economies

Recent Topics in Economic Research

Feature Papers for Cerebrating the 10th Anniversary of Economies

Edited by

Ralf Fendel, Robert Czudaj and Sajid Anwar

Printed Edition of the Special Issue Published in *Economies*

www.mdpi.com/journal/economies

MDPI

Recent Topics in Economic Research – Feature Papers for Cerebrating the 10th Anniversary of Economies

Recent Topics in Economic Research – Feature Papers for Cerebrating the 10th Anniversary of Economies

Editors

Ralf Fendel
Robert Czudaj
Sajid Anwar

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

Editors
Ralf Fendel
WHU - Otto Beisheim School of Management
Germany

Robert Czudaj
Technical University of Freiberg
Germany

Sajid Anwar
University of the Sunshine Coast
Australia

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 *Economies* (ISSN 2227-7099) (available at: https://www.mdpi.com/journal/economies/special_issues/10th_Anniversary_Economies).

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 978-3-0365-6892-8 (Hbk)
ISBN 978-3-0365-6893-5 (PDF)

Contents

About the Editors

Ralf Fendel

Ralf Fendel is Professor of Economics at WHU – Otto Beisheim School of Management since 2008. Previously, he held positions at the Deutsche Bundesbank, the European Business School, the University of Erfurt, and the University of Konstanz. His research interest include monetary economics and policy, macroeconomics, international finance as well European integration.

Robert Czudaj

Robert Czudaj is Full Professor for Economics, in particular (Monetary) Macroeconomics at Technische Universität Bergakademie Freiberg. Previously, he was affiliated at the Ludwig Maximilian University of Munich, at Chemnitz University of Technology, and at the University of Duisburg-Essen. His research interests include Applied Econometrics, Monetary Economics, Macroeconomics, Time Series Analysis, Financial Econometrics, Financial Markets, and Exchange Rates.

Sajid Anwar

Sajid Anwar is professor at University of the Sunshine Coast (Queensland, Australia) since January 2009. Prior to this, he has held academic positions at University of South Australia (Adelaide, Australia), University of Adelaide (Adelaide, Australia), James Cook University (Townsville and Cairns, Australia), Charles Sturt University (Bathurst, Australia), University of Canberra (ACT, Australia), University of Macau (Taipa, Macau) and Northern Territory University (Darwin, Australia). His areas of research and teaching include international economics, corporate finance and international business.

economies

Article

Efficiencies of Faith and Secular Microfinance Institutions in Regions of Asia, Africa, and Latin America: A Two-Stage Dual Efficiency Bootstrap DEA Approach

Adithya Kiran Kolloju [1,*] and Michele Meoli [2]

[1] Department of Management, Information and Production Engineering, University of Bergamo, 24100 Bergamo, Italy
[2] Department of Management, Information and Production Engineering, CCSE-Cisalpino Institute for Comparative Studies in Europe, University of Bergamo, 24100 Bergamo, Italy; michele.meoli@unibg.it
* Correspondence: adithya.kolloju@unibg.it; Tel.: +39-339-240-6828

Abstract: Purpose: the objective is to measure the financial and social performance of 127 microfinance institutions (MFIs) and observe the effects with explanatory factors such as "type", "geography region", and "secular and faith" variables. Design/methodology/approach: The time-series performance analysis of microfinance institutions is determined in two stages. In the first stage, both the social and financial efficiencies are measured with Data Envelopment Analysis (DEA) approach. The two explanatory factors along with faith and secular variables show the effect on these determined efficiencies by the second stage of the Tobit regression Random effect Model. Findings: Financial performance is greater than the social performance from the first stage analysis. When considering the explanatory variables, the social performances are not significant with religious factors. When the regression is performed in a group, the financial score is more significant with religious and other explanatory variables. Faith-based and secular-based microfinance institutions are strongly significant if the performances (efficiencies) are highly maintained. Originality/Value: faith and secular variables are identified based on the background/history information of each microfinance institution (MFI).

Keywords: DEA-approach; non-parametric; Tobit; random-effect; religion; conventional; time-series; efficiency performance; microfinance; region

Citation: Kolloju, Adithya Kiran, and Michele Meoli. 2022. Efficiencies of Faith and Secular Microfinance Institutions in Regions of Asia, Africa, and Latin America: A Two-Stage Dual Efficiency Bootstrap DEA Approach. *Economies* 10: 66. https://doi.org/10.3390/economies10030066

Academic Editor: Andreia Dionisio

Received: 30 December 2021
Accepted: 8 March 2022
Published: 16 March 2022

1. Introduction

Financial services for an individual ameliorate inclusion and performing business activities for obtaining a smooth economy (Li et al. 2019). Organizational performance promotes good outcomes irrespective of size, type, or faith, but is achieved by good loan repayment rates. The self-employed destitute have strong capabilities and commitments to structure their living and the microfinance institutions (MFIs) extend the frontiers with small credits and other financial services such as insurance, policies, mutual funds, etc. (Alimukhamedova 2013), while the access to traditional banks for the same destitute has been difficult for financial advancement activities. MFI finances with low-security risks and has a critical role in the progress of the economic system in emerging countries by showing their significant financial services (such as group lending, self-monitoring, short-repayment methods, etc.) to millions of poor households globally (Daley-Harris 2009; Hadi and Kamaluddin 2015) in addition with the digitalized services (Agrawal and Sen 2017). It is observed that MFIs emerged from traditional banks to calculate and function their administration based on the profit efficiency, operating costs, and returning capital constraints which leads to economical profit being more important than social outreach (Kar 2012). Although one or group of efficiency (financial, social, and technological) is important in the functioning of a financial organizations to sustain bankruptcy or shutting down. At a certain point, the financial, social, and technological efficiencies are interconnected. Few dimensions of MFIs (NBFC[1], credit unions, micro-credits, rural banks, NGOs[2]) were taken into consideration for the influence of their services and functionalities. Few operate with the spirit of capitalism (private ownership) but always abide by the reserve bank guidelines.

The capital structuring for an MFI (faith or secular) is obtained from sources such as government subsidiaries, public-private investments (shareholders), debt capital, profits, and third-party donations (Tchuigoua 2015). Apart from the public and the commercial funders, both provide funds as a charity for certain types of microfinance organizations for global development (Cobb et al. 2016). The level of profits, revenues, ROA[3], and ROE[4] ratios from operating costs, loan size, cost per borrower, etc., describes the performance of an MFIs. For better sustainability of MFI, it is important to maintain the average returns from the clients on timely basis agreements. It is also important to merge the economic and social gaps for proper and smooth functionality (Liñares-Zegarra and Wilson 2018). The profit-based institutions extend links with capital markets to ensure maintaining profitability and high-rate efficiency. The traditional MFIs are profit-based because the drawn-out monetary usefulness is a higher priority than the social effort (Leite et al. 2019; Liñares-Zegarra and Wilson 2018; Schwarz et al. 2015).

We drew the financial data from the World Bank catalog with the Mix-market collaboration and the data were extracted from 25 countries with 127 microfinance institutions and similar data were used in several studies (Aggarwal et al. 2015; Leite et al. 2019; Navajas and Tejerina 2006; Sainz-Fernandez et al. 2018; Wijesiri et al. 2015, 2017). The dual efficiency is a technique to understand the organization's performances and may estimate by diverse variables while we considered financial and social variables that reflect the performance factors of each microfinance institution. The data envelopment analysis (DEA) (Charnes et al. 1978), is a non-parametric approach bootstrap method that analyses the sensitivity of measured efficiency score by decision-making units (DMUs) (Simar and Wilson 2007). The DMUs have non-specified measurable properties and are neither classified in categories nor groups of faith and secular but the assumption is made for all MFIs in the analysis.

It is necessary to explore the marginal changes and impact of environmental indicators on the dual efficiencies. Several studies (Lebovic 2004; Wijesiri et al. 2015) used a censored model to investigate the determinants of efficiency estimates. The Tobit regression truncated random effect method is preferred as the correlation exists with an error term for both input and output variables (Amemiya 1984), the censored regression is used to analyze the impact of indicators such as type, religion, and region on efficiency performances. It is believed that an indirect relationship exists between the social/economical factors with the organizational principles (type, region, and religion). The religion and region indicators are organized by faith, secular institutions by the life cycle of MFIs, and they were identified and represented by three regions (Asia, Africa, and Latin America). The straightforward time-series Tobit censored linear regression model is assumed with a minimum of two autonomous variables/indicators individually for dual efficiencies. The second stage regression exhibits the indicators over efficiencies, and it is significantly seen that faith-adopted MFIs show better social outreach than traditional institutions.

The study is organized in sections and the following literature (Section 2) describes the important classification that is concentrated to determine the efficiencies, the effect of efficiencies over the observatory variable on women in microfinance, religious performance, effectiveness. Section 3 has empirical literature followed with the data and Section 4 has the methodology, Section 5 has the results of the performances and impaction with the regression model. Section 6 has the discussion and conclusions.

2. Literature Review

Microfinance institutions serve the development of society assisting the poor with turning out to be little business visionaries by micro-credits/grants without accepting interest rates (faith-operated) or little interest rates (traditionally operated). The clustered faith-based institutions are differentiated with the functionalities. To strike the difference between outreach and poverty alleviation, the institutional performance varies on the technical operations, attractiveness, strategies, methodologies, etc., and these categories have importance depending on the secular-based microfinance institutions (SB-MFIs) or the faith-based microfinance institutions (FB-MFIs). The cross-country MFIs are some laid out privately but are regulated under government policies with suitable regulatory acts, credit information companies, Reserve/Central bank, and Microfinance Regulatory Authority. Loan portfolios, recovering activities, are managed by financial managers and legal teams, apart from tradition/religion, MFIs raise funds with savings mobilizations (Wijesiri et al. 2015), donations (with the risk of liquidity management, this is considered as the minimal investment), deposits, and other cumulative surpluses. Partially, few MFIs are controlled and managed by women with profiting terms and policies, this is considered as a business activity or entrepreneurial establishment. We further discuss the role of gender (women) in microfinance.

2.1. Microfinance and Women

Business with opportunity-driven strategies in low-income countries satisfies basic needs and enhance the welfare of the poor (Schwarz et al. 2015). Gender orientation contrasts in microfinance are particularly critical because women have a crucial impact on money-related progressions, especially in the turn of advancements. Women are granted lower importance than men because of their educational levels, discrimination, etc. Women in microfinance promote and strengthen the International Labor Organizations (ILOs) which has a decent work agenda. MFIs are established in urban/rural areas, where the women borrowing rate is increasing in annual periods (from the cross-country data). In the rural areas, the increase in small enterprises by women allows to mitigate patriarchal practices and make women empowerment grow in significant proportions. Trust is the main concern that the financial institution believes, and MFIs assist with entrepreneurial training skills to women in some regions and provide confidence in the establishment of small-scale industries in rural areas with low-interest rates (Hadi and Kamaluddin 2015). There is a 54.6% increase in active women borrowers since 2009, which increased the financial inclusion in microfinance. The combination of microfinance and social assert, brings norms, trust, and networks to connect people and coordinate with their actions. The average borrowing rate of female borrowers and only 16 countries have at least 50% and 9 countries have 70% access in rural areas, the average rate is much increased from the year 2008. MFIs grant capital with low-interest rates for entrepreneurial enterprises, especially female borrowers are better with loan repayment, and analytically the financial outreach is significantly high (Aggarwal et al. 2015; Belwal et al. 2012; Ul-Hameed et al. 2018). Indeed, we chose "women borrower" as a significant indicator for determining the social efficiency for regions in Asia, Africa, Latin America, and the Caribbean, where women's participation plays a decisive role in the analysis. From the religious point of view women are playing a more valuable role than in traditional institutions.

2.2. Religion Performance and Effectiveness

Religion rehearses impact and changes the human mind despite literacy (Barro and Mccleary 2003). Non-profit and regional MFIs primarily focus on developments in poverty alleviation. On the other hand, faith MFIs follow the community and religiously implemented principles and the results are more explored from the conviction-based institutions. In Islamic law, there is a prohibition to apply interest rates on loans (Imam and Kpodar 2016) and there are more potential in attaining social performance with counterparts with conventional organizations, and fundings are made with profits (Fersi and Boujelbéne 2016). Catholic microfinance institutions are better with financial performance in developing efforts (Mersland et al. 2013). It is empirically analyzed that women are more concerned with religious practices and ethics (Gyapong et al. 2021). The religious institutions are risk aversion and have low stock returns. In an Islamic country such as Bangladesh, social efficiency is 22% more than financial efficiency (Khan and Sulaiman 2015; Mia and Govindaraju 2016).

We took the religion factor (irrespective of belief) as a dependent variable to differentiate the efficiencies socially and financially with profit/non-profit microfinance organizations.

The faith-based institutions follow the principles imposed by their religion and run administration to reach social and financial objectives. In countries such as Pakistan, Afghanistan, Bangladesh, and Middle Eastern countries, religion (Islam) is very important and treated in almost every activity. Mansori et al. (2020) collected questionaries from each Muslim and the gender classification (majority men have favor towards religion) showed importance. Spotted the difference in the risks between FB-MFIs and conventional Islamic MFIs cross-sectionally from the period 1998–2014 from the MIX database, the results show financial stability with SB-MFIs with low risk, evidence from OSS regression. The DEA-CRS[5] analysis on 72 MFIs (NGOs and NBFIs) in MENA regions where MFI performances in Islamic nations are significant socially and financially, and these efficiencies are concerning to age (Mature, Young, and New) (Ben Abdelkader and Mansouri 2019). It is feasible to identify institutions and Table 1 shows the institution type with the religious status of each country. The background and historical analysis are perfectly performed to explore the religious status of the microfinance institution itself, it is observed many MFIs are integrated with Vision fund in Latin American regions (Catholic), and East Asian countries.

Table 1. Microfinance country, type, and religious status.

Country	Institution Type	Religion Status
Afghanistan	NGOs	Secular
Azerbaijan	Credit corporative, Bank, and NBFC	Secular, Faith
Bolivia	Bank, and Credit corporative	Secular, Faith
Brazil	Bank, and Credit corporative	Secular, Faith
Bangladesh	NGO, Bank	Secular, Faith
Costa Rica	NGO, Bank, and Credit corporative	Secular, Faith
Colombia	Microcredit	Secular
Cambodia	NGO, and Bank	Secular, Faith
El Salvador	NGO	Secular, Faith
Egypt	NGO	Secular, Faith
Ecuador	Bank, Credit corporative, Microcredit, and NGO	Secular
Guatemala	Bank, NGO, Credit, NBFI corporative, and Microcredit	Secular, Faith
Ghana	Bank	Secular
India	NGO, Credit corporative, NBFI, and Bank	Secular, Faith
Kosovo	NGO	Secular
Kazakhstan	Credit corporative	Secular
Jordon	Bank	Secular, Faith
Mexico	NBFI, and Bank	Secular, Faith
Morocco	Bank, and Credit corporative	Secular, Faith
Nicaragua	NBFI, and Credit corporative	Secular, Faith
Peru	Bank, Credit corporative, NGO, and NBFI	Secular, Faith
Pakistan	NBFI, and Bank	Secular, Faith
Palestine	Credit corporative	Faith
Tanzania	Bank	Secular
Tajikistan	Bank, and NGOs	Secular

Source: Mix Market, Background Historical status of MFIs.

3. Research Design

3.1. Research Questions

The research questions are developed in two parts:

1. To determine the social and financial efficiency of 129 cluster MFIs (faith and secular) with a frontier non-parametric methodology (Simar and Wilson 2007; Seiford and Zhu 2014; Wijesiri et al. 2015). This helps to identify the microfinance status in established countries;

2. To determine the effect of external variables on these efficiency variables through both cluster[6] and two-model[7] groups, the time series as the Tobit regression model (Wijesiri et al. 2015). This helps to explore which variable can affect the MFI performance.

3.2. Data Selection

The performance and the selection of variables were inspired by the contributions in the microfinance industry (Gutiérrez-Nieto et al. 2009; Gyapong et al. 2021; Mersland et al. 2013; Schwarz et al. 2015; Wijesiri et al. 2015; Wijesiri and Meoli 2015), etc. The sample of 127 MFIs was selected from the World Bank financial database (Mix market) with annual periods 2008–2018. The specific variable "Female borrowers" is inspiring with the increase in clients annually in many countries. Table 2 discusses the MFIs list and the percentage of borrowers, investments, and equity in each type of MFIs.

Table 2. Financial funding of different microfinance institutions.

	NBFI	NGO	Bank	Credit Union	Rural Bank
Borrowings	36	30	21	8	10
Deposits/investments	43	28	58	79	69
Equity	21	42	20	13	21

Source: Mix market 2019, World Bank's data catalog (%).

3.3. Variable Selection

Table 3 describes the explanation of dependent and independent variables for efficiency measurement with definitions. The selection of indicators is influenced by the previous works published in the articles where cross-country data have importance. The analysis is not based on specific country performance instead, it is cross-country productivity impaction.

Table 3. Selection and description of input/output variables for the first stage of analysis.

Variables	Indicators	Term Definitions	Data Explanation
Input variables	Loan Officers	Dealing with authorizing loans to the clients, monitoring interest and security, responsible for outgoing credits and outstanding loans.	The database consists of 127 well-performing microfinance institutions. A total of 30 MFIs are identified as faith-based organizations and 97 MFIs are secularly operating microfinance institutions. The 30 Faith institutions are related to Catholic- and Islamic-based MFIs. These are identified by the establishment and collaborations and merging with dominating microfinance institutions. All the data are drawn from Mix.org merged with World Bank Data in the year 2019.
	Total Assets	The absolute estimation of assets is constrained by the financial organization because of past occasions and from which future monetary advantages are relied upon to stream to the MFI.	
	Cost per borrower	It is the cost imposed on a borrower from loan application to the total loan repayment.	
Output variables	Financial Revenue (Financial efficiency)	Revenues generated from loans, assets, other financial accumulations.	
	Female borrowers (Social efficiency)	The percentage of women borrowers is consistently increasing periodically. The social outreach is well enough.	

The financial revenue is an output variable to determine the financial performance of MFIs. The financial performance tends to raise if there are more dependents and good repayment rates irrespective in rural or urban regions, for this purpose the variable borrowers are used to determine the social outreach performance of MFI. (*Microfinance-Barometer-2019_web-1*, n.d.) report specifies there is an incremental increase with women entrepreneurs annually for a decade so, women/female borrowers help to determine the social efficiency. The following Table 4 has numerical statistics of input and output variables along with the correlations between the variables are in Table 5.

Table 4. Mean and standard deviation of input/output variables.

Variable	Mean	Std. Dev.
Total Assets (USD MM *)	19.71	270.50
Personal (Count)	208	2190
Cost per Borrower (USD)	137	1065
Financial Revenue (USD MM *)	4.60	51.2
Female borrowers (Count)	15.709	681,944

* USD MM represents value Million USD.

Table 5. Correlations of input/output variables.

Correlations	Assets	Financial Revenue	Personal (s)	Female Borrower(s)	Cost per Borrower(s)
Assets	1.00				
Financial revenue	0.73	1.00			
Personal(s)	0.27	0.49	1.00		
Female borrower(s)	0.22	0.39	0.64	1.00	
Cost per borrower(s)	0.023	0.024	−0.02	0.64	1.00

3.4. Borrowing Rate

Women workers in developing countries all through the world contribute their development in financial turn and microfinance helps to empower them and helps with making their commitment (Somavia 2007). According to the data, women clients are up to 73% on average. The average female borrowing rate is classified in Table 6, from the countries which are considered in the analysis. Ghana and Pakistan which have faith-based MFIs are dominating in the second position and Azerbaijan, Tanzania takes the last position in the female borrowing rate.

Table 6. Percentage of female borrowers in 25 countries.

Country	Avg Female Borrower (%)
Afganisthan	52
Azerbaijan	29
Bolivia	40
Brazil	46
Bangladesh	96
Costa Rica	43
Colombia	58
Combodia	68
El Salvador	70
Egypt	100
Ecuador	46
Guatemale	75
Ghana	75
India	100
Kosovo	22
Kazaksthan	64
Jordon	84
Mexico	93
Morocco	57
Nicaragua	59
Peru	43
Pakistan	95
Palestine	100
Tanzania	32
Tajikistan	34

Source: Mix Market 2019.

4. Methodology

Data envelopment analysis (DEA) is a non-parametric bootstrap approach that requires simple time-series data (with no financial gaps) to analyze the competence scores without any complex assumptions in any industry. Charnes et al. 1978 first developed the efficiency measuring model, later applied to identify the organizational performances in the banking sector (Wijesiri et al. 2015, 2017). Bootstrapping has the simple notion of repeated data generating process (DGP) and applying the original simulant to each outcome model (Wijesiri et al. 2015). The selection of appropriate input and output variables depends on the input-oriented and output-oriented methods. In this analysis, we restricted the non-parametric, output-orientated, and constant return to scale DEA approach to determine the production performances. The MFI identities are disclosed and are represented by the normal series from 1, 2, 3 . . . ., 127 which act as the decision-making units to produce efficiency scores (Simar and Wilson 2010). We assume for the development of efficiencies from the decision-making units "p" (DMUs), it is necessary to select "m" different inputs and "n" output/s from the available dataset. Each DMU ("p") has one positive input and one positive output.

We assume:

$$x_{ij} \geq 0 \; for\; i\; input \tag{1}$$

$$y_{kj} \geq 0 \; for\; k\; output \tag{2}$$

The virtual output to the virtual input to maximize the efficiency of each DMU (p) is formulated and should be less than or equal to zero.

$$\text{i.e., } maxh_o(u,v) = \frac{\Sigma_r u_r y_{ro}}{\Sigma_i v_i x_{io}} \leq 1 \tag{3}$$

where $j = 1, \ldots, n;\ u_r v_i \geq 0$ for all i and r.

The censored Tobit truncated random effect regression (as we have time-series data) was discussed to exhibit the relation between the measured dual efficiency scores (act as dependent variables) in the first stage of DEA analysis on the explanatory factors. Tobit model aids in determining the marginal changes on dependent variables by concerning lower and upper impediments. It represents the effect of an independent variable over the conditional variable. The regression provides the validation by considering the dual efficiencies and the overall empirical performance of the microfinance industry is obtained and controls different parameters. The linear regression is followed:

$$\theta_{(i,j,p)*} = \beta_0 + \beta_1 var1 + \beta_2 var2 + \varepsilon;\ i,j = 1, \ldots\ldots, p \tag{4}$$

$$\theta_{(i,j,p)*} = \beta_0 + \beta_1 var1 + \varepsilon;\ i,j = 1, \ldots\ldots, p \tag{5}$$

$$\theta_{(i,j,p)*} = \beta_0 + \beta_1 var2 + \varepsilon;\ i,j = 1, \ldots\ldots, p \tag{6}$$

where i represents the social efficiency score, j represents the financial efficiency score, the β_0, β_1, β_2 are the coefficients for the explanatory variables and ε is the statistical error term, where p represents the DMSs/MFIs (127). Equation (3) represents the social efficiency depending on all the three explanatory indicators. Whereas Equations (4)–(6) represent two explanatory indicators. The indicators imply the religious status, geographic region, and type of financial organization. The religious status, geographic position, and type of the institutions are imported as dummy variables and represent binary numbers. For instance, if religion is a faith, I considered it as 1 otherwise 0. The same with the location and type. We tabulate the results from the analysis in Table 6.

5. Results

5.1. First Stage Non-Parametric Constant Rail Scale DEA Approach

Table 7 reports the efficiency scores of 127 MFIs globally. The aggregate efficiency score of MFI in 10 years and the individual scores are calculated separately. It is observed that financial outreach is dominating with social outreach.

Table 7. The DEA social and financial efficiencies.

MFI (p)	Social		Financial		MFI	Social		Financial	
	θ	θ_a	θ	θ_a		θ	θ_a	θ	θ_a
1	0.45	0.43	0.86	0.83	65	0.72	0.71	0.90	0.57
2	0.92	0.67	0.22	0.62	66	0.56	0.50	0.65	0.32
3	0.93	0.73	0.36	0.45	67	0.38	0.43	0.69	0.54
4	0.13	0.14	0.21	0.30	68	0.02	0.02	0.96	0.86
5	0.01	0.39	0.19	0.64	69	0.02	0.52	0.34	0.34
6	0.01	0.47	0.73	0.57	70	0.22	0.51	0.59	0.33
7	0.00	0.49	1.00	0.58	71	0.36	0.50	0.43	0.57
8	1.00	0.45	0.48	0.44	72	0.37	0.73	0.06	0.45
9	0.80	0.60	0.51	0.51	73	0.89	0.59	0.67	0.48
10	0.61	0.44	0.90	0.37	74	1.00	0.61	1.00	0.44
11	0.31	0.52	0.22	0.42	75	0.98	0.64	0.56	0.72
12	0.48	0.33	0.76	0.75	76	0.52	0.53	0.37	0.67
13	0.48	0.57	0.67	0.43	77	0.08	0.46	0.79	0.65
14	0.88	0.21	0.00	0.37	78	0.17	0.54	0.86	0.41
15	0.02	0.57	0.05	0.46	79	0.88	0.59	0.17	0.41
16	0.19	0.27	0.12	0.14	80	0.11	0.62	0.42	0.41
17	0.00	0.62	0.00	0.30	81	0.87	0.94	0.19	0.27
18	0.28	0.61	0.00	0.60	82	0.59	0.42	0.23	0.50
19	0.58	0.51	0.94	0.31	83	0.76	0.57	0.07	0.36
20	1.00	0.86	0.94	0.43	84	0.02	0.59	0.81	0.49
21	0.83	0.56	0.75	0.64	85	0.54	0.51	0.63	0.59
22	0.00	0.45	0.19	0.59	86	0.30	0.51	0.09	0.55
23	0.94	0.57	0.20	0.38	87	0.00	0.52	0.89	0.59
24	1.00	0.84	0.90	0.71	88	0.11	0.43	1.00	0.55
25	0.01	0.19	0.21	0.29	89	0.46	0.45	0.56	0.46
26	0.14	0.07	0.70	0.53	90	0.44	0.58	0.49	0.47
27	0.62	0.22	1.00	0.52	91	0.29	0.65	0.83	0.51
28	1.00	0.21	1.00	0.62	92	0.94	0.64	0.41	0.27
29	0.01	0.01	0.26	0.33	93	0.70	0.90	0.10	0.75
30	0.43	0.14	0.89	0.61	94	0.00	0.18	0.35	0.57
31	0.77	0.61	0.42	0.37	95	0.00	0.41	0.50	0.61
32	0.31	0.48	0.59	0.50	96	0.00	0.41	0.00	0.37
33	0.49	0.47	0.08	0.32	97	0.79	0.73	0.99	0.68
34	0.01	0.29	0.00	0.16	98	1.00	0.87	1.00	0.87
35	0.00	0.27	0.31	0.36	99	0.93	0.63	0.69	0.55
36	0.35	0.51	0.03	0.50	100	0.98	0.63	0.20	0.47
37	0.22	0.18	0.55	0.46	101	0.85	0.43	0.71	0.59
38	0.57	0.62	0.71	0.61	102	0.07	0.49	0.63	0.57
39	0.69	0.59	0.93	0.71	103	0.62	0.46	3.63	0.90
40	0.19	0.11	0.39	0.31	104	0.11	0.57	0.21	0.44
41	0.41	0.50	0.87	0.70	105	1.00	0.67	0.04	0.32
42	0.57	0.53	1.00	0.58	106	0.12	0.20	0.50	0.44
43	0.27	0.61	0.69	0.51	107	0.26	0.46	0.63	0.47
44	0.38	0.40	0.42	0.52	108	0.22	0.17	0.83	0.68
45	0.80	0.47	0.99	0.79	109	0.73	0.65	0.36	0.40
46	0.02	0.49	0.03	0.30	110	0.81	0.44	0.43	0.56
47	0.33	0.52	0.01	0.40	111	0.38	0.58	0.45	0.47
48	0.49	0.63	0.32	0.53	112	0.10	0.49	0.74	0.46

Table 7. *Cont.*

MFI (p)	Social		Financial		MFI	Social		Financial	
	θ	θ_a	θ	θ_a		θ	θ_a	θ	θ_a
49	0.13	0.61	0.47	0.27	113	0.28	0.56	0.68	0.35
50	0.26	0.36	0.98	0.56	114	0.84	0.48	0.92	0.48
51	0.59	0.51	0.98	0.62	115	0.62	0.39	0.76	0.56
52	0.13	0.53	0.54	0.54	116	0.06	0.36	0.89	0.60
53	0.61	0.51	0.50	0.53	117	0.21	0.05	0.70	0.38
54	0.16	0.46	0.33	0.46	118	0.00	1.00	0.00	1.00
55	0.05	0.47	0.38	0.38	119	0.91	0.64	0.3	0.21
56	0.69	0.23	0.64	0.44	120	0.50	0.40	0.68	0.64
57	0.14	0.39	0.88	0.67	121	0.00	0.57	0.00	0.41
58	0.15	0.59	0.97	0.43	122	0.58	0.61	0.00	0.19
59	0.10	0.07	0.33	0.76	123	0.80	0.95	0.73	0.57
60	0.02	0.29	0.28	0.57	124	0.04	0.11	0.13	0.34
61	0.32	0.47	0.57	0.45	125	0.55	0.61	0.80	0.56
62	0.43	0.25	0.79	0.31	126	0.00	0.36	0.62	0.70
63	1.00	0.51	0.89	0.51	127	0.01	0.22	0.10	0.50
64	0.61	0.54	0.40	0.56					

The efficiency scores, aggregate efficiency (θ_a), determined efficiency (θ) are bounded between 0 and 1. The MFIs with the efficiency scores of 1 are assumed developed or highly productive with high outreaches either socially or financially or both. Hence, we had clustered data with both religious and non-religious institutions, it is a bit challenging to spot the differences. The graphical representation of dual efficiency scores of faith and traditional institutions is exhibited in Figure 1.

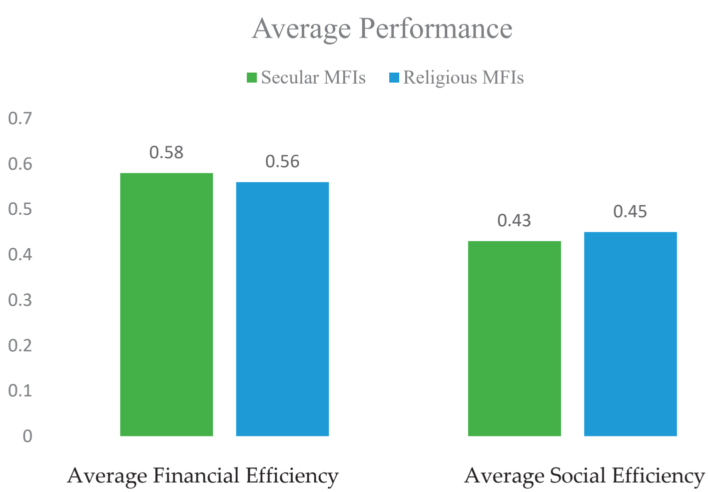

Figure 1. Faith and secular MFIs efficiency scores representation.

5.2. Second Stage Tobit Random Effect Model Regression

The Tobit regression model is an alternative to ordinary least squares regression and is employed when the dependent variable is bounded from below or above or both with positive probability pileup at the interval ends, either by being censored or by being corner solution. The regression with random effect model (unobservable effects are uncorrelated with the observed exploratory variables) shows strong/weak significance levels, the same effect on financial efficiency is vice versa and shows negative relation. The dependent variable aggregate social efficiency (SE) shows an effect with the independent indicators to test the significance of the type of institution and region. The type of order is followed by NBFC, Bank, Microcredit, Credit corporation, and NGOs which are the five types of microfinance institutions located in three different regions, Asia, Africa, Latin America, and the Caribbean. Table 8 represents the censored regression values, how the religious status depends on the explanatory factors.

Table 8. Censored regression to test social efficiency as a response variable.

Social Efficiency (θ_a) (127 MFIs)	**Coefficient**	**Social Efficiency (θ_a)**	**Coefficient**	**Social Efficiency (θ_a)**	**Coefficient**
Religion (faith)	0.04	Religion (faith)	−0.07 **	Religion (faith)	−0.05 *
Region	0.16 *** 0.2 *** −0.09 **	Type	−0.01 −0.00 0.02	Region	0.12 *** 0.15 ** −0.06
Type	−0.02 0.03 −0.06 −0.09 ** −0.17 **		−0.04 −0.05		
Cons	0.63 ***		0.51 ***		0.57 ***

Legend: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

6. Discussion and Conclusions

This study examines the cross-country efficiency analysis of 127 MFIs followed by the religious beliefs and principles from 25 countries. In the literature section, we discussed the importance of women in microfinance and the religiosity effects from the published studies. The DEA is a meta-frontier approach that gained attention in determining the efficiencies with technological gaps in time series with the decision-making units (DMUs). The DMUs are the identities of organizations (financial institutions in our case) with heterogeneous financial properties from various groups and types of institutions (Walheer 2018), here the methodology is applied for the European data in different sectors. The cross-country dual efficiency scores from the first stage results show that social performance is weaker in traditional institutions and financial performance is greater and vice versa. It is also observed that the change in selecting variables does not lead to a change in the efficiency score. This observation leads to the general suggestion that management is inefficient to maintain the outreaches accurately. The conclusion also reaches the management is well established to collect the revenues (loans, interest, profits, etc.) or the clients are more responsible with their repayment to main the health of the economy. The Tobit random effect model uses the efficiency scores (social efficiency, in Table 7) to check the effectiveness of the explanatory indicators. The religion, type of the MFI, and location are artificially created dummy variables chosen in groups to establish a relation between the explanatory factors and the efficiencies (SE). The observation is made with the change in the group, there is a change in the significance level. The main intention is to determine the faith and secular performances, and it is observed with the cross-sectional efficiency scores the financial efficiencies show positive significance level at 0.95 and 0.90 but there is no strong significance level at 0.99. If the microfinance institution is healthily maintained with both social and financial outreaches, there is a high significance to determine the religious status will have stronger significance either financially or socially.

The religious factors have positive aspects to certain limits in accessing the loans, indeed the household poor can gain in their business establishments (women). It is well-known from the literature studies that women are more in accessing loans and good in their repayments. When it comes to traditional organizations, there are several formalities for loan assessment and therefore social outreach can damage in the rural areas where there is an absence of absolute empowerment. The financial organization has strategic moves but if there is no sufficient repayment rate in both secular/faith-based MFIs it leads to negative financial health. Therefore, the efficiency determination should be explored by consideration of the repayment rate in MFIs in future studies.

Microfinance organizations are vanishing the barriers with digital approaches. Although we analyzed with other variables (input and output factors), the results lead to the same conclusions with efficiencies. There should be a combination of dual efficiencies along with technical efficiency which would conclude better results in minimizing the resources in religious and secular microfinance institutions.

The efficiency analysis and regression dealt with the religious, locality, and type of institution factors. There are only a few studies concentrated purely on religious institutions because the count (of religious MFIs) is small in some countries. There is a scope to analyze the performances by considering and concentrating on religious institutions irrespective of religiosity, along with the technical performance.

