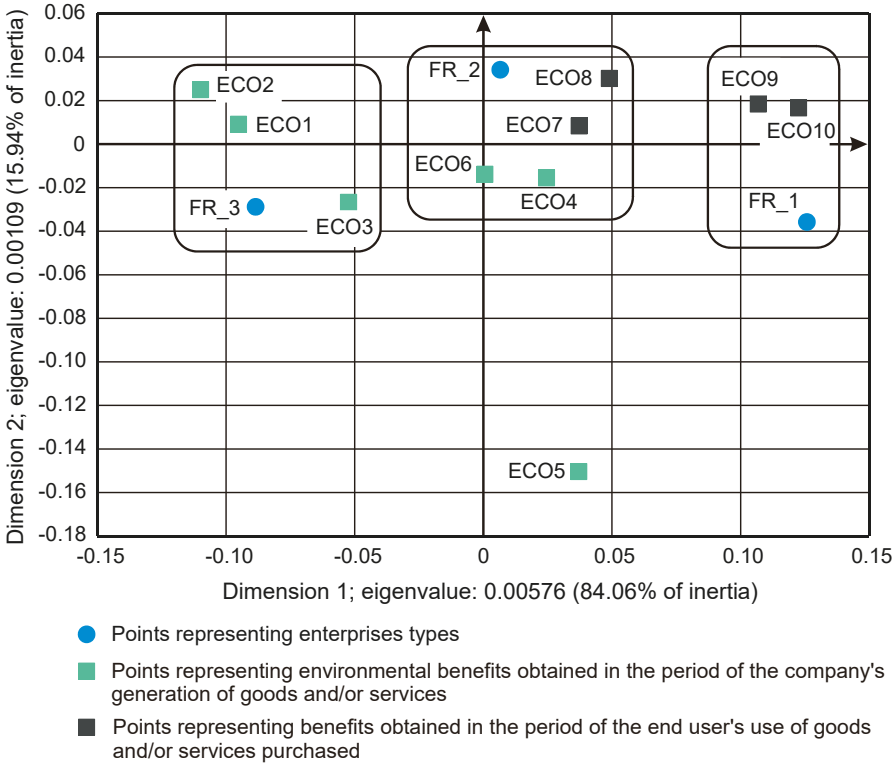


Figure 2. Correspondence map showing the co-occurrence of eco-innovation forms and types of enterprises in the period 2012–2014 (dimensions 1–2; 100% of total inertia).

Table 11. Variables describing the goals of innovative activity in the years 2008–2010, their degrees of importance for innovative activity of enterprises with regard to product or process innovation in 2008–2010, and the method of coding.

Goals of Innovative Activity Scale: 1—High; 2—Medium; 3—Low; 4—Irrelevant	Codes	Degree of Importance
Increase of the product or service assortment	G1	1, 2, 3, 4
Replacement of obsolete products or processes	G2	1, 2, 3, 4
Entering into new markets or increasing the existing market share	G3	1, 2, 3, 4
Improvement of the quality of products or services	G4	1, 2, 3, 4
Improvement in production flexibility	G5	1, 2, 3, 4
Increase of production capacity	G6	1, 2, 3, 4
Reduction of labor costs per unit of product	G7	1, 2, 3, 4
Reduction of consumption of materials and energy per unit of product	G8	1, 2, 3, 4
Reduction of harmfulness to the environment	G9	1, 2, 3, 4
Improvement of work safety	G10	1, 2, 3, 4

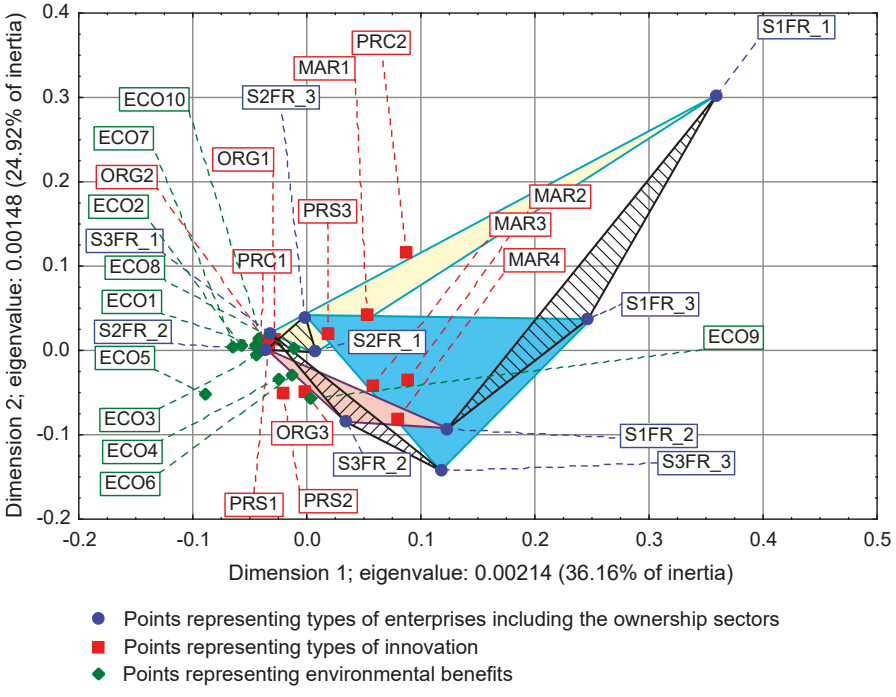


Figure 3. Correspondence map showing the co-occurrence of the types of enterprises including the ownership sectors, the types of innovation, and the environmental benefits in the period 2012–2014 (dimensions 1–2; 61.08% of total inertia).

Table 12. Total inertia of the Polish industrial processing section.

Total Inertia		
2004–2006	**2008–2010**	**2012–2014**
0.0126	0.01801	0.00593

Table 13. List of assumptions and calculations necessary to verify the hypothesis regarding the relationship between the type and ownership sector of an enterprise and the goals of its innovative activity with the eco-innovations as supplementary points (2012–2014).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	**The type and ownership sector of enterprises have no effect on the goals of innovative activity, taking into account eco-innovations as supplementary points**
Alternative hypothesis (H_1)	The type and ownership sector of enterprises have an effect on the goals of innovative activity, taking into account eco-innovations as supplementary points
χ^2 **statistics value**	65.248
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.96687$
Decision	Since $p > \alpha$, there are no grounds for rejecting H_0

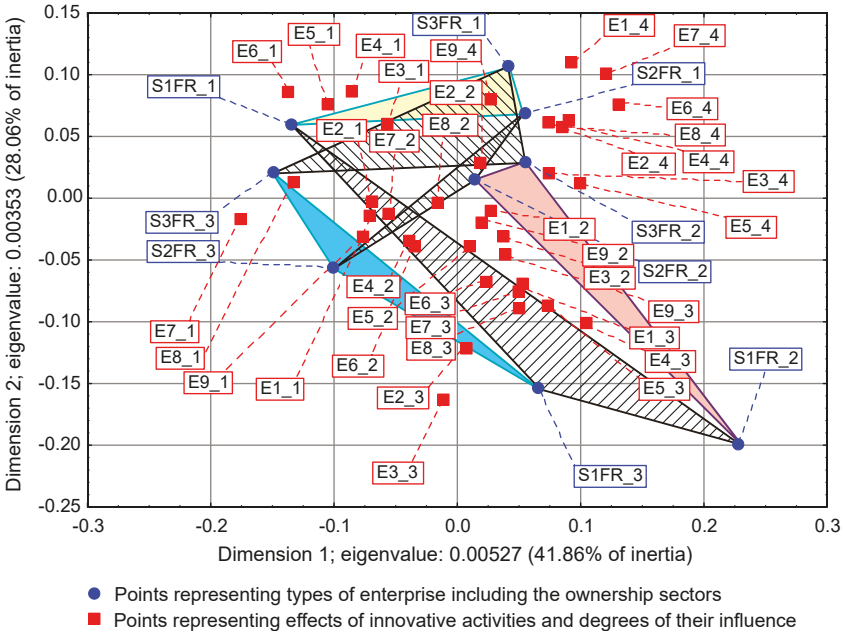


Figure 4. Correspondence map showing the co-occurrence of the types of enterprises including the ownership sectors, the effects of innovative activity, and degrees of their influence on enterprises in period 2004–2006 (dimensions 1–2; 69.92% of total inertia).

Table 14. List of assumptions and calculations necessary to verify the hypothesis regarding the relationship between the type and ownership sector of an enterprise and the eco-innovations (2012–2014).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	The type and ownership sector of enterprises have no effect on the activity of a firm concerning eco-innovation
Alternative hypothesis (H_1)	**The type and ownership sector of enterprises have an effect on the activity of a firm concerning eco-innovation**
χ^2 **statistics value**	311.44
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.0000$
Decision	H_0 hypothesis should be rejected in favour of H_1

3.2. Low Level of Innovativeness of Public Enterprises

In the three periods examined, public enterprises (S1FR_1, S1FR_2, and S1FR_3) showed significantly less innovative activities than enterprises from other ownership sectors. In the first period (2004–2006), the effects of their innovative activities were weak (Figure 4) [37], and in the next two periods (2008–2010 and 2012–2014), the objectives of innovative activities were not substantially achieved (Figures 3 and 5) [37]. If any effects or goals were achieved, then degrees of influence or importance were low or irrelevant. This means that the innovations introduced by enterprises in the first period had virtually no impact on the activities of companies at the end of 2006 and that the goals achieved in the second period had little effect on innovation activities in the field of product

and process innovations. Therefore, if we consider the relationships (3), (5), and (7), then it can be stated that in the case of public enterprises, there were no positive feedbacks. Although there were some differences in individual periods, they did not affect the trend described above. After adopting the prosperity period as a benchmark, it can only be said that in the years of the global financial crisis, the situation of small (S1FR_1) and medium-sized (S1FR_2) public enterprises deteriorated and the situation of large public enterprises (S1FR_3) improved (Figures 4 and 5) [37]. However, during the recovery period, there were no significant changes except that the innovation activity of medium-sized enterprises (S1FR_2) improved, and the innovation activity of the large enterprises (S1FR_3) deteriorated (Figure 3) [37].

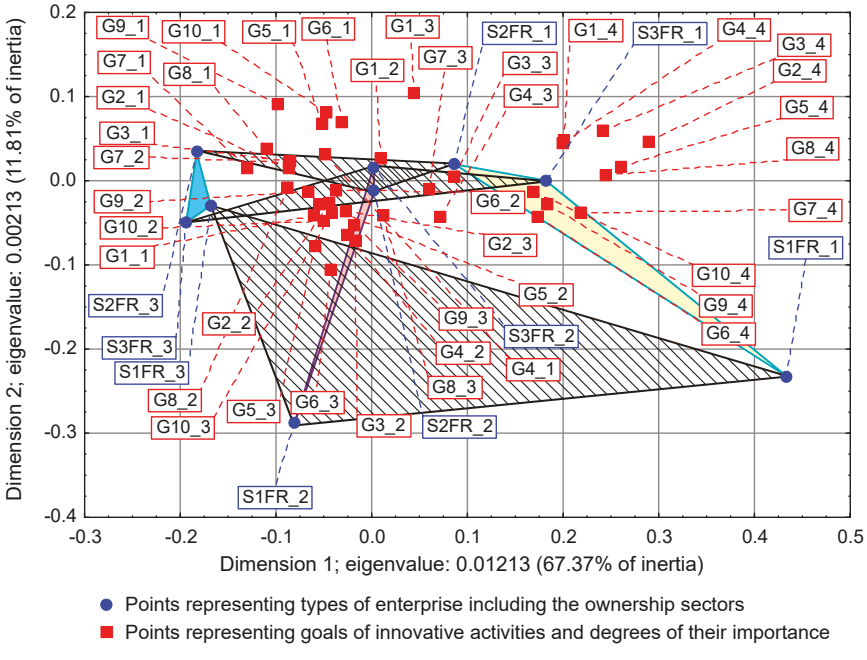


Figure 5. Correspondence map showing the co-occurrence of the types of enterprises including the ownership sectors, the goals of innovative activity, and their degrees of importance for enterprises in period 2008–2010 (dimensions 1–2; 79.18% of total inertia).

Table 15. List of assumptions and calculations necessary to verify the hypothesis about the relationship between the type and sector of enterprise ownership and the effects of its innovative activities (2004–2006).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	The type and ownership sector of the enterprise have no impact on the effects of innovative activity
Alternative hypothesis (H_1)	**The type and ownership sector of the enterprise have an impact on the effects of innovative activity**
χ^2 **statistics value**	426.05
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.0000$
Decision	H_0 hypothesis should be rejected in favour of H_1

Table 16. List of assumptions and calculations necessary to verify the hypothesis about the relationship between the type and sector of enterprise ownership and the goals of its innovative activities (2008–2010).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	The type and ownership sector of the enterprise have no impact on the goals of innovative activity
Alternative hypothesis (H_1)	**The type and ownership sector of the enterprise have an impact on the goals of innovative activity**
χ^2 **statistics value**	668.581
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.0000$
Decision	H_0 hypothesis should be rejected in favour of H_1

In general, the low innovativeness of public enterprises does not seem to be a big economic problem, since the share of the public sector in the entire manufacturing sector was small in the first period, and in the following ones, it showed a decreasing trend. In periods of prosperity, crisis, and recovery, this share was 4.37%, 1.18%, and 1.10%, respectively (Tables 3–5) [37]. However, when considering the interactions between ownership sectors and types of enterprises, it cannot be excluded that the public sector may have an adverse effect on enterprises belonging to other ownership sectors. On the other hand, one should also take into account the positive impact of companies from the private and mixed sectors on the public sector. It seems that in the manufacturing sector, there may be some kind of a dynamic balance between ownership and the type (size) of enterprises, which changes in particular phases of the business cycle. Perhaps in this way, the adverse effects of cyclical fluctuations on the innovation activities of enterprises are somewhat neutralized. There may also be impacts from factors operating for even longer periods. Nevertheless, this issue requires further in-depth research, and some related proposals are presented in the next two parts of the article.

3.3. The Red Queen Effect

The Red Queen effect is a metaphor derived from Alice's adventures in Wonderland where this Queen reigned [43]. In the Queen's land, one would have to run as fast as one could in order to keep in the same place. However, if one would wish to get somewhere else, one would have to run at least twice as fast as that. Indeed, after a long and exhausting run, Alice noticed with amazement that she was in the same place as before. Initially, this hypothesis was used in biology to explain the law of extinction, which states that organisms in any adaptive zone die with a stochastically constant rate [44]. The justification for this principle is that adaptation to certain living conditions of one species may change the selection pressure to other species and lead to positive feedback between species. The properties of communities will then change in a directional manner. This is similar to an arms race, which takes place both between species and within them. Organisms must be in constant motion and adapt to change in order to survive. Thus, existence and survival (i.e., being in the same place) requires constant running. It soon became apparent that the Red Queen hypothesis could also explain many economic phenomena [37].

Schumpeter's creative destruction theory includes the Red Queen effect. According to Schumpeter, economic development is associated with the growth and collapse of companies and entire branches of industry [9,11]. Enterprises cannot last forever, and the reason for their collapse is almost always the lack of adequate capacity to implement innovation. Each innovation success is rewarded with a bonus in the form of profit, which is inherently temporary and tends to decrease during competition and adaptation processes. Therefore, no enterprise is safe against bankruptcy.

The Red Queen effect applies to public sector enterprises (S1FR_1, S1FR_2, and S1FR_3) as their share in the manufacturing section is constantly decreasing (Tables 3–5) [37]. Therefore, they cannot cope with competition from private and mixed sector companies. If proper preventive action is not

taken, public enterprises may permanently disappear from the manufacturing section. This can entail not only economic, but also political consequences.

3.4. Insensitivity of the Effects and Goals of Innovative Activities of Enterprises to the Business Cycle Phases

The conducted research shows that in each of the examined periods, the vast majority of enterprises achieved the assumed effects and goals of innovative activities, which contributed to increasing their subsequent activities in this field to a high or medium degree. In the years 2004–2006, the innovations introduced by enterprises had a great impact on the activities of enterprises at the end of 2006, and in the years 2008–2010, most of the achieved goals contributed significantly to increasing the activity of enterprises in the field of product and process innovations. Therefore, there were positive feedback loops that show relationships (3) and (5). These phenomena are visible on correspondence maps (Figures 4 and 5) in the form of short distances between points representing ownership sectors and types of enterprises, and points responsible for the effects or objectives of innovative activities and the degrees of their influence or importance [37]. A similar phenomenon could not be found in 2012–2014, but only because the statistical form PNT-02 did not contain such information. There is, however, other indirect evidence that, in the third period, there was indeed positive feedback, illustrated by the relationship (7), between innovations undertaken at different times. This is evidenced by the near zero inertia of the entire industrial processing section (Table 12), which signifies that almost all enterprises achieved the assumed goals of innovative activities (Figure 3) [37].

In general, the innovation activity of enterprises varies depending on the phases of the business cycle, but this is not always as expected. In the first period, points representing ownership sectors and types of enterprises as well as points responsible for the effects of innovative activities and high or medium degrees of their influence formed a joint cluster on the correspondence map (Figure 4) [37]. This signifies that most enterprises achieve the intended effects and that positive feedback occurs in accordance with relationship (3). When the period of prosperity is taken as the basis for comparative analysis, it should be noted that during the global financial crisis, there were changes that consisted in the formation of two separate clusters of points on the biplot (Figure 5) corresponding to this period [37]. In the first cluster, which covered the vast majority of enterprises, the average distances between points representing types and ownership sectors of enterprises and points corresponding to the objectives of innovative activity and high or medium degrees of their importance for innovative activity in the field of product and process innovation during this period decreased noticeably. Thus, during the crisis, there was a positive feedback presented by the relationship (5). The second cluster was very small and was relatively far away from the first. The enterprises in this cluster showed very little innovative activity, and if they were already achieving some goals, their degrees of importance were low or irrelevant. The cluster contained a yellow triangle, so it mainly represented small businesses (S1FR_1, S2FR_1, and S3FR_1). It seems that the crisis has sifted companies, and thus separated companies with high innovation activity from those that were not very innovative. This is paradoxical, as the position of more innovative enterprises improved during the crisis, and the position of less innovative companies deteriorated. This observation can be justified by referring again to the Red Queen effect (i.e., an explanation indicating the selection pressure created during the crisis that the best companies exerted on all others). In the third period, ownership sectors and types of enterprises had no impact on the objectives of their innovation activities, which included both four basic types of innovation (i.e., process, product, organizational, and marketing innovations) as well as eco-innovations (Table 9 and Figure 3) [37]. There was a further decrease in average distances between points representing ownership sectors and types of enterprises and points related to the objectives of innovative activity. This demonstrates that almost all companies achieved their objectives and proves the existence of positive feedback assumed by dependence (7). A large increase in the innovative activity of enterprises in this period indicates that they considered introducing innovations as a necessary condition for their development [37].

The observations indicate the existence of a certain trend, which consists in a steady increase in the innovative activity of most enterprises and appears to be largely independent of the phases of the business cycle. The global financial crisis proved to be only a selection mechanism that separated the group of the best companies from the weakest and at the same time improved the situation of the former and worsened the situation of the latter. The first group includes the vast majority of enterprises from the industrial processing section. This leads to the conclusion that this trend may have more to do with secular factors than with cyclical factors.

Interestingly, the bad situation of public sector enterprises also does not seem to be connected with cyclical factors. In addition, political factors do not seem to be responsible for this state of affairs as political changes can be faster than cyclical changes. The indication of the Red Queen effect as a reason therefore has an additional justification in the form of long-term selection pressure exerted on public sector companies by more innovative enterprises from other ownership sectors. This would prove that selection pressure is not cyclical, but secular, which seems reasonable.

3.5. The Effect of the Polish Green Island

The next stage of research was to consider the impact of interactions between ownership sectors and types of companies on innovation barriers and the ability of these barriers to inhibit the innovative activity of enterprises (degrees of their influence in the first period or degrees of importance in the second and third period). The aim was to determine the impact of innovation barriers on the innovation activities of enterprises. In each of the three examined periods, the combined effect of two feedback loops was considered, one of which examined the interdependencies between (a) the interaction of ownership sectors and types of enterprises and (b) the effects or objectives of innovative activities and the degrees of their influence or importance, while the other concerned the interdependencies between (c) cooperation of ownership sectors and types of enterprises and (d) barriers to innovation and the degree of their influence or importance. In other words, a total of two cybernetic schemes were examined in each period: (3) and (4) in the years of prosperity 2004–2006, (5) and (6) during the period of the global financial crisis 2008–2010 as well as (7) and (8) in the years of recovery 2012–2014 [36].

The results will be presented chronologically from the first period, which will be the benchmark of comparative analysis. In the years of prosperity, the results of the χ^2 independence test confirmed the existence of statistically significant relationships, which concerned both the impact of mutual interactions of ownership sectors and types of enterprises on the effects of innovative activities (Table 15) as well as the impact of mutual interactions of ownership sectors and types of enterprises on barriers to innovation (Table 17) [36]. As a result of using correspondence analysis, a correspondence map (Figure 6) was obtained that captured the co-occurrence of phenomena in more detail [36]. It contained eighty-nine points representing the states of individual variables: nine of them represented ownership sectors and types of enterprises, thirty-six were responsible for the effects of innovative activities of enterprises and degrees of their influence (Table 10), while forty-four were for barriers to innovation and degrees of their influence (Table 18) [36,37]. Points representing barriers to innovation and the degrees of their influence were located at a relatively large distance from a cluster of points consisting of both points representing ownership sectors and types of enterprises as well as points describing the effects of innovation activities and the degree of their influence. Innovation barriers were not overly troublesome for enterprises.

During the global financial crisis, the results of applying the χ^2 independence test were identical because they showed the dependence of goals of innovation activity and innovation barriers on the combined impact of ownership sectors and types of enterprises (Tables 16 and 19) [36]. The correspondence map (Figure 7) showed the co-occurrence between ninety-three points representing the states of individual variables, nine of which represented ownership and size of companies, forty for the goals of innovation activities and their degrees of importance (Table 11), and forty-four barriers to innovation and their degrees of importance (Table 18) [36,37]. In comparison to the previous period, it can be observed that the distance between the two clusters of points representing (1) ownership sectors

and types of enterprises as well as the objectives of innovation activities and (2) barriers to innovation and their significance levels increased. This demonstrates that during the crisis, the importance of all innovation barriers decreased quite significantly.

Table 17. List of assumptions and calculations necessary to verify the hypothesis about the relationship between the type and sector of enterprise ownership and innovation barriers (2004–2006).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	The type and ownership sector of the enterprise have no impact on innovation barriers
Alternative hypothesis (H_1)	**The type and ownership sector of the enterprise have an impact on innovation barriers**
χ^2 **statistics value**	1519.68
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.0000$
Decision	H_0 hypothesis should be rejected in favour of H_1

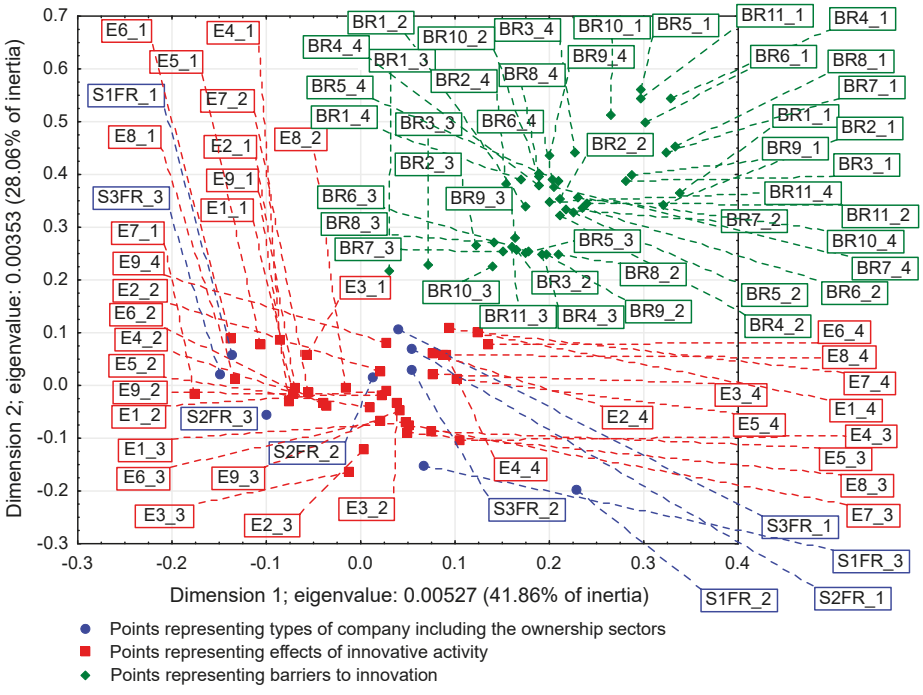


Figure 6. Correspondence map showing the co-occurrence of the types of enterprises including the ownership sectors, the effects of innovative activity, and the barriers to innovation in the period 2004–2006 (dimensions 1–2; 69.92% of total inertia).

In the years of recovery from 2012 to 2014, some stronger alterations could be observed as the χ^2 independence test indicated the dependence of the goals of innovative activity on the combined impact of ownership and size of enterprises (Table 20), but did not confirm the existence of the relationship between innovation barriers and the combined impact of ownership and size (Table 21) [36]. The goals of innovative activity (fifteen in total) included product and process innovations as well as eco-innovations (PRC1, PRC2, PRS1–PRS3, ECO1–ECO10). A similarity to the previously obtained result (Table 14) was

observed here, which confirmed the importance of eco-innovations for the examined enterprises [37]. Furthermore, it has been found that eco-innovations in Poland are more related to product and process innovations than to organizational and marketing innovations. During this period, the significance of the goals of innovation activities was not considered. Sixty-eight points were included on the correspondence map (Figure 8) depicting the co-occurrence of phenomena, nine of which represented ownership and size, fifteen the objectives of innovative activity (Table 6), and forty-four the barriers to innovation and their degrees of importance (Table 22) [36,37]. Compared to the crisis period, the cluster of points containing ownership and size as well as the objectives of innovative activity is further away from the cluster of points representing barriers to innovation and their degrees of importance. It can be concluded that during the recovery period, nothing prevented the innovative activity of enterprises.

Table 18. Variables describing innovation barriers in the years 2004–2006 and 2008–2010, degrees of their influence on enterprises and coding.

Type of Barrier	Factors Impeding Innovative Activity Scale: 1—High; 2—Medium; 3—Low; 4—Irrelevant	Codes	Degree of Influence
Economic factors	Lack of financial resources in your company or in your group of enterprises	BR1	1, 2, 3, 4
	Lack of financial resources from external sources	BR2	1, 2, 3, 4
	Too high costs of innovation	BR3	1, 2, 3, 4
Knowledge factors	Lack of qualified staff	BR4	1, 2, 3, 4
	No information about technology	BR5	1, 2, 3, 4
	No information on markets	BR6	1, 2, 3, 4
	Difficulties in finding partners for cooperation in the field of innovative activity	BR7	1, 2, 3, 4
Market factors	Market split by dominant enterprises	BR8	1, 2, 3, 4
	Uncertain demand for innovative (new) products	BR9	1, 2, 3, 4
Other factors	No need to run innovative activity due to the introduction of innovations in previous years	BR10	1, 2, 3, 4
	No demand for innovation	BR11	1, 2, 3, 4

Table 19. List of assumptions and calculations necessary to verify the hypothesis about the relationship between the type and sector of enterprise ownership and innovation barriers (2008–2010).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	The type and ownership sector of the enterprise have no impact on innovation barriers
Alternative hypothesis (H_1)	**The type and ownership sector of the enterprise have an impact on innovation barriers**
χ^2 **statistics value**	3174.84
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
***p*-value** (p)	$p = 0.0000$
Decision	H_0 hypothesis should be rejected in favour of H_1

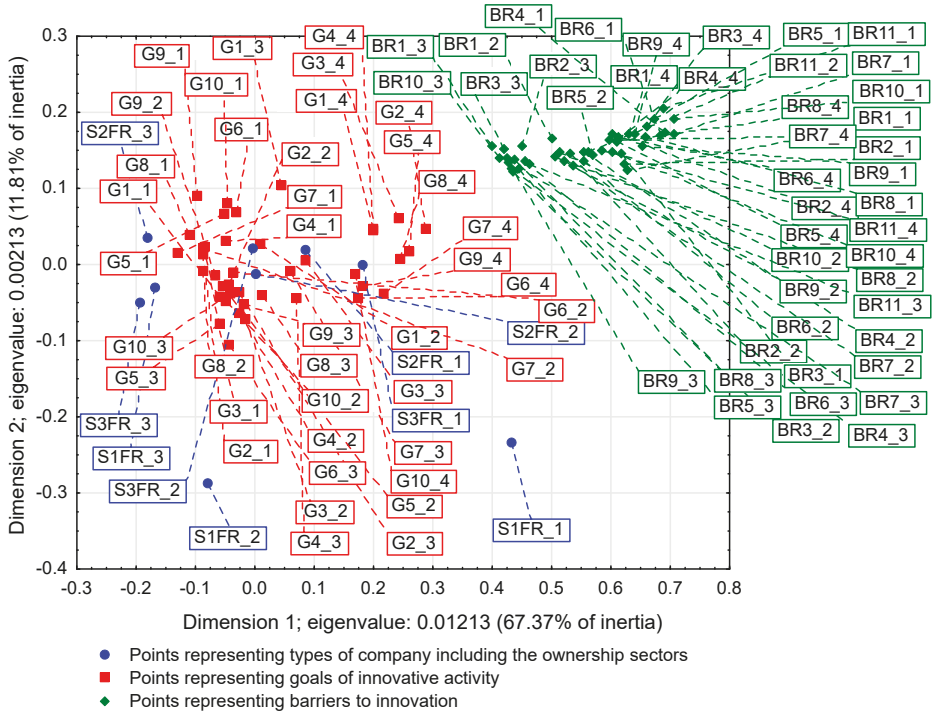


Figure 7. Correspondence map showing the co-occurrence of the types of enterprises including the ownership sectors, the goals of innovative activity, and the barriers to innovation in the period 2008–2010 (dimensions 1–2; 79.18% of total inertia).

Summing up the research, it can be stated that in the three studied periods, there was a constant tendency to shorten the distance between the points in the two clusters: (1) representing ownership sectors and types of enterprises as well as effects or goals of innovative activity along with degrees of influence or degrees of importance and (2) corresponding only to innovation barriers and degrees their importance. These changes were accompanied by a gradual increase in the distance between these two clusters. It can be observed that these two clusters gradually thickened and at the same time moved away from each other. The trend emerging from these processes showed two simultaneously occurring phenomena: (1) an increase in the innovative activity of enterprises, which is reflected in the implementation of their goals of innovative activities and the appearance of related positive feedback and (2) a gradual decrease in the importance of innovation barriers and the emergence of positive feedbacks related to this process, which led to the almost complete lack of significance of these barriers for enterprises. This long-term trend is independent of the business cycle phases. This signifies that we are dealing with phenomena shaped by pro-developmental secular factors. Most likely, their impact is not limited to the industrial processing section, but applies to the entire economy, which must therefore have a solid, strong foundation for economic growth and development. In order to justify this claim, it is necessary to study the latest economic history of the country.

In 2010, the Polish government presented to the public a map of Europe on which individual countries were attributed the actual economic growth rates they achieved in 2009. Poland had a positive growth rate and was marked in green, while in all surrounding countries, the growth rates were negative, which is why those countries were marked in red. In this way, Poland was presented as the Green Island of economic growth against the background of Europe in crisis [45,46]. In the critical

year 2009 for the Polish economy, the growth rate decreased to 1.7%, while the EU average at that time was negative and amounted to –4.2% [47] (p. 91). According to the latest data presented at the Economic Forum in Krynica-Zdrój in 2019, Poland has been recording uninterrupted economic growth since 1992 with an average annual growth rate of over 4%. Over the past 27 years, only Australia has achieved a similar result among the OECD (Organisation for Economic Co-operation and Development) countries. In the years 1990–2018, GNP tripled, and the Polish economy is currently the seventh largest economy in the European Union and the twenty-third in the world [48]. In addition, forecasts show that by 2025, the Polish economy may become one of the strongest engines of growth in Europe and a significant force in the global market [49]. For these reasons, we have called the long-term trend discussed above as the Polish Green Island effect. The map from 2010 was therefore not associated with the government's propaganda, as was often presented, but showed an actual economic success.

Table 20. List of assumptions and calculations necessary to verify the hypothesis regarding the relationship between the type and ownership sector of an enterprise and the goals of its innovative activity in the years 2012–2014 (PRC1, PRC2, PRS1–PRS3, ECO1–ECO10).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	The type and ownership sector of the enterprise have no impact on the goals of innovative activity
Alternative hypothesis (H_1)	**The type and ownership sector of the enterprise have an impact on the goals of innovative activity**
χ^2 **statistics value**	2361.7
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
p (p)	$p = 0.0000$
Decision	H_0 hypothesis should be rejected in favour of H_1

3.6. The Global Financial Crisis as the Turn of the Fifth and Sixth Kondratieff Waves

The long-term trends described above apparently lead to some culmination in the period 2012–2014. We observed a decreasing inertia of the industrial processing section, which during the recovery period 2012–2014 became close to zero (Table 12) [37]. This is tantamount to achieving the assumed goals of innovative activity by most enterprises. Moreover, there is a constant decrease in the importance of innovation barriers, until their almost complete disappearance in the third period (Figures 6–8) [36]. The second of these phenomena affects the first, which is why they undoubtedly form a certain systematic integrity. It should also be noted that as part of the observed trend, the total impact of ownership sectors and types of enterprises on the innovation activity of companies is significantly reduced (Table 9, Table 15, and Table 16) [36,37]. Therefore, it seems that ownership and size are the factors determining the innovativeness of companies only in relatively short periods appropriate for traditional business cycles. The emergence of this trend can only be explained by secular factors, which indicates the need to interpret the results presented here as part of Kondratieff long-wave theory. Therefore, at least changes in technology and adaptation processes of the Polish economy to the conditions of the world economy should be taken into account. When it comes to technological innovations, it should be emphasized that they are almost always the result of technological revolutions. In this context, it is justified to refer to the modern interpretation of K-waves.

The theory of the Kondratieff cycle was developed and adapted to modern conditions by Šmihula. In the modern age, counted from 1600, he distinguished the following six K-waves [8]:

(1) financial-agricultural revolution (1600–1780; 180);
(2) industrial revolution (1780–1880; 100);
(3) technological revolution (1880–1940; 60);
(4) scientific-technological revolution (1940–1985; 45);
(5) information and telecommunications revolution (1985–2015; 30); and

(6) post-information technological revolution, in other words, the biomedical-hydrogen revolution (2015–2035; 20).

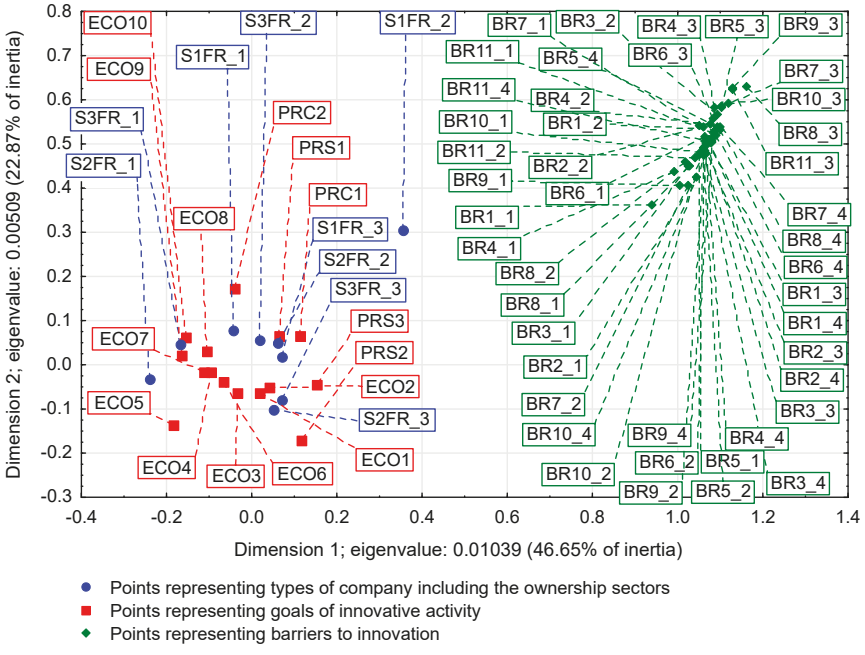


Figure 8. Correspondence map showing the co-occurrence of the types of enterprises including the ownership sectors, the goals of innovative activity, and the barriers to innovation in the period 2012–2014 (dimensions 1–2; 69.52% of total inertia).

Table 21. List of assumptions and calculations necessary to verify the hypothesis regarding the relationship between the type and ownership sector of an enterprise and the reasons for the lack of innovation and barriers to innovation (2012–2014).

Pearson's χ^2 Test of Independence	
Null hypothesis (H_0)	**The enterprise type and ownership sector have no impact on the reasons for the lack of innovation and barriers to innovation**
Alternative hypothesis (H_1)	The type and ownership sector of the enterprise have an impact on the reasons for the lack of innovation and barriers to innovation
χ^2 **statistics value**	251.602
Critical region	right-tailed
Level of significance (α)	$\alpha = 0.05$
$\boldsymbol{p}$ (p)	$p = 0.9999$
Decision	Since $p > \alpha$, there are no grounds for rejecting H_0

The duration and length of each wave are given in brackets. A characteristic feature of the presented concept is the shortening of the length of each subsequent wave, which is explained by the acceleration of scientific and technological progress. Obviously, the latest wave is predictive, however, it should be noted that its beginning was dated to 2015, so it almost coincided with the period of significant structural changes in the Polish manufacturing industry.

Table 22. Variables describing the reasons for the lack of innovation and barriers to innovation in the years 2012–2014, their degrees of importance for enterprises, and the method of coding.

Reasons for a Lack of Innovation	Factors Impeding Innovative Activity Scale: 1—High; 2—Medium; 3—Low; 4—Irrelevant	Codes	Degree of Importance
No compelling reason for introducing innovation	Low demand for innovation on market	BR_1	1, 2, 3, 4
	No need to implement innovation due to earlier innovations	BR_2	1, 2, 3, 4
	No need to implement innovation due to low competition on the market	BR_3	1, 2, 3, 4
	Lack of good ideas for innovation	BR_4	1, 2, 3, 4
The implementation of innovations was considered, but the barriers proved to be too high	Lack of financing opportunities for innovation from the company's internal sources	BR_5	1, 2, 3, 4
	Lack of financing for innovation from external sources – loans or funds under private equity financing (including venture capital)	BR_6	1, 2, 3, 4
	No staff with the right skills in your company	BR_7	1, 2, 3, 4
	Difficulties in obtaining public grants or subsidies for innovation	BR_8	1, 2, 3, 4
	No partners to cooperate with	BR_9	1, 2, 3, 4
	Uncertain market demand for your ideas for innovation	BR_10	1, 2, 3, 4
	Too much competition on the market	BR_11	1, 2, 3, 4

According to Šmihula, the typical end of every K-wave is the economic crisis, which is characterized by stagnation caused by technological stalemate and increased demand for new inventions and innovations [8]. The crisis ending the application phase creates good conditions for the emergence of new inventions, but it takes some time for a new technological revolution to start and technological innovations capable of stimulating investment growth to appear. In his opinion, these changes are practically impossible to be proven by statistical methods due to strong relativism in the assessment of inventions and innovations.

The results presented in this article seem to coincide with the concept of modern K-waves that end and begin with an economic crisis. Considering the two factors mentioned by Kondratieff [7], namely changes in technology and the assimilation of new countries into the world economy, it can be assumed that the transitions occurring in the Polish manufacturing sector are part of a larger whole. If we consider the Fifth Kondratieff wave (i.e., information and telecommunications revolution), it seems that many signs of the end of its application phase can be observed. As a result of continuous improvement of the related innovations, it can be observed that information technology has long become a part of everyday economic life. It is likely that greater profits and revolutionary inventions may appear soon in other industries. There are also many indications that the Sixth Kondratieff wave will be associated with the biomedical-hydrogen revolution. With this approach to the problem, the discovered trend is a fragment of a global phenomenon that is associated with the breakthrough between the Fifth and Sixth Kondratieff waves [37]. The Polish economy is already so integrated with the global economy and included in international supply chains that it can reflect global trends. Therefore, the breakthrough can be determined by statistical methods, however, secular factors have to be considered.

4. Discussion

This paper is a summary of the research carried out so far in the area of innovation in Polish manufacturing enterprises [35–37]. The obtained results prove that the cybernetic approach is of great importance in researching the innovative activity of enterprises. The ownership sector and the type of enterprise are among the most important factors of innovative activity, however, considering their impact separately may cause that some phenomena remain unrecognized. Examining the mutual influence of these variables has a significant impact on results, as demonstrated by the example of the third period. Focusing only on the types of enterprises, their statistical significant impact on the goals of innovative activity was confirmed, however, after including ownership sectors in the considerations and after taking into account the interactions between ownership and the size of the enterprises, the situation changed radically. This means that the long-term trend referred to in the article could remain undiscovered.

Regarding the poor performance of public enterprises in the field of innovation, it was assumed that one of the reasons may be the political criteria for the selection of managerial staff in these enterprises. Another explanation points to significant pay disparities between employees in the public and private sectors that are working to the disadvantage of the public sector. However, research shows that in terms of innovation, public sector enterprises do not have to perform worse than private ones. An example is Chinese state-owned enterprises, which have gained an advantage over private companies in the field of process innovation. It should not be forgotten, however, that in China after 2000, over 90% of government-owned corporations adopted the Modern Enterprise System, which consisted of implementing corporate or shareholding reforms and adopting a sound corporate structure, as a result of which the boards of shareholders, directors, supervisors, and managers were created in them [34]. Perhaps similar reforms are required by Polish industrial processing enterprises operating in the public sector.

The Red Queen hypothesis indicates that the source public enterprises' problems may be the strong selection pressure exerted on them by a more innovative environment. The gradual decline in the share of public enterprises in the manufacturing section is worrying, as it may be a source of some political and economic perturbations and have a negative impact on the entire economy. Some enterprises should certainly remain public due to the need to achieve certain social goals.

The insensitivity of the effects and goals of the innovative activity of enterprises to the business cycle phases most likely means that in the total span of the three studied periods, the importance of secular factors was greater than that of business cycles. From this point of view, the global financial crisis should be treated as a phenomenon that strengthened and accelerated the operation of the selection mechanism, as a result of which most innovative enterprises improved their market position. Some less innovative enterprises could also benefit as the Red Queen effect could have forced them to have a reverse reaction in the form of increased innovation efforts. The appearance of positive feedback would mean that the crisis probably accelerated the innovation race, thus improving the situation of most companies in the third period, as evidenced by the near zero inertia of the system under study.

The effect of the Polish Green Island signifies that very good results of the majority of industrial processing enterprises in the field of innovation depend to a small extent on the phases of the business cycle. In the three examined periods, the majority of enterprises did not have problems with achieving the assumed effects or goals of innovative activity, and barriers to innovation gradually disappeared [35–37]. This is demonstrated by Poland's economic successes in the last thirty years, which would not have been possible without the great innovation activity of enterprises. In addition to the strong and sustained economic growth mentioned earlier, there are other important achievements to be mentioned including economic opening to the world, reduction of inflation, and increase in welfare [48]. When the studied periods are examined in a comprehensive way, it can be observed that the Polish economy is more affected by secular trends than by cyclical fluctuations. Entrepreneurs have a good understanding of these phenomena, which indicates the appropriate use of information sources for innovative activities.

Research on innovation has provided convincing statistical evidence to support the claim that the global financial crisis heralded a breakthrough that was the end of the Fifth and the beginning of the Sixth Kondratieff wave. The following two discoveries should be mentioned here: the trend associated with the decreasing inertia of the industrial processing section, which has recently been close to zero, and the essential importance of ownership, as its inclusion in the analysis has completely changed the results. The implementation of the goals set by most enterprises in 2012–2014 may indicate that the application phase of the Fifth K-wave, associated with the information and telecommunications revolution, is slowly coming to an end. The level of competition in the manufacturing industry gradually increased, so in the third period, it could already be close to the maximum. The deadlock can be broken only due to new, breakthrough technological innovations that will open new opportunities for economic growth, profit increase, and competition. Information technology has already become an integral part of everyday economic life and will certainly be further used and refined, but it has apparently already used its potential as a driving force for economic growth and development. Many new inventions have appeared on the horizon that are related to biotechnology, nanotechnology, biomedicine, and hydrogen as the fuel of the future. Soon, they can become the basis of a new technological revolution that will initiate the Sixth K-wave. The results presented in this article are consistent with the forecast of Šmihula [8], which dated the beginning of the biomedical-hydrogen revolution for 2015. Therefore, it is quite possible that we are already living in the Sixth K-wave without knowing anything about it. We are reminded in this respect of Monsieur Jourdain, the hero of a five-act comédie-ballet *The Middle Class Gentleman* written by Molière, who said to the Master of Philosophy [50]: *By my faith! For more than forty years I have been speaking prose without knowing anything about it, and I am much obliged to you for having taught me that*.

5. Conclusions

The presented research focused on the impact of ownership sectors and types of enterprises on the innovation activities of companies, considering changes in the economic environment in which these activities happen. In the three examined periods, a gradual increase in innovation activity was found and it was not disturbed by cyclical fluctuations [35–37]. All enterprises, apart from public ones, achieved the assumed effects and goals of innovation activities, which contributed to the creation of positive feedback loops leading to further growth of innovation. There was also another phenomenon that is strongly associated with the first, consisting of a gradual decrease in the importance of innovation barriers, until they became practically imperceptible to enterprises in the last of the examined periods. The Red Queen effect indicates strong competition between enterprises in the field of innovation, and data show that most of them have met this challenge. The long-term growth of innovative activity was not even slowed down by the global financial crisis, which became a selection mechanism for enterprises and mobilized them to increase their efforts in the field of innovation. This contributed to even better corporate performance during the recovery period of 2012–2014. These discoveries are irrefutable evidence that the Polish Green Island effect is a real phenomenon, not a government propaganda trick. They also explain the reasons for Poland's incredible economic success over the past twenty-seven years, which include strong and uninterrupted economic growth, opening of the economy to the world, controlled inflation, and reduction in unemployment.