Author Contributions: A.K.K.—Data curation: Creating and organizing the variables for clear visualization, Data extracting form world bank (Mix market database), selection of variables; Formal analysis: The prior study and understanding of methodology along with study of variables and relation between them; software: Stata 13; Investigation: Study of results, troubleshoots, technical errors; Original draft preparation: Theory and literature design with sectional structures; and conceptualization: The draft concept, ordering data, research questions, tables, figures, and sections. M.M.—Conceptualization: The review, design ideas, methodologies; review and editing: The table structures, definitions, concepts, tables structure (internally); Visualization: sectional organization, study of results, rectifying error terms; Project administration: The resources with assistance; Software: Stata 13; Investigation: Study of results, troubleshoots, technical errors; Methodology: The teaching and explanation of appropriate study to conduct and perform the analysis. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data was extracted from the Mix market database (world bank).

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

Notes

1 NBFC—Non-bank Financial corporation.

2 NGO—Non-Government Organization.

3 ROA—Return on Assets.

4 ROE—Return on Equity.

5 Data envelopment analysis—constant return to scale.

6 All three external variables.

7 Two external variables individually.

References

Aggarwal, Raj, John W. Goodell, and Lauren J. Selleck. 2015. Lending to women in microfinance: Role of social trust. *International Business Review* 24: 55–65. [CrossRef]

Agrawal, Pawan, and Shayan Sen. 2017. Digital economy and microfinance. *The MIBM Research Journal* 5: 27–35.

Alimukhamedova, Nargiza. 2013. Contribution of microfinance to economic growth: Transmission channel and the ways to test it. *Business and Economic Horizons* 9: 27–43. [CrossRef]

Amemiya, Takeshi. 1984. Tobit Models: A Survey. *Journal of Econometrics* 24: 3–61. [CrossRef]

Barro, Robert J., and Rachel Mccleary. 2003. Religion and economic growth across countries. *American Sociological Review* 68: 760–781. [CrossRef]

Belwal, Rakesh, Misrak Tamiru, and Gurmeet Singh. 2012. Microfinance and Sustained Economic Improvement: Women Small-Scale Entrepreneurs in Ethiopia. *Journal of International Development* 24: 84–99. [CrossRef]

Ben Abdelkader, Ines, and Faysal Mansouri. 2019. Performance of microfinance institutions in the MENA region: A comparative analysis. *International Journal of Social Economics* 46: 47–65. [CrossRef]

Charnes, Abraham, William W. Cooper, and Edwardo Rhodes. 1978. Measuring the efficiency of decision making units. *Company European Journal of Operational Research* 2: 429–44. [CrossRef]

Cobb, J. Adam, Tyler Wry, and Eric Yanfei Zhao. 2016. Funding Financial Inclusion: Institutional Logics and the Contextual Contingency of Funding for Microfinance Organizations. *Academy of Management Journal* 59: 2103–31. [CrossRef]

Daley-Harris, S. 2009. State of the Microcredit Summit Campaign Report 2009 Sam Daley-Harris. Available online: https://www.researchgate.net/publication/265114804_State_of_the_Microcredit_Summit_Campaign_Report_2009 (accessed on 11 February 2022).

Fersi, Marwa, and Mouna Boujelbéne. 2016. The Determinants of the Performance and the Sustainability of Conventional and Islamic Microfinance Institutions. *Economics World* 4: 197–215. [CrossRef]

Gutiérrez-Nieto, Begoña, Carlos Serrano-Cinca, and Cecilio Mar Molinero. 2009. Social efficiency in microfinance institutions. *Journal of the Operational Research Society* 60: 104–19. [CrossRef]

Gyapong, Ernest, Daniel Gyimah, and Ammad Ahmed. 2021. Religiosity, borrower gender and loan losses in microfinance institutions: A global evidence. In *Review of Quantitative Finance and Accounting*. New York: Springer, vol. 57. [CrossRef]

Hadi, Nabawiyah Abdul, and Amrizah Kamaluddin. 2015. Social Collateral, Repayment Rates, and the Creation of Capital among the Clients of Microfinance. *Procedia Economics and Finance* 31: 823–828. [CrossRef]

Imam, Patrick, and Kangni Kpodar. 2016. Islamic banking: Good for growth? *Economic Modelling* 59: 387–401. [CrossRef]

Kar, Ashim Kumar. 2012. Does capital and financing structure have any relevance to the performance of microfinance institutions? *International Review of Applied Economics* 26: 329–48. [CrossRef]

Khan, Zahoor, and Jamalludin Sulaiman. 2015. Social and Financial Efficiency of Microfinance Institutions in Pakistan. *The Pakistan Development Review* 54: 389–403. [CrossRef]

Lebovic, James H. 2004. Uniting for peace? Democracies and United Nations peace operations after the cold war. *Journal of Conflict Resolution* 48: 910–36. [CrossRef]

Leite, Rodrigo de Oliveira, Layla Mendes, and Luiz Claudio Sacramento. 2019. To profit or not to profit? Assessing financial sustainability outcomes of microfinance institutions. *International Journal of Finance & Economics*, 1–13. [CrossRef]

Li, Linyang, Niels Hermes, and Aljar Meesters. 2019. Convergence of the performance of microfinance institutions: A decomposition analysis. *Economic Modelling* 81: 308–24. [CrossRef]

Liñares-Zegarra, J., and John O. S. Wilson. 2018. The size and growth of micro finance institutions. *The British Accounting Review* 50: 199–213. [CrossRef]

Mansori, Shaheen, Meysam Safari, and Zarina Mizam Mohd Ismail. 2020. An analysis of the religious, social factors and income's influence on the decision making in Islamic microfinance schemes. *Journal of Islamic Accounting and Business Research* 11: 361–76. [CrossRef]

Mersland, Roy, D'Espallier Bert, and Magne Supphellen. 2013. The Effects of Religion on Development Efforts: Evidence from the Microfinance Industry and a Research Agenda. *World Development* 41: 145–56. [CrossRef]

Mia, Md Aslam, and V. G. R Chandran Govindaraju. 2016. Measuring Financial and Social Outreach Productivity of Microfinance Institutions in Bangladesh. *Social Indicators Research* 127: 505–27. [CrossRef]

Navajas, Sergio, and Luis Tejerina. 2006. Microfinance in Latin America and the Caribbean: How Large Is the Market? Available online: http://www.asbasupervision.com/en/bibl/financial-inclusion/microfinance/223-microfinance-in-latin-america-and-the-caribbean-how-large-is-the-market/file (accessed on 11 February 2022).

Sainz-Fernandez, Isabel, Begoña Torre-Olmo, Carlos López-Gutiérrez, and Sergio Sanfilippo-Azofra. 2018. Development of the Financial Sector and Growth of Microfinance Institutions: The Moderating Effect of Economic Growth. *Sustainability* 10: 3930. [CrossRef]

Schwarz, Susan, Alexander Newman, and Abu Zafar Shahriar. 2015. Profit orientation of microfinance institutions and provision of financial capital to business start-ups. *International Small Business Journal* 34: 532–52. [CrossRef]

Seiford, Lawrence M., and Joe Zhu. 2014. Data Envelopment Analysis: History, Models, and Interpretations Chapter 1 (Issue August 2011). Available online: https://www.researchgate.net/publication/226038831_Data_Envelopment_Analysis_History_Models_and_Interpretations (accessed on 11 February 2022).

Simar, Léopold Leopold, and Paul W. Wilson. 2007. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics* 136: 31–64. [CrossRef]

Simar, Léopold, and Paul W. Wilson. 2010. A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics* 27: 779–802. [CrossRef]

Somavia, J. 2007. Why Target Women? Promoting Gender Equality and Decent Work through Microfinance in Tanzania. Available online: www.ilo.org/gender (accessed on 29 December 2021).

Tchuigoua, Hubert Tchakoute. 2015. Capital Structure of Microfinance Institutions. *Journal of Financial Services Research* 43: 313–40. [CrossRef]

Ul-Hameed, Waseem, Hisham Bin Mohammad, and Hanita Binti Kadir Shahar. 2018. Microfinance institute's non-financial services and women-empowerment: The role of vulnerability. *Management Science Letters* 8: 1103–16. [CrossRef]

Walheer, Barnabé. 2018. Aggregation of metafrontier technology gap ratios: The case of European sectors in 1995–2015. *European Journal of Operational Research* 269: 1013–26. [CrossRef]

Wijesiri, Mahinda, and Michele Meoli. 2015. Journal of Retailing and Consumer Services Productivity change of micro finance institutions in Kenya: A bootstrap Malmquist approach. *Journal of Retailing and Consumer Services* 25: 115–21. [CrossRef]

Wijesiri, Mahinda, Laura Viganò, and Michele Meoli. 2015. Efficiency of microfinance institutions in Sri Lanka: A two-stage double bootstrap DEA approach. *Economic Modelling* 47: 74–83. [CrossRef]

Wijesiri, Mahinda, Jacob Yaron, and Michele Meoli. 2017. Journal of Multinational Financial Assessing the financial and outreach efficiency of microfinance institutions: Do age and size matter? *Journal of Multinational Financial Management* 40: 63–76. [CrossRef]

economies

Article

Firm Survival and Gender of Firm Owner in Times of COVID-19: Evidence from 10 European Countries

Joachim Wagner [1,2]

[1] Kiel Institute for the World Economy, Leuphana University of Lüneburg, 21314 Lüneburg, Germany; wagner@leuphana.de
[2] IZA Institute of Labour Economics, 53113 Bonn, Germany

Abstract: This paper uses firm level data from World Bank Enterprise surveys conducted in 2019, and COVID-19 follow-up surveys conducted in 2020, in ten European countries to investigate the link between the gender of the firm's owner and the firm's survival until 2020. The empirical investigation uses econometric models that control for the firm's characteristics that are known to be related to firm survival. The estimated effect of female ownership is positive ceteris paribus. Furthermore, the size of this estimated effect can be considered to be large on average. Having a female owner helped firms to survive.

Keywords: gender; female owned firms; firm survival; COVID-19; World Bank Enterprise surveys

JEL Classification: D22; L20; L25; L29

Citation: Wagner, Joachim. 2022. Firm Survival and Gender of Firm Owner in Times of COVID-19: Evidence from 10 European Countries. *Economies* 10: 98. https://doi.org/10.3390/economies10050098

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 7 March 2022
Accepted: 18 April 2022
Published: 21 April 2022

1. Motivation

When the coronavirus and COVID-19 reached Europe in the first quarter of 2020, firms were hit by negative demand shocks due to quarantine and lockdown measures. Furthermore, supply chains were damaged and this led to negative supply shocks. These shocks had a negative impact on many dimensions of firm performance. Waldkirch (2021) reports evidence on the impact of the COVID-19 pandemic on firms around the world based on the so-called COVID-19 follow-up surveys to the World Bank's Enterprise Surveys conducted in 2020.

Some firms were hit so hard by these negative exogenous shocks that they decided to close down permanently. Waldkirch (2021, p. 6) reports that in samples of firms, collected in 21 countries in Europe, Latin America, Africa, and Asia in 2019/2020, 4.1 percent of the businesses had permanently closed. In the sample of firms used here in this study, the share of exits is 4.59 percent (see Table 1).

An important question is which characteristics of firms help many of them to survive the pandemic. In addition to the usual suspects, which have been discussed at length in the literature over recent decades, and include firm demographics such as firm age, firm size, exports, productivity, and innovation (and that will be looked at in more detail in Section 2 of this paper), one firm characteristic that has, to the best of my knowledge, not been considered yet in the context of the COVID-19 pandemic, is the gender of the firm's owner. In their comprehensive survey of the literature, Jennings and Brush (2013, p. 671) argue that many empirical studies find that with respect to many standard economic indicators, female-led firms tend to exhibit inferior performance; however, Jennings and Brush (2013, p. 672) point out that a small yet growing number of studies challenge whether firms that are female-owned invariably underperform relative to male-owned firms. There are studies that reveal mixed results, no significant differences, or even an advantage to female-led firms. "In terms of survival, for instance, the evidence is mixed." (Jennings and Brush 2013, p. 672).

Table 1. Descriptive evidence on share of exits by gender of owner in 10 European countries, 2019/2020.

Country	No. of Firms	Share of Exits (Percent)	Share of Female Owned Firms (Percent)	Share of Exits among Female Owned Firms (Percent)	Share of Exits among Male Owned Firms (Percent)
All countries	6013	4.59	36.37	3.98	4.94
Bulgaria	552	6.70	37.86	5.74	7.29
Croatia	349	2.58	32.09	0.89	3.38
Czech Rep.	398	1.76	32.91	1.53	1.87
Hungary	623	1.77	49.12	0.98	2.52
Italy	439	7.74	20.96	3.26	8.93
Poland	887	2.93	39.80	3.68	2.43
Portugal	795	9.69	43.65	7.49	11.38
Romania	518	3.47	36.68	5.26	2.44
Russia	1116	3.94	29.66	4.83	3.57
Slovak Rep.	336	3.87	34.52	0.86	5.45

Source: Own calculations based on the World Bank Enterprise surveys; for details, see text.

Although we have (to the best of my knowledge) no empirical evidence on the difference of firm survival by gender of firm owner during the COVID-19 pandemic, one may hypothesize that female owned firms have a lower rate of closure because women tend to be more risk adverse than men, and therefore, firms lead by men might be engaged in more risky businesses and have a higher risk of failure, especially in times of negative demand and supply shocks.

This paper contributes to the literature by using firm level data from ten European countries collected in the World Bank's Enterprise Surveys in 2019, and from the COVID-19 follow-up surveys conducted in 2020, to investigate the link between gender of the firm's owner and firm survival, controlling for other determinants of firm exit.

The rest of the paper is organized as follows: Section 2 introduces the data used and discusses the variables that are included in the empirical model to test for the role of the gender of the firm's owner in firm survival. Section 3 reports descriptive evidence and results from the econometric investigation. Section 4 concludes.

2. Data and Discussion of Variables

The firm level data used in this study are taken from the World Bank's Enterprise Surveys in 2019 and from the COVID-19 follow-up surveys conducted in 2020.[1] These surveys were conducted in a large number of countries all over the world. In this study we focus on countries from Europe. All countries with complete data for at least five firms, that took part in the 2019 survey, and reported in the 2020 follow-up survey that they had permanently closed down, are included in the study. This leaves us with data for ten countries: Bulgaria, Croatia, Czech Republic, Hungary, Italy, Poland, Portugal, Romania, Russia, and the Slovak Republic.[2]

The classification of firms as survivors or exits is based on question B.0[3] in the follow-up survey from 2020. Firms that participated both in the regular 2019 survey and in the follow-up survey were asked "[c]urrently is this establishment open, temporarily closed (suspended services or production), or permanently closed?" Firms that answered "permanently closed" are classified as exits, the other firms are considered to be survivors.

The classification of firms into female-owned firms and male-owned firms is based on the answer to question B4 in the regular survey of 2019. When there are females amongst the owners of the firm, a firm is considered to be a female-owned firm (and a male-owned firm otherwise).

Descriptive evidence on the share of firm exits by gender of the firms' owners in the total sample and by country is reported in in Table 1. Although the overall share of exits among female owned firms is 3.98 percent, it is about one percentage point higher among male-owned firms (4.94 percentage points). This raw difference in favor of female-owned firms can be considered large. Results differ by country, but in seven out of ten countries, the exit rate is smaller among female-owned firms.

In the empirical investigation of the link between the gender of the firm's owner and firm survival, a number of firm characteristics that are known to be correlated with firm exit (and that might be related to the gender of the firm's owner as well) are controlled for.

Their link to firm survival, and the way they are measured here, is discussed below (see Wagner 2021).

Firm size: Audretsch (1995, p. 149) mentions as a stylized fact from many empirical studies on exits that the likelihood of a firm exit apparently declines with firm size (usually measured by the number of employees in a firm). This is theoretically linked to the hypothesis of "liability of smallness" from organizational ecology. A small size can be interpreted as a proxy variable for a number of unobserved firm characteristics, including disadvantages of scale, higher restrictions on the capital market leading to a higher risk of insolvency and illiquidity, disadvantages of small firms in the competition for highly qualified employees, and lower talent of management (Strotmann 2007). For Germany, Fackler et al. (2013) show that the mortality risk falls with establishment size, which confirms the liability of smallness.

Firm size is measured as the number of permanent, full-time individuals that worked in the establishment at the end of the last complete fiscal year at the time of the regular 2019 enterprise survey (see question I.1).

Firm age: Audretsch (1995, p. 149) mentions as another stylized fact from many empirical studies on exits that the likelihood of a firm exit apparently declines with firm age, too. This positive link between firm age and probability of survival is labelled "liability of newness" and it is related to the fact that older firms are "better" because they spent a longer time in the market, during which they learned how to solve the range of problems facing them in day-to-day business. For Germany, Fackler et al. (2013) find that the probability of exit is substantially higher for young establishments which are not more than five years old, thus confirming the liability of newness.

Firm age is measured as follows. In question B.5 of the regular survey in 2019, firms were asked "[i]n what year did this establishment begin operation?". Firm age is the difference between 2019 and the founding year.

Exports: Exporting can be considered as a form of risk diversification through spread of sales over different markets with different business cycle conditions or in a different phase of the product cycle; therefore, exports might provide a chance to substitute sales at home by sales abroad when a negative demand shock hits the home market and would force a firm to close down otherwise (see Wagner 2013). Furthermore, Baldwin and Yan (2011, p. 135) argue that non-exporters are, in general, less efficient than exporters (younger, smaller, and less productive) and that, as a result, one expects that non-exporters are more likely to fail than exporters.

A number of recent empirical studies look at the role of international trade activities in shaping the chances for survival of firms; Wagner (2012, p. 256ff.) summarizes this literature. As a rule, the estimated chance of survival is higher for exporters, and this holds after controlling for firm characteristics that are positively associated with both exports and survival (such as firm size and firm age). This might point to a direct positive effect of exporting on survival.

The firm is considered to be an exporter if it reports any direct exports in question D.3 of the regular enterprise survey in 2019.

Productivity: In theoretical models for the dynamics of industries with heterogeneous firms, productivity differentials play a central role for entry, growth, and exit of firms. In terms of the equilibrium growing and shrinking, exiting and entering firms that have different productivities are found in an industry. These models lead to hypotheses that can be tested empirically. Hopenhayn (1992) considers a long-run equilibrium in an industry with many price-taking firms producing a homogeneous good. Output is a function of inputs, and a random variable that models a firm specific productivity shock. These shocks are independent between firms and are the reason for the heterogeneity of firms. There are sunk costs to be paid at entry, and entrants do not know their specific shock in advance. Incumbents can choose between exiting or staying in the market. When firms realize their productivity shock, they decide the profit maximizing volume of production. The model assumes that a higher shock in $t + 1$ has a higher probability the higher the shock is in t. In equilibrium, firms will exit if, for given prices of output and input, the productivity shock is smaller than a critical value, and production is no longer profitable.

Farinas and Ruano (2005, p. 507f.) argue that this model leads to the following testable hypothesis: firms that exit in year t were $t-1$ less productive than firms that continue to produce in t. They test this hypothesis using panel data for Spanish firms. The hypothesis is supported by the data. Wagner (2009) replicates the study by Farinas and Ruano (2005) with panel data for West and East German firms from manufacturing industries. For the cohorts of exit from 1997 to 2002, the results are in line with the results for Spain.

Unfortunately, however, there is no suitable measure of productivity in the World Bank Enterprise survey, so productivity cannot be controlled for in the empirical models that test for a link between the gender of the firm's owner and firm survival; however, productivity is controlled for indirectly by including the information on the exporter status of the firm, because it is a stylized fact that has been found in hundreds of empirical studies from countries all over the world that exporters tend to be much more productive than non-exporters from the same narrowly defined industry (see Wagner 2007 for a survey).

Foreign ownership: Baldwin and Yan (2011) argue that from a theoretical point of view, the relationship that should be expected between foreign ownership and firm exit is not clear. On the one hand, foreign owned firms may have access to superior technologies belonging to their foreign owners that might increase their efficiency and lower the risk of exit. Their greater propensity to invest in R&D might lead to more innovations, improve their competitiveness at home and in foreign markets, and might therefore increase the chance of survival. On the other hand, Baldwin and Yan (2011) point out that foreign owned firms are less rooted in the host country economy, and they can shift their activities to another country when the local economy deteriorates. This should increase the probability of shutdown compared with nationally owned firms.

With a view on the COVID-19 pandemic, Waldkirch (2021, p. 4) argues that "on the one hand, multinational companies may be better able to weather the storm, as they are more financially stable or have access to multiple sources of inputs, thereby minimizing disruptions to the supply chain. On the other hand, these firms may also be exposed to the pandemic's impacts on a larger scale, in multiple countries, and at different times given the differential timing of the virus's spread and mandated quarantines and shutdowns in different countries".

A number of recent micro-econometric studies use firm level data for foreign owned firms and domestically controlled firms to investigate the (ceteris paribus) relationship between foreign ownership and firm survival. The Wagner and Weche Gelübcke (2012) survey of 26, mainly country specific studies, uses data from 17 developed and developing countries, two of which use data on affiliates worldwide. The big picture emerging from the findings of these studies can be summarized as follows. Results are highly country dependent. Foreign affiliates were found to be more likely to exit as compared with their domestic counterparts in Ireland, Belgium, Spain, and Indonesia, but less likely to exit in Canada, Italy, Taiwan, and the US. No significant differences in closure rates due to foreign ownership could be revealed for Japan, Turkey, and the UK.

In the regular survey in 2019, firms were asked what percentage of this firm is owned by private foreign individuals, companies, or organizations (see question B2). Firms that reported a positive amount here are considered as (partly) foreign owned firms.

Innovation: Josef Schumpeter (1942, p. 84) argued some 80 years ago that innovation plays a key role for the survival of firms, because it "strikes not at the margins of the profits and the outputs of the existing firms but at their foundations and their very lives". Baumol (2002, p. 1) called innovative activity "a life-and-death matter for the firm". This positive link between innovation and firm survival is found in a number of empirical studies. For example, Cefis and Marsili (2005) show that firms benefit from an innovation premium in that ceteris paribus extends their life expectancy; process innovation in particular seems to have a positive effect on firm survival.

In the regular survey in 2019, firms were asked whether during the last three years, this establishment has introduced new of improved products and services (see question H1). Firms that answered in the affirmative are considered to be product innovators. Similarly, firms were asked whether during the last three years, this establishment introduced any new or improved process, including methods of manufacturing products or offering services; logistics, delivery, or distribution methods for inputs, products, or services; or supporting activities for processes (see question H5). Firms that answered in the affirmative are considered as process innovators.

Web presence, i.e., having a website where potential customers can learn about, and order, goods or services, when personal contacts are not possible due to quarantine, and lockdown is often mentioned in the business press as a factor that might help firms to survive in the pandemic. Wagner (2021) uses the same data used here to test this hypothesis. He finds that the estimated effect of web presence on firm survival is positive, statistically significant, and large.

In the regular 2019 survey, firms were asked in question C22b, "[a]t present time, does this establishment have its own website or social media page?" Firms that answered "yes" are classified as a firm with web presence.

Furthermore, firms are divided by broad sectors of activity (manufacturing, retail/wholesale, construction, hotel/restaurant, and services) based on their answer to the question for the establishment's main activity and product, measured by the largest proportion of annual sales (see question D1a1).

Descriptive information on the difference between female and male owned firms in the sample are reported in Table 2. Compared with male owned firms, female owned firms have a slightly lower web presence, are smaller, are product innovators more often, are foreign owned firms less often, and are exporters less often. Furthermore, they are more often in the retail/wholesale and hotel/restaurant industry, and less often in construction and services.

Table 2. Difference between female and male owned firms in the sample table.

Variable	Female Owned Firms	Male Owned Firms
Web-presence (Dummy; 1 = yes)	0.6936 (0.46)	0.7357 (0.44)
Firm age (Years)	21.84 (16.53)	20.60 (14.50)
Firm size (Number of employees)	62.75 (135.69)	92.43 (432.36)
Product innovator (Dummy; 1 = yes)	0.2437 (0.42)	0.2062 (0.40)
Process innovator (Dummy; 1 = yes)	0.1280 (0.33)	0.1270 (0.33)
Foreign owned firm (Dummy; 1 = yes)	0.0453 (0.21)	0.0844 (0.28)
Exporter (Dummy; 1 = yes)	0.2515 (0.43)	0.2882 (0.44)
Manufacturing (Dummy; 1 = yes)	0.6255 (0.48)	0.6349 (0.48)
Retail/Wholesale (Dummy; 1 = yes)	0.2318 (0.42)	0.1832 (0.39)
Construction (Dummy; 1 = yes)	0.0366 (0.19)	0.0659 (0.25)
Hotel/Restaurant (Dummy; 1 = yes)	0.0425 (0.20)	0.0311 (0.17)
Services (Dummy; 1 = yes)	0.0636 (0.24)	0.0849 (0.28)
Number of firms	2187	3826

The table reports the mean values of the variables used in the estimation by gender of firm owner (standard deviations in brackets).

Descriptive statistics for all variables used in the empirical investigation are reported for the whole sample in the Appendix A Table A1.

3. Testing for the Role of Gender of the Firm's Owner in Firm Survival

To test for the role of the gender of the firm's owner in firm survival, empirical models are estimated with an indicator variable for firm survival, or not, until 2000 as the endogenous variable, an indicator variable for female-owned firms, or not, in 2019 as the exogenous variable, and various sets of control variables. All models are estimated by Probit, and average marginal effects with prob-values to indicate their statistical significance are reported.

Four different variants of empirical models are estimated. Model 1 has only the indicator variable for female-owned firms as an exogenous variable; Model 2 adds a set of country dummy variables, Model 3 adds a set of sector dummy variables, and Model 4 includes all control variables detailed in Section 2, too. Results are reported in Table 3.

Table 3. Gender of firm owner and firm exit in 10 European countries, 2019/2020: results from econometric models. Method: Probit (Average Marginal Effects); Dependent variable: Firm exit (1 = yes).

Model Variable		**1**	**2**	**3**	**4**
Female-owned firm	Average marginal effect	−0.00962	−0.00885	−0.00873	−0.00928
(Dummy; 1 = yes)	*p*-value	0.078	0.106	0.112	0.088
Web-presence	Average marginal effect				−0.0252
(Dummy; 1 = yes)	*p*-value				0.000
Firm age	Average marginal effect				−0.00067
(Years)	*p*-value				0.003
Firm size	Average marginal effect				−0.01163
(Log Number of employees)	*p*-value				0.000
Exporter	Average marginal effect				−0.0139
(Dummy; 1 = yes)	*p*-value				0.042
Foreign owned firm	Average marginal effect				0.0163
(Dummy; 1 = yes)	*p*-value				0.286
Product innovator	Average marginal effect				−0.0114
(Dummy; 1 = yes)	*p*-value				0.090
Process innovator	Average marginal effect				−0.0210
(Dummy; 1 = yes)	*p*-value				0.006
Country dummy variables		no	yes	yes	yes
Sector dummy variables		no	no	yes	yes
Number of observations		6013	6013	6013	6013

Source: Own calculations with data from World Bank Enterprise surveys; for details see text.

The most important result is that the estimated average marginal effect of being a female-owned firm, on firm exit, is negative and statistically significant at a ten percent level in all four empirical models. Irrespective of the control variables included, the model of female ownership in 2019 reduces the probability of firm exit until 2020.

With regard to the control variables included in Model 4, all of the estimated average marginal effects have the theoretically expected sign (as discussed in Section 2 above), and are statistically different from zero at an error level of 8 percent or much better, the only exception being the indicator for a foreign owned firm (where no clear theoretical hypothesis is found in the literature according to the discussion in Section 2 above).

Note that the estimated average marginal effect of female ownership on the chance to survive is about constant over all four models, so adding control variables does not change the results much. Furthermore, the size of this estimated effect can be considered as being large on average—the estimated average reduction in the probability of exit is slightly less than one percentage point, and this is really large compared with the overall exit probability of 4.59 percent in the sample reported in Table 1. Having a female owmer helped firms to survive the negative shocks during the pandemic.

4. Concluding Remarks

This paper demonstrates that having a female owner is positively related to the probability of survival for firms facing negative demand and supply shocks during the COVID-19 pandemic. The estimated effect is statistically significant at a reasonable level ceteris paribus after controlling for various firm characteristics that are known to be positively related to survival. Furthermore, the size of this estimated effect can be considered as being large on average. Female owners helped firms to survive.

This finding might be explained by different degrees of risk aversion by gender. If women tend to be more risk averse than men, and firms led by men are engaged in more risky businesses, it follows that they will have a higher risk of failure, especially in times of negative demand and supply shocks; however, the data at hand are not rich enough to investigate whether this is the case. The empirical findings presented here are nevertheless interesting on its own.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in this article.

Acknowledgments: I thank four anonymous referees for helpful comments on an earlier version. The data from the World Bank Enterprise surveys are available after registration from the website https://www.enterprisesurveys.org/portal/login.aspx, accessed on 20 February 2022. Stata code used to produce the empirical results reported in this note is available from the author.

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

Appendix A

Table A1. Descriptive statistics for sample used in estimations.

Variable	Mean	Std. Dev.
Female-owned firm (Dummy; 1 = yes)	0.363	0.481
Firm exit (Dummy; 1 = yes)	0.046	0.209
Web-presence (Dummy; 1 = yes)	0.720	0.449
Firm age (Years)	21.05	15.28
Firm size (Number of employees)	81.64	354.73
Product innovator (Dummy; 1 = yes)	0.220	0.414
Process innovator	0.127	0.333

Table A1. *Cont.*

Variable	Mean	Std. Dev.
(Dummy; 1 = yes)		
Foreign owned firm (Dummy; 1 = yes)	0.070	0.255
Exporter (Dummy; 1 = yes)	0.256	0.436
Manufacturing (Dummy; 1 = yes)	0.631	0.482
Retail/Wholesale (Dummy; 1 = yes)	0.201	0.401
Construction (Dummy; 1 = yes)	0.055	0.228
Hotel/Restaurant (Dummy; 1 = yes)	0.035	0.184
Services (Dummy; 1 = yes)	0.077	0.267

Notes

1. The data from the World Bank Enterprise surveys are available free of charge after registration from the website https://www.enterprisesurveys.org/portal/login.aspx, accessed on 20 February 2022.
2. Not included are Albania, Cyprus, Estonia, Greece, Latvia, Lithuania, Malta, and Slovenia.
3. The questionnaires of the regular 2019 survey and the follow-up survey conducted in 2020 are available from the World Bank's Enterprise Survey web site referred to above.

References

Audretsch, David B. 1995. *Innovation and Industry Evolution*. Cambridge: Cambridge University Press.

Baldwin, John, and Beiling Yan. 2011. The death of Canadian manufacturing plants: Heterogeneous responses to changes in tariffs and real exchange rates. *Review of World Economics* 147: 131–67. [CrossRef]

Baumol, William J. 2002. *The Free-Market Innovation Machine: Analyzing the Growth Miracle of Capitalism*. Princeton: Princeton University Press.

Cefis, Elena, and Orietta Marsili. 2005. A matter of life and death: Innovation and firm survival. *Industrial and Corporate Change* 14: 1167–92. [CrossRef]

Fackler, Daniel, Claus Schnabel, and Joachim Wagner. 2013. Establishment exits in Germany: The role of size and age. *Small Business Economics* 41: 683–700. [CrossRef]

Farinas, Jose C., and Sonia Ruano. 2005. Firm productivity, heterogeneity, sunk costs and market selection. *International Journal of Industrial Organization* 23: 505–34. [CrossRef]

Hopenhayn, Hugo. 1992. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica* 60: 1127–50. [CrossRef]

Jennings, Jennifer E., and Candida G. Brush. 2013. Research on Women Entrepreneurs: Challenges to (and from) the Broader Entrepreneurship Literature? *The Academy of Management Annals* 7: 663–715. [CrossRef]

Schumpeter, Joseph A. 1942. *Capitalism, Socialism and Democracy*. New York: Harper & Row.

Strotmann, Harald. 2007. Entrepreneurial Survival. *Small Business Economics* 28: 87–104. [CrossRef]

Wagner, Joachim. 2007. Exports and Productivity: A survey of the evidence from firm level data. *The World Economy* 30: 5–32. [CrossRef]

Wagner, Joachim. 2009. Entry, Exit and Productivity: Empirical Results for German Manufacturing Industries. *German Economic Review* 11: 78–85. [CrossRef]

Wagner, Joachim. 2012. International trade and firm performance: A survey of empirical studies since 2006. *Review of World Economics* 148: 235–67. [CrossRef]

Wagner, Joachim. 2013. Exports, Imports and Firm Survival: First evidence for manufacturing enterprises in Germany. *Review of World Economics* 149: 113–30. [CrossRef]

Wagner, Joachim. 2021. With a little help from my website. Firm survival and web presence in times of COVID-19—Evidence from 10 European countries. *Economics Bulletin* 41: 1898–906.

Wagner, Joachim, and John Phillipp Weche Gelübcke. 2012. Foreign Ownership and Firm Survival: Fist Evidence for Enterprises in Germany. *International Economics* 132: 117–39. [CrossRef]

Waldkirch, Andreas. 2021. Firms around the World during the COVID-19 Pandemic. *Journal of Economic Integration* 16: 1–19. [CrossRef]

economies

Article

The China Shock Impact on Labor Informality: The Effects on Brazilian Manufacturing Workers

Lourenço S. Paz

Department of Economics, Baylor University, Waco, TX 76798, USA; lourenco_paz@baylor.edu

Abstract: The vigorous growth of the Chinese economy together with its increasingly successful role in international trade may have profoundly impacted developing countries. This study examines the large increase in the international trade exposure of the Brazilian economy during 2000–2012 to assess the impacts of import competition on its manufacturing formal and informal labor markets. In this period, import penetration grew by more than 20 percent in Brazil, and the share of the import penetration originating in China increased from 3 to 20 percent. At the same time, the share of informal workers in manufacturing declined from 27 to approximately 15 percent. Employing a switching regression model and Brazilian household survey data, this study finds that a greater industry-level Chinese and 'rest of the world' import penetration increases the likelihood of jobs becoming informal at different intensities, and these effects are smaller in unskilled-labor intensive industries and manufacturing states. Additionally, both types of import penetration positively impact the average informal wage. In contrast, the estimates suggest that a larger Chinese import penetration reduces average formal wages, while imports from elsewhere have the opposite effect. The results also indicate that the magnitude of the effects on wages are moderated by the unskilled labor intensity of the industry and whether the worker is located in a manufacturing state.

Keywords: Brazil; China; import penetration; informal labor markets; wages

Citation: Paz, Lourenço S. 2022. The China Shock Impact on Labor Informality: The Effects on Brazilian Manufacturing Workers. *Economies* 10: 109. https://doi.org/10.3390/economies10050109

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 28 March 2022
Accepted: 4 May 2022
Published: 7 May 2022

1. Introduction

China has experienced an impressive economic transformation in terms of fast economic growth and of increasing participation in international trade since the late 1970s. Indeed, China's share of world trade increased from four percent in 2000 to ten percent in 2012, while the world trade flows expanded by 75 percent (United Nations 2003). This unprecedented expansion in China's exports both in absolute and in relative terms became known as the China shock. This rapid ascension of China as a major manufacturing powerhouse—about 90 percent of its exports are made of manufactured goods (Paz 2018)—raised fears of deindustrialization in Latin American and other developing countries. Such concerns are grounded on the fact that China has a huge labor endowment that makes it a labor-abundant country relative to those in the developing world. Moreover, its large domestic market leads to economies of scale that are important in several manufacturing industries (Moreira 2006). These features provide China with a strong competitive edge in world markets.

According to many observers, manufacturing not only generally pays higher wages than agriculture or services, but is also a key contributor to economy-wide productivity growth. Additionally, developing countries typically display a large share of informal workers in manufacturing. These informal workers are usually less productive, paid lower wages, and account for at least a quarter of the manufacturing labor force in countries like Brazil, Colombia, and Mexico (Paz 2014; Dávalos 2019). Since the manufacturing sector is the most exposed sector to import competition, it is paramount to study the effects of globalization on developing countries' manufacturing, separately for formal and informal workers.

In this vein, the case of Brazil is interesting because it is Latin America's largest economy with a sizable manufacturing industry (Araújo and Paz 2014). Moreover, Brazil exhibited during 2000–2012 an increase in its manufacturing import penetration, in excess of 20 percent. At the same time, China's accession to the WTO granted Chinese products better access to foreign markets (Chandra 2014). In fact, the Chinese share of Brazilian imports increased six-fold, from 3 to 20 percent, which made China the largest exporter to Brazil. Interestingly, Facchini et al. (2010) point out that Chinese manufacturing goods seem to be close substitutes of those produced in Brazil. In the 2000–2012 period, the share of informal workers in manufacturing declined from 27 percent to approximately 15 percent in 2012. This is in stark contrast to the increase in informality of 8 percentage points that took place between 1989 and 2000 during the 1990s trade liberalization in Brazil (Paz 2014). In view of these disparate responses to increased import competition, the China shock seems to be a good candidate to explain the distinct response of informality to trade liberalization in the 2000s.

Unfortunately, there are few studies focusing on the impacts of the China shock on formal and informal workers in manufacturing, especially for Latin America. This study represents a step towards, filling this gap by studying how imports from China and from the rest of the world (ROW) affected Brazil's manufacturing labor market in the 2000–2012 period. It combines two strands of the literature. The first strand investigates the effects of international trade on informality and uncovers mixed results, e.g., the cross-country studies by Dávalos (2019) for Latin American countries, Aleman-Castilla (2020) for Mexico, Paz (2014) and Almeida et al. (2022) for Brazil. The other strand of the literature looks at the heterogeneous effect of trade according to the source of the imports. Pierola and Sanchez-Navarro (2019) for Peru and Paz (2018, 2019b) for Brazil find that Chinese imports did impact differently in labor markets in Latin America when compared with imports from elsewhere.

This study conducts a rigorous empirical analysis to examine how the informality of manufacturing employment in Brazil was affected by the changes in the industry-level import penetration of goods sourced in China and in the ROW. This analysis utilizes household-level data from the Brazilian demographic census and the Pesquisa Nacional por Amostra de Domicilios (PNAD), which comprise detailed employment and demographic information about both formal and informal workers. The methodology employed here follows Paz (2014) by estimating the effects of Chinese and ROW import penetrations on the likelihood of holding an informal job via an IV Probit model, and then employing a switching regression model for estimation of the effects of these types of import penetration on both average formal and informal wages. This specification has the merit of addressing both worker self-selection into formal and informal jobs, and the potential endogeneity of trade policy.

The empirical results indicate that greater industry-level Chinese and ROW import penetrations increase informal job likelihood, albeit at different intensities. Furthermore, these effects are modulated by the unskilled labor intensity of the industry and by the degree of industrialization of the specific Brazilian state. Indeed, in unskilled-labor intensive industries, ROW import penetration has a negative effect on informality likelihood and the Chinese import penetration has no effect, while both types of import penetration have positive impact on the remaining industries. An increase in Chinese import penetration reduces informality whereas ROW import penetration increases it in manufacturing states. In contrast, both types of import penetration positively affect informality in non-manufacturing states.

The effects of import competition on average formal and informal wages are more complex. A larger ROW import penetration decreases the average formal wage, except for in manufacturing states or in non-unskilled-labor intensive industries. In contrast, an increase in Chinese import penetration raises the average formal wage, except in manufacturing states. Both Chinese and the ROW import penetration have positive effects on the average informal wage, but of different magnitudes. Additionally, both have negative effects on the informal wages of workers located in non-manufacturing states. Taken together, these results suggest that Chinese and ROW import effects on Brazilian manufacturing workers are different, and such heterogeneity depends on the location and unskilled labor intensity of the industry.

The remainder of this paper is organized as follows. The next section provides an overview of trade related policies in Brazil since the 1990s, and describes the data used and their descriptive statistics. The theoretical framework and its corresponding empirical methodology are discussed in Section 3. The empirical estimates are displayed and analyzed in Section 4. Finally, Section 5 presents the conclusions.

2. Policy Background, Data, and Theoretical Framework

This section provides a brief overview of the changes in trade-related policies that have taken place in Brazil since the 1990s. This is followed by a description of the original sources of each component of the dataset employed here and the respective cleaning and assembly procedure used. Next, descriptive statistics on the evolution of the import competition in the Brazilian economy and its labor market outcomes during the 2000–2012 period are presented. Finally, the theoretical framework that motivates the analysis is laid out and its testable hypotheses are discussed.

2.1. Policy Background

Inaugurated in 1990, the Collor de Mello administration implemented a series of economic reforms aiming to reintegrate Brazil into world markets. Such reforms eliminated hurdles in the foreign currency market and implemented changes in the trade policy to reduce protection levels (Baumann 2001). In fact, the new president suddenly and drastically reduced the non-tariff measures of protection (NTMs) and also scheduled nominal tariff cuts that were heterogeneous across industries, to be implemented between 1990 and 1994 (Kume et al. 2003, 2008). The protection of the manufacturing sector declined substantially from a 40 percent average tariff in 1989 to a 17 percent average tariff in 2000. Accordingly, manufacturing imports grew by more than 200 percent between 1990 and 2000, and the import penetration in manufacturing almost tripled, growing from its initial level of 5.7 percent in 1990 to 14 percent in 2000.

The trade protection reforms implemented in the 2000s by the da Silva and the Roussef administrations were considerably different than the reforms of the 1990s (Paz 2018). In 2004, the da Silva administration granted market economy status to China in November of 2004. This decision came in the aftermath of China's accession to the WTO in 2001 (Chandra 2014). It conceded most favored nation tariffs to Chinese imports and reduced the ability of the Brazilian government to impose safeguards countervailing duties and anti-dumping against Chinese exporters. These reforms resulted in a greater openness of the Brazilian economy that is illustrated by an increase in the overall manufacturing import penetration from 14% in 2000 to 18% in 2012.

2.2. Data Description

The database assembled for this study comprises information on international trade flows, on Brazilian national accounts, and on household surveys. The bilateral international trade data are downloaded from the Comtrade system (United Nations 2003) for the period between 2000 and 2012 using the six-digit 1996 version of the harmonized system. These are used to build industry-level Brazilian imports from China and from the remaining countries of the world (hereafter called ROW) series, and also the excluded instruments, as discussed in the next section.

The Brazilian national accounts data (IBGE 2015, 2016) provide information on total output level, on employment level, imports and exports at IBGE's level 56 industry classification. The worker-level data come from the PNAD-Pesquisa Nacional por Amostra de Domicilios (Brazilian household survey) and from the Brazilian demographic censuses of 2000 and 2010, as the PNAD household surveys do not take place in census years. These surveys ask similar questions about workers' observable characteristics such as earnings, hours worked in a week, job formality status, industry affiliation, education, gender, age, marital status, race, and Brazilian state of residence. The period under analysis ends in 2012 because in 2013 a major change in the social security contribution incidence was enacted by Federal Law 12546. PNAD's methodology substantially also changed in the 2015.

This study only considers employed workers. Employers, self-employed, and unemployed people are excluded from the analysis. Moreover, an informal job is defined as the employment relationship in which the employer does not comply with the social security contributions, as in Paz (2014). The different industry classifications used in the original data were harmonized by means of correspondence tables from the CONCLA-IBGE website (https://concla.ibge.gov.br/, accessed on 19 April 2021). The classification used by the National Accounts data is the most cursory in this study. Hence, it dictated the industry classification used here, which consists of a modified version of the Nível 56 classification with 26 manufacturing industries.

2.3. Raw Data Patterns

This subsection starts with an overview of the trade relationship between China and Brazil and of the import competition experienced by the Brazilian manufacturing sector. In 2000, Chinese imports made up 2.7 percent of Brazilian imports, which ranked it as the tenth largest exporter to Brazil. Yet, these figures were radically different in 2012, when Chinese imports amounted to approximately 20 percent of Brazilian imports, of which more than 90 percent are manufactured goods. Moreover, other labor abundant countries like India, Indonesia, and Vietnam accounted for less than one percent of Brazilian imports in this period (Paz 2019a). Figure 1 shows that imports from China and from high-income countries account for most Brazilian imports. Thus, ROW imports are mainly driven by imports from high income countries. Figure 1 also displays a growing Chinese share of Brazilian imports and a shrinkage of the high-income countries' share. Indeed, both the Chinese and the ROW volume of imports grew over time, albeit at a different pace. Hence, the Brazilian experience in the 2000s cannot be summarized merely into a case of substitution of suppliers.

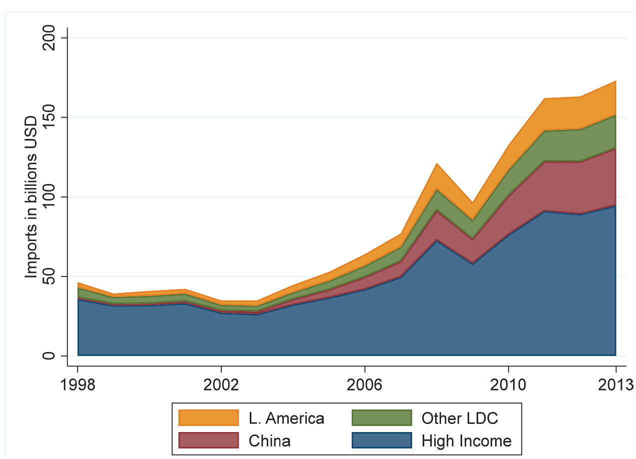

Figure 1. Brazilian manufacturing imports by source. Notes: L. America–Latin America. Other LDC–developing countries other than China and those in Latin America.

The import competition measure of Brazilian firms used in this study is the industry-level import penetration. This is the ratio between imports and the apparent consumption (production plus imports minus exports). In contrast with tariffs, the import penetration also captures the effects of NTMs, such as import licenses, quotas, and anti-dumping duties. Moreover, the import tariff between 2000 and 2012 shows little variability across industries and over time, despite the large variations in import volumes and in import penetration (Paz 2018). As a result, import tariffs are not recommended for the analysis carried out in this study.

Turning to the descriptive statistics at the industry level, Table 1 reports the 2000- and the 2012-level, the average, and the standard deviation of ROW and Chinese import penetrations. We can see that 19 out of 26 industries exhibited an increase in import penetration, and in most of these cases the increase was in excess of 20 percent. Most importantly, these industries employ more than half of the workers in manufacturing. Additionally, the Chinese import penetration grew in 24 out of 26 industries. Although the average Chinese import penetration is smaller than the ROW import penetration, the former has a significantly larger coefficient of variation due to the increase in Chinese participation in the Brazilian imports in this period. Together with Figure 1, these statistics indicate that this growth in Chinese import penetration was not simply a case of substitution of ROW imports. Table 1 also displays descriptive statistics for industry-level informality share. They show shares below 5 percent in industries such as automobiles, steel, and biofuels, while industries such as apparel and wood products exhibited an informality share close to 30 percent.

Table 1. Industry-level trade exposure and labor market characteristics of manufacturing industries in Brazil during 2000–2012.

Industry/Year	**ROW Import Penetration (%)**				**Chinese Import Penetration (%)**				**Informal Share (%)**			
	2000	**2012**	**Mean**	**S. dv.**	**2000**	**2012**	**Mean**	**S. dv.**	**2000**	**2012**	**Mean**	**S. dv.**
Food and Beverages	4.33	3.53	3.87	0.42	0.05	0.65	0.21	0.17	20.40	21.05	19.45	39.58
Tobacco	27.61	27.38	26.90	1.49	0.22	2.19	1.48	1.15	6.96	11.51	10.69	30.91
Textiles	9.45	9.20	8.88	0.81	0.23	6.44	2.38	2.26	17.78	46.67	25.16	43.39
Apparel	2.90	6.31	3.98	1.00	0.17	5.45	2.20	1.78	22.28	53.31	29.26	45.50
Footwear and leather products	5.67	5.67	5.08	0.48	0.56	2.66	1.49	0.78	20.16	19.59	17.12	37.67
Wood products	2.54	1.34	2.00	0.63	0.07	0.47	0.18	0.12	25.00	51.95	28.20	45.00
Paper products	10.59	7.45	8.97	1.58	0.01	0.85	0.21	0.23	11.25	10.34	9.51	29.33
Printing and Publishing	0.62	1.50	0.97	1.01	0.00	0.15	0.04	0.05	25.81	24.86	25.64	43.66
Petroleum refining	0.00	4.33	9.65	1.84	0.00	0.00	0.18	0.12	6.79	12.86	7.21	25.87
Biofuel	20.64	26.63	0.59	1.09	0.43	1.65	0.00	0.00	7.36	1.17	3.31	17.89
Pharmaceutical products	19.62	26.01	23.17	2.66	0.00	0.97	0.93	0.47	8.90	6.08	8.05	27.21
Cleaning products	7.02	6.60	20.98	1.84	0.01	0.21	0.28	0.29	15.63	8.63	15.41	36.11
Paint and varnish	9.56	10.79	6.10	1.04	0.14	2.55	0.06	0.05	9.13	5.83	9.00	28.63
Rubber and plastic products	6.83	9.64	9.77	0.43	0.10	2.88	0.88	0.78	11.52	12.75	9.91	29.88
Steel	22.54	24.80	8.07	1.61	0.27	1.58	1.59	1.45	6.18	5.47	4.83	21.43
Non-ferrous metals	7.61	8.20	22.78	1.97	0.30	3.05	0.73	0.48	12.28	5.78	10.29	30.39
Metal products	26.11	21.63	7.42	0.28	0.26	4.63	1.25	0.84	17.42	31.57	16.95	37.52
Machinery and equipment	1.36	0.91	22.68	3.14	0.24	2.43	1.97	1.66	13.76	12.42	12.60	33.18
Appliances	24.71	22.11	0.99	0.32	0.02	1.30	0.99	0.77	6.57	5.06	6.06	23.87
Auto Parts	55.64	32.86	21.02	3.97	0.26	3.61	0.34	0.34	11.01	6.53	8.49	27.88
Other transp. Equipment	4.90	4.60	39.86	12.09	0.12	2.38	0.96	0.71	21.33	14.60	14.78	35.49
Non-metallic min. and products	45.08	32.43	4.47	0.56	2.07	18.22	0.87	0.77	30.31	24.00	29.79	45.73
Office, electrical, electronic, optical, precision, and communication equipment	13.73	15.37	39.94	5.83	0.00	0.39	8.36	6.72	13.97	13.89	12.95	33.58
Automobiles, trucks, and buses	26.69	30.16	12.93	1.55	0.59	3.45	0.18	0.19	5.20	3.60	4.18	20.02
Other chemical products	4.84	4.03	26.75	1.38	1.27	5.31	1.45	0.96	13.44	11.02	14.91	35.62
Furniture and other products	4.33	3.53	4.04	0.58	0.05	0.65	3.28	1.54	28.04	44.86	24.41	42.96

Notes: Informal workers are those without social security coverage. Number of observations is 312. Household survey weights used for informal share.

Table 2 presents industry-level average characteristics of manufacturing workers according to their formality status. The hourly wage consists of the monthly wage divided by 4.3 times the number of hours worked in a week. The inflation adjustment is conducted according to Corseuil and Foguel (2002) using inflating factors from IPEADATA (2017). The natural logarithm of the real hourly wage also exhibits heterogeneity across industries, being larger in skilled-labor intensive industries. These figures also indicate that formal workers earn substantially higher hourly wages and are more likely to be male. Formal workers are slightly older than informal workers. The industry-level average education level shows substantial cross-industry variation, which tend to be higher in skilled-labor-intensive industries such as pharmaceutical products. Even though formal workers are more educated on average, the difference in the average of years of schooling range from half year in footwear to three years in other chemical products. This study now turns to the presentation of the theoretical framework that motivates the description of the empirical methodology used here.

Table 2. Formal and Informal workers' average characteristics at industry level.

Industry/Year	Log (Hourly Wage)		Age		Female Share (%)		Years of Schooling	
	Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal
Food and Beverages	1.468	0.917	33.085	32.326	0.266	0.370	7.062	5.737
Tobacco	1.611	0.933	36.011	33.143	0.393	0.525	7.638	6.226
Textiles	1.338	0.573	33.154	33.985	0.499	0.717	7.309	6.139
Apparel	1.289	1.017	36.285	35.975	0.837	0.843	7.094	6.840
Footwear and leather products	1.332	1.076	30.934	30.661	0.477	0.514	7.260	6.766
Wood products	1.340	1.042	34.910	31.956	0.168	0.129	5.687	5.077
Paper products	1.796	1.312	32.940	30.561	0.201	0.317	8.548	7.594
Printing and Publishing	1.866	1.467	32.534	30.062	0.277	0.297	9.361	8.914
Petroleum refining	2.446	2.204	36.709	35.769	0.134	0.193	10.287	10.262
Biofuel	1.625	1.270	33.767	33.830	0.109	0.085	6.776	5.702
Pharmaceutical products	2.131	1.743	32.530	31.317	0.437	0.437	10.864	10.066
Cleaning products	1.750	1.375	32.278	31.475	0.409	0.475	9.129	8.016
Paint, varnish, and laqueur	1.957	1.591	32.758	31.393	0.178	0.202	9.203	8.645
Rubber and plastic products	1.722	1.346	32.162	30.531	0.249	0.292	8.302	7.498
Steel	2.080	1.616	34.845	32.308	0.079	0.162	9.202	8.198
Non-ferrous metals	1.892	1.368	33.914	31.185	0.124	0.147	8.856	7.558
Metal products	1.816	1.302	34.089	31.526	0.105	0.085	7.799	6.902
Machinery and equipment	1.920	1.544	34.092	32.706	0.108	0.127	8.668	7.687
Appliances	1.832	1.396	31.336	29.532	0.276	0.276	9.642	8.454
Auto Parts	1.895	1.463	32.339	30.278	0.170	0.177	8.938	7.919
Other transportation equipment	1.995	1.496	34.004	33.641	0.111	0.108	9.038	7.167
Non-metallic minerals and products	1.476	0.941	33.044	29.233	0.112	0.102	6.454	5.105
Office, electrical, electronic, optical, precision, and communication equipment	1.860	1.538	31.589	30.646	0.315	0.255	9.629	8.894
Automobiles, trucks, and buses	2.298	1.773	33.389	30.367	0.128	0.225	10.015	8.932
Other chemical products	1.991	1.064	34.097	32.576	0.207	0.217	9.285	5.809
Furniture and other products	1.601	1.087	33.541	31.521	0.258	0.239	7.615	6.811

Notes: Number of observations is 669,966. Informal workers are those without social security coverage. Household survey weights used.

2.4. Theoretical Framework

The analysis of the effects of the changes in the trade openness of the Brazilian economy on the formal and informal labor market outcomes is guided by the theoretical model and empirical framework developed by Paz (2014)[1]. This is a heterogeneous firm model based on Melitz (2003), in which there is monopolistic competition and firms are heterogeneous in two dimensions: productivity and wages. Firms decide whether to enter the market or not considering their intrinsic characteristics and the current level of market competitiveness. If they decide to do so, what type of labor do firms offer to workers: formal or informal? By hiring formal workers, firms are subject to a fixed cost per worker to comply with labor regulations and a variable cost related to social security contributions. An informal labor contract has no such costs, but firms incur the risk of being audited (with positive probability) and fined if they are caught employing informal workers. In equilibrium, larger (more productive and higher wage) firms will hire only formal workers. In contrast, smaller (less productive and lower wage) firms will hire informal workers.

An increase in import penetration fosters competitiveness in domestic markets. This makes the smallest firms—who typically employ informal workers and pay lower wages—experience negative profits and exit the market. This leads to a reduction in the employment of informal workers and an increase in the average informal average wage. Yet the firms that were previously indifferent about hiring either formal or informal workers switch from formal to informal employment in response to this increased competitiveness. This raises the level of informal employment. Such a change may increase or decrease average formal and informal wages depending on the specific firm's joint distribution of productivity and wages. In sum, an increase in import penetration leads to ex ante ambiguous effects of import penetration on informality and on average formal and informal wages.

An emerging literature has uncovered evidence of more nuanced effects of import competition depending on the source country of the imports and on the impacted industry characteristics. Facchini et al. (2010) found that the elasticity of substitution between Chinese and Brazilian manufacturing products was higher than that between Brazilian and high-income country products. This finding implies that imports from China are closer substitutes for Brazilian made goods. As a result, Chinese imports exert a stronger competitive pressure on Brazilian producers than imports from the rest of the world (especially those from high-income countries). This leads to the first testable prediction:

Hypothesis 1. *The effect of Chinese import penetration has a greater magnitude than that of ROW import penetration.*

Moreira (2006) points out that China's accession to the WTO exposed its trade partners to an unparalleled trade shock due to the uniqueness of the labor abundance of the Chinese economy relative to almost all countries in the world (Brazil included). Hence, according to the Heckscher-Ohlin model, unskilled-labor intensive imports from China apply a stronger competitive pressure than other imports. This is supported by Ashournia et al. (2014), who found that low-skill intensive firms in Denmark were heavily impacted by Chinese imports. Conversely, the competitiveness of ROW imports is milder in unskilled-labor intensive industries relative to that in other industries. This leads to the second testable prediction.

Hypothesis 2. *Increased Chinese import penetration has a larger effect on unskilled-labor intensive industries relative to its impact on other industries. Conversely, increased ROW import penetration affects other industries more than unskilled-labor intensive industries.*

Another factor that can modulate the effects of import competition on labor market outcomes is the location of manufacturing activity. Indeed, this exhibits a significant spatial heterogeneity in Brazil, as 7 out of 26 states account for approximately 80 percent of all manufacturing activity (Paz 2019b). These seven states are those that make up the Southeastern and Southern regions of Brazil. According to Kapri and Paz (2019), such a spatial concentration makes manufacturing workers experience a different exposure to trade-induced shocks depending on their location. Thus, the effects of import penetration are expected to be different on workers in states with a large manufacturing sector. This leads to the third and last testable hypothesis:

Hypothesis 3. *Increased industry-level import penetration has a distinct impact on workers in manufacturing states relative to those in non-manufacturing states.*

3. Empirical Methodology

The empirical methodology used to assess the testable hypotheses exploits the industry-level variation in competition induced by imports from China and from the ROW on the worker-level informality likelihood and on average formal and informal wages. Paz (2014) points out that decisions about the type of labor contract used (formal or informal) and the wage paid are made simultaneously. Hence, overlooking this simultaneity leads to biased estimates of the effects of import competition on wages. This was found to be the case in Brazil during the 1990s by Paz (2014) and in Peru during the 2000s by Pierola and Sanchez-Navarro (2019). Paz (2014) proposes addressing this simultaneity by means of the two-step switching regression framework from Maddala (1983). The first step in this framework is the regime selection equation that models the choice of the employment type using a Probit discrete choice specification, as depicted by Equation (1).

$$informal_{ijst} = \alpha + \beta_1 IP_{j,\ t-1}^{China} + \beta_2 IP_{j,\ t-1}^{ROW} + \Psi X_{ijst} + otherformal_{ijst} + \gamma_j + \theta_s + \delta_t + u_{ijst} \quad (1)$$

$$informal_{ijst}^{*} = 1\ if\ informal_{ijst} > 0,\ informal_{ijst}^{*} = 0\ otherwise.$$

where $informal_{ijst}^{*}$ is an indicator dependent variable that is "1" if worker i in industry j in state s and year t holds an informal job, and is "0" if worker holds a formal job; and $informal_{ijst}$ is the respective latent variable. $IP_{j,\ t-1}^{China}$ is the Chinese import penetration. $IP_{j,\ t-1}^{ROW}$ is the ROW import penetration. X_{ijst} is a vector of worker i's observable characteristics, namely age, age squared, number of years of education, and indicators for gender, race, marital status, high school degree, and college degree. γ_j are industry fixed effects that control for the industry-specific and time-invariant characteristics. θ_s are state fixed effects which capture state specific and time-invariant characteristics, such as being landlocked. δ_t are year effects that account for time-varying factors that affect industries equally, such as business cycles. u_{jt} is the error term.

For identification purposes, the switching-regression framework requires the selection equation to contain at least one relevant variable that does not affect the wage earned by the worker. Following Paz (2014), the variable fulfilling this role is $otherformal_{ijst}$. This is a dummy variable that is "1" if another person in the household has a formal job, and "0" otherwise. The idea is that the tradeoff between a formal and an informal job experienced by a worker is affected by another household member having a formal job, since informal labor contracts are used by both firms and workers to evade taxation. In this vein, having another household with a formal job may increase the likelihood of detection of income tax evasion, and this reduces the incentives for choosing an informal job. This suggests a negative coefficient for this indicator variable in the selection equation.

The second step of the switching regression framework is to model the average wage through a Mincer-type wage (y_{ijst}) equation for each type of job using the inverse Mills ratio to control for the worker self-selection into that regime. The inverse Mills ratio ($\Lambda(z)$) is defined as $\Lambda(z) \equiv \phi(z)/\Phi(z)$, where $\phi(z)$ is the standard normal distribution function and $\Phi(z)$ is the standard normal cumulative distribution function. It is calculated using the predicted values of the informality likelihood ($\widehat{informal}_{ijst}$). These wage equations are depicted by Equation (2) for the formal worker and by Equation (3) for the informal worker.

$$y_{ijst}^{for} = \alpha^{for} + \beta_1^{for} IP_{j,\ t-1}^{China} + \beta_2^{for} IP_{j,\ t-1}^{ROW} + \Psi^{for} X_{ijst} + \gamma_j^{for} + \theta_s^{for} + \delta_t^{for} + \sigma^{for} \Lambda\left(-\widehat{informal}_{ijst}\right) + v_{ijst}^{for} \quad (2)$$

$$y_{ijst}^{inf} = \alpha^{inf} + \beta_1^{inf} IP_{j,\ t-1}^{China} + \beta_2^{inf} IP_{j,\ t-1}^{ROW} + \Psi^{inf} X_{ijst} + \gamma_j^{inf} + \theta_s^{inf} + \delta_t^{inf} + \sigma^{inf} \Lambda\left(\widehat{informal}_{ijst}\right) + v_{ijst}^{inf} \quad (3)$$

The selection into an informal or a formal labor contract takes place as a non-zero correlation among the error terms u_{ijst}, v_{ijst}^{for}, and v_{ijst}^{inf}. Should this be the case, the estimated coefficients of the inverse Mills ratios will be statistically significant. Accordingly, the omission of these terms will then lead to biased estimates of the effects of the import penetrations on wages. Note that the standard errors for Equations (2) and (3) are estimated by means of a 500-repetition bootstrap using household survey weights because the inverse Mills ratio is a generated regressor.

The calculation of the marginal effects of the import penetrations on wages ought to consider the effects of the import penetrations on the regime selection. This means that the effects of the import penetrations on wages will, in a non-trivial way, depend on the workers characteristics via the selection equation. The marginal effect of Chinese import penetration on the formal and the informal wages are given by Equations (4) and (5), respectively. The marginal effects for ROW import penetration are calculated in a similar fashion.

$$\frac{\partial wage_{ijst}^{for}}{\partial IP_{j,\,t-1}^{China}} = \widehat{\beta_1^{for}} + \widehat{\sigma^{for}}\Lambda\left(-\widehat{informal_{ijst}}\right)\widehat{\beta_1}\left[\widehat{informal_{ijst}} - \Lambda\left(-\widehat{informal_{ijst}}\right)\right] \quad (4)$$

$$\frac{\partial wage_{ijst}^{inf}}{\partial IP_{j,\,t-1}^{China}} = \widehat{\beta_1^{inf}} - \widehat{\sigma^{inf}}\Lambda\left(\widehat{informal_{ijst}}\right)\widehat{\beta_1}\left[\widehat{informal_{ijst}} + \Lambda\left(\widehat{informal_{ijst}}\right)\right] \quad (5)$$

To assess the predictions of Hypothesis 2, Equations (1)–(3) are augmented to include new regressors that are interactions between the import penetration measures and LI_j, which is an indicator variable that is "1" for unskilled-labor intensive industries, and "0" otherwise. As in Paz and Ssozi (2021), the unskilled-labor intensive industries are the seven industries with the lowest average of year of schooling in 2000, namely food and beverages, textiles, apparel, footwear and leather products, wood products, non-metallic minerals and products, and furniture and other products. Similarly, Hypothesis 3 is evaluated through the addition to Equations (1)–(3) of interaction terms between the import penetrations and an indicator variable ($Manufacturing\ state_s$) that is "1" if state s is in the South or Southeast of Brazil, and "0" otherwise.

There are some concerns about the empirical methodology that deserve careful discussion. First, Brazilian producers may react slowly to changes in market conditions. This can be addressed by employing lagged import penetrations. The second concern is that both the import penetration measures and wages are simultaneously determined because the latter is part of the value added that is used in the calculation of the import penetration. This concern can also be alleviated by using the first lag of the import penetration measures.

The third concern is the omitted variable bias. More precisely, this could be the case of omitted time variant industry-specific shocks that affect both import penetrations and outcomes, and this biases the estimates. Examples of such omitted variables are demand or supply shocks in the Brazilian economy. For instance, suppose that a larger than expected import penetration increase leads the Brazilian government to impose higher import tariffs, safeguards, or countervailing duties. This behavior can be seen in the number of antidumping procedures initiated in Brazil that reached almost 100 cases in the 2000s, of which a quarter were against Chinese producers (cf. WTO Antidumping Gateway 2016). These product-level non-tariff protection measures cannot be accounted for in the empirical specification, and thus can bias the estimates. The use of an instrumental variable (IV) approach can address this concern, as explained next.

The excluded instruments used in the IV strategy are based upon Iacovone et al.'s (2013) idea of using supply-driven shocks as an instrument for import penetration.[2] Thus, the excluded instrument for Chinese import penetration is the Chinese share of imports in third countries. The third countries considered have very small trade ties with Brazil and are located in Latin America, namely Mexico, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Guyana, Jamaica, Nicaragua, Panama, and Peru. The correlation between the Chinese import penetration and Chinese share of imports in third countries is 0.574. The same endogeneity concern applies to the ROW import penetration; therefore, an additional excluded instrument is required. In a similar fashion, this is the high-income countries' share in the imports of the above mentioned third countries. These high-income countries are Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Czechia, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, USA, and United Kingdom. The simple correlation between the ROW import penetration and the high-income share of imports is 0.316. In the specifications that include interaction terms between the import penetrations and indicator variables, the additional excluded instruments are built by interacting the excluded instruments mentioned above with the respective indicator variable.

Table 3 reports the estimated coefficients of the regressions of the endogenous explanatory variables (Chinese and ROW import penetrations) on the excluded instruments and other control variables used in the selection Equation (1). We can see that the Chinese share of imports in Latin America is statistically significant when the Chinese import penetration is the dependent variable, column (1), while the other excluded instrument is not significant. Similarly, the high-income countries' share of imports in Latin America is significant in the specification for the ROW import penetration in column (2), whereas the Chinese share of imports in Latin America is not significant in this regression. None of the workers' characteristics are statistically significant. These estimates show that the excluded instruments have predictive power over the endogenous regressors. The *F*-statistic of these regressions in Table 3 is above 200, which suggests that a weak instrument is not the problem.

Table 3. First-stage OLS regressions of the endogenous regressors.

Regressors/Dependent Variable	(1) Chinese Import Penetration$_{t-1}$	(2) ROW Import Penetration$_{t-1}$	(3) Chinese Import Penetration$_{t-1}$	(4) ROW Import Penetration$_{t-1}$
Chinese share of imports in LA$_{t-1}$	9.831 **	−8.861	9.780 **	−8.794
	(4.452)	(6.576)	(4.356)	(6.376)
High Income countries share of imports in LA$_{t-1}$	−4.108	−10.112 **	−4.238	−9.743 **
	(2.466)	(5.027)	(2.523)	(4.495)
Age	−0.001	0.006 *	−0.001	0.005
	(0.002)	(0.003)	(0.002)	(0.003)
Age^2	0.000	−0.000 *	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Female	−0.009	−0.008	−0.010	−0.003
	(0.011)	(0.014)	(0.009)	(0.011)
Married	−0.005	−0.056 *	−0.005	−0.055 *
	(0.019)	(0.032)	(0.018)	(0.031)
Black	0.000	0.009	0.004	0.002
	(0.005)	(0.013)	(0.008)	(0.014)
Asian	−0.008	−0.013	−0.012	0.005
	(0.018)	(0.086)	(0.015)	(0.074)
Years of schooling	−0.003	0.007	−0.002	0.004
	(0.003)	(0.006)	(0.002)	(0.004)
High school	0.004	−0.023	−0.001	−0.005
	(0.005)	(0.019)	(0.006)	(0.015)
College	0.020	−0.040	0.017	−0.033
	(0.019)	(0.040)	(0.016)	(0.035)
Other household has a formal job$_{ijst}$	0.009	−0.008	0.007	−0.004
	(0.007)	(0.011)	(0.007)	(0.010)
Industry, state, and year fixed effects	Yes	Yes	No	No
Industry and state-year fixed effects	No	No	Yes	Yes
F-statistic	392.14	570.5	201.36	335.88

Number of observations is 671,134. **, and * indicate statistical significance at the 5%, and 10% levels, respectively. Standard errors clustered at industry level are reported in parenthesis. Household survey weights used.

4. Results and Discussion

This section begins with the estimates of the selection specification based on equation (1), which is the first step of the switching regression framework. Table 4 displays the estimated coefficients obtained using Probit in columns (1)–(3) and those obtained using IVProbit in columns (4)–(6). The worker's characteristics' estimated coefficients do not vary across specifications. They indicate that older, female, married, and Asian workers are more likely to hold an informal job, while black workers are less likely. The likelihood of having an informal job is smaller for those with a high school degree and with a larger number of years of schooling. Interestingly, the college degree indicator is not statistically significant at the 5 percent level in any specification. The other household member has a formal job indicator estimated coefficient which is negative and statistically significant in all specifications of Table 4, with a similar magnitude across specifications.

Table 4. Worker-level Probit (1–3) and IVProbit (4–6) estimates of the effects of industry-level import penetration on the informal status indicator, using Equation (1).