The cybernetic approach, consisting of considering the combined impact of ownership sectors and types of enterprises on innovation activities, has contributed to the discovery of a long-term trend of a steady decrease in the inertia of the entire industrial processing section. In the last of the examined periods, the inertia was already close to zero. This means that the activity of enterprises was more influenced by secular changes than by cyclical fluctuations. This effect was discovered after including the ownership sectors in the considerations, and it did not occur when only the impact of the size of enterprises on their innovativeness was examined. The secular changes determining the innovative activity of companies were rather external to the Polish economy and were the result of the impact of global trends. It is hard to imagine that the source of these changes might lie in internal conditions

because endogenous changes would be too weak compared to exogenous influences. The industrial processing sector simply reflected global trends due to the fact that the Polish economy was already an integral part of the world economy.

Identification of the carriers of impulses from the global economy to the Polish manufacturing sector is a very important issue. This role was undoubtedly played by the ownership because its inclusion in the considerations revealed the trend of decreasing inertia of the examined system. In order to justify this view, it is necessary to establish the ownership structure in the private and mixed sectors, and in particular, to separate the share of foreign ownership in these sectors. Available data show that in 2010, the percentage of foreign capital in basic (share) capital in the Polish processing industry was 47.9%, while in banking, it exceeded as much as 75% [51] (p. 16). This confirms not only that foreign ownership plays the role of a communication channel transferring knowledge from the world economy to the Polish industrial processing section, but also gives an idea of the channel's high capacity.

The long-term trend of decreasing inertia can serve as evidence that the global financial crisis was associated with the turn of the Fifth and Sixth K-waves (i.e., the transition from the information and telecommunications revolution to the biomedical-hydrogen revolution). Such crisis is a typical phenomenon because it results from the technological deadlock related to the exhaustion of the investment potential of the technology used so far. The technical progress that has taken place in recent years has resulted in many inventions, and these may, in a relatively short time, turn into technological innovations and usher in a new technological revolution. The rate of profit associated with new innovations can be much higher than the one currently provided by information technologies. It is natural, therefore, that during such a crisis, investments decrease and enterprises are looking for new opportunities to develop and overcome competition. When applying this reasoning to the Polish economy, it should be noted that despite its stunning successes as discussed above, the investment rate is decreasing, which is interpreted as a threat to the economic growth in the following years. Of particular concern are the investments of enterprises in machines and devices [48,52]. However, it is possible to look at this problem differently. It cannot be ruled out that entrepreneurs in Poland, knowing about global trends, are refraining from investing in information technologies to take a good position in the upcoming Sixth K-wave. A typical phenomenon is that at the beginning of each wave of technological innovations, there are a lot of relatively small enterprises that use many different technological methods. After some time, as a result of concentration processes, only a few semi-monopolistic enterprises remain on the market, and the number of technological methods decreases to the few most efficient. The current investment problem may therefore be whether to invest in old technologies or new ones.

Author Contributions: Conceptualization, A.J. and D.R.; Methodology, A.J. and D.R.; Software, A.J. and D.R.; Validation, A.J. and D.R.; Formal analysis, A.J. and D.R.; Investigation, A.J. and D.R.; Resources, A.J. and D.R.; Data curation, A.J. and D.R.; Writing—original draft preparation, A.J. and D.R.; Writing—review and editing, A.J. and D.R.; Visualization, A.J. and D.R.; Supervision, A.J. and D.R.; Project administration, A.J. and D.R.; Funding acquisition, A.J. and D.R. All authors have read and agreed to the published version of the manuscript.

Funding: The study presented here was conducted as part of the Opus 9 project entitled "Różnicowanie się strategii innowacji polskich przedsiębiorstw przetwórstwa przemysłowego pod wpływem zmian uwarunkowań makroekonomicznych" (Differentiation of innovation strategies of Polish manufacturing firms as an effect of changes in macroeconomic environment) funded by the National Science Centre, Poland, under contract no. UMO-2015/17/B/HS4/02742. The statistical data used in the calculations originated from the Statistical Office in Szczecin, which assumes no responsibility for the conclusions reached in the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wagemann, E. *Economic Rhythm: A Theory of Business Cycles*; McGraw-Hill Book Company: New York, NY, USA, 1930.

2. Besomi, D. Naming crises: A note on semantics and chronology. In *Crises and Cycles in Economic Dictionaries and Encyclopaedias*; Besomi, D., Ed.; Routledge—Taylor & Francis Group: New York, NY, USA, 2012; pp. 54–132. ISBN 978-0-415-49903-3.
3. Braudel, F. *Civilization and Capitalism, 15th–18th Century: The Perspective of the World*; University of California Press: Berkeley, CA, USA, 1992; Volume 3, ISBN 978-0520081161.
4. Brødsgaard, K.E.; Rutten, K. *From Accelerated Accumulation to Socialist Market Economy in China: Economic Discourse and Development from 1953 to the Present*; Koninklijke Brill NV: Leiden, The Netherlands, 2017; ISBN 978-90-04-33008-5. [CrossRef]
5. Peck, J. *Constructions of Neoliberal Reason*; Oxford University Press: New York, NY, USA, 2010; ISBN 978-0-19-958057-6. [CrossRef]
6. Thomson, J.K.J. Variations in industrial structure in pre-industrial Languedoc. In *Manufacture in Town and Country before the Factory*; Berg, M., Hudson, P., Sonenscher, M., Eds.; The Press Syndicate of the University of Cambridge: Cambridge, UK, 2002; pp. 61–91. ISBN 0-521-89359-3.
7. Kondratieff, N.D. The long waves in economic life. *Rev. Econ. Stat.* **1935**, *17*, 105–115. [CrossRef]
8. Šmihula, D. The waves of the technological innovations of the modern age and the present crisis as the end of the wave of the informational technological revolution. *Studia Politica Slovaca* **2009**, *2*, 32–47. Available online: https://www.sav.sk/journals/uploads/04121424SPS_1_2009_%20D%20Smihula.pdf (accessed on 25 November 2019).
9. Schumpeter, J.A. *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*; McGraw-Hill Book Company: New York, NY, USA, 1939; Volumes 1–2.
10. Schumpeter, J.A. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*; Harvard University Press: Cambridge, MA, USA, 1949.
11. Schumpeter, J.A. *Capitalism, Socialism, and Democracy*, 2nd ed.; Harper and Brothers Publishers: New York, NY, USA, 1942.
12. Beinhocker, E.D. *The Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*; Harvard Business School Press: Boston, MA, USA, 2006; ISBN 1-57851-777-X.
13. Arthur, W.B. *Complexity and the Economy*; Oxford University Press: New York, NY, USA, 2015; ISBN 978-0-19-933429-2.
14. Świadek, A. Sales range and innovative activity in the manufacturing system of Poland. *Equilibrium* **2018**, *13*, 725–740. [CrossRef]
15. Spitsin, V.; Mikhalchuk, A.; Chistyakova, N.; Spitsyna, L.; Pavlova, I. Development of innovative industries in Russia under unfavourable external environment. *Equilibrium* **2018**, *13*, 467–485. [CrossRef]
16. Lewandowska, A.; Stopa, M. Do SME's innovation strategies influence their effectiveness of innovation? Some evidence from the case of Podkarpackie as peripheral region in Poland. *Equilibrium* **2019**, *14*, 521–536. [CrossRef]
17. Pisar, P.; Bilkova, D. Controlling as a tool for SME management with an emphasis on innovations in the context of Industry 4.0. *Equilibrium* **2019**, *14*, 763–785. [CrossRef]
18. Gorączkowska, J. Influence of business support organizations on innovation activity in manufacturing companies in the Masovian Voivodeship in Poland. *Equilibrium* **2018**, *13*, 741–759. [CrossRef]
19. Lewandowska, A.; Stopa, M. SMEs innovativeness and institutional support system: The local experiences in qualitative perspective. Polish case study. *Oecon. Copernic.* **2018**, *9*, 333–351. [CrossRef]
20. Zygmunt, A. External linkages and intellectual assets as indicators of firms' innovation activities: Results from the Czech Republic and Poland. *Oecon. Copernic.* **2019**, *10*, 291–308. [CrossRef]
21. Świadek, A.; Dzikowski, P.; Tomaszewski, M.; Gorączkowska, J. Sectoral patterns of innovation cooperation in Polish industry. *Equilibrium* **2019**, *14*, 183–200. [CrossRef]
22. Kijek, T.; Matras-Bolibok, A. The relationship between TFP and innovation performance: Evidence from EU regions. *Equilibrium* **2019**, *14*, 695–709. [CrossRef]
23. Cieślik, A.; Michałek, J.J. Process and product innovations, multi-product status and export performance: Firm-level evidence from V–4 countries. *Equilibrium* **2018**, *13*, 233–250. [CrossRef]
24. Derevianko, O. Reputation stability vs anti-crisis sustainability: Under what circumstances will innovations, media activities and CSR be in higher demand? *Oecon. Copernic.* **2019**, *10*, 511–536. [CrossRef]
25. Peters, E.E. *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*; John Wiley & Sons: New York, NY, USA, 1994; ISBN 0-471-58524-6.

26. Gell-Mann, M. *The Quark and the Jaguar: Adventures in the Simple and the Complex*, 8th ed.; W.H. Freeman and Company: New York, NY, USA, 2002; ISBN 0-7167-2725-0.
27. Bryant, J. *Entropy Man*; VOCAT International Ltd.: Harpenden, UK, 2015; ISBN 978-0-9562975-4-9.
28. Gell-Mann, M.; Lloyd, S. Information measures, effective complexity, and total information. *Complexity* **1996**, *2*, 44–52. [CrossRef]
29. Brossard, O.; Lavigne, S.; Sakinç, M.E. Ownership structures and R&D in Europe: The good institutional investors, the bad and ugly impatient shareholders. *Ind. Corp. Chang.* **2013**, *22*, 1031–1068. [CrossRef]
30. Falk, M. Effects of foreign ownership on innovation activities: Empirical evidence for twelve European countries. *Natl. Inst. Econ. Rev.* **2008**, *204*, 85–97. [CrossRef]
31. Minetti, R.; Murro, P.; Paiella, M. Ownership structure, governance, and innovation. *Eur. Econ. Rev.* **2015**, *80*, 165–193. [CrossRef]
32. Ortega-Argilés, R.; Moreno, R.; Caralt, J.S. Ownership structure and innovation: Is there a real link? *Ann. Reg. Sci.* **2005**, *39*, 637–662. [CrossRef]
33. Bitler, M.P.; Moskowitz, T.J.; Vissing-Jørgensen, A. Testing agency theory with entrepreneur effort and wealth. *J. Finance* **2005**, *60*, 539–576. [CrossRef]
34. Huang, Y.; Salike, N.; Yin, Z.; Zeng, D.Z. Enterprise innovation in China: Does ownership or size matter? *RIEI Work. Pap. Ser. 2017-06*. Xi'an Jiaotong-Liverpool University, Research Institute for Economic Integration, 2017. Available online: http://58.210.89.21/RePEc/xjt/working-papers/RIEI-WP_2017-06.pdf (accessed on 15 December 2019).
35. Jakimowicz, A.; Rzeczkowski, D. Diversification of innovation strategies of Polish industrial processing enterprises depending on their size after the global financial crisis. *JEMI* **2019**, *15*, 35–76. [CrossRef]
36. Jakimowicz, A.; Rzeczkowski, D. Do barriers to innovation impact changes in innovation activities of firms during business cycle? The effect of the Polish green island. *Equilibrium* **2019**, *14*, 631–676. [CrossRef]
37. Jakimowicz, A.; Rzeczkowski, D. Firm ownership and size versus innovation activities over the business cycle: Near-zero inertia as a sign of the transition from the fifth to the sixth Kondratieff wave. *Oecon. Copernic.* **2019**, *10*, 689–741. [CrossRef]
38. Commission Regulation (EU) No. 651/2014 of 17 June 2014 declaring certain categories of aid compatible with the internal market in application of Articles 107 and 108 of the Treaty. *OJEU* L 187/1, 26.6.2014. Available online: https://eur-lex.europa.eu/eli/reg/2014/651/oj (accessed on 12 March 2020).
39. McHugh, M.L. The Chi-square test of independence. *Biochem. Med.* **2013**, *23*, 143–149. [CrossRef] [PubMed]
40. Greenacre, M. *Correspondence Analysis in Practice*, 2nd ed.; Chapman & Hall/CRC—Taylor & Francis Group: Boca Raton, FL, USA, 2007; ISBN 978-1-58488-616-7.
41. Nenadić, O.; Greenacre, M. Correspondence analysis in $\mathbb{R}$, with two- and three-dimensional graphics: The ca package. *J. Stat. Softw.* **2007**, *20*, 1–13. [CrossRef]
42. Jakimowicz, A.; Rzeczkowski, D. Prosumption in the public administration sector. *Acta Phys. Pol. A* **2016**, *129*, 1011–1017. [CrossRef]
43. Carroll, L. *Through the Looking-Glass, and what Alice Found There*; Macmillan and Co.: London, UK, 1872.
44. Van Valen, L. A new evolutionary law. *Evol. Theory* **1973**, *1*, 1–30.
45. Tomescu-Dubrow, I.; Słomczyński, K.M.; Domański, H.; Dubrow, J.K.; Sawiński, Z.; Przybysz, D. *Dynamics of Class and Stratification in Poland*; Central European University Press: Budapest, Hungary, 2018; ISBN 978-963-386-155-4.
46. Tomescu-Dubrow, I.; Dubrow, J.K.; Kiersztyn, A.; Andrejuk, K.; Kołczyńska, M.; Slomczynski, K.M. *The Subjective Experience of Joblessness in Poland*; Springer Nature Switzerland AG: Cham, Switzerland, 2019; ISBN 978-3-030-13646-8. [CrossRef]
47. *Poland. Country Study Guide: Strategic Information and Developments*; International Business Publications: Washington, DC, USA, 2013; Volume 1, ISBN 1-4387-7534-2.
48. Antoniak, A.; Mrowiec, M.; Piękoś, P. *Polska gospodarka i wyzwania na najbliższe dekady*; The report was presented on September 3, 2019 during the XXIX Economic Forum in Krynica-Zdrój; Bank Pekao: Krynica-Zdrój, Poland, 2019. (In Polish)
49. Bogdan, W.; Boniecki, D.; Labaye, E.; Marciniak, T.; Nowacki, M. *Poland 2025: Europe's New Growth Engine*; McKinsey & Company: Warsaw, Poland, 2015.
50. Molière. In *The Middle Class Gentleman*; Available online: https://www.gutenberg.org/files/2992/2992-h/2992-h.htm (accessed on 12 November 2019).

51. Soroka, P. Deindustrializacja po 1989 roku i potrzeba reindustrializacji w Polsce. *Przegląd Geopolityczny* **2019**, *29*, 9–24. Available online: http://przeglad.org/wp-content/uploads/2019/09/XXIX-01-Soroka.pdf (accessed on 10 April 2020). (In Polish).
52. *Inwestycje w Polsce. Szanse i zagrożenia*; Związek Przedsiębiorców i Pracodawców: Warsaw, Poland, 2019. Available online: https://zpp.net.pl (accessed on 21 January 2020). (In Polish).

entropy

MDPI

Article

Allometric Scaling of Mutual Information in Complex Networks: A Conceptual Framework and Empirical Approach

Eduardo Viegas [1,*,†], Hayato Goto [1,2,†], Yuh Kobayashi [2,3], Misako Takayasu [2,3], Hideki Takayasu [2,4] and Henrik Jeldtoft Jensen [1,2]

1 Centre for Complexity Science and Department of Mathematics, Imperial College London, London SW7 2AZ, UK; h.goto@imperial.ac.uk (H.G.); h.jensen@imperial.ac.uk (H.J.J.)
2 Institute of Innovative Research, Tokyo Institute of Technology, 4259, Nagatsuta-cho, Yokohama 226-8502, Japan; kobayashi.y.bz@m.titech.ac.jp (Y.K.); takayasu.m.aa@m.titech.ac.jp (M.T.); takayasu.h.aa@m.titech.ac.jp (H.T.)
3 Department of Mathematical and Computing Science, School of Computing, Tokyo Institute of Technology, Yokohama 226-8502, Japan
4 Sony Computer Science Laboratories, 3-14-13, Higashi-Gotanda, Shinagawa-ku, Tokyo 141-0022, Japan
* Correspondence: e.viegas11@imperial.ac.uk
† These authors contributed equally to this work.

Received: 20 January 2020; Accepted: 10 February 2020; Published: 12 February 2020

Abstract: Complexity and information theory are two very valuable but distinct fields of research, yet sharing the same roots. Here, we develop a complexity framework inspired by the allometric scaling laws of living biological systems in order to evaluate the structural features of networks. This is done by aligning the fundamental building blocks of information theory (entropy and mutual information) with the core concepts in network science such as the preferential attachment and degree correlations. In doing so, we are able to articulate the meaning and significance of mutual information as a comparative analysis tool for network activity. When adapting and applying the framework to the specific context of the business ecosystem of Japanese firms, we are able to highlight the key structural differences and efficiency levels of the economic activities within each prefecture in Japan. Moreover, we propose a method to quantify the distance of an economic system to its efficient free market configuration by distinguishing and quantifying two particular types of mutual information, total and structural.

Keywords: complexity science, information theory, economic complexity, evolutionary dynamics, network theory

1. Introduction

One can argue that statistical physics and theoretical computing are the common roots feeding the science branches of complexity and information theory, as attested by the early exchanges of ideas between von Neumann and Shannon. Whereas the latter was more preoccupied in quantitatively measuring the encoding and transmission of information [1], the former (as articulated by the automata theory) had its focus on information replication with mutation but without generating tendencies [2] (i.e., self replication, or 'evolution') as well as the processing functions at an individual and aggregated level [3] (i.e., general automata and basic organs, or 'emerging scaling properties of a network structure'). Since then, these aforementioned fields of science have progressed significantly, and developed to an extent that they seem to bear little in common. Yet, significant insight can be obtained if one were to recombine these fields and develop a framework articulating the link between the emerging structural properties of a network and the flow, or encoding, of information within [4,5].

Given the importance of evolution and scaling to such framework, useful mathematical methods can be applied by borrowing concepts from the biological, natural world, in particular the diversity of species and allometric scaling [6–8].

In precis, this is the core motivation and aim of our research. Here, we create a method within the context of a network flow of resources to measure two fundamental quantities underpinning information theory, namely entropy and mutual information. This method is then wrapped into a framework that draws parallels to the biological context of body growth and allometric scaling so that the meaning and significance of mutual information within this construction can be better understood and intuitively rationalised.

1.1. The Context

From an empirical perspective, we make use of the real 'Interfirm Business Transaction Network' within Japan, consisting of detailed granular level transaction data among over 600,000 firms during a 25-year period, 1994 to 2018, provided by Teikoku Databank. This rich dataset allows the breakdown of an extensive network into smaller subgraphs, at a prefecture level, so that a dynamic comparison between different segments of the network can be carried out. Previous works on the real trade Japanese network [9–11] are centred on the system dynamics surrounding the formation of the network, as well as the structural analysis [12] including studies on allometric scaling [13] of quantities such as sales and income. Our research, however, adopted a fundamentally distinct approach as it seeks to answer different questions. Here, we are less preoccupied with the dynamics of network formation. Instead, our focus is mostly on constructing a framework based on mutual information and resource usage efficiency, akin to metabolism, that allows for a direct comparison of different regional economic activities, in this case of 47 Japanese prefectures. We emphasise that we have in mind that mutual information represents effective cash flow (i.e., the movement of resources) between companies, since the scaling properties of the real trade network, referred above, allow us to make use of the degree of a company as a proxy that can be measured directly from the data, so that we avoid issues surrounding cash flow estimation.

Furthermore, we note that our work bears some similarity to existing ecological network analysis for economic systems, where resources defined as currency cash flows are used as the basis for calculating entropy and mutual information [14]. However, data, methods, and objectives fundamentally differ. Importantly, we are not preoccupied with measuring economic development. Instead, our focus is on the understanding of structural evolution of trade networks as described above.

Since our work is applied to a real countrywide economic system and financial network, it is a requirement for the conceptual framework to be adapted in order to incorporate key fundamental economic principles. This is to ensure that the concept of mutual information is aligned not only to biological metabolic rates but to the specifics and concrete elements of the network studied as well. Essentially, this means that the analogy to metabolic rate is further extended to define the average resources used by a company to generate new trades (and the related cashflows and income). In very simplified terms, since the focus here is not on detailed finance and economics, the metabolic rate can be generically equated to a cost to acquire new trades [15].

1.2. The Complexity Framework: Allometric Nature of Mutual Information

From a conceptual perspective, our framework can be regarded as a triangulation between concepts arising from three different fields of study: Network science, information theory, and economics.

Firstly, from a network science perspective, previous academic works show that the distribution of nodes and edges for the Japanese inter-firm trade network follows a power law distribution governed by mechanisms associated with a cumulative advantage [16] and preferential attachment [17]. Essentially, these mechanisms tend to lead, but not inevitably, to the formation of a disassortative

network, essentially meaning that the average number of nodes connected to a specific selected node tends to decrease as the degree of the latter gets larger. Regardless of the specific mechanisms of a network, the power law structure will always tend to lead to a level of disassortativeness [18].

From an information theory perspective, it follows that an amount of mutual information will always be different from zero if the network is disassortative, simply as a result of the functional forms of entropy and mutual information. Therefore, within our framework, it is possible to break the computation of the mutual information into two separate but related components: The structural mutual information, SMI, and the total mutual information, I. The former solely relates to the degree distribution of the nodes within a given network, whereas the latter encompasses both the node degree distribution as well as the disassortativeness of the network.

Such distinction also fits well within the economics and finance perspective since SMI can be related to a theoretical 'free-market', stock market-type configuration, whereas I is not only naturally associated with, but reflects, the real world situation.

Once the above is addressed, we overlay the biological dimension of our framework.

1.2.1. Structural Mutual Information: SMI

The structural mutual information SMI is intended to capture the basic quantities held by the network simply as a result of the power law like degree distribution of companies, and their related sizes, within a given network. Essentially, we make use of the term 'structural' to refer to the basic existence of the nodes without taking into account the dynamics of the preferential attachment and cumulative advantage mechanism. The method is inspired by the allometric scaling and power laws in ecological systems. In particular, we make use here of the allometric scaling equation leading to an analogy whereby SMI and I for each prefecture can be related to the metabolic rates B of an individual which is known to scale with body size M as:

$$B = Q_1 R M^b \tag{1}$$

where the exponent $b = 3/4$ has been suggested to describe a range of biological cases [6,8]. The two other elements, $Q_1 R$ capture, essentially, the variability in resource supply rates as well as variables affecting body size and density.

1.2.2. Total Mutual Information, I

Further extending our analogy, SMI is akin to the resting, or basal, metabolic rate. In contrast, the total mutual information I contains the additional thermal food and physical effects. Within our framework, these two additional effects represent activities comparable to the way that companies express trading preferences among themselves (i.e., the dynamics of preferential attachment and cumulative advantage). Therefore, these dynamics act as a multiplier to the core, structural mutual information.

1.3. The Economic Dimension

Given that companies always aim to increase profits by maximising income and minimising costs, it is only natural to reason that the dynamics of preferential attachment and cumulative advantage become a natural feature of general business dynamics [19]. Specifically, small companies with very limited resources would tend to be most efficient when selling all their output to a single (or at least very few) company in order to reduce costs. In contrast, larger companies with additional resources will be driven by income expansion and therefore are willing to trade across as many agents as possible.

Here, without making any judgement about merits and disadvantages of distinct economic systems, we note that a centralised style communist system can be regarded as an extreme case of preferential attachment since virtually almost all market agents will almost solely trade with the largest entities (i.e., governments and large public companies). Yet, the current Western 'capitalist'

system also tends to lead to a virtual monopoly by the largest companies [9,20]. Therefore, one can reasonably argue that Western-style developed economies are no longer structurally capitalist, typically underpinned by a free market configuration. It is important to recognise, however, the fact that there is no single, or commonly accepted, definition of a 'free market economy' [21,22] within the field of economics. Therefore, an element of constrained licentia poetica will inevitably be required when attempting to define a theoretical free efficient market within the economic context. To mitigate the effects of such an issue, we specifically, and narrowly, interpret 'free market economies' to be those that structurally resemble the dynamics of organised markets, such as the stock markets. Within these markets, each unit traded, such as a single quantity stock or the minimum denomination of a bond is not dependent on any trait of the buyer or sell, or any other trade activity. Therefore, the preferential attachment mechanism is virtually absent, since the identity of the buyer is unlikely to be known by the seller and vice versa. Yet, a higher probability that a small agent will trade with larger agents, such as pension and investment funds, are still likely to exist simply due to sheer size. It is a fact that a large entity will have a higher number of trades, but will also be subject to some scaling of costs due to a higher activity.

Adapting the analogy for metabolic rates to the economic dimension, the structural mutual information SMI captures the mutual information solely arising as a result of the size of companies as if these companies were theoretically trading on a configuration similar, or akin, to stock markets, or 'free markets'. In order to measure this component, we randomised the real network, preserving the structure of the nodes (i.e., to total degree distributions) while minimising the effects of preferential attachment (i.e., degree correlations). In contrast, the mutual information I is directly derived from the real network configuration, which includes both the effect captured by the structural mutual information as well as the additional quantities arising from the dynamics of preferential attachment and cumulative advantage. By making this distinction, we are able to compare the efficiency of different Japanese prefectures and better understand the structure and activities of these prefectures.

2. Results

Our results are presented in three sections. Firstly, under the Data Analysis section, we present an analysis of the evolution of the entropies and mutual information for the Japanese prefectures over a 25-year period. This analysis is then enhanced and further analysed by layering the geographical dimension of the prefectures across Japan. The second section covers the results of applying the framework based on the allometric scaling of mutual information and related analogies to metabolism and biological systems to real world data. Thirdly, we zoom into a more microscopic level, specifically local interaction, analysing the formation and contribution to mutual information at a pointwise level, where the effects of the preferential attachment and cumulative advantage mechanisms can be clearly observed.

2.1. Data Analysis

By applying the grouping and coarse graining process as described in the Methods section, the values for entropy H, joint entropy J, and mutual information I for each prefecture were directly measured from real Japanese interfirm trade network data by making use of Equations (5), (6), and (7), respectively.

2.1.1. Macro Features of Entropy and Mutual Information

Consistently with established literature [18], the entropy, and the joint entropy, of the Japanese trade network in its totality or within each subgraph, i.e., the prefecture level, will tend to be higher as the system grows in size (i.e., increasing the number of nodes) as $H \sim \alpha \log N$. Moreover, the rate of growth α tends to be similar for all prefectures. Such a fact can clearly be observed in the left and centre plots within Figure 1a,b, where five representative prefectures are highlighted, from largest to smallest in terms of GDP (gross domestic product) size and selected on similar ranking intervals.

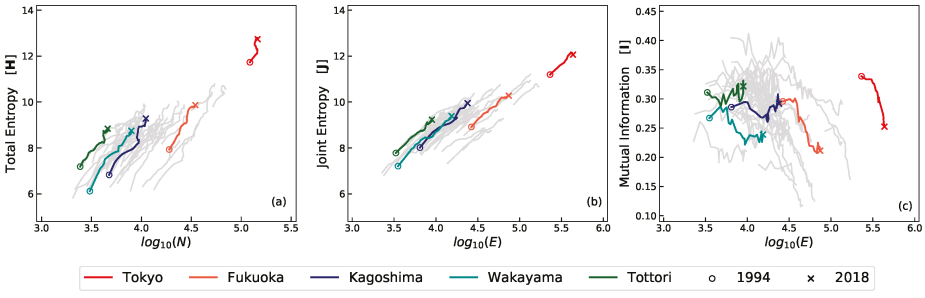


Figure 1. Entropy and mutual information prefectures in Japan between 1994 and 2018. Plot (**a**) shows the evolution of the total entropy H as a function of the total number of companies (nodes) $|\mathcal{C}|$ during the period 1994–2018. Similarly, plots (**b**,**c**) show the equivalent joint entropy J and mutual information I as a function of total number of edges E. Each grey line represents the path taken single prefecture in Japan during the period 1994–2018. The coloured lines highlighting representative prefectures, selected from the largest (i.e., Tokyo) to smallest (i.e., Tottori) in terms of GDP (gross domestic product) and maintaining similar ranking intervals in between (Fukuoka, Kagoshima, and Wakayama). The circles point to the year 1994 whereas the x-cross relates to 2018.

In contrast, the mutual information I exhibits a more complex, non linear, behaviour as shown in plot (c). One can note that I tends to decrease for the very large prefectures (i.e., Tokyo, Fukuoka) as the number of companies, (i.e., nodes) and trades (i.e., edges) increase. However, similar behaviour is not fully replicated in smaller prefectures (i.e., Kagoshima, Wakayama, and Tottori), where I may be increasing, stable, or decreasing. Moreover, and distinctly from H or J, the specific and numerical value of I bear a much weaker, albeit yet existing, relation to the system size as will be demonstrated further on, within the subsection allometric nature of the mutual information.

2.1.2. The Geographical Perspective of Mutual Information

As described above, comparison among prefectures of the numerical value of the total mutual information I at a given time gives little way of immediate insight, and their relative significances can only be appreciated once the association with metabolism as described by the framework within the next subsection, the allometric nature of the mutual information, is in place. However, a clear picture also emerges once a geospatial perspective is combined with a time series vector analysis for I, where the average rate of decline for each prefecture is linearly obtained by fitting $I_t = a + b(t - t_0)$.

The heatmap of Japan within Figure 2 shows the geographical distribution of the average rate of decline ($b < 0$) or increase ($b > 0$) to the mutual information I over the period in study, 1994–2018. It is easy to visualise that the highest rates of declines (i.e., red areas) are almost totally associated with the prefectures and urban conurbations of Japan's largest cities, with the sole exceptions of (a) Sapporo, a large city in a very large rural prefecture (Hokkaido) and (b) Oita where no immediate explanation can be found. In a consistent manner, the lowest rates of decline (or slight increase) are associated with the smallest prefectures, in economic terms as measured by the GDP, such as Tottori and Ehime. These results suggest that a time series analysis of the evolution of the mutual information provides a measure (the linear slope 'b' or the average changes to I), which reflects the level of the economic activity, or urbanism, for a given region. Essentially this approach can be feasibly used to potentially define economic clusters and conurbations in a quantitative manner through a single unit measure.

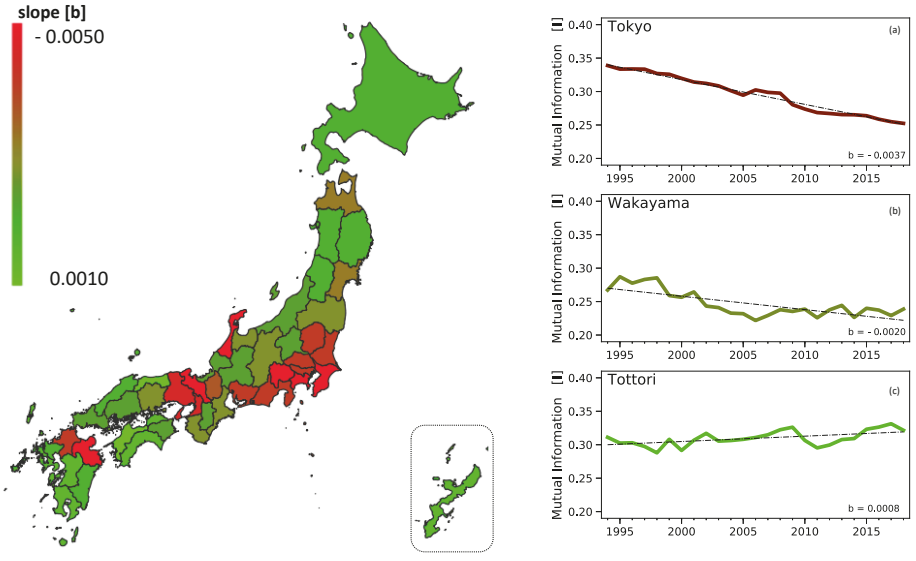


Figure 2. Average decline/increase rate of the mutual information within prefectures in Japan, 1994 to 2018. The map on the left consists of a geographical heatmap for the average yearly rate of decline (**red**) or increase (**light green**) of the mutual information for each of the 47 prefectures as approximated by a linear fitting $I_t = a + b(t - t_0)$. Each graph on the side corresponds to the evolution of the total mutual information I (y-axis) over the period 1994–2018 (x-axis) for (**a**) Tokyo, (**b**) Wakayama and (**c**) Tottori prefectures.

3. The Allometric Nature of Mutual Information

The results shown by each of the panels within Figure 3 lend important weight to the validation of a framework to quantify and evaluate mutual information within networks through the prism of biological metabolism and allometric scaling.

Firstly the distinction between structural mutual information SMI (or $\tilde{I}$ as explained within the Methods section) and total mutual information I and an analogy to basal metabolism and physical metabolism provide a useful description of the important differences between information (a) arising simply as a result of the existence of a node or company type and (b) that results from the dynamics and interaction between agents.

Within this context, it is to a certain extent remarkable to note the emergence of the 3/4 allometric scaling coefficient, as indicated (see Methods) by the results within Figure 3, and that by applying the same coefficients to both types of mutual information, SMI and I, one can observe that the distribution of variance is larger for the 'total metabolism' and are relatively small for the 'resting metabolism'. Moreover, such distributions of variance fit reasonably well to normally distributed curves (albeit with some differences towards the tail values) which indicate that such a variance is consistent with a generic random stochastic process. Here, we note that the number of datapoints are relatively limited, around 280, which can exaggerate the effects towards the tail.

Secondly, the framework and scaling of mutual information provide us with a valuable insight in terms of the economic structure of the prefectures: The 'structural efficiency' of a prefecture is not determined by size in isolation (as measured by the entropy) but by the diversity of the agent types within the system (which can be captured by SMI). Here it is important to pause and explain 'structural efficiency'.

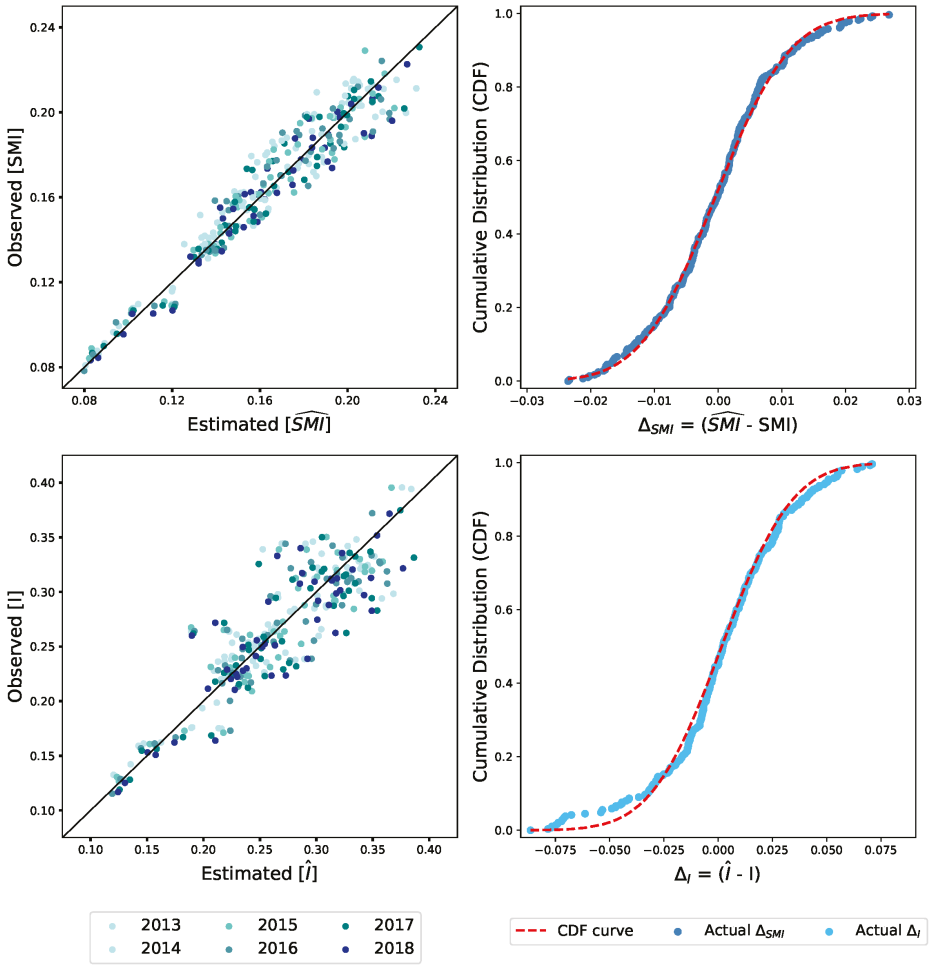


Figure 3. Observed and estimated values for the structural mutual information SMI and total mutual information I for the years 2013–2018. The **top left** panel shows the comparative results between the values observed for structural mutual information SMI (or $\tilde{I}$ as described within Methods) on the y-axis against those estimated by making use of Equation (11) on the x-axis. Each dot consists of a single prefecture within Japan at a specific year, with years colour mapped from lighter to darker shades, older to the most recent. The diagonal line represents the point where $y = x$. Similarly, the **bottom left** panel shows the observed mutual information I on the y-axis, calculated in accordance with Equation (7), against those estimated by the model by making use of Equation (12) on the x-axis. On the **right side**, the differences between the estimates to the actual values (x-axes) are ranked and plotted against the cumulative function of a normally distributed curve as shown by the red lines.

Within any business environment, a company would ideally like to sell products to every other company as it would increase sales and profits. However, resource limitation and costs of trading with various parties lead to a selection of business partners or 'preference to trade'. Therefore, from a narrow perspective, the more a company sells to a single partner, the lower the acquisition costs. However, one can argue that such an approach also leads to significant inefficiency within an economic system since opportunities for better and innovative trading and new links are reduced. Therefore,

SMI and I can be effectively viewed as an indicator as to the distance of a prefecture to a theoretical free market configuration.

Entropy and Mutual Information Micro Features

Whereas analysing the evolution of the mutual information from a macro level provides an insight with regards to the structural economic activity of prefectures, by zooming into the local, micro level, structures of interactions it is possible to better understand the impact of the essential mechanisms underpinning the network in study.

The effect of the preferential attachment and cumulative advantage mechanisms as catalysers to the generation of mutual information can be clearly observed by analysing Figure 4. The pointwise contribution to the mutual information heatmaps for the real networks (left side), within Osaka and Kagoshima prefectures, show the larger absolute values to be concentrated at the left and top borders of the panels. In contrast, the equivalent maps (maintaining the same colour coding scales) for the randomised network show a much more homogenous distribution of values across the heatmaps, with contributions to the mutual information and pointwise values tending towards zero. Moreover, the zoomed maps for the real network (central panels) provide a neat illustration of the core relationship between the cumulative advantage and preferential attachment mechanisms and mutual information: By 'preferring' to attach to larger companies, smaller entities tend to 'repel' its own kind. As a result, the pointwise mutual information turns negative within the small to small region. In a consistent manner, higher levels of pointwise values tend to be stronger at the preference region (i.e., smaller to larger companies).

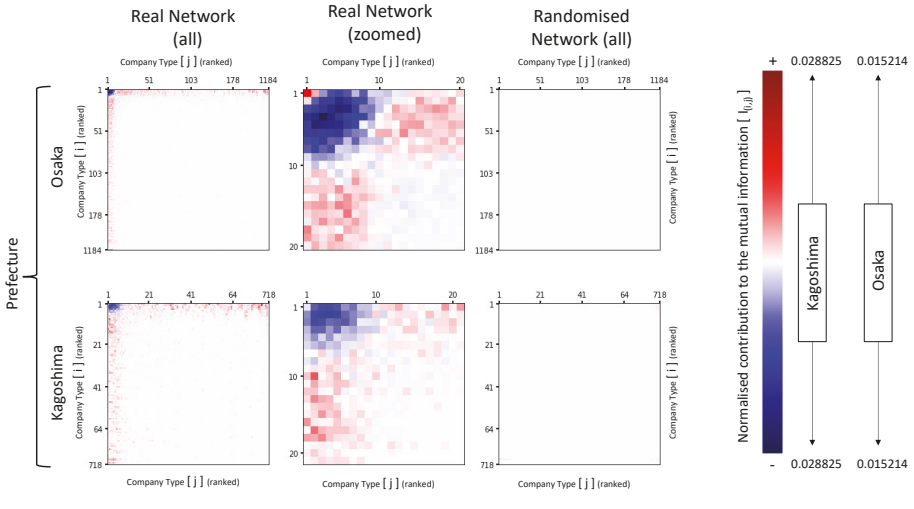


Figure 4. Pointwise contribution to the mutual information, $I_{(i,j)}$, for real and randomised networks. Each panel represents a heatmap of the pointwise contribution to mutual information for the directional edge combination 'i' (vertical axis) to 'j' (horizontal axis), calculated in accordance with Equation (8). Both axes are equal in value, consisting of the ranked sequence of the total degree distribution of companies for the relevant representative prefectures, Osaka in the top row and Kagoshima in the bottom row. The **left** (all degrees) and **centre** (zoomed degrees up to 20) panels show the contribution to the mutual information for the real network, whereas the **right** panels show contribution related to the randomised network. The colour maps on the right show the intensity of the contribution, with different scales by prefecture, but the same for all panels for the selected prefecture. Darker colours are associated with higher numbers with blue being negative values, red being positive, and totally white being zero.

However, it is important to note at this stage that by preserving the degree distribution of the nodes and at the same time maintaining the same number of edges, the randomisation process significantly reduces, but does not fully eliminate, the disassortative structure of the network, as shown in (a.2) and (c.2) within Figure 5. This is due to the fact that the neutral degree correlation under a power law degree distribution can only be achieved if the condition described in Equation (9) is satisfied. Therefore, the remaining level of disassortativeness can be regarded as 'structural', being the consequence of the power law distribution, since it is simply a fact that such a distribution of companies and company types do not allow for the probabilities of selection of all source and target nodes to be equal.

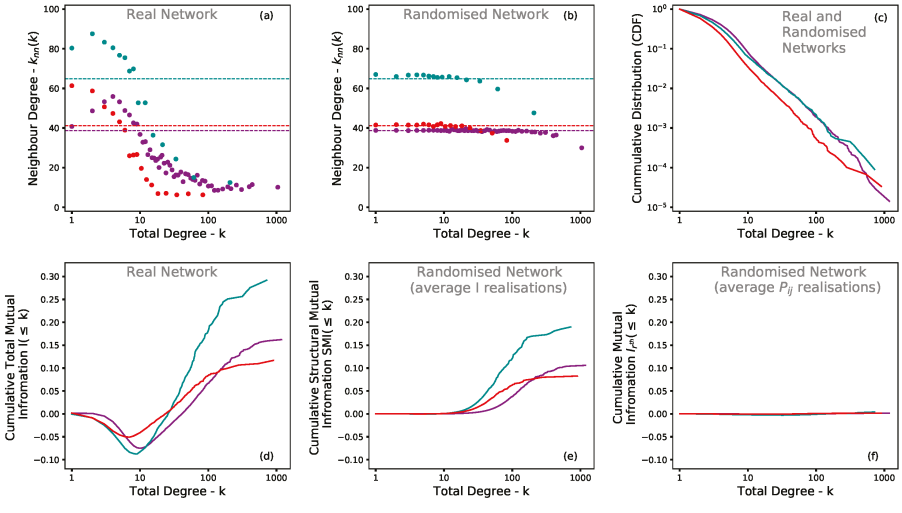


Figure 5. Average degree, population distribution, and cumulative mutual information values for selected Japanese prefectures in 2018. Panels (**a**,**b**) show the average degree of the neighbouring nodes [23] '$k_{nn}(k)$' (y-axis) of companies with total degree 'k' (x-axis, on a lognormal scale) for three selected prefectures: Osaka (magenta), Kagoshima (turquoise), and Saitama (red). Each dot represents the aggregate of companies of total degree 'k' and the average of their neighbours '$k_{nn}(k)$' generated through a binning process with a minimum of 1000 edges (i.e., datapoints) per bin. Whereas (**a**) relates to data extracted directly from the real network, (**b**) shows the average values for 1000 randomised realisations. Panel (**c**) consists of the total degree distribution of companies for the selected prefectures plotted on a log-log scale. The bottom panels (**d**,**e** and **f**) show the cumulative of the total mutual information I within the y-axis as a function of the degree distributions of companies, within the x-axis, on a lognormal scale. The left panel (**d**) relates to data from the real network, whereas the centre panels (**e**) consists of the average value of the mutual information I for each of the 1000 realisations adopting $P_{ij} = w_{ij}$ as calculated by Equations (3) and (7). In contrast, the right panel (**f**) consists of the calculation of a single value for mutual information $I_{r^{th}}$ for all aggregated realisations of P_{ij} as described in Equation (10).

4. Conclusions and Discussion

The results of this research indicated that the allometric scaling of the mutual information within the Japanese interfirm business networks to be akin and analogous to the metabolic rates of biological systems, providing further substance to the metaphor proposed by West [24] when researching the scaling of phenomena of cities and economies. Moreover, the 3/4 scaling exponent found in biological systems [6], as well some of the dynamics within cities and economies [7], fitted very well within

our complexity framework when applied to the Japanese economy at a national as well as regional prefecture level.

By measuring the mutual information at a national and regional prefecture, levels under our framework and method, and evaluating over an extensive time series, it is possible to appreciate the relationship between mutual information and the level of economic activity and urbanism of these prefecture, and therefore to place them into a comparative scale.

Moreover, we identified the structural mutual information SMI as the contribution arising as a result of the structure of the nodes, and segregated from the total mutual information I, which also includes the dynamics of interaction between agents. In doing so, we were able to clearly articulate that these quantities essentially represent the distance of a given economic structure from a theoretical free market configuration.