Regressors	(1)	(2)	(3)	(4)	(5)	(6)
Chinese import penetration$_{t-1}$	−0.010	−0.003	0.000	0.002	0.085 **	0.031 ***
	(0.014)	(0.010)	(0.015)	(0.007)	(0.034)	(0.008)
ROW import penetration$_{t-1}$	−0.006	−0.002	−0.007	0.017	0.069 ***	0.045 ***
	(0.005)	(0.004)	(0.005)	(0.013)	(0.026)	(0.014)
Chinese imp. penet.$_{t-1}$ × L. int.$_j$		−0.015			−0.083 ***	
		(0.017)			(0.030)	
ROW imp. penet.$_{t-1}$ × L. int.$_j$		−0.004			−0.113 *	
		(0.024)			(0.065)	
Chinese imp. penet.$_{t-1}$ × Manuf.$_s$			−0.014			−0.039 ***
			(0.010)			(0.006)
ROW imp. penet.$_{t-1}$ × Manuf.$_s$			0.001			−0.035 ***
			(0.004)			(0.003)
Age	−0.120 ***	−0.120 ***	−0.120 ***	−0.120 ***	−0.119 ***	−0.119 ***
	(0.008)	(0.008)	(0.008)	(0.002)	(0.002)	(0.002)
Age^2	0.002 ***	0.002 ***	0.002 ***	0.002 ***	0.002 ***	0.002 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	0.463 ***	0.463 ***	0.463 ***	0.463 ***	0.459 ***	0.456 ***
	(0.091)	(0.091)	(0.091)	(0.008)	(0.008)	(0.008)
Married	0.133 ***	0.132 ***	0.133 ***	0.135 ***	0.131 ***	0.135 ***
	(0.035)	(0.036)	(0.035)	(0.008)	(0.008)	(0.008)
Black	−0.079 ***	−0.079 ***	−0.079 ***	−0.079 ***	−0.079 ***	−0.080 ***
	(0.023)	(0.023)	(0.023)	(0.013)	(0.013)	(0.013)
Asian	0.196 ***	0.196 ***	0.197 ***	0.197 ***	0.196 ***	0.211 ***
	(0.049)	(0.049)	(0.049)	(0.054)	(0.053)	(0.054)
Years of schooling	−0.028 ***	−0.028 ***	−0.028 ***	−0.028 ***	−0.028 ***	−0.028 ***
	(0.006)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)
High school	−0.176 ***	−0.175 ***	−0.175 ***	−0.175 ***	−0.173 ***	−0.172 ***
	(0.018)	(0.018)	(0.018)	(0.011)	(0.011)	(0.011)
College	0.018	0.017	0.019	0.018	0.018	0.036 *
	(0.033)	(0.033)	(0.034)	(0.020)	(0.020)	(0.020)
Other household has a formal job$_{ijst}$	−0.345 ***	−0.345 ***	−0.345 ***	−0.345 ***	−0.343 ***	−0.344 ***
	(0.020)	(0.020)	(0.020)	(0.007)	(0.007)	(0.007)
Exogeneity test				7.724 **	36.47 ***	126.6 ***
				(0.021)	(0.000)	(0.000)

Notes: Number of observations is 671,134. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Standard errors clustered at the industry level. Sample weights from PNAD/Census used. State, industry, and year fixed effects included in the estimated model. The excluded instruments are the Latin American countries' Chinese share of imports, the Latin American countries' high-income countries share of imports, and their interactions.

The Probit specifications do not display statistically significant coefficients for either the import penetrations or for their interactions. Additionally, notice that the null hypothesis of exogeneity of the endogenous regressors is rejected at the 5 percent level in all IVProbit specifications. This suggests that the focus of the analysis should be on the IV Probit specifications since they account for the endogeneity of the import penetrations and their interactions.

The IVProbit specification of column (4) displays no statistically significant coefficient of the Chinese and the ROW import penetration, therefore this result renders no support for Hypothesis 1. The specification in column (5) is designed to assess Hypothesis 2. It presents the Chinese and the ROW import penetrations with positive and statistically significant coefficients. The interaction terms of the import penetrations with the unskilled-labor intensive indicator are negative for both import penetrations, and statistically significant at the 5 percent level for the Chinese and at the 10 percent level for the ROW import penetration. These coefficients imply that an increase in any import penetration raises informality in non-unskilled-labor intensive industries. For unskilled-labor intensive industries, Chinese import penetration has no effect on informality, while ROW import penetration reduces informality in unskilled-labor intensive industries. These results do support that the Chinese and the ROW import penetrations have different impacts on informality (Hypothesis 1), and such impacts also depend on the unskilled-labor intensity of the industry. These results are not in line with Hypothesis 2. The estimates to assess Hypothesis 3 are in column (6), where both import penetrations are positive and significant as before. This means that increased import penetration leads to more informality in non-manufacturing states. The interaction terms with manufacturing state indicators are negative and significant. Accordingly, for manufacturing states Chinese import penetration reduces (albeit with a smaller magnitude) informality, whereas ROW import penetration has a positive effect. These results do support Hypothesis 3.

The results for ROW imports are somewhat distinct from those obtained by Paz (2014) using Brazilian data for the 1990s. His results indicate that greater imports—most from high income countries—increased the manufacturing informality share. On the one hand, Pierola and Sanchez-Navarro (2019) had IVProbit results of no effect regarding Chinese imports on informality for Peru in the 2000s, in line with this study findings. On the other hand, their findings that Chinese imports do increase informality among unskilled workers are at odds with the results of Table 4.

Table 5 reports the OLS and IV estimates employing equation (2) for the formal workers' average wage in columns (1)–(3) and (4)–(6), respectively. The selection terms (the inverse Mills ratios) were computed using the IVProbit estimates of the specification that contained the same explanatory variables (except for the other formal indicator). The estimated coefficients of the selection terms are statistically significant in all columns. These coefficients are positive except in column (5). This is evidence that selection effects are taking place and should not be overlooked. The estimated coefficients for workers' characteristics are stable across specifications. Formal workers' wages are increasing with age and number of years of schooling, with considerable premium for high school and college degrees. Females and blacks earn a lower wage on average. Additionally, married and Asian workers have higher average wages. Notice that the OLS and the IV estimates for some variables are substantially different. In fact, the null of exogeneity of the import penetrations (and their interaction terms) is rejected at the 5 percent level in every column. Hence, the discussion will focus on the IV specifications results.

In column (4), the estimated coefficients of the Chinese import penetration are positive and statistically significant at the 5 percent level of confidence and the coefficient for the ROW import penetration is zero. As discussed in the previous section, the estimated coefficients of the import penetrations and their interactions are different from their marginal effects on the average wage due to the selection effect. The marginal effect is calculated for the average formal and informal worker, i.e., the estimated selection equation fitted at the average value of its regressors, as shown in Equation (4). The effects on the average formal wage of a percentage point increase in the Chinese import penetration is a 1.09 percent increase and for the ROW import penetration a 0.06 percent increase, respectively. These figures suggest that the Chinese and ROW import penetrations had a distinct impact on formal workers' average wage, which supports Hypothesis 1.

The specification in column (5) shows that both the Chinese and the ROW import penetrations are positive and significant. The estimated coefficients of the interaction terms with the unskilled-labor intensive indicators were negative, though only the interaction with the ROW import penetration is significant. This means that the total effect of Chinese import penetration on the formal workers' average wage for unskilled-labor intensive industries is positive, while the effect for the ROW import penetration in negative. For non-unskilled-labor intensive industries, the marginal effect on the average formal wage of a percentage point increase in the Chinese import penetration is a 3.47 percent increase, and for the ROW import penetration a 3.04 percent increase. For unskilled-labor intensive industries, these marginal effects are a 2.01 percent increase and a 4.28 percent decrease, respectively. These results are at odds with the predictions of Hypothesis 2. These results for average informal wages are at variance with the findings of Paz (2014) and Pierola and Sanchez-Navarro (2019). The former found a negative effect of ROW imports, and the latter found no effect of Chinese imports for the average worker, though their results for unskilled workers show a negative impact.

Table 5. Worker-level IV estimates of the effects of industry-level import penetration on the wages of formal workers using Equation (2).

Regressors	(1)	(2)	(3)	(4)	(5)	(6)
Chinese import penetration$_{t-1}$	−0.003 **	−0.009 ***	0.002	0.011 ***	0.028 **	0.009 ***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.012)	(0.003)
ROW import penetration$_{t-1}$	0.001	−0.003 **	−0.002 **	−0.000	0.025 **	−0.012 **
	(0.001)	(0.001)	(0.001)	(0.005)	(0.012)	(0.005)
Chinese imp. penet.$_{t-1}$ × L. int.$_j$		0.028 ***			−0.008	
		(0.003)			(0.020)	
ROW imp. penet.$_{t-1}$ × L. int.$_j$		−0.013 ***			−0.066 **	
		(0.003)			(0.028)	
Chinese imp. penet.$_{t-1}$ × Manuf.$_s$			−0.006 ***			0.015 ***
			(0.001)			(0.004)
ROW imp. penet.$_{t-1}$ × Manuf.$_s$			0.004 ***			0.021 ***
			(0.000)			(0.002)
Age	0.053 ***	0.055 ***	0.053 ***	0.055 ***	0.043 ***	0.015 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)
Age^2	−0.000 ***	−0.001 ***	−0.000 ***	−0.001 ***	−0.000 ***	0.021 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Female	−0.285 ***	−0.292 ***	−0.282 ***	−0.290 ***	−0.247 ***	0.078 ***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)	(0.002)
Married	0.028 ***	0.027 ***	0.028 ***	0.027 ***	0.035 ***	−0.001 ***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)
Black	−0.080 ***	−0.078 ***	−0.081 ***	−0.079 ***	−0.088 ***	−0.375 ***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)
Asian	0.151 ***	0.148 ***	0.151 ***	0.151 ***	0.166 ***	0.011 ***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.003)
Years of schooling	0.047 ***	0.047 ***	0.047 ***	0.047 ***	0.044 ***	−0.061 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)
High school	0.107 ***	0.109 ***	0.106 ***	0.108 ***	0.093 ***	0.110 ***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.015)
College	0.700 ***	0.699 ***	0.700 ***	0.699 ***	0.707 ***	0.054 ***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.001)
Inverse Mills ratio	0.080 ***	0.118 ***	0.065 ***	0.104 ***	−0.124 ***	0.573 ***
	(0.013)	(0.013)	(0.013)	(0.014)	(0.033)	(0.050)
Endogeneity test				39.24 ***	22.32 ***	115.62 ***
				(0.000)	(0.000)	(0.000)
Technique	OLS	OLS	OLS	IV	IV	IV

Notes: Number of observations is 533,584. *** and ** indicate statistical significance at the 1% and 5% levels, respectively. Standard errors are bootstrapped with 500 repetitions. Sample weights from PNAD/Census used. State, industry, and year fixed effects included in the estimated model.

The estimates in column (6) display a positive coefficient for the Chinese import penetration and a negative for ROW import penetration. Both coefficients are statistically significant. The estimated coefficients of the interaction terms between the manufacturing state indicator and import penetrations are positive and significant. This means that, for manufacturing states, the total effects of both import penetrations are positive. For a formal worker located in a non-manufacturing state, the marginal effect on the average formal wage of a percentage point increase in the Chinese import penetration is a 0.1 percent decrease and for the ROW import penetration a 2.66 percent decrease. In manufacturing states, in contrast, these marginal effects are a 2.6 percent increase and a 0.64 percent increase, respectively. These estimates corroborate Hypothesis 3.

Table 6 reports the estimates for the average informal wage based on Equation (3). The OLS estimates are displayed in columns (1)–(3), and the IV estimates in columns (4)–(6). As in the previous table, the selection term estimated coefficients indicate the presence of selection effects. They are statistically significant in every specification, except in column (5). The estimated coefficients of the workers' observable characteristics are stable across specifications. They present signs that are comparable to those in Table 5; however, the coefficients for female and years of schooling have a larger magnitude, while the coefficients for Asian, high school, and college indicators exhibit a smaller magnitude. As before, the null of exogeneity of the import penetrations is rejected at the 5 percent level in the IV specifications.

Table 6. Worker-level IV estimates of the effects of industry-level import penetration on the wages of informal workers using Equation (3).

Regressors	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Chinese import penetration$_{t-1}$	0.012 ***	−0.002	0.015 ***	0.005	0.046 *	−0.019
	(0.004)	(0.006)	(0.005)	(0.008)	(0.025)	(0.019)
ROW import penetration$_{t-1}$	0.010 ***	−0.001	−0.001	0.032 ***	0.046 *	−0.025
	(0.002)	(0.004)	(0.003)	(0.009)	(0.025)	(0.018)
Chinese imp. penet.$_{t-1}$ × L. int.$_j$		0.021 ***			−0.045	
		(0.007)			(0.028)	
ROW imp. penet.$_{t-1}$ × L. int.$_j$		0.006			−0.018	
		(0.007)			(0.030)	
Chinese imp. penet.$_{t-1}$ × Manuf.$_s$			−0.007 **			0.010
			(0.003)			(0.014)
ROW imp. penet.$_{t-1}$ × Manuf.$_s$			0.009 ***			0.058 ***
			(0.001)			(0.014)
Age	0.054 ***	0.067 ***	0.070 ***	0.051 ***	0.052 ***	0.138 ***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.006)	(0.020)
Age^2	−0.001 ***	−0.001 ***	−0.001 ***	−0.000 ***	−0.000 ***	−0.002 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	−0.374 ***	−0.436 ***	−0.450 ***	−0.363 ***	−0.363 ***	−0.766 ***
	(0.010)	(0.012)	(0.013)	(0.011)	(0.030)	(0.093)
Married	0.026 ***	0.015 **	0.013 **	0.029 ***	0.027 ***	−0.033 **
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.015)
Black	−0.076 ***	−0.064 ***	−0.061 ***	−0.078 ***	−0.077 ***	0.004
	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)	(0.022)
Asian	0.089 **	0.060	0.050	0.097 **	0.092 **	−0.118 *
	(0.039)	(0.039)	(0.039)	(0.039)	(0.041)	(0.063)
Years of schooling	0.060 ***	0.063 ***	0.064 ***	0.059 ***	0.059 ***	0.083 ***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.007)
High school	0.019 **	0.048 ***	0.055 ***	0.015	0.014	0.203 ***
	(0.009)	(0.010)	(0.010)	(0.009)	(0.014)	(0.033)
College	0.376 ***	0.374 ***	0.370 ***	0.378 ***	0.375 ***	0.333 ***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.032)
Inverse Mills ratio	−0.110 ***	0.113 ***	0.166 ***	−0.153 ***	−0.149	1.309 ***
	(0.020)	(0.034)	(0.035)	(0.024)	(0.095)	(0.304)
Endogeneity Test				7.31 **	14.14 ***	19.12 ***
				[0.026]	[0.001]	[0.000]
Technique	OLS	OLS	OLS	IV	IV	IV

Notes: Number of observations is 136,382. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Standard errors are bootstrapped with 500 repetitions. Sample weights from PNAD/Census used. State, industry, and year fixed effects included in the estimated model.

We can see in column (4) that the only statistically significant import penetration coefficient is for the ROW, which is positive. The marginal effect on the average informal wage of a percentage point increase in the Chinese import penetration is a 0.52 percent increase and for the ROW import penetration is a 3.39 percent increase. These figures support Hypothesis 1. The estimates in column (5) for both the Chinese and the ROW import penetrations are positive and significant at the 10 percent level. The coefficients of the interaction terms with the unskilled-labor intensive indicator are negative and not significant. Focusing on non-unskilled-labor intensive industries, the marginal effect on the average informal wage of a percentage point increase in the Chinese import penetration is a 5.4 percent increase and for the ROW import penetration a 5.3 percent increase. For the unskilled-labor intensive industries, these figures are a 0.12 percent increase and a 2.26 percent increase, respectively. These effects are at variance with the predictions of Hypothesis 2. These results for average formal wages are comparable to those of Paz (2014), who found a positive effect of ROW imports. Yet, Pierola and Sanchez-Navarro (2019) found no effect of Chinese imports on the average informal worker, and a negative effect for the unskilled informal worker.

The estimates in column (6) reveal that both import penetrations are negative, albeit not statistically significant. Their interactions with the manufacturing state indicator are positive, but only the interaction with the ROW import penetration is statistically significant. The marginal effect on the average informal wage of a percentage point increase in the Chinese import penetration is a 4.72 percent decrease, and for the ROW import penetration a 6.61 percent decrease for non-manufacturing states. The effects for manufacturing states are a 0.14 percent decline and a 2.23 percent increase, respectively. These different impacts on the average informal wage support Hypothesis 3.

Robustness Checks

The robustness check conducted here employs state-by-year fixed effects in lieu of state and year effects. This new specification is used to account for state specific policies such as changes in state-level educational systems, state-level minimum wages, and labor regulations enforcement (Almeida et al. 2022). Moreover, since commodity (iron ore or soybeans, for instance) production in Brazil is geographically concentrated in a few states, these state-by-year effects can also pick-up the effects of the increased Chinese demand for these primary commodities. Table 7 reports the Probit and IV Probit estimates of the selection equation using this new set of fixed effects. We can see that the results are very similar to those in Table 3. Table 8 reports the IV estimates for the average wage of informal workers in columns (1)–(3) and for the average wage of formal workers in columns (4)–(6). The estimates in columns (1)–(3) are very similar to those in columns (4)–(6) in Table 6. Moreover, these new estimates are more statistically significant in many cases. The specifications for the average wage of formal workers present estimated coefficients that are very similar in magnitude to those in columns (4)–(6) in Table 5. Nevertheless, their statistical significance has slightly declined. In sum, these estimates obtained using a different specification of fixed effects corroborates the results obtained with the main specifications.

Table 7. Worker-level IVProbit estimates of the effects of industry-level import penetration on the informal status indicator using Equation (1) and state, year and industry fixed effects.

Regressors	(1)	(2)	(3)	(4)	(5)	(6)
Chinese import penetration$_{t-1}$	−0.009	−0.003	−0.003	0.004	0.069 *	0.029 ***
	(0.012)	(0.009)	(0.016)	(0.007)	(0.038)	(0.009)
ROW import penetration$_{t-1}$	−0.007	−0.003	−0.007	0.016	0.055 *	0.046 ***
	(0.004)	(0.004)	(0.005)	(0.014)	(0.030)	(0.014)
Chinese imp. penet.$_{t-1}$ × L. int.$_j$		−0.017			−0.069 **	
		(0.016)			(0.034)	

Table 7. *Cont.*

Regressors	(1)	(2)	(3)	(4)	(5)	(6)
ROW imp. penet.$_{t-1}$ × L. int.$_j$		−0.001			−0.083	
		(0.022)			(0.070)	
Chinese imp. penet.$_{t-1}$ × Manuf.$_s$			−0.009			−0.032 ***
			(0.013)			(0.007)
ROW imp. penet.$_{t-1}$ × Manuf.$_s$			0.000			−0.037 ***
			(0.004)			(0.003)
Age	−0.120 ***	−0.120 ***	−0.120 ***	−0.120 ***	−0.120 ***	−0.119 ***
	(0.008)	(0.008)	(0.008)	(0.002)	(0.002)	(0.002)
Age^2	0.002 ***	0.002 ***	0.002 ***	0.002 ***	0.002 ***	0.002 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	0.464 ***	0.464 ***	0.463 ***	0.464 ***	0.461 ***	0.456 ***
	(0.092)	(0.092)	(0.092)	(0.008)	(0.008)	(0.008)
Married	0.134 ***	0.133 ***	0.134 ***	0.136 ***	0.132 ***	0.135 ***
	(0.036)	(0.036)	(0.036)	(0.008)	(0.008)	(0.008)
Black	−0.079 ***	−0.079 ***	−0.079 ***	−0.079 ***	−0.079 ***	−0.080 ***
	(0.023)	(0.023)	(0.023)	(0.013)	(0.013)	(0.013)
Asian	0.201 ***	0.201 ***	0.201 ***	0.201 ***	0.199 ***	0.215 ***
	(0.048)	(0.048)	(0.048)	(0.054)	(0.054)	(0.054)
Years of schooling	−0.028 ***	−0.028 ***	−0.028 ***	−0.028 ***	−0.028 ***	−0.028 ***
	(0.006)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)
High school	−0.175 ***	−0.175 ***	−0.175 ***	−0.175 ***	−0.174 ***	−0.172 ***
	(0.018)	(0.018)	(0.018)	(0.011)	(0.011)	(0.011)
College	0.019	0.019	0.019	0.019	0.019	0.036 *
	(0.033)	(0.033)	(0.033)	(0.020)	(0.020)	(0.020)
Other household has a formal job$_{ijst}$	−0.344 ***	−0.344 ***	−0.344 ***	−0.344 ***	−0.343 ***	−0.343 ***
	(0.020)	(0.020)	(0.020)	(0.007)	(0.007)	(0.007)
Exogeneity test				7.552 **	29.22 ***	142.1 ***
				[0.023]	[0.000]	[0.000]

Notes: Number of observations is 671,134. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Standard errors clustered at the industry level. Sample weights from PNAD/Census used. The excluded instruments used in all IV estimates are the Latin American countries' Chinese share of imports, the Latin American countries' high-income countries share of imports, and their interactions.

Table 8. Worker-level IV estimates of the effects of industry-level import penetration on the wages of informal and formal workers using Equations (2) and (3), respectively.

Regressors	(1)	(2)	(3)	(4)	(5)	(6)
Chinese import penetration$_{t-1}$	−0.004	0.053 **	0.001	0.011 ***	0.010	0.000
	(0.008)	(0.022)	(0.008)	(0.003)	(0.012)	(0.003)
ROW import penetration$_{t-1}$	0.035 ***	0.055 **	0.015 *	−0.001	0.007	−0.009 **
	(0.009)	(0.022)	(0.009)	(0.005)	(0.013)	(0.004)
Chinese imp. penet.$_{t-1}$ × L. int.$_j$		−0.063 **			0.012	
		(0.026)			(0.019)	
ROW imp. penet.$_{t-1}$ × L. int.$_j$		−0.024			−0.026	
		(0.025)			(0.027)	
Chinese imp. penet.$_{t-1}$ × Manuf.$_s$			−0.016 **			0.018 ***
			(0.007)			(0.003)
ROW imp. penet.$_{t-1}$ × Manuf.$_s$			0.023 ***			0.012 ***
			(0.006)			(0.002)
Age	0.051 ***	0.048 ***	0.070 ***	0.053 ***	0.051 ***	0.062 ***
	(0.002)	(0.002)	(0.005)	(0.001)	(0.001)	(0.001)
Age^2	−0.000 ***	−0.000 ***	−0.001 ***	−0.000 ***	−0.000 ***	−0.001 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	−0.361 ***	−0.348 ***	−0.448 ***	−0.284 ***	−0.279 ***	−0.316 ***
	(0.011)	(0.013)	(0.024)	(0.003)	(0.004)	(0.005)

Table 8. *Cont.*

Regressors	(1)	(2)	(3)	(4)	(5)	(6)
Married	0.030 ***	0.031 ***	0.015 **	0.028 ***	0.029 ***	0.023 ***
	(0.006)	(0.006)	(0.006)	(0.002)	(0.002)	(0.002)
Black	−0.077 ***	−0.079 ***	−0.058 ***	−0.081 ***	−0.082 ***	−0.074 ***
	(0.011)	(0.011)	(0.009)	(0.004)	(0.004)	(0.003)
Asian	0.100 ***	0.101 ***	0.050	0.153 ***	0.154 ***	0.136 ***
	(0.038)	(0.039)	(0.032)	(0.015)	(0.015)	(0.012)
Years of schooling	0.059 ***	0.058 ***	0.064 ***	0.046 ***	0.046 ***	0.049 ***
	(0.001)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)
High school	0.016	0.010	0.053 ***	0.110 ***	0.107 ***	0.122 ***
	(0.010)	(0.010)	(0.010)	(0.004)	(0.004)	(0.003)
College	0.377 ***	0.376 ***	0.363 ***	0.701 ***	0.702 ***	0.691 ***
	(0.026)	(0.026)	(0.022)	(0.009)	(0.009)	(0.008)
Inverse Mills ratio	−0.160 ***	−0.208 ***	0.158 **	0.078 ***	0.050 **	0.262 ***
	(0.024)	(0.033)	(0.074)	(0.013)	(0.022)	(0.027)
Endogeneity test	13.45 ***	26.27 ***	16.21 ***	39.37 ***	4.15	92.07 ***
	[0.001]	[0.000]	[0.000]	[0.000]	[0.125]	[0.000]
Job type	Informal	Informal	Informal	Formal	Formal	Formal

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Standard errors are bootstrapped with 500 repetitions. Sample weights from PNAD/Census used.

5. Conclusions

China is one of the most populous countries in the world, and it entered the 21st century not only as one of the largest and fast-growing economies but also as a major player in world trade. This rather swift ascension together with its cost advantage in manufacturing production prompted several concerns in developing countries as to whether they would still be able to sustain a dynamic manufacturing sector in view of this Chinese competitive edge. Such concern is built on the fact that many observers perceive a strong manufacturing sector as a key driver of economic growth and as a provider of higher wage jobs relative to those available in agriculture and services.

A good case study to assess such concerns is the increase in import competition experienced by the Brazilian manufacturing sector in 2000–2012. In this period, the import penetration increased by more than 20 percent and the Chinese share of such imports increased from 3 to 20 percent. Brazil is also the largest economy of Latin America and has a large and diverse manufacturing sector with ubiquitous informal employment.

This study employed Brazilian household data to examine the impacts of the increasing Chinese and rest of the world import penetrations on the likelihood of informal employment and on the average wage of formal and informal workers in the manufacturing sector for 2000–2012. The empirical methodology employs a switching regressions model as in Paz (2014), which accounts for worker self-selection into formal and informal jobs and for the potential endogeneity of trade policy.

The empirical results indicate that greater industry-level Chinese and ROW import penetrations increase the informal job likelihood at different intensities. Furthermore, these effects are heterogeneous and modulated by the unskilled labor intensity of the industry and by the degree of industrialization of the Brazilian states. Indeed, the ROW import penetration has a negative effect, and the Chinese import penetration has no effect on the informality likelihood in unskilled-labor intensive industries. In contrast, both import penetrations have positive impact on informality likelihood in the remaining industries. An increase in the Chinese import penetration reduces informality while ROW import penetration increases it in manufacturing states. Nevertheless, both forms of import penetration positively affect informality in non-manufacturing states.

The effects on the average formal and informal wages are more nuanced. An increase in Chinese import penetration raises the average formal wage, except in manufacturing states. Greater ROW import penetration decreases the average formal wage, except for in manufacturing states or in non-unskilled-labor intensive industries. For the average informal wage, both the Chinese and the ROW import penetration have positive effects of different magnitudes, and negative effects for informal workers located in non-manufacturing states.

The evidence amassed in this study suggests that the effects of international trade on labor market outcomes are moderated by the country of origin of imports and, at the same time, by the unskilled labor intensity of industries and by regional characteristics. The important nuances uncovered by this study should not be overlooked in the design of public policies to address potential harmful effects of increased import competition, especially because the most vulnerable workers seem to experience a negative impact from this trade.

Moreover, this study's estimates are at variance with those of the extant literature for different countries and periods of time. This strongly suggests that such trade effects are highly heterogeneous. Unfortunately, data scarcity that plagues the entire literature—such as the lack of employee–employer matched data covering informal jobs—is a major limitation of this study because it precludes an investigation of the role of either the unobservable characteristics of workers or firm characteristics. Given the available data, a promising avenue for future research is the use of cross-country data to investigate whether country-specific institutions are behind these disparate results.

Funding: This research received no external funding.

Data Availability Statement: The data used in this article can be found at www.ibge.gov.br and ipeadata.gov.br.

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

Notes

1 This type of ambiguity is also present in other models of informal labor like Goldberg and Pavcnik (2003) and Aleman-Castilla (2006).

2 This type of excluded instrument is also used by Paz (2021).

References

Aleman-Castilla, Benjamin. 2006. *The Effect of Trade Liberalization on Informality and Wages: Evidence from Mexico*. CEP Discussion Paper No. 763. London: Centre for Economic Performance, London School of Economics and Political Science.

Aleman-Castilla, Benjamin. 2020. *Trade and Labour Market Outcomes Theory and Evidence at the Firm and Worker Levels*. ILO Working Paper 12. Geneva: ILO.

Almeida, Rita, Lourenço S. Paz, and Jennifer P. Poole. 2022. *Precarization or Protection? The Impact of Trade and Labour Policies on Informality*. Helsinki: The United Nations University World Institute for Development Economics Research (UNU-WIDER).

Araújo, Bruno César, and Lourenço S. Paz. 2014. The Effects of Exporting on Wages: An Evaluation Using the 1999 Brazilian Exchange Rate Devaluation. *Journal of Development Economics* 111: 1–16. [CrossRef]

Ashournia, Damoun, Jakob Roland Munch, and Daniel Nguyen. 2014. *The Impact of Chinese Import Penetration on Danish Firms and Workers*. IZA Discussion Paper No. 8166. Bonn: Institute for the Study of Labor.

Baumann, Renato. 2001. Brazil in the 1990s: An economy in transition. *CEPAL Review* 73: 147–69. [CrossRef]

Chandra, Piyush. 2014. WTO subsidy rules and tariff liberalization: Evidence from accession of China. *The Journal of International Trade & Economic Development* 23: 1170–205.

Corseuil, Carlos Henrique, and Miguel Nathan Foguel. 2002. *Uma Sugestão de Deflatores Para Rendas Obtidas a Partir de Algumas Pesquisas Domiciliares do IBGE*. Textos Para Discussão, IPEA TD 0897. Brasília: Ipea.

Dávalos, Jorge. 2019. Trade openness effects on informality and the real exchange rate channel. *Applied Economics Letters* 26: 506–10. [CrossRef]

Facchini, Giovanni, Marcelo Olarreaga, Peri Silva, and Gerald Willmann. 2010. Substitutability and Protectionism: Latin America's Trade Policy and Imports from China and India. *The World Bank Economic Review* 24: 446–73. [CrossRef]

Goldberg, Pinelopi Koujianou, and Nina Pavcnik. 2003. The response of the informal sector to trade liberalization. *Journal of Development Economics* 72: 463–96. [CrossRef]

Iacovone, Leonardo, Ferdinand Rauch, and L. Alan Winters. 2013. Trade as an engine of creative destruction: Mexican experience with Chinese competition. *Journal of International Economics* 89: 379–92. [CrossRef]

IBGE. 2015. Banco de Dados Agregados. Sistema IBGE de Recuperação Automática—SIDRA. 2015. Available online: https://sidra.ibge.gov.br/home/pimpfbr/brasil (accessed on 11 April 2021).

IBGE. 2016. Tabelas Sinóticas Retropoladas 2000–2013. Available online: https://www.ibge.gov.br/estatisticas/economicas/comercio/9052-sistema-de-contas-nacionais-brasil.html (accessed on 12 April 2021).

IPEADATA. 2017. Instituto de Pesquisa Economica Aplicada. Available online: http://ipeadata.gov.br (accessed on 11 April 2021).

Kapri, Kul Prasad, and Lourenço S. Paz. 2019. The Effects of the Chinese imports on Brazilian Manufacturing Workers. *Economies* 7: 76.

Kume, Honório, Guida Piani, and Carlos Souza. 2003. Instrumentos de Política Comercial no Periodo 1987–1998. In *A Abertura Comercial Brasileira nos anos 1990: Impactos Sobre Emprego e Salário*. Edited by Carlos Henrique Corseuil and Honório Kume. Rio de Janeiro, Brazil: Ipea.

Kume, Honorio, Guida Piani, and Pedro Miranda. 2008. Política Comercial, Instituições e Crescimento Economico no Brasil. In *Crecimiento Económico, Instituciones, Política Comercial y Defensa de la Competencia en el Mercosur*. Edited by Honorio Kume. Montevidéu: Red Mercosur, vol. 11.

Maddala, Gangadharrao S. 1983. *Limited-Dependent and Qualitative Variables in Economics*. Cambridge: Cambridge University Press.

Melitz, Marc J. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71: 1695–725. [CrossRef]

Moreira, Mauricio Mesquita. 2006. *Fear of China: Is There a Future for Manufacturing in Latin America?* Intal-ITD, Occasional paper 36. Buenos Aires: Intal.

Paz, Lourenço S. 2014. The Impacts of Trade Liberalization on Informal Labor Markets: A Theoretical and Empirical Evaluation of the Brazilian Case. *Journal of International Economics* 92: 330–48. [CrossRef]

Paz, Lourenço S. 2018. The effect of import competition on Brazil's Manufacturing Labor Market in the 2000s: Are imports from China different? *The International Trade Journal* 32: 76–99. [CrossRef]

Paz, Lourenço S. 2019a. Chinese imports' impacts on Brazil's Inter-Industry Wage Premium. *Journal of Economic Studies* 46: 1052–64. [CrossRef]

Paz, Lourenço S. 2019b. *The Impact of the China Shock on the Manufacturing Labor Market in Brazil*. IDB WORKING PAPER SERIES N° IDB-WP-01085. Washington, DC: Inter-American Development Bank.

Paz, Lourenço S. 2021. How does import competition impact job type? *Economics Bulletin* 41: 2063–20.

Paz, Lourenço S., and John Ssozi. 2021. The Effects of Chinese Imports on Female Workers in the Brazilian Manufacturing Sector. *The Journal of Development Studies* 57: 807–23. [CrossRef]

Pierola, Martha Denisse, and Dennis Sanchez-Navarro. 2019. *Import Competition in the Manufacturing Sector in Peru: Its Impact on Informality and Wages*. IDB WORKING PAPER SERIES N° IDB-WP-01093. Washington, DC: Inter-American Development Bank.

United Nations. 2003. *Statistical Division. UN Comtrade*. New York: United Nations.

WTO Antidumping Gateway. 2016. World Trade Organization, Geneva, Switzerland. Available online: https://www.wto.org/english/tratop_e/adp_e/adp_e.htm (accessed on 10 April 2021).

economies

Article

Aggregate Online Brand Name Pharmacy Price Dynamics for the United States and Mexico

Thomas M. Fullerton, Jr. [1,*] **and Steven L. Fullerton** [2]

[1] Department of Economics & Finance, University of Texas at El Paso, El Paso, TX 79968-0543, USA
[2] Border Region Modeling Project, University of Texas at El Paso, El Paso, TX 79968-0543, USA; slfullerton@utep.edu
* Correspondence: tomf@utep.edu; Tel.: +1-915-747-7747; Fax: +1-915-747-6282

Abstract: Virtual cross-border medical tourism allows many residents in the United States to purchase brand name medicines from companies in Mexico without travelling there. Monthly economic reports indicate that the online brand name pharmaceutical product prices in Mexico are noticeably lower than the corresponding internet prices in the United States. There have been very few econometric studies on how these prices are linked and the dynamic nature of those relationships. Results in this study indicate that online medicine prices in Mexico respond very rapidly to online prices changes in the high-price market.