Such a finding helps to articulate a paradox which is essentially a core to today's economic analysis [25]. Whereas markets are most efficient when all agents are equally informed and have equal competitive chances (essentially there is no existence of preferential attachment and cumulative advantage mechanisms), these dynamics embedded within a capitalist system lead to a monopolistic configuration. Therefore, one could reason that in order to promote and protect free markets, governments and related agencies must actually intervene to mitigate the impact of the above mechanisms, and therefore be compelled to negate the more extreme interpretations of 'invisible hand' and 'laissez faire economics'.

From a micro level perspective, the analysis of the pointwise contribution to the mutual information showed that small companies tended to 'repeal' each other and be dependent on large entities. Again, one could argue that such a configuration is contrary to the efficient, free market configuration, and therefore has a potential focus to enhanced economic policy.

5. Materials and Methods

5.1. Measuring Entropy and Mutual Information

Individual companies c within the set of companies $\mathcal{C}$ are aggregated into groups S_k of companies with same total degree k:

$$S_k = \{c | k_c = k, c \in \mathcal{C}\} \tag{2}$$

where k_c is the total degree of company c.

By making an analogy to the real biological ecosystems, one may regard these groups as representing the average body size or mass of the individuals within that given group, since the total degree of companies scale in accordance with their sizes [11,13] within the business context, as measured by the number of employees, income, or total assets.

The expectation of edges between two groups S_i and S_k, or body sizes, is:

$$w_{ij} = \frac{E_{ij}}{E} \tag{3}$$

where E_{ij} is the total number of edges within the network from the source group i to the target group to j. The sum of these edges represents a proxy for the direct flow of resources between different groups. E is the total number of edges within the network. Within this configuration, the expectation w is taken as $P_{ij} \approx w_{ij}$. Therefore, in a similar manner, the probability of encountering an edge starting from a node of degree i within the distribution of the total population of E edges is:

$$P_i^{(\text{out})} \approx W_i = \frac{\sum_{x=1}^{k_{max}} E_{ix}}{E}, \quad P_i^{(\text{in})} \approx W_i = \frac{\sum_{x=1}^{k_{max}} E_{xi}}{E}. \tag{4}$$

The source $H^{(\text{out})}$, target $H^{(\text{in})}$, and total H entropies for both real and randomised networks in Japan and each of its 47 prefectures in isolation, are calculated in accordance with the classic Shannon construct:

$$H^{(\text{out})} = -\sum_i P_i^{(\text{out})} log_2 P_i^{(\text{out})}, \quad H^{(\text{in})} = -\sum_i P_j^{(\text{in})} log_2 P_j^{(\text{in})}, \quad H = H^{(\text{out})} + H^{(\text{in})}. \tag{5}$$

The logarithm base of two intends to represent the discrete and binary nature of undirected and unweighted edges, i.e., either two selected companies transact to each other or not.

In a similar manner, the joint entropy for two groups i trading with j, is given by:

$$J = -\sum_{i,j} P_{ij} log_2 P_{ij} \tag{6}$$

with the corresponding mutual information is given by:

$$I = -\sum_{i,j} P_{ij} log_2 \frac{P_{ij}}{P_i^{(\text{out})} P_j^{(\text{in})}} \tag{7}$$

and each pointwise contribution to the mutual information being:

$$I_{(i,j)} = -P_{ij} log_2 \frac{P_{ij}}{P_i^{(\text{out})} P_j^{(\text{in})}}. \tag{8}$$

5.2. Network Randomisation and Rewiring Process

Here, we make use of the symbol ($\tilde{}$) to denote quantities and corresponding outputs from Equations (2)–(8) associated with the randomised and rewired network. The process consists in generating a directional edge between two companies, namely the source and target nodes, for each step until the total number of edges $\tilde{E} = E$ is achieved. A constraint is placed whereby the total degree of each node within the population is maintained so that $\tilde{k_c} = k_c$ and $\tilde{S_k} = S_k$. In this manner, the degree distribution is preserved but the degree correlation of nodes is effectively and mostly randomised. We note, however, that the process as a whole is not totally random (i.e., without any form of correlations) since the node population constraint leads to slight distortions on the probability of edge selections, since the condition:

$$NP_\alpha \alpha = NP_\beta \beta = ... = NP_\omega \omega, \text{ where } \begin{cases} \text{N is the total number of nodes} \\ \{\alpha, \beta, ..., \omega\} \text{ is the set of the total degree of the nodes} \\ \qquad \text{within the population} \\ \{P_\alpha, P_\beta, ..., P_\omega\}, \text{ the related set of the probability of} \\ \qquad \text{nodes for a given total degree} \end{cases} \tag{9}$$

is not satisfied for every and any power law distribution represented by $P_k \sim k^{-\gamma}$, where $\gamma \neq 1$.

Therefore, we define the observed structural mutual information SMI as being the mutual information for the rewired network $\tilde{I}$ (i.e., $SMI = \tilde{I}$) computed in accordance with Equations (3)–(8) above with the equivalent quantities of the rewired network.

Furthermore, we also estimate the probabilities from the outcomes of $\rho = 1000$ realisations in the following way:

$$P_{ij} = \frac{\sum_{r=1}^{R} E_{ij}^{r}}{E * \rho}, \quad \text{where } E_{ij}^{r} \text{ is the value of } E_{ij} \text{ in the } \rho^{th} \text{ realisation.} \tag{10}$$

As shown in (c.3) within Figure 5, the degree correlation and related mutual information $I_{r^{th}}$ in this circumstance tends to zero as expected. The comparison between finite statistic and continuum realisation methods illustrates the effect of applying the analysis of a single realisation only to a real world network.

5.3. Methods Underpinning the Complexity Framework

As previously described within the Introduction section, the framework draws parallels between the allometric scaling of metabolic rates and the mutual information obtained from the method applied to the network. This is done in context of a dataset underpinning a real financial and economic system.

5.3.1. Structural Mutual Information, SMI

As described within the Introduction, the structural mutual information SMI is intended to capture the basic quantities held by the network simply as a result of the power law like degree distribution of companies, and their related sizes, through a method inspired by the allometric scaling in ecological systems.

Therefore, our method adapts the allometric scaling Equation (1) to the context of our research, specifically a trade network between companies in different territories. We regard the average degree of each network, namely the total number of trade links E (i.e., the total money flows) divided by the total number of companies $|\mathcal{C}|$ to be equivalent to the supply rate, and therefore $R \propto E/|\mathcal{C}|$. In a similar manner, the variability affecting body size can be represented respectively by the ratio between the diversity of species (i.e., total number of groups) and the number of links within the network, and therefore $Q_1 \propto S/E$. Lastly, we take the equivalent of the body size quantity of a prefecture to be proportional to the largest total degree (i.e., the sum of a node's in and out degrees) in which in turn it is proportional to the number of groups, and therefore $M \propto k_{max} \propto S$. By substituting these elements into Equation (1), using a single proportional constant λ, and adding a minimum floor parameter, we obtain:

$$\widehat{SMI} = \frac{\lambda}{|\mathcal{C}|} S^{7/4} + \tau \tag{11}$$

where $\lambda \sim 0.6$ and $\tau \sim 0.0275$ are empirically derived from the data, and $b \sim 3/4$ is also corroborated by the data.

5.3.2. Total Mutual Information I

As described within the Introduction, the total mutual information I can be regarded a multiplier to the core, structural mutual information SMI. Therefore, it can be mathematically represented as:

$$\widehat{I} = \kappa \widehat{SMI} + (1 - \kappa)\tau = \frac{\kappa\lambda}{|\mathcal{C}|} S^{7/4} + \tau \tag{12}$$

where it is empirically found that $\kappa \sim 7/4$.

Author Contributions: E.V. wrote the manuscript. E.V. and H.G. contributed equally in performing experiments, testing hypothesis, and interpreting results. Y.K. performed initial experiments and hypothesis. M.T., H.T., and H.J.J. provided advice on the method analysis of results. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript

Funding: This research received no external funding.

Acknowledgments: The authors are highly thankful to Teikoku Databank Ltd. for providing both access to its databases as well as the financial support.

Conflicts of Interest: The authors declare no conflict of interest affecting this research.

Abbreviations

The following abbreviations are used in this manuscript:

SMI	Structural Mutual Information
I	Total Mutual Information
H	Entropy
J	Total Joint Entropy
GDP	Gross Domestic Product

References

1. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Techn. J.* **1948**, *27*, 379–423. [CrossRef]
2. von Neumann, J. *Papers of John Von Neumann on Computing and Computer Theory*; MIT Press: Cambridge, MA, USA, 1987.
3. von Neumann, J. *Collected works / Vol. V, Design of Computers, Theory of Automata and Numerical Analysis*; Pergamon Press: Oxford, UK, 1963.
4. Dehmer, M.; Emmert-Streib, F.; Mehler, A. The Central Role of Information Theory in Ecology. In *Towards an Information Theory of Complex Networks: Statistical Methods and Applications*, 1st ed.; Birkhäuser Boston: Boston, MA, USA, 2011; pp. 153–167.
5. Ulanowicz, R.E. Reckoning the nonexistent: Putting the science right. *Ecol. Model.* **2014**, *293*, 22–30. [CrossRef]
6. West, G.B. A general model for the origin allometric scaling laws in biology. *Science* **1997**, *276*, 122–126. [CrossRef] [PubMed]
7. Bettencourt, L.M.A.; Lobo, J.; Helbing, D.; Kühnert, C.; West, G.B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7301. [CrossRef] [PubMed]
8. West, G. B. Santa Fe Institute studies in the sciences of complexity. In *Scaling in Biology*; Oxford University Press: Oxford, UK, 2000.
9. Goto, H.; Viegas, E.; Jensen, H.J.; Takayasu, H.; Takayasu, M. Appearance of Unstable Monopoly State Caused by Selective and Concentrative Mergers in Business Networks. *Sci. Rep.* **2017**, *7*, 5064. doi:10.1038/s41598-017-05362-5. [CrossRef] [PubMed]
10. Goto, H.; Viegas, E.; Jensen, H.; Takayasu, H.; Takayasu, M. Smoluchowski Equation for Networks: Merger Induced Intermittent Giant Node Formation and Degree Gap. *J. Stat. Phys.* **2018**, *172*, 1086–1100. [CrossRef]
11. Takayasu, M.; Sameshima, S.; Watanabe, H.; Ohnishi, T.; Iyatomi, H.; Iino, T.; Kobayashi, Y.; Kamehama, K.; Ikeda, Y.; Takayasu, H.; et al. *Massive Economics Data Analysis by Econophysics Method-The Case of Companies' Network Structure*; Annual Report of the Earth Simulator Center; Earth Simulator Center: Yokohama, Japan, 2008; pp. 237–241.
12. Viegas, E.; Goto, H.; Takayasu, H.; Takayasu, M.; Jensen, H. Assembling real networks from synthetic and unstructured subsets:the corporate reporting case. *Sci. Rep.* **2019**, *9*, 1–10. [CrossRef] [PubMed]
13. Watanabe, H.; Takayasu, H.; Takayasu, M. Relations between allometric scalings and fluctuations in complex systems: The case of Japanese firms. *Physica A* **2013**, *392*, 741–756. [CrossRef]
14. Huang, J.; Ulanowicz, R. Ecological Network Analysis for Economic Systems: Growth and Development and Implications for Sustainable Development. *PLoS ONE* **2014**, *9*, e100923. [CrossRef] [PubMed]
15. Arkolakis, C. *Market Penetration Costs and the New Consumers Margin in International Trade*; Working Paper No. 14214; NBER: Washington, DC, USA, Augest 2008.
16. Price, D.D.S. A general theory of bibliometric and other cumulative advantage processes. *J. Assoc. Inf. Sci. Technol.* **1976**, *27*, 292–306. [CrossRef]
17. Barabási, A.L.; Albert, R. Emergence of Scaling in Random Networks. *Science* **1999**, *286*, 509–512. [CrossRef] [PubMed]

18. Sanchirico, A.; Fiorentino, M. Scale-free networks as entropy competition. *Phys. Rev. E* **2008**, *78*, 046114. [CrossRef] [PubMed]
19. Sutton, J. In *Sunk Costs and Market Structure : Price Competition, Advertising, and the Evolution of Concentration*; MIT Press: Cambridge, MA, USA, 1991.
20. Viegas, E.; Cockburn, S.P.; Jensen, H.J.; West, G.B. The dynamics of mergers and acquisitions: ancestry as the seminal determinant. *Proc. Math. Phys. Eng. Sci.* **2014**, *470*, 20140370. [CrossRef] [PubMed]
21. Aitsahlia, F.; Yoon, J.H. Information stages in efficient markets. *J. Bank. Finance* **2016**, *69*, 84–94. [CrossRef]
22. Ito, M.; Noda, A.; Wada, T. The evolution of stock market efficiency in the US: A non-Bayesian time-varying model approach. *Appl. Econ.* **2016**, *48*, 621–635. [CrossRef]
23. Lee, S.H.; Kim, P.J.; Jeong, H. Statistical properties of sampled networks. *Phys. Rev. E* **2006**, *73*, 016102. doi:10.1103/PhysRevE.73.016102. [CrossRef] [PubMed]
24. West, Geoffrey B.A. *Scale : The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*; Penguin Press: New York, NY. USA, 2017.
25. Boettke, P.J. What Happened to "Efficient Markets"? *Indep. Rev.* **2010**, *14*, 363–375.

entropy

MDPI

Article

New Measure of Economic Development Based on the Four-Colour Theorem

Aleksander Jakimowicz [1,*] and Daniel Rzeczkowski [2]

[1] Department of World Economy, Institute of Economics, Polish Academy of Sciences, Palace of Culture and Science, 1 Defilad Sq., 00-901 Warsaw, Poland

[2] Department of Market and Consumption, Faculty of Economic Sciences, University of Warmia and Mazury in Olsztyn, 1/327 Cieszyński Sq., 10-720 Olsztyn, Poland; daniel.rzeczkowski@uwm.edu.pl

* Correspondence: ajakimowicz@inepan.waw.pl

Received: 18 November 2020; Accepted: 28 December 2020; Published: 31 December 2020

Abstract: The location quotient is one of the basic quantitative tools for identifying the regional poles and the turnpikes of economic growth in spatial economy. The disadvantage of this traditional measure is the limited scope of economic information contained in it. The new measure of economic development proposed in the article encompasses a complex spectrum of phenomena in one number, as it takes into account the influence of the public administration sector, as well as top technology in the form of ICT and its practical business models. It also takes into account the digital prosumption and the platforms for participation. The participation platforms in the public administration sector are the websites of municipal public administration offices. A cluster analysis was used to distinguish four quality classes of these websites. These classes were assigned four different colours, which were then used to draw up a map of the selected province. Each municipality is marked with a colour that corresponds to the quality class of the website of the state administration office operating on its territory. The colour system resulting from the four-colour theorem and the corresponding dual graph play the role of a reference system in relation to each empirical colour distribution and another dual graph related to it. The measure of the economic development of a region is the degree of reduction of the dual graph corresponding to the empirical distribution of colours, which identifies the actual growth poles and determines the routes of growth. The presented indicator better and more precisely identifies poles and routes of economic growth than the traditional location quotient.

Keywords: measure of economic development; websites; public administration sector; municipality; four-colour theorem; prosumption; platforms for participation; location quotient; dual graph; Euler characteristic

1. Introduction

Development of new technologies has an enormous impact on economic growth. John Hicks in the classic work *Capital and Growth* notes that each production technique has a corresponding rate of return, which is fully determined only by this technique [1]. If it is possible to choose a production technique, then there should be a technique with the highest rate of return, achievable along a growth equilibrium path, within the technology. Such a technique is labelled as top technique. It corresponds to the maximum rate of sustainable growth in the known John von Neumann model [2]. Contemporarily, this role is played by the digital technique, which transformed the traditional economy into the digital economy. Therefore, its inclusion is necessary in all studies on economic growth and development.

The impact of digital computing technologies on the economy is at the centre of the dynamically developing research trend called wikinomics [3,4]. It was established that there was a significant change in business rules, which led to the creation of completely new business models. The most important of them include prosumption and participation platforms. The first implies active involvement of consumers (recipients) who are a non-standard source of innovation and creativity, in the design and manufacture of new products (services) [5,6]. Participation platforms include products and technology infrastructure made available to large communities of partners. Access to these products and technologies enables the creation of new values and initiates innovative ventures. Prosumption is increasingly becoming digital and participation platforms are a medium for the flow of economic information.

Participation platforms take the form of websites that help enterprises collaborate with business partners, clients and the wider economic environment. They enable practical implementation of such principles of wikonomics as openness, peering, sharing, and acting globally. Participation platforms in the private sector have become the engine of economic development, as they have contributed to the increase in production capacity without the need of incurring additional fixed costs. In addition, the creation of open participation platforms greatly facilitates innovation activities. Web services mashups are becoming increasingly popular, as they combine various ready-made services or applications, made available by other websites, into a new whole which has improved quality and functionality.

The possibilities of using participation platforms are virtually limitless, they can be the basis for development for many different products or services, literally everything that can be controlled by software. Therefore, participation platforms are not limited to commercial applications, but can also perform many useful functions in the public administration sector. There are large databases in this sector that are hardly used. They could become a source of many new services stimulating economic growth and development. Thus, societies do not make full use of all the opportunities resulting from technical progress to improve their standard of living. Websites in public administration should take the form of platforms for grassroots action and include platforms for public disclosure and platforms for neighbourhood knowledge [3] (pp. 199–205).

The purpose of the article is to develop a new, synthetic measure of economic growth in the regional perspective, which would include in one number abundant information relevant from the point of view of local entrepreneurship. The object of research is the Warmia and Mazury Province, consisting of 116 municipalities, located in north-eastern Poland. This province is one of the least developed in Poland [7,8], but in recent years many economic initiatives have been undertaken in order to accelerate the development and reduce the distance to the rest of the country [9,10]. The proposed measure meets Hicks's condition for top technique because it is based on wikinomics business models in the form of prosumption and participation platforms which are characteristic for digital economy. In addition, this measure considers the spatial diversity of municipalities, the barriers they encounter in mutual economic cooperation, and it is also useful for locating regional poles and turnpikes of growth. As it is based on digital technologies, it is not limited to one industry, effectively covers all types of economic activities that can be controlled by software. This is an advantage over classic measures of local growth and development like the location quotient.

Currently, websites of municipal public administration offices—basic units of the local government in Poland—are much less developed than platforms for commerce, which means that not all factors of regional economic growth are available to municipal communities. From the point of view of wikinomics, these websites should be treated as the basic sources of entrepreneurship, as their role should be to initiate and develop business activities in municipalities. Empirical research confirms the impact of the size of local government administration at a municipal level on entrepreneurship, but this impact is not clear. It has been observed that the increase in the size of administration affects entrepreneurship in a negative way [11]. This justifies the need to observe this issue from the digital technique point of view. Therefore, the

starting point of the study was to assess the functionality of websites of municipal public administration offices. The evaluation criterion was the degree of their fulfilment of the role of wikinomics platforms for participation. By means of cluster analysis, these websites were grouped into four distinct quality classes: low, medium, high and very high. Four different colours were assigned to these classes, which were used to map the evaluated province. Each municipality was marked with a colour corresponding to the quality class of its website. As the map of the studied region is a normal map, it meets the conditions of the four-colour theorem. In this way, two types of colour systems can be distinguished. The first results from the application of the theorem itself, while the second includes empirical colour distributions reflecting the quality classes of websites in specific periods. Based on each map, a corresponding dual graph was prepared. Capitals of all municipalities were marked on the map of the province, and then the capitals of the neighbouring municipalities were connected by roads crossing their common borders. The dual graph meeting the conditions of the four-colour theorem is a reference frame for the research. On empirical maps, the edges of the graph, indicate barriers to cooperation between individual municipalities. The new measure of economic growth is determined by comparing the number of edges of empirical graphs with the reference graph. For this purpose, the generalisation of the Euler characteristic made by Augustin-Louis Cauchy was used. In addition, the article discusses the problem of dividing websites into four quality classes in the light of René Thom's classification theorem. The impact of the complexity of geographical lines forming boundaries between municipalities on possibilities of application of the four-colour theorem is also being studied. It can be observed that municipalities in the Warmia and Mazury Province are fractals, similar to natural structures such as coastlines or clouds and the four-colour theorem can be used to describe infinitely complex fractal structures.

This paper precisely defines all the methods used in the research to accurately identify the sources of complexity emerging in the examined system and to facilitate their economic description. The research belongs to complexity economics, which implies that it is also an element of a wider discipline—the science of complexity. Advances in information technology have contributed to significant changes in economic life and have resulted in the need to modify classic economic laws and to discover new ones. This study combines several mathematical, physical and economic methods applied to obtain an overall measure of the economic performance of any region consisting of sub-units, which would allow determining growth and development based on the rules of wikinomics and, in the same time, take into account top technology as understood by Hicks. The four-colour theorem plays an essential role in the measure of development proposed in the paper. Number four is not a random number, but an attractor of the economic system, as indicated by the catastrophe theory formulated by René Thom. The subject of the research is the dynamics of objects in three-dimensional physical space–time, which consists of two spatial dimensions and one temporal dimension. These objects are dual graphs based on five maps of the examined province, one of which was coloured according to the four-colour theorem principles and provides a system of reference, while the other four are of an empirical nature and present the distribution of website quality classes of all municipalities in the four periods under consideration. All events taking place in this space–time continuum consist of the transformations of two-dimensional empirical dual graphs over time. What does not change is only the dual graph corresponding to the map of the province, coloured according to the four-colour theorem principles. Following the classification theorem, which underlies catastrophe theory, in three-dimensional physical space–time there are only five elementary catastrophes, which represent all possible manifestations of dynamic phenomena. Therefore, all empirical distributions of the municipality websites quality classes must coincide with four elementary catastrophes, while the fifth catastrophe represents the dynamics of the environment, i.e., the rest of the national economy since the examined province is an open system. Consequently, the empirical division of municipality websites into four quality classes, resulting from the use of the *k*-means algorithm, is correct in a topological sense. In addition, the province of Warmia and Mazury—comprising one hundred and sixteen municipalities—is a system so

diverse that it provides an excellent research subject ensuring an unbiased test verifying the correctness of the presented method for economic growth measurement. In other words, this province is complex enough to allow empirical demonstration of the attractor in the form of the number four. It is also possible to note significant mathematical links between the catastrophe theory and the four-colour theorem. They lead to the conclusion that the classification theorem provides topological proof of the four-colour theorem. This fact has not been recognised in science so far and we are probably the first researchers to discover it.

If two adjacent municipalities have websites marked with different colours, i.e., belonging to different quality classes, then the edge crossing the common border and connecting their capitals symbolises the absence of significant economic cooperation between them. The existence of cooperation would mean the unification of websites and their transfer to high or very high quality classes. On the other hand, if neighbouring municipalities have websites marked with the same colours and therefore belong to the same quality class, the absence of edges represents economic cooperation between them. The four-colour theorem is the reference system for the research since the resulting dual graph represents the total absence of growth poles and cooperation between municipalities. In this case, every two adjacent municipalities are connected by edges. For Warmia and Mazury, the number of these edges determined based on Cauchy's theorem is 277. This value always appears in the denominator of Equation (16), which defines the proposed measure of economic development, while the numerator contains the changing numbers of the edges of the empirical dual graphs determined for each year, i.e., 63, 49, 106 and 58. Thus, the defined measure is an absolute measure, which makes it more objective compared to relative measures commonly used in economics (e.g., the method for calculating real national income based on price deflators).

The methodological procedure used to develop a new measure of economic development is complex as it is based on a combination of many different approaches. Its individual stages can be summarised in the following points:

1. Selection of the province for which the measure is to be designated. Such a region should be sufficiently diversified economically and it should consist of a large number of municipalities.
2. Development of qualitative criteria for the evaluation of municipal websites on the basis of wikinomics guidelines.
3. Binary evaluation of websites, which consists in assigning each website the value of one for meeting a given criterion. If the specified criterion is not met, the assigned value is zero.
4. Calculation of the quality indicator for each evaluated website according to the sum of the values received for the fulfilled criteria. The maximum number of points signifies that all qualitative criteria are met.
5. Division of websites into homogeneous subsets called quality classes with the use of the *k*-means clustering method. According to the classification theorem by René Thom, the attractor of this division is the number four, provided the region is sufficiently differentiated.
6. Assignment of colour to each quality class and preparation of an empirical map of the province with the inclusion of the colour coding of the municipalities according to the quality classes of their websites.
7. Preparation of an absolute reference system in the form of a map of the province coloured according to the four-colour theorem.
8. Preparation of the dual graphs for the empirical maps and reference map. Comparison of the number of edges of the empirical dual graph with the number of edges of the reference dual graph is the essence of the proposed measure of economic development. The edges of the graphs represent the lack of economic cooperation between neighbouring municipalities.

The evaluation of municipalities on the basis of the quality of the websites is not the only way to determine the proposed measure of economic development. This measure may be of a more general

nature. It could be based on a different socio–geographic–economic indicator that would apply to all municipalities in a given province and would have four elements in the internal structure.

2. The Location Quotient

The location quotient is one of the basic quantitative tools used to identify regional poles and turnpikes of growth in spatial economy. It is used to compare two spatial units described by percentage indicators, one of which refers to the characteristics of a given region (municipality, province), while the other refers to a higher spatial unit (province or country). The location quotient measures the region's industrial specialisation in relation to a larger geographic unit [12–14]. Any base that is significant for the problem and region under study can be used. For this reason, the location quotient can find application even in sciences such as criminology [15]. Sometimes bases such as earnings or GDP by metropolitan area are used, however employment is most often accepted as the most relevant base. Employment in various sectors of industry is distinguished, moreover trade and services are also included. The location quotient (LQ_{i_n}) is given by [16,17]:

$$LQ_{i_n} = \frac{E_{i_n}}{E_{T_n}} \Big/ \frac{\sum_{n=1}^{N} E_{i_n}}{\sum_{n=1}^{N} E_{T_n}}, \tag{1}$$

where: n—small area under study, N—total number of areas, E_i—employment in industry i, and E_T—total employment in all industries.

The location quotient is used in economic research as in [18] (pp. 173–176):

(1) a measure of relative concentration or location advantage/disadvantage of specific industries in the regional economy;
(2) a proxy for input–output coefficients used to assess regional or subregional multiplier effects.

As a measure of relative concentration, it demonstrates the strengths and weaknesses of various industries in the region. If the location quotient for the i-th industry is greater than one, then the region is assumed to be exporting the products or services of that industry. Then, in a given area, relatively more employees are employed in a specified sector of industry compared to a larger area which implies that the given area produces more goods or services than are consumed by the area's residents. The excess can therefore be exported. When the location quotient is less than or equal to unity, the industry does not export outside the region, because it is not self-sustainable. In other words, Equation (1) shows whether the region is an exporter or importer of goods produced by the i-th industry. For $LQ_{i_n} > 1$ the export activity (X_{i_n}) of regional employment in i-th industry (E_{i_n}) is calculated as follows [16]:

$$X_{i_n} = [1 - (1/LQ_{i_n})]\, E_{i_n}. \tag{2}$$

Multiplier effects appear due to the fact that a change in regional export activity leads to a total change in the regional economy. The economic base multiplier is greater than one by the proportion of local to export activity. It is a special case of the traditional Keynesian multipliers [19]. A total change in the regional economy is a product of the multiplier and the export change. The economic base multiplier can be determined from the location quotient as follows [16]:

$$X_{i_n} = \left(\frac{E_{i_n}}{E_{T_n}} - \frac{\sum_{n=1}^{N} E_{i_n}}{\sum_{n=1}^{N} E_{T_n}} \right) E_{T_n}. \tag{3}$$

Typically, the economic base multiplier is calculated by determining the export employment on the basis of Equation (2) for all industries for which $LQ_{i_n} > 1$, by summing the export employment of all those industries, and dividing the sum by total employment. Formula (3) shows that this traditional calculation

process can be simplified. The calculation of export employment involves subtracting the share of the nation's economy accounted for by *i*-th industry from the share of the region's economy accounted for by *i*-th industry, and the resulting difference is multiplied by the region's total employment.

This article examines municipalities as the basic spatial units. The superior unit is a province and, therefore, interpretation of the Formula (1) is affected. If we understand small areas as municipalities in the Warmia and Mazury Province and tourism (i) is assumed to be a given branch of the economy, then the individual symbols should be understood as: E_{i_n}—employment in tourism in n-th municipality, E_{T_n}—total employment in n-th municipality, $\sum_{n=1}^{N} E_{i_n}$—employment in tourism in the entire province, $\sum_{n=1}^{N} E_{T_n}$—total employment in a given province. According to the above approach, the measure expressed by Formula (1) can be used to compare relative concentrations of employment in tourism of two municipalities.

The main disadvantage of the location quotient is the limited amount of economic information it conveys. It is designated only for one economic base and does not match modern wikinomics business models, such as digital prosumption and participation platforms. Nowadays, a new, synthetic measure of economic activity is needed, which would be appropriate for the digital economy. Such a measure should be based on teleinformatics and it needs to consider the regional diversity of municipalities, their mutual cooperation and it ought to allow the identification of regional poles and turnpikes of economic growth. It is also advisable that it not be limited to one branch of industry but includes all branches if possible. It is also necessary to consider municipal public administration offices, which—as the smallest units of the local government—are the initiators of economic activity in their subordinate areas. In this article we propose a measure that brings together all these key issues into one number.

3. The Four-Colour Theorem

Usually, the four-colour theorem is presented as follows: regions on each map drawn on the plane can be marked with only four colours in such a way that each two adjacent regions have different colours [20–22]. Adjacency signifies that the areas are bordering each other along a line. Cases of bordering in one point or even at a finite number of points are omitted. Region is understood as a fragment of the plane that is connected, and therefore consisting of only one area.

For normal maps, there is an upper limit on the number of neighbours that regions can border with. As Alfred Bray Kempe noted, any normal map plotted on a plane meets the following equation [23]:

$$4p_2 + 3p_3 + 2p_4 + p_5 - p_7 - 2p_8 - 3p_9 - \cdots - (N-6)\,p_N = 12, \tag{4}$$

where p_n represents the number of regions on the map that have exactly n neighbours, and N represents the largest number of neighbours any region can have. The formula starts with p_2, because cases when $n = 0$ *i* $n = 1$ are omitted. Normal maps do not have enclaves or islands. Each p_n is either zero or positive, which is true for $n < 6$. Equation (4) demonstrates that, at least one of the numbers p_2, p_3, p_4 or p_5 has to be positive, so that the left side of the equation is positive and corresponds to the right side. Therefore, one of the regions must have either two or three or four or five neighbours. There cannot be a normal map on a plane where each region would have six or more neighbours.

Each non-normal map can be assigned a normal map that can be drawn with at least the same number of colours. The truthfulness of the four-colour theorem for normal maps entails its truthfulness for all maps. The map of the United States is a good example as it is not a normal map because it contains the quadripoint, i.e., a single point at which the boundaries of the four states meet: Colorado, Utah, Arizona, and New Mexico. Yet, four colours are enough to colour it [21] (p. 5).

Two concepts appear in the context of the four-colour theorem: the unavoidable set and the reducible configurations. A set is called unavoidable if it contains a configuration set consisting of a region with

two neighbours, a region with three neighbours, a region with four neighbours, or a region with five neighbours. The name of the unavoidable set results from the fact that each normal map must contain at least one of the four configurations listed. The reducible configuration occurs when it cannot appear on the minimal five-chromatic normal map, and thus the smallest normal map requiring the use of five colours. It is enough to find the unavoidable set of reducible configurations to prove the four-colour theorem.

The problem of four colours was formulated in 1852 by Francis Guthrie, who was then a student of mathematics at University College in London. During the colouring of the map of England, he noticed that four colours are sufficient to meet the condition that each two neighbouring counties have different colours. However, proving the four-color theorem turned out to be very difficult and, despite many attempts, it took mathematicians 125 years. The correct proof was announced in 1976 by Kenneth Appel and Wolfgang Haken, who based it on the construction of the unavoidable set of 1936 reducible configurations. Later it occurred that the number of these configurations was higher and finally amounted to 1482 [24]. The most interesting fact was that the proof was obtained on the basis of computer calculations. In the following years the proof was improved, but always required the use of a computer [25–29]. Validation of this proof requires a new type of reasoning, quite different from the one previously used in mathematics, which, perhaps, goes beyond the capabilities of the human mind. It sparked a discussion about the nature of a mathematical proof [30]. This problem even initiated some philosophical discussions around the four-colour theorem, which indicated that it was not an a priori truth, as required by classical mathematical tradition, but an a posteriori truth [31–33]. The map colouring problem seems to be a task worth a quantum computer [34].

In recent years, simpler proofs of the four-colour theorem have emerged, as exemplified by the proposals of the Polish mathematician Antoni Smoluk or John W. Oller, Jr. [35,36], but—so far—these works have not been the subject of wider scientific discussion. It has even been suggested that some of these proofs do not require the use of a computer as they can be fully verified by human mathematicians [37]. It is worth noting in this context that the complexity of the four-colour theorem corresponds to that of the Riemann hypothesis and is almost four times greater than the complexity of Fermat's last theorem [38].

4. The Dual Graph

In this part, we will discuss the basic concepts of graph theory that will be needed in further research. A planar graph is understood as such a graph that can be drawn on the plane in such a way that no two edges intersect geometrically, except for the vertex, where they combine. As such a graph is represented on a plane, it is often called a plane graph. A connected graph is a graph in one piece, so any two vertices are connected by a path. For a planar graph G located on a plane, another graph G^*, called the (geometric) dual of G can be specified. Its construction takes place in two stages [39,40]:

(1) inside each face F of graph G we choose the point V^*; all these points will create the vertices of the new G^* graph;
(2) corresponding to each edge E of the graph G we draw the line E^* which intersects only the edge E (and no other edge of the graph G) and connects the vertices V^* located on the faces F adjoining to E; these lines form the edges of the dual graph G^*.

The concept of duality has been known for a long time, it was noticed in antiquity that the dual of a cube is an octahedron, and that the dual of a dodecahedron is an icosahedron. The fact that the graph G is both plane and connected demonstrates, that a dual graph G^* has the same properties. Moreover if a plane connected graph G has N vertices, E edges and F faces, and its geometric dual G^* has N^* vertices, E^* edges and F^* faces, then between the numbers of vertices, edges and faces of G and G^* we observe the following relations: $N^* = F$, $E^* = E$, and $F^* = N$.

A map on a plane is a type of planar graph, so such a map also has a corresponding dual graph. The political map of South America was selected in order to present the idea of constructing a dual graph. The map was coloured according to the conditions of the four-colour theorem. The map includes 13 countries, within which there are 13 capitals. This map is a type of planar graph, in which the faces represent individual countries, the edges are inter-state borders, and the vertices form the points where the borders meet. If we connect the capitals of neighbouring countries with roads passing through their common borders, then we receive a dual graph of the original map. Capitals of the countries denote the vertices of the dual graph, the roads connecting the capitals of neighbouring countries are its edges, and the triangles resulting from the connection of capitals are faces. Figure 1 presents both the political map of South America and the corresponding dual graph. It should be emphasised that the edges of the dual graph are not real routes connecting national capitals, but conventional connections in the form of straight line segments.

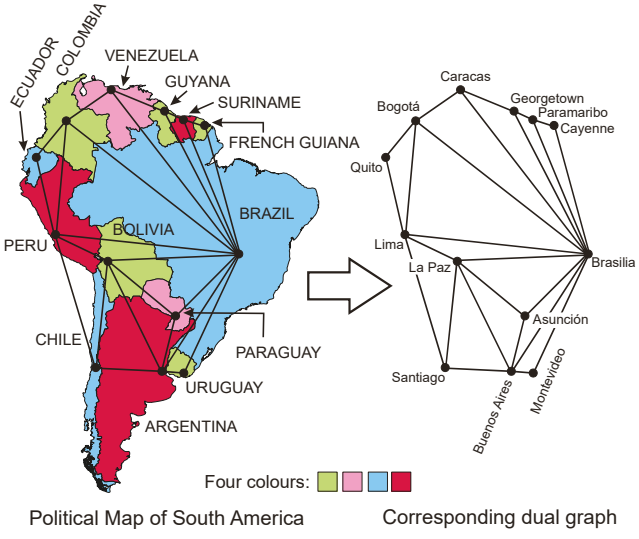


Figure 1. The political map of South America and the corresponding dual graph.

5. The Model Graph Being the Reference System for Research

In spite of appearances the four-colour theorem does not have significant applications in cartography and geography. In fact, it can be stated that its role in science is twofold. First of all, it is an extremely important and unique mathematical achievement, especially with regard to the proof of the theorem, which initiated a philosophical discussion concerning cognitive capabilities of the human mind. So far, there is no chance that a human being could personally check the reducibility of all of the configurations in the unavoidable set. Secondly, from the point of view of other sciences, this theorem is still only a mathematical curiosity, because non-mathematical sciences lack its significant applications, i.e., those that would solve a major problem. So far, only a few attempts have been made to apply this theorem in various fields of science, but these endeavours have not led to significant progress [41–46]. We decided to break this deadlock. In 2018, we were able to prove that the four-colour theorem can be a useful tool for identifying spatial regional poles and turnpikes of economic growth [47]. Presently, we wish to proceed a

step further and present that the theorem can be used to develop tools for measuring regional growth in the digital economy.

Dual graphs play a major role in the proposed method for identifying spatial regional poles and turnpikes of economic growth. The conversion of the map into a dual graph allows the study of a very significant economic problem, which is the cooperation of municipalities. Wikinomics participation platforms, i.e., websites of municipal public administration offices, are the basis of cooperation. On one hand, these websites allow municipalities, as the basic units of the local government, to support entrepreneurship in their subordinate areas, while on the other they enable the manifestation of creativity and innovation of the inhabitants of these municipalities in the form of digital prosumption. That is why the quality of websites plays a key role in the modern digital economy.

Two reference systems are used to assess economic growth in the Warmia and Mazury Province. The first is the dual graph, called the model graph, which corresponds to the map coloured according to the conditions of the four-colour theorem and is synonymous with a complete lack of growth poles. The second occurs when the websites of all municipalities belong to the highest quality class. In such a case, the dual graph is reduced to one vertex, and the whole province becomes one big growth pole and dominates the national economy. These two reference systems, although they have little chance of occurring in the real world, form a natural framework defining the scope of research. Everything that occurs in reality must be between these two cases. However, they point to the essence of the presented method used to identify spatial regional poles and turnpikes of economic growth. The method involves the reduction in dual graphs corresponding to real maps and comparing them with the dual graph corresponding to the reference map. Municipalities on real maps merge into clusters, which denote the quality classes of websites run by these basic local government units.

According to these guidelines, at the beginning the map of the province was coloured in such a way that each two neighbouring municipalities were marked with different colours. This allowed the plotting of the corresponding dual graph. Subsequently, the capitals of all municipalities were marked on the map of the province, and then the capitals of adjacent municipalities were connected by roads passing through their common borders. Figure 2 presents the map of the province coloured according to the conditions of the theorem and the corresponding dual graph. The capitals of municipalities form the vertices of this graph, while the roads are its edges. The dual graph is plotted in such a way that all its edges are sections of straight lines. These edges divide the plane into areas called faces. The original map—after eliminating cases of complete surrounding of one municipality by another municipality—is a normal map on which each vertex joins precisely three edges. All dual graph faces are triangles because the faces are connected to the vertices of the original map. Consequently, such a dual graph is called triangulation.

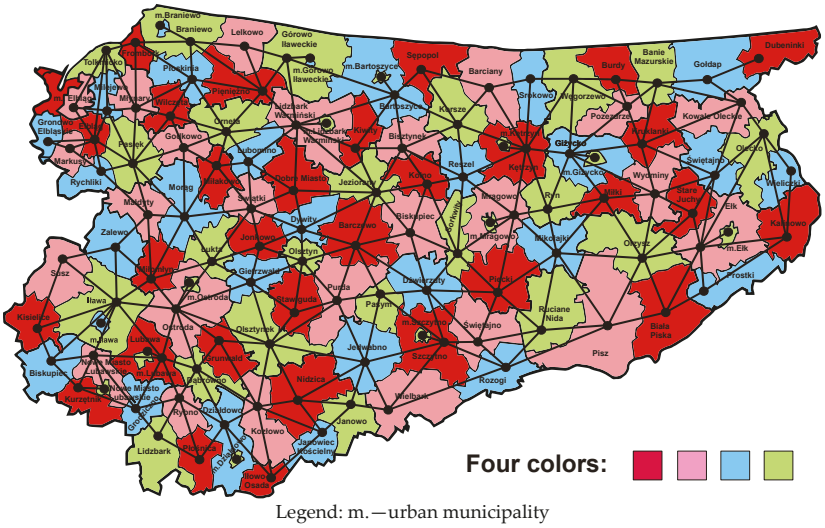


Figure 2. Map of the municipalities of the Warmia and Mazury Province coloured in accordance with the four-colour theorem and a corresponding dual graph. It is a frame of reference without growth poles.

In the dual graph, the number of edges ending in a given vertex is called the degree of the vertex and is equal to the number of country borders that on the original map correspond to that vertex. The circuit of the graph is called a path of edges, which has a beginning and an end in the same vertex, does not cross itself and divides the graph into two parts: its interior and its exterior. In the vocabulary of dual graphs, a configuration is understood as a part of triangulation, which consists of a set of vertices and all edges connecting them. The boundary circuit is called the ring of the configuration [48] (p. 166). Figure 2 shows that on the original map of the Warmia and Mazury Province there are thirty-nine external municipalities surrounding its internal part. After plotting the dual graph, they were replaced with the ring of the configuration, i.e., the boundary circuit, consisting of thirty-nine vertices and thirty-nine edges. The vertices of first degree, such as the municipality of Dubeninki, were omitted for obvious reasons. The configuration presented in Figure 2 in dual form is called the thirty-nine ring configuration, since its ring has thirty-nine vertices. It corresponds to the ring of thirty-nine municipalities that circle the original configuration.

6. Evaluation of Websites of Municipal Public Administration Offices

The main point of research was the assessment of websites of municipal public administration offices in the Warmia and Mazury Province in terms of their role as wikinomics participation platforms. The assessment consisted of sixteen points, which are presented in Table 1 [49,50]. The points were coded with symbols from A01 to A16. In order to obtain information about the quality of websites, a binary method was used, which consisted in assigning the value one to the website when the given criterion was met, or zero otherwise. Afterwards, all values assigned to individual websites were added together. Thus, the webpage functionality index included a number from the closed interval from 0 to 16.

Table 1. Criteria for the functionality of the municipal websites in 2009, 2012, 2015 and 2019.

Code	Website Functionality Criteria
A01	Website is updated on a regular basis
A02	The postal address of the office is included, directions are provided
A03	The office publishes chat lines and/or discussion lists for the citizens
A04	The structure of the office has been posted
A05	Current information is published on a regular basis
A06	There is a possibility to search for necessary information
A07	A calendar of posts is published
A08	The user can fill and send a form online
A09	Other than Polish language versions are available
A10	Website provides icons that help the user to use the website
A11	The website address of the office is intuitive
A12	Archive exists
A13	A map of the municipality is published
A14	Tourist attractions are indicated
A15	Link to "digital office" provided
A16	Link to ePUAP (Electronic Platform of Public Administration Services) provided

The next stage of research involved the use of the *k*-means algorithm to divide websites into homogeneous subsets representing their quality classes [51]. The term cluster analysis was first used in 1939 by Robert Tryon, who was a pioneer in the study of various data classification algorithms [52]. James B. MacQueen started fundamental studies on the *k*-means clustering method [53], although the idea itself came from the Polish–Jewish mathematician Hugo Steinhaus [54], who dealt with this problem in 1956. In the 1970s, John A. Hartigan and M.A. Wong made a significant contribution to the improvement of the *k*-means clustering method [55,56], by proposing a suitable algorithm, finding recognition even today [57]. The well-known standard algorithm was developed by Stuart P. Lloyd in 1957, but the article on this topic was not published in the scientific journal until 1982 [58]. The problem consists of the division of the data set $X = (x_1, x_2, \cdots, x_n)$ into a predetermined number of k clusters of greatest possible distinction. As a result of calculations, we obtain a set of k cluster centroids and an assignment of each point X to one cluster in such a way that the distances of all points belonging to a given cluster from the corresponding centroid are smaller than their distances from any other centroid. From the mathematical perspective, it is presented as follows:

$$\operatorname{argmin} \sum_{i=1}^{k} \sum_{x_n \in C_k} \|x_n - \mu_k\|^2, \tag{5}$$

where C_k and μ_k denote clusters and centroids, respectively. The division of the set of observations into $k \leq n$ cluster is based upon the minimisation of the within-cluster sum of squares, and thus variance.

It should be noted here that *k*-means clustering belongs to NP-hard problems [59–63]. In computational complexity theory NP (non-deterministic, polynomial time) it is the complexity class containing the set of decision problems that can be solved by a non-deterministic Turing machine in polynomial time [64] (p. 56). Most of the difficulties arising from the fact that *k*-means clustering belongs to the class NP-hard problems can be overcome by using an iterative method known as Lloyd's algorithm. It consists of an alternate performance of two operations: (1) when a set of centroids μ_k is determined, the clusters C_k are actualised by reducing—inside each cluster—the distance of points from the centroid; (2) based on the set of clusters, centroids which are the means of all points belonging to individual clusters are recalculated. These operations take the following form:

$$C_k = \{x_n : \|x_n - \mu_k\| \leq \text{all } \|x_n - \mu_l\|\}, \tag{6}$$

$$\mu_k = \frac{1}{C_k} \sum_{x_n \in C_k} x_n. \tag{7}$$

This procedure is continued until the assignments of clusters and centroids no longer change. The algorithm shows the convergence in few steps, but the solution can be in the form of a local minimum. In the worst case, Lloyd's algorithm needs $i = 2^{\Omega(\sqrt{n})}$ iterations, and thus its running time is superpolynomial [65]. The used big-Ω notation regards the asymptotic lower bounds, and thus the limit of the growth of the running time of the algorithms from below for large enough input sizes [66]. For the Lloyd algorithm, the time complexity varies from $\Omega(n)$ to $2^{\Omega(\sqrt{n})}$.

Table 2 presents the results of using the Lloyd's algorithm to divide the websites of municipal public administration offices into wikinomics quality classes. Four clusters representing websites of low quality, average quality, high quality and very high quality were naturally separated. On the basis of this division, it is possible to determine not only empirical dual graphs, the edges of which represent barriers to inter-municipal cooperation, but also complementary graphs, the vertices and edges of which define, respectively, digital growth poles and development axes [67]. The study was conducted four times, i.e., in 2009, 2012, 2015 and 2019 [47,50]. Table 2 shows that most municipalities have high and very high quality websites, which demonstrates that these websites should be treated as the seeds of wikinomics platforms for participation. After some transformations these websites could perform the role of platforms for grassroots action, which would include platforms for public disclosure and platforms for neighbourhood knowledge. In this way, information held by public administrations would contribute to the creation of new values and services that would benefit both residents and entrepreneurs. Such actions would certainly lead to the development of entrepreneurship and an increase in the regional economic growth rate and the level of prosperity. In addition, the high quality of websites would contribute to the development of digital prosumption, which would enable the use of the natural innovation and creativity of people. In the processes of prosumption, the difference between the producer and the consumer, the service provider and the recipient disappears, which allows the passive party, the consumer or the recipient, to participate in the design, creation and production of goods or services. However, new business models in the form of prosumption and participation platforms can only function well if the four basic principles of wikinomics are met: openness, peering, sharing, and acting globally. Under these conditions, it is necessary to give prosumer communities some control over the product or service. In this case, it involves co-creating some public administration services by the local government, residents and entrepreneurs.

Table 2. Four clusters representing the wikinomics categories of quality of the websites of the municipal administrative authorities in 2009, 2012, 2015, and 2019.

Characteristics	Low Quality	Average Quality	High Quality	Very High Quality
Year		**2009**		
Points (min–max)	0–4	5–9	10–11	12–16
Centroid	0	7.72	10.45	13.35
Number of municipalities	17	22	37	40
Year		**2012**		
Points (min–max)	0–6	7–10	11–13	14–16
Centroid	0	9	11.75	14.45
Number of municipalities	6	11	49	50
Year		**2015**		
Points (min–max)	0–6	7–11	12–13	14–16
Centroid	0.47	8.69	11.75	14.65
Number of municipalities	10	32	35	39
Year		**2019**		
Points (min–max)	0–1	8–11	12	13–15
Centroid	1	10.46	12.00	13.87
Number of municipalities	3	30	17	66

7. Determination of Growth Poles and Development Axes According to the Dual Graph Reduction Method

Empirical maps and the resulting dual graphs were created in a similar way to the reference map and the reference dual graph. The only difference is that each municipality received a colour corresponding to the quality class of its website run by the administration office operating in its area. After the map of the examined province was coloured in accordance with the quality classes of websites, it occurred that some neighbouring municipalities have the same colours. If these colours correspond to classes of high or very high quality, then it is assumed that the given municipalities form spatial poles or even regional axes of growth and development when there are more municipalities. Generally speaking, the same colour of neighbouring municipalities, i.e., having websites of the same quality, is interpreted as a form of cooperation between the municipalities. The empirical dual graphs of the studied province were created on the assumption that a set of neighbouring municipalities, marked with the same colour, is treated as one region. In this case, the contractual capitals of all regions or even individual municipalities are marked with points, provided that they are separated from the surrounding, and then connected with sections of straight lines passing through the common borders of these regions or municipalities. In this way a dual graph corresponding to the real map is created. It is reduced in comparison to the model dual graph corresponding to the map coloured according to the four-colour theorem. The reduction in the edges of empirical dual graphs indicates inter-municipal economic cooperation. However, the existence of edges connecting neighbouring municipalities indicates the existence of certain growth barriers, because the websites belong to different quality classes.

Enterprises and sectors of the economy which initiate economic activity in a given area and contribute to better economic results of enterprises and industries operating in their environment are propulsion units that form regional poles of growth. The growth pole theory was developed in the 1950s by the French economist François Perroux [68], who included in his research many innovative elements that are the basis of today's complexity economics [69]. According to wikinomics, a characteristic feature of a growth pole is the initiation of economic activity with use of advanced ICT tools. A website of the municipal office of public administration, which acts as a participation platform and integrates and strengthens the economic

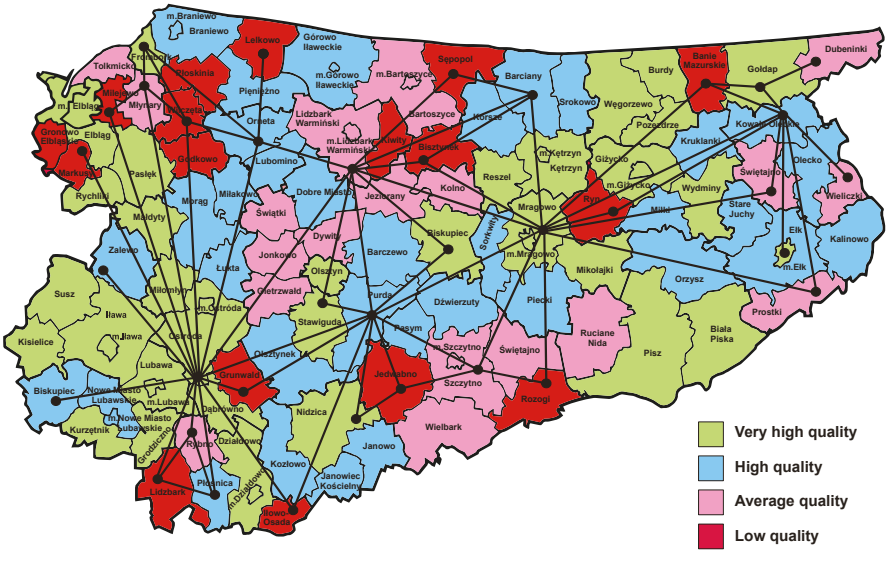
forces operating in the territory of a given municipality can be such a pole. The higher the quality of a website, the greater the chance that it will become a local growth pole.

The theory of development axes elaborated in 1963 by Pierre Pottier is closely connected with the theory of growth poles, according to which economic development is spreading along trade routes and transport networks connecting major industrial centres [70,71]. Territories located between growth poles and providing transport communication receive additional growth impulses due to the increased flow of goods, the spread of innovation and the development of infrastructure. Therefore, they transform into development axes (corridors), which together with the growth poles define the spatial framework for economic growth of a large region or country. The axis concept helps to connect the transport network with the urban hierarchy theories and growth centres into a single unit. These concepts can be easily adapted to the research regarding the Warmia and Mazury Province. Currently, the role of the development axes is played not only by transport networks, but also by computer networks. In the digital economy, the role of ICT is at least as important as traditional trade routes. The spatial distribution of municipalities that have high or very high quality websites form wikinomics poles of growth, indicates through which paths the economic growth spreads in space and where barriers to growth appear. In this way, digital growth turnpikes or development axes can be identified in the economic space. The lack of edges between neighbouring municipalities means that the municipalities have websites belonging to the same quality class. The high or very high quality of these websites means that development axes are created between these municipalities. However, if there are edges between neighbouring municipalities, it can be said that there are growth barriers between them.

Figure 3 presents a map of the Warmia and Mazury Province, where municipalities are marked with colours representing the quality classes of websites maintained by municipal public administration offices in 2009 [50]. A comparison of dual graphs corresponding to the reference map based on the four-colour theorem and the empirical map from 2009 indicates that the impact of local growth poles on the area is associated with a reduction in the number of vertices and the number of edges of the graph formed on the basis of the actual map. The dual graph shown in Figure 3 has significantly fewer vertices and edges than the dual graph associated with the reference map with no growth poles. In this way essential information regarding regional economic growth is presented. The dual graph associated with the empirical map from 2009 indicates a particular feature of the province in question. Overall, it is dominated by historically and politically conditioned infrastructural growth poles, that are closely related to road, rail and inland waterway infrastructure existing in the Warmia and Mazury Province [47]. Historical considerations result from the fact that this infrastructure is the effect of the work of many previous generations that have built it over the last several hundred years.

The infrastructure currently existing in the studied province is also the result of political history, which dates back to at least the 13th century. In the years 1226–1466 the area was ruled by the Teutonic Order (Monastic State of the Teutonic Knights). The order contributed to the economic development of these lands by introducing technical progress in the form wind milling improving the process of grinding and replacing the previously used hand-mills [72]. As a result of the Second Peace of Thorn, which was signed in 1466, the State of the Teutonic Order was split into Royal Prussia and Teutonic Prussia (since 1525 Ducal Prussia). Both parts of Prussia were actually attached to Poland, but the latter was dependent on fiefs [73,74]. The period from the 15th to the 18th century (and also a little later) was a golden period of development of windmills in Warmia and Mazury. At that time, these areas were formally or informally part of the Polish economy. The location of windmills in this area was consistent with records from the turn of the 19th and 20th centuries, and closely correlated with the existing transport infrastructure. Thus, windmills have marked the basic, existing to this day, trade and communication routes. To conclude these historical remarks, it should be noted that a significant part of the Warmia and Mazury Province later belonged to East Prussia and it was only in 1945 that it was incorporated into Poland as part of the

Recovered Territories. The studied province currently borders on the north with the Kaliningrad Region belonging to Russia. To sum up, the infrastructure currently existing in this area is not only the result of a specific economic history, but also of political history.



Legend: m.—urban municipality

Figure 3. The map of the municipalities of the Warmia and Mazury Province divided into categories of the quality of the websites maintained by municipal public administration offices in 2009 and a corresponding dual graph.

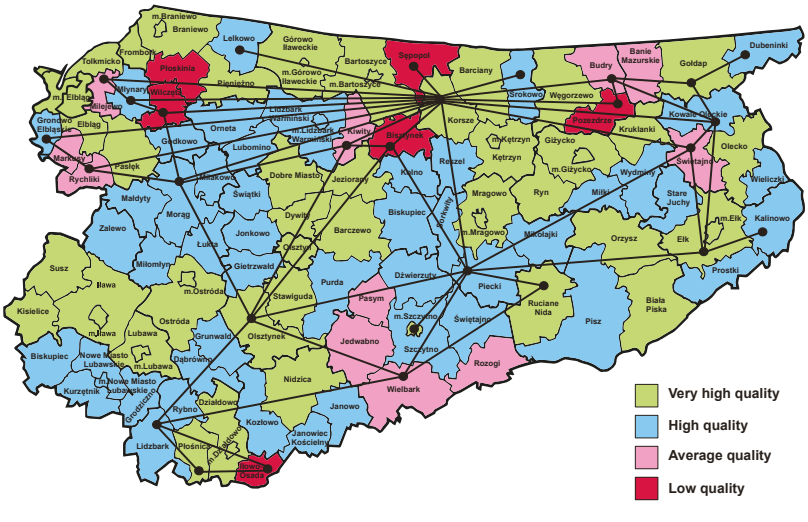
As noted above, high and very high quality municipal websites are closely linked to the transport infrastructure created in previous years. A cluster of municipalities with very high quality websites is in the western part of the province and it creates a development axis located on the north–south line. The transport infrastructure in this area comprises of four elements:

(1) the Ostróda–Elbląg Canal, which was built in the 19th century;
(2) Elbląg port, serving cargo barges floating on the Vistula Lagoon, improving trade exchange between Poland and Russia;
(3) road connections Nidzica–Olsztynek–Ostróda–Pasłęk–Elbląg–Braniewo–Królewiec with two road border crossings Gronowo–Mamonowo and Grzechotki–Mamonowo II;
(4) Malbork–Elbląg–Bogaczewo–Braniewo–Królewiec railway line with the Braniewo–Mamonowo railway border crossing.

The second cluster of municipalities with the highest quality of websites is located in the eastern part of the province and creates a development axis also running along the north–south line. There are two national roads in that part of the province. The first of them (No. 63) runs through Pisz, Orzysz, Giżycko, Węgorzewo and before 1945 led to today's Kryłów in the Kaliningrad Region (on the Russian side). Currently there is no border crossing there. The second road (No. 65) runs through the cities of Ełk, Olecko, Gołdap and ends with the Polish–Russian road border crossing Gołdap–Gusiew. On the

map presented in Figure 3 it is possible to notice the third cluster in the form of a strip of municipalities with high and very high quality sites located in the central part of the province, which is located on the southwest–northeast line. The national road No. 16, which connects towns such as Iława, Ostróda, Olsztyn, Mrągowo, Orzysz and Ełk, and then crosses the Podlasie Province and runs up to the border with Lithuania in Ogrodniki forms the main communication route in this region. This cluster has a railway line No. 353 connecting the towns of Iława–Ostróda–Olsztyn–Czerwonka–Korsze, which ends with the Skandawa–Żeleznodorożnyj railway border crossing. The fourth cluster, slightly less visible, is connected with the national road No. 51, which connects the towns of Olsztynek, Olsztyn, Dobre Miasto, Lidzbark Warmiński, Bartoszyce and ends with the Bezledy–Bagrationowsk border crossing. To sum up, the empirical map from 2009 together with the corresponding dual graph (Figure 3) allows the location of four clusters of municipalities with regional growth poles connected by a transport network. The Warmia and Mazury Province is therefore one of the best examples of the correctness of the development axes concept proposed by Pierre Pottier. It is also worth noting that the barriers to growth and development, represented by the edges of the dual graph, are located near the northern and southern borders of the province.

Figure 4 presents a map of the Warmia and Mazury Province, in which municipalities are marked with colours representing the quality classes of websites operated by municipal public administration offices in 2012 [50]. It can be observed that in comparison to 2009, the number of municipalities with high and very high quality websites has increased significantly. The dual graph from 2012 has significantly fewer edges than the dual graph from 2009. This demonstrates the emergence of new growth poles and development axes that have contributed to the elimination of many barriers to inter-municipal cooperation. Barriers located near the northern border of the province could not be removed, so investment and innovation should be continued in this area.



Legend: m.—urban municipality

Figure 4. The map of the municipalities of the Warmia and Mazury Province divided into categories of the quality of the websites maintained by municipal public administration offices in 2012 and a corresponding dual graph.

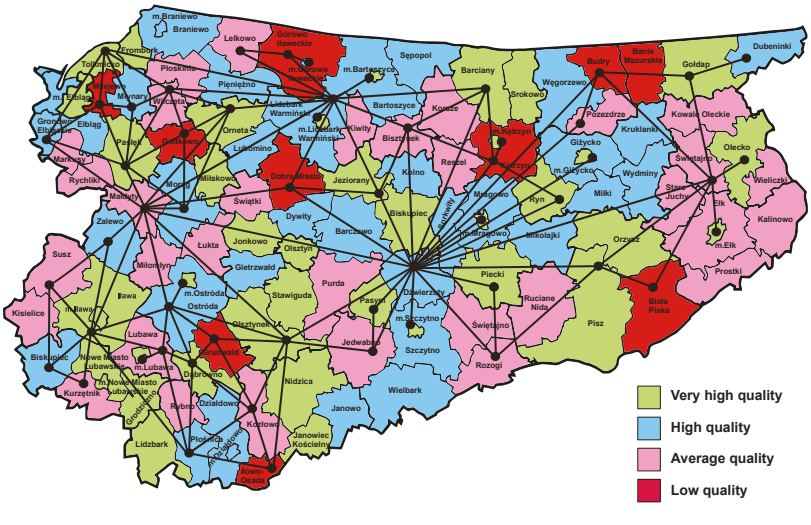
The investment in the construction of fibre-optic backbone distribution network with a length of 2292 km along with the appropriate infrastructure significantly contributed to the rapid progress in the development of municipal websites in 2012 [75,76]. The network consists of 216 distribution nodes, 10 backbone nodes, 2 contact points to a higher-level fibre-optic network and two network management centres. In the distribution layer, the network provides access with a bandwidth of at least 30 Mbit/s, and in the backbone layer at least 100 Gbit/s, thus meeting the requirements of next generation access (NGA) networks. The purpose of this project was to provide access to broadband internet to residents, enterprises and government institutions in the Warmia and Mazury Province. This investment was carried out in the years 2008–2015 and its value amounted to PLN 327,041,042.07 (approx. EUR 71,559,490.65). It was financed by the European Regional Development Fund, the state budget and the province's own contribution. The project was implemented on the basis of public–private partnerships as part of the design, build, operate, transfer (DBOT) model. The Warmia and Mazury Province is the owner of the created infrastructure, while the private partner was obliged to design, build, manage and operate the telecommunications infrastructure until the end of 2025. After this time, all generated assets will be transferred to the public partner.

Investment in the development of telecommunications infrastructure is crucial for the region's economy, as it will accelerate its development and facilitate the introduction of innovative technologies. Computer networks form the basis of digital economy, contribute to the intensification of modernisation processes of existing enterprises and stimulate the development of new products and services provided with use of the Internet. This type of growth refers to the well-known economic turnpike theory [77,78]. This name comes from an American English word meaning highway. Let us assume that we want to put the economy on the sustainable growth path, which is the maximum rate of growth in the sense of von Neumann. This path is associated with the top technique, which gives the highest rate of return. The solution resembles a dilemma of a driver who wants to reach his destination as soon as possible and has a choice between using the turnpike and the minor roads. Most often, the best choice turns out to be the use of a turnpike, even if it involves incurring costs at intermediate stages. This development method was chosen in the studied province, because according to wikinomics, the information and communication technique is now the top technology. Thus turnpikes can appear both in economic spaces, described by mathematical growth models, and in physical spaces, creating development axes defined by Pierre Pottier.

Figure 5 presents a map of the Warmia and Mazury Province, in which municipalities are marked with colours representing the quality classes of websites maintained by municipal public administration offices in 2015. It can be observed that the number of municipalities with high and very high website classes has heavily decreased. Thus, the number of municipalities with websites belonging to low or medium quality class has increased. At the same time, in comparison to 2012, the number of vertices and dual graph edges associated with the empirical map of 2015 increased, which can be interpreted as the emergence of new development barriers hampering the region's economic growth. The reasons for this state of affairs should be sought in the sphere of economic policy. The deterioration in the quality of websites may result from the fact that on 15 September 2015 all construction works related to the broadband network discussed above were completed. The network is backbone-distribution; therefore, its implementation does not signify an immediate improvement in internet access for residents, institutions and enterprises. It is also necessary to build the last mile networks, i.e., access networks, which will connect individual end users to the main network. This task was entrusted to local telecommunications operators, who on an equal basis can provide internet access services to interested entities from the Warmia and Mazury Province.

The decrease in the quality of many websites of municipal public administration offices can be explained by the fact that delays related with the construction of the last mile network were not taken into account and existing internet connections, such as radio networks, were too lightly abandoned. To prevent delays, the Digital Plan 2025 for Warmia and Mazury was implemented, which aims to

eliminate or reduce economic barriers that would hinder the construction of the last mile networks in the studied region [79,80]. This plan is a unique agreement in nationwide scale signed between the Province Board and the authorities of districts and municipalities from the region. As part of it, fees for placing telecommunications infrastructure along public roads have been reduced and unified, and this infrastructure has been exempted from property tax, in order to reduce the price of internet access for the end user. This programme is treated as a flywheel for the socio–economic development of the region, as it is intended to support the development of an innovative, low-carbon economy and direct entrepreneurship to services and products offered through a global network. It is worth noting that the internet network in the Warmia and Mazury Province is strongly correlated with the locations of windmills in the past centuries, most of which are gone today. As mentioned, windmills had a decisive impact on the formation of the transport network in the region. Over time, old trade and merchant routes were transformed into today's public roads. The development of modern road infrastructure was based on historical trade routes, which proved to be an economically advantageous solution. This is due to the geographical diversity of the land surface in Warmia and Mazury, where there are many rivers, streams, forests and thousands of lakes. Construction of new roads overcoming these obstacles would be too expensive, although such projects are undertaken in special cases, such as bypasses of major cities. The fact that the telecommunications infrastructure runs along ancient roads shows that the development axes in the region under review are extremely durable. In other words, not only growth poles, but also development axes are the result of historical and political considerations.

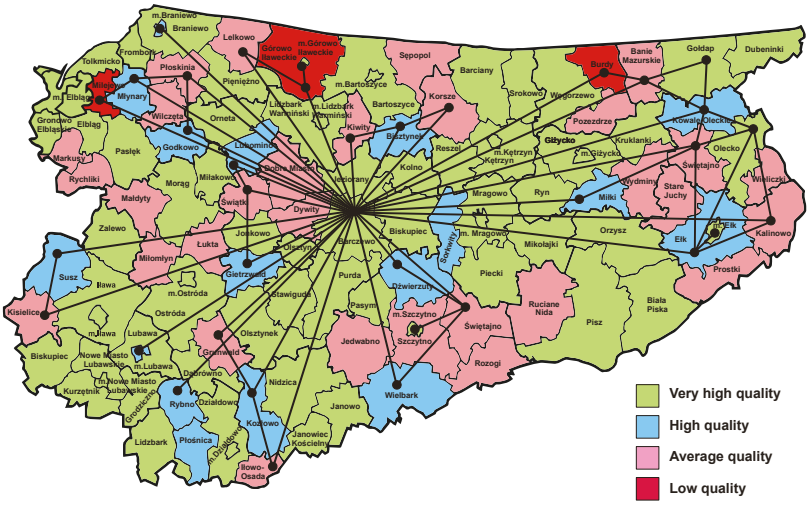


Legend: m.—urban municipality

Figure 5. The map of the municipalities of the Warmia and Mazury Province divided into categories of the quality of the websites maintained by municipal public administration offices in 2015 and a corresponding dual graph.

Figure 6 presents a map of the Warmia and Mazury Province, in which municipalities are marked with colours representing the quality classes of websites maintained by municipal public administration offices in 2019. As is easy to see, the number of municipalities with websites included in the classes of

high and very high quality has significantly increased. Some barriers that prevented inter-municipal cooperation in previous years have also disappeared. This signifies that new digital growth poles and development axes have emerged in the surveyed province. These changes prove that the Digital Plan 2025 for Warmia and Mazury is being implemented gradually. Nevertheless, the empirical dual graph from 2019 demonstrates that there are still many barriers to development, especially in the northern and eastern parts of the region.



Legend: m.—urban municipality

Figure 6. The map of the municipalities of the Warmia and Mazury Province divided into categories of the quality of the websites maintained by municipal public administration offices in 2019 and a corresponding dual graph.

8. Descartes' Theorem, Euler Characteristic, and Cauchy Modification

The determination of the measure of economic growth of the Warmia and Mazury Province requires designating the relationship between the numbers of vertices, edges and faces in the dual graph. René Descartes (1596–1650) was the first to study the general properties of polyhedra. He presented the basic theorem regarding the problem in the work *Progymnasmata de Solidorum Elementis* (Exercises on the Elements of Solids) [81]. It reads as follows:

Theorem 1. *The sum of the deficiencies of the solid angles in a polyhedron is always eight right angles.*

The deficiency of a solid angle is the amount by which the sum of its face angles is smaller than 360°. The size of deficiency determines the sharpness of the solid angle in such a way that the greater the deficiency the more acute the angle. In a cube, the deficiency of each solid angle is 90°, whereas in a dodecahedron the deficiency of each solid angle is 36°. It is worth noting that in both cases the deficiencies of all solid angles amount to 720°, which equals eight right angles (the whole sphere) [82] (pp. 187–189). Based on this theorem, Descartes formulated two statements:

1. There are always twice as many plane angles as sides [E] on the surface of a solid body, for one side is always common to two faces.
2. I always take α [V] for the number of solid angles and ϕ [F] for the number of faces. The actual number of plane angles is $2F + 2V - 4$.

Symbols V, E, and F in square brackets are abbreviations of modern equivalents of terms used by Descartes: numbers of vertices, edges, and faces of a polyhedron accordingly. The following relationship results from the above:

$$Number\ of\ plane\ angles = 2F + 2V - 4 = 2E, \tag{8}$$

which gives the basic result:

$$F + V - E = 2. \tag{9}$$

Perhaps Descartes was unaware of the existence of such a relationship in a polyhedron [83] (p. 515), but it is not certain. There is no proof of this theorem in his work. However, Descartes did not publish the work *Progymnasmata*, its content survived as a result of several unusual events. The history of this manuscript is described by Peter Cromwell in his book *Polyhedra* [82] (pp. 181–182):

> *In the autumn of 1649, Descartes went to Stockholm at the invitation of Queen Christina of Sweden, but the severity of the climate was too much for him and he died six months later. His belongings were shipped back to France but suffered accident on route, the box carrying his manuscripts ending up in the Seine at Paris. The papers were rescued from the river, separated and dried. Later, some were published and the remainder were made available for consultation. In 1676 Gottfried Wilhelm Leibniz (one of the founders of the calculus) made copies of several of the latter manuscripts including the work on polyhedra. Descartes' original manuscript has vanished and it is only through the copy that the work is preserved. Even so, it remained unknown until 1860 when the copy was discovered by Comte Foucher de Careil among a collection of uncatalogued Leibniz papers.*

Leonhard Euler (1707–1783)—Swiss mathematician and physicist—was another researcher who began studying the general properties of polyhedra. At the time mathematicians used to describe their findings in letters to friends. This was due to the long waiting time for publication of results. Euler corresponded for many years with the German mathematician Christian Goldbach (1690–1764). In a letter dated 14 November 1750 Euler wrote to Christian Goldbach that he had begun studying polyhedra. The Swiss scholar aimed to categorise the properties of various and seemingly unrelated solids. Below is an excerpt from his letter [84] (p. 76):

> *Recently it occurred to me to determine the general properties of solids bounded by plane faces, because there is no doubt that general theorems should be found for them, just as for plane rectilinear figures, whose properties are:*
>
> *(1) that in every plane figure the number of sides is equal to the number of angles, and*
> *(2) that the sum of all the angles is equal to twice as many right angles as there are sides, less four.*
>
> *Whereas for plane figures only sides and angles need to be considered, for the case of solids more parts must be taken into account . . .*

Euler wrote two papers on the polyhedral formula that were published in 1758. The first contains the statement of the theorem [85], while the second contains the poof [86]. The theorem given by Euler has the form:

Theorem 2. *In every solid enclosed by plane faces, the number of faces along with the number of solid angles exceeds the number of edges by two.*

Using previously proposed symbols the relation can be presented as follows:

$$F + V = E + 2, \tag{10}$$

hence

$$\chi = V - E + F = 2, \tag{11}$$

where χ stands for the Euler characteristic. The number χ is a topological invariant, as it describes a topological space's shape or structure. Euler aimed to classify all polyhedra, by developing the theory of stereometry (solid geometry) in the image of existing planimetry (planar geometry). However, in the above theorem, he indicates that the considered polyhedra are enclosed by planes, therefore convex polyhedra, which is a sufficient condition for the truthfulness of his formula. Euler, however, was not able to provide a rigorous proof of the theorem [87–89].

In 1813, French mathematician Augustin-Louis Cauchy (1789–1857) generalised the Euler characteristic by projecting the polyhedron onto a plane. One of the faces of the polyhedron is assumed as the basis, and then all the other vertices are transferred to it without changing their number, which gives a set of polygons within this chosen face. In other words, the result of flattening of the polyhedron is a plane network of polygons. In this way Euler's formula can be presented in the form of planar graphs or equivalently as maps on the plane. Cauchy's theorem has the following form [90] (p. 77):

Theorem 3. *If a polyhedron is decomposed into as many others as we choose, by taking at will new vertices in the interior, and if the number of new polyhedra so formed is denoted by P, the total number of vertices including those of the original polyhedron by S [V] the total number of faces by F, and the total number of edges by A [E], then*

$$V + F = E + P + 1, \tag{12}$$

that is, the sum of the number of vertices and that of the faces exceeds by one unit the sum of the number of edges and that of the polyhedra.

Euler's theorem is a special case of Cauchy's theorem. After assuming that all the polyhedra are reduced to a single one, which is equivalent to $P = 1$, the Equation (10) is obtained. Another theorem that applies to plane geometry can also be derived from Cauchy's theorem. Let all polyhedra be reduced to one. Let us assume that the last polyhedron will be transformed so that all the other vertices will be projected onto the selected face, without changing their number. This means substitution $P = 0$ in Equation (12), which gives

$$V + F = E + 1. \tag{13}$$

This leads to the conclusion that the sum of the number of vertices and the number of faces exceeds by one unit the number of edges. Theorem (13) in plane geometry is the equivalent of the general theorem (12) in the geometry of polyhedra. Both theorems have been proved by Cauchy [84] (pp. 79–83).

9. Measurement of Economic Development in the Province of Warmia and Mazury

The proposed method of economic growth measurement is based on dual graphs, one of which—corresponding to the model map—plays the role of a reference system, while the others contain information about growth poles and development axes in individual years. The essence of economic growth measurement consists in comparing information contained in empirical graphs with the model graph. The latter is an absolute reference system, which makes our measure resistant to relative comparisons. It is something like absolute time in the sense of Isaac Newton. This situation in science

is extremely rare, because all reality is constantly changing. The four-colour theorem gives us a kind of anchor that establishes dynamic and lasting links between an objective mathematical being—which is the four-colour theorem—and the changing economic situation. Usually in science we are constricted to relative comparisons. Calculation of real amounts of national income when we have nominal amounts of this income in individual years is a good example of this problem. The only way to solve this dilemma is to adopt the national income from a given year as the base value, and then determine the income in the remaining years at prices from the base year. This is an assumption that must be made and on which the calculation result depends. In the measure of economic growth presented here, this assumption is unnecessary.

The next step of the proposed method is the reduction of the dual graph. It happens when two bordering municipalities have web pages belonging to the same quality class. The removal of the edge means that the barrier of local growth has been overcome, which allows both municipalities to grow together which in turn may allow them to become the seed of a new growth pole. The edges of the dual graph correspond to the growth barriers, so they should be calculated precisely. For this purpose, the Euler characteristics $\chi(S)$ for convex polyhedra can be used as follows:

$$\chi(S) = V - E + F = 2, \tag{14}$$

Due to the need for economical interpretation of the dual graph it is better to apply the Cauchy interpretation of the above formula, which is applicable in plane geometry:

$$\chi(G) = V - E + F = 1. \tag{15}$$

In this way, we focus only on the dual graph itself, which is a representative of the contribution of the public administration sector to the economy of the Warmia and Mazury Province. In this case, we are not interested in the environment, and thus in the calculation of $\chi(G)$, the infinite (unbounded) face outside the dual graph, which contains a point in infinity, is not counted. Table 3 presents the results of calculating the number of vertices, faces and edges of all the dual graphs discussed above with the use of the Cauchy formula.

Table 3. Characteristics of dual graphs corresponding to the political map of South America, the reference map and four empirical maps of the Warmia and Mazury Province.

Type of Dual Graph	Number of Vertices (*V*)	Number of Faces (*F*)	Cauchy's Formula Interpretation	
			Number of Edges (*E*=*V*+*F*−1)	Euler Characteristic (χ(*G*)=*V*−*E*+*F*)
Political map of South America	13	13	25	1
The reference graph (four-colour theorem)	116	162	277	
The empirical graph 2009	36	28	63	
The empirical graph 2012	28	22	49	
The empirical graph 2015	55	52	106	
The empirical graph 2019	36	23	58	

The definition of a new measure of the region's economic development (M_D) based on the four-colour theorem is as follows:

$$M_D^{Year} = 1 - \frac{Sum\ of\ edges\ of\ the\ dual\ graph\ on\ the\ real\ map}{Sum\ of\ edges\ of\ the\ dual\ graph\ on\ the\ reference\ map}. \tag{16}$$

The limit values of the fraction in the above formula are as follows:

0—(no edges) dual graph has only one vertex ⇒ all municipalities cooperate with each other ⇒ $M_D^{Year} = 1$.

1—(all edges) no municipality cooperates with any other ⇒ empirical map = reference map ⇒ $M_D^{Year} = 0$.

This measure, apart from being economic in nature and referring to the equipment of the public administration sector in the latest telecommunications technologies, also takes into account the spatial aspects of economic development. It can be applied to any region that consists of smaller and simultaneously autonomous spatial units.

The proposed measure in one number contains a complex set of information, and at the same time it is simple to use. It can therefore be used in interregional research, and it also allows intertemporal analyses. It has both a static and dynamic character. It can be used to identify spatial growth poles and development axes in the public administration sector, and thus also in the entire regional economy, in this case the province economy. However, it is required that there are public administration offices in the smallest surveyed units (municipalities) belonging to a higher-order unit (province).

The calculation method of spatial economic growth measures in individual years in accordance with the Formula (16) and the data contained in Table 3 is as follows:

$$M_D^{2009} = 1 - \frac{63}{277} = 0.7726, \tag{17}$$

$$M_D^{2012} = 1 - \frac{49}{277} = 0.8231, \tag{18}$$

$$M_D^{2015} = 1 - \frac{106}{277} = 0.6173, \tag{19}$$

$$M_D^{2019} = 1 - \frac{58}{277} = 0.7906. \tag{20}$$

The result can be rounded to four decimal places. Establishing certain thresholds defining the development phases of the region requires similar research to be carried out in all provinces, i.e., on the scale of the entire national economy.

Interpretation of the indicators points to a conclusion that that despite the increased financial outlays, public administration in the Warmia and Mazury Province is a weak factor of economic growth. Taking as a basis the assessment of the starting situation from 2009, we can see a great progress in 2012, intensified inter-municipal cooperation, as evidenced by the fact that many edges of the dual graph are reduced. New local growth poles appear and new ways of spreading economic growth are formed. However, in 2015 regress occurs and the situation becomes worse than in 2009. This should be connected with the completion of the construction of a fibre-optic network, which allowed access to broadband internet for businesses in almost the entire region. This means that public administration is a very sensitive sector for changes in financing. The source of this financing is mainly government spending. It is true that the network was built, but apparently it was forgotten that it is not a one-time investment and that it requires further government spending to properly use it. However, a remedy in the form of the Digital Plan 2025 for Warmia and Mazury was found relatively quickly [79]. The situation improved significantly in 2019, which is a sign of clear progress in the construction of the last mile networks.

There are interconnections between the growth poles and development axes in the form of the municipal digital platforms and the transport infrastructure. One is a mirror image of the other. Strictly speaking, there are feedbacks between them, which can be both positive and negative. Certainly, the construction of the fibre-optic backbone distribution network will have a great impact on the expansion of

the transport infrastructure and thus contribute to the economic development of the region. However, infrastructural changes are slow, so the economic development measure expressed by the Equation (16) may be a leading indicator. According to the latest research, in Warmia and Mazury the synthetic index of regional innovation development has shown a steady decline since 2009. In 2009, this index was 0.27, in 2014 it was 0.24, and in 2019 it decreased to 0.22 [91]. After comparing it with the new measure of economic development introduced in this article, it is clear that digital growth poles and development axes precede investments in the expansion of transport networks, which form the main routes to prosperity and welfare in the region. As infrastructure-based economic growth is slow, the time gap between the two indicators can range from several months to even several years.

10. Discussion

The application of *k*-means clustering led to the creation of four quality classes for websites of municipal public administration offices in the four examined years. As presented below, this classification of websites is natural in a topological sense. Moreover, the diversification of the examined objects from a static point of view (one point in time, i.e., a specific year) and dynamic point of view (four points in time or more) allows the drawing of more far-reaching conclusions. Changes taking place in quality classes—represented by the transformation of dual graphs—prove that the Warmia and Mazury Province is a complex adaptive system. A complex adaptive system should be understood as an object that exchanges energy, matter and information with the environment; furthermore it functions based on cognitive schemas and has the ability to tune to the edge of chaos [92]. It is therefore an open system showing emergence. This conclusion can be justified on the basis of the catastrophe theory, which is a method of classifying stable forms [93–101].

Catastrophe theory is based on Thom's classification theorem [102–104]. It indicates that any singularity of the catastrophe map is equivalent with one type of singularity belonging to a finite family of types, which are called elementary catastrophes. In addition, the catastrophe map is locally stable at all manifold points due to small perturbations. The number of elementary catastrophes depends only on the codimension (the number of control variables). Elementary catastrophes describe all possible ways of manifesting discontinuities in dynamical systems. Considering the examined problem of the division of websites into four quality classes, the case of codimension = 3 is essential. Then the control space can be interpreted as physical space–time, which consists of two spatial dimensions and one time dimension. All events that occur in this space–time rely on the transformations of the dual graph on the plane, which occur over time. The catastrophe map then has five types of singularities. Four elementary catastrophes can be attributed to four quality classes of websites of municipal public administration offices, while the fifth catastrophe is related to the environment, and thus to what is located outside of the province under study. This interpretation of reality seems to be confirmed by life itself. The division of a three-dimensional entity, consisting of a two-dimensional object located on a plane and changing in time, into four quality classes is therefore natural. In other words, the compliance of the division of websites into four quality classes with Thom's classification theorem proves that this division was made correctly. The fifth catastrophe, representing the space located outside of the examined province, emphasises the fact that the examined province is an open system that exchanges matter, energy and information with the environment, i.e., with the rest of the national economy or even the world economy. It is worth noting that monetary flows that enable the functioning of municipal public administration offices and initiate the economic development of the province originate from the space outside of the province.

The studied province is a complex adaptive system because it can be described by a measure called total information, which is the sum of effective complexity and an entropy term [105]. Effective complexity is the length of a compact description of the regularities identified in the examined system, and thus

describes the object's rule-based features. The effective complexity of the studied province is represented by dual graphs. Whereas the entropy term measures the information necessary to describe the random aspects of the entity. Entropy can also be understood as an ignorance measure based on Shannon informational entropy [106,107]. In this case, ignorance should be understood as the degree of unawareness of the national economic centre regarding the investment needs of the region.

One of the issues that need clarification is the impact of the geographical line complexity on results obtained through the four-colour theorem. Borders between countries, provinces and municipalities are not always smooth lines, sometimes they can have a very complex shape, depending on the terrain. The Warmia and Mazury Province is very diverse in terms of nature, has over 3000 lakes, many rivers and streams, and more than a third of its surface is covered by forests. The topography is mostly lowland, but hills and valleys occur in many places. In such conditions, it is often the case that the geographical lines separating individual municipalities are complex. It seems logical that the longer the common border between the two municipalities, the greater the likelihood that they will undertake economic cooperation and thus transform into one pole of growth. However, there may be exceptions to this rule when neighbouring municipalities compete for some common resources.

English researcher Lewis F. Richardson was the first to apply a topological approach to study armed conflicts between states [108,109]. He considered various causes of wars, and one of them was spatial relations. In particular, he studied the impact of common borders between the two countries on the emergence of conflicts. Richardson used the Euler characteristic for this purpose, which presents the relationship between the numbers of vertices, edges and faces in the polyhedron. It is established that during the period he studied, i.e., the years 1820–1950, there were about 60 stable nations and empires in the world. This allowed him to demonstrate that for any plausible arrangement of nations, the average number of neighbours for each of these countries should be around six. So, if countries seeking war would select their enemies entirely at random, there would be a ten percent chance that every pair at war would have a common border. Richardson's research, however, showed that in the period under review, out of 94 international wars, which had only two participants, only 12 cases concerned combatants that had no common borders. This leads to the conclusion that close neighbourhood can be one of the main causes of conflicts.

Richardson found that the result of measuring the length of complex geographic lines, such as sea coasts or borders between countries, depends on the unit of measurement [110] (p. 26). The measured length is the greater the smaller the yardstick. This is because a shorter ruler measures more accurately the sinuosity of coastlines or land borders than a longer ruler because it takes greater account of the roughness. According to intuition, one would expect that taking smaller and smaller units of measurement will cause the total of the measurements to tend to a certain finite number representing the true length of the geographical line. However, it has been established to be otherwise. If the length of the meter decreases to zero, then the length of the shoreline or boundary tends to infinity. This phenomenon is now known as the Richardson effect. In this context, the coastline paradox is also discussed. It can be concluded that the length is not an adequate feature to describe a geographical line, because its length is simply undefinable. Geographical lines are actually fractals that can be described by the fractal dimension which is determined by the slope of a straight line obtained by plotting the length of the ruler versus the measured length of the geographical line on a log–log plot [111,112]. This is how Richardson discovered fractals [113] (p. 260).

There are similarities between the measure of economic development of a separate spatial object (province) based on the four-colour theorem proposed in this article and Richardson's research on the causes of armed conflicts. Richardson focused on the study of spatial relations, which prompted him to use dual graphs to analyse the data collected on international wars. Modern reconstructions show that these graphs were very complicated [114]. War was marked on the map with an edge connecting the countries involved. In addition, the thickness of the edge reflected the magnitude of the conflict.

Richardson classified wars and other quarrels on the basis of a method borrowed from astronomy, which involved the use of common logarithms, i.e., the base-10 logarithms. Hence a terror campaign, during which 1000 people die has a magnitude of 3 $\left(\log 10^3 = 3\right)$, and a war with 10 million casualties has a magnitude of 7 $\left(\log 10^7 = 7\right)$.

In Richardson's research, the edges of dual graphs denote armed conflicts and their magnitudes, and thus—generally speaking—the lack of peace, while in our research the edges signify a lack of cooperation. Thus, Richardson's conclusions can be at least partially extrapolated to municipalities. The main difference lies in the fact that people die during wars, while the lack of cooperation between municipalities slows down regional economic development. However, it may be concluded that logical homologies identified with use of the above method indicate a relatively new and interesting direction of regional research. Inter-municipal cooperation does not have to be treated as either there or not, but it can be gradable.

Richardson noted that administratively designated internal borders usually look very different from the natural external borders that separate countries. Usually, internal borders take the form of straight lines that are clearly and directly drawn on the maps, typically regardless of the natural terrain. In contrast, external borders often use natural features such as rivers or mountain ranges. His conclusions regarding the internal and external borders do not match the Warmia and Mazury Province. Very often municipal boundary lines use the natural terrain and run along rivers, lakes, forests and arable fields. Probably such a division was intended to avoid competition for the share of the same resources. On the other hand, the international border, which separates the examined province from the Kaliningrad Region, which belongs to Russia, has a much simpler course. Since the administrative lines separating the municipalities have such a complicated track, it means that the municipalities are fractals, i.e., objects showing statistical self-similarity. This is because they contain reduced copies of themselves. Therefore, the fractal dimension should be used to describe the boundaries between municipalities, not the usual length. In this way, complexity is manifested in the four-colour theorem. Therefore, this theorem can be used to colour each map, including the infinitely complex fractal patterns drawn on the complex plane [36].

11. Conclusions

Changes in growth poles and development axes in the studied province are represented by empirical dual graphs, which were determined for 2009, 2012, 2015 and 2019. The measure of development consists in comparing information contained in empirical graphs with the model dual graph, which was created on the basis of a map corresponding to the conditions of the four-colour theorem. Graph edges are interpreted as barriers to cooperation between neighbouring municipalities. The reference graph has the maximum number of edges, and therefore represents the case of a complete absence of growth poles and development axes. It plays the role of the absolute reference system in the research. As demonstrated, the public administration sector at the local level is very sensitive to changes in financing, which can be seen especially in the example of the implementation of telecommunications technologies. Most projects in this sector depend on the level of government spending. In the thermoeconomics terms, changes in empirical dual graphs can be treated as a result of the flow of monetary entropy through the Warmia and Mazury Province [115]. The studied region is a dissipative system that intakes low entropy from the environment, which increases the level of organisation of municipal public administration offices. This is reflected in the transition of websites to higher quality classes. Money allocated for the development of telecommunication infrastructure is the medium of low entropy. Lack of money is associated with a decrease in the quality of websites and their stabilisation in the lower classes. In this way entropy increases in the whole system. The inflow of money reduces the internal entropy of the system, which is associated with the formation of new growth poles and development axes. Maintenance of existing infrastructure at the 2012 level is therefore conditioned by the inflow of additional funds. Province authorities are well

aware of this, which is why they launched the Digital Plan 2025 for Warmia and Mazury. The goal of the programme is the proper operation of the fibre-optic backbone distribution network, and thus providing access to the network to the end subscribers.

The proposed measure of economic growth contains in one number a rich collection of economic and spatial information. It includes Hicks' condition for top technique because it incorporates wikinomics business models such as participation platforms and prosumption that could not exist without computer networks. Additionally, this measure enables the location of regional growth poles and development axes and considers the spatial diversity of municipalities and barriers in their mutual cooperation. It is based on digital technologies, so it is not limited to one industry, but covers all those industries that are controllable by software. Therefore, this measure has an advantage over the location quotient and all other classic measures of regional growth and development. Furthermore, it should be emphasised that the new measure applies to the developing economy, in which the share of high and very high quality municipal websites is significant. The four-colour theorem can be used as a reference system for empirical studies only if it corresponds to the Thom classification theorem. This signifies adequate representation of elementary catastrophes by the quality classes of the municipal websites. The proposed measure also takes into account the spatial complexity of municipalities, because in the Warmia and Mazury Province, most of them are fractals. It follows that both the management of the entire province and individual municipalities is a complex problem and only fractal organisations can cope with it [116–118]. Therefore, all public administration reforms should aim at giving the administration fractal features, which applies in particular to municipal public administration offices as basic units of the local government.