Keywords: online pharmacies; brand name medicines; price dynamics; applied econometrics

JEL Classification: I11 Health Markets; L81 E-Commerce; M21 Business Economics

Citation: Fullerton, Thomas M., Jr., and Steven L. Fullerton. 2022. Aggregate Online Brand Name Pharmacy Price Dynamics for the United States and Mexico. *Economies* 10: 112. https://doi.org/10.3390/economies10050112

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 24 March 2022
Accepted: 10 May 2022
Published: 12 May 2022

1. Introduction

In 2018, prescription drug expenditures in the United States exceeded USD 335 billion (CDC 2021). Due to high prices, consumers from the United States often purchase brand name pharmaceutical products in Mexico (Dalstrom et al. 2020). Those purchases are transacted both in-person and online (Fullerton and Miranda 2011; Fullerton et al. 2014). Most research related to this form of cross-border medical tourism employs point-in-time cross-sectional data sets (Dalstrom et al. 2021).

This study examines an overlooked aspect of the online brand name medicine trade. Rather than examine price differences for a single point-in-time, this effort examines whether there is a dynamic relationship between online medicine prices in the two countries. It achieves this using a data set that monitors monthly prices for the top 50 online brand name medicines sold in Mexico and the United States (Fullerton and Fullerton 2022).

2. Literature Review

Medication price studies tend to rely on cross sectional data (Daalen et al. 2021). Some studies have documented links between market structures and competitive effect pricing patterns (Barigozzi and Jelovac 2020; Granlund 2022). Potential time series linkages among cross-country industry aggregate prices have not received as much attention. While regulatory and other international boundary barriers may limit the strength of those pricing links, very little empirical evidence regarding that possibility has been previously conducted.

Lower priced medicines are sought by consumers due to budget constraints. As shown by the "generics paradox", there is no guarantee that the availability of lower priced drugs will cause prices to decline (Vandoros and Kanavos 2013). However, cross-sectional price data indicate that United States consumers can save more than USD 1000 annually by purchasing prescription medicines online in Mexico (Fullerton et al. 2014). There is substantial evidence that consistently confirms that brand name pharmaceuticals are substantially more expensive in the United States than in Mexico and most other countries

(Mulcahy et al. 2021), but no study of how those prices evolve over time or if cross-country movements in those prices are correlated.

Brand name medicines tend to be more expensive in lower- and middle-income countries once income and cost of living differences are taken into account, but not in dollar terms based on currency market exchange rates (Moye-Holz and Vogler 2022). Patients in the United States are able to take advantage of the latter with respect to pharmaceutical products purchased in Mexico either in-person (Fullerton and Miranda 2011) or online (Fullerton et al. 2014). This study examines whether the online prices for medicines sold to customers from both countries are correlated over time.

3. Data and Methodology

The University of Texas at El Paso Border Region Modeling Project collects prices every month for brand name medicines that can be purchased online in Mexico and the United States. These prices do not include value-added taxes in Mexico, sales taxes in the United States, or shipping and handling costs, nor do they include the cost of doctor appointments. International medicines can be legally imported by individual consumers in the United States as long as the amounts purchased do not exceed 90-day supplies. Furthermore, those purchases can only be for personal use (CBP 2011; Fullerton et al. 2014).

The sample period is from January 2007 through December 2021. Prices for the 50 top selling medicines are for equal dosages on a per unit basis, exclusive of shipping fees, handling charges, and taxes. Websites in both countries advertise their services in English and quote product prices in dollars. The internet sites that sell the medicines do not always offer all 50 brand name products, so multiple sites have to be sampled. Unweighted averages for both sets of prices are calculated from the data collected for each month.

Table 1 reports summary statistics for both average price variables. In line with conventional wisdom, the average price in Mexico is noticeably lower than that of the United States. However, the north-of-the-border price index is somewhat more variable than that observed for Mexico. The standard deviation of the United States online price index is double the magnitude of the southern online price index. While that is impressive, it should be noted that the difference in the magnitudes of the coefficients of variation is much more moderate. The data for both indices are slightly left-skewed, but fairly close to being symmetric. Both sets of price data are notably platykurtic.

Table 1. Summary Statistics.

Variable Name	AVG_$_MEX	AVG_$_USA
Mean	11.28	18.85
Median	13.09	23.17
Std. Dev.	5.28	10.89
Maximum	18.71	30.51
Minimum	3.85	3.76
Skewness	−0.341	−0.361
Kurtosis	1.510	1.370
CV	0.468	0.578

Notes: Sample Period: January 2007—December 2021. Unit of Measure: United States Dollars. AVG_$_MX—Average Brand Name Medicines Price in Dollars for Mexico Internet Sites. AVG_$_USA—Average Brand Name Medicines Price in Dollars for USA Internet Sites. Std. Dev.—Standard Deviation. CV—Coefficient of Variation. Data Web Sites—goodrx.com accessed on 15 December 2021, healthwarehouse.com accessed on 15 December 2021, medicinesmexico.com accessed on 15 January 2020, medsmex.com accessed on 15 December 2018, mexmedsforyou.com accessed on 15 December 2021.

A linear transfer function autoregressive moving average (LTF ARIMA) methodology is employed to examine the dynamic relation between the two indices. For monthly time series data, the LTF ARIMA approach has proven useful in other contexts involving international economic data for the United States and Mexico (Fullerton and Solis 2020). Due to the magnitude of the United States market for pharmaceutical products, causality between the two price indices is expected to run north-to-south. If that were not the case, then an alternative method such as vector autoregression would be required (Diebold 2007).

4. Empirical Results

A series of F-tests at various lags indicate weak north-to-south unidirectional causality. Given all of the regulatory barriers separating the two markets, that is not surprising. The data indicate that changes in United States brand name online drug prices are accompanied by similar changes in Mexico internet prices, but not vice versa. That implies that medicine prices in Mexico are responsive to price fluctuations in the higher income market.

A cross-correlation function is used to determine the potential lag structure governing the linkages between the two series, (Diebold 2007). The outcome indicates that the reaction of online prices in Mexico is contemporaneous with no subsequent, statistically reliable lagged responses. Parameter estimation outcomes are summarized in Table 2.

Table 2. Generalized Least Squares Estimation Results.

Variable Name	Coefficient	Std. Error	t-Statistic	Probability
Constant	−0.0195	0.0358	−0.5460	0.5857
d(AVG_$_USA)	0.4610	0.0492	9.3618	0.0000
MA(4)	−0.1082	0.0763	−1.4191	0.1576
R-squared	0.3276	Dependent Variable Mean	0.0475	
Adj. R-squared	0.3200	Dep. Var. Std. Deviation	0.6341	
Pseudo R-sq.	0.9951	Std. Error of Regression	0.5229	
Sum Sq. Residuals	48.1226	Log Likelihood	−136.4433	
F-statistic	42.8730	Prob. (F-statistic)	0.0000	
Adjusted Sample:	2007M02–2021M12	Included Observations:	179	
Convergence achieved after 3 iterations.				
Coefficient covariance computed using outer product of gradients.				
Inverted MA Roots	0.57	0.00–0.57i	0.00–0.57i	−0.57

The constant term in Table 2 indicates that Mexico online medicine prices tend to lose ground at a rate of roughly USD 0.02 per month. The standard error for that coefficient is fairly large, so the reliability of that estimate is not very strong. Most of the sample period corresponds to a period of relative currency market weakness for the peso, especially after 2015. That may account for some of the steady erosion of the south-of-the-border prices that occur in Table 2.

The slope coefficient indicates that every USD 1 increase (decrease) in United States online prices is accompanied by a USD 0.46 increase (decrease) in Mexico online prices for brand name pharmaceuticals. In the sample means, the magnitude of the coefficient indicates that the elasticity of online prices in Mexico with respect to those of the United States is 0.770. That estimate implies a fairly high degree of cross-border price sensitivity for brand name medicines sold online in Mexico. The computed t-statistic for this parameter estimate is 9.362 with a 0.000 p-value.

The overall diagnostics in Table 2 are relatively favorable. The pseudo coefficient of determination is 0.995. Given that, it is not surprising that the computed F-statistic of 42.873 has a p-value of 0.000. The specification does not, however, account for all systematic variation in the dependent variable. Residual serial correlation necessitates the inclusion of moving average term at lag 4. That coefficient has a somewhat large standard error, but higher log-likelihood statistic results for the equation when it is included.

As a robustness check, the sample period was shortened by seven years to cover only January 2007 through December 2014. That period is selected because it pre-dates the 2016

US presidential campaign that began in 2015, wherein several major candidates criticized trade with Mexico. Those results also indicate that the internet pharmacy prices charged in Mexico react very quickly to any north-of-the-border price variations and the slope coefficient is almost identical to that reported in Table 2. The coefficients of determination for the shorter period are larger than those reported in Table 2, potentially due to a less friendly trade environment between the two countries and due to the advent of the global pandemic in 2020 (Komkova 2019; Ceylan et al. 2020).

5. Conclusions

Empirical research on global pharmaceutical prices has uncovered numerous interesting commonalities and differences across international markets. This study examines dynamic aggregate price movements for brand name medicines sold over the internet in the United States and Mexico. Although brand name medicine distribution is tightly regulated in both countries, it is legal for consumers to import limited quantities for personal usage.

Over the course of the 15-year sample period, internet medicine prices in Mexico are, on average, 40 percent below the online prices charged in the United States. The prices in Mexico react very quickly to any variation in the prices in the higher-income market. Every USD 1 change in the United States average price index is matched by a USD 0.46 change in the Mexico average price index.

Based on the results reported in this exploratory effort, it seems clear that more research on this topic is warranted. This study employs simple average price measures for both economies. A logical next step would be to examine dynamic patterns among prices for these brand name medicines within a panel setting. The outcomes noted above indicate that cross-border linkages between individual online brand name pharmaceutical prices may be fairly strong.

Author Contributions: Conceptualization, T.M.F.J. and S.L.F.; Methodology, T.M.F.J.; Validation, T.M.F.J. and S.L.F.; Formal analysis, T.M.F.J. and S.L.F.; Investigation, T.M.F.J. and S.L.F.; Data curation, S.L.F.; Writing—original draft preparation, T.M.F.J. and S.L.F.; Writing—review and editing, T.M.F.J. and S.L.F.; Supervision, T.M.F.J. and S.L.F.; Project administration, T.M.F.J.; Funding acquisition, T.M.F.J. and S.L.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by El Paso Water, National Science Foundation Grant DRL-1740695, Texas Department of Transportation ICC 24-0XXIA001, TFCU, UTEP Institutional Advancement, and the UTEP Center for the Study of Western Hemispheric Trade.

Institutional Review Board Statement: Ethical review and approval are not required for this study due to the fact that anonymous data are used that are not traceable to individuals at any time.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data employed for this study are available upon request from tomf@utep.edu and slfullerton@utep.edu.

Acknowledgments: Helpful comments and suggestions were provided by Dan Pastor and two anonymous referees.

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

References

Barigozzi, Francesca, and Izabela Jelovac. 2020. Research Funding and Price Negotiation for New Drugs. *Health Economics* 29: 83–96. [CrossRef] [PubMed]

CBP. 2011. *FDA Enforcement Policy Regarding the Personal Import of Violative Drugs*. Washington, DC: U.S. Customs and Border Protection.

CDC. 2021. *Health, United States, 2019*. Hyattsville: National Center for Health Statistics.

Ceylan, Rahmiye Figen, Burhan Ozkan, and Esra Mulazimogullari. 2020. Historical Evidence for Economic Effects of COVID-19. *Europen Journal of Health Economics* 21: 817–23. [CrossRef] [PubMed]

Daalen, Jules M. Janssen, Anouk den Ambtman, Mark Van Houdenhoven, and Bart J. F. van den Bemt. 2021. Determinants of Drug Prices: A Systematic Review of Comparison Studies. *BMJ Open* 11: e046917. [CrossRef] [PubMed]

Dalstrom, Matthew, Ryan Chung, and Lynette Castronovo. 2020. Impacting Health through Cross-Border Pharmaceutical Purchases. *Medical Anthropology* 39: 182–95. [CrossRef] [PubMed]

Dalstrom, Matthew, Ryan Chung, and Lynette M. Castronovo. 2021. Strategies to Reduce Healthcare Costs. *Journal Pharmaceutical Health Services Research* 12: 87–89. [CrossRef]

Diebold, Francis X. 2007. *Elements of Forecasting*. Mason: Thomson South-Western.

Fullerton, Thomas M., Jr., and Omar Solis. 2020. Borderplex Bridge Wait Time Headache Reactions. *Journal of Transport Economics & Policy* 54: 58–78.

Fullerton, Thomas M., Jr., and Osvaldo Miranda. 2011. Borderplex Brand Name Medicine Price Differences. *Applied Economics* 43: 929–39. [CrossRef]

Fullerton, Thomas M., Jr., and Steven L. Fullerton. 2022. Borderplex Top 10 Economic Predictions for 2022. *Borderplex Business Barometer* 6: 1–58.

Fullerton, Thomas M., Jr., Francisco J. Pallares, and Adam G. Walke. 2014. Are Online Pharmacy Prices Really Lower in Mexico? *Asian Economic & Financial Review* 4: 416–31.

Granlund, David. 2022. The Price Effects of Competition from Parallel Imports and Therapeutic Alternatives. *Review of Industrial Organization* 60: 63–92. [CrossRef]

Komkova, Elena G. 2019. USMCA instead of NAFTA. *Mirovaya Ekonomika* 63: 50–58. [CrossRef]

Moye-Holz, Daniela, and Sabine Vogler. 2022. Comparison of Prices and Affordability of Cancer Meidicines in 16 Countries in Europe and Latin America. *Applied Health Economics & Health Policy* 20: 67–77. [CrossRef]

Mulcahy, Andrew W., Christopher M. Whaley, Mahlet Gizaw, Daniel Schwam, Nathaniel Edenfield, and Alejandro U. Becerra-Ornelas. 2021. *International Prescription Drug Price Comparisons: Current Empirical Estimates and Comparisons with Previous Studies*. Santa Monica: RAND Corporation.

Vandoros, Sotiris, and Panos Kanavos. 2013. The Generics Paradox Revisited: Empiricasl Evidence from Regulated Markets. *Applied Economics* 45: 3230–39. [CrossRef]

economies

Article

Does the Obesity Problem Increase Environmental Degradation? Macroeconomic and Social Evidence from the European Countries

Matheus Koengkan [1] and José Alberto Fuinhas [2,*]

1 Governance, Competitiveness and Public Policies (GOVCOPP), Department of Economics, Management, Industrial Engineering and Tourism (DEGEIT), University of Aveiro, 3810-193 Aveiro, Portugal; matheuskoengkan@ua.pt
2 Centre for Business and Economics Research (CeBER), Faculty of Economics, University of Coimbra, 3004-512 Coimbra, Portugal
* Correspondence: fuinhas@uc.pt

Abstract: The macroeconomic effect of the obesity epidemic on environmental degradation was examined for panel data from thirty-one European countries from 1991 to 2016. The quantile via moments model (QVM) was used to realize our empirical investigation. The empirical results indicate that the obesity epidemic, electricity consumption, and urbanisation encourage environmental degradation by increasing CO_2 emissions, while economic growth decreases them. Moreover, we identify that the obesity epidemic raises the environmental degradation problem in three ways. First, the obesity epidemic is caused by the increased consumption of processed foods from multinational food corporations. The increase in food production will positively impact energy consumption from non-renewable energy sources. Second, obesity reduces physical and outdoor activities, increasing the intensive use of home appliances and motorized transportation and screen-viewing leisure activities, consequently increasing energy consumption from non-renewable energy sources. A third possible way can be related indirectly to economic growth, globalization, and urbanisation. This empirical investigation will contribute to the literature and for policymakers and governments. Therefore, this investigation will encourage the development of initiatives to mitigate the obesity problem in European countries and accelerate the energy transition process. Finally, this investigation will open a new topic in the literature regarding the correlation between the obesity epidemic and environmental degradation.

Keywords: CO_2 emissions; energy consumption; environmental degradation; European region; food production; health problem; macroeconomics; obesity

Citation: Koengkan, Matheus, and José Alberto Fuinhas. 2022. Does the Obesity Problem Increase Environmental Degradation? Macroeconomic and Social Evidence from the European Countries. *Economies* 10: 131. https://doi.org/10.3390/economies10060131

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 28 March 2022
Accepted: 30 May 2022
Published: 6 June 2022

1. Introduction

Climate change caused by the increase in carbon dioxide emissions (CO_2) is one of the biggest concerns in the European region (Bianco et al. 2019). CO_2 emissions are the most significant contributor to increased greenhouse gas emissions (GHGs), contributing 77% of GHGs. In contrast, other gases such as methane (CH_4), nitrous oxide (N_2O), and ozone (O_3) contribute with 14%, 8%, and 1%, respectively (e.g., Koengkan and Fuinhas 2021a, 2021b; Khan et al. 2014).

Several initiatives have emerged to mitigate climate change (e.g., the United Nations Framework Convention on Climate Change (UNFCCC), the Earth Summit (1992), the Kyoto Protocol (1997), the 21st Conference of the Parties (COP 21) (2015), and the 26th Conference of the Parties (COP 26) (2021)). These initiatives aim to substantially limit the increase in temperature levels during this century to lower than 2 °C and limit that increase to 1.5 °C. These initiatives will take temperatures to pre-industrial levels. In addition, all countries that align with this agreement will move towards a low-carbon economy. Indeed, as has long been known, global GHGs, mainly CO_2 emissions, have been increasing since the 1970s (e.g., World Bank Open Data 2022; Koengkan and Alberto Fuinhas 2021b).

However, from 1990 to 2016, these emissions grew dramatically, and in 1990, CO_2 emissions were 3.0991 (metric tons per capita) and reached 4.6807 (metric tons per capita) in 2016. During this period, these emissions grew from 33 megatons of CO_2 equivalent (MtCO2eq) in 1990 to 47 MtCO2eq in 2016, an annual increase of 1.5% during this period (Bárcena et al. 2019). The energy, industry, transport, and building sector have increased emissions since the 2000s. In 2010 the energy sector contributed 25%, AFOLU (agriculture, forestry and other land use) 24%, industry 21%, transport 14%, other energy sources 10%, and the building sector contributed 6% to this growth (Koengkan et al. 2020).

Most of these emissions are caused by the production of electricity and heat, which emanate from the industrial and residential sectors. GHGs come about through direct emissions from fossil fuel combustion for providing power, cooling, heating, and cooking (Khan et al. 2014). As stated above, in 2010, energy consumption was responsible for 25% of global GHGs. The growth in global energy use is accountable for increasing CO_2 emissions. Energy use has been rising since the 1970s when energy use was 1337.00 (kg of oil equivalent per capita) in 1971 and reached 1897.25 in 2016 (e.g., World Bank Open Data 2022; Koengkan and Fuinhas 2021a).

Indeed, 94% of this energy use in 1970 came from fossil fuels worldwide, and only 6.45% came from renewable energy. However, in 2016, the contribution of fossil fuels decreased slightly, reaching 85% of the total energy use. Indeed, this reduction is related to the increase in the share of renewable energy sources, which reached 14.35% in 2016 (Our World in Data 2022).

In the European region, CO_2 emissions in 1971 were 8.0244 (metric tons per capita) and reached a value of 6.4684 (metric tons per capita) in 2016 (World Bank Open Data 2022). Therefore, between 1990 and 2004, these emissions in the European region remained relatively unchanged. However, due to the 10.8% decrease in primary energy consumption, CO_2 emissions dropped sharply between 2005 and 2016 (IEA 2020). For example, in 1990 the energy consumption was 1641 million tonnes of oil equivalent (Mtoe), while in 2004 it had already reached 1789 Mtoe. However, between 2005 and 2016, this consumption decreased and fell to 1598 Mtoe in the year 2016. The energy efficiency improvements that increased the share of renewable energy sources in the energy matrix and the changes in climate conditions were the causes for this decrease in the primary energy consumption between 2005 to 2016 for most European region countries (e.g., the European Environment Agency 2019; Eurostat 2020).

In the European region, 93% of this energy use in 1970 came from fossil fuels, and only 6.90% came from renewable energy. However, in 2016 this value decreased slightly, reaching 75% of total energy use. Indeed, this reduction is related to the increase in the share of renewable energy sources in energy use, where 25% was reached in 2016 (e.g., Our World in Data 2022; Koengkan and Alberto Fuinhas 2021b).

As has long been known, various drivers have been influencing the increase of CO_2 emissions. Economic growth, globalisation, trade, financial liberalisation, urbanisation, population growth and energy prices have gained notoriety. However, the literature has given little consideration to a possible relationship between the obesity epidemic problem and the increase in environmental degradation. To the best of our knowledge, the first study to address the link between obesity and climate change was made by Edwards and Roberts (2009). However, this link is very complex and is not exempt from criticism (e.g., Gallar 2010).

Nevertheless, the literature remains scarce and primarily focused on the effect of obesity on climate change via CO_2 emissions. Their connections are associated with oxidative metabolic demands, food production, and fossil fuels. The links between obesity and climate change also include processed foods from fast-food and multinational supermarket chains, multinational food corporations, food production on farms, transportation of goods, retail processing and storage of processed food. These approaches also emphasise that the intensive use of motor vehicles and modern household appliances reduces physical effort in the context of a sedentary lifestyle (e.g., Magkos et al. 2019; Furlow 2013; Viscecchia et al. 2012; Breda et al. 2011; Edwards and Roberts 2009).

Obesity is defined as abnormal or excessive fat accumulation that may impair health, that is, individuals that have a mean body mass index (BMI) $\geq$ 30.0, as defined by the World Health Organization (WHO). The organisation also defines 'overweight' as BMI $\geq$ 25.0 (Our World in Data 2022). In 2016, about 39% (2.0 billion) of adults aged 18 years and older, 38% of men and 40% of women worldwide, were overweight or obese (Our World in Data 2022). Indeed, this chronic disease is a significant risk factor for people with many other diseases.

The obesity epidemic has increased significantly over the past three decades. In 2014, over 600 million adults, or 13% of the total adult population, were classified as obese worldwide. Of these 600 million obese adults, 11% are men, and 15% are women (Pineda et al. 2018). It is estimated that 25.6% of the total adult population (18 and over) can be classified as obese in the European region. This disease has almost doubled since the late 1980s. In 1985, the percentage of obese adults that are obese was 12.60%, and this value reached 23.30% in 2016 (Our World in Data 2022). Indeed, it has hit the world's richest countries, regardless of individuals' income levels.

Indeed, the obesity epidemic is caused by several factors: genetic, social, economic, environmental, political, and physiological, which have interacted to varying degrees over time (Wright and Aronne 2012). Moreover, other factors, such as the globalisation process, urbanisation and technological progress, have caused an increase in the obesity epidemic (e.g., Fox et al. 2019; Toiba et al. 2015; Popkin 1998). The weight increase caused by the factors mentioned earlier contributes to making people physically less active (less physical activity also contributes to increasing obesity). Consequently, it leads to using more motorised vehicles and modern household appliances that reduce physical effort. In addition, it contributes to weight gain due to the lower caloric expenditure of individuals, as well as to increased consumption of processed foods, mainly produced by (i) multinational food companies, (ii) multinational supermarkets, and (iii) fast-food chains. All of these factors contribute to the increase in energy consumption from non-renewable energy sources and negatively impact the environment.

In the literature, the impact of the obesity problem on environmental degradation using a macroeconomic approach is not advanced in the literature. For this reason, this investigation opted to use similar studies related to this topic (e.g., Koengkan and Alberto Fuinhas 2021b; Cuschieri and Agius 2020; Magkos et al. 2019; Swinburn 2019; Webb and Egger 2013; Viscecchia et al. 2012; Breda et al. 2011; Davis et al. 2007; Higgins 2005). These investigations pointed out that the obesity problem increases environmental degradation. However, none of these studies realised an analysis using a macroeconomic approach. They used the percentage of adults that are overweight or obese as a proxy for obesity and CO_2 emissions as a proxy for environmental degradation as well as the quantile via moments (QvM) method. Furthermore, these studies do not use the urban population and globalisation as independent variables. Moreover, none of these studies investigated the European countries. That is, there are gaps in the literature that need to be filled.

In order to fill the gaps that were mentioned above, this investigation will identify the macroeconomic effect of the obesity epidemic on environmental degradation. Indeed, to identify this effect, this empirical investigation will study a group of thirty-one countries from the European region between 1991 to 2016 that have experienced a rapid increase in the obesity epidemic and social, economic, and environmental transformations. Certainly, to carry out this empirical investigation, the quantile via moments (QvM) approach, which Machado and Silva (2019) developed, will be used.

This investigation will introduce a new analysis regarding the macroeconomic impact of the obesity problem on environmental degradation in European countries. This topic of research has never been approached before in the literature. Therefore, this study can open new opportunities for studying the correlation between obesity and environmental degradation through a macroeconomic aspect. Furthermore, this investigation is innovative in that it uses econometric and macroeconomic approaches to identify the possible effect of the obesity problem on ecological degradation. It is the first time this methodology approach has been employed in this kind of investigation.

Moreover, this investigation will contribute to the literature for several reasons: (i) it introduces a new analysis regarding the effect of the obesity epidemic on environmental degradation in European countries. This topic of investigation is new and can open new issues of inquiry regarding the relationship between health and the environment using a macroeconomic approach; (ii) this investigation will contribute to introducing the QvM model; and (iii) the results of this study will help governments and policymakers develop more initiatives to reduce the obesity problem in the European countries, in addition to policies to reduce the consumption of non-renewable energy sources and environmental degradation.

This study is ordered as follows. Section 2 presents the literature review regarding the effect of the obesity epidemic on environmental degradation. Section 3 provides the data and the methodology approach. Section 4 presents the results and a brief discussion. Finally, Section 5 presents the conclusions and limitations of the study.

2. Literature Review

As mentioned before in the introduction, the literature has given little attention to a possible connection between the obesity epidemic problem and the increase in environmental degradation. Due to this, our investigation opted to use the few existing pieces of literature that approached this topic of investigation and which are similar (e.g., Koengkan and Alberto Fuinhas 2021b; Cuschieri and Agius 2020; Magkos et al. 2019; Swinburn 2019; Webb and Egger 2013; Viscecchia et al. 2012; Breda et al. 2011; Davis et al. 2007; Higgins 2005).

Koengkan and Alberto Fuinhas (2021b) investigated the impact of the overweight epidemic on energy consumption in thirty-one countries in the European region from 1990 to 2016. The authors find that being overweight increases the consumption of energy from fossil fuels and consequently increases the emissions of CO_2. Moreover, according to the authors, the increase in energy consumption and CO_2 emissions by the overweight epidemic is related to the increased consumption of processed foods from fast-food and multinational supermarket chains and multinational food corporations. Indeed, this process positively affects farm production, fast-food and multinational supermarket chains, and multinational food corporations to attend to the demand for processed foods. This increase affects the consumption of energy from non-renewable energy sources.

Magkos et al. (2019) explored the effect of obesity on climate change. The authors point out that this health problem can aggravate climate change with increased CO_2 emissions in three ways: (i) oxidative metabolic demands; (ii) food production; and (iii) fossil fuels use. The increase of oxidative metabolic demands caused by the higher body mass associated with obesity is responsible for 7% of total GHGs. Indeed, the rise in production driven by the need to provide higher energy caloric intake is responsible for 52% of total GHGs. In contrast, the increase in fossil fuel consumption caused by transport and food production is responsible for 41% of these emissions. Thus, the authors estimated that the obesity epidemic adds the equivalent of 700 megatons of extra carbon dioxide to emissions per year or about 1.6% of the total global emissions. This idea is also shared by Swinburn (2019), who investigated the same topic.

Other authors also share similar ideas, such as Breda et al. (2011), who studied the relationship between climate change and obesity in four regions of Karakalpakstan in Uzbekistan. According to the authors, there is strong evidence that being overweight contributes more to climate change, where overweight influences food consumption and production. Those categories contribute more to climate change by consuming processed foods from fast-food and multinational supermarket chains, multinational food corporations, and farms. The authors add that the food sector accounts for 7% of CO_2 emissions, 43% of CH_4 emissions, and 50% of N_2O emissions produced across the entire economy. Viscecchia et al. (2012) investigated the relationship between obesity and climate change in Italy. The authors opted to use the ordinary least squares method to undertake this investigation. The authors found that the increase in food consumption with low energy content has a twofold effect on reducing obesity and climate change mitigation. Moreover, the increase in food consumption with low energy caloric content reduces the obesity rate from 9.68 to 7.04% and avoids 5,406,000 tons of CO_2 emissions per year.

Cuschieri and Agius (2020) investigated the link between diabetes caused by obesity and climate change. The increase in the demand for processed food caused by obesity also has an adverse effect on the climate. The authors highlight that this effect is caused by the increased transportation of goods, retail processing, and processed food storage. Webb and Egger (2013) also studied the link between obesity and climate change and point out that some behaviours connected with obesity also affect emissions of GHGs associated with climate change. The authors show that consuming processed food and non-renewable energy sources results from intensive motor vehicles and modern household appliances that reduce physical effort.

Davis et al. (2007) investigated the interactions between cars, obesity, and climate change in the United Kingdom from 1974 to 2004. Their results precede the idea developed later by Webb and Egger (2013). According to the authors, the intensive use of motor vehicles in the United Kingdom has reduced physical activity and increased obesity and CO_2 emissions by increasing non-renewable energy sources. Higgins (2005), in an investigation that investigated whether "exercise-based transportation reduces oil dependence, carbon emissions and obesity", points out that the use of the automobile as a means of transport also contributes to a sedentary lifestyle and the obesity epidemic and poor health. The author adds that these problems consume 27% of global oil production and produce 25% of global carbon emissions.

The summary of the literature presented in this section has discussed some of the most consequential investigations that directly approached the impact of obesity on environmental degradation and similar investigations. However, none of these studies realised an analysis using a macroeconomic approach. Instead, they used the percentage of adults that are overweight or obese as a proxy for obesity and CO_2 emissions as a proxy for environmental degradation, as well as the quantile via moments (QvM) method. Furthermore, these studies do not use the urban population and globalisation as independent variables. Moreover, none of these studies investigated the European countries. Therefore, there are gaps in the literature that need to be filled. The following section will show the data and methods used in this investigation.

3. Data and Methodology

This section is organised into two parts. The data, including the variables, is presented first, and the second part describes the methodology used in this study.

3.1. Data

As mentioned before, this section will present the data used in this empirical investigation. Thirty-one countries from the European region were used, namely Austria (AT), Belgium (BE), Bulgaria (BG), Croatia (HR), Czech Republic (CZ), Denmark (DK), Estonia (EE), Finland (FI), France (FR), Germany (DE), Greece (GR), Hungary (HU), Iceland (IS), Ireland (IE), Italy (IT), Latvia (LV), Lithuania (LT), Luxembourg (LU), Malta (MT), the Netherlands (NL), Norway (NO), Poland (PL), Portugal (PT), Republic of Cyprus (CY), Romania (RO), Slovakia (SK), Slovenia (SI), Spain (ES), Sweden (SE), Turkey (TR), and the United Kingdom (UK). Moreover, as mentioned earlier, this group of countries was selected because they have experienced a rapid increase in the obesity epidemic and social, economic, and environmental transformations that have facilitated this problem in the last three decades.

Data for the period between 1991 and 2016 was utilised in this investigation. Please note that data for the variable Y only begins in 1991. Moreover, the time series of this investigation goes until 2016 due to data availability for the variable OBESE (see, Our World in Data 2022). The variables used in this empirical investigation are shown in Table 1 below. The variables EC, Y, and CO_2, were first transformed into per capita values. Per capita values allow disparities to be controlled for population growth over time and within countries (e.g., Fuinhas et al. 2022; Koengkan et al. 2020). After this, all variables were transformed into natural logarithms ("Log").

Table 1. Description of variables and summary statistics.

Description of Variables				Summary Statistics				
	Variable	Definition	Source	Obs	Mean	Std Dev	Min	Max
Dependent variable	$LogCO_2$	CO_2 emissions (kg per capita 2011 PPP $ of GDP).	World Bank Open Data (2022)	806	−1.3226	0.4362	−2.6267	0.2197
Independent variables	LogOBESE	Percentage of adults that are obese. Obese is defined as having a body-mass index (BMI) equal to or greater than 30. BMI is a person's weight in kilograms divided by their height in metres squared.	Our World in Data (2022)	806	4.027	0.0902	3.7612	4.2535
	LogEC	Electric power consumption (kWh per capita) from fossil fuels.	World Bank Open Data (2022)	806	8.6668	0.6540	6.8724	10.9433
	LogY_PC	GDP per capita, PPP (constant 2011 international $).	World Bank Open Data (2022)	806	10.2095	0.5013	8.5276	11.421
	LogUP	Urban population (% of the total population). This variable is a proxy for urbanisation.	World Bank Open Data (2022)	806	4.2624	0.1727	3.8809	4.5841
	LogGLOBA	Globalisation index. This index is compounded by de facto economic, social and political components of globalisation and is scaled from 1 to 100. Thus, this variable encompasses three main factors of globalisation.	KOF Globalization Index (2022)	802	4.9901	0.1616	4.2648	5.2038

Notes: Obs., Std.-Dev., Min. and Max denote the number of observations, the standard deviation, the minimum and the maximum, respectively.

To capture the effect of obesity on environmental degradation, the econometric model has to include other variables that also explain pieces of the explained variable, the so-called control variables. Thus, the model uses variables that are in line with economic theory. Furthermore, the variables have support from the literature. For instance, variables Y_PC, EC, and UP have been used to justify the increase in CO_2 emissions, and GLOBA has been used as a proxy for environmental degradation (e.g., Koengkan and Fuinhas 2021a; Hdom and Fuinhas 2020; Wang et al. 2018). On the other hand, the variable OBESE, taken from the Our World in Data (2022), has not been used in literature to capture the rise in environmental degradation. Therefore, our study, by including this variable, is pioneering. The panel of countries and the variables used in this investigation are presented in this subsection. The methodology pursued in this investigation will follow the strategy presented in Figure 1 below.

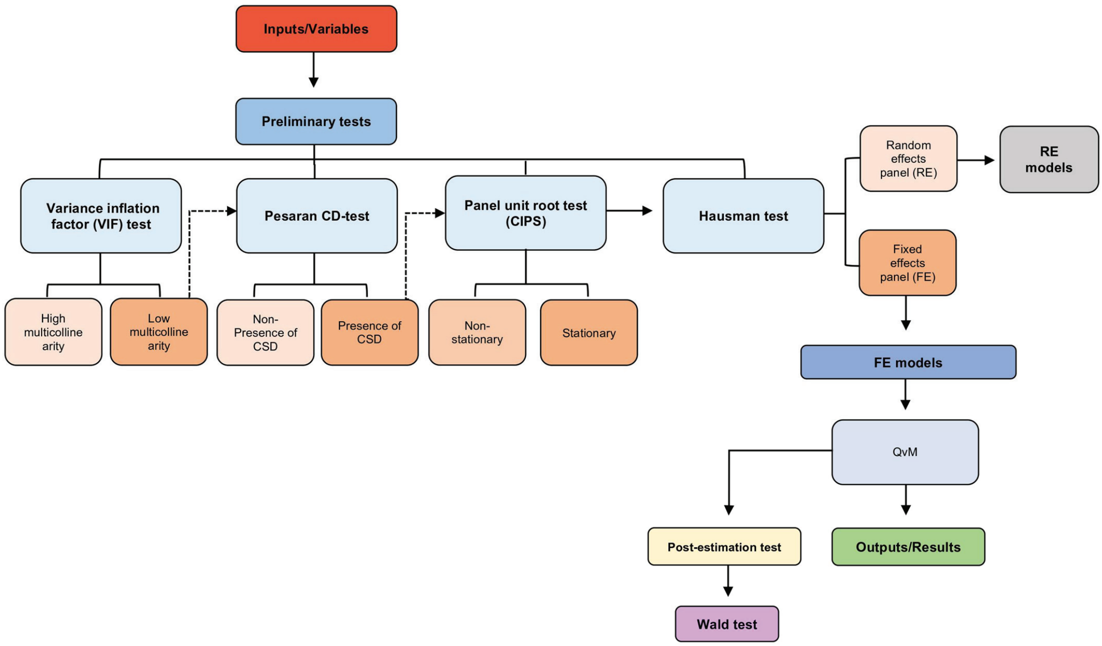

Figure 1. Methodology strategy. The authors created this figure.