Author Contributions: Conceptualisation, A.J. and D.R.; methodology, A.J. and D.R.; software, A.J. and D.R.; validation, A.J. and D.R.; formal analysis, A.J. and D.R.; investigation, A.J. and D.R.; resources, A.J. and D.R.; data curation, A.J. and D.R.; writing—original draft preparation, A.J. and D.R.; writing—review and editing, A.J. and D.R.; visualisation, A.J. and D.R.; supervision, A.J. and D.R.; project administration, A.J. and D.R.; funding acquisition, A.J. and D.R. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hicks, J.R. *Capital and Growth*; Oxford University Press: New York, NY, USA, 1965; ISBN 978-0198281504.
2. Von Neumann, J. A Model of General Economic Equilibrium. *Rev. Econ. Stud.* **1945**, *13*, 1–9. [CrossRef]
3. Tapscott, D.; Williams, A.D. *Wikinomics: How Mass Collaboration Changes Everything*; Portfolio: New York, NY, USA, 2006; ISBN 978-1-59184-138-8.
4. Tapscott, D.; Williams, A.D. *Macrowikinomics: Rebooting Business and the World*; Portfolio/Penguin: New York, NY, USA, 2012; ISBN 978-1-59184-356-6.
5. Toffler, A. *The Third Wave*; William Morrow and Company: New York, NY, USA, 1980; ISBN 0-688-03597-3.
6. Toffler, A.; Toffler, H. *Revolutionary Wealth: How It Will Be Created and How It Will Change Our Lives*; Doubleday: New York, NY, USA, 2007; ISBN 978-0-385-52207-6.
7. Czyżewska-Misztal, D.; Golejewska, A. The least innovative regions in Poland in the process of smart specialization. *Optimum* **2016**, *5*, 123–137. [CrossRef]
8. Wichowska, A. Shrinking municipalities and their budgetary revenues on the example of the Warmian-Masurian Voivodeship in Poland. *Oeconomia Copernicana* **2019**, *10*, 419–432. [CrossRef]
9. Wierzbicka, W. Information infrastructure as a pillar of the knowledge-based economy—an analysis of regional differentiation in Poland. *Equilibrium* **2018**, *13*, 123–139. [CrossRef]
10. Wierzbicka, W. Socio-economic potential of cities belonging to the Polish National Cittaslow Network. *Oeconomia Copernicana* **2020**, *11*, 203–224. [CrossRef]

11. Harasym, R.; Rodzinka, J.; Skica, T. The size of local government administration at a municipal level as a determinant of entrepreneurship. *JEMI* **2017**, *13*, 5–31. [CrossRef]
12. Isard, W. *Methods of Regional Analysis: An Introduction to Regional Science*; The MIT Press: Cambridge, MA, USA, 1960; ISBN 9780262090032.
13. Edwards, M.E. *Regional and Urban Economics and Economic Development: Theory and Methods*; Auerbach Publications: Boca Raton, FL, USA, 2007; ISBN 978-0-8493-8317-5.
14. McCann, P. *Modern Urban and Regional Economics*, 2nd ed.; Oxford University Press: Oxford, UK, 2013; ISBN 978-0-19-958200-6.
15. Brantingham, P.L.; Brantingham, P.J. Mapping crime for analytic purposes: Location quotients, counts, and rates. In *Crime Mapping and Crime Prevention*; Weisburd, D., McEwen, T., Eds.; Criminal Justice Press: Monsey, NY, USA, 1997; pp. 263–288. ISBN 1-881798-08-9.
16. Isserman, A.M. The location quotient approach to estimating regional economic impacts. *J. Am. Plann. Assoc.* **1977**, *43*, 33–41. [CrossRef]
17. Klosterman, R.E.; Brail, R.K.; Bossard, E.G. *Spreadsheet Models for Urban and Regional Analysis*; Center for Urban Policy Research: New Brunswick, NJ, USA, 1993; ISBN 0-88285-142-X.
18. Skidmore, Owings and Merrill. Guidance Notebooks for the Environmental Assessment of Airport Development Projects. In *Notebook 2: Environmental Assessment Techniques*; University of Michigan Library: Ann Arbor, MI, USA, 1978.
19. Keynes, J.M. The general theory of employment, interest and money. In *The Collected Writings of John Maynard Keynes*; Robinson, A., Moggridge, D., Eds.; Cambridge University Press: Cambridge, UK, 2013; Volume 7, ISBN 978-1-107-67373-1.
20. Fritsch, R.; Fritsch, G. *The Four-Color Theorem. History, Topological Foundations, and Idea of Proof*; Springer-Verlag: New York, NY, USA, 1998; ISBN 0-387-98497-6.
21. Wilson, R.J. *Four Colors Suffice: How the Map Problem Was Solved, Revised Color edition*; Princeton University Press: Princeton, NJ, USA, 2014; ISBN 978-0-691-15822-8.
22. Barnette, D. *Map Coloring, Polyhedra, and the Four-Color Problem*; The Mathematical Association of America: Washington, DC, USA, 1983; ISBN 0-88385-309-4.
23. Kempe, A.B. On the geographical problem of the four colours. *Amer. J. Math.* **1879**, *2*, 193–200. [CrossRef]
24. Appel, K.; Haken, W. *Every Planar Map is Four Colorable*; American Mathematical Society: Providence, RI, USA, 1989; ISBN 0-8218-5103-9.
25. Robertson, N.; Sanders, D.P.; Seymour, P.; Thomas, R. A new proof of the four-colour theorem. *Electron. Res. Announc. Amer. Math. Soc.* **1996**, *2*, 17–25. [CrossRef]
26. Robertson, N.; Sanders, D.; Seymour, P.; Thomas, R. The four colour theorem. *J. Combin. Theory Ser. B* **1997**, *70*, 2–44. [CrossRef]
27. Gonthier, G. A Computer-Checked Proof of the Four Colour Theorem. Microsoft Research Cambridge. Available online: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/gonthier-4colproof.pdf (accessed on 28 July 2020).
28. Gonthier, G. Formal proof—the four-color theorem. *Not. AMS* **2008**, *55*, 1382–1393. Available online: http://www.ams.org/notices/200811/tx081101382p.pdf (accessed on 26 July 2020).
29. Steinberger, J.P. An unavoidable set of *D*-reducible configurations. *Trans. Amer. Math. Soc.* **2010**, *362*, 6633–6661. [CrossRef]
30. Krantz, S.G. *The Proof is in the Pudding: The Changing Nature of Mathematical Proof*; Springer Science + Business Media: New York, NY, USA, 2011; ISBN 978-0-387-48908-7. [CrossRef]
31. Swart, E.R. The philosophical implications of the four-color problem. *Am. Math. Mon.* **1980**, *87*, 697–707. [CrossRef]
32. Tymoczko, T. The four-color problem and its philosophical significance. *J. Philos.* **1979**, *76*, 57–83. [CrossRef]
33. Bondecka-Krzykowska, I. The four-color theorem and its consequences for the philosophy of mathematics. *Ann. UMCS Inform.* **2004**, *AI 2*, 5–14. [CrossRef]

34. Dahl, E.D. Programming with D-Wave: Map Coloring Problem. D-Wave White Papers, D-Wave Systems: The Quantum Computing Company, Burnaby, BC, Canada. November 2013. Available online: https://www.dwavesys.com/resources/publications?type=white (accessed on 26 April 2020).
35. Smoluk, A. A simple proof of the four-colors theorem. *Didact. Math.* **2016**, *13*, 35–38. [CrossRef]
36. Oller, J.W., Jr. The four-color theorem of map-making proved by construction. *Open Access Libr. J.* **2016**, *3*, e3089. [CrossRef]
37. Barbosa, A.L. A Human-Checkable Four-Color Theorem Proof. Available online: https://arxiv.org/abs/1708.07442v19 (accessed on 24 July 2020).
38. Calude, C.S.; Calude, E. The complexity of the four colour theorem. *LMS J. Comput. Math.* **2010**, *13*, 414–425. [CrossRef]
39. Wilson, R.J. *Introduction to Graph Theory*, 4th ed.; Addison-Wesley Longman: Harlow, UK, 1996; ISBN 0-582-24993-7.
40. Ore, O. *The Four-Color Problem*; Academic Press: New York, NY, USA, 1967; ISBN 9780080873398.
41. Ahmed, S. Applications of graph coloring in modern computer science. *Int. J. Comput. Inf. Technol.* **2012**, *3*, 130101. Available online: http://www.ijcit.org/ijcit_papers/vol3no2/IJCIT-130101.pdf (accessed on 27 March 2020).
42. Bajerowski, T.; Gerus-Gościewska, M. Zagadnienie czterech barw jako inspiracja badań użytkowania przestrzeni wiejskiej. *Zeszyty Naukowe Akademii Rolniczej im. Hugona Kołłątaja w Krakowie* **1999**, *353*, 57–63.
43. Gerus-Gościewska, M.; Bajerowski, T. Zagadnienie czterech barw a teoria i praktyka doboru funkcji planistycznych. In *Koncepcje Teoretyczne i Metody Badań Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej*; Rogacki, H., Ed.; Bogucki Wydawnictwo Naukowe: Poznań, Poland, 2001; pp. 283–291. ISBN 83-88163-70-1.
44. Kainen, P.C. Quantum Interpretations of the Four Color Theorem. Technical Report. 11 May 1999. Available online: http://faculty.georgetown.edu/kainen/qtm4ct.pdf (accessed on 16 March 2020).
45. Popescu, M.-C.O.S.; Mastorakis, N.E. Applications of the four color problem. *Int. J. Appl. Math. Inform.* **2009**, *3*, 17–26. Available online: http://www.naun.org/main/UPress/ami/19-180.pdf (accessed on 21 March 2020).
46. Timofte, R.; Van Gool, L. Efficient loopy belief propagation using the four color theorem. In *Advanced Topics in Computer Vision*; Farinella, G.M., Battiato, S., Cipolla, R., Eds.; Springer: London, UK, 2013; pp. 313–339. ISBN 978-1-4471-5520-1. [CrossRef]
47. Jakimowicz, A.; Rzeczkowski, D. Application of the four colour theorem to identify spatial regional poles and turnpikes of economic growth. *Acta Phys. Pol. A* **2018**, *133*, 1362–1370. [CrossRef]
48. Appel, K.; Haken, W. The four-color problem. In *Mathematics Today: Twelve Informal Essays*; Steen, L.A., Ed.; Springer: New York, NY, USA, 1979; pp. 153–180. ISBN 0-387-90305-4.
49. Jakimowicz, A.; Rzeczkowski, D. Prosumption in the public administration sector. *Acta Phys. Pol. A* **2016**, *129*, 1011–1017. [CrossRef]
50. Rzeczkowski, D. *Potencjał Innowacyjności Sektora Administracji Publicznej*; Wydawnictwo Naukowe PWN: Warsaw, Poland, 2014; ISBN 978-83-01-17459-0.
51. Everitt, B.S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*, 5th ed.; John Wiley and Sons, Ltd.: Chichester, UK, 2011; ISBN 978-0-470-74991-3.
52. Tryon, R.C. *Cluster Analysis: Correlation Profile and Orthometric (Factor) Analysis for the Isolation of Unities in Mind and Personality*; Edwards Brothers, Inc., Lithoprinters and Publishers: Ann Arbor, MI, USA, 1939.
53. MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability: Statistics. Held at the Statistical Laboratory University of California 21 June–18 July 1965 and 27 December 1965–7 January 1966*; Le Cam, L.M., Neyman, J., Eds.; University of California Press: Berkeley, CA, USA, 1967; Volume 1, pp. 281–297.
54. Steinhaus, H. Sur la division des corps matériels en stron. *Bull. Acad. Polon. Sci. Cl. III* **1956**, *4*, 801–804.
55. Hartigan, J.A. *Clustering Algorithms*; John Wiley and Sons, Inc.: New York, NY, USA, 1975; ISBN 0-471-35645-X.
56. Hartigan, J.A.; Wong, M.A. A *k*-means clustering algorithm. *J. Royal Stat. Soc. Ser. C Appl. Stat.* **1979**, *28*, 100–108. [CrossRef]

57. Telgarsky, M.; Vattani, A. Hartigan's method: *K*-means clustering without Voronoi. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2010, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010; Available online: http://proceedings.mlr.press/v9/telgarsky10a/telgarsky10a.pdf (accessed on 8 May 2020).
58. Lloyd, S.P. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–137. [CrossRef]
59. Garey, M.; Johnson, D.; Witsenhausen, H. The complexity of the generalized Lloyd—Max problem (corresp.). *IEEE Trans. Inf. Theory* **1982**, *28*, 255–256. [CrossRef]
60. Kleinberg, J.; Papadimitriou, C.; Raghavan, P. A microeconomic view of data mining. *Data Min. Knowl. Discov.* **1998**, *2*, 311–324. [CrossRef]
61. Aloise, D.; Deshpande, A.; Hansen, P.; Popat, P. NP-hardness of Euclidean sum-of-squares clustering. *Mach. Learn.* **2009**, *75*, 245–248. [CrossRef]
62. Dasgupta, S.; Freund, Y. Random projection trees for vector quantization. *IEEE Trans. Inf. Theory* **2009**, *55*, 3229–3242. [CrossRef]
63. Mahajan, M.; Nimbhorkar, P.; Varadarajan, K. The planar k-means problem is NP-hard. In Proceedings of the WALCOM: Algorithms and Computation, Third International Workshop, WALCOM 2009, Kolkata, India, 18–20 February 2009; Das, S., Uehara, R., Eds.; Springer: Berlin, Germany, 2009; pp. 274–285, ISBN 978-3-642-00201-4. [CrossRef]
64. Zhao, Y.; Pan, J.Z.; Thomas, E.; Jekjantuk, N.; Ren, Y. Ontology languages and description logics. In *Ontology-Driven Software Development*; Pan, J.Z., Staab, S., Aßmann, U., Ebert, J., Zhao, Y., Eds.; Springer: Berlin, Germany, 2013; pp. 51–68. ISBN 978-3-642-31225-0. [CrossRef]
65. Arthur, D.; Vassilvitskii, S. How slow is the k-means method? In Proceedings of the Twenty-Second Annual Symposium on Computational Geometry (SCG'06), Sedona, AZ, USA, 5–7 June 2006; The Association for Computing Machinery: New York, NY, USA, 2006; pp. 144–153, ISBN 1-59593-340-9. [CrossRef]
66. Knuth, D.E. Big Omicron and big Omega and big Theta. *ACM SIGACT News* **1976**, *8*, 18–24. [CrossRef]
67. Jakimowicz, A.; Rzeczkowski, D. Municipality digital platforms and growth poles in the Warmia and Mazury region: Implications for smart specialization. In *Partnerships for Regional Innovation and Development: Implementing Smart Specialization in Europe*; Gancarczyk, M., Ujwary-Gil, A., González López, M., Eds.; Routledge: New York, NY, USA, 2021.
68. Perroux, F. *L'Économie du XXe Siècle*; Presses Universitaires de France: Paris, France, 1964.
69. Crabbé, P. François Perroux et Ilya Prigogine: Systèmes complexes et science économique. *Études Int.* **1998**, *29*, 405–421. [CrossRef]
70. Pottier, P. Axes de communication et développement économique. *Rev. Économique* **1963**, *14*, 58–132. [CrossRef]
71. Geyer, H.S. The terminology, definition and classification of development axes. *SA Geogr.* **1989**, *16*, 113–129.
72. Chodkowska, W.; Sabljak-Olędzka, M. *O Wiatrakach Warmii i Mazur i Młynarzu z Daleka...*; Muzeum Budownictwa Ludowego—Park Etnograficzny w Olsztynku: Olsztynek, Poland, 2016; ISBN 978-83-943479-4-9.
73. Achremczyk, S. *Historia Warmii i Mazur*; Ośrodek Badań Naukowych im. Wojciecha Kętrzyńskiego w Olsztynie: Olsztyn, Poland, 2010–2011; Volume 1: Pradzieje–1772, ISBN 978-83-60839-33-1. Volume 2: 1772–2010, ISBN 978-83-60839-46-1.
74. Małłek, J. *Dwie Części Prus. Studia z Dziejów Prus Książęcych i Prus Królewskich w XVI i XVII Wieku*; Wydawnictwo Pojezierze: Olsztyn, Poland, 1987; ISBN 83-7002-302-9.
75. Krzyżanowski, P. Koniec Budowy sieci Szerokopasmowej na Warmii i Mazurach. *Komputer Świat* **2015**. Available online: https://www.komputerswiat.pl/artykuly/redakcyjne/koniec-budowy-sieci-szerokopasmowej-na-warmii-i-mazurach/6pxxn4m (accessed on 17 May 2020).
76. Sieć Szerokopasmowa Polski Wschodniej. Regionalny Portal Informacyjny Wrota Warmii i Mazur. Available online: https://www.warmia.mazury.pl/rozwoj-regionu/informatyzacja/158-siec-szerokopasmowa-polski-wschodniej (accessed on 18 May 2020).
77. Dorfman, R.; Samuelson, P.A.; Solow, R.M. *Linear Programming and Economic Analysis*; Dover Publications: New York, NY, USA, 1987; ISBN 0-486-65491-5.
78. McKenzie, L.W. Turnpike theory. *Econometrica* **1976**, *44*, 841–865. [CrossRef]

79. Plan Cyfrowy 2025 dla Warmii i Mazur—Stan na 6 Marca 2017. Regionalny Portal Informacyjny Wrota Warmii i Mazur. Available online: https://warmia.mazury.pl/rozwoj-regionu/informatyzacja/165-plan-cyfrowy-2025-dla-warmii-i-mazur-stan-na-6-marca-2018 (accessed on 19 May 2020).
80. O rozwoju sieci „Ostatniej Mili" na Warmii i Mazurach. Regionalny Portal Informacyjny Wrota Warmii i Mazur. Available online: https://warmia.mazury.pl/rozwoj-regionu/informatyzacja/162-o-rozwoju-sieci-ostatniej-mili-na-warmii-i-mazurach (accessed on 20 May 2020).
81. Descartes, R. Progymnasmata de solidorum elementis. In *Oeuvres de Descartes. Physico-Mathematica, Compendium Musicae, Regulae ad Directionem Ingenii, Recherche de la verité, Supplément à la Correspondance*; Adam, C., Tannery, P., Eds.; Léopold Cerf, Imprimeur-Éditeur: Paris, France, 1908; Volume 10, pp. 265–276.
82. Cromwell, P.R. *Polyhedra*; Cambridge University Press: Cambridge, UK, 1997; ISBN 0-521-55432-2.
83. Wilson, R.J. Graph theory. In *History of Topology*; James, I.M., Ed.; Elsevier B.V.: Amsterdam, The Netherlands, 1999; pp. 503–529. ISBN 0-444-82375-1.
84. Biggs, N.L.; Lloyd, E.K.; Wilson, R.J. *Graph Theory, 1736–1936*; Clarendon Press: Oxford, UK, 1998; ISBN 0-19-853916-9.
85. Euler, L. Elementa doctrinae solidorum. *Novi Comment. Acad. Sci. Imp. Petropol.* **1752**, *4*, 109–140.
86. Euler, L. Demonstratio nonnullarum insignium proprietatum quibus solida hedris planis inclusa sunt praedita. *Novi Comment. Acad. Sci. Imp. Petropol.* **1758**, *4*, 140–160.
87. Sandifer, C.E. *How Euler Did It*; The Mathematical Association of America: Washington, DC, USA, 2007; ISBN 978-0-88385-563-8.
88. Richeson, D. The polyhedral formula. In *Leonhard Euler: Life, Work and Legacy*; Bradley, R.E., Sandifer, C.E., Eds.; Elsevier B.V.: Amsterdam, The Netherlands, 2007; pp. 421–439. ISBN 978-0-444-52728-8.
89. Lakatos, I. *Proofs and Refutations: The Logic of Mathematical Discovery*; Worrall, J., Zahar, E., Eds.; Cambridge University Press: Cambridge, GB, USA, 1976; ISBN 0-521-21078-X.
90. Cauchy, A.L. Recherches sur les polyèdres, première et seconde parties. *J. Éc. Polytech.* **1813**, *9*, 68–86.
91. Kogut-Jaworska, M.; Ociepa-Kicińska, E. Smart specialisation as a strategy for implementing the regional innovation development policy—Poland case study. *Sustainability* **2020**, *12*, 7986. [CrossRef]
92. Gell-Mann, M. *The Quark and the Jaguar: Adventures in the Simple and the Complex*, 8th ed.; W.H. Freeman and Company: New York, NY, USA, 2002; ISBN 0-7167-2725-0.
93. Casti, J.L. *Reality Rules: I. Picturing the World in Mathematics: The Fundamentals*; A Wiley-Interscience Publication: New York, NY, USA, 1997; Volume 1, ISBN 0-471-57021-4.
94. Jakimowicz, A. Catastrophes and chaos in business cycle theory. *Acta Phys. Pol. A* **2010**, *117*, 640–646. [CrossRef]
95. Jakimowicz, A.; Kulesza, S. The mechanism of transformation of global business cycles into dynamics of regional real estate markets. *Acta Phys. Pol. A* **2018**, *133*, 1351–1361. [CrossRef]
96. Poston, T.; Stewart, I. *Catastrophe Theory and its Applications*; Dover Publications: Mineola, NY, USA, 1996; ISBN 978-0-486-69271-5.
97. Rosser, J.B., Jr. *From Catastrophe to Chaos: A General Theory of Economic Discontinuities: Mathematics, Microeconomics, Macroeconomics, and Finance*, 2nd ed.; Springer Science + Business Media: New York, NY, USA, 2000; Volume 1, ISBN 978-94-017-1615-4. [CrossRef]
98. Wilson, A.G. *Catastrophe Theory and Bifurcation. Applications to Urban and Regional Systems*; Routledge: New York, NY, USA, 2011; ISBN 978-0-415-68782-9.
99. Zeeman, E.C. On the unstable behaviour of stock exchanges. *J. Math. Econ.* **1974**, *1*, 39–49. [CrossRef]
100. Arnold, V.I. *Catastrophe Theory, 2nd*; Springer: Berlin, Germany, 1986; ISBN 3-540-16199-6.
101. Stewart, I.N. Catastrophe theory. *Math. Chron.* **1977**, *5*, 140–165.
102. Thom, R. *Structural Stability and Morphogenesis. An Outline of a General Theory of Models*; W.A. Benjamin: Reading, MA, USA, 1975; ISBN 0-8053-9276-9.
103. Trotman, D.J.A.; Zeeman, E.C. The classification of elementary catastrophes of codimension ≤ 5. In Structural Stability, the Theory of Catastrophes, and Applications in the Sciences. In *Proceedings of the Conference Held at Battelle Seattle Research Center 1975*; Hilton, P., Ed.; Springer: Berlin, Germany, 1976; pp. 263–327. ISBN 3-540-07791-X.

104. Zeeman, E.C. Applications of catastrophe theory. In *Manifolds-Tokyo 1973: Proceedings of the International Conference on Manifolds and Related Topics in Topology, Tokyo, 1973*; Hattori, A., Ed.; Published for the Mathematical Society of Japan by University of Tokyo Press: Tokyo, Japan, 1975; pp. 11–23. ISBN 978-0860081098.
105. Gell-Mann, M.; Lloyd, S. Information measures, effective complexity, and total information. *Complexity* **1996**, *2*, 44–52. [CrossRef]
106. Gell-Mann, M.; Lloyd, S. Effective complexity. In *Nonextensive Entropy: Interdisciplinary Applications*; Gell-Mann, M., Tsallis, C., Eds.; Oxford University Press: New York, NY, USA, 2004; pp. 387–398. ISBN 0-19-515976-4.
107. Jakimowicz, A. The role of entropy in the development of economics. *Entropy* **2020**, *22*, 452. [CrossRef]
108. Richardson, L.F. *Statistics of Deadly Quarrels*; Wright, Q., Lienau, C.C., Eds.; The Boxwood Press: Pittsburgh, PA, USA, 1960; ISBN 9780910286107.
109. Richardson, L.F. The problem of contiguity: An appendix to statistics of deadly quarrels. In *General Systems: Yearbook of the Society for General Systems Research*; Rapoport, A., von Bertalanffy, L., Meier, R.L., Eds.; Published by the Society for General Systems Research: Ann Arbor, MI, USA, 1961; Volume 6, pp. 139–187.
110. Smith, R.P. The influence of the Richardson arms race model. In *Lewis Fry Richardson: His Intellectual Legacy and Influence in the Social Sciences*; Gleditsch, N.P., Ed.; Springer International Publishing: Cham, Switzerland, 2020; pp. 25–34. ISBN 978-3-030-31588-7. [CrossRef]
111. Mandelbrot, B.B. *The Fractal Geometry of Nature*; W.H. Freeman and Company: New York, NY, USA, 1983; ISBN 0-7167-1186-9.
112. Mandelbrot, B.B. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science* **1967**, *156*, 636–638. [CrossRef]
113. Ashford, O.M. *Prophet—or Professor? The Life and Work of Lewis Fry Richardson*; Adam Hilger: Bristol, UK, 1985; ISBN 0-85274-774-8.
114. Hayes, B. Computing science: Statistics of deadly quarrels. *Am. Sci.* **2002**, *90*, 10–15. Available online: https://www.jstor.org/stable/27857587 (accessed on 16 April 2020). [CrossRef]
115. Bryant, J. *Entropy Man*; VOCAT International Ltd.: Harpenden, UK, 2015; ISBN 978-0-9562975-4-9.
116. Warnecke, H.-J. *The Fractal Company: A Revolution in Corporate Culture*; Springer: Berlin, Germany, 1993; ISBN 978-3-642-78126-1. [CrossRef]
117. Hoverstadt, P. *The Fractal Organization: Creating Sustainable Organizations with the Viable System Model*; John Wiley & Sons: Chichester, UK, 2008; ISBN 978-0-470-06056-8.
118. Malik, P. *The Fractal Organization: Creating Enterprises of Tomorrow*; Sage Publications, Inc.: Thousand Oaks, CA, USA, 2015; ISBN 978-93-515-0244-9.

entropy

MDPI

Concept Paper

What Motivates Speculators to Speculate?

Bedane S. Gemeda [1], Birhanu G. Abebe [1], Andrzej Paczoski [2], Yi Xie [3] and Giuseppe T. Cirella [2,*]

[1] Ethiopian Institute of Architecture, Building Construction and City Development, Addis Ababa University, 120 Addis Ababa, Ethiopia; bedanes@yahoo.com (B.S.G.); mesi.bire@gmail.com (B.G.A.)
[2] Faculty of Economics, University of Gdansk, 81-824 Sopot, Poland; andrzej.paczoski@ug.edu.pl
[3] School of Economics and Management, Beijing Forestry University, No. 35, Qinghua East Road, Haidian District, Beijing 100083, China; yixie@bjfu.edu.cn
* Correspondence: gt.cirella@ug.edu.pl

Received: 21 November 2019; Accepted: 28 December 2019; Published: 31 December 2019

Abstract: Land speculation that occurs on the urban border can be very problematic to the healthy development of cities—critical to economic growth. Speculative land investors, concerned with profits from trading in landed property, can especially affect developing countries where regulation is often poorly controlled and overly bureaucratic. An investigation into the factors motivating land speculators operating in the urban fringe of the city of Shashemene, Ethiopia is examined. The paper, in addition to contributing to the literature, is the second-known attempt and extension of the authors' pilot research to study the behavior of land speculators in the urban fringe of a growing Ethiopian city. A theoretical framework and conceptual breakdown are put together with historical reference to early land speculation examples. Two questionnaires were separately administered with a representative random sample of 159 members from the local land developer association (i.e., investors) and 24 senior officials from the study area. A principal component analysis categorized the most significant dynamics in controlling land speculation procurements. Results indicated motivational reasoning as the prime cause for speculative activities. Evidence indicated that land speculation is a critical dynamic for self-worth especially with business-oriented persons. Entropy, the disorder of the communicative data, suggests a possible rethinking of the way government should intervene in the urban property market. As such, developmental smart cities in Ethiopia must thoroughly consider the dynamisms of speculative activities and its effects on local housing as it moves forward–in the 2020s.

Keywords: speculation; land acquisition; motivation; real estate; development; Ethiopia

1. Introduction

Land speculation is regarded as a critical issue in Ethiopia; thus, it is most welcome to enact the 1993 Land Use Act to address the land speculation problem and curb the incidence of growing land prices that are largely due to speculator activity. Nonetheless, this issue is far from being completely resolved as the issue of major acquisitions and land scarcity is still prevalent in many of the country's urban and urban fringe areas [1,2]. Land management is essential in retrospect of city planning, land use development, and land speculation to create a successful cyclic system. Speculative land investors are those concerned with income from landed property trading, rather than with its use or earning capacity after growth, particularly in regard to housing which appears to be the highest and best use of the urban border [3–5]. This clearly implies that almost any investment aimed at a rapid accumulation of assets may be regarded as speculative [6]. Non-speculative investors, on the other hand, are those concerned with deriving profits from either property development or cash flow earnings received as a long-term investment from holding land [7,8]. Siegel [7] has stated that speculative investment can lead to an overvaluation of assets (i.e., securities or real estate) within a given area, which can often lead to an effect of property bubble-and-burst. Mohamed [9] described speculative activities as a form

of fulfillment, a set of sub-optimal targets for which the desire of investors among developers has been widely observed. This occurrence tends to become a continuum as the policy responses needed to turn aside or alleviate a tightened situation are not as readily effective as anticipated. The main outcome of satisfying manners, as argued by Mohamed [9], are that many developers prefer projects located on green field sites (i.e., amenity-based or agricultural land) that take a relatively short time to dispose of in order to recoup the invested capital at a practical level of profit. This type of economic resolve can be more readily seen as urban building encroaches outwardly onto the urban border [10,11].

According to Gemeda, et al. [12], Shashemene's much-acclaimed deficit of over 1000 housing units is projected to intensify the need for urgent land acquisition and growth to meet the current deficit. Nevertheless, some researchers (e.g., Golland and Boelhouwer [13] and Wang and Hua [14]) proclaim that developers' land acquisition is not a specific response to demand; rather, developers take advantage of the general level of housing market activity and, in particular, "data signals" [13]. The State's position in controlling land acquisition appears to affect the motivation of acquisitions by speculators, either due to the fact that land can provide a source of profit through land banking or land trading or as a development factor. Recently, from observation, investment activities tend to have increased in frequency due to rising urban populations and anticipation and hype surrounding urban fringe areas (i.e., mostly as a result of new road infrastructure). Other drivers of speculative activities include close proximity of bare land to an urban center, excitement and hype surrounding the location's commercial viability, overestimation of land demand, and the notion that the location is a viable growth investment area within the development plan of the city. Eventually, speculators may buy the land and use it immediately for a completely different purpose, which defies logic in order to lay claim to it, such as establishing a mechanical factory, operating a vegetable farm, erecting a perimeter fence, or building on-site quarters for servants (e.g., security guards). These are some provisional activities frequently carried out by land speculators after valid purchase in order to "legitimize" their proprietorship of the land in line with Ethiopia's jurisdictional definition of development.

Dealings in the land market can be an indication of the idiosyncrasy of the environment as well as the area where it is situated. Two market types exist as both central and peri-urban [15,16]—classified as sale and rental markets. According to the World Bank [16], a sales market is when freehold interest can be sold while a rental market refers to when "usage rights" are temporarily reassigned for a limited period. While land speculators are inclined towards the sales market for transferability of freehold interest, it should be noted that many societies also allow for both partial and total disposal of property leasehold interest. This study, in an umbrella-like effect, oversees such dynamics motivating land speculators operating in Shashemene's urban fringe. As such, Ethiopia's Proclamation No. 721/2011 of its Land Use Act allows a lessee to assign their unexpired interest in land with consent from the State; thus, allowing the market to be either formal and structured or informal and operating extra-legally [17–19]. A formal land market follows a set structured rules and regulations while participation in the market abides by legal procedures both for acquisition and disposal. On the other hand, the informal market often outweighs land nationalization since it is not always fixated with strict pertinent rules commonly found in many developing countries where property rights are often constricted [20–22].

In the city of Shashemene, the legal occupier of any land will not only have lawful title to it (i.e., a freehold, leasehold, or license from the actual landowner), but also need formal planning approval for any buildings or other erected structures permanently affixed to the land, including land use of which those structures pertain to. As defined by the Business Dictionary [23], development is "the carrying out of the building, engineering, mining, or other operations in, on, over, or under land, or the making of any material change in the use of any buildings or other land." According to the statutory definition, building operations refer to the demolition of buildings, rebuilding, structural alterations of, or additions to buildings and other operations such as "excavation or other works carried out on the land, including fencing" [17]. Thus, in accordance with Ethiopia's Land Use Act, it is important to note that, since a land speculator can effectively prevent their land from being repossessed by the

original seller (i.e., for failing to develop within two years), they "employ somewhat deceptive-like activity" [17]. Experience indicates that unprincipled speculators can ensue and process a genuine (i.e., legal) land title and acquire planning permission for a structure or improvement (i.e., even an ordinary fence) all in an attempt to retain legal ownership and procure capital appreciation.

The motivation of the research is focused on property market dynamism. It increasingly has become intertwined with speculative financial flows and has shaped (and reshaped) urban land under the stimulus of capital over accumulation with the intention of absorbing surplus (i.e., a process that ironically can be a source of economic macro crises as well as a share of the property market as gross domestic product increases). In this context, Ethiopia is certainly not excluded from urban development and speculation despite its ever-increasing level of urbanization—especially since the 1960s. In light of this, land speculators are harvesting huge revenue from urban and peri-urban areas by keeping land vacant. Land speculation in this regard is seen as a critical issue in Ethiopia; hence, it is appropriate to enact the Land Use Act to address the problem and curb the incidence of rising land prices that by in large spike due to speculative activity. However, this issue is far from being fully resolved as the problem of large acquisitions and land deprivation is still endemic in throughout the country's urban fringes [12]. To resolve this problem and increase the benefit to society, this study aimed at looking into the motivational factors behind land speculators using Ethiopia as a case study. We look at the motives that motivate land speculators in the urban fringe with the position of making policy recommendations in line with speculation-based best practices. A breakdown of the paper is structured as follows: Section 2 explores the theoretical frame and conceptual clarification of land speculation, Section 3 contains the methodology, Section 4 elucidates the results and discussion, and Section 5 provides the conclusions.

2. Theoretical Frame and Conceptual Clarification

The theoretical frame is provided by insights gleaned from the classical concentric zone model [24], sector or axial development model [25], and Alonso's [26] bid rent model—all which seek to clarify why detailed land use is situated where they are in the cityscape. Burgess' [24,27] model portrays residential land use structures by examining how one might plan the layout of a city [28]; out-of-date, it overlooks the importance of transport routes, site, and physical uniqueness in shaping a city's evolution and urban border zones. In terms of the concentric model, weaknesses are remedied by land use zones that focalize on key urban areas associated along central arterial transportation lines. From this theoretical aspect, and as already identified in the analysis part by principal component analysis, land speculators are motivated by regulatory lapses, location preference, informality of land title, and inexpensive land. Those variables vary depending on the model. For instance, within the context of transportation routes, where infrastructure is usually found, speculators are motivated to hoard more land. Moreover, as described by Hoyt [25,29] in which the major land use zones in the urban area is "aligned along the major arterial transportation routes" [29], land speculators in parallel follow infrastructure (i.e., transportation, roads, etc.) when modeling city advancement. Likewise, with axial theory, current major transport facilities also control urban development via two factors: speed and pattern. This has meant that, while an area may be remote, once it has direct access to major roads (i.e., in terms of time taken to and from the central business district), a transition zone is likely to materialize which, potentially, can act as a target for land speculators. In practice, however, the impacts of externalities such as rising urban population, rising demand for land, and the desire for business profit can oblige land speculators to consider holding the land for a block of time in order to ensure its appreciation—that is, introducing the notion of "mastering the silent game" [30] of a risk-return trade-off in order to break-even (Figure 1).

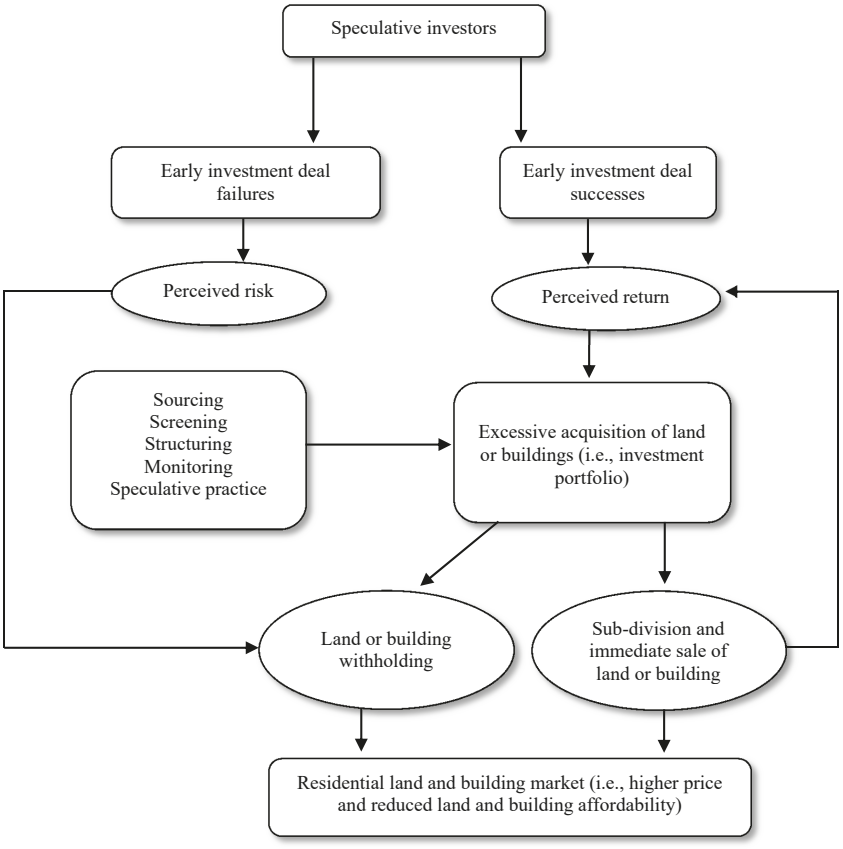


Figure 1. A cognitive model of speculative investors' operations in property sectors, adapted from Valliere and Peterson [31].

Conceptually, land speculation may be defined as a function of the ineptitude of land policies and the low believability of land transactions taking place in an urban tassel. For instance, land policies and their efficacy rely on urban and regional planning procedures and interrelated policies and strategies that regulate and control the economic and social practices on property and land assignment. This is achieved through proper enforcement of "building regulations, planning standards, and the zoning bylaw guiding [specific land uses and] infrastructure provision" [32], the lack of which may result in confounded and poor development patterns along the urban border. Accordingly, best practices would make certain that planning tools are judiciously practical as both precautionary and counteractive measures for ensuring proper and effective land use deter sub-standard, natural, and chaotic development and pointless delays in project development completion. As a result, the physical, spatio-temporal and social uniqueness of a community are vital criteria for thoughtfulness in reforming the urban border [32]. A follow-up notion, within this framing, is land transactions (i.e., mortgage transactions that use property as collateral for bank loans). Specifically, land transactions experiencing changes regarding their arbitrative status, beginning from "an intention to acquire and proceed to the procurement and certification of land title" [33]. Urban fringe land is often perceived as marginal land, available but "uncultivated" [34] and "suitable for investments" [35] other than agriculture [36,37]. As a result, its availability in an uncultivated, large-scale form—outside the scope of agriculture—prompted the conceptual framing of Figure 2. As such, the real cause of speculation is the specified expectation

of the augmentation of land value, which occurs in all advancing countries from the stable increase of rent, which leads to speculation, or the "holding of land for a higher price than it would then otherwise bring" [38]. In contrast, the consequences of land speculation are "tenantry and debt to farms, and slums to luxury in cities" [39].

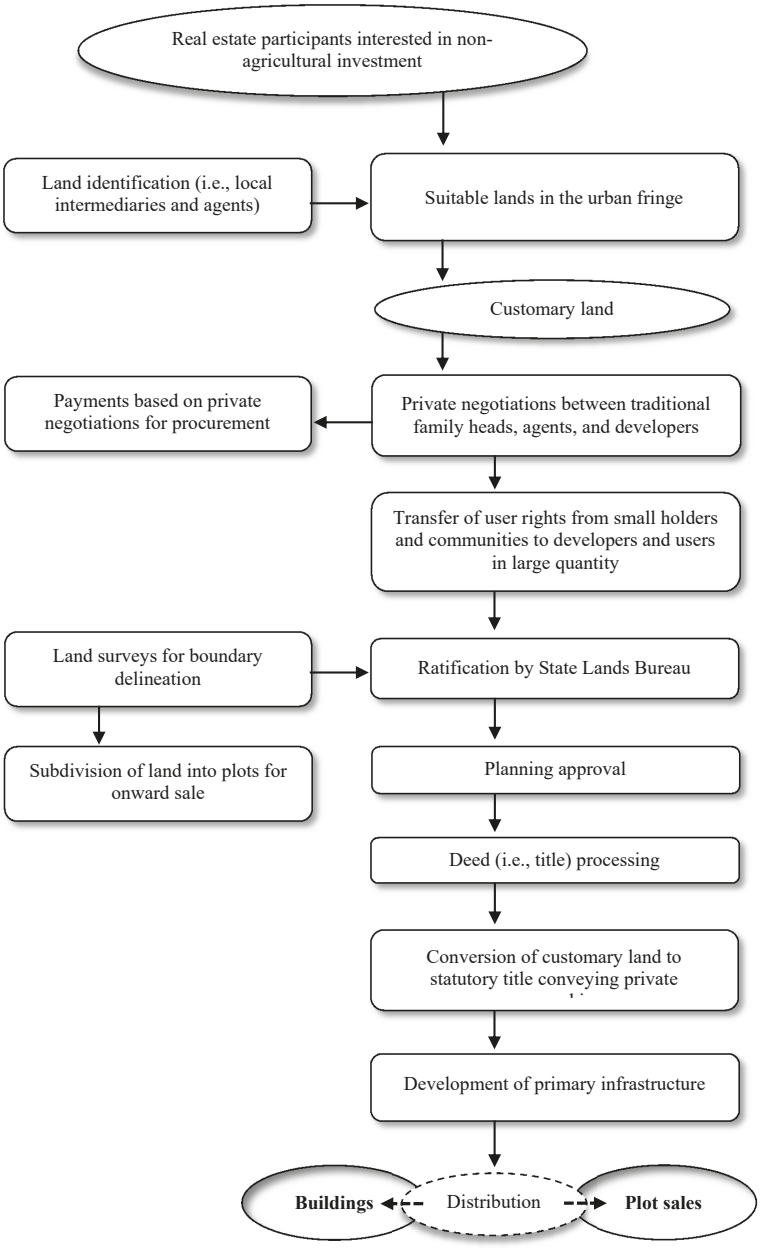


Figure 2. Land acquisition and legalization procedures in Shashemene.

Large-scale land hoarding, throughout the developing world, stresses competition for land and results in land disputes due to escalating food prices and the necessity to cultivate more land for agriculture, "uncontrolled rural-urban immigration leading to higher population densities" [40], and increasing demand for shelter and other forms of accommodation [3,33,40,41]. According to Colin [42], numerous large land acquisitions along the urban border involve prolonged negotiation processes which revolve around hoarding of long-term rights of ownership. In Ethiopia, long leaseholds of 99 years or other agreeable titles are only decided by the State, in observance with the provisions of the Land Use Act No. 4 of 1993. The Act, as observed by Hurni, et al. [43], was promulgated following a report submitted to the Federal Government by which a land use panel examined the current system of land occupancy and advocated measures for curbing the activities of land speculators, "by making it easier for the government to acquire land" [43] for development projects nationwide, including the capital. Typical large land acquisition involves identification, negotiation, procurement, title, and distribution or subdivision. Suitable privately-owned land is first recognized through consultation with local intermediaries (i.e., traditional authorities), in which mostly local agents act as intercessors. In Shashemene's urban fringe, some developers with the support of local go-betweens (i.e., partners) instigate direct communication with local traditional chiefs (i.e., including heads of families) when they wish to acquire land which usually instigated via a back-and-forth. It should be noted that many land deals in Ethiopia's urban tassel are classified as "semi-formal and customary" [42] rather than formal tenure arrangements—prevalent in many African countries [44,45].

After settling the contracted purchase agreement, the owner is directed to submit it to the Lands Bureau for registration of title. Additional documents that must be succumbed including a survey plan indicating the beacons and boundaries of the land, copy of the approved building plan, and, if necessary, an environmental impact assessment report on the proposed project for the site. Once the submission process is ratified, land and building certification can be issued. Recently published research by Gemeda, et al. [12] is a first attempt to piece together some of the roles land speculators pose within the country. This paper acts as an expansion to those findings by furthering the research and nature of land speculation in Ethiopia's major cities (i.e., in respect to urban and urban fringe areas). A core focus is to make policy recommendations in the interest of city livability and sustainability.

Speculators, according to Andreasson, et al. [46], are likened to "high-risk traders almost of the same kind to gamblers" [46], whereas lower risk investments grounded on core research and analysis fall into the category of desirable investments. Subject to an in-depth analysis, a good investment should "swear the safety of the capital invested in addition to ample income and yield return" [46]. On the other hand, investment processes not up to scratch with these requirements are termed as speculative in nature. In terms of systemic risk, the possibility of an event triggering severe instability or collapse of an economy, in retrospect, plays with the speculative idea of "too big to fail". Often systemic risk can be used as a justification for government to intervene in the economy. The basis for this intervention is the belief that government can reduce or minimize the ripple effect during turbulent economic times [47,48]. As such, an important framing of variable risk in relation the number of properties owned (i.e., held) is correlative to total risk which starts off as non-systemic risk and augments to systemic risk as owned properties numbers increase (Figure 3). Sometimes, however, it has been found that non-intervention in the economy can be beneficial solely based upon the fact that if variable risk is too volatile, not doing anything (i.e., leaving the land market to sort itself) is best. This is more often the exception than the rule since it can destabilize the land market, more than projected, due to speculator sentiment [47–50].

Investor characteristics differ from those of a speculator, in so that, an investor "acquires land as a factor of production" [51], whereas a speculator attains land in the hope of profiting from an upsurge in its market value (i.e., capital appreciation at the end of a holding period). Speculation is not caused by a shift in demand due to change in taste, fashion, consumer need, or supply. As argued by Knittel and Pindyck [51], a shift in fundamentals can frequently trigger a transformation in price, an outcome not necessarily a determinant of land market speculation. As such, speculative demand and pressure

can force changes in the property market due unavoidable external influences that simultaneously structure urban fringe areas regardless of government policy.

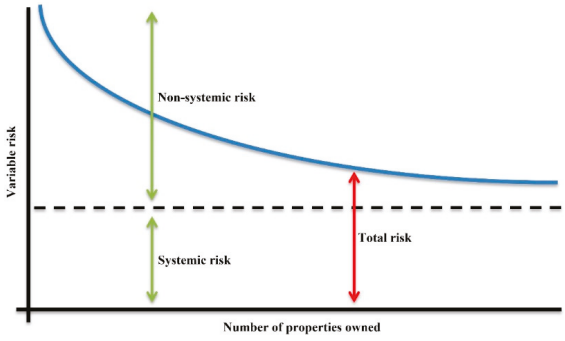


Figure 3. Variable risk versus number of properties owned as a measure of non-systemic and systemic risk (i.e., total risk).

Participants involved in speculative practices are grouped into three classified levels by Triantafyllopoulos [52]. First, there are those who are "informed speculators" [52] in the sense that they have access to both public and private information. They are otherwise known as public officials in charge of land allocation (i.e., plots and buildings). Second, there are those who are "un-informed speculators" [52]—these are persons who only have access to public information and are otherwise known as either investors or developers. Third, there are "private purchasers" [52] who are individuals that are not information-driven. In practice, the different functions played by each of the actors can particularly alter the cityscape and overall layout of how cities interact. For example, developers may be regarded as investors who initiate and carry out land development projects which can play an active and leading role in the development of the urban fringe, perceived as the prime "sculptors of spatial structures" [53] of any open space. Jonas and Wilson [54], Logan and Molotch [55], and Molotch [6] refer to these developers as "growth machines" and directly link speculative practices with the conditions for how cities grow.

Generally, land speculation depends on endogenous and exogenous factors. Endogenous factors can be related to the institutional framework of any given country (i.e., property rights protect by law, enterprises dealing in land markets, and types of land markets—e.g., regulated or free). Exogenous factors are much more complex. They reflect the international position of any given country, situation, and perspective of its economy, political stabilization, social and human capital (i.e., its approach and control of nationally-owned land), and foreign investment [56–58]. These influences forecast concerns and potential land and building prices as well as project an overall perception on the type of land speculation at the city-level. As such, popular and scholarly housing debates, for example in London, are concentrated on the super-rich as stated by Atkinson [59], Hay [60], and Hay and Muller [61] in terms of patterns of consumption, economic power, and political control [62,63]. Moreover, parks are considered safe havens for park property [64] and produce rental income that often outgrows inflation compared to lower interest rates on savings accounts and weaker returns on investment of financial products [65]. There is evidence that, since the early 1990s, middle-class buyers from Hong Kong have been investing in London's main property market as noted by Ho and Atkinson [65]. Although Hong Kong's first wave of investors primarily focused on property for their own use, second and third wave investors emerging between 2000 and 2009 were interested in financial returns (i.e., short-term investment or long-term rentals). The scale of this investment activity has become somewhat significant since Hong Kong shareholders, by 2012, have already purchased about one in six new residential properties sold in central London [66]. Significantly, an estimated 514,000 Hong Kongers invested in property outside Hong Kong in 2016 [67], representing 7% of Hong Kong's population at

around 7.4 million [68], a trending figure that continues to increase. Similar issues (e.g., in regards to housing affordability in Sydney, Australia [69–71] and Vancouver, Canada [56,72,73]) parallel the global crisis of urban land speculation practices in which "accelerated (re)urbanization of capital and people [has led to] the provision of cheap credit and the rise of intra-society inequality" [56]. Nonetheless, global real estate is considered a class of capital for which investors "diversify their investment portfolios" [74] often at the expense of local community housing affordability interrelated via political and economic spheres.

3. Methodology

There are six main urban fringes in the greater city area of Shashemene (Figure 4). Of these areas, a number of peri-urban villages (i.e., Awasho, Arada, Abosto, and Bulchana) exist. These peri-urban areas stretch adjacently along the Awasho asphalt road in the east, Bulchana in the southeast, and Alelu, Dida-Boke, and Arada in the neighboring areas of the northwest. The Awasho urban fringe in the southeast was selected as the study site due to its fast-growing hub-like industrial and residential activities. Awasho is also a major transportation route currently coping with the impact of rapid urbanization where investors and foreign industrialists are favoring to locate. Pressure and demand for residential land, even though it is unaffordable due to unnecessary land withholding and price increase (i.e., created by land speculators), is seen to be physically available.

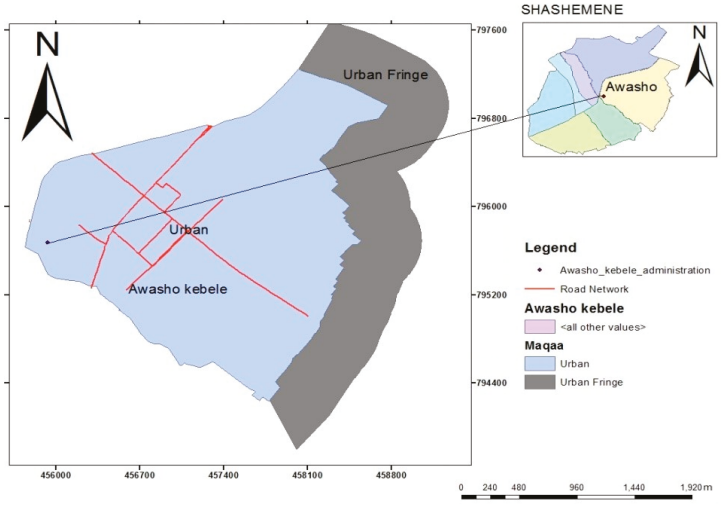


Figure 4. Map of the study area.

Primary data were collected through a combination of methods including self-administered questionnaires, direct observations, and in-depth interviews, using two types of questionnaires. The first set of questionnaires was randomly given out to the local land developer association (i.e., investors) in which the sampling frame was 159 investors—confirming the sample size determination was free of bias or error of misclassification [75,76]. The structure focused on the motivational dynamics for large-scale land purchases, reasons for their holdings (i.e., after procurement), period of time between procurement and actual development (i.e., if any), land status (i.e., procured by developers), legal title, and period between procurement and further subdivision (i.e., for onward sale without further development).

The second set of questionnaires was directed at 24 senior officials of the Shashemene City Land Agency who reported upon whether any effective policies were put in place for limiting (i.e., controlling) developer activities with a tendency towards speculative activity. Further inquiry looked at what approaches and policies had been used as well as their success rate. The samples represented

the target population and not the total population. To regulate the average mean of the developers, we employed a five-point Likert-type rating scale that considered cause and control of land speculation city-wide. Reliability of the questionnaire included a Guttman Split-half test to control for internal consistency. To end, principal components analysis (PCA) was exercised on vital factors germane to large land procurements. Specifically, the PCA method used orthogonal transformation to convert the set of land acquisition variables into a set of values of principal components.

4. Results and Discussion

4.1. Factors Motivating Land Speculators

Essentially, land speculation is holding land primarily to meet future demand and not present need. The aim of the speculator is to create a synthetic value that in itself is unacceptable. Recent activities leading to large-scale land acquisitions and withholding in many urban fringes have triggered price rises in the market, intensified the inelasticity of supply of developable land, and complicated the already disorganized housing issue (i.e., affordability for low-income earners). The provision of affordable housing cannot be left exclusively to private developers as government has a major role in ensuring equity and effectiveness in the allocation of land resources by way of externalities [75]. In analyzing the motive behind speculative activities and land use control in the urban and urban fringe of Shashemene, the following dimensions were discussed as a means of contributing to knowledge: prominent characteristics of land acquisition, range of time between actual land acquisition and development by speculators, and factors responsible for speculative land acquisition.

4.2. Prominent Characteristics of Land Acquisition

Respondents to the land developer questionnaire were predominately private developers (i.e., 85%). The status of land bought by developers on the average shown to be left vacant and underutilized (i.e., undeveloped) was more than seventeen years. Moreover, only four of the lands bought have adequate legal title in the form of the governor's consent. Others either have an illegal receipt of purchase or have been registered in the State gazette as evidence of government approval, while others do not even have any certification. Some of the developers indicated that they usually subdivide their land immediately after purchase (i.e., 25.2%), while 65.6% subdivide within a two-year period. This suggests most of the respondents did not have an aim in commencing concrete development at the time of purchasing land. The location of acquired land is spread over the whole of the study area with respect to individual developers' needs and requirements, usually dependent on accessibility and proximity to infrastructural services.

The oldest land was 400 m^2 purchased in 1991 that has remained vacant for 29 years since its purchase. The most recent and largest purchase of 6900 m^2 was done in 2015 with government consent, but has remained vacant ever since. The cheapest acquisition was secured from a family by private treaty (i.e., negotiation) for 875 Ethiopian Birr (i.e., US$ 30.13) per m^2, while the costliest acquisitions closed at 5000 Ethiopian Birr (i.e., US$ 172.18), backed by government consent. Only four of the acquired lands are either fully or partially developed for housing (Table 1). As such, it can be inferred that developers (i.e., local speculators) hoarding land and taking it off market for long periods of time also have a relational connection to housing rental income and the overall housing market. As it is commonly portrayed in urban economics, the demand for land is driven by the demand for housing [56,63]. As the price of land increases, the price of housing follows and vice versa, which was commonly seen in Ethiopia's land market. Moreover, land markets would thus have a positive impact on improving land access by land-poor households. As long as imperfections affect only one market, all relevant actors still have the opportunity to cultivate the same amount of land per capita. However, credit market imperfections can offset or even eliminate supervision cost advantages of family farmers.

Table 1. The status of lands bought by speculative developers.

Land Bought (m^2)	Location	Year Purchased	Year Developed	† Price per m^2	Document Title When Purchased	Status	Vacant Period (Years)
160	Awasho	1997	not developed	2000	illegal receipt	Vacant	23
220	Awasho	1998	not developed	1400	illegal receipt	Vacant	22
320	Awasho	2000	not developed	1000	illegal receipt	Vacant	19
350	Awasho	2001	not developed	2500	illegal receipt	Underutilized	18
200	Awasho	2004	not developed	4000	illegal receipt	Vacant	16
6900	Awasho	2015	not developed	5000	legal receipt	Vacant	5
350	Alelu	2001	not developed	4000	family receipt	Vacant	19
250	Alelu	2005	not developed	3500	illegal receipt	Vacant	15
140	Alelu	2013	not developed	2500	legal receipt	Vacant	7
500	Arada	1997	not developed	3500	illegal receipt	Vacant	23
450	Arada	1995	2013	4500	illegal receipt	developed	19
400	Bulchana	1991	not developed	875	family receipt	Underutilized	29
320	Bulchana	1998	2016	3500	illegal receipt	Developed	19
140	Dida Boke	1997	2005	4500	illegal receipt	Developed	9
250	Dida Boke	2000	2014	3500	illegal receipt	Developed	15
250	Dida Boke	2004	not developed	980	legal receipt	Vacant	16
200	Kuyera	2011	not developed	1050	legal receipt	Vacant	9

† US$ 1.00 = 28.7 Ethiopian Birr; Source: field survey.

4.3. Range of Time between Actual Land Acquisition and Development by Speculators

In reference to Table 1, the range of time between actual land acquisition and development by speculators shows a revealing reality that the four developed properties bought in 1995, 1997, 1998, and 2000 were developed a whopping 19, 9, 19, and 15 years later, respectively. These numbers correlate with the average, to date, mean status of land bought by developers that has been left vacant and underutilized (i.e., seventeen years). This confirms our earlier misgiving that most of the land speculators in the study area do not seriously intend on undertaking development within the short-term following their acquisition. This ability to hold on to land for long periods of time points at the low risk presented suggests that land speculators tend to delay development in the expectation that a buyer will show up eventually. As such, it has been noted that, when speculators' hope of finding a buyer is reduced, they resort to the legitimization of their ownership by spurious or unscrupulous land enhancements, including the erection of just an ordinary perimeter fence.

4.4. Factors Responsible for Speculative Land Acquisition

Speculators were asked to indicate the factors motivating large land acquisition in which more than 20 reasons were ranked using a Likert scale. Responses obtained were analyzed using PCA to signal reasons for significant factors relevant to large land acquisitions. As such, seven principal components were identified during this process (Table 2). As a widely used statistical technique, PCA is useful for determining the latent variables of the obvious variables [75]. The results of the principal components are sensitive to the relative scaling of the original variables.

Table 2. Principal components matrix of reasons for large-scale land acquisitions ‡.

† Variable	Comp1	Comp2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7	Unexplained
Q1	−0.2115	0.6684	−0.1562	0.1136	−0.1684	−0.6653	0.0144	0
Q2	−0.3811	0.0490	0.4494	−0.1927	0.7660	−0.1616	0.0190	0
Q3	−0.1903	0.1367	0.7892	0.2840	−0.4580	0.1774	0.0201	0
Q4	−0.2884	0.5696	−0.2687	0.0696	0.1431	0.7032	0.0192	0
Q5	0.2873	−0.0128	−0.0217	0.8822	0.3697	−0.0423	−0.0089	0
Q6	0.5470	0.3424	0.2116	−0.2120	0.0973	0.0461	−0.6984	0
Q7	0.5607	0.3004	0.1827	−0.2032	0.0918	0.0383	0.7147	0

† Q1: good profiteering business and high possibility of rising price, Q2: good investment to buy land and keep it (i.e., as long as the value keeps increasing) before selling it, Q3: bought land due to ease of access of obtaining land title, Q4: privilege information on zoning and planning of a new development scheme, Q5: excess demand not met by government allocation, Q6: bought land due to its close proximity to a developed area, Q7: regulatory lapses makes it possible to buy land cheaply; ‡ data displayed factorability potential based on Bartlett's test of sphericity, employing a chi-square value of 192.773 at 21 degrees of freedom, significant at 0.01 showing correlations among the chosen variables, hence a supportive criterion for factorability.

From the PCA, two types of findings can be noted. First, anti-image diagonals illustrated sampling adequacy via the Kaiser–Meyer–Olkin measure (i.e., with a value of 0.60 for each variable), while Bartlett's test of sphericity indicated a significance p-value of 0.000—both demonstrating a significant level of sample tolerability for this analysis (Table 3). Second, a total of seven principal components were recognized with a collective variance of 69.59% in which specific high impact findings originate from the first, second, and third components (Table 4). Out of the three factors studied, the first two together explained 54.42% of variation compared to 69.59 % explained by unrotated factors, as indicated in the Scree plot in Figure 5.

Table 3. Two calculations: Kaiser–Meyer–Olkin measure and Bartlett's test.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		**0.60**
Bartlett's test of sphericity	approx. chi-square	192.773
-	df	21
-	sig	0.000

Table 4. Total variance explained for the initial eigenvalues.

Component	Initial Eigenvalues		Cumulative Percentage of Variance
	Total	Percent of Variance	
1	2.48345	35.48	35.48
2	1.32565	18.94	54.42
3	1.0619	15.17	69.59
4	0.907426	12.96	82.55
5	0.636939	9.10	91.65
6	0.544361	7.78	99.42
7	0.0402671	0.58	100.00

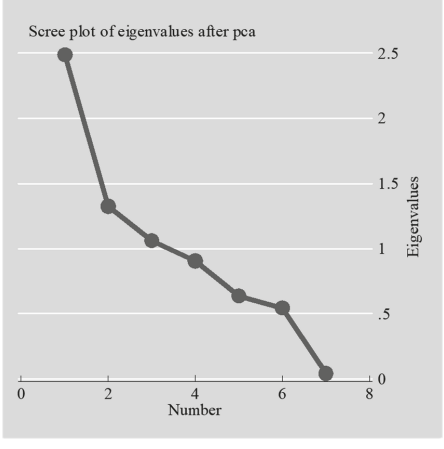


Figure 5. Scree plot of land speculation in the Shashemene urban fringe.

Correlating the Scree plot of land with Table 2 shows the rotated component factors. The rotated component matrix specifies sufficient loading on all seven components. There are two dominant variables on the first component (i.e., PCA 1) due to speculators who bought land in close proximity to a developed area as well as regulatory lapses make it possible to buy land cheaply. The dominant variables on the second component (i.e., PCA 2) were good profiteering business efforts and the high possibility of rising prices for profit maximization. In the third component (i.e., PCA 3), the dominant variable was spectators' ability to buy land due to ease of access of obtaining a land title. Next, privilege information on zoning and planning of a new development scheme best denoted the fourth component (i.e., PCA 4). Good investment for buying land and being able to keep it (i.e., as long as the value keeps increasing) before selling it as well as privilege information on zoning and planning of a new development scheme were the dominant variables in PCA 5 and PCA 6, respectively. Lastly, regulatory lapses making it possible to buy land cheaply was dominant in PCA 7. The common variables that constantly occur among the components were: (1) privilege information on zoning and planning of a new development scheme and (2) regulatory lapses that made it possible to buy land cheaply. However, it is unusual that, of all the rated factors, 88% of the responses did not feel that the demand was being met by government allocation. In other words, the usual acclaimed motivational reason for land acquisition of excess demand not met by government allocation rated very low (i.e., low probability) and did not even feature more than once (i.e., except in PCA 4). This entropy, disorder of communicative data, suggests disparity between the hidden intention of speculators [77,78] and what really exists physically on the ground. According to Zhang [79], surprisal information on "price histories can be used to predict near future returns with a probability better than [that of] random chance". As such, since market participants are separated between producers and speculators—the

former postulates "negative entropy into the price, upon which the latter feed" [79]. Residual negative entropy suggests an urgent need for more effective government intervention in allocating land in the urban land market, particularly due to land scarcity and relating price mechanism ill-equipped in maintaining a standard.

In reference to Mercer's [80] definition of livability (i.e., "a concept that assesses which locations around the world provide the best or the worst living conditions" [80]), livability combines a range of benefits that include benchmarking perceptions of development to assigning hardship allowance as part of expatriate relocation packages. Critics of this view have argued that "no city in the world is really excellent and that livability is only a relative term" [80]. In the context of this paper, it is contended that land speculation is detrimental to a city's livability (i.e., sustainability) in several ways as proclaimed by Swierenga's [81] research on land speculation and impact on American economic growth and welfare. He states that land speculation gives rise to green pockets that are easily transformed into high crime areas fostering slum development by way of rural-urban migration and accentuating a lack of housing as a result of holding developable land. At length, Swierenga's [81] findings parallel, to some degree, Shashemene's reduced property tax base since a number of American municipality's often do not tax vacant land or are not subject to property tax (i.e., they are subject to a tenement rate), which can result in a loss of potential sources of revenue.

5. Conclusions

We conclude that, for the city of Shashemene, the government is fully aware of speculative practices occurring along its border areas. Government policy put in place for regulating excessive land acquisition is still limited to the Land Use Act of 1993. It is not clear how this legislation will deal with restricting or ridding speculative activities throughout the city's limits. As of now, over the past few decades, legislative implementation or political clout has failed to decrease this phenomenon. Government has not effectively controlled or penalized excessive land acquisitions that are not for immediate use. Our results verify a relatively low level of property rights formalization, mainly due to bureaucratic practices and corruption that encourage large acquisitions of land with impunity. This problem demonstrates a disturbing trend in the dynamic of the city's fringe areas mostly due to repeated failures of government in regulating land use activities. As a result, there is an urgent need for more effective government intervention in allocating land in the urban land market, particularly due to land scarcity and relating price mechanism ill-equipped in maintaining a standard. In addition, there is the need for strategic measures, such as the institutionalization of public–private joint ventures that will not only work in partnership with landowners and developers, but also work to develop a "speedy, efficient, and sustainable management" [46] and development of the fringes in terms of provision of cost-effective infrastructure and increased public information on land market opportunities in the fringe areas. Toward this end, the creation of a government holding division to procure vacant lands is believed to be a needed priority.

To minimize land speculation seen in urban and peri-urban areas, the following two points are recommended: (1) policy implementation of "one-man-on plot" rule, which implies that the number of plots given to residents must be restricted and (2) there should be imposition of a land value tax (LVT). These two points would significantly curb excessive land speculation and allow for controlled development to occur. As a result, a number of benefits on enforcing a high LVT, include: high annual overhead cost to the land speculator that would either lower present values or provide encouragement for quick disposal, especially if landowners could not meet the overhead costs and pessimistic future appreciation of land value. Moreover, Clawson [82] argues that sprawl supplemented by land speculation is fruitless in that it absorbs "capital, manpower, and entrepreneurial skills without proportionate public gains" [82]. Based on our findings, it is contended that the need for strict regular site inspections and a corruption-free regulatory regime are clearly critical in monitoring land speculative activities in Shashemene. Underlying these suggestions is the idea that the government, at some level, retains significant powers for influencing, if not controlling, the future form of the outer

edges. To achieve positive goals, speculation on urban fringe lands must be significantly decreased or eliminated. The net effect of these various instruments would be to greatly alter the general probability (i.e., prospect) of land actors on speculation. Entropy theory suggests disparity between the hidden intention of speculators [52,53] in which the historicity of price can be used to predict near future earnings with a probability that is better than random chance [54]. As a result, the timely implementation of an LVT policy could lead to more predictable land prices, extemporaneous housing, or other land use advances. In this process, some of the private validations for land speculation would be overcome. Additionally, since the eventual purpose of speculation is short-term capital gains, one way to stop speculation could be through effective government's association with land-based "pressure groups" [82] or non-profit organizations regularly monitoring the border areas located in the city's domain. Furthermore, support for landowners to initiate direct development within a sensible period after the purchase could be a secondary approach. Finally, lending institutions (i.e., banks and insurance companies) should be vigorously involved in the push towards easing conditions placed on mortgage applicants.

Author Contributions: Conceptualization and Data Curation, B.S.G.; Methodology, Software, Validation, and Writing—Original Draft Preparation, B.G.A. and B.S.G.; Formal Analysis, Investigation, Resources, and Writing—Review and Editing, A.P., B.G.A., B.S.G., G.T.C., and Y.X.; Visualization and Supervision, B.G.A. and G.T.C.; Project Administration and Funding Acquisition, A.P and G.T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dimuna, K.O. Enhancing Land Acquisition for Individual Housing Development in Nigeria: A Case Study of Benin Metropolis Edo State, Nigeria. *Int. J. Res. Innov. Appl. Sci.* **2016**, *1*, 1–9.
2. Nwoko, K.C. Land ownership versus development in the era of globalisation: A trajectory of conflict and wealth accumulation in Southern Nigeria. *J. Afr. Transform.* **2016**, *1*, 77–94.
3. Chang, V.; Newman, R.; Walters, R.J.; Wills, G.B. Review of economic bubbles. *Int. J. Inf. Manag.* **2016**, *36*, 497–506. [CrossRef]
4. Makkonen, T.; Williams, A.M. Border region studies: The structure of an 'offbeat' field of regional studies. *Reg. Stud. Reg. Sci.* **2016**, *3*, 355–367. [CrossRef]
5. Scott, J.W.; Sohn, C. Place-making and the bordering of urban space: Interpreting the emergence of new neighbourhoods in Berlin and Budapest. *Eur. Urban Reg. Stud.* **2018**, 096977641876457. [CrossRef]
6. Molotch, H. The City as a Growth Machine: Toward a Political Economy of Place. *Am. J. Sociol.* **1976**, *82*, 309–332. [CrossRef]
7. Siegel, J.J. What Is an Asset Price Bubble? An Operational Definition. *Eur. Financ. Manag.* **2003**, *9*, 11–24. [CrossRef]
8. Kittrell, K. Impacts of Vacant Land Values. *Transp. Res. Rec. J. Transp. Res. Board* **2012**, *2276*, 138–145. [CrossRef]
9. Mohamed, R. The Psychology of Residential Developers. *J. Plan. Educ. Res.* **2006**, *26*, 28–37. [CrossRef]
10. Fauser, M. The Emergence of Urban Border Spaces in Europe. *J. Borderl. Stud.* **2017**, *34*, 605–622. [CrossRef]
11. Joshua, P.B.; Glanda, G.G.; Ilesanmi, F.A. The Effects of Land Speculation on Urban Planning and Development in Bajabure Area, Girei Local Government, Adamawa State. *J. Environ. Earth Sci.* **2016**, *6*, 128–133.
12. Gemeda, B.S.; Abebe, B.G.; Cirella, G.T. The Role of Land Speculators around the Urban Edge of Shashemene City, Ethiopia. *Eur. J. Sustain. Dev. Res.* **2019**, *4*, em0108. [CrossRef]
13. Golland, A.; Boelhouwer, P. Speculative housing supply land and housing, markets: A comparison. *J. Prop. Res.* **2002**, *19*, 231–251. [CrossRef]
14. Wang, X.L.; Hua, S. Influence Factors of Chinese Real Estate. *Appl. Mech. Mater.* **2013**, *405–408*, 3391–3395. [CrossRef]

15. Westerink, J.; Haase, D.; Bauer, A.; Ravetz, J.; Jarrige, F.; Aalbers, C.B.E.M. Dealing with Sustainability Trade-Offs of the Compact City in Peri-Urban Planning Across European City Regions. *Eur. Plan. Stud.* **2013**, *21*, 473–497. [CrossRef]
16. World Bank. *Urbanization beyond Municipal Boundaries*; The World Bank: Washington, DC, USA, 2013; ISBN 978-0-8213-9840-1.
17. Ambaye, D.W. Land Rights and Expropriation in Ethiopia. Ph.D. Thesis, School of Architecture and the Built Environment, Royal Institute of Technology, Stockholm, Sweden, 2013.
18. NechaSungena, T.; Serbeh-Yiadom, K.; Asfaw, M. Towards an Efficient Implementation of the Land Lease Policy of Ethiopia: A case-study of Hawassa. *Dev. Ctry. Stud.* **2014**, *4*, 155–166.
19. Abdo, M. Legislative Protection of Property Rights in Ethiopia: An Overview. *Mizan Law Rev.* **2014**, *7*, 165. [CrossRef]
20. Barbier, E.B.; Hochard, J.P. Does Land Degradation Increase Poverty in Developing Countries? *PLoS ONE* **2016**, *11*, e0152973. [CrossRef]
21. Unruh, J.D. Land Tenure and the "Evidence Landscape" in Developing Countries. *Ann. Assoc. Am. Geogr.* **2006**, *96*, 754–772. [CrossRef]
22. Diergarten, Y. Indigenous or Out of Scope? Large-scale Land Acquisitions in Developing Countries, International Human Rights Law and the Current Deficiencies in Land Rights Protection. *Hum. Rights Law Rev.* **2019**, *19*, 37–52. [CrossRef]
23. Business Dictionary Online Business Dictionary. Available online: http://www.businessdictionary.com/ (accessed on 20 June 2019).
24. Park, R.E.; Burgess, E.W.; McKenzie, R.D. *The City*; University of Chicago Press: Chicago, IL, USA, 1925; ISBN 0226646114.
25. Adams, J.S.; Hoyt, H. 1939: The structure and growth of residential neighborhoods in American cities. Washington, DC: Federal Housing Administration. *Prog. Hum. Geogr.* **2005**, *29*, 321–325. [CrossRef]
26. Alonso, W. The Historic and the Structural Theories of Urban Form: Their Implications for Urban Renewal. *Land Econ.* **1964**, *40*, 227. [CrossRef]
27. Burgess, E.W. The Growth of the City: An Introduction to a Research Project. In *Urban Ecology*; Springer: Boston, MA, USA, 2008; pp. 71–78.
28. Motoro, P. A Comparative Analysis of Residential Property Rental Markets in Informal Settlements and Formal Areas of Lae and Port Moresby. Master's Thesis, Addis Ababa University, Addis Ababa, Ethiopia, 2016.
29. Hoyt, H. *One Hundred Years of Land Values in Chicago: The Relationship of the Growth of Chicago to the Rise of Its Land Values, 1830–1933*; Beard Books: Washington, DC, USA, 2000; ISBN 9781587980169.
30. Perman, R.; Perman, R. *Natural Resource and Environmental Economics*; Pearson Addison Wesley: Harlow, UK, 2011; ISBN 0321417534.
31. Valliere, D.; Peterson, R. Venture Capitalist Behaviours: Frameworks for Future Research. *Ventur. Cap.* **2005**, *7*, 167–183. [CrossRef]
32. Oduwaye, L. Challenges of Sustainable Physical Planning and Development in Metropolitan Lagos. *J. Sustain. Dev.* **2009**, *2*, 159. [CrossRef]
33. Osabuohien, E.S. Large-scale agricultural land investments and local institutions in Africa: The Nigerian case. *Land Use Policy* **2014**, *39*, 155–165. [CrossRef]
34. Ariyo, J.A.; Ogbonna, D.O. The effects of land speculation on agricultural production among peasants in Kachia local government area of Kaduna State, Nigeria. *Appl. Geogr.* **1992**, *12*, 31–46. [CrossRef]
35. Cotula, L. The international political economy of the global land rush: A critical appraisal of trends, scale, geography and drivers. *J. Peasant Stud.* **2012**, *39*, 649–680. [CrossRef]
36. Nesheim, I.; Verburg, R.; Abdeladhim, M.A.; Bursztyn, M.; Chen, L.; Cissé, Y.; Feng, S.; Gicheru, P.; Jochen König, H.; Novira, N.; et al. Causal chains, policy trade offs and sustainability: Analysing land (mis)use in seven countries in the South. *Land Use Policy* **2014**, *37*, 60–70. [CrossRef]
37. Sharmina, M.; Hoolohan, C.; Bows-Larkin, A.; Burgess, P.J.; Colwill, J.; Gilbert, P.; Howard, D.; Knox, J.; Anderson, K. A nexus perspective on competing land demands: Wider lessons from a UK policy case study. *Environ. Sci. Policy* **2016**, *59*, 74–84. [CrossRef]
38. Brown, H.G. Wealth and Want: Speculation. Available online: http://www.wealthandwant.com/themes/Speculation.html (accessed on 18 June 2019).

39. Sinclair, U. Wealth and Want: Speculation. Available online: http://www.wealthandwant.com/docs/Sinclair_CoLS.html (accessed on 18 June 2019).
40. Malpezzi, S.; Wachter, S.M. The Role of Speculation in Real Estate Cycles. *J. Real Estate Lit.* **2005**, *13*, 143–164. [CrossRef]
41. Messerli, P.; Giger, M.; Dwyer, M.B.; Breu, T.; Eckert, S. The geography of large-scale land acquisitions: Analysing socio-ecological patterns of target contexts in the global South. *Appl. Geogr.* **2014**, *53*, 449–459. [CrossRef]
42. Colin, J.-P. Securing rural land transactions in Africa. An Ivorian perspective. *Land Use Policy* **2013**, *31*, 430–440. [CrossRef]
43. Hurni, H.; Berhe, W.A.; Chadhokar, P.; Daniel, D.; Gete, Z.; Grunder, M.; Kassaye, G. *Soil and Water Conservation in Ethiopia: Guidelines for Development Agents*; Centre for Development and Environment (CDE), University of Bern, with Bern Open Publishing (BOP): Bern, Switzerland, 2016; ISBN 9783906813134.
44. Nolte, K. Large-scale agricultural investments under poor land governance in Zambia. *Land Use Policy* **2014**, *38*, 698–706. [CrossRef]
45. Petrick, M.; Wandel, J.; Karsten, K. Rediscovering the Virgin Lands: Agricultural Investment and Rural Livelihoods in a Eurasian Frontier Area. *World Dev.* **2013**, *43*, 164–179. [CrossRef]
46. Andreasson, P.; Bekiros, S.; Nguyen, D.K.; Uddin, G.S. Impact of speculation and economic uncertainty on commodity markets. *Int. Rev. Financ. Anal.* **2016**, *43*, 115–127. [CrossRef]
47. Ermolieva, T.; Havlík, P.; Ermoliev, Y.; Mosnier, A.; Obersteiner, M.; Leclère, D.; Khabarov, N.; Valin, H.; Reuter, W. Integrated Management of Land Use Systems under Systemic Risks and Security Targets: A Stochastic Global Biosphere Management Model. *J. Agric. Econ.* **2016**, *67*, 584–601. [CrossRef]
48. Brunnermeier, M.K.; Cheridito, P.; Brunnermeier, M.K.; Cheridito, P. Measuring and Allocating Systemic Risk. *Risks* **2019**, *7*, 46. [CrossRef]
49. Singh, G. Is India Hedged Against Systemic Risk? An Attempt at an Answer. *Rev. Mark. Integr.* **2013**, *5*, 83–129. [CrossRef]
50. Huang, X.; Vodenska, I.; Havlin, S.; Stanley, H.E. Cascading failures in bi-partite graphs: Model for systemic risk propagation. *Sci. Rep.* **2013**, *3*, 1219. [CrossRef]
51. Knittel, C.; Pindyck, R. *The Simple Economics of Commodity Price Speculation*; The MIT Press: Cambridge, MA, USA, 2013.
52. Triantafyllopoulos, N. On the origins of tourist urbanisation in Greece: Land speculation and property market (in)efficiency. *Land Use Policy* **2017**, *68*, 15–27. [CrossRef]
53. Maruani, T.; Amit-Cohen, I. Open space planning models: A review of approaches and methods. *Landsc. Urban Plan.* **2007**, *81*, 1–13. [CrossRef]
54. Jonas, A.E.G.; Wilson, D. *The Urban Growth Machine: Critical Perspectives Two Decades Later*; SUNY Press: Albany, NY, USA, 1999; pp. 1–312.
55. Logan, J.R.; Molotch, H.L. *Urban Fortunes: The Political Economy of Place*; University of California Press: Berkeley, CA, USA, 1987; ISBN 0520063414.
56. Wetzstein, S. The global urban housing affordability crisis. *Urban Stud.* **2017**, *54*, 3159–3177. [CrossRef]
57. Paczoski, A.; Abebe, S.T.; Cirella, G.T. Debt and Deficit Growth Rate Reporting for Post-Communist European Union Member States. *Soc. Sci.* **2019**, *8*, 173. [CrossRef]
58. Paczoski, A. Różnice kulturowe wobec wzrostu gospodarczego: Ze szczególnym uwzględnieniem przykładów państw UE (Translated from Polish: "Cultural differences and economic growth: With emphasized examples from the EU"). *Polityka Gospod.* **2014**, *22*, 91–118.
59. Atkinson, R. Limited exposure: Social concealment, mobility and engagement with public space by the super-rich in London. *Environ. Plan. A Econ. Sp.* **2016**, *48*, 1302–1317. [CrossRef]
60. Hay, I. *Geographies of the Super-Rich*; Edward Elgar: Cheltenham, UK, 2013; ISBN 0857935682.
61. Hay, I.; Muller, S. 'That Tiny, Stratospheric Apex That Owns Most of the World'-Exploring Geographies of the Super-Rich. *Geogr. Res.* **2012**, *50*, 75–88. [CrossRef]
62. Lee, K.N.H. Residential property price-stock price nexus in Hong Kong: New evidence from ARDL bounds test. *Int. J. Hous. Mark. Anal.* **2017**, *10*, 204–220. [CrossRef]
63. Ioannides, Y.M.; Rosenthal, S.S.; Ioannides, Y.; Rosenthal, S. Estimating the Consumption and Investment Demands for Housing and Their Effect on Housing Tenure Status. *Rev. Econ. Stat.* **1994**, *76*, 127–141. [CrossRef]

64. Fernandez, R.; Hofman, A.; Aalbers, M.B. London and New York as a safe deposit box for the transnational wealth elite. *Environ. Plan. A Econ. Sp.* **2016**, *48*, 2443–2461. [CrossRef]
65. Ho, H.K.; Atkinson, R. Looking for big 'fry': The motives and methods of middle-class international property investors. *Urban Stud.* **2018**, *55*, 2040–2056. [CrossRef]
66. Knight Frank International Residential Investment in London: International Project Marketing 2013. Available online: https://content.knightfrank.com/research/503/documents/en/2013-1217.pdf (accessed on 10 December 2019).
67. Zheng, S. Record Complaints as Overseas Sales Sting Hong Kong Investors. Available online: https://www.scmp.com/news/hong-kong/economy/article/2065077/record-complaints-overseas-sales-sting-hong-kong-investors (accessed on 15 December 2019).
68. Census and Statistics Department Mid-Year Population for 2018. Available online: https://www.info.gov.hk/gia/general/201808/14/P2018081400417.htm (accessed on 10 December 2019).
69. Bangura, M.; Lee, C.L. The differential geography of housing affordability in Sydney: A disaggregated approach. *Aust. Geogr.* **2019**, *50*, 295–313. [CrossRef]
70. Lee, C.L. An examination of the risk-return relation in the Australian housing market. *Int. J. Hous. Mark. Anal.* **2017**, *10*, 431–449. [CrossRef]
71. Lee, M.-T.; Lee, C.L.; Lee, M.-L.; Liao, C.-Y. Price linkages between Australian housing and stock markets. *Int. J. Hous. Mark. Anal.* **2017**, *10*, 305–323. [CrossRef]
72. Revington, N.; Townsend, C. Market Rental Housing Affordability and Rapid Transit Catchments: Application of a New Measure in Canada. *Hous. Policy Debate* **2016**, *26*, 864–886. [CrossRef]
73. Fleming, T.; Damon, W.; Collins, A.B.; Czechaczek, S.; Boyd, J.; McNeil, R. Housing in crisis: A qualitative study of the socio-legal contexts of residential evictions in Vancouver's Downtown Eastside. *Int. J. Drug Policy* **2019**, *71*, 169–177. [CrossRef]
74. Rogers, D.; Koh, S.Y. The globalisation of real estate: The politics and practice of foreign real estate investment. *Int. J. Hous. Policy* **2017**, *17*, 1–14. [CrossRef]
75. Joseph, L.; Belisle, P. Bayesian Sample Size Determination for Case-Control Studies When Exposure May be Misclassified. *Am. J. Epidemiol.* **2013**, *178*, 1673–1679. [CrossRef]
76. Krejcie, R.V.; Morgan, D.W. Determining Sample Size for Research Activities. *Educ. Psychol. Meas.* **1970**, *30*, 607–610. [CrossRef]
77. Bekiros, S.; Nguyen, D.K.; Sandoval Junior, L.; Salah Uddin, G. Information Diffusion, Cluster formation and Entropy-based Network Dynamics in Equity and Commodity Markets. *Eur. J. Oper. Res.* **2015**, *256*, 945–961. [CrossRef]
78. Harremoës, P.; Topsøe, F. Maximum Entropy Fundamentals. *Entropy* **2001**, *3*, 191–226. [CrossRef]
79. Zhang, Y.-C. Toward a Theory of Marginally Efficient Markets. *Phys. A Stat. Mech. Its Appl.* **1999**, *269*, 30–44. [CrossRef]
80. Mercer. *Western European Cities–Top Quality of Living Ranking*; Mercer: London, UK, 2016.
81. Swierenga, R.P. Land Speculation and Its Impact on American Economic Growth and Welfare: A Historiographical Review. *West. Hist. Q.* **1977**, *8*, 283. [CrossRef]
82. Clawson, M. Urban Sprawl and Speculation in Suburban Land. In *A Geography of Urban Places*; Putnam, R.G., Taylor, F.J., Kettle, P.K., Eds.; Routledge: London, UK, 1970; pp. 329–342.

entropy

MDPI

Article

The Value of Information Searching against Fake News

José Martins [1,*,†] and Alberto Pinto [2,†]

[1] LIAAD-INESC TEC and School of Technology and Management, Polytechnic of Leiria, Campus 2, Morro do Lena-Alto do Vieiro, 2411-901 Leiria, Portugal

[2] LIAAD-INESC TEC and Faculty of Sciences, University of Porto, R Campo Alegre, 4169-007 Porto, Portugal; aapinto@fc.up.pt

* Correspondence: jmmartins@ipleiria.pt

† These authors contributed equally to this work.

Received: 13 November 2020; Accepted: 19 November 2020; Published: 3 December 2020

Abstract: Inspired by the Daley-Kendall and Goffman-Newill models, we propose an Ignorant-Believer-Unbeliever rumor (or fake news) spreading model with the following characteristics: (i) a network contact between individuals that determines the spread of rumors; (ii) the value (cost versus benefit) for individuals who search for truthful information (learning); (iii) an impact measure that assesses the risk of believing the rumor; (iv) an individual search strategy based on the probability that an individual searches for truthful information; (v) the population search strategy based on the proportion of individuals of the population who decide to search for truthful information; (vi) a payoff for the individuals that depends on the parameters of the model and the strategies of the individuals. Furthermore, we introduce evolutionary information search dynamics and study the dynamics of population search strategies. For each value of searching for information, we compute evolutionarily stable information (ESI) search strategies (occurring in non-cooperative environments), which are the attractors of the information search dynamics, and the optimal information (OI) search strategy (occurring in (eventually forced) cooperative environments) that maximizes the expected information payoff for the population. For rumors that are advantageous or harmful to the population (positive or negative impact), we show the existence of distinct scenarios that depend on the value of searching for truthful information. We fully discuss which evolutionarily stable information (ESI) search strategies and which optimal information (OI) search strategies eradicate (or not) the rumor and the corresponding expected payoffs. As a corollary of our results, a recommendation for legislators and policymakers who aim to eradicate harmful rumors is to make the search for truthful information free or rewarding.

Keywords: fake news; rumor spreading; Nash equilibrium; evolutionarily stable strategies; evolutionary information search dynamics

1. Introduction

The theory of rumor (or fake news) spreading proposed by Daley and Kendall [1,2] became known as the DK model, in which a population is divided into three different groups: ignorants—people who are ignorant concerning the rumor; spreaders—people who actively spread the rumor; and stiflers—people who have heard the rumor but are no longer interested in spreading it. Goffman and Newill [3] also published a paper in 1964 that generalized epidemic theory and provided a clear analogy between the spreading of infectious disease and the transmission of ideas. In the subsequent years, several authors developed the theory of rumor spreading, proposing new models using complex networks [4], transitions capable of describing different issues in the transmission process [5],

and Lévy noise [6]. In this paper, we develop a game-theoretical approach for a network-extended version of the rumor spreading models proposed in [1,3], using the ideas developed by Bauch and Earn [7] for the well-known Susceptible-Infected-Recovered (SIR) epidemiological model. Furthermore, the topic of fake news is quantified in terms of the value of searching for truthful information (learning), the impact of believing the fake rumor, and the individual's payoff, which is of paramount importance in academia.

Rumors and fake news can be considered a form of cheating. Individuals might be pushed toward risk-seeking or loss aversion on the basis of their feelings (see [8]). Political news can have a strong effect on stock prices (see [9]). In terms of the outbreak of COVID-19, information on social media can lead to numerous negative behaviors that can reduce vaccination coverage and the use of COVID alert applications (see [10]). On the other hand, rumors and fake news do not necessarily have negative impacts. An extreme example occurred during the Cold War: the propaganda machines from the American and Soviet sides spread numerous rumors about (i) the intentions of their rivals and (ii) the achievements of their countries in several areas (e.g., science, business, and industry). These rumors served the purpose of contributing to improvements in the well-being of both populations. If we see a rumor as an exaggerated piece of information with an essence of truth, then it can have a positive impact on the population. Another example occurred during World War II, when many rumors were spread concerning Nazi Germany. Although some news was fake, it served the purpose of boosting the morale of the Allied population and the troops. Hence, fake news can have either a negative or a positive impact on an individual's behavior.

In the Ignorant-Believer-Unbeliever (IBU) rumor (or fake news) spreading dynamical model, individuals are spatially distributed in a network and can be either ignorants, believers, or unbelievers regarding a certain rumor. When a rumor appears in a population, individuals will act differently depending on their beliefs about the rumor. If an individual believes the rumor, then he/she will spread the rumor to his/her neighbors. On the other hand, individuals who do not believe the rumor will not act as active spreaders. This spreading dynamical model is fully inspired by the SIR epidemiological model. The impact measure y of the rumor evaluates the gains and losses resulting from individuals' decisions, provoked by their beliefs in the rumor. The value v of searching for truthful information (learning), instead of just believing the rumor, has natural benefits and costs to the individual. Each ignorant individual has his/her own information search strategy S based on his/her probability of searching for truthful information per unit of time. The population's information search strategy s is the proportion of ignorant individuals who will choose to search for truthful information per unit of time. For instance, if all ignorant individuals follow the same strategy S (homogeneous strategy), then $s = S$. For an ignorant individual, we introduce the expected information search payoff, which depends on (i) his/her information search strategy S; (ii) the population information search strategy s; (iii) the value v of searching for truthful information; (iv) the impact measure y of the rumor; and (v) the spread dynamics of the rumor.

A population information search strategy S is a Nash equilibrium if not a single individual has an incentive to change his/her information search strategy to any other strategy $S' \neq S$ (see [11]). A population information search strategy S is an evolutionarily stable information search strategy if any small group of individuals that tries to adopt a different strategy S' obtains a lower payoff than those adopting the original strategy S (see [11]). Evolutionarily stable information search strategies are Nash strategies that are practiced by individuals in non-cooperative environments. A population information search strategy S is an optimal information search strategy if it maximizes the payoffs of individuals. Optimal information search strategies are practiced by individuals in (eventually forced) cooperative environments. Here, we fully characterize the triples (v, y, S), where S is (i) a Nash strategy, (ii) an evolutionarily stable information search strategy, or (iii) an optimal information search strategy; v is the value of searching for information; and y is the impact measure of believing a false rumor (fake news). Finally, we introduce evolutionary information searching dynamics following the replicator dynamics theory [11–13], where the search strategies evolve over time to increase the payoffs

of individuals. Evolutionarily stable information search strategies are the attractors of the dynamics; i.e., over time, the population search strategy tends toward the evolutionarily stable information search strategy (and not necessarily to the optimal information search strategies).

For rumors that are advantageous to the population (positive impact $y > 0$), three distinct scenarios occur, depending on the value of searching for truthful information: (i) for high positive values of searching, both evolutionarily stable information (ESI) and optimal information (OI) search strategies coincide, and all individuals search for truthful information, eradicating the rumor; (ii) (bi-stability) for small positive values of searching, there are two ESI search strategies: either all individuals search for truthful information (eradicating the rumor in non-cooperative environments), or no one searches for truthful information (persistence of the rumor in non-cooperative environments), and the OI search strategy jumps (at the right-boundary of this bi-stability region) from no one searching (persistence of the rumor in cooperative environments) to all individuals searching (eradicating the rumor in cooperative environments); (iii) for negative values of searching, we show that both ESI and OI search strategies coincide, and no individuals search for truthful information, and thus, the rumor persists.

For rumors that are harmful to the population (negative impact $y < 0$), we show the existence of three distinct scenarios that occur, depending on the value of searching for truthful information: (i) for positive values of searching, both ESI and OI search strategies coincide, and all individuals search for truthful information, eradicating the rumor; (ii) for small negative values of searching, the OI search strategy coincides with the critical probability that is necessary to eradicate the rumor, and thus, the rumor is eradicated in cooperative environments, but the ESI search strategy is less successful than the OI search strategy, so, unfortunately, the rumor persists in non-cooperative environments; (iii) for highly negative values of searching, both ESI and OI search strategies coincide, and no individuals search for truthful information, and thus, the rumor persists. Hence, a recommendation for legislators and policymakers who aim to eradicate harmful rumors is to make the search for truthful information free or rewarding, i.e., information search value $v \geq 0$. For instance, truthful public social media campaigns can help by making the information easily available.