Therefore, after presenting the variables and the methodology strategy that this investigation will follow, it is also necessary to present the methodological approach for our empirical analysis.

3.2. Methodology

As mentioned before, our empirical analysis will use the QvM model approach. This method was developed by Machado and Silva (2019) as an alternative for quantile regression. Consistent with Kazemzadeh et al. (2022) and Koengkan et al. (2020), this method can differentiate individual effects in panel data models. Moreover, Machado and Silva (2019) also add that this method can explain how the regressor affects the entire conditional distribution. Therefore, this method can be adapted to provide estimates in cross-sectional models with endogenous variables. Koengkan et al. (2020) reveal that his method is based on moment conditions, not on conditional means, to identify the conditional means under exogeneity and that it allows the identification of the exact structural quantile function. Thus, given these advantages indicated by Koengkan et al. (2020) and Machado and Silva (2019), this empirical investigation opted to use this methodological approach.

After briefly presenting the methodological approach and its advantages, it is time to show the equation where the QvM is constructed. For this, Equation (1) is presented:

$$Y_{it} = a_i + X'_{it}\beta + \left(\delta_i + Z'_{it}\gamma\right)U_{it}\,, \tag{1}$$

where $\left\{\,(Y_{it},\ X'_{it})'\right\}$ comes from a panel of n individuals $i = 1, \ldots, n$ over T periods, with $P\{\delta_i + Z'_{it}\gamma > 0\} = 1$. The parameters (α_1, δ_i), $i = 1, \ldots, n$, catch the individual i fixed-effects, and Z is a k-vector of known differentiable (with probability 1) transformations of the components of X, with the element l given by $Z_l = Z_l(X)$, $l = 1, \ldots, k$. The sequence $\{\,X_{it}\}$ is *i.i.d.*, for any fixed I, and independent across t. U_{it} are *i.i.d.*, across i and t, statistically independent of X_{it}, and normalised to satisfy the moment condition $E(U) = 0 \wedge E(|U|) = 1$ (e.g., Koengkan et al. 2020).

Before estimating the model regression, it is advised to assess the statistical proprieties of variables. Therefore, a battery of preliminary tests is applied (see Table 2 below).

Table 2. Preliminary tests.

Test	Objective
Variance inflation factor (VIF) (Belsley et al. 1980)	This test verifies the presence of multicollinearity between the variables of the model.
Cross-section dependence (CSD) (Pesaran 2004)	This test verifies the presence of cross-sectional dependence (CSD) in model variables.
Panel unit root test (CIPS) (Pesaran 2007)	This test verifies the presence of unit roots in the model's variables.
Hausman test	This test verifies the presence of heterogeneity, i.e., whether the panel has random effects (RE) or fixed-effects (FE) in the model regression.

Indeed, after the regression of the QvM model, it is necessary to apply the post-estimation tests to identify if the models are adequate. Table 3 below shows the post-estimation tests that will be used in this empirical investigation.

Table 3. Post-estimation tests for the QvM model.

Post-Estimation Tests for the QvM Model	
Test	**Objective**
Wald test (Agresti 1990)	This test verifies the global significance of the estimated models.

Notes: The authors created this table.

Stata Commands

After presenting the preliminary tests, the QvM model and the post-estimation test, we must show the Stata commands we used in this empirical investigation. Table 4 below shows the Stata commands used.

Table 4. Stata commands.

Preliminary Tests	
Test	**Stata Command**
Descriptive statistics of variables	*sum*
Variance inflation factor (VIF) test	*vif*
Cross-sectional dependence (CSD) test	*xtcd*
Panel unit root test (CIPS)	*multipurt*
Hausman test	*hausman* (with the option, *sigmamore*)
QvM model	
QvM	*xtqreg, quantile (0.25 0.50 0.75)*
Post-estimation test	
Wald test	*testparm*

Indeed, the preliminary tests, model regression, and post-estimation tests will be accomplished using Stata 17.0. The following section will show the results and discussions.

4. Results and Discussions

As previously explained, this section will present the results and the possible explanations for the macroeconomic impact of the obesity epidemic on environmental degradation. The preliminary tests indicated that the variables used have characteristics such as (i) low-multicollinearity among independent variables (as shown in Table A1 in Appendix A); (ii) cross-sectional dependence in the logs of variables (as shown in Table A2 in Appendix A); (iii) variables with orders of integration borderline I(0) and I(1) (as shown

Table A3 in Appendix A); and (iv) the presence of panel fixed-effects (as shown Table A4 in Appendix A). This last result is significant because fixed effects are required in the QvM model. Therefore, the fixed-effects estimator is the most suitable for accomplishing this empirical analysis. Therefore, the empirical results of these tests are vital to identifying the characteristics of the group of countries under study and the possible methodologies to be applied.

The next step after the preliminary tests is to carry out the model regression. The 25th, 50th, and 75th quantiles were calculated to assess the non-linearities of the effect of the obesity epidemic on environmental degradation. We utilised these quantiles to simplify the exhibition of empirical results. Furthermore, we used several quantiles (e.g., 5th, 10th, 15th, and others). It can be seen that there is no information loss, as all independent variables pointed to the same effect of the dependent variable.

Moreover, a dummy variable was added to the model because, during the analysis, the European countries suffered some shocks, such as economic and political. Indeed, if these shocks are not considered, it could produce inaccurate results and misinterpretations during model regression. Therefore, this empirical analysis added dummy variables that represent these shocks. In the literature, the inclusion of dummy variables needs to follow the following triple criterion of choice developed by Fuinhas et al. (2017). For example, (i) the potential relevance of recorded economic and political events at the country level; (ii) the occurrence of international events known to have disturbed the European countries; and (iii) a significant disturbance in the estimated residuals. Therefore, the dummy variables added to the model regression are IDEUROPE_2015 (Europe, the year 2015). The dummy variables called "IDEUROPE_2015" represent a decrease in all countries' GDP in the model. This event was caused by the persistent effects of the European debt crisis (often also referred to as the eurozone crisis or the European sovereign debt crisis (Koengkan and Fuinhas 2021a). Table 5 below shows the QvM model regression results with the dummy variable's inclusion. The QvM model results without including the dummy variable can be seen in Table A5 in Appendix A.

Table 5. QvM estimation (controlling for shocks).

Independent Variables	Dependent Variable ($LogCO_2$)					
	Quantiles					
	25th		50th		75th	
LogOBESE	1.2695	**	1.2152	***	1.1743	***
LogEC	0.2568	***	0.2942	***	0.3225	***
LogY_PC	−0.5582	***	−0.6068	***	−0.6434	***
LogUP	0.6176	*	0.5724	***	0.5383	**
IDEUROPE2015	−0.1522	***	−0.0916	***	−0.0459	*
TREND	−0.0343	***	−0.0332	***	−0.0324	***
Obs	806		806		806	

Notes: ***, ** and * denote statistically significant at the 1%, 5%, and 10% levels, respectively.

The QvM model regression results indicate that in the 25th, 50th and 75th quantiles, the obesity epidemic, electricity consumption, and urbanisation process increase CO_2 emissions. That is, they encourage environmental degradation by increasing CO_2 emissions. Nevertheless, economic growth decreases emissions of CO_2 in the European region. Moreover, the results from the QvM model also show non-linear behaviour. The estimated coefficients values vary as we go up (or down) in the quantiles' regression. Thus, the empirical results answer the central question of our empirical investigation. Moreover, the post-estimation tests (e.g., the Modified Wald test) indicate that the estimator of this study is adequate (as shown in Table 6 below).

Table 6. Post-estimation test.

Wald Test	Quantiles					
	25th		50th		75th	
	Chi2(3) = 78.33	***	Chi2(3) = 228.78	***	Chi2(4) = 325.10	***

Notes: *** denotes statistically significant at the 1% level; H_0 of Wald test: the coefficients for all variables are jointly equal to zero.

After finding the positive effect of the obesity epidemic on environmental degradation, it is necessary to ascertain whether the results found by the QvM model regression are reliable and robust when we perform a change in the econometric method approach. Indeed, this approach to finding if the model is robust or not was previously used by Koengkan et al. (2020) and Fuinhas et al. (2017).

After identifying that the obesity epidemic encourages environmental degradation by increasing CO_2 emissions, the next step is to answer the following question: What is the possible explanation for this phenomenon? One possible way of explaining this effect is that the obesity epidemic is caused by the increased consumption of processed foods from multinational food corporations, fast-food chains and multinational supermarket chains, as well as the food production on farms, as indicated by some authors (e.g., Koengkan and Alberto Fuinhas 2021b; Fox et al. 2019; Gerbens-Leenes et al. 2010; Popkin 1998). The increased consumption of processed foods from multinational food corporations and farms will positively affect energy consumption from non-renewable energy sources.

Another explanation for this phenomenon is related to the reduction of outdoor activities, which exacerbates the problem of obesity. Consequently, this reduction will encourage intensive motorised transportation, screen-viewing leisure activities, and the use of home appliances, as indicated by some authors (e.g., Koengkan and Alberto Fuinhas 2021b; Bell et al. 2002; Sobal 2001). Indeed, Koengkan and Alberto Fuinhas (2021b) also add that the increase in the use of home appliances and motorised transportation has implications for the energy demand from fossil fuel energy sources, where the consumption of these kinds of sources increases considerably.

Additionally, the positive impact of the obesity epidemic on CO_2 emissions is also indirectly related to economic growth, globalisation, and urbanisation. Therefore, the levels of obesity have been related to increasing economic activity, where economic development causes effects on dietary changes (e.g., Springmann et al. 2016). Indeed, the transition from low to high income caused by this process tends to induce some individuals to consume fatty and energy-dense foods of animal origin. Therefore, the increase in income contributes to the rise in obesity levels, except in countries where home-produced food is predominant (e.g., Roskam et al. 2010; Gerbens-Leenes et al. 2010).

The globalisation process also causes an increase in obesity by the dietary changes. According to Popkin (1998), Fox et al. (2019), Koengkan and Alberto Fuinhas (2021b) and Koengkan et al. (2021), the process of globalisation will contribute to the food chain extension. As mentioned above, this extension will enable economies of scale in food production processes. Consequently, the economies of scale in food production processes will allow a diet rich in energy-caloric foods. Food consumption with high sugar and salt contents is less expensive and accessible to lower-income classes. Moreover, the unhealthy supply of processed foods is related to the increase in multinational food companies, supermarkets, and fast-food chains caused by globalisation.

The urbanisation process also plays a role in the increase in the obesity problem. The process of urbanisation allows better accessibility to food due to supermarkets, multinational supermarkets, and fast-food chains offering a ready supply of processed foods, which consequently causes the decline of farm stands and open markets with healthier foods (Reardon et al. 2003). This same process also exposes people to the mass media marketing of food and beverages that influence traditional diets (Hawkes 2006). Moreover, urbanisation increases car use and reduces walking or biking for transportation or leisure, contributing to obesity, and obesity increases car use. Moreover, all these explanations align with the findings of Koengkan and Alberto Fuinhas (2021b). They found that economic growth, globalisation, and urbanisation positively affect the overweight problem in the European region. They consequently encourage energy consumption from non-renewable energy sources and subsequently increase CO_2 emissions/environmental degradation. Furthermore, the explanations above align with the results from the complementary analysis that this investigation carried out (as shown in Table A6, in Appendix A).

Indeed, the positive effect of electric power consumption on CO_2 emissions could be related to some factors highlighted in this empirical investigation. On the one hand, it can be linked to electricity consumption in groups of not environmentally responsible countries. Consequently, electricity generation from fossil fuels is linked with increases in CO_2 emissions. These findings may also signal that the panel of countries under study could depend on fossil fuel sources for economic growth. On the other hand, it may be linked to the inefficiency of renewable energy policies that stimulate the development and consumption of renewable energy sources. Several authors have already found this impact (e.g., Fuinhas et al. 2021; Ozcan et al. 2020; Muhammad et al. 2020; Adedoyin et al. 2020; Koengkan and Alberto Fuinhas 2021b; Yazdi and Dariani 2019; Salahuddin et al. 2019; Fuinhas et al. 2017).

However, the negative effect of economic growth on CO_2 emissions could be related to some factors highlighted in this empirical investigation. A justification for the negative impact can be linked to a strong depression/recession affecting people's consumption behaviour. Accordingly, it affected energy-intensive sectors, electricity consumption, and, finally, the emissions of CO_2. A U-shaped relationship between economic growth and CO_2 emissions could be another possible explanation for this negative impact. An increase in economic growth initially leads to a decline in CO_2 emissions levels, consequently reaching a threshold. Indeed, economic activity intensifies environmental degradation. Indeed, the country's industrialisation increases pollution. The policies limiting the levels of industrial pollution can be another explanation. Those policies promote the embracing of environmentally friendly techniques and processes of production.

Consequently, environmentally friendly technologies were promoted, and this promotion contributes to producing and consuming renewable energy by industries and families. However, some authors found a negative impact on economic growth and CO_2 emissions (e.g., Koengkan and Alberto Fuinhas 2021b; Muhammad et al. 2020; Aye and Edoja 2017).

Moreover, the positive effect of urbanisation on CO_2 emissions could be related to increased urban populations, positively affecting the demand for energy from fossil fuel energy sources, households, the transport sector and industries. Additionally, this positive effect could be related to the low energy efficiency improvement caused by the slow introduction of new energy technologies, low diversification of energy sources and low environmental regulation efficiency. It encourages industries' and families' acquisition of environmentally friendly technologies (Koengkan and Alberto Fuinhas 2021b). Figure 2 below summarises the effect of independent variables on the dependent variable.

This section showed the results from the primary model and the robustness check, the possible explanations for the impact of the obesity epidemic on environmental degradation, and a brief explanation of the impact of other variables. The next section will show the conclusions of this experimental investigation.

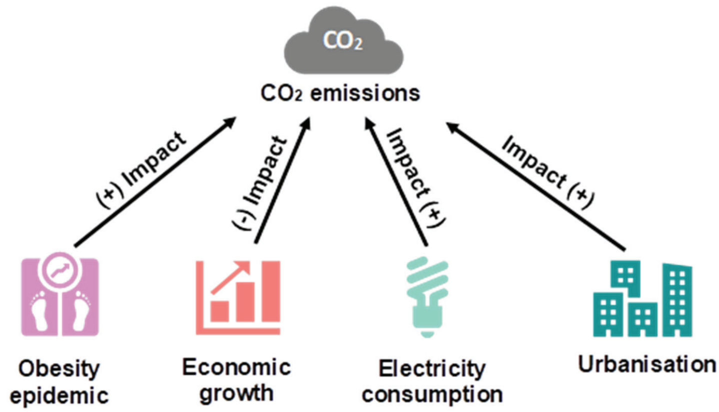

Figure 2. Summarization of the effect of the independent variables on the dependent variable.

5. Conclusions

This empirical investigation approached the macroeconomic effect of the obesity epidemic on environmental degradation using CO_2 emissions as a proxy in a panel of thirty-one European countries from 1991 to 2016. As stated above, the QvM model was used to carry out an empirical study. This study's preliminary tests indicated that the variables used have low-multicollinearity characteristics, cross-sectional dependence in logarithms, and variables have I(0) or borderline I(1) order of integration, non-presence cointegration between the variables, and also fixed-effects.

The QvM and fixed-effect models regression results show that the obesity epidemic, electricity consumption, and urbanisation encourage environmental degradation by increasing CO_2 emissions. At the same time, economic growth decreases the emissions of CO_2 in the European region. Indeed, the post-estimation test results for the QvM model indicate that this study's estimator is adequate.

The obesity epidemic increases the environmental degradation problem in three ways. First, the obesity epidemic is caused by increased consumption of processed foods from fast-food and multinational supermarket chains and multinational food corporations. The increase in food production will positively impact fossil fuel energy sources' energy consumption. Second, obesity reduces physical and outdoor activities, increasing the intensive use of motorised transportation, home appliances, and screen-viewing leisure activities, consequently increasing energy consumption from non-renewable energy sources. Finally, a third possible way can be related indirectly to economic growth, globalisation, and urbanisation. This empirical evidence leads to a supplementary research question: What can be done to reverse the influence of the obesity epidemic on environmental degradation in the European region?

Several initiatives need to be created to reduce the effect of the obesity epidemic on environmental degradation. The first initiative is related to policies that reduce the sale of foods with high calorie-energy close to schools; the second initiative is to create policies that restrict the sale and consumption of unhealthy foods through taxation. Unhealthy food will cost more with the introduction of taxes and, consequently, encourage healthier foods. The third initiative is related to developing policies that encourage the generalised practice of physical activity and its importance. The fourth initiative is related to reducing lobbying by multinational food corporations through policies encouraging local producers. Finally, the fifth and last initiative is associated with creating policies that encourage the food sector to produce foods that are more healthy and have the least possible impact on the environment.

However, this problem is not limited to reducing the obesity problem. It is necessary to change production and consumption in the European region. Indeed, several initiatives have already been implemented to reduce the consumption of non-renewable energy in the region. However, it is essential to do more to reverse this situation. For example, policy-makers need to create more measures to reduce the barriers to products and technologies that improve energy efficiency and produce green energy. This reduction could benefit households and industries by acquiring renewable energy technologies and reducing these products' prices.

Regarding food production, it is necessary to introduce policies that encourage, (i) better productivity, where the efficiency improvements can lead to a 33% reduction in land use, a 12% reduction in water use, and a 16% reduction in production emissions; (ii) reduce livestock emissions, where the increase in productivity and efficiency gains can reduce land use, feed requirements, and GHG emissions per gallon of milk or pound of meat; (iii) reduce the consumption of fertiliser, as the use of these substances emits nitrous oxide, a potent greenhouse gas (the introduction of techniques including nitrification inhibitors can replace applications of fertilisers); (iv) introduce renewable energy and energy efficiency technologies supported by fiscal and financial instruments that help farmers gain access to renewable energy and energy-efficient technologies; and (v) reduce waste and food loss, with support of fiscal and financial instruments to help farmers to improve their equipment and energy efficiency in farm buildings. In addition, these policies can reduce the consumption of fossil fuels and emissions.

This study is a kick-off regarding the effect of the obesity epidemic on environmental degradation and other aspects such as energy consumption, economic growth, and urbanisation. Therefore, this investigation is in the initial stages of maturation, which will supply a solid foundation for second-generation research regarding this topic.

Limitations of the Study

As we already know, this empirical investigation is not free of limitations. Indeed, the preliminary limitations of this empirical investigation stem from (i) the presence of a short period. In this investigation we used the period between 1991 and 2016. Indeed, this period was used due to data availability for the variable OBESE. Moreover, more time is necessary to capture the dynamic effects of the variables OBESE, EC, Y_PC, and UP; (ii) the inexistence of literature that approaches the macroeconomic effect of the obesity epidemic on environmental degradation. The lack of this kind of literature makes difficult the elaboration of deeper discussions regarding the results found; and (iii) the European countries are firmly integrated and are mainly developed ones. This former characteristic limits the generalisation of our results to diverse contexts.

The limitations mentioned above are usually found in investigations in their early stages of maturation. Developing second-generation studies regarding this topic is essential to overcoming these limitations. Despite the limitations in this investigation, this study could draw meaningful conclusions.

Author Contributions: J.A.F.: Writing—review and editing, Supervision, Funding acquisition, and Project administration. M.K.: Conceptualisation, Writing—Original draft, Supervision, Validation, Data curation, Investigation, Formal analysis, and Visualisation. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the research unit on Governance, Competitiveness and Public Policy, UIDB/04058/2020 and UIDP/04058/2020, funded by national funds through FCT—Fundação para a Ciência e a Tecnologia; and by the CeBER R&D unit, funded by national funds through FCT—Fundação para a Ciência e a Tecnologia, I.P., project UIDB/05037/2020.

Data Availability Statement: Data available on request from the corresponding author.

Conflicts of Interest: The authors declare that they have no conflict of interest.

Appendix A

Table A1. VIF-test.

	Dependent Variable ($LogCO_2$)		
Independent variables	LogOBESE	1.54	0.6487
	LogEC	2.59	0.3863
	LogY_PC	3.78	0.2644
	LogUP	1.71	0.5832
	LogGLOBA	3.39	0.2949
Mean VIF		2.60	<6

Table A2. Pesaran CD-test.

Variables		CD-Test	*p*-Value	
Dependent variable	$LogCO_2$	91.88	0.000	***
Independent variables	LogOBESE	109.22	0.000	***
	LogEC	61.94	0.000	***
	LogY_PC	99.91	0.000	***
	LogUP	24.62	0.000	***
	LogGLOBA	105.10	0.000	***

Notes: *** denotes statistical significance at the 1% level; Ho for CD-test: cross-section independence.

Table A3. Panel Unit Root test (CIPS-test).

Variables		Panel Unit Root Test (CIPS) (Zt-Bar)				
		Without Trend			With Trend	
		Lags	Zt-Bar		Zt-Bar	
Dependent variables	$LogCO_2$	1	−5.400	***	94.123	***
Independent variables	LogOBESE	1	1.384		67.492	
	LogEC	1	−6.660	***	84.082	**
	LogY_PC	1	−1.774	**	43.646	**
	LogUP	1	43.899		101.639	***
	LogGLOBA	1	−2.531	***	24.089	***

Notes: *** and ** denote statistically significant at the 1% and 5% levels, respectively.

Table A4. Hausman test.

	Dependent Variable ($LogCO_2$)				
Independent variables	LogOBESE	−2.0925	−1.8408	−0.2517	0.0639
	LogEC	0.3207	0.2833	0.0374	0.0127
	LogY_PC	−0.7426	−0.7778	0.0352	0.0116
	LogUP	1.6847	1.2572	0.4274	0.1125
Chi2 (4)			21.80 ***		

Notes: *** denotes statistically significant at the 1% level.

Table A5. QvM estimation without dummy variable.

Independent Variables	**Dependent Variable ($LogCO_2$)**					
	Quantiles					
	25th		**50th**		**75th**	
LogOBESE	1.4425	***	1.3335	***	1.2524	***
LogEC	0.2597	***	0.2929	***	0.3176	***
LogY_PC	−0.5563	***	−0.5899	***	−0.6150	***
LogUP	0.5592	*	0.5304	***	0.5090	***
TREND	−0.0371	***	−0.0354	***	−0.0341	***
Obs	806		806		806	
F/Wald test	chi2(3) = 91.30 ***		chi2(3) = 249.03 ***		chi2(3) = 206.52 ***	

Notes: *** and * denote statistically significant at the 1% and 10% levels, respectively.

Table A6. QvM estimations (complementary analysis).

Independent Variables	**Dependent Variable (LogOBESE)**					
	Quantiles					
	25th		**50th**		**75th**	
LogGLOBA	0.2397	***	0.2439	***	0.2494	***
LogY_PC	0.0905	***	0.0855	***	0.0792	***
LogUP	0.9900	***	0.9978	***	1.0076	***
LogEC	0.0641	***	0.0637	***	0.0631	***
Obs	829		829		802	
F/Wald test	chi2(3) = 1182.55 ***		chi2(4) = 7155.18 ***		chi2(4) = 3313.36 ***	
	Dependent variable (LogEC)					
LogOBESE	0.4711	**	0.8690	***	1.2596	***
LogGLOBA	0.3514	***	0.0742		−0.1979	
LogY_PC	0.1997	***	0.1481	***	0.0975	*
LogUP	0.8091	**	0.6018	**	0.3983	*
Obs	829		829			
F/Wald test	chi2(4) = 361.24 ***		chi2(4) = 204.63 ***		chi2(4) = 61.24 ***	
	Dependent variable (LogY_PC)					
LogOBESE	0.9484	**	1.0261	***	1.0847	***
LogGLOBA	1.3978	***	1.1282	***	0.9248	***
LogEC	0.0590		0.1387		0.1988	***
LogUP	−1.3186	***	−1.3223	***	−1.3251	***
Obs	829		829		829	
F/Wald test	chi2(4) = 156.71 ***		chi2(4) = 524.13 ***		chi2(4) = 762.55 ***	
	Dependent variable (LogUP)					
LogOBESE	0.4266	***	0.4318	***	0.4365	***
LogGLOBA	−0.0500	***	−0.0663	***	−0.0809	***
LogY_PC	−0.0495	***	−0.0475	***	−0.0456	***
LogEC	0.0185	**	0.0197	***	0.0208	***
Obs	829		829		829	
F/Wald test	chi2(4) = 339.41 ***		chi2(4) = 606.57 ***		chi2(4) = 318.66 ***	
	Dependent variable (LogGLOBA)					
LogOBESE	0.7550	***	0.6667	***	0.5971	***
LogY_PC	0.2737	***	0.2609	***	0.2507	***
LogEC	0.0047		0.0188		0.0298	
LogUP	−0.4515	**	−0.4161	***	−0.3882	***
Obs	829		829		829	
F/Wald test	chi2(4) = 559.54 ***		chi2(4) = 1257.30 ***		chi2(4) = 827.28 ***	

Notes: ***, ** and * denote statistically significant at the 1%, 5%, and 10% levels, respectively.

References

Adedoyin, Festus Fatai, Moses Iga Gumede, Festus Victor Bekun, Mfonobong Udom Etokakpan, and Daniel Balsalobre-Lorente. 2020. Modelling coal rent, economic growth and CO_2 emissions: Does regulatory quality matter in BRICS economies? *Science of The Total Environment* 710: 136284. [CrossRef]

Agresti, Alan. 1990. *Categorical Data Analysis*. New York: John Wiley and Sons, ISBN 0-471-36093-7.

Aye, Goodness C., and Prosper Ebruvwiyo Edoja. 2017. Effect of economic growth on CO_2 emission in developing countries: Evidence from a dynamic panel threshold model. *Cogent Economics & Finance Journal* 5: 1–23. [CrossRef]

Bárcena, Alicia, Joseluis Samaniego, Luis Miguel Galindo, Jimy Ferrer Carbonell, José Eduardo Alatorre, Pauline Stockins, Orlando Reyes, Luis Sánchez, and Jessica Mostacedo. 2019. A economia da mudança climática na América Latina e no Caribe. *CEPAL* 1–61. Available online: https://repositorio.cepal.org/bitstream/handle/11362/44486/1/S1801217_pt.pdf (accessed on 11 February 2022).

Bell, A. Colin, Keyou Ge, and Barry M. Popkin. 2002. The Road to Obesity or the Path to Prevention: Motorized Transportation and Obesity in China. *Obesity Research* 10: 277–83. [CrossRef]

Belsley, David A., Edwin Kuh, and Roy E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley. [CrossRef]

Bianco, Vincenzo, Furio Cascetta, Alfonso Marino, and Sergio Nardini. 2019. Understanding energy consumption and carbon emissions in Europe: A focus on inequality issues. *Energy* 170: 120–30. [CrossRef]

Breda, Joao, Zulfia Atadjanova, Arustan Joldasov, and Pernille MalbergDyg. 2011. Climate Change and Its Impact on Food and Nutrition Security: Report on An Assessment Conducted in Four Regions of the Autonomous Republic of Karakalpakstan, Uzbekistan. pp. 1–89. Available online: https://www.unicef.org/equity/archive/files/Climate_Change_and_Food_and_Nutrition_Security-final_report.pdf (accessed on 11 February 2022).

Cuschieri, Sarah, and Jean Calleja Agius. 2020. The interaction between diabetes and climate change—A review on the dual global phenomena. *Early Human Development* 155: 105220. [CrossRef] [PubMed]

Davis, Adrian, Carolina Valsecchi, and Malcolm Fergusson. 2007. Unfit for purpose: How car use fuels climate change and obesity. *The National Academics of Sciences Engineering Medicine*. Available online: https://trid.trb.org/view/847092 (accessed on 11 February 2022).

Edwards, Phil, and Ian Roberts. 2009. Population adiposity and climate change. *International Journal of Epidemiology* 38: 1137–40. [CrossRef]

European Environment Agency. 2019. *Total Greenhouse Gas Emission Trends and Projections in Europe*. Copenhagen: European Environment Agency, pp. 1–37. Available online: https://www.eea.europa.eu/downloads/6f120fbcec964495b3c693f5f0a635c0/1576746405/assessment-3.pdf (accessed on 11 February 2022).

Eurostat. 2020. *Greenhouse Gas Emission Statistics—Emission Inventories: Statistics Explained*; Luxembourg: Eurostat, pp. 1–7. Available online: https://ec.europa.eu/eurostat/statistics-explained/pdfscache/1180.pdf (accessed on 11 February 2022).

Fox, Ashley, Wenhui Feng, and Victor Asal. 2019. What is driving global obesity trends? Globalisation or "modernisation"? *Globalisation and Health* 15: 32. [CrossRef]

Fuinhas, José Alberto, António Cardoso Marques, and Matheus Koengkan. 2017. Are renewable energy policies upsetting carbon dioxide emissions? The case of Latin America countries. *Environmental Science and Pollution Research* 24: 15044–54. [CrossRef]

Fuinhas, José Alberto, Matheus Koengkan, Nuno Carlos Leitão, Chinazaekpere Nwani, Gizem Uzuner, Fatemeh Dehdar, Stefania Relva, and Drielli Peyerl. 2021. Effect of Battery Electric Vehicles on Greenhouse Gas Emissions in 29 European Union Countries. *Sustainability* 13: 13611. [CrossRef]

Fuinhas, José Alberto, Matheus Koengkan, Nuno Silva, Emad Kazemzadeh, Anna Auza, Renato Santiago, Mônica Teixeira, and Fariba Osmani. 2022. The Impact of Energy Policies on the Energy Efficiency Performance of Residential Properties in Portugal. *Energies* 15: 802. [CrossRef]

Furlow, Bryant. 2013. Food production and obesity linked to climate change. *The Lancet Respiratory Medicine* 1: 187–88. [CrossRef]

Gallar, Manuel. 2010. Obesity and climate change. *International Journal of Epidemiology* 39: 1398–99. [CrossRef] [PubMed]

Gerbens-Leenes, Winnie, Sanderine Nonhebel, and Martinus S. Krol. 2010. Food consumption patterns and economic growth. Increasing affluence and the use of natural resources. *Appetite* 55: 597–608. [CrossRef]

Hawkes, Corinna. 2006. Uneven dietary development: Linking the policies and processes of globalisation with the nutrition transition, obesity and diet-related chronic diseases. *Globalisation and Health* 2: 1–18. [CrossRef] [PubMed]

Hdom, Hélde A. D., and José Alberto Fuinhas. 2020. Energy production and trade openness: Assessing economic growth, CO_2 emissions and the applicability of the cointegration analysis. *Energy Strategy Reviews* 30: 100488. [CrossRef]

Higgins, Paul A. T. 2005. Exercise-based transportation reduces oil dependence, carbon emissions and obesity. *Environmental Conservation* 32: 197–202. Available online: https://www.jstor.org/stable/44521867 (accessed on 11 February 2022). [CrossRef]

IEA. 2020. Available online: https://www.iea.org/data-and-statistics/charts/total-primary-energy-supply-for-the-eu-and-selected-countries-1971-2017 (accessed on 11 February 2022).

Kazemzadeh, Emad, Matheus Koengkan, and José Alberto Fuinhas. 2022. Effect of Battery-Electric and Plug-In Hybrid Electric Vehicles on PM2.5 Emissions in 29 European Countries. *Sustainability* 14: 2188. [CrossRef]

Khan, Muhammad Azhar, Muhammad Zahir Khan, Khalid Zaman, and Lubna Naz. 2014. Global estimates of energy consumption and greenhouse gas emissions. *Renewable and Sustainable Energy Reviews* 29: 336–44. [CrossRef]

Koengkan, Matheus, and José Alberto Fuinhas. 2021a. Is gender inequality an essential driver in explaining environmental degradation? Some empirical answers from the CO_2 emissions in European Union countries. *Environmental Impact Assessment Review* 90: 106619. [CrossRef]

Koengkan, Matheus, and José Alberto Fuinhas. 2021b. Does the overweight epidemic cause energy consumption? A piece of empirical evidence from the European region. *Energy*, 1–19. [CrossRef]

Koengkan, Matheus, José Alberto Fuinhas, and Nuno Silva. 2020. Exploring the capacity of renewable energy consumption to reduce outdoor air pollution death rate in Latin America and the Caribbean region. *Environmental Science and Pollution Research*, 1–19. [CrossRef]

Koengkan, Matheus, José Alberto Fuinhas, and Celso Fuinhas. 2021. Does Urbanisation Process Increase the Overweight Epidemic? The Case of Latin America and the Caribbean Region. *SSRN*, 1–7. [CrossRef]

KOF Globalization Index. 2022. Available online: https://www.kof.ethz.ch/en/forecastsand-indicators/indicators/kof-globalisation-index.html (accessed on 11 February 2022).

Machado, José A. F., and J. M. C. Santos Silva. 2019. Quantiles via Moments. *Journal of Econometrics*. [CrossRef]

Magkos, Faidon, Inge Tetens, Susanne Gjedsted Bügel, Claus Felby, Simon Rønnow Schacht, James O. Hill, Eric Ravussin, and Arne Astrup. 2019. The environmental foodprint of obesity. *Obesity* 29: 73–79. [CrossRef]

Muhammad, Sulaman, Xingle Long, Muhammad Salman, and Lamini Dauda. 2020. Effect of urbanisation and international trade on CO_2 emissions across belt and road initiative countries. *Energy* 196: 11702. [CrossRef]

Our World in Data. 2022. Obesity. Available online: https://ourworldindata.org/obesity (accessed on 11 February 2022).

Ozcan, Burcu, Panayiotis G. Tzeremes, and Nickolaos G. Tzeremes. 2020. Energy consumption, economic growth and environmental degradation in OECD countries. *Economic Modelling* 84: 203–13. [CrossRef]

Pesaran, M. Hashem. 2004. *General Diagnostic Tests for Cross-Section Dependence in Panels*. Cambridge Working Papers in Economics n. 0435. Cambridge: Faculty of Economics, The University of Cambridge. [CrossRef]

Pesaran, M. Hashem. 2007. A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics* 22: 256–312. [CrossRef]

Pineda, Elisa, Luz Maria Sanchez-Romero, Martin Brown, Abbygail Jaccard, Jo Jewell, Gauden Galea, Laura Webber, and João Breda. 2018. Forecasting Future Trends in Obesity across Europe: The Value of Improving Surveillance. *Obesity Facts* 11: 360–71. [CrossRef]

Popkin, Barry M. 1998. The nutrition transition and its health implications in low-income countries. *Public Health Nutrition* 1: 5–21. [CrossRef] [PubMed]

Reardon, Thomas, C. Peter Timmer, Christopher B. Barrett, and Julio Berdegué. 2003. The rise of supermarkets in Africa, Asia, and Latin America. *American. Journal of Agricultural Economics* 85: 1140–46. Available online: https://www.jstor.org/stable/1244885 (accessed on 11 February 2022). [CrossRef]

Roskam, Albert-Jan R., Anton E. Kunst, Herman Van Oyen, Stefaan Demarest, Jurate Klumbiene, Enrique Regidor, and Uwe Helmert. 2010. Comparative appraisal of educational inequalities in overweight and obesity among adults in 19 European countries. *International Journal of Epidemiology* 39: 392–404. [CrossRef]

Salahuddin, Mohammad, Jeff Gow, Md Idris Ali, Md Rahat Hossain, Khaleda Shaheen Al-Azami, Delwar Akbar, and Ayfer Gedikli. 2019. Urbanization-globalization-CO_2 emissions nexus revisited: Empirical evidence from South Africa. *Heliyon* 5: 01974. [CrossRef]

Sobal, Jeffery. 2001. Commentary: Globalisation and the epidemiology of obesity. *International Journal of Epidemiology* 30: 1136–37. [CrossRef]

Springmann, Marco, H. Charles J. Godfray, Mike Rayner, and Peter Scarborough. 2016. Analysis and valuation of the health and climate change cobenefits of dietary change. *Proceedings of the National Academy of Sciences USA* 113: 4146–51. [CrossRef] [PubMed]

Swinburn, Boyd. 2019. The Obesity and Climate Change Nexus. *Obesity* 28: 1–2. [CrossRef]

Toiba, Hery, Wendy J. Umberger, and Nicholas Minot. 2015. Diet Transition and Supermarket Shopping Behaviour: Is there a link? *Bulletin of Indonesian Economic Studies* 51. [CrossRef]

Viscecchia, Rosaria, Antonio Stasi, and Maurizio Prosperi. 2012. Health and environmental benefits from combined control of obesity and climate changes. *ECOG* 15: 1527. [CrossRef]

Wang, Shaojian, Guangdong Li, and Chuanglin Fang. 2018. Urbanization, economic growth, energy consumption, and CO_2 emissions: Empirical evidence from countries with different income levels. *Renewable and Sustainable Energy Reviews* 81: 2144–59. [CrossRef]

Webb, Gary James, and Garry Egger. 2013. Obesity and Climate Change: Can We Link the Two and Can We Deal With Both Together? *American Journal of Lifestyle Medicine*, 1–12. [CrossRef]

World Bank Open Data. 2022. Available online: http://www.worldbank.org/ (accessed on 11 February 2022).

Wright, Suzanne M., and Louis J. Aronne. 2012. Causes of obesity. *Abdominal Radiology* 37: 730–32. [CrossRef] [PubMed]

Yazdi, Soheila Khoshnevis, and Anahita Golestani Dariani. 2019. CO_2 emissions, urbanisation and economic growth: Evidence from Asian countries. *Economic Research* 32: 1–22. [CrossRef]

economies

MDPI

Review

A Review Research on Tourism in the Green Economy

Diego R. Toubes [1,*] and Noelia Araújo-Vila [2]

1 Department of Business Organization, Business Administration and Tourism School, University of Vigo, 32004 Ourense, Spain
2 Financial Economics and Accounting Department, University of Vigo, 32004 Ourense, Spain; naraujo@uvigo.es
* Correspondence: drtoubes@uvigo.es

Abstract: Following the publication of the 2011 United Nations Environment Programme (UNEP) Green Economy Report, the UNWTO published the Tourism in the Green Economy report as clear evidence of the tourism sector's growing influence on green economy. This paper provides the first narrative review on green tourism economy based on all available published research. The paper synthesizes existing knowledge, reviews responses and initiatives on green tourism, and suggests future research and methodological approaches to help advance this field. The Scopus and WoS databases were used for the bibliometric study. The results are structured into three sections—challenges, opportunities, and enabling conditions—which comprehensively cover the main themes.

Keywords: green economy; tourism; sustainable development; hospitality; energy

Citation: Toubes, Diego R., and Noelia Araújo-Vila. 2022. A Review Research on Tourism in the Green Economy. *Economies* 10: 137. https://doi.org/10.3390/economies10060137

Academic Editors: Ralf Fendel, Sajid Anwar and Robert Czudaj

Received: 11 May 2022
Accepted: 6 June 2022
Published: 9 June 2022

1. Introduction

The 1992 United Nations (UN) Conference on Environment and Development held in Rio de Janeiro formally adopted the sustainable development concept that the 1987 Brundtland Commission Report defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (Brundtland and Khalid 1987). Twenty years later, the UN's Rio+20 conference drew attention to the green economy approach, or more precisely, how economies can achieve "green growth" to transition away from business as usual, take action to end poverty, address environmental destruction and build a bridge to a sustainable future (Barbier 2012). The green economy concept that has entered the mainstream of policy debates does not replace that of sustainable development but rather encompasses it, expressing the need for the right economy to achieve sustainable development (Loiseau et al. 2016; UNEP 2011). Green economy is an umbrella concept that includes elements from areas such as circular economy and bioeconomy. Some of the most cited articles in the field of green economy are included in the category of bioeconomy, for example, D'Amato et al. (2017), which to date, is cited 327 times. Specifically, circular economy and bioeconomy focus on resources, while the green economy encompasses all processes and incorporates the social dimension, for example, aspects relating to local communities and ecotourism.

Tourism could significantly drive global economic growth. Before the COVID-19 pandemic, in 2019, travel and tourism was one of the world's largest sectors, accounting for 10.4% of the global GDP (USD 9.2 trillion), 10.6% of all jobs (334 million), and was responsible for creating one in four of all new jobs across the world (WTTC 2021). Moreover, the transversal nature and connection of tourism with many sectors in the destinations where it operates, and its global scope, mean that changes in tourism practices towards greening may have significant effects beyond the direct and immediate impact on tourism.

"Tourism in the green economy refers to tourism activities that can be maintained, or sustained, indefinitely in their social, economic, cultural, and environmental contexts:

sustainable tourism" (UNWTO 2012). Sustainable tourism is not a particular form of tourism; rather, all forms of tourism can strive to address this approach, taking into account the needs of visitors, industry, the environment and host communities (OECD 2020).

There is growing recognition of the tourism sector's potential contribution to the green economy through more sustainable practices, climate change mitigation and ecotourism (Reddy and Wilkes 2015). Greening tourism involves policies and programmes that take into account not only tourists' expectations for the responsible management of natural resources but also the needs of the communities that tourism schemes support or affect. Therefore, sustainable tourism aims to increase efficiency in the use and conservation of energy, to minimize waste and to increase awareness and support for the sustainable use of natural resources such as water, to conserve biodiversity, cultural heritage and traditional values, and to generate local income and strengthen local communities with a view to improving livelihoods and reducing poverty (UNWTO 2012).

The contribution and influence of the tourism industry on the green economy is gaining attention from academics, practitioners and policymakers. In this sense, research and the scientific literature have produced a substantial and growing body of knowledge about tourism in the green economy that highlights the challenges and opportunities that open up to the green economy and its specific applications.

In this context, this paper examines and evaluates the existing literature on tourism in the green economy. Through a systematic literature review, 327 articles from the Web of Sciences (WoS) and Scopus databases are analysed and synthesized into an integrative framework that links the green economy of tourism with its background and consequences. The bibliometric analysis and the narrative review that is carried out on the articles in the sample offers a twofold contribution. First, it provides scholars and practitioners with a synthesis of the state of the art of the issue, in which the main lines of knowledge and the development towards greener tourism are indicated. Second, it helps to identify gaps and research directions that academics need to address in order to further develop the existing body of knowledge on tourism in the green economy.

2. Methodology

This paper aims to analyse academic and scientific research on tourism in the green economy. Figure 1 shows the steps of the data collection procedure for articles on green tourism economy (GTE) in the Scopus and WoS databases. Publications other than the articles such as book chapters and conference proceedings were not included in the study with the aim of analysing the most relevant works published in high-impact journals. All articles published and accepted in print proof from 1990 to 2022 were collected. A search with the string of words "green tourism economy" in the titles, abstracts and keywords yielded 238 articles on Scopus and 397 on WoS. After eliminating duplications, the total sample amounted to 327 works. The 327 papers have been published in 159 journals. Most of the articles found were published after 2010, which shows the novelty of the topic (see Table 1).

Data analysis follows two methods: (1) bibliometric analysis for which the variables of keywords and abstracts have been considered; and (2) analysis of deductive content applied to the articles' titles, abstracts and keywords and analysis of correlation of the most studied topics. The analysis process involved two main steps: developing a coding matrix and coding the data into the categories of the matrix. Words were adopted as the unit of record, and frequency was the main enumeration rule for coding. QSR NVivo 11 was used in the analysis process. Finally, an analysis of intraclass correlation coefficient (ICC) was performed using the Pearson ρ statistic. This coefficient is used as the probability of establishing a linear equation between two variables; for each unit change in one of them, a correlative unit change in the other is expected, without taking into account either the variables' magnitude or the measurement scale (Zhou et al. 2016). The ICC exists to quantify the concordance between different measurements of a numerical variable and extends its use to cases where more than two observations per subject are available. It indicates the reliability of a single measurement determined by the following expression:

$$\rho = \frac{\text{Var}\ (\pi)}{\text{Var}\ (\pi) + \text{Var}\ (\varepsilon)} \tag{1}$$

The ICC values can range from 0 to 1 so that the maximum possible agreement corresponds to a value of ICC = 1. The ICC = 0 value is obtained when the observed concordance is equal to that which would occur only by chance.

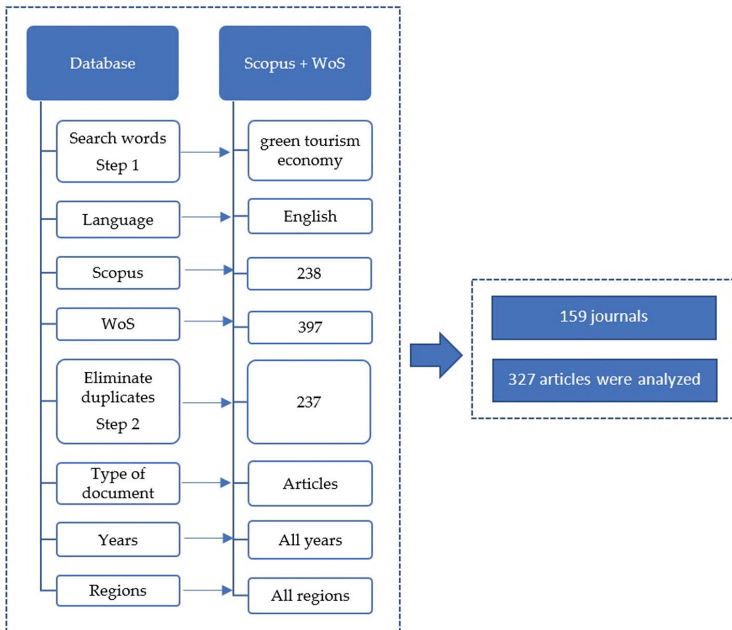

Figure 1. The data collection steps.

Table 1. Search performed in Scopus and WoS to delimit the study sample.

	Scopus		WoS	
Year	**Articles**	**%**	**Articles**	**%**
2022	13	5.5%	14	3.5%
2021	42	17.6%	91	22.9%
2020	28	11.8%	66	16.6%
2019	14	5.9%	47	11.8%
2018	24	10.1%	42	10.6%
2017	20	8.4%	32	8.1%
2016	18	7.6%	25	6.3%
2015	17	7.1%	25	6.3%
2014	10	4.2%	14	3.5%
2013	9	3.8%	9	2.3%
2012	8	3.4%	8	2.0%
1991–2011	35	14.7%	24	6.0%
Total	238		397	

3. Results

First, the most productive journals on the subject under study are analysed. Ten journals published more than five articles, with *Sustainability* standing out from the others with 36 articles on GTE. Of the group of the 10 most productive journals, 4 of them are included in the journal citation reports (JCR), all included in the second or first quartile of this ranking (see Table 2). Regarding publishers, only MDPI has two journals in this top ten.

Table 2. Journals with the largest number of publications in the area of study.

#	Journals	Cite Score	Cuartil	Articles	Category	Editorial
1	*Sustainability*	3.251	Q2 (JCR)	36	Environmental studies; Green and sustainable science and technology	MDPI
2	*Environmental Science and Pollution Research*	0.85	Q2 (SJR)	13	Environmental science	Springer Science + Business Media
3	*International Journal of Environmental Research and Public Health*	3.390	Q2 (JCR)	11	Public, environmental and occupactional health; Environmental sciences	MDPI
4	*Journal of Cleaner Production*	1.94	Q1 (SJR)	11	Business, management and accounting; Energy; Engineering; Environmental science	Elsevier Ltd.
5	*Journal of Sustainable Tourism*	7.968	Q1 (JCR)	10	Hospitality, leisure, sport and tourism; Green and sustainable science and technology	Taylor and Francis Ltd.
6	*Monthly Notices of the Royal Astronomical Society*	2.06	Q1 (SJR)	7	Earth and planetary sciences; Physics and astronomy	Oxford University Press
7	*Fresenius Environmental Bulletin*	0.18	Q4 (SJR)	6	Environmental science	Parlar Scientific Publications
8	*Astronomy Astrophycs*	-	-	5	-	
9	*Tourism Economics*	4.438	Q1 (JCR)	5	Economics; Hospitality, leisure, sport and tourism	SAGE Publications Inc.
10	*Journal of Environmental Management and Tourism*	0.3	Q3 (SJR)	5	Business, management and accounting; Economics, econometrics and finance; Environmental science	ASERS Publishing House

The authors, affiliations and countries with the most works in the sample are analysed (Table 3), as well as the most repeated keywords (Figure 2). Regarding the authors, Tsai, S. B. (China) and Walker, A. R. (United Kingdom) are the most productive, publishing five articles each. Research is highly fragmented regarding authors' affiliations; for example, the academic institutions with the highest number of published papers are University of California System (16 papers), University of Illinois System (13), University of Illinois Urbana–Champaign (13), Chinese Academy of Sciences (12) and University of Johannesburg (12). Regarding territory, China stands out in research on tourism in the green economy, with 32.1% of articles published, followed by the United States, with 25% of publications. The next countries in the ranking are England, Spain, Germany and Australia, all with more than 30 papers published (Figure 2).

Table 3. Ranking of authors and affiliations with more publications on GTE.

Authors	No. Articles	Affiliation	No. Articles
Tsai, S. B.	5	University of California System	16
Woosnam, K. M.	5	University of Illinois System	13
Law, A.	4	University of Illinois Urbana–Champaing	13
Chien, F. S.	4	Chinese Academy of Sciences	12
Filimonau, V.	4	University of Johannesburg	12
Lee, C. H.	4	Max Planck Society	11
Le, X.	4	University of London	11
Lipman, G.	4	Texas A&M University College Station	10
		Texas A&M University System	10
		University of California Berkeley	10
		University System of Georgia	10

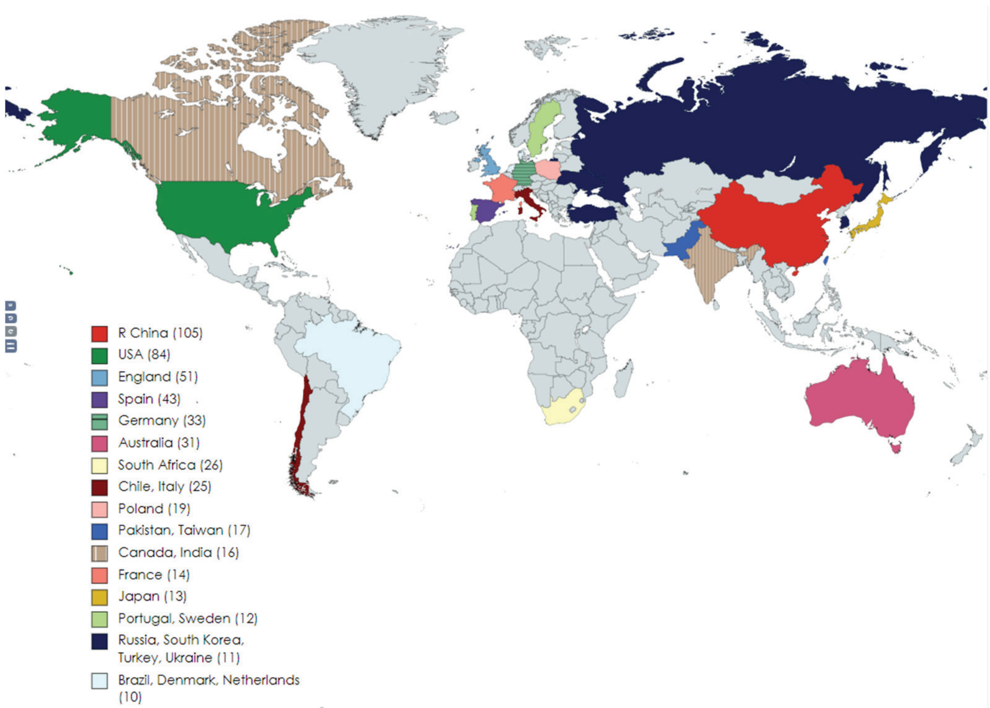

Figure 2. Ranking of countries/regions with the highest number of publications on GTE.

Table 4 shows the 10 most cited articles in the study area, with a brief summary of each, and the number of citations received at the time of writing.

Table 4. Ten most cited articles on GTE.

Article	Summary	Cites
Smetacek, V., & Zingone, A. (2013). Green and golden seaweed tides on the rise. *Nature*, 504(7478), 84–88.	Sudden beaching of huge seaweed masses smother the coastline. These 'seaweed tides' can harm tourism-based economies or disrupt traditional artisanal fisheries. Harvesting the macroalgae, a valuable raw material, before they beach could well be developed into an effective solution.	417
D'Amato, D., Droste, N., Allen, B., Kettunen, M., Lähtinen, K., Korhonen, J., . . . & Toppinen, A. (2017). Green, circular, bio economy: A comparative analysis of sustainability avenues. *Journal of Cleaner Production*, 168, 716–34.	This article comprehensively analyses the diversity within and between the concepts of Circular Economy, Green Economy and Bioeconomy. The results are drawn from a bibliometric review of almost two thousand scientific articles published within the last three decades.	327
Kang, K. H., Stein, L., Heo, C. Y., & Lee, S. (2012). Consumers' willingness to pay for green initiatives of the hotel industry. *International Journal of Hospitality Management*, 31(2), 564–72.	This paper examines hotel guests' willingness to pay a premium for environmentally friendly and sustainable practices of the U.S. hotel industry. Guests with higher degrees of environmental concerns declare a higher willingness to pay premiums for hotels' green initiatives.	311
Hjalager, A. M. (1996). Agricultural diversification into tourism: Evidence of a European Community development programme. *Tourism Management*, 17(2), 103–11.	This article discusses the impact of rural tourism on agricultural holdings. It is shown that the financial returns most often do not measure up either to the expectations of the politicians or to that of the farmers.	169
Pham, N. T., Tučková, Z., & Jabbour, C. J. C. (2019). Greening the hospitality industry: How do green human resource management practices influence organizational citizenship behavior in hotels? A mixed-methods study. *Tourism Management*, 72, 386–99.	Drawing on Ability-Motivation-Opportunity theory, this study develops and tests direct and interactive effects of green human resource management practices on organizational citizenship behavior for the environment	131
Hens, L., Block, C., Cabello-Eras, J. J., Sagastume-Gutierez, A., Garcia-Lorenzo, D., Chamorro, C., . . . & Vandecasteele, C. (2018). On the evolution of "Cleaner Production" as a concept and a practice. *Journal of Cleaner Production*, 172, 3323–33.	This paper provides a review of essentials that contributed to the fundamental changes in "Cleaner Production" (CP). The links between CP and green and circular economy are indicated, and CP for sustainable tourism is discussed in more detail.	102
Prideaux, B., Thompson, M., & Pabel, A. (2020). Lessons from COVID-19 can prepare global tourism for the economic transformation needed to combat climate change. *Tourism Geographies*, 22(3), 667–78.	For the tourism industry to thrive in a post-COVID-19 world, it must look beyond the temptation of adopting strategies based on a return to the pre-pandemic normal of the past and instead seek to understand how it should respond to the emerging transformation of the global economy to carbon neutrality.	91
Cazcarro, I., Hoekstra, A. Y., & Chóliz, J. S. (2014). The water footprint of tourism in Spain. *Tourism Management*, 40, 90–101.	This study complements the water footprint estimations for Spain, incorporating insights of the process analysis and input–output analysis. The virtual water trade of agricultural and industrial products is evaluated, but also of services, especially through tourism.	67
Park, E., & Boo, S. (2010). An assessment of convention tourism's potential contribution to environmentally sustainable growth. *Journal of Sustainable Tourism*, 18(1), 95–113.	This research assesses the current environmental position of the convention industry in the U.S. and formulates suggestions for future direction in regard to "green" concepts by examining behavioral intentions of three groups: convention attendees, meeting planners, and convention suppliers.	64
Duffy, R. (2015). Nature-based tourism and neoliberalism: Concealing contradictions. *Tourism Geographies*, 17(4), 529–43.	This paper examines the claims around tourism as 'green economy' can produce environmentally sustainable economic growth. It is discussed how nature-based tourism, simultaneously produces and conceals the contradictions of capitalism.	63

It is also interesting to identify the keywords most commonly used in order to obtain information on research areas related to GTE (Figure 3). "Tourism" is the most common keyword, followed by "economy" and "sustainable/sustainability", as well as "management", "environment" or "energy", which are terms closely linked to the subject of study.

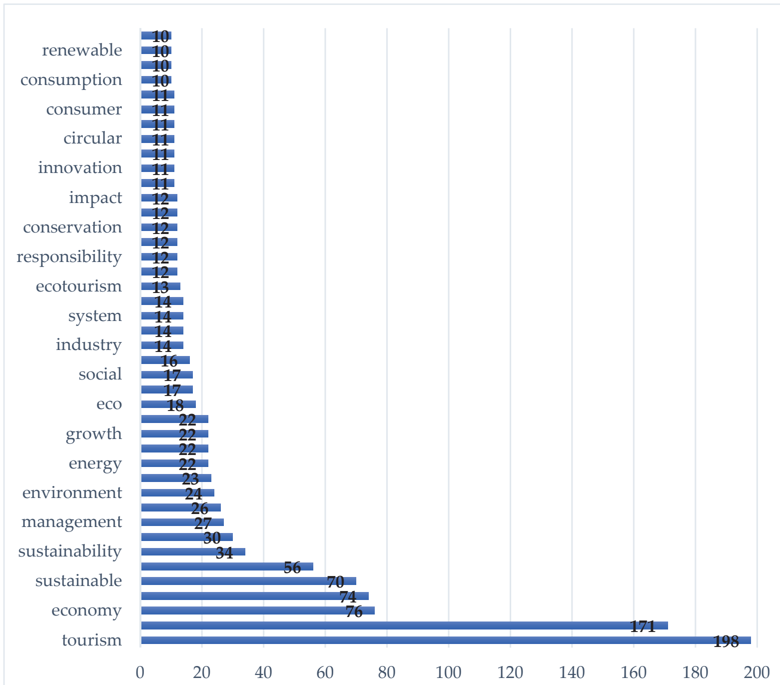

Figure 3. Most common keywords in research on GTE.

After a frequency analysis of the words in the titles of the total sample, it is observed that the main terms used are the same as in the keywords: tourism, green, economy, development, sustainable and environmental (Figure 4). Therefore, to obtain more detailed information on the subject, a content analysis of the articles' abstracts by the most prolific authors in this area is carried out (Table 5). Thus, we found that, regarding tourism in the green economy, the main topics studied focus on the following.

- Reducing energy consumption and carbon emissions in the logistics chain (air and rail transport).
- The importance of including green development in rural and mountain development policies.
- Analysis of the local population's perception of tourism, tourist rental housing and economic advantages/disadvantages that tourism brings to destinations.
- Analysis of tourists' perceptions of hotel chains implementing green measures.
- Transformation of tourism towards green economies, both in rural areas and in small and large tourism enterprises.
- Waste management in hotel chains and tourist destinations.
- Transition to a green economy in the tourism sector as well as supporting decision-making.

Figure 4. Word cloud of the articles' titles in the sample.

Table 5. Content analysis of the articles of the most prolific authors on the subject under study.

Author	Topics
Tsai, S. B.	Air transport, energy conservation, emission reduction, local economy, economic revenues, carbon emissions, energy consumption in logistics, high-speed rail network, supply chain.
Woosnam, K. M.	Perception of the local community about tourism, advantages and disadvantages, economic and non-economic factors in the perception of tourism, residents' perceptions about the holiday rental business.
Law, A.	Green economy in tourist destinations, economic crisis in rural areas, analysis of economic indicators in rural areas, transformation of tourism towards green economy.
Chien, F. S.	Economic growth and carbon emissions, role of human capital, green energy and economic growth, fiscal policies and energy efficiency, environmental awareness, barriers to green economy in small and medium-sized enterprises.
Filimonau, V.	Tourism growth and waste management, energy conservation in large hotels, food and waste management
Lee, C. H.	Development plans in mountain villages, waste management, sustainable production.
Le, X.	Guests' perceptions of green practices in hotels.
Lipman, G.	Transition to green economy, green economy as part of tourism management, green growth as support for tourism decision-making.

A more detailed analysis of keywords is also carried out, focusing on those that show some link with the green economy and tourism (Table 6). Thus, in addition to the word "tourism", the themes that have attracted the greatest interest in this field are "development", "sustainability", "rural areas", "economic growth", "energy use and consumption", "carbon consumption", "governance", "hotel business" and "green economic development in urban areas", among others.

Finally, once the topics most studied in GTE were identified, possible correlations or thematic associations between them were analysed. This methodology has been applied in previous research in other areas of economics (Porte et al. 2018). After applying Pearson's ρ statistic, the results closest to 1 indicate the strongest correlations (see Table 7). It is observed that the works dealing with tourism are about green, economy and environment. Articles that explicitly address the subject of green also explore tourism and energy. Articles focusing on rural areas address carbon emissions, management, energy consumption and ecosystems. In addition, one of the concerns of hotel management is the environment. Figure 5 shows graphically the thematic correlations of the different areas of study.

Table 6. Frequency analysis of keywords linked to tourism in the green economy (n = 2288). Ranking of words with frequency greater than 7.

Rank	Keyword	Frequency	Rank	Keyword	Frequency
1º	tourism	143	23º	change	11
2º	green	136	24º	system	11
3º	economy	61	25º	industry	11
4º	development	54	26º	climate	11
5º	sustainable	50	27º	blue	10
6º	environmental	37	28º	analysis	10
7º	sustainability	26	29º	circular	9
8º	rural	23	30º	city	9
9º	economic	22	31º	innovation	9
10º	environment	18	32º	conservation	8
11º	growth	17	33º	land	7
12º	carbon	17	34º	impact	7
13º	management	16	35º	emissions	7
14º	energy	16	36º	waste	7
15º	services	15	37º	resources	7
16º	ecological	14	38º	data	7
17º	eco	14	39º	panel	7
18º	hotel	14	40º	marketing	7
19º	urban	14	41º	tourist	7
20º	ecosystem	14	42º	renewable	7
21º	ecotourism	13	43º	index	7
22º	model	12			

Table 7. Correlations between the most common research topics in GTE.

	T	G	EC	DE	SU	EN	SUST	RU	ECO	ENV	GRO	CAR	MAN	ENE	SER	ECOL	HO	UR	ECOS	ECOT
T	1.00	1.00	1.00	0.99	0.96	1.00	0.97	0.98	0.97	0.96	0.94	0.97	0.99	0.99	0.98	0.98	0.96	0.97	0.97	0.61
G	1.00	1.00	0.99	0.99	0.94	0.99	0.97	0.99	0.98	0.94	0.94	0.98	1.00	1.00	0.99	0.99	0.94	0.99	0.99	0.56
EC	1.00	0.99	1.00	0.99	0.98	1.00	0.97	0.96	0.96	−0.06	0.93	0.95	0.97	0.97	0.96	0.97	0.98	0.95	0.95	0.67
DE	0.99	0.99	0.99	1.00	0.95	0.99	0.99	0.97	0.93	0.93	0.98	0.96	0.97	0.97	0.97	0.97	0.93	0.97	0.97	0.55
SU	0.82	0.94	0.98	0.95	1.00	0.97	0.97	0.88	0.91	0.99	0.89	0.86	0.90	0.90	0.89	0.90	0.99	0.87	0.87	0.77
EN	0.96	0.99	1.00	0.99	0.97	1.00	0.97	0.97	0.96	0.96	0.94	0.96	0.98	0.98	0.97	0.98	0.97	0.96	0.96	0.64
SUST	0.97	0.97	0.97	0.99	0.97	0.97	1.00	0.95	0.89	0.89	0.99	095	0.95	0.95	0.96	0.94	0.89	0.96	0.96	0.48
RU	0.98	0.99	0.96	0.97	0.88	0.97	0.95	1.00	0.97	0.89	0.93	1.00	1.00	1.00	1.00	1.00	0.89	1.00	1.00	0.47
ECO	0.97	0.98	0.96	0.93	0.91	0.96	0.89	0.97	1.00	0.95	0.84	0.96	0.98	0.98	0.97	0.98	0.95	0.95	0.95	0.64
ENV	0.96	0.94	−0.06	0.93	0.99	0.96	0.89	0.89	0.95	1.00	0.83	0.87	0.91	0.91	0.89	0.91	1.00	0.86	0.86	0.82
GRO	0.94	0.94	0.93	0.98	0.89	0.94	0.99	0.93	0.84	0.83	1.00	0.93	0.92	0.92	0.93	0.91	0.83	0.95	0.95	0.39
CAR	0.97	0.98	0.95	0.96	0.86	0.96	0.95	1.00	0.96	0.87	0.93	1.00	0.99	0.99	1.00	0.99	0.87	1.00	1.00	0.43
MAN	0.99	1.00	0.97	0.97	0.90	0.98	0.95	1.00	0.98	0.91	0.92	0.99	1.00	1.00	1.00	1.00	0.92	0.99	0.99	0.52
ENE	0.99	1.00	0.97	0.97	0.90	0.98	0.95	1.00	0.98	0.91	0.92	0.99	1.00	1.00	1.00	1.00	0.92	0.99	0.99	0.52
SER	0.98	0.99	0.96	0.97	0.89	0.97	0.96	1.00	0.97	0.89	0.93	1.00	1.00	1.00	1.00	1.00	0.90	1.00	1.00	0.47
ECO	0.98	0.99	0.97	0.97	0.90	0.98	0.94	1.00	0.98	0.91	0.91	0.99	1.00	1.00	1.00	1.00	0.92	0.99	0.99	0.53
HO	0.96	0.94	0.98	0.93	0.99	0.97	0.89	0.89	0.95	1.00	0.83	0.87	0.92	0.92	0.90	0.92	1.00	0.87	0.87	0.81
UR	0.97	0.99	0.95	0.97	0.87	0.96	0.96	1.00	0.95	0.86	0.95	1.00	0.99	0.99	1.00	0.99	0.87	1.00	1.00	0.42
ECOS	0.97	0.99	0.95	0.97	0.87	0.96	0.96	1.00	0.95	0.86	0.95	1.00	0.99	0.99	1.00	0.99	0.87	1.00	1.00	0.42
ECOT	0.61	0.56	0.67	0.55	0.77	0.64	0.48	0.47	0.64	0.82	0.39	0.43	0.52	0.52	0.47	0.53	0.81	0.42	0.42	1.00

T: tourism; G: green; EC: economy; DE: development; SU: sustainable; EN: environmental; SUST: sustainability; RU: rural; ECO: economic; ENV: environment; GRO: growth; CAR: carbon; MAN: management; ENE: energy; SER: services; ECOL: ecological; HO: hotel; UR: urban; ECOS: ecosystem; ECOT: ecotourism.

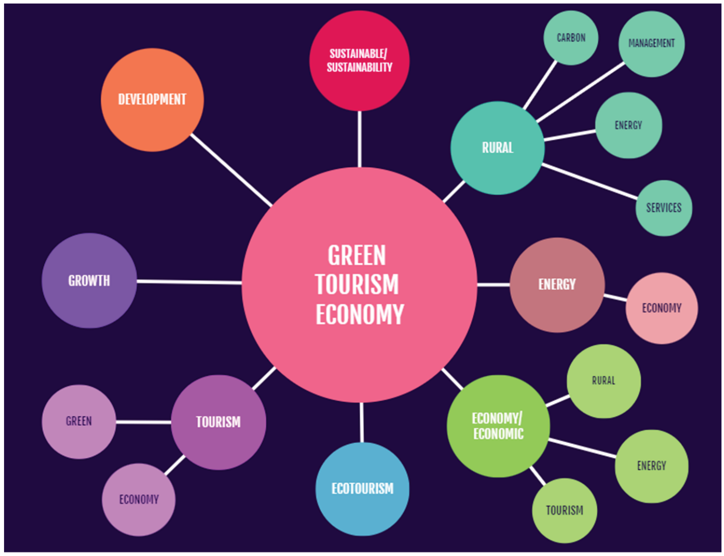

Figure 5. Main thematic correlations in the area of study.

4. Discussion and Conclusions

After the bibliometric study, a narrative review was carried out based on the 327 articles published on GTE. We explain the existing knowledge on the subject and discuss the findings presented in the most recent research work. We can group the articles that have addressed the topic into three broad areas.

The first group of articles address the contribution and interaction of tourism with the green economy as a real challenge. The role of tourism in the green economy is in its early stages, and consequently, research in this area is still very recent. For example, about 80% of the sample articles in this study were published between 2015 and 2022. The tourism industry faces a number of challenges that it must address to move towards a green economy. Some of these challenges are common to other productive areas, such as energy consumption, waste management and carbon emissions, but others are clearly specific to the tourism sector, such as those related to passenger transport or the management and use of natural resources.

The second group of articles that have been identified in the narrative review addresses the opportunities for green economy commitments, both for the tourism sector itself and for strengthening the roadmap to a green economy in general. In this sense, the global scope of the tourism sector and its transversal nature makes it an enabler for others in transforming the economy and the green transition. Conversely, a significant number of studies address the potential of the green economy applied to tourism for social development and ultimately for sustainable development of local communities. For example, 28.4% of the studies in the sample address the sustainable development as one of the main issues. Many of these studies (13.1% of the articles) focus on the rural context, thus pointing out the opportunities for rural tourism companies to adopt green economy practices.

The third largest group of articles are those regarding issues related to the enabling elements or conditions for implementing the green economy in tourism. In this respect, the papers emphasize the involvement and commitment of multi-stakeholders, for example, tourists' perceptions and their involvement in using green actions, cooperation structures between the different actors, the use of quality certification, governance measures, legal and planning aspects and in general, all types of policies aimed at the main objective of establishing a green economy in the tourism sector. In this sense, the practical attitude of the research is evident; for example, 43% of the articles analysed contain policy proposals of various kinds—fiscal, legal, financial and social.

Below we carry out a more detailed analysis of these three fields.

4.1. Challenges

Five key challenges have been identified for tourism to meet the proposed objectives of a green economy. These challenges directly relate to the natural environment, and given the strong human behaviour component of tourism, the different possibilities for social interaction will bring benefits or potential damage to terrestrial and marine biodiversity. These challenges are as follows:

1. energy consumption and treatment and adopting renewable energy sources;
2. studies related to greenhouse gas (GHG) emissions and the main role of tourism in the passenger transport industry;
3. efficient waste management;
4. excessive water consumption; and
5. emerging risks and vulnerabilities climate change causes.

Energy consumption and treatment is a key issue within the challenges of the green economy in tourism. Energy is a relevant concept in about 35% of the articles in the sample. Specifically, the concepts of "energy consumption", "renewable energy", "energy conservation" and "energy efficiency" are addressed in 11.6% of the articles. Numerous articles address the challenge of saving and conserving energy in tourism, which is closely related to the analysis of carbon emissions and low-carbon technology schemes (Lu et al. 2018; Ma et al. 2022; Salehi et al. 2021). Energy saving analyses are applied both at the general level and in specific areas such as rural areas (Dzhusibalieva et al. 2016; Majdak and de Almeida 2022; Llinàs et al. 2021; Li et al. 2005). In general, green energy plays a leading role in transitioning to a green economy (Potts et al. 2019), particularly in tourist destinations in emerging economies (Tazikhina et al. 2022; Huang et al. 2021). Several articles study various initiatives in the adoption of renewable energy sources (Khan et al. 2022; Janjua et al. 2021; Lucrezi and Saayman 2017); with 102 citations, Hens et al. (2018) authored the most cited articles in the area. On the demand side, the behaviour of energy consumption is also a frequently analysed aspect (Ohajionu et al. 2022), with specific studies carried out on energy consumption during the COVID-19 pandemic (Strielkowski et al. 2021).

As we have mentioned, energy conservation analyses are closely related to studies of carbon emissions. Data on GHG emissions from tourism was identified as one of the critical gaps (Law et al. 2016, 2017). In this sense, it is worth highlighting the studies on the effects of tourism development and green innovation on economic growth (Ghosh 2022; Razzaq et al. 2021; Deng et al. 2020). Tourism's main role in transport and the problems of mobility and infrastructure resulting from the intense international flow of passengers is also one of the outstanding challenges the research addresses (Scorza and Fortunato 2021; Sharif et al. 2020; Salvino et al. 2018). A number of papers address CO_2 emissions related to traffic and transport (Lee et al. 2018; Ma et al. 2022) and cruise tourism (Ye et al. 2019).

Several studies explore mass tourism's responsibility for polluting nature reserves and coastal areas and the damage it causes to biological and marine ecosystems. In this sense, efficient waste management is one of the main challenges to face (Mena-Nieto et al. 2021; Voukkali et al. 2021); this problem is more evident in developing countries, which are highly dependent on tourism (Manomaivibool 2015; Filimonau and Tochukwu 2020). Catering businesses face the problem of food waste management; several studies analyse the determinants of consumers and suppliers' food waste mitigation (Filimonau et al. 2019, 2020). Another major challenge is managing excessive water consumption. Tourism is often a water-intensive sector (Mimbs et al. 2020; Hause 2016) with an important water footprint (Cazcarro et al. 2014). Managing water conservation and preservation is present in recent studies (Torres-Bagur et al. 2020; Gabarda-Mallorquí et al. 2018).

Finally, the studies analysed address the emerging risks and vulnerabilities climate change causes. The anthropogenic origin of GHG emissions places tourism as an important sector regarding its effect on nature and biodiversity, and the potential risks to society and local ecosystems (Lee et al. 2018; Karani and Failler 2020; Cashman et al. 2012). Regarding adaptation, economic policies and the development of management plans should help to provide the most exposed stakeholders with the tools to minimise threats and capitalise on opportunities in a sustainable way in the three areas—economic, social and environmental (Prideaux et al. 2020; Pearce and Schott 2010; Olcina Cantos and Vera-Rebollo 2016).

4.2. Opportunities

There is a very close relationship between the green economy and the concept of sustainable development, which incorporates aspects of ecological, social and economic development that do not harm the biological diversity and environmental quality of the tourist destination and the socio-cultural fabric of the host community (UNWTO 2012). In line with the green economy approach, gains from using protected areas should be reinvested in local livelihoods and employment (Catibog-Sinha 2015).

The tourism industry is a favourable sector for proposing the change to a green economy, and it can even be considered as a standard bearer of the transition process in search of wellbeing opportunities and possibilities for sustainable development (Prideaux et al. 2020; Arnedo et al. 2021; McKinley et al. 2019; Sevastiyanov et al. 2014; Law et al. 2012; Mestanza-Ramón et al. 2019). Green tourism can ensure viable long-term economic operations in host communities in ways that bring wealth to all stakeholders and that socio-economic benefits are fairly distributed. This implies stable employment, income-generating opportunities and social services that contribute to poverty alleviation. The closed loop perspective applied to community-based tourism has a strong potential to facilitate the transformation towards a green economy as closed loop systems can lead to significant efficiencies in production and sustainable consumption of resources.

Numerous studies in the area of tourism in different locations focus on the efforts to be made in the process of transitioning to a green economy (Sharif et al. 2020; Ahmad et al. 2020; Pongsakornrungsilp and Pongsakornrungsilp 2021; Pimonenko et al. 2021). This transformation is not straightforward and requires a number of challenges for tourism stakeholders. Various articles propose how to reconcile the vocation of tourism's continuous growth with the strengthening towards an economy of sustainable development. In this sense, there is a need for a much more radical change (Law et al. 2016; Stroebel 2015; Denona Bogovic and Grdic 2020; Hein 2014).

The term sustainability is found in the topic of 55% of the sample studies (title, abstract and keywords), and the concept "sustainable development" is in 28.3% of the articles. The focus on the economy is clearly seen as an opportunity for social development, as evidenced by the abundant number of studies focusing on local communities in various world regions. Specifically, the number of studies carried out in rural areas and oriented towards poverty alleviation is relevant (Drăgoi et al. 2017; Hoefle 2016; de Sousa and Kastenholz 2015; Anand et al. 2012; Hashimoto and Telfer 2010; Ye et al. 2003; Holm et al. 2013).

4.3. Enabling Conditions

Realising tourism's potential in the green economy requires a number of enabling mechanisms that provide appropriate instruments for achieving sustainable development. Government administrations' involvement through implementing development plans and strategies is key in this process. In this area, the narrative review highlights the importance of studies on green governance, planning and roadmap strategy design (Law et al. 2016; Koide and Akenji 2017; Loia et al. 2021) and the development of green marketing practices in tourism (Hanna et al. 2018; Temperini et al. 2012).

Regarding the private sector, the increasing importance of corporate social responsibility exemplifies the leading role of tourism businesses in the goal of sustainable development. This is especially the case for small businesses, as the predominance of small and medium-sized enterprises in tourism, their central role in human activities and their growing importance in sustainable tourism development suggest that these entities can help tourist destinations move towards sustainability goals (Toubes et al. 2021).

A considerable number of research studies focus on the adoption of quality certifications (eco-certification and eco-labelling programmes). These certificates follow certain standards; for example, the ISO has established a number of general global quality and environmental management standards for manufacturing and service industries that are also applied in the hospitality industry, such as ISO 9001 (quality management), ISO 14001 (environmental management), OHSAS 18001 (occupational health and safety) and ISO 22000 (food safety standard). In general, the benefits of green certification in tourism to environmentally sustainable growth are acknowledged in the research (Lebe and Vrečko 2015), mainly in the application of eco-certification in international tourist hotels (Chen 2019; Bandara et al. 2018; Abdou et al. 2020; Trišić et al. 2021). Voluntary compliance of tourism businesses with sustainability certification is also a way to gain competitiveness against local competitors (Panzer-Krause 2017; DeBoer et al. 2017).

Several studies address consumers and users' perceptions regarding the application of the requirements of a green economy (Lucrezi et al. 2017; Lee et al. 2016; Tao and Chen 2016). This research explores the adoption of green measures in hotels and holiday apartments, the growth in consumer demand for green tourism products and services and tourists' willingness to pay an extra for these services (Nelson et al. 2021; Fudurich and Mackay 2020). The role of green human resource practices (training, empowerment and rewarding for pro-environmental behaviors) in fostering employees' green recovery performance is also a key element that makes an appreciable contribution to maintaining or restoring environmental quality (Luu 2018; Pham et al. 2019; Nhamo 2010)].

The green economy could contribute to human wellbeing by fostering a more holistic approach to tourism development, quality oriented, cautious and compliant with social and environmental factors, more humanized and committed to protecting the cultural and natural values of the territories. The implementation of a new model of sustainable and circular tourism will only be possible with the involvement of all the actors that make up the economy and are interconnected: public administrations, businesses, tourists and the local community.

4.4. Gaps in Existing Studies for Potential Future Research

A central debate is on the compatibility of a green economy's objective with indiscriminate economic growth. Some authors argue for the need for major changes and a global economic structure that ensures environmental boundaries are not transgressed (Stroebel 2015). A genuine green economic future for tourism is based on the idea of steady-state economics. Steady-state tourism is a tourism system that encourages qualitative development but not aggregate quantitative growth to the detriment of natural capital (Hall 2010). According to this approach, the focus of responding to climate and environmental change must be on development, which is a measure of quality, rather than growth, which is a measure of quantitative change. The COVID-19 epidemic has highlighted the greater need to address changes to achieve a true green economy (Gössling et al. 2020).

As early as 1973, Schumacher (1973) concluded that government efforts should focus on sustainable development. The bioregional economy also focuses on strengthening local values and resources, such as local businesses, training local workers, improving the natural and cultural environment, to achieve a better place to live (Cato 2013). Holden (2013) argued that routine measures are not enough to address the environmental challenges facing the tourism industry, and that without a stronger environmental ethic in the market, it will be difficult to impose controls on tourist behaviour designed for environmental conservation. Therefore, there is a need for further research on ethical issues and to deepen the debate on social justice, inequities in global trade and the risk of "greenwashing" (Hall 2015; Newton 2015).

Deepening the sustainability challenges and effects of climate change is also an important area for future research. Specifically, the urgency of implementing climate change mitigation and adaptation measures to minimize the harmful impact on ecosystems and biodiversity on which personal and economic well-being depends. There is strong evidence that the global tourism system is highly wasteful of resources and that significant mitigation efforts in addressing climate change would have no overall negative economic effect (Gössling 2020). Understanding the implications of climate change for a destination requires developing an integrated systems approach that holistically considers supply and demand effects and adaptation options. Future research must also consider how climate change will interact with other major influencing variables in the tourism sector, such as fuel prices, increasing travel safety and health concerns, aging populations in industrialised countries, increased environmental and cultural awareness, advances in transportation technology and environmental constraints, water supply and pollution, among others.

Finally, significant gaps in knowledge remain about how environmental effects and measures are perceived. We find that the study of tourists and visitors' environmental behaviour is a research area that needs to be addressed in order to more accurately estimate potential changes in long-term tourism demand.

Author Contributions: Conceptualization, D.R.T.; methodology, N.A.-V.; validation, D.R.T.; formal analysis, D.R.T.; investigation, D.R.T. and N.A.-V.; resources, N.A.-V.; data curation, N.A.-V.; writing—original draft preparation, D.R.T.; writing—review and editing, D.R.T. and N.A.-V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

References

Abdou, Ahmed, Thowayeb Hassan, and Mohammed El Dief. 2020. A Description of Green Hotel Practices and Their Role in Achieving Sustainable Development. *Sustainability* 12: 9624. [CrossRef]

Ahmad, Norshafinas Binti, Mohamad Rosni Othman, Mohd Saiful Izwaan Saadon, and Dina Azleema Mohamed Nor. 2020. Sustainable development goal of the recreation port: The case study of the Duyong marina & resort, Terengganu, Malaysia. *Journal of Critical Reviews* 7: 1449–54.

Anand, Anupam, Pankaj Chandan, and Ram Babu Singh. 2012. Homestays at Korzok: Supplementing Rural Livelihoods and Supporting Green Tourism in the Indian Himalayas. *Mountain Research and Development* 32: 126–36. [CrossRef]

Arnedo, Esther González, Jesús Alberto Valero-Matas, and Antonio Sánchez-Bayón. 2021. Spanish Tourist Sector Sustainability: Recovery Plan, Green Jobs and Wellbeing Opportunity. *Sustainability* 13: 11447. [CrossRef]

Bandara, Chamali, D. M. P. P. Dissanayake, Gayani Karunasena, and Nadun Madhusanka. 2018. Mitigation of challenges in sustaining green certification in the Sri Lankan hotel sector. *Built Environment Project and Asset Management* 8: 515–27. [CrossRef]

Barbier, Edward B. 2012. The Green Economy Post Rio+20. *Science* 338: 887–88. [CrossRef]

Brundtland, Gro Harlem, and Mansour Khalid. 1987. *Our Common Future*. Oxford: World Commission on Environment and Development.

Cashman, Adrian, Janice Cumberbatch, and Winston Moore. 2012. The effects of climate change on tourism in small states: Evidence from the Barbados case. *Tourism Review* 67: 17–29. [CrossRef]

Catibog-Sinha, Corazon. 2015. The role of nature-based tourism. A broader perspective for conservation and sustainability in the Philippines. In *Tourism in the Green Economy*. Edited by Maharaj Vijay Reddy and Keith Wilkes. London and New York: Routledge, pp. 57–70.

Cato, Molly Scott. 2013. *The Bioregional Economy: Land, Liberty and Pursuit of Happiness*. Abingdon: Routledge.

Cazcarro, Ignacio, Arjen Y. Hoekstra, and Julio Sánchez-Chóliz. 2014. The water footprint of tourism in Spain. *Tourism Management* 40: 90–101. [CrossRef]

Chen, Long-Fei. 2019. Green certification, e-commerce, and low-carbon economy for international tourist hotels. *Environmental Science and Pollution Research* 26: 17965–73. [CrossRef]

D'Amato, Dalia, Nils Droste, B. Allen, M. Kettunen, Katja Lähtinen, Jaana Korhonen, P. Leskinen, Brent Matthies, and Anne Toppinen. 2017. Green, circular, bio economy: A comparative analysis of sustainability avenues. *Journal of Cleaner Production* 168: 716–34. [CrossRef]

de Sousa, Ana João Gaspar, and Elizabeth Kastenholz. 2015. Wind farms and the rural tourism experience—Problem or possible productive integration? The views of visitors and residents of a Portuguese village. *Journal of Sustainable Tourism* 23: 1236–56. [CrossRef]

DeBoer, Jennifer, Rajat Panwar, and Jorge Rivera. 2017. Toward A Place-Based Understanding of Business Sustainability: The Role of Green Competitors and Green Locales in Firms' Voluntary Environmental Engagement. *Business Strategy and the Environment* 26: 940–55. [CrossRef]

Deng, Qizhong, Xiaohui He, and Keming Yu. 2020. Re-evaluating the load and identifying the influencing factors of tourism's carbon emissions in major Chinese cities. *Journal of Environmental Protection and Ecology* 21: 379–89.

Denona Bogovic, Nada, and Zvonimira Grdic. 2020. Transitioning to a Green Economy—Possible Effects on the Croatian Economy. *Sustainability* 12: 9342. [CrossRef]

Drăgoi, Mihaela Cristina, Irina-Eugenia Iamandi, Sebastian Madalin Munteanu, Radu Ciobanu, Ramona Iulia Dieaconescu, Lădaru Țarțavulea, and Raluca Georgiana. 2017. Incentives for Developing Resilient Agritourism Entrepreneurship in Rural Communities in Romania in a European Context. *Sustainability* 9: 2205. [CrossRef]

Dzhusibalieva, Aruzhan, Galiya Mauina, Meiramkul Kasimbekova, Saule Serikbayeva, Kuralay Balginova, Gaukhar Zhanibekova, Lazzat Dusembaeva, Aigul Tasmambetova, Toty Bekzhanova, Aliya Aktymbayeva, and et al. 2016. Prospects for a green economy in tourism. *Journal of Internet Banking and Commerce* 21: 228.

Filimonau, Viachaslau, and Cynthia Oluchi Tochukwu. 2020. Exploring managerial approaches to mitigating solid waste in hotels of Lagos, Nigeria. *Journal of Cleaner Production* 270: 122410. [CrossRef]

Filimonau, Viachaslau, Hafize Fidan, Iordanka Alexieva, Stefan Dragoev, and Denitsa Dimitrova Marinova. 2019. Restaurant food waste and the determinants of its effective management in Bulgaria: An exploratory case study of restaurants in Plovdiv. *Tourism Management Perspectives* 32: 100577. [CrossRef]

Filimonau, Viachaslau, Jorge Matute, Magdalena Kubal-Czerwińska, Kinga Krzesiwo, and Mirosław Mika. 2020. The determinants of consumer engagement in restaurant food waste mitigation in Poland: An exploratory study. *Journal of Cleaner Production* 247: 119105. [CrossRef]

Fudurich, Geoffrey, and Kelly Mackay. 2020. Greening the guest bedroom: Exploring hosts' perspectives on environmental certification for Airbnb. *Tourism and Hospitality Research* 20: 423–34. [CrossRef]

Gabarda-Mallorquí, Fraguell Ariadna, Rosa Maria, and Anna Ribas. 2018. Exploring Environmental Awareness and Behavior among Guests at Hotels That Apply Water-Saving Measures. *Sustainability* 10: 1305. [CrossRef]

Ghosh, Sudeshna. 2022. Effects of tourism on carbon dioxide emissions, a panel causality analysis with new data sets. *Environment, Development and Sustainability* 24: 3884–906. [CrossRef]

Gössling, Stefan. 2020. *Carbon Management in Tourism: Mitigating the Impacts on Climate Change*. London: Routledge.

Gössling, Stefan, Daniel Scott, and C. Michael Hall. 2020. Pandemics, tourism and global change: A rapid assessment of COVID-19. *Journal of Sustainable Tourism* 29: 1–20. [CrossRef]

Hall, C. Michael. 2010. Changing Paradigms and Global Change: From Sustainable to Steady-state Tourism. *Tourism Recreation Research* 35: 131–43. [CrossRef]

Hall, C. Michael. 2015. Economic greenwash. On the absurdity of tourism and green growth. In *Tourism in the Green Economy*. Edited by Maharaj Vijay Reddy and Keith Wilkes. Abingdon and Oxford: Routledge, pp. 339–58.

Hanna, Paul, Xavier Font, Caroline Scarles, Clare Weeden, and Charlotte Harrison. 2018. Tourist destination marketing: From sustainability myopia to memorable experiences. *Journal of Destination Marketing & Management* 9: 36–43. [CrossRef]

Hashimoto, Atsuko, and David J. Telfer. 2010. Developing sustainable partnerships in rural tourism: The case of Oita, Japan. *Journal of Policy Research in Tourism, Leisure and Events* 2: 165–83. [CrossRef]

Hause, Heather. 2016. Fountains of youth: Medical tourism in Green Cove Springs, Florida, 1845–1900. *Journal of Tourism History* 8: 1–21. [CrossRef]

Hein, Carola. 2014. Port cities and urban wealth: Between global networks and local transformations. *International Journal of Global Environmental Issues* 13: 339. [CrossRef]

Hens, Luc, Chantal Block, Juan José Cabello-Eras, Alexis Sagastume-Gutierez, D. Garcia-Lorenzo, C. Chamorro, Ketty Herrera Mendoza, Dries Haeseldonckx, and Carlo Vandecasteele. 2018. On the evolution of "Cleaner Production" as a concept and a practice. *Journal of Cleaner Production* 172: 3323–33. [CrossRef]

Hoefle, Scott William. 2016. Multi-functionality, juxtaposition and conflict in the Central Amazon: Will tourism contribute to rural livelihoods and save the rainforest? *Journal of Rural Studies* 44: 24–36. [CrossRef]

Holden, Andrew. 2013. Tourism and the Green Economy: A Place for an Environmental Ethic? *Tourism Recreation Research* 38: 3–13. [CrossRef]

Holm, Jesper, Line Maria Bram Pedersen, and Søren Sørensen. 2013. Experiencing organic farms and food by regional tourism innovation. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science* 63: 85–93. [CrossRef]

Huang, Shi-Zheng, Fengsheng Chien, and Muhammad Sadiq. 2021. A gateway towards a sustainable environment in emerging countries: The nexus between green energy and human Capital. *Economic Research-Ekonomska Istraživanja* 1–18. [CrossRef]

Janjua, Laeeq Razzak, Fida Muhammad, Patthaphon Sukjai, Abdul Rehman, and Zhang Yu. 2021. Impact of COVID-19 pandemic on logistics performance, economic growth and tourism industry of Thailand: An empirical forecasting using ARIMA. *Brazilian Journal of Operations and Production Management* 18: e2021999. [CrossRef]

Karani, Patrick, and Pierre Failler. 2020. Comparative coastal and marine tourism, climate change, and the blue economy in African Large Marine Ecosystems. *Environmental Development* 36: 100572. [CrossRef]

Khan, Syed Abdul Rehman, Danish Iqbal Godil, Zhang Yu, Farwa Abbas, and Muhammad Asif Shamim. 2022. Adoption of renewable energy sources, low-carbon initiatives, and advanced logistical infrastructure—An step toward integrated global progress. *Sustainable Development* 30: 275–88. [CrossRef]

Koide, Ryu, and Lewis Akenji. 2017. Assessment of Policy Integration of Sustainable Consumption and Production into National Policies. *Resources* 6: 48. [CrossRef]

Law, Alexandra, Terry De Lacy, G. Michael McGrath, Paul A. Whitelaw, Geoffrey Lipman, and Geoff Buckley. 2012. Towards a green economy decision support system for tourism destinations. *Journal of Sustainable Tourism* 20: 823–43. [CrossRef]

Law, Alexandra, Terry De Lacy, Geoffrey Lipman, and Min Jiang. 2016. Transitioning to a green economy: The case of tourism in Bali, Indonesia. *Journal of Cleaner Production* 111: 295–305. [CrossRef]

Law, Alexandra, Terry De Lacy, and G. Michael McGrath. 2017. A green economy indicator framework for tourism destinations. *Journal of Sustainable Tourism* 25: 1434–55. [CrossRef]

Lebe, Sonja Sibila, and Igor Vrečko. 2015. Eco-labels and Schemes: A Requisitely Holistic Proof of Tourism's Social Responsibility? *Systems Research and Behavioral Science* 32: 247–55. [CrossRef]

Lee, Hyeryeon, Tun-Min (Catherine) Jai, and Xu Li. 2016. Guests' perceptions of green hotel practices and management responses on TripAdvisor. *Journal of Hospitality and Tourism Technology* 7: 182–99. [CrossRef]

Lee, Su-Hsin, Shih-Chieh Wu, and Anna Li. 2018. Low-carbon tourism of small islands responding to climate change. *World Leisure Journal* 60: 235–45. [CrossRef]

Li, Zhimin, Runsheng Tang, Chaofeng Xia, Huilong Luo, and Hao Zhong. 2005. Towards green rural energy in Yunnan, China. *Renewable Energy* 30: 99–108. [CrossRef]

Llinàs, O. Fullana, Enric Tello Aragay, Ivan Murray Mas, Gabriel Jover-Avellà, and Joan Marull López. 2021. Socio-ecological transition in a Mediterranean agroecosystem: What energy flows tell us about agricultural landscapes ruled by landlords, peasants and tourism (Mallorca, 1860–1956–2012). *Ecological Economics* 190: 107206. [CrossRef]

Loia, Francesca, Nunzia Capobianco, and Roberto Vona. 2021. Towards a resilient perspective for the future of offshore platforms. Insights from a data driven approach. *Transforming Government: People, Process and Policy* 16: 218–30. [CrossRef]

Loiseau, Eléonore, Laura Saikku, Riina Antikainen, Nils Droste, Bernd Hansjürgens, Kati Pitkänen, Pekka Leskinen, Peter Kuikman, and Marianne Thomsen. 2016. Green economy and related concepts: An overview. *Journal of Cleaner Production* 139: 361–71. [CrossRef]

Lu, Ping, Xuan Yang, and Zhou-Jing Wang. 2018. Fuzzy Group Consensus Decision Making and Its Use in Selecting Energy-Saving and Low-Carbon Technology Schemes in Star Hotels. *International Journal of Environmental Research and Public Health* 15: 2057. [CrossRef]

Lucrezi, Serena, and Melville Saayman. 2017. Sustainable scuba diving tourism and resource use: Perspectives and experiences of operators in Mozambique and Italy. *Journal of Cleaner Production* 168: 632–44. [CrossRef]

Lucrezi, Serena, Martina Milanese, Vasiliki Markantonatou, Carlo Cerrano, Antonio Sarà, Marco Palma, and Melville Saayman. 2017. Scuba diving tourism systems and sustainability: Perceptions by the scuba diving industry in two Marine Protected Areas. *Tourism Management* 59: 385–403. [CrossRef]

Luu, Trong Tuan. 2018. Employees' green recovery performance: The roles of green HR practices and serving culture. *Journal of Sustainable Tourism* 26: 1308–24. [CrossRef]

Ma, Huiqiang, Jiale Liu, and Jianchao Xi. 2022. Decoupling and decomposition analysis of carbon emissions in Beijing's tourism traffic. *Environment, Development and Sustainability* 24: 5258–74. [CrossRef]

Majdak, Piotr, and Antonio M. M. de Almeida. 2022. Pre-Emptively Managing Overtourism by Promoting Rural Tourism in Low-Density Areas: Lessons from Madeira. *Sustainability* 14: 757. [CrossRef]

Manomaivibool, Panate. 2015. Wasteful tourism in developing economy? A present situation and sustainable scenarios. *Resources, Conservation and Recycling* 103: 69–76. [CrossRef]

McKinley, Emma, Oscar Aller-Rojas, Caroline Hattam, Celine Germond-Duret, Inés Vicuña San Martín, Charlotte Rachael Hopkins, Héctor Aponte, and Tavis Potts. 2019. Charting the course for a blue economy in Peru: A research agenda. *Environment, Development and Sustainability* 21: 2253–75. [CrossRef]

Mena-Nieto, Angel, Charles Estay-Ossandon, and Sergio Pereira dos Santos. 2021. Will the Balearics and the Canary Islands meet the European Union targets for municipal waste? A comparative study from 2000 to 2035. *Science of the Total Environment* 783: 147081. [CrossRef]

Mestanza-Ramón, Carlos, Maritza Sanchez Capa, Hilter Figueroa Saavedra, and Juana Rojas Paredes. 2019. Integrated Coastal Zone Management in Continental Ecuador and Galapagos Islands: Challenges and Opportunities in a Changing Tourism and Economic Context. *Sustainability* 11: 6386. [CrossRef]

Mimbs, Benjamin Prangle, B. Bynum Boley, J. M. Bowker, Kyle M. Woosnam, and Gary T. Green. 2020. Importance-performance analysis of residents' and tourists' preferences for water-based recreation in the Southeastern United States. *Journal of Outdoor Recreation and Tourism* 31: 100324. [CrossRef]

Nelson, Katherine M., Stefan Partelow, Moritz Stäbler, Sonya Graci, and Marie Fujitani. 2021. Tourist willingness to pay for local green hotel certification. *PLoS ONE* 16: e0245953. [CrossRef]

Newton, Adrian C. 2015. Defining the green economy and the potential role of green tourism. In *Tourism in the Green Economy*. Edited by Maharaj Vijay Reddy and Keith Wilkes. London and New York: Routledge, pp. 32–45.

Nhamo, G. 2010. Green economies and green jobs: Implications for South Africa. *Environmental Economics and Investment Assessment III* 131: 257–68. [CrossRef]

OECD. 2020. *OECD Tourism Trends and Policies 2020*. Paris: Organisation for Economic Co-Operation and Development, OECD Publishing. [CrossRef]

Ohajionu, Uchechi Cynthia, Bright Akwasi Gyamfi, Murat Ismet Haseki, and Festus Victor Bekun. 2022. Assessing the linkage between energy consumption, financial development, tourism and environment: Evidence from method of moments quantile regression. *Environmental Science and Pollution Research* 29: 30004–18. [CrossRef]

Olcina Cantos, Jorge, and José Fernando Vera-Rebollo. 2016. Adaptation of the tourism sector to climate change in Spain. The importance of action at local level and in tourism businesses. *Anales de Geografia de la Universidad Complutense* 36: 321–52. [CrossRef]

Panzer-Krause, Sabine. 2017. Un-locking unsustainable tourism destination paths: The role of voluntary compliance of tourism businesses with sustainability certification on the island of Rügen. *Zeitschrift für Wirtschaftsgeographie* 61: 174–90. [CrossRef]

Pearce, Douglas G., and Christian Schott. 2010. Tourism and climate change: Public and private sector responses in New Zealand. *Bridging Tourism Theory and Practice* 3: 165–86.

Pham, Nhat Tan, Zuzana Tučková, and Charbel José Chiappetta Jabbour. 2019. Greening the hospitality industry: How do green human resource management practices influence organizational citizenship behavior in hotels? A mixed-methods study. *Tourism Management* 72: 386–99. [CrossRef]

Pimonenko, Tetyana, Oleksii Lyulyov, and Yana Us. 2021. Cointegration between Economic, Ecological and Tourism Development. *Journal of Tourism and Services* 12: 169–80. [CrossRef]

Pongsakornrungsilp, Pimlapas, and Siwarit Pongsakornrungsilp. 2021. Mindful tourism: Nothing left behind–creating a circular economy society for the tourism industry of Krabi, Thailand. *Journal of Tourism Futures* 1–15. [CrossRef]

Porte, Marcelo, Irina Saru-Amaral, and Carlos Pinho. 2018. Pesquisa em auditoria: Principais temas. *Revista Contabilidade & Fianançasa* 29: 41–59.

Potts, Tavis, Piotr Niewiadomski, and Katrin Prager. 2019. The Green Economy Research Centre—Positioning geographical research in Aberdeen to address the challenges of green economy transitions. *Scottish Geographical Journal* 135: 356–70. [CrossRef]

Prideaux, Bruce, Michelle Thompson, and Anja Pabel. 2020. Lessons from COVID-19 can prepare global tourism for the economic transformation needed to combat climate change. *Tourism Geographies* 22: 667–78. [CrossRef]

Razzaq, Asif, Tehreem Fatima, and Murshed Muntasir. 2021. Asymmetric effects of tourism development and green innovation on economic growth and carbon emissions in top 10 GDP countries. *Journal of Environmental Planning and Management* 1–30. [CrossRef]

Reddy, Maharaj Vijay, and Keith Wilkes, eds. 2015. *Tourism in the Green Economy*. London and New York: Routledge.

Salehi, Mohammad, Viachaslau Filimonau, Zahed Ghaderi, and Jamshid Hamzehzadeh. 2021. Energy conservation in large-sized hotels: Insights from a developing country. *International Journal of Hospitality Management* 99: 103061. [CrossRef]

Salvino, Robert, Trevor Tarleton, Nicole Kuhn, and Raegan Kauffman. 2018. Regional Growth and Transportation Infrastructure from Greenways to Public Transit. *Review of Business* 38: 74–86.

Schumacher, Ernst Friedrich. 1973. *Small Is Beautiful. Economics as If People Mattered*. Oxford: Blond & Briggs.

Scorza, Francesco, and Giovanni Fortunato. 2021. Cyclable Cities: Building Feasible Scenario through Urban Space Morphology Assessment. *Journal of Urban Planning and Development* 147: 05021039. [CrossRef]

Sevastiyanov, Dmitry V., Alfred Colpaert, Eugene Korostelyov, Oleg Mulyava, and Larissa Shitova. 2014. Management of tourism and recreation possibilities for the sustainable development of the north-western border region in Russia. *Nordia Geographical Publications* 43: 27–38.

Sharif, Arshian, Sahar Afshan, Sesokunthideth Chrea, Aguir Amel, and Syed Abdul Rehman Khan. 2020. The role of tourism, transportation and globalization in testing environmental Kuznets curve in Malaysia: New insights from quantile ARDL approach. *Environmental Science and Pollution Research* 27: 25494–509. [CrossRef] [PubMed]

Strielkowski, Wadim, Irina Firsova, Inna Lukashenko, Jurgita Raudeliūnienė, and Manuela Tvaronavičienė. 2021. Effective management of energy consumption during the COVID-19 pandemic: The role of ICT solutions. *Energies* 14: 893. [CrossRef]

Stroebel, Melanie. 2015. Tourism and the green economy: Inspiring or averting change? *Third World Quarterly* 36: 2225–43. [CrossRef]

Tao, Zhuo, and Shuangshuang Chen. 2016. Green innovation, environmental protection perception and tourism decision-making behavior: Evidence from tourism in China. *International Journal of Earth Sciences and Engineering* 9: 1683–88.

Tazikhina, Tatiana, Vladimir Kvasha, Yulia Solovova, and Igbal Guliev. 2022. Green energy in the caribbean: Influence on tourism and economic development. *Problemy Ekorozwoju* 17: 279–87. [CrossRef]

Temperini, Valerio, Gian Luca Gregori, Silvio Cardinali, and Andrea Perna. 2012. Common issues in the development of "green marketing" practices in tourism: Experiences from Marche region, Italy. *Quality—Access to Success* 13: 103–7.

Torres-Bagur, Maria, Anna Ribas, and Josep Vila-Subirós. 2020. Understanding the Key Factors That Influence Efficient Water-Saving Practices among Tourists: A Mediterranean Case Study. *Water* 12: 2083. [CrossRef]

Toubes, Diego R., Noelia Araújo-Vila, and José A. Fraiz-Brea. 2021. Organizational Learning Capacity and Sustainability Challenges in Times of Crisis: A Study on Tourism SMEs in Galicia (Spain). *Sustainability* 13: 11764. [CrossRef]

Trišić, Igor, Snežana Štetić, Donatella Privitera, Marko Petrović, Marija Maksin, Slavoljub Vujović, Zoran Jovanović, and Marija Kalinić. 2021. Perspectives on Sustainable Tourism Development in the Hotel Industry—A Case Study from Southern Europe. *Sustainability* 13: 5563. [CrossRef]

UNEP. 2011. *Towards a Green Economy: Pathways to Sustainable Development and Poverty Eradication—A Synthesis for Policy Makers*. St-Martin-Bellevue: United Nations Environment Programme.

UNWTO. 2012. *Tourism in the Green Economy—Background Report*. Madrid: United Nations Environment Programme (UNEP) and World Tourism Organization (UNWTO).

Voukkali, Irene, Pantelisa Loizia, Jose Navarro Pedreño, and Antonis A. Zorpas. 2021. Urban strategies evaluation for waste management in coastal areas in the framework of area metabolism. *Waste Management & Research: The Journal for a Sustainable Circular Economy* 39: 448–65. [CrossRef]

WTTC. 2021. *Travel & Tourism Economic Impact 2021*. Global Economic Impact and Trends. London: World Travel & Tourism Council.

Ye, X. L., L. Wang, Z. F