This paper is organized as follows. In Section 2, we introduce the IBU rumor spreading model for networks. In Section 3, we introduce a utility for individuals that depends on the value of information and the impact of believing the rumor. Nash and evolutionarily stable information search strategies are completely characterized. In Section 4, optimal information search strategies are deduced for different values of information. In Section 5, we introduce evolutionary information search dynamics and study its attractors. Section 6 provides the conclusions of the paper and directions for future research work.

2. The IBU Spreading Model on Regular Networks

Inspired by the work in [1,3], we propose the Ignorant-Believer-Unbeliever (IBU) dynamic model for rumor spreading based on the classical Susceptible-Infected-Recovered (SIR) epidemic model (see also [14,15]). Individuals can be either *Ignorants*, *Believers*, or *Unbelievers* of a certain rumor. The IBU model is directly analogous to the SIR model:

S—*Susceptibles* correspond to *I*—*Ignorants*,
I—*Infected* individuals correspond to *B*—*Believers* of the rumor,
R—*Recovered* individuals correspond to *U*—*Unbelievers* of the rumor.

Individuals who believe the rumor are the active spreaders: i.e., they are the individuals who transmit the rumor to ignorant individuals. Once a believer stops believing the rumor and becomes an unbeliever, he/she will stop transmitting the rumor. Hence, unbelievers are not active spreaders. As in epidemiology [10], a transition corresponding to vaccination is introduced in the model. This transition

is due to information search activities that can be voluntarily adopted by an ignorant individual. State transitions in the IBU model are illustrated in Figure 1 and defined by the following reaction scheme:

$$\begin{aligned} \mathcal{I}_i + \mathcal{B}_j &\xrightarrow{\beta} \mathcal{B}_i + \mathcal{B}_j \\ \mathcal{B}_i &\xrightarrow{\gamma} \mathcal{U}_i \\ \mathcal{I}_i &\xrightarrow{\nu} \mathcal{U}_i \\ \mathcal{I}_i, \mathcal{B}_i, \mathcal{U}_i &\xrightarrow{\mu} \mathcal{I}_i. \end{aligned}$$

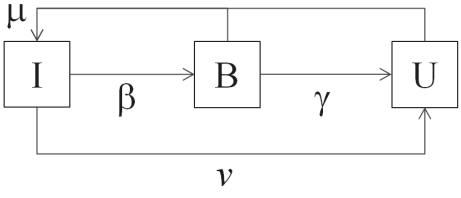


Figure 1. The compartmental Ignorant-Believer-Unbeliever (IBU) rumor spreading model.

The individual state variables $\mathcal{I}_i$, $\mathcal{B}_i$, and $\mathcal{U}_i \in \{0, 1\}$ identify the state of individual i, restricted to the condition that the individual belongs to one of the three classes. Hence,

$$\mathcal{I}_i + \mathcal{B}_i + \mathcal{U}_i = 1.$$

The parameters of the model have the following interpretation: β is the rate at which one believer individual spreads the rumor; μ is the mean birth and death rates, and thus, $1/\mu$ is the mean life expectancy at birth; γ is the rate at which a believer stops believing the rumor and stops spreading it, and thus, $1/(\gamma + \mu)$ is the mean believing/spreading period; ν is the information search rate, i.e., the rate at which an ignorant individual searches for real information and become an unbeliever.

Let us assume that a population is fixed in size with N individuals; hence,

$$\sum_{i=1}^{N} (\mathcal{I}_i + \mathcal{B}_i + \mathcal{U}_i) = N.$$

To describe the neighbor structure of individuals in the population, we consider the $N \times N$ adjacency matrix J with elements $J_{i,j} \in \{0, 1\}$ such that: if individual i is a neighbor of j, then $J_{i,j} = 1$, and if individual i is not a neighbor of j, then $J_{i,j} = 0$. The matrix J is symmetric with zero elements in the diagonal. Let $\{\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., \mathcal{I}_i, \mathcal{B}_i, \mathcal{U}_i, ..., \mathcal{U}_N\}$ denote a certain state of the population, and let $p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., \mathcal{I}_i, \mathcal{B}_i, \mathcal{U}_i, ..., \mathcal{U}_N, t)$ be the probability of that state occurring at time t. The time evolution of $p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., \mathcal{I}_i, \mathcal{B}_i, \mathcal{U}_i, ..., \mathcal{U}_N, t)$ is described by a master equation [16] given by an ordinary differential equation (ODE) system that models the probabilistic combination of states and the switching between those states depending on the transition rates of the mathematical model and the spatial structure of the population. Following Glauber's Ising spin dynamics [17] or Stollenwerk et al.'s reinfection SIRI model [18,19], the master equation for the IBU spreading model is given by

$$
\begin{aligned}
\frac{d}{dt} \quad & p \quad (\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., \mathcal{I}_i, \mathcal{B}_i, \mathcal{U}_i, ..., \mathcal{U}_N, t) \\
= \quad & \sum_{i=1}^{N} \beta \left(\sum_{j=1}^{N} J_{ij} \mathcal{B}_j \right) (1 - \mathcal{I}_i)\ p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., 1 - \mathcal{I}_i, 1 - \mathcal{B}_i, \mathcal{U}_i ..., \mathcal{U}_N, t) \\
& + \sum_{i=1}^{N} \gamma\ (1 - \mathcal{B}_i)\ p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., \mathcal{I}_i, 1 - \mathcal{B}_i, 1 - \mathcal{U}_i ..., \mathcal{U}_N, t) \\
& + \sum_{i=1}^{N} \nu (1 - \mathcal{I}_i)\ p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., 1 - \mathcal{I}_i, \mathcal{B}_i, 1 - \mathcal{U}_i ..., \mathcal{U}_N, t) \\
& + \sum_{i=1}^{N} \mu \Big[(1 - \mathcal{I}_i)\ p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., 1 - \mathcal{I}_i, \mathcal{B}_i, \mathcal{U}_i ..., \mathcal{U}_N, t) \\
& \qquad + (1 - \mathcal{B}_i)\ p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., 1 - \mathcal{I}_i, 1 - \mathcal{B}_i, \mathcal{U}_i ..., \mathcal{U}_N, t) \\
& \qquad + (1 - \mathcal{U}_i)\ p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, ..., 1 - \mathcal{I}_i, \mathcal{B}_i, 1 - \mathcal{U}_i ..., \mathcal{U}_N, t) \Big] \\
& - \sum_{i=1}^{N} \left[\beta \left(\sum_{j=1}^{N} J_{ij} \mathcal{B}_j \right) \mathcal{I}_i + \gamma\ \mathcal{B}_i + \nu \mathcal{I}_i + \mu\ (\mathcal{I}_i + \mathcal{B}_i + \mathcal{U}_i) \right]\ p(...\mathcal{I}_i, \mathcal{B}_i, \mathcal{U}_i ...).
\end{aligned}
\tag{1}
$$

The expectation value for the total number of ignorant individuals in the population at a given time t is defined by

$$
\langle I \rangle \quad = \quad \sum_{\mathcal{I}_1=0}^{1} \sum_{\mathcal{B}_1=0}^{1} \sum_{\mathcal{U}_1=0}^{1} \sum_{\mathcal{I}_2=0}^{1} ... \sum_{\mathcal{U}_N=0}^{1} \left(\sum_{i=1}^{N} \mathcal{I}_i \right) \cdot p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, \mathcal{I}_2, ..., U_N, t), \tag{2}
$$

and its time evolution is given by

$$
\frac{d}{dt} \langle I \rangle \quad = \quad \sum_{\mathcal{I}_1=0}^{1} \sum_{\mathcal{B}_1=0}^{1} \sum_{\mathcal{U}_1=0}^{1} \sum_{\mathcal{I}_2=0}^{1} ... \sum_{\mathcal{U}_N=0}^{1} \left(\sum_{i=1}^{N} \mathcal{I}_i \right) \cdot \frac{d}{dt} p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, \mathcal{I}_2, ..., U_N, t). \tag{3}
$$

Inserting the master equation into Equation (3), after some computations, we obtain the dynamic equation for the mean quantity of ignorant individuals in the population:

$$
\frac{d}{dt} \langle I \rangle \quad = \quad -\beta\ \langle IB \rangle_1 - \nu \langle I \rangle - \mu \langle I \rangle + \mu (\langle I \rangle + \langle B \rangle + \langle U \rangle). \tag{4}
$$

Similarly, for the expectation value of the total number of believers and unbelievers, we obtain the following dynamic equations:

$$
\frac{d}{dt} \langle B \rangle \quad = \quad \beta\ \langle IB \rangle_1 - \gamma \langle B \rangle - \mu \langle B \rangle \tag{5}
$$

$$
\frac{d}{dt} \langle U \rangle \quad = \quad \gamma \langle B \rangle + \nu \langle I \rangle - \mu \langle U \rangle. \tag{6}
$$

The dynamics of the first moments depend on the second moment:

$$
\langle IB \rangle_1 \quad = \quad \sum_{\mathcal{I}_1=0}^{1} \sum_{\mathcal{B}_1=0}^{1} \sum_{\mathcal{U}_1=0}^{1} \sum_{\mathcal{I}_2=0}^{1} ... \sum_{\mathcal{U}_N=0}^{1} \left(\sum_{i=1}^{N} \sum_{j=1}^{N} (J^1)_{ij} \mathcal{I}_i \mathcal{B}_j \right) \cdot p(\mathcal{I}_1, \mathcal{B}_1, \mathcal{U}_1, \mathcal{I}_2, ..., U_N, t)
$$

which is the mean number of ignorant and believer neighbors. We can now proceed by computing the dynamic equation for the second moment $\langle IB\rangle_1$, or we can close the ODE system (4)–(6) by approximating $\langle IB\rangle_1$ by a mathematical formula involving only the first moments $\langle I\rangle$, $\langle B\rangle$, and $\langle U\rangle$. Here, we close the ODE system (4)–(6) using the mean-field approximation.

Let us assume that the individuals in the population are distributed in a regular network, where all individuals have the same number of neighbors Q, and hence,

$$\sum_{j=1}^{N} J_{ij} = Q\,.$$

In the mean-field approximation, the exact number of believers who are neighbors of a certain individual i is approximated by the average of the number of believers in the entire population:

$$\sum_{j=1}^{N} J_{ij}\mathcal{B}_j \approx Q\frac{\langle B\rangle}{N}, \quad \forall\, i = 1, ..., N.$$

Hence, the second moment $\langle IB\rangle_1$ is approximated by

$$\langle IB\rangle_1 \approx \frac{Q}{N}\langle I\rangle\langle B\rangle,$$

and the ODE system (4)–(6) transforms into the closed system

$$\frac{d}{dt}\langle I\rangle = -\beta\,\frac{Q}{N}\langle I\rangle\langle B\rangle - \nu\langle I\rangle - \mu\langle I\rangle + \mu(\langle I\rangle + \langle B\rangle + \langle U\rangle) \tag{7}$$

$$\frac{d}{dt}\langle B\rangle = \beta\,\frac{Q}{N}\langle I\rangle\langle B\rangle - \gamma\langle B\rangle - \mu\langle B\rangle \tag{8}$$

$$\frac{d}{dt}\langle U\rangle = \gamma\langle B\rangle + \nu\langle I\rangle - \mu\langle U\rangle. \tag{9}$$

We observe that more complex ODEs can be obtained by using higher-order moment closures (see [18,20]).

Next, let the normalized state variables $I(t) = \langle I\rangle/N$, $B(t) = \langle B\rangle/N$ and $U(t) = \langle U\rangle/N$ denote the mean densities of ignorant, believer, and unbeliever individuals in the population; then, we normalize the time scale $\tau = (\gamma+\mu)t$ by the mean believing/spreading period $1/(\gamma+\mu)$. Hence, $I(\tau) + B(\tau) + U(\tau) = 1$ and Equations (7)–(9) are rescaled to the following ODE system:

$$\frac{dI}{d\tau} = -R_0 BI - (s+f)I + f \tag{10}$$

$$\frac{dB}{d\tau} = R_0 BI - B \tag{11}$$

$$\frac{dU}{d\tau} = (1-f)B + sI - fU\,; \tag{12}$$

where

(a) $f = \mu/(\gamma+\mu) > 0$, typically very small, is the mean birth and death rates in the time unit given by the mean believing/spreading period (τ);

(b) $s = \nu/(\gamma+\mu)$ is the information search rate in the time unit (τ); and

(c) $R_0 = \beta Q/(\gamma+\mu)$ is the so-called basic reproductive number R_0 (see [21]) in epidemiological models: i.e., R_0 is the rate at which the expected number of ignorant individuals become believers through the influence of the expected number of believer/spreader individuals in the time unit (τ).

Stable Stationary States

Let the stationary values of ignorants, believers, and unbelievers of the rumor be denoted by I^*, B^*, and U^*, respectively.

The stationary states of the ODE system (10)–(12) are given by

$$I_0^* = \frac{f}{s+f}, \quad B_0^* = 0 \quad \textit{and} \quad U_0^* = \frac{s}{s+f}, \tag{13}$$

and by

$$I^* = \frac{1}{R_0} \tag{14}$$

$$B^* = f\left(1 - \frac{1}{R_0}\right) - \frac{s}{R_0} \geq 0 \tag{15}$$

$$U^* = (1-f)\left(1 - \frac{1}{R_0}\right) + \frac{s}{R_0}. \tag{16}$$

From Equation (15), we observe that the believers' stationary state decreases linearly with the information search rate s (see also Figure 2), and the *critical information search rate*, which is the rate at which the believers' stationary state vanishes, is

$$s_C = f(R_0 - 1). \tag{17}$$

Since f is a small number, we assume in this paper that $0 < s_C = f(R_0 - 1) < 1$. We observe that the stationary states (I^*, B^*, U^*) only hold for $s \leq s_C$ because of the natural restriction that $B^* \geq 0$. If $s = s_C$, then there is a single equilibrium $(I_0^*, B_0^*, U_0^*) = (I^*, B^*, U^*)$.

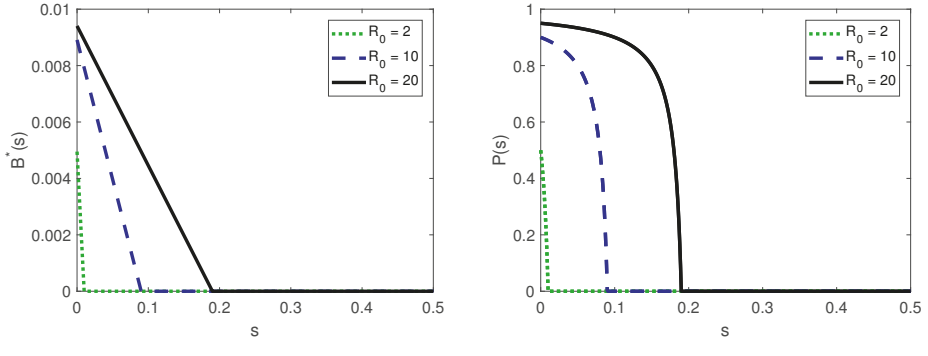


Figure 2. (**left**) The stationary value of believers $B^*(s)$ and (**right**) the probability that an ignorant individual does not search for real information to become believer $\mathcal{P}(s)$, which depends on the information search rate s. The other parameter is $f = 0.01$.

Lemma 1. *For $s < s_C$, the stationary states (I_0^*, B_0^*, U_0^*) are unstable, and the stationary states (I^*, B^*, U^*) are stable. Furthermore, for $s > s_C$, the stationary states (I_0^*, B_0^*, U_0^*) are stable.*

Proof. The Jacobean matrix of the ODE system (10)–(12) is given by

$$J(I,B,U) = \begin{bmatrix} -R_0\,B - s - f & -R_0\,I & 0 \\ R_0\,B & R_0\,I - 1 & 0 \\ s & 1-f & -f \end{bmatrix}$$

The eigenvalues of the Jacobean matrix $J(I_0^*, B_0^*, U_0^*)$ are

$$\lambda_1 = -f\,, \qquad \lambda_2 = -s - f \qquad \text{and} \qquad \lambda_3 = \frac{f(R_0-1)-s}{s+f}\,.$$

Hence, all eigenvalues have a negative real part if and only if $s > f(R_0 - 1) = s_C$. The eigenvalues of the Jacobean matrix $J(I^*, B^*, U^*)$ are

$$\lambda_1 = -f\,, \qquad \lambda_2 = -1/2\,fR_0 - 1/2\,\sqrt{f^2{R_0}^2 + 4\,s + 4\,f - 4\,fR_0} \qquad \text{and}$$
$$\lambda_3 = -1/2\,fR_0 + 1/2\,\sqrt{f^2{R_0}^2 + 4\,s + 4\,f - 4\,fR_0}\,.$$

Hence, all eigenvalues have a negative real part if and only if

$$f^2R_0^2 > f^2{R_0}^2 + 4\,s + 4\,f - 4\,fR_0.$$

This is equivalent to $s < f(R_0 - 1) = s_C$. □

3. Nash and Evolutionarily Stable Information Search Strategies

In this section, we consider a game in which individuals have to decide between searching and not searching for real information to avoid believing the false rumor. Here, we define the Nash and evolutionarily stable information search strategies (see [7,10,11]).

S denotes the probability that an ignorant individual will choose to search for information. This probability S is the individual's information search strategy in the game. The uptake level of searching for information in the population is the proportion of individuals who will choose to search for real information, i.e., the mean of all information search strategies. We denote the uptake level of searching for information by s, i.e., the population information search strategy.

Let b_{L} and c_{L} denote the benefits and the costs of searching for information, respectively, and let $v = b_{\mathrm{L}} - c_{\mathrm{L}}$ denote the value of the information search. We define the payoff of an ignorant individual who searches for real information and does not believe in the false rumor by v.

Let b_{B} and c_{B} denote the benefits and the costs of believing the rumor, respectively, and let $y = b_{\mathrm{B}} - c_{\mathrm{B}}$ denote the impact measure that assesses the risk of believing the rumor.

Let $\mathcal{P}(s)$ denote the probability that an ignorant individual, who does not search for real information, becomes a believer for a proportion s of individuals in the population who search for information. The probability $\mathcal{P}(s)$ uses the stable stationary states of ignorant and believer individuals computed in Lemma 1:

$$\mathcal{P}(s) \;=\; \frac{R_0B^*I^*}{R_0B^*I^* + fI^*} = \frac{f(R_0-1)-s}{fR_0-s}\,, \qquad if \quad s < s_C. \tag{18}$$

If $s \geq s_C$, then $B^* = 0$, and thus, $\mathcal{P}(s) = 0$ (see Figure 2). In particular, $\mathcal{P}(0) = (R_0 - 1)/R_0$. We define the payoff of an ignorant individual who does not search for real information and believes the rumor by

$$y\mathcal{P}(s).$$

The *expected information search payoff* $E(S,s)$ of an individual with an information search strategy S in a population with an information search strategy s is

$$\begin{aligned} E(S,s) &= vS + y\mathcal{P}(s)(1-S) \\ &= y\mathcal{P}(s) + S(v - y\mathcal{P}(s)). \end{aligned} \tag{19}$$

Nash and evolutionarily stable information search strategies are the typical strategies studied in game theory (see [10,12]).

A population information search strategy $s = S^*$ is an *information search Nash equilibrium* if

$$\Delta_{S^* \to S'} = E(S', S^*) - E(S^*, S^*) \leq 0\,, \tag{20}$$

for every strategy $S' \in [0,1]$. By Equation (19), an information search strategy S^* is a Nash equilibrium if and only if

$$(S' - S^*)(v - y\mathcal{P}(S^*)) \leq 0.$$

Let $W \equiv y(R_0 - 1)/R_0$ be the threshold for believing a rumor, where $(R_0 - 1)/R_0 = \mathcal{P}(0)$. The remark below follows, for instance, from Lemma 1 in [10].

Remark 1. *An information strategy S^* is a Nash equilibrium if and only if S^* satisfies one of the following conditions:*

(a) $S^* = 0$ *and* $v \leq W$, *with*

$$E(0,0) = W; \quad \text{or}$$

(b) $S^* \in (0,1)$ *and* $v = y\mathcal{P}(S^*)$, *with* $\mathcal{P}(S^*) < \mathcal{P}(0)$ *and*

$$E(S^*, S^*) = y\mathcal{P}(S^*) = v; \quad \text{or}$$

(c) $S^* = 1$ *and* $v \geq 0$, *with*

$$E(1,1) = v.$$

Hence, for every $S^* > 0$, $E(S^*, S^*) = v$ is constant, with $|v| < |W|$. We observe that (i) for every $S^* \in (0, s_C)$, $\mathcal{P}(S^*) > 0$, and (ii) for every $S^* \in [s_C, 1)$, $\mathcal{P}(S^*) = 0$, and thus, $E(S^*, S^*) = 0 = v$. In Figure 3, we plot the Nash information search strategies $s = S^*$ for each mixed Nash strategy with the value of information $v = y\mathcal{P}(s)$ and for pure Nash strategies $S^* = 0$ and $S^* = 1$.

To define an evolutionarily stable information search strategy, we start by assuming that all individuals in the population opt for an individual information search strategy S. If a group of size ε chooses a different individual information search strategy S', then the population information search strategy becomes

$$s(\varepsilon) = (1 - \varepsilon)S + \varepsilon S'.$$

A population information search strategy S^* is a *left evolutionarily stable information search strategy* if there is a $\varepsilon_0 > 0$ such that for every $\varepsilon \in (0, \varepsilon_0)$ and for every $S' < S^*$,

$$\Delta E_{S^* \to S'}(s(\varepsilon)) = E(S', s(\varepsilon)) - E(S^*, s(\varepsilon)) = (S' - S^*)\,(v - y\mathcal{P}(s(\varepsilon))) < 0\,.$$

The definition of a *right evolutionarily stable information search strategy* is similar. A population information search strategy S^* is an *evolutionarily stable information search strategy* if it is a left and right evolutionarily stable strategy.

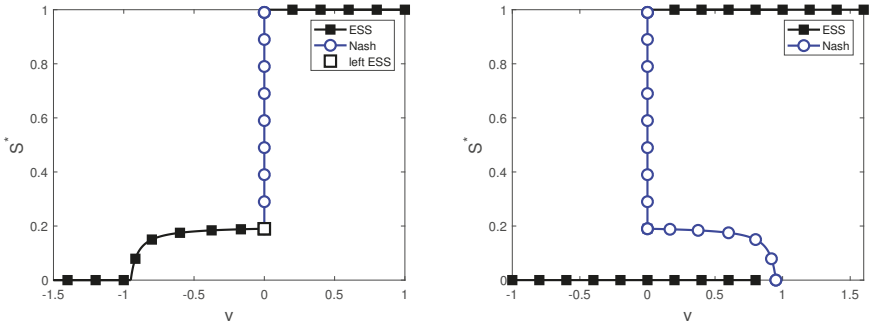


Figure 3. Nash and evolutionarily stable information search strategies s, depending on the information search values v. On the (**left**), a negative impact measure $y = -1$ is considered, and on the (**right**), a positive impact measure $y = 1$ is considered. The blue line corresponds to Nash equilibria (that are not ESI strategies), and the black line corresponds to evolutionarily stable information (ESI) search strategies. Other parameter values: $R_0 = 20$, $f = 0.01$, and $s_C = 0.19$.

Theorem 1. *A Nash search strategy S^* is an evolutionarily stable information (ESI) search strategy if and only if S^* satisfies one of the following conditions:*

(i) For positive impact measures $y \geq 0$,

- *(a) $S^* = 0$ and $v < y\mathcal{P}(0)$; or*
- *(b) $S^* = 1$ and $v > 0$.*

(ii) For negative impact measures $y \leq 0$,

- *(a) $S^* = 0$ and $v \leq y\mathcal{P}(0)$; or*
- *(b) $S^* \in (0, s_C)$ and $v = y\mathcal{P}(S^*)$; or*
- *(c) $S^* = 1$ and $v > 0$.*

Moreover, S^* is a Nash equilibrium and a left (and not a right) evolutionarily stable information search strategy if and only if $S^* = s_C$, $v = 0$, and $y > 0$.

Hence, S^* is a Nash equilibrium and not an evolutionarily stable information search strategy if $S^* \in [0, 1]$ and $v = y\mathcal{P}(S^*)$ and $y > 0$.

In Figure 3, we plot the evolutionarily stable information search strategies $s = S^*$ for each value v of searching for information.

Proof. The proof follows from Lemma 2 in [10], noting that v is negative and $\mathcal{P}(S^*)$ is strictly decreasing for $S^* \in (0, s_C)$. □

4. Optimal Strategies

In this section, we compute the optimal information (OI) search strategy for every value of searching for information and every value of the rumor impact measure, under the assumption that all individuals adopt the same information search strategy $s = S$ (homogeneous strategy). Let $s_C = f(R_0 - 1)$ be the critical information search strategy. Let

$$\tilde{E}(s) \equiv \tilde{E}(s; v) = vs + y\mathcal{P}(s)(1 - s),$$

for $0 \leq s \leq s_C$.

Lemma 2. *Assume that f is sufficiently small, where $f < 1/R_0$.*

(i) $\tilde{E}''(s) < 0$ *for positive impact measure values* $y > 0$;
(ii) $\tilde{E}''(s) = 0$ *for null impact measure values* $y = 0$; *and*
(iii) $\tilde{E}''(s) > 0$ *for negative impact measure values* $y < 0$.

Proof. We have

$$\tilde{E}''(s) \quad = \quad y\left(\mathcal{P}''(s)(1-s) - 2\mathcal{P}'(s)\right).$$

We observe that $\tilde{E}''(s)/y < 0$ is equivalent to $2\mathcal{P}'(s) > \mathcal{P}''(s)(1-s)$. By Equation (18), we have $\mathcal{P}'(s) = -f/(fR_0 - s)^2$ and $\mathcal{P}''(s) = -2f/(fR_0 - s)^3$. Thus, $2\mathcal{P}'(s) > \mathcal{P}''(s)(1-s)$ is equivalent to

$$-2f/(fR_0 - s)^2 > -2f/(fR_0 - s)^3.$$

Hence, we conclude that $\tilde{E}''(s)/y < 0$ is equivalent to $fR_0 < 1$. □

By Equation (19), the expected information search payoff is given by

$$E(s;v) \equiv E(s,s;v) = \begin{cases} \tilde{E}(s;v) & if\ s \le s_C \\ vs & if\ s > s_C \end{cases}. \tag{21}$$

Since $\mathcal{P}(s_C) = 0$, we note that $\tilde{E}(s_C;v) = v\,s_C$, and thus, E is a continuous function (see also Figure 4). The optimal information (OI) search strategy (or strategies, eventually) is

$$s_O \equiv s_O(v) = \arg\max_{0 \le s \le 1} E(s;v).$$

The expected payoff of the optimal information search strategy is $E_O(v) = E(s_O(v);v)$. Let $s_{ESI}(v)$ denote the evolutionarily stable information search strategy (or strategies, eventually). The expected payoff of the evolutionarily stable information search strategy is $E_{ESI}(v) = E(s_{ESI}(v);v)$. Let $s_{Nash}(v)$ denote the Nash search strategy (or strategies, eventually) that are not evolutionarily stable information search strategies. The Nash expected payoff is $E_{Nash}(v) = E(s_{Nash}(v);v)$.

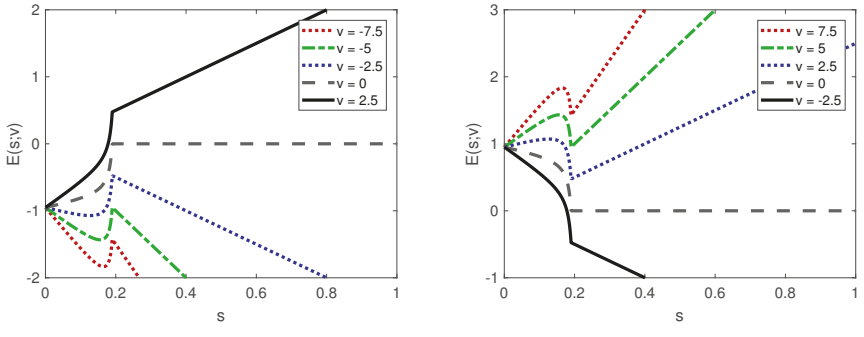


Figure 4. The expected information search payoff, depending on the information search strategy s, for different information search values v. On the (**left**), a negative impact measure $y = -1$ is considered, and on the (**right**), a positive impact measure $y = 1$ is considered. Other parameter values: $f = 0.01$, $R_0 = 20$, and $s_C = 0.19$.

4.1. The OI Search Strategy for a Positive Impact Measure

Throughout this section, let us assume that the impact measure is positive $y > 0$. Hence, by Lemma 2 (see also Figure 4), for $v \neq 0$, $\tilde{E}$ is strictly concave, and thus, the optimal information

search strategy is a pure strategy (0 or 1) or a mixed strategy s_C or $s_M(v)$, where $s_M(v)$ is the interior maximum point of $\tilde{E}(s;v)$ (when it exists).

Let $U = y(fR_0(R_0-1)+1)/(fR_0^2)$ be the positive information search threshold. Note that $0 < W < U$.

Lemma 3. *Assume that f is small, where $f < 1/R_0$. For a positive impact measure $y > 0$, the optimal information (OI) search strategy is*

(a) for $v < W$, $s_O(v) = 0$, with $E(s_O(v)) = W$;
(b) $s_O(W) \in \{0,1\}$, with $E(s_O(W)) = W$; and
(c) for $v > W$, $s_O(v) = 1$, with $E(s_O(W)) = v$.

For a null impact measure $y = 0$, optimal information (OI) search strategies are similar to those described above, observing that $s_O(0) \in [0,1]$ (note that $W = 0$).

Proof. Since $\tilde{E}$ is strictly concave, if $E'(0) \leq 0$, then 0 or 1 is the maximum of E. Hence, let us compute the following for $E'(0) \leq 0$. The first derivative of the expected payoff is

$$E'(s) = v - y(\mathcal{P}(s) + (s-1)\mathcal{P}'(s)).$$

Since $\mathcal{P}(0) = (R_0-1)/R_0$ and $\mathcal{P}'(0) = -1/(fR_0^2)$, we have $E'(0) = v - U$. Hence, $E'(0) \leq 0$ if and only if $v \leq U$. Therefore, for $v \leq U$, 0 is the maximum point of E when $E(0) \geq E(1)$, and 1 is the maximum point of E when $E(0) \leq E(1)$. Recall that $E(0) = W$ and $E(1) = v$. Hence, $E(0) \geq E(1)$ if and only if $v \leq W$.

Finally, for $v > U > 0$, let us prove that 1 is the maximum of E. This follows from the confirmation that $E(s) < E(1)$ for every $s \leq s_C$. We observe that $E(s) < E(1)$ if and only if

$$s < fR_0 + \frac{yf}{v-y}.$$

Since $s_C = f(R_0-1)$, we confirm that the equivalence between

$$s_C = f(R_0-1) < fR_0 + \frac{yf}{v-y}$$

and $-y(1-f) < fv$ holds because of $-y(1-f) < 0 < fv$. Hence,

$$s \leq s_C < fR_0 + \frac{yf}{v-y},$$

which concludes the proof. □

Remark 2. *Assume that f is small, where $f < 1/R_0$, and the impact measure is positive $y > 0$.*

(a) $s_{ESI}(v) = s_O(v) = 0$, for $v < 0$;
(b) $0 = s_O(0) < s_{Nash}(0) \in [s_C, 1]$;
(c) $0 = s_O(v) < s_{Nash}(v) < 1$ and $s_{ESI}(v) \in \{0,1\}$, for $0 < v < W$;
(d) $0 = s_{Nash}(W) < s_{ESI}(W) = 1$ and $s_O(W) \in \{0,1\}$; and
(e) $s_{ESI}(v) = s_O(v) = 1$, for $v > W$.

For the null impact measure $y = 0$, the comparison is similar to that described above, observing that $s_O(0), s_{Nash}(0) \in [0,1]$ (note that $W = 0$).

By Lemma 3 and Remark 2, (i) for small values of the information search $v \leq W$, the optimal strategy $s_O = 0$ coincides with the evolutionarily stable information search strategy $s_{ESI} = 0$, in which

individuals never search for truthful information; (ii) for positive values of the information search $v \geq W$, the optimal strategy $s_O = 1$ coincides with the evolutionarily stable information search strategy $s_{ESI} = 1$, in which individuals always search for truthful information. In Figure 5, we compare the expected payoff of the evolutionarily stable information search $E_{ESI}(v)$ with that of the optimal information search, denoted by $E_O(v)$.

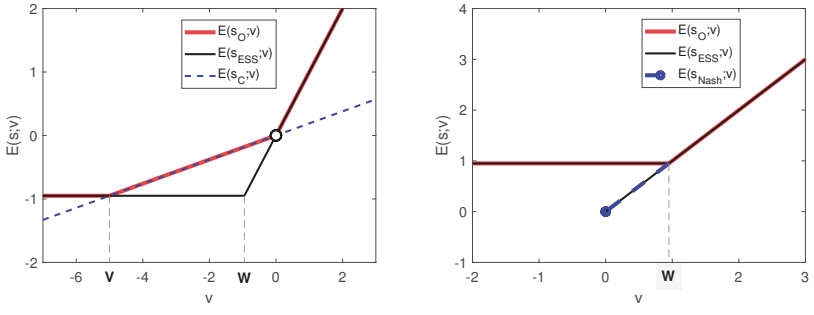


Figure 5. The expected information search payoff $E(s;v)$, depending on the value of the information search v for different search strategies s: critical search strategy s_C, OI search strategy s_O, ESI search strategy s_{ESI}, and Nash strategy s_{Nash}. On the (**left**), a negative impact measure $y = -1$ is considered, and on the (**right**), a positive impact measure $y = 1$ is considered. The other parameter values are $f = 0.01$ and $R_0 = 20$. Hence, $V = -5$ and $W = -0.95$ (*left*) or $W = 0.95$ (*right*).

4.2. The OI Search Strategy for a Negative Impact Measure

Throughout this section, let us assume that the impact measure is negative $y < 0$. Hence, by Lemma 2 (see also Figure 4), $\tilde{E}$ is strictly convex, and thus, the optimal information search strategy is a pure strategy (0 or 1) or a mixed strategy s_C for $v \neq 0$. Let $V = y/(fR_0) < 0$ be the negative information search threshold.

Lemma 4. *Assume that f is small, where $f < 1/R_0$. For a negative impact measure $y < 0$, the optimal information (OI) search strategy is*

(a) *for $v < V$, $s_O(v) = 0$, with $E(s_O(v)) = W$;*
(b) *$s_O(V) \in \{0, s_C\}$, with $E(s_O(V)) = W$;*
(c) *for $V < v < 0$, $s_O(v) = s_C$, with*

$$E(s_O(v)) = vs_C = vf(R_0 - 1) = vW/V;$$

(d) *$s_O(0) \in [s_C, 1]$, with $E(s_O(0)) = 0$; and*
(e) *for $v > 0$, $s_O(v) = 1$, with $E(s_O(v)) = v$;*

Proof. Since $\tilde{E}(s;v)$ is a linear function in v, there is only one value $V = y/(fR_0) < 0$ such that $\tilde{E}(0;V) = \tilde{E}(s_C;V)$. Furthermore, $E(0;v) > E(s_C;v)$ if and only if $v < V$. (a) If $v < V < 0$, $E(0;v) > E(s_C;v)$ and, by linearity, $E(s_C;v) > E(1;v)$. Hence, $s_O(v) = 0$. (b) If $v = V < 0$, $E(0;V) = E(s_C;V)$ and, by linearity, $E(s_C;V) > E(1;V)$. Hence, $s_O(V) = 0$ or $s_O(V) = s_C$. (c) If $V < v < 0$, $E(s_C;v) > E(0;v)$ and, by linearity, $E(s_C;v) > E(1;v)$. Hence, $s_O(v) = s_C$. (d) If $v = 0$, $E(s_C;v) > E(0;v)$ and, by linearity, $E(s;0) = E(s_C;0)$ for all $s \in [s_C, 1]$. Hence, $s_O(0) \in [s_C, 1]$. (e) If $v > 0$, $E(s_C;v) > E(0;v)$ and, by linearity, $E(1;v) > E(s_C;v)$. Hence, $s_O(v) = 1$. □

Since $V = y/(fR_0) < y(R_0 - 1)/R_0 = y\mathcal{P}(0) \equiv W$, we state the following remark.

Remark 3. *Assume that f is small, where $f < 1/R_0$ and the impact measure is negative $y < 0$.*

(a) $s_{ESI}(v) = s_O(v) = 0$, *for* $v < V$;
(b) $s_{ESI}(V) = 0$, $s_O(V) \in \{0, s_C\}$;
(c) $s_{ESI}(v) = 0 < s_C = s_O(v)$, *for* $V < v < W$;
(d) $0 < s_{ESI}(v) = \mathcal{P}^{-1}(v) < s_C = s_O(v)$, *for* $W < v < 0$;
(e) $s_{Nash}(0), s_O(0) \in [s_C, 1]$; *and*
(f) $s_{ESI}(v) = s_O(v) = 1$, *for* $v > 0$;

By Lemma 4 and Remark 3, (i) for small values of searching for information $v \leq V$, the optimal strategy $s_O = 0$ coincides with the evolutionarily stable information search strategy $s_{ESI} = 0$, in which individuals never search for truthful information; (ii) for positive values of searching for information $v > 0$, the optimal strategy $s_O = 1$ coincides with the evolutionarily stable information search strategy $s_{ESI} = 1$, in which individuals always search for truthful information; (iii) for intermediate values of searching for information $V < v < 0$, the optimal strategy coincides with the critical information search rate s_C, which eradicates the rumor. This value is above the value given by the evolutionarily stable information search strategy s_{ESI} that is not able to eradicate the rumor and yields a lower expected information search payoff $E_{ESI}(v) < E_O(v)$ (see Figure 5).

5. Evolutionary Information Search Dynamics

Evolutionary information search dynamics is introduced here (see [11–13]), under the assumption that all individuals adopt the same information search strategy $s = S$ (homogeneous strategy).

Consider a case in which a small group of individuals of size ε modify their search strategy from the population information search strategy S to $S + \Delta S$. The change in the expected information search payoff satisfies

$$\frac{\Delta E_{S\to(S+\Delta S)}}{\Delta S} = \frac{E(S+\Delta S, s(\varepsilon)) - E(S, s(\varepsilon))}{\Delta S} = v - y\mathcal{P}(s(\varepsilon)), \tag{22}$$

where $s(\varepsilon) = (1-\varepsilon)S + \varepsilon(S+\Delta S) = S + \varepsilon\Delta S$ defines the new population search strategy.

Let $s(\tau)$ be the population information search strategy adopted at time τ. Hence, we define the *evolutionary information search dynamics* by

$$\frac{ds}{d\tau} = \eta(s) \lim_{\Delta S\to 0} \frac{\Delta E_{S\to(S+\Delta S)}}{\Delta S} = \eta(s)(v - y\mathcal{P}(s)), \tag{23}$$

where $\eta(s) \geq 0$ is a smooth map that measures the *information search strategy adaptation speed* of the population.

A point s is a *dynamic equilibrium* of the evolutionary information search dynamics if and only if $ds/d\tau = 0$. Hence, a point s is a dynamic equilibrium if and only if

$$\text{(i)} \quad \eta(s) = 0 \quad \text{or} \quad \text{(ii)} \quad v = y\mathcal{P}(s).$$

Recall that f is assumed to be small, and thus, $s_C = f(R_0 - 1) < 1$; $\mathcal{P}(1) = 0$, and $W = y\mathcal{P}(0)$ is the rumor belief threshold. As usual (see [10]), we assume the following for η: (i) $\eta(s) > 0$, for all $0 < s < 1$; (ii) if $v < W$, then $\eta(0) = 0$ and $\eta'(0) > 0$; (iii) if $v > W$, then $\eta(0) > 0$; (iv) if $v > 0$, then $\eta(1) = 0$ and $\eta'(1) < 0$; and (v) if $v < 0$, then $\eta(1) > 0$.

We use the standard definition of left, right, and global attractors for a dynamic equilibrium p (see [10]).

Theorem 2. *Assume that f is small, where $f(R_0 - 1) < 1$.*

(i) For negative impact measures $y \leq 0$, the dynamic equilibria of the evolutionary information search dynamics are as follows:

- *(a) for $v < 0$, the evolutionarily stable information search strategy $s_{ESI}(v)$ is a global attractor;*
- *(b) for $v = 0$, the Nash information search strategies $s_{Nash}(v) \in [s_C, 1]$ are equilibria points, and s_C is a left (and not right) attractor;*
- *(c) for $v > 0$, the evolutionarily stable information search strategy $s_{ESI}(v) = 1$ is a global attractor.*

(ii) For positive impact measures $y \geq 0$, the dynamic equilibria of the evolutionary information search dynamics are as follows:

- *(a) for $v < W$, the evolutionarily stable information search strategy $s_{ESI}(v) = 0$ is an attractor (also global for $v < 0$);*
- *(b) for $0 \leq v \leq W$, the Nash information search strategies $s_{Nash}(v)$ are dynamical equilibria, but not attractors; and*
- *(c) for $v > 0$, the evolutionarily stable information search strategy $s_{ESI}(v) = 1$ is an attractor (also global for $v > W$).*

In Figure 6, we show the dynamics described above. For advantageous rumors, we observe the existence of a bi-stability region, where the evolutionarily stable information search strategies in which no one searches (persistence of the rumor) or everyone searches (eradication of the rumor) are the attractors, and the Nash equilibria form the boundary of the basins of attraction of the two attractors. Hence, the Nash equilibria are unstable equilibria and are thus not observed (at least for large periods), but have the interesting property of determining the basin of attraction of the attractors. For harmful rumors, we observe that for negative values of the information search $v < 0$, the evolutionary information search dynamic drives the population search strategy to an evolutionarily stable information search strategy that is lower than the critical information search rate $s_{ESI} < s_C$. Hence, to eradicate the rumor, a forcing mechanism must be implemented to increase the population search strategy to (or above) the critical information search rate s_C. For positive values of the information search $v > 0$, the evolutionary information search dynamic drives the population search strategy to the evolutionarily stable information search strategy, in which individuals always search for truthful information $s_{ESI} = 1$, and thus, the rumor is eradicated. Hence, a recommendation for legislators and policymakers who aim to eradicate harmful rumors is to make the search for truthful information free or rewarding, i.e., information search value $v \geq 0$. Truthful public social media campaigns can help by facilitating access to information.

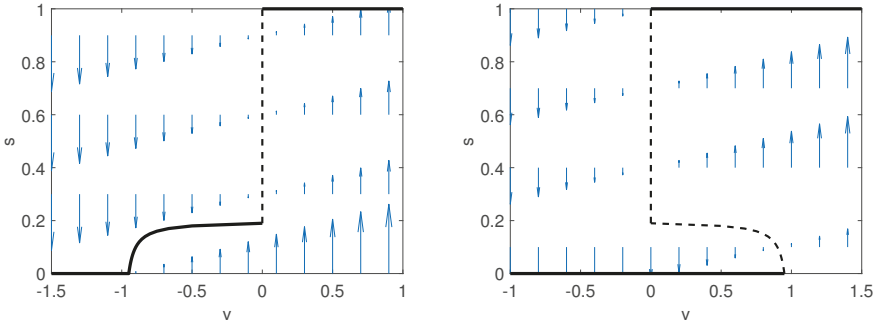


Figure 6. The stable (solid line) and the unstable (dashed line) equilibria of the evolutionary information search dynamics. On the (**left**), a negative impact measure $y = -1$ is considered, and on the (**right**), a positive impact measure $y = 1$ is considered. Parameter values: $f = 0.01$ and $R_0 = 20$.

The proof of the above theorem follows similarly to the proofs of Theorems 6–8 in [10].

Proof. Let $y < 0$ (the proof follows similarly for $y > 0$). To simplify the presentation of the proof, let us introduce the function

$$F(s) = \eta(s)(v - y\mathcal{P}(s))$$

such that $ds/d\tau = F(s)$.

(a) If $v \leq y\mathcal{P}(0)$, then $\eta(0) = 0$, and thus, $F(0) = 0$. Hence, $s = 0$ is an equilibrium point. Since $\mathcal{P}(s)$ is decreasing, for every $s' \in (0, 1]$, $\mathcal{P}(s') < \mathcal{P}(0)$, and thus, $F(s') < 0$. Hence,

$$\lim_{\tau \to \infty} s(\tau; s') = 0\,,$$

and therefore, $s = 0$ is a global attractor.

(b) If $y\mathcal{P}(0) < v < 0$ and s^* is such that $v = y\mathcal{P}(s^*)$, then $F(s^*) = 0$, and s^* is an equilibrium point. Since $\mathcal{P}(s)$ is decreasing, for every $s' \in [0, s^*)$, $\mathcal{P}(s') > \mathcal{P}(s^*)$, and thus, $F(s') > 0$. Hence,

$$\lim_{\tau \to \infty} s(\tau; s') = s^*\,,$$

and therefore, s^* is a left attractor in $[0, s^*)$. For every $s' \in (s^*, 1]$, $\mathcal{P}(s') < \mathcal{P}(s^*)$, and thus, $F(s') < 0$. Hence,

$$\lim_{\tau \to \infty} s(\tau; s') = s^*\,,$$

and therefore, s^* is a right attractor in $(s^*, 1]$. Hence, s^* is a global attractor.

(c) For every $s^* \in [s_C, 1]$, $\mathcal{P}(s^*) = 0$, and thus, $F(s^*) = 0$ if $v = 0$. Hence, $s^* \in [s_C, 1]$ are equilibria points. Since $\mathcal{P}(s)$ is decreasing, for every $s' \in [0, s_C)$, $\mathcal{P}(s') > \mathcal{P}(s_C) = 0$, and thus, $F(s') > 0$. Hence,

$$\lim_{\tau \to \infty} s(\tau; s') = s_C\,,$$

and therefore, s_C is a left attractor in $[0, s_C)$.

(d) If $v > 0$, then $\eta(1) = 0$, and thus, $F(1) = 0$. Hence, $s = 1$ is an equilibrium point. Since $\mathcal{P}(s)$ is decreasing, for every $s' \in [0, 1)$, $\mathcal{P}(s') > \mathcal{P}(1)$, and thus, $F(s') > 0$. Hence,

$$\lim_{\tau \to \infty} s(\tau; s') = 1\,,$$

and therefore, $s = 1$ is a global attractor. □

6. Conclusions

In this paper, we present a rumor spreading model with potential information searching in a population where individuals can be ignorants, believers, or unbelievers of the rumor. Depending on whether the impact measure, which assesses the risk of believing the rumor (fake news), is positive or negative and on the value of searching for information, we introduce an expected payoff or utility for the individuals. We derive all of the Nash and all of the evolutionarily stable information search strategies. Furthermore, we introduce evolutionary information search dynamics, whose attractors are evolutionarily stable information search strategies.

For advantageous rumors, we observe the existence of a bi-stability region, where the evolutionarily stable information search strategies are either to fully search for truthful information or not search at all. For harmful rumors, we observe that there is a single evolutionarily stable information search strategy by which individuals decide, or not, to search for information. When the benefits of searching for information outweigh the costs, i.e., the value of information search is positive, the evolutionarily stable information search strategy is to search for information with a probability of 1. However, when the value of the information search is negative, the evolutionarily stable information search strategy is smaller than the optimal information search strategy that eradicates the rumor.

In this case, unfortunately, the rumor persists. The persistence of false rumors may be quite dangerous and lead to extensive damage to the individual, as well as to all of society. For example, when a disease is spreading, some cases of vaccination with moderate side-effects can be inflated by social media, provoking fear in the population and leading to a large proportion of individuals deciding against vaccination. In an outbreak with a large transmission rate, such as COVID-19, decisions against vaccination are a major contributor to the spread of the disease and so are quite harmful to all. A recommendation for legislators and policymakers who aim to eradicate harmful rumors is to make the search for truthful information free or rewarding.

The population is assumed to be distributed in a regular spatial network, where all individuals have the same number of neighbors, and thus, all of them can equally spread the rumor. This model will be the basis for future works that involve different and more complex spatial networks, heterogeneous strategies, and higher moment closure approximations and encompass the routes of modern social media transmission.

Author Contributions: Conceptualization, J.M. and A.P.; formal analysis, J.M. and A.P.; writing—original draft, J.M. and A.P.; writing—review and editing, J.M. and A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by LIAAD-INESC TEC and FCT Fundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology) within project UID/EEA/50014/2019 and by National Funds through the FCT within Project "Modelling, dynamics and games", with reference PTDC/MAT-APL/31753/2017.

Acknowledgments: The authors thank the referees for their invaluable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Daley, D.J.; Kendall, D.G. Epidemics and rumours. *Nature* **1964**, *204*, 1118. [CrossRef] [PubMed]
2. Daley, D.J.; Kendall, D.G. Stochastic rumours. *IMA J. Appl. Math.* **1965**, *1*, 42–45. [CrossRef]
3. Goffman, W.; Newill, V.A. Generalization of epidemic theory: An application to the transmission of ideas. *Nature* **1964**, *204*, 225–228 [CrossRef] [PubMed]
4. Moreno, Y.; Nekovee, M.; Pacheco, A.F. Dynamics of rumor spreading in complex networks. *Phys. Rev. E* **2004**, *69*, 066130 [CrossRef] [PubMed]
5. Zhang, N.; Huang, H.; Duarte, M.; Zhang, J. Risk analysis for rumor propagation in metropolises based on improved 8-state ICSAR model and dynamic personal activity trajectories. *Physica A* **2016**, *451*, 403–419. [CrossRef]
6. Jia, F.; Lv, G.; Zou, G. Dynamic analysis of a rumor propagation model with Lévy noise. *Math. Meth. Appl. Sci.* **2018**, *41*, 1661–1673. [CrossRef]
7. Bauch, C.T.; Earn, D.J.D. Vaccination and the theory of games. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13391–13394. [CrossRef] [PubMed]
8. Huynh, T.L.D. Replication: Cheating, loss aversion, and moral attitudes in Vietnam. *J. Econ. Psychol.* **2020**, *78*, 102277. [CrossRef]
9. Burggraf, T.; Fendel, R.; Huynh, T.L.D. Political news and stock prices: Evidence from Trump's trade war. *Appl. Econ. Lett.* **2020**, *27*, 1–4. [CrossRef]
10. Martins, J.; Pinto, A. Bistability of Evolutionary Stable Vaccination Strategies in the Reinfection SIRI Model. *Bull. Math. Biol.* **2017**, *79*, 853–883. [CrossRef] [PubMed]
11. Maynard-Smith, J. *Evolution and the Theory of Games*; Cambridge University Press: Cambridge, UK, 1982.
12. Hofbauer, J.; Sigmund, K. *Evolutionary Games and Population Dynamics*; Cambridge University Press: Cambridge, UK, 1998.
13. Nowak, M. *Evolutionary Dynamics: Exploring the Equations of Life*; The Belknap Press of Harvard University Press: Cambridge, MA, USA, 2006.
14. Kermack, W.O.; McKendrick, A.G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond.* **1927**, *115*, 700–721.

15. Pinto, A.; Martins, J. A game theoretical analysis in a rumor spreading model based on the SIR epidemic model. In Proceedings of the 17th International Conference on Computational and Mathematical Methods in Science and Engineering (CMMSE2017), Cadiz, Spain, 4–8 July 2017; Volume 1689–1694.
16. Van Kampen, N.G. *Stochastic Processes in Physics and Chemistry*; North-Holland: Amsterdam, The Netherlands, 1992.
17. Glauber, R.J. Time-dependent statistics of the Ising model. *J. Math. Phys.* **1963**, *4*, 294–307. [CrossRef]
18. Stollenwerk, N.; Martins, J.; Pinto A. The phase transition lines in pair approximation for the basic reinfection model SIRI. *Phys. Lett. A* **2007**, *371*, 379–388. [CrossRef]
19. Stollenwerk, N.; van Noort, S.; Martins, J.; Aguiar, M.; Hilker, F.; Pinto, A.; Gomes, M.G. A spatially stochastic epidemic model with partial immunization shows in mean field approximation the reinfection threshold. *J. Biol. Dyn.* **2010**, *4*, 634–649. [CrossRef] [PubMed]
20. Martins J.; Pinto, A.; Stollenwerk, N. Stationarity in moment closure and quasi-stationarity of the SIS model *Math. Biosci.* **2012**, *236*, 126–131. [CrossRef] [PubMed]
21. Heesterbeek, J.A.P. A brief history of R0 and a recipe for its calculation. *Acta Biotheor.* **2002**, *50*, 189–204. [CrossRef] [PubMed]

entropy

MDPI

Article

The Gender Productivity Gap in Croatian Science: Women Are Catching up with Males and Becoming Even Better

Dorian Wild [1,2], Margareta Jurcic [1] and Boris Podobnik [1,2,3,*]

[1] Department of Finance, Zagreb School of Economics and Management, 10000 Zagreb, Croatia; dorian.wild7@gmail.com (D.W.); mjurcic2@zsem.hr (M.J.)
[2] Department of Finance, Luxembourg School of Business, 2453 Luxembourg, Luxembourg
[3] Faculty of Civil Engineering, University of Rijeka, 51000 Rijeka, Croatia
* Correspondence: bp@phy.hr

Received: 25 August 2020; Accepted: 23 October 2020; Published: 26 October 2020

Abstract: How much different genders contribute to citations and whether we see different gender patterns between STEM and non-STEM researchers are questions that have long been studied in academia. Here we analyze the research output in terms of citations collected from the Web of Science of males and females from the largest Croatian university, University of Zagreb. Applying the Mann–Whitney statistical test, for most faculties, we demonstrate no gender difference in research output except for seven faculties, where males are significantly better than females on six faculties. We find that female STEM full professors are significantly more cited than male colleagues, while male non-STEM assistant professors are significantly more cited than their female colleagues. There are ten faculties where females have the larger average citations than their male colleagues and eleven faculties where the most cited researcher is woman. For the most cited researchers, our Zipf plot analyses demonstrate that both genders follow power laws, where the exponent calculated for male researchers is moderately larger than the exponent for females. The exponent for STEM citations is slightly larger than the exponent obtained for non-STEM citations, implying that compared to non-STEM, STEM research output leads to fatter tails and so larger citations inequality than non-STEM.

Keywords: power law; Zipf law; gender productivity gap

1. Introduction

The gender productivity gap in academia, well known as the 'productivity puzzle'—that on average men publish more papers than women—still persists across the countries, but the trend is ramping down over time and the gap considerably varies between different subfields of science [1,2]. Many theories explain the gender productivity difference ranging from family responsibilities [3,4] to career absences [5], to mention a few.

In contrast to the OECD finding that there are more female than male undergraduate and graduate students in most countries [6], there are relatively few female than male full professors in many countries. According to the US National Science Foundation, female scientists earn almost half the PhDs in science and engineering (STEM fields) in the US, but comprise only 21% of full science professors and surprisingly only 5% of full professors in engineering [7]. There is persisting gender inequality in hiring [8], female scientists continue to face discrimination in earnings [7], funding [9], patenting [10], prizes and awards [11], and grant applications [12]. Moreover, female scientists publish significantly fewer papers in areas in which research is expensive, and it is much less likely that females are listed as either first or last author on a paper [2]. Holman, Stuart-Fox, and Hauser estimate that men are invited by journals to submit papers at approximately double the rate of women [13].

Analyses accomplished in the Quebec region reported that after women have passed the age of about 38, they receive less funding for research than men, and are at a slight disadvantage in terms of the scientific impact (measured by citations) of their publications [14]. However, a few papers reported that the gender gap is diminishing [1,15], e.g., Xie and Shauman reported that the gender differences in research productivity declined overtime with the female-to-male ratio increasing from about 60 percent in the late 1960s to 75 to 80 percent in the late 1980s and early 1990s [15]. Le Moine shows that the concentration of women among researchers who publish a single article is greater than for men, while their representation among 'star' scientists is less [16].

How children and marriage affect men and women is not conclusive. Most studies report the positive effect of marriage on scientific productivity, but Prpic [17] shows that men more than women experience benefitsdue to the presence of a spouse. Fox [3] reports that unmarried men are the least productive of all. Married women, particularly those married for the second or third time, exhibit a higher level of productivity. Stack shows that women with preschool-aged children publish less than other women [18] . For example, the divorce rate among tenured females is 50% higher than that of tenured men [19]. Tenured female scientists are almost three times more likely than male colleagues to be single without children [20]. Obviously, the time, energy, and money devoted to raising children can reduce time devoted to science. Due to thinking that raising kids is more a female responsibility, men with children make them more productive than women with children [17]. Moreover, female postdocs who plan or become parents decide to abandon research careers up to twice as often as men [21], confirming that children and marriages do affect carriers in science, but not equally males and females.

In terms of citations, it was revealed that the smaller number of citations received by females is first because on average, men like to cite their own papers 56% more than women [22]. Second, females generally publish less [1]. However, on average, papers written by females receive more citations than papers written by males [23], implying that women often have a higher impact (citations) per publication [24–26]. Duch et al. hypothesized and confirmed that the higher the resource requirements on research, the greater is the difference in the publication rates between females and males in favor of males. However, the gender differences in publication rate and citations are discipline-specific [26].

To stress gender differences across different research areas, for the Italian academia the authors showed that females are more productive than male colleagues in medical sciences, agriculture, veterinary sciences, and earth sciences, while men dominate in industrial and information engineering, chemical sciences, physical sciences, and mathematics and information sciences [27]. Abramo et al. concluded that there is a significant number of scientific fields where women's performance cannot be considered to be inferior compared to men's.

2. Materials and Methods

To test whether the male and female researchers comprise two distinct subgroups and not a unique, in our paper we apply the Mann–Whitney U test that is intended to measure the difference between two populations, in our analyses the difference in ranks of male and female citations. Generally, when the test is applied in practice, the total citations data (males and females collected together) must be first sorted in ascending order. Combining all citations values in a single array, but keeping information of which sample each observation comes from, the U test first ranks all scientists according to their citations from smallest to largest and then separately sums up all female and male ranks, thus, taking the gender into account. We denote these sums by R_1 and R_2, whereby N_1 and N_2 we denote the respective sample sizes, in our case representing the numbers of male and female researchers. Since the the test is conceived to incorporate both citations and the sample size which is the number of professors of a particular gender, the test is actually designed to evaluate citations per capita of a given gender. The test statistic quantifying the difference between the rank sums is defined as:

$$U_1 = N_1 N_2 + \frac{N_1(N_1+1)}{2} - R_1. \quad (1)$$

or

$$U_2 = N_1 N_2 + \frac{N_2(N_2+1)}{2} - R_2. \tag{2}$$

When the two-tail test is applied, the two-tail test statistic U is taken to be the smaller of U_1 and U_2.

The distribution of U test is symmetrical where a mean and variance equal

$$\mu_U = \frac{N_1 N_2}{2}, \tag{3}$$

$$\sigma_U = \frac{N_1 N_2 (N_1 + N_2 + 1)}{12}. \tag{4}$$

If N_1 and N_2 are at least equal to 8, the distribution of U is approximately Gaussian so that

$$z = (U - \mu_U)/\sigma_U \tag{5}$$

is standardized Gaussian distributed with mean zero and variance 1.

Here, the null hypothesis is that the distributions of citations per capita of the two subgroups (men and women) are the same. The null hypothesis is rejected for values of the test statistic falling into either tail of its sampling distribution. If a one-tailed test is performed, the alternative hypothesis suggests that the variable of one group is larger than the other group.

3. Results

3.1. Empirical Evidence

Here, we analyze the publishing career of 3331 active scientists from their publication record in theWeb of Science (WoS) database who have the affiliation of University of Zagreb, which is the largest Croatian university. As a proxy for research excellence, we take the the number of citations according to WoS. With the criterion that only affiliation matters, we exclude the total number of citations the researchers may have collected being on other institutions. The data had been collected between 5th and 10th November 2019. This choice may serve as a proxy for the real contribution of scientists in their university rankings. Due to the importance of physics, biology, and chemistry in science in general, we analyze these subfields of natural sciences individually even though they all belong to a single institution, the Faculty of Natural Sciences. For each scientist we identify the gender, the number of publications, citations, and whether he/she belongs to a STEM field (science, technology, engineering and mathematics) or not. We put particular focus on STEM field (science, technology, engineering, and mathematics) or not. We put particular focus on STEM fields because the recent studies indicate the prominence of STEM for the country's growth. Specifically, the World Economic Forum [28] and National Academies [29,30] studies indicate that STEM fields are key in economic development.

Among the researchers, we first find 1837 male and 1485 female scientists among the full, associate, and assistant professors. However the gender contribution in different subgroups clearly reveal that academia is becoming more open to females. Specifically, while male full professors substantially outnumber female colleagues (≈50%), i.e., 778 vs. 518, the gender number gap is not only diminishing at the level of associate professor, where there are 415 males and 343 females (≈20% more males), but even brings more females than males at the lowest professor rank, where there are 587 male assistant professors and 613 female assistant professors.

Our focus on bibliometric data limits our analysis to only publishing careers that are easier to analyze than teaching careers. Nevertheless, our efforts constitute an attempt to quantify gender inequality in STEM/non-STEM publications and citations at the largest Croatian university in a former socialist country.

The total numbers of male and female scientists hide an underlying disciplinary differences, as the fraction of women, is as low as 9% in physics, 30% mathematics, and 16% computer science (Faculty of Electrical Engineering and Computing), but female scientists constitute the majority in biology and

chemistry with 65% and 56%, respectively (see Table 1). The Faculty of Natural Sciences comprising all-natural sciences, physics, biology, chemistry, including mathematics have 49.49% of male professors.

Table 1. Men vs. Women—citation comparison across different University constituents. Shown are z-scores of Equation (5) and corresponding p-values for both one- and two-tailed statistics where in brackets we indicate if there is dominance of males or females.

	% **of Females**	z**-Value**	p**-Value**
Catholic Fac of Theology	34.4	0.022	0.982
Fac Agriculture	48.4	−1.95	0.05 (0.025) (M)
Fac. Architecture	31.3	0.872	0.383
Fac Chem Eng & Tech	56.3	0.255	0.799
Fac Civil Eng	44.2	0.344	0.730
Fac Econ & Business	51.6	0.128	0.896
Fac Education & Rehabilitation Sci	87.3	1.069	0.285
Fac Electri Eng & Comp	16.2	0.359	0.720
Fac Food Tech & Biotech	69.4	0.202	0.840
Fac Forestry	18.5	0.065	0.948
Fac Geodesy	14.3	2.268	0.023 (0.11) (M)
Fac Geotechnical Engineering	26.1	1.330	0.182
Fac Graphic Arts	63.3	0.367	0.714
Fac Humanities & Social Sciences	51.9	0.297	0.766
Fac Kinesiology	44.3	0.951	0.341
Fac Law	54.1	−2.259	0.023 (0.12) (F)
Fac Mech Eng & Naval Arch	14.2	0.567	0.571
Fac Metallurgy	47.8	−1.848	0.064
Mining, Geology & Petroleum Eng	36.3	1.874	0.061
Fac Organization & Informatics	41.1	0.226	0.821
Fac of Pharmacy & Biochemistry	81.3	0.791	0.429
Fac Political Science	38.7	1.376	0.169
Fac Science—Physics	9.1		
Fac Science—Biology	65.4	0.987	0.322
Fac Science—Chemistry	56.1	0.631	0.529
Fac Science—Geography	38.5	0.847	0.397
Fac Science—Geology	54.5	2.078	0.038 (0.19) (M)
Fac Science—Geophysics	50	1.123	0.317
Fac Science—Mathematics	29.6	0.786	0.432
Fac Teacher Education	63.5	2.108	0.035 (0.17) (M)
Fac Textile Technology	60.9	1.150	0.250
Fac Transport & Traffic Sciences	25.4	2.321	0.020 (0.01) (M)
Fac Veterinary Medicine	43.5	0.822	0.411
Fac Dental Medicine	53.6	0.910	0.362
Fac Medicine	46	−2.711	0.006 (0.003) (M)

Specifically, in the US, male scientists are the majority in all STEM fields, in Croatia at the largest university we find that females scientists outnumber their male colleagues in biology and chemistry and this gender equality seems to be a general trend in former socialist countries. For example, Huang, Gates, Sinatra, and Barabasi [31] reports recently a finding that the the proportion of female scientists worldwide can be as low as 28% in Germany and reaches almost 50% in Russia. In support that socialism was helpful in reaching the gender equality in academia, note that according to Shen, in Lithuania, female PhDs recipients in science make 63% of all recipients [7]. Similarly, as on the West, female scientists at the Croatian the largest university has a much larger proportion in social than natural sciences, whereas in some disciplines in social sciences females constitute even the majority in contrast to the US where females' fraction in social sciences are larger than in STEM fields,

but males still substantially dominate. For example, while in the US, females reach 33% in psychology at University of Zagreb female psychologists constitute 71%. Similarly, an interesting result is that at the Faculty of economics females constitute the majority, i.e., 52% (see Table 1).

3.2. Empirical Evidence: Males vs. Females at Faculty Level

For each faculty, we apply the Mann–Whitney U test and for most of them, we accept the null hypothesis, thus confirming that at a 5% confidence level there is no gender difference in citations. However, the gender difference in research productivity measured through citations per capita is confirmed for the following faculties (see Table 1): Faculty of Medicine, Faculty of Transport and Traffic Sciences, Faculty of Geodesy, Faculty of Teacher Education, and Faculty of Science-Geology, and Faculty of Law. With a single exception of Faculty of Law, male researchers on these faculties are significantly better than the female. In agreement with Duch et al. our results also reveal that the gender differences in citations are discipline-specific [26].

3.3. Empirical Evidence: Males vs. Females at University Level at Different Professor's Rank

Although women have made significant progress in catching up males concerning the gender productivity gap, on average, female scientists around the world continue to face discrimination. However, a situation on gender equality today is supposed to be much better than a few decades ago. Therefore, next, we test the hypothesis that the gender productivity gap in Croatian academy is diminishing over time. Gender inequality can be nicely captured analyzing the productivity and impact differences quantified by citations between the genders. As a good dynamic proxy for females' catching up males in the following analyses, we make gender comparison separately for the assistant, associate, and full professors. In Table 2 for the top 500 researchers we find that while on average, male full professors publish 23 papers during their active career, female full professors publish almost the same number of publications, 20, resulting in a small gender gap in total productivity. At a rank of associate professor, we find no difference in total productivity between the genders, but as an interesting result, we find that females have on average more citations than the male colleagues (see Table 3), an outcome to our knowledge not reported for any other country. Again, a difference between the genders appears between the male and female assistant professors, but this is most likely since raising a family is still more female than male responsibility.

Table 2. Men vs. Women—Publications comparison. Both STEM and non-STEM.

	Average No. Pubs (M)	Average No. Pubs (F)
Assistant Prof	12	8.9
Associate Prof	15	14.7
Full Prof	23.4	19.8

Table 3. Men vs. Women—Citation comparison. Both STEM and non-STEM.

	Average Citations (M)	Average Citations (F)
Assistant Prof	97.7	64.9
Associate Prof	127	144.7
Full Prof	258.2	207

3.4. Empirical Evidence: STEM Males vs. Females at University Level at Different Professor's Rank

Hence, as humans' societies are developing the difference in science performance between males and females is diminishing. However, it is not clear whether females are approaching males in their scientific results equally in STEM and non-STEM fields. Holman, Stuart-Fox and Hauser estimate that the gender gap appears likely to persist for generations, particularly in surgery, computer science, physics, and maths [13] .

Making comparison first for STEM fields, we obtain that male colleague exhibit significantly better results than their female colleagues only for a rank of full professors, while the gender difference is fading away for the associate and assistant professors. Please note that in the Mann–Whitney test, the larger the difference between the two populations, the larger the difference between the U values. To test whether STEM female and male researchers comprise two different subgroups, we apply the Mann–Whitney U test that quantifies the difference between two populations based on the difference between the ranks of their citations. Here the null hypothesis is that the distributions of the two subgroups are the same. We find that for full professors the test statistics produce z score $= -3.79$ (Table 4), rejecting the null hypothesis and confirming that at a 5% confidence level female STEM full professors are significantly more cited than their male colleagues. However, the gender gap we do not find for the other two ranks of a professorship, implying that STEM fields presently are more popular among females than a few decades ago.

Table 4. STEM: Men vs. Women—citation comparison.

	No. (M)	No. (F)	U (M)	U (F)	z Score	p-Value
Assistant Prof	331	248	40,284	41,804	0.38	0.704
Associate Prof	214	174	19,980	17,256	−1.24	0.215
Full Prof	511	266	79,219	56,706	−3.79	0.00

3.5. Empirical Evidence: Non-STEM Males vs. Females at University Level at Different Professor's Rank

The similar comparison between males and females we perform in non-STEM fields. We find no gender differences at the levels of full and associate professorship. However, we obtain that for the youngest level of the professorship, males exhibit significantly better results than their female colleagues. From Table 5 we find that for assistant professors the z score $= 2.54$, rejecting the null hypothesis that the distributions of the two subgroups for males and female are the same and confirming that at a 5% confidence level non-STEM male assistant professors are significantly more cited than their female colleagues. The result is not surprising because female assistant professors are at the age when as young women they spend a considerable amount of time with their kids. This result is in agreement with the conclusions by Stack, who found that women with preschool-aged children publish less than others [18].

Table 5. Non-STEM: Men vs. Women—citation comparison.

	No. (M)	No. (F)	U (M)	U (F)	z Score	p-Value
Assistant Prof	256	365	41,286	52,156	2.54	0.01
Associate Prof	201	169	16,739	17,230	0.24	0.808
Full Prof	267	252	33,256	34,027	0.22	0.819

3.6. Empirical Evidence: Pareto Inequality, STEM vs. Non-STEM

Since the work of Pareto on the application of power law to income inequality, we know that the tail part of an income distribution follows a power law [32]. Since that discovery, a huge literature has been published to demonstrating how distributions in economics and generally social science follow power-law tails and contribute to inequality [33–37]. As Stiglitz realized [36], the small fraction of superstar firms that dominate entire sectors of the economy (e.g., Amazon, Apple, Google, Microsoft) are the main drivers of wealth disparity. Indeed, it seems they are outliers, driven by STEM. Here we can hypothesize that STEM fields generate the fatter tails than non-STEM fields even in academia, when citations are considered. Motivated by Redner's finding that citations follow power law [38], here we apply the power-law formalism to test the research inequality between male and female scientists at the University of Zagreb, with particular focus on the difference between STEM and non-STEM fields.

Figure 1 shows the log-log Zip plots of citations versus rank [39,40], which is an alternative representation of the Pareto distribution [39,40]. When a probability distribution is asymptotically represented by a Pareto (power-law) distribution with exponent $1 + \alpha$, then a Zipf plot of size s versus rank R asymptotically follows a power law with exponent ξ relating the Pareto exponent α as [41]

$$\xi = 1/\alpha. \tag{6}$$

Generally, the smaller the α value, the fatter the tail of the power-law distribution.

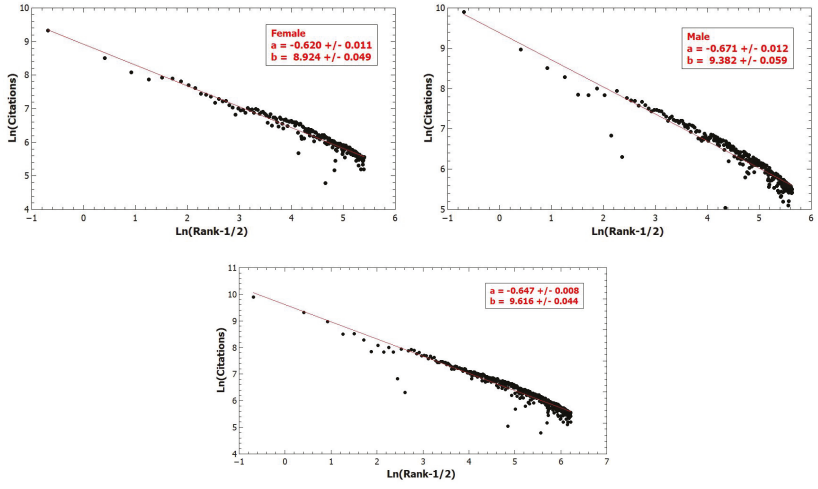


Figure 1. No substantial difference in research output measured by citations between males and females. However, the Zipf plot of citations for males exhibit a slightly fatter tail than the females'.

Here we apply the Gabaix-Ibragimov $R - 1/2$ method of fitting Zipf plots [40]. For year 2020, we perform a Zipf plot to both female and male researchers, no matter whether they are in STEM and non-STEM fields, and find the Zipf exponent for females $\xi = 0.62 \pm 0.01$ corresponding to $\alpha = 1.61$, whereas for males $\xi = 0.67 \pm 0.01$ corresponding to $\alpha = 1.49$. Here we note that both α values are within a range $\alpha \in (0, 2)$ characteristic for Levy distributions [41], which are characterized by infinite variance for which is known that the famous *Central Limit Theorem* does not hold.

Next we extend our analysis on research output, measured by citations, to separate STEM and non-STEM fields. Figure 2 shows a small difference between the slopes of the Zipf plot of STEM citations and non-STEM citations, $\xi_S = 0.66 > \xi_{nS} = 0.61$, thus STEM researchers generate fatter tails than non-STEM researchers (see Equation (6)).

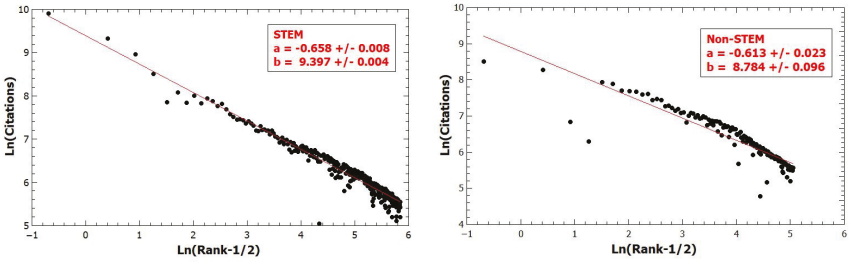


Figure 2. STEM researchers contribute to a slightly fatter tails thansdcc non-STEM researchers.

Statistics of citations for all male and female researchers we show in Table 6, while in Tables 7 and 8 we show statistics for STEM and non-STEM, respectively. In STEM, female associate professors have larger mean and kurtosis than their male colleagues. Kurtosis is generally larger for STEM than for non-STEM fields.

Table 6. Men vs. Women—Statistics. Both STEM and non-STEM.

	Mean	Median	Skewness	Kurtosis
All	161.6	31	21.5	687.8
All male	180.5	35	19.4	508.9
All female	138.2	26	6.7	117.9
Assistant Prof (M)	97.7	20	21.4	494.9
Assistant Prof (F)	64.9	10	7.6	89.5
Associate Prof (M)	126.9	33	4.3	24.2
Associate Prof (F)	144.7	35	9.3	11.9
Full Prof (M)	258.2	59	16.3	341.9
Full Prof (F)	207.0	61.5	4.04	124.0

Table 7. Men vs. Women—Statistics. STEM.

	Mean	Median	Skewness	Kurtosis
All male	208.7	51	17.9	395.9
All female	197.9	68.5	6.3	160.5
Assistant Prof (M)	122.7	30	16.8	297.6
Assistant Prof (F)	94.9	30	7.5	78.1
Associate Prof (M)	141.7	44.5	3.7	116.9
Associate Prof (F)	181.8	70	9.2	101.7
Full Prof (M)	281.4	73	15.4	280.6
Full Prof (F)	278.7	142	3.5	120.8

The Pareto principle has been found in large number of datasets. This principle, accomplished for citations, states that roughly 80% of consequences come from 20% of the causes [32,42]. However this principle does not hold for any power law, but for $\alpha = 1.16$ for which 80% of effects come from 20% of causes.

From Table 9 for male STEM researchers we find that 20% of males bring 80% of citations in well agreement with Pareto principle. However, 27.2% of STEM females bring 80% citations. In case of non-STEM researchers we find substantial difference from the Pareto principle.

Shannon's entropy of information is a foundational concept in information theory quantifying the amount of information embedded in the variable or the amount of storage expressed by the number of bits required to store the variable [43]. The larger the entropy, the larger the amount of surprise comprised in the data. When all values of discrete variable (R in total) are equally probable, each probability equals $1/R$ and then it is easy to show that the Shannon entropy equals $ln(R)$. If the probability of a particular variable value approaches 1, the Shannon entropy tends to zero.

Here we apply Shannon's entropy for a randomly chosen citation where it can be assigned to either males or females. In Table 10 the last column reports the entropy values which are close to the entropy (equal to 1) which one would obtain for a regular coin with two equally likely outcomes for which there is no way to predict the outcome of the coin toss. The entropy equal to 0 would represent a case when one of the probabilities equals 1 and any foredooming outcome can be predicted perfectly. We demonstrate no gender difference except small differences at the level of full professors, both STEM and non-STEM, and at the level of STEM Assistant Professors.

Table 8. Men vs. Women—Statistics. non-STEM.

	Mean	Median	Skewness	Kurtosis
All male	138.6	9	6.8	62.8
All female	85.4	6	6.2	51.8
Assistant Prof (M)	65.4	6	4.3	23.9
Assistant Prof (F)	44.5	3	5.4	35.9
Associate Prof (M)	111.3	19	4.9	28.6
Associate Prof (F)	106.5	9	4.5	25.8
Full Prof (M)	213.7	11	5.2	37.0
Full Prof (F)	131.2	12.5	5.3	34.6

Table 9. Men vs. Women—Pareto principle. Both STEM and non-STEM.

STEM (M)	STEM (F)	Non-STEM (M)	Non-STEM (F)
20.5	27.2	14.13	15.2

Table 10. Men vs. Women—Shannon's Entropy. Both STEM and non-STEM.

	Probability (M)	Prob (F)	Shannon's Entropy
STEM Assist Prof	0.6331	0.3668	0.9482
STEM Associate Prof	0.4892	0.5107	0.9996
STEM Full Prof	0.6594	0.3401	0.9249
Non-STEM Assist Prof	0.5076	0.4923	0.9998
Non-STEM Associate Prof	0.5540	0.4459	0.9915
Non-STEM Full Prof	0.6330	0.3669	0.9482
Assistant Prof	0.5904	0.4095	0.9762
Associate Prof	0.5148	0.4851	0.9993
Full Prof	0.6520	0.3479	0.9322

4. Discussion and Conclusions

At the largest Croatian University, applying the Mann–Whitney test, we demonstrate no gender difference in number of citations except for seven faculties, where males are significantly better than females on six faculties. We report that female STEM full professors are significantly more cited than their male colleagues, while male non-STEM assistant professors are significantly more cited than their female colleagues. In Table 11 as a comparison between male and female researchers, we report the average citations across different University members. In contrast to the US and many Western nations, at the largest Croatian University there are several faculties where the most cited researcher is woman and these faculties are the following, both STEM and non-STEM: Faculty of Chemical Engineering and Technology, Faculty of Economics and Business, Faculty of Food Technology and Biotechnology, Faculty of Geotechnical Engineering, Faculty of Graphic Arts, Faculty of Kinesiology, Faculty of Organization and Informatics, Faculty of Political science, Faculty of Science-Biology, Faculty of Science-Geography, and Faculty of Teacher Education. Thus, at individual level, at some particular faculties and scientific fields, the most cited female is better off than the most cited male. Moreover, at collective level, there are ten faculties where females have the larger average citations than their male colleagues: Faculty of Chemical Engineering, Faculty of Economics, Faculty of Food Technology, Faculty of Forestry, Faculty of Geotechnical Engineering, Faculty of Graphical Arts, Faculty of Organization and Informatics, Faculty of Political science, Faculty of Science-Physics, and Faculty of Dental Medicine.

Table 11. Men vs. Women—average citation across different University constituents.

	Average Citations for Males	Females
Catholic Fac of Theology	18	1.0
Fac Agriculture	146.1	89.8
Fac. Architecture	4.5	3.3
Fac Chem Eng & Tech	411.7	528.9
Fac Civil Eng	46.2	31.4
Fac Econ & Business	13.6	19.2
Fac Education & Rehabilitation Sci	23.1	7.7
Fac Electri Eng & Comp	129.9	111.0
Fac Food Tech & Biotech	259.0	320.8
Fac Forestry	75.4	77.4
Fac Geodesy	23.9	3.6
Fac Geotechnical Engineering	16.8	82.5
Fac Graphic Arts	23.6	49.9
Fac Humanities & Social Sciences	24.4	11.7
Fac Kinesiology	186.6	160.6
Fac Law	18.5	11.8
Fac Mech Eng & Naval Arch	116.3	91.2
Fac Metallurgy	129.1	33.5
Mining, Geology & Petroleum Eng	89.5	44.4
Fac Organization & Informatics	14.2	21.4
Fac of Pharmacy & Biochemistry	640.3	387.8
Fac Political Science	6.2	13.3
Fac Science—Physics	1931	2365
Fac Science—Biology	372.8	361.6
Fac Science—Chemistry	488.2	390.4
Fac Science—Geography	22.3	21.4
Fac Science—Geology	206.1	83.3
Fac Science—Geophysics	611.7	215.7
Fac Science—Mathematics	167.4	87.6
Fac Teacher Education	16.4	7.6
Fac Textile Technology	254.7	74.9
Fac Transport & Traffic Sciences	17.5	4.4
Fac Veterinary Medicine	203.0	149.2
Fac Dental Medicine	140.2	150.7
Fac Medicine	442.3	285.5

Wealthy countries such as Japan, Germany, and Switzerland has fewer women authors than poorer ones [13]. Besides this asymmetry, there is a significant difference in gender productivity between the US and the poorer countries. However, would it not be reasonable to assume that the more developed a country, the smaller the gender differences in their performance? Partially the gender difference in the US academia appears because in the US, in contrast to the EU, the large majority of universities are private and the private educational system recognizes only quality and does not care much about gender or even racial equality. Comparing the US and developing countries, due to globalization where the US is the attractor of best students all over the world, the level of competition in the US academia is much higher than in any other country. Therefore, females to compete with males in the US academia must compete not only with the best US males, but with the most brilliant world candidates. Since raising a family is even in the US still considered to be more female than male responsibility, being a female makes their position much harder. Clearly, if a female is a genius like Nobel Prize Laureate Marie Curie, she can achieve excellent academic results in parallel with being an excellent mother and even raise her daughter to become another Nobel Prize Laureate. However, there are not so many geniuses of that kind. In large educational systems such as the US,

fluctuations are substantially smaller than in small countries such as Croatia, and so events much different than expectations are more likely.

Briefly, we expect that in a limit where children are equally educated regardless of gender and where family obligations are equally distributed among both parents, it is meaningful that men and women should perform equal results, not only in science. The closer a society to this limit, the larger the similarity in their performance. It is a challenge to test this obvious hypothesis. To this end, in academia, the gender research productivity gap will not diminish without substantial reforms in education, mentoring, and academic publishing.

For a country to be competitive in both science and business, in agreement with Cobb-Douglas production function extended for human capital, maximization of its intellectual capital particularly its females' part is a top country's priority. In the short run, policymakers should launch new programs required to stimulate international collaboration for female researchers, because international collaboration is one of the pillars for excellence in science. We suggest that the best female researchers should generally receive more grants and funding than their best male colleagues as compensation for spending more time in raising a family. The pay gap in science should become an illegal practice. In the grant review process, gender equality among reviewers should be always respected, and if not, the call should be canceled. We even suggest that kindergartens should be open at least within large Universities and institutes. In long run, we need to rapidly boost the interest in science and particularly STEM disciplines among the girls.

Author Contributions: Conceptualization, B.P.; Data curation, D.W. and M.J.; Formal analysis, D.W. and M.J.; Funding acquisition, B.P.; Investigation, B.P., M.J. and D.W.; Resources, M.J.; Software, D.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received external funding. B.P. acknowledged support of this work by the Croatian Research Agency (KK.01.1.1.01.0009, DATACROSS), ZSEM, and the University of Rijeka. M.J. received funding from ZSEM. D.W. received funding from ZSEM and LSB.

Acknowledgments: B.P. acknowledge the gracious support of this work by ZSEM and the University of Rijeka.

Conflicts of Interest: The authors declare no conflict of interest.

Reference

1. Cole, J.R.; Zuckerman, H. The productivity puzzle: Persistence and changes in patterns of publication of men and women scientists. *Adv. Motiv. Achiev.* **1984**, *2*, 217–258.
2. Lariviere, V.; Ni, C.; Gingras, Y.; Cronin, B.; Sugimoto, C.R. Global gender disparities in science. *Nature* **2013**, *504*, 211–213. [CrossRef] [PubMed]
3. Fox, M.F. Gender, family characteristics, and publication productivity among scientists. *Soc. Stud. Sci.* **2005**, *35*, 131–150. [CrossRef]
4. Carr, P.L. Relation of family responsibilities and gender to the productivity and career satisfaction of medical faculty. *Ann. Intern. Med.* **1998**, *129*, 532–538. [CrossRef] [PubMed]
5. Cameron, E.Z.; White, A.M.; Gray, M.E. Solving the productivity and impact puzzle: Do men outperform women, or are metrics biased? *BioScience* **2016**, *66*, 245–252. [CrossRef]
6. Organisation for Economic Co-Operation and Development. *Education at a Glance 2012*; OECD: Paris, France, 2012.
7. Shen, H. Inequality quantified: Mind the gender gap. *Nat. News* **2013**, *495*, 22–24. [CrossRef]
8. Moss-Racusin, C.A.; Dovidio, J.F.; Brescoll, V.L.; Graham, M.J.; Handelsman, J. Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16474–16479. [CrossRef]
9. Holden, C. General Contentment Masks Gender Gap in First AAAS Salary and Job Survey. *Science* **2001**, *294*, 396–411. [CrossRef]
10. Ding, W.W.; Murray, F.; Stuart, T.E. Gender Differences in Patenting in the Academic Life Sciences. *Science* **2006**, *313*, 665–667. [CrossRef]
11. Lincoln, A.E.; Pincus, S.; Koster, J.B.; Leboy, P.S. The Matilda effect in science: Awards and prizes in the US, 1990s and 2000s. *Soc. Stud. Sci.* **2012**, *42*, 307–320. [CrossRef]

12. Ley, T.J.; Hamilton, B.H. The gender gap in NIH grant applications. *Science* **2008**, *322*, 1472–1474. [CrossRef] [PubMed]
13. Holman, L.; Stuart-Fox, D.; Hauser, C.E. The gender gap in science: How long until women are equally represented? *PLoS Biol.* **2018**, *16*, e2004956. [CrossRef] [PubMed]
14. Lariviere, V.; Vignola-Gagne, E.; Villeneuve, C.; Gelinas, P.; Gingras, Y. Sex differences in research funding, productivity and impact: An analysis of Quebec university professors. *Scientometrics* **2011**, *87*, 483–498. [CrossRef]
15. Xie, Y.; Shauman, K.A. Sex differences in research productivity: New evidence about an old puzzle. *Am. Socio. Rev.* **1998**, *63*, 847–870. [CrossRef]
16. Lemoine, W. Productivity patterns of men and women scientists in Venezuela. *Scientometrics* **1992**, *24*, 281–295. [CrossRef]
17. Pripic, K. Gender and productivity differentials in science. *Scientometrics* **2002**, *55*, 27–58. [CrossRef]
18. Stack, S. Gender, Children and research Productivity. *Res. High. Educ.* **2004**, *45*, 891–920. [CrossRef]
19. Mason, M.A.; Goulden, M. Do Babies Matter? The Effect of Family Formation on the Lifelong Career of Academic Men and Women. *Academe* **2002**, *88*, 21. [CrossRef]
20. National Science Foundation. *2010 Survey of Doctorate Recepients*; NSF: Arlington, TX, USA, 2010.
21. Goulden, M.; Frasch, K.; Mason, M.A. *Staying Competitive*; Center for American Progress: Washington, DC, USA, 2009.
22. Chawla, D.S. Self-citation rates higher for men. *Nature* **2016**, *535*, 212. [PubMed]
23. Long, J.S. Measures of sex differences in scientific productivity. *Soc. Forces* **1992**, *71*, 159–178. [CrossRef]
24. Goulden, M.; Mason, M.A.; Frasch, K. Keeping women in the science pipeline. *Ann. Am. Acad. Polit. Soc. Sci.* **2011**, *638*, 141–162 [CrossRef]
25. Symonds, M.R.E.; Gemmell, N.J.; Braisher, T.L.; Gorringe, K.L.; Elgar, M.A. Gender differences in publication output: Towards an unbiased metric of research performance. *PLoS ONE* **2006**, *1*, e127. [CrossRef] [PubMed]
26. Duch, J.; Zeng, X.H.T.; Sales-Pardo, M.; Radicchi, F.; Otis, S.; Woodruff, T.K.; Nunes, A.L.A. The possible role of resource requirements and academic career-choice risk in gender differences in publication rate and impact. *PLoS ONE* **2012**, *7*, e51332. [CrossRef] [PubMed]
27. Abramo, G.; D'Angelo, C.A.; Caprasecca, A. Gender differences in research productivity: A bibliometric analysis of the Italian academic system. *Scientometrics* **2009**, *79*, 517–539. [CrossRef]
28. World Economic Forum. *The Human Capital Report 2016*; World Economic Forum: Colony, Switzerland, 2016.
29. Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future (2007), Chapter: Front Matter. Available online: https://www.nap.edu/read/11463/chapter/1 (accessed on 22 October 2020).
30. National Academy of Sciences; National Academy of Engineering and Institute of Medicine. *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*; The National Academies Press: Washington, DC, USA, 2007.
31. Huang, J.; Gatesa, A.J.; Sinatrad, R.; Barabasi, A.L. Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proc. Natl. Acad. Sci. USA* **2019**, *117*, 4609–4616. [CrossRef] [PubMed]
32. Pareto, V. *Cours D'conomie Politique*; Librairie Droz: Geneva, Switzerland, 1896.
33. Piketty, T.; Saez, E. Income Inequality in The United States, 1913–1998. *Q. J. Econ.* **2003**, *118*, 1–41. [CrossRef]
34. Acemoglu, D.; Robinson, J.A. The Rise and Decline of General Laws of Capitalism. *J. Econ. Perspect.* **2015**, *29*, 3–28. [CrossRef]
35. Gabaix, X.; Lasry, J.M.l.; Lions, P.L.; Moll, B. The Dynamics of Inequality. *Econometrica* **2016**, *84*, 2071–2111. [CrossRef]
36. Stieglitz, J.E. *People, Power, and Profits*; W. W. Norton & Company: New York, NY, USA, 2019.
37. Dragulescu, A.; Yakovenko, V.M. Exponential and power-law probability distributions of wealth and income in the UK and the USA. *Phys. A* **2001**, *299*, 213–221. [CrossRef]
38. Redner, S. How popular is your paper? An empirical study of the citation distribution. *Eur. Phys. J. B* **1998**, *4*, 131–134. [CrossRef]
39. Zipf, G.K. *Human Behavior and the Principle of Least Effort*; Addison-Wesley Press: Cambridge, MA, USA, 1949.
40. Gabaix, X.; Ibragimov, R. Rank—1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents. *J. Bus. Econ. Stat.* **2011**, *29*, 24–39. [CrossRef]

41. Podobnik, B.; Valentincic, A.; Horvatic, D.; Stanley, H.E. Asymmetric Levy Flight in Financial Ratios. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17883–17888. [CrossRef] [PubMed]
42. Pareto, V. *Manual of Political Economy*, Variorum ed.; Montesano, A., Zanni, A., Bruni, L., Chipman, J.S., McLure, M., Eds.; Oxford University Press: Oxford, UK, 2014.
43. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Entropy Editorial Office
E-mail: entropy@mdpi.com
www.mdpi.com/journal/entropy

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-0365-0795-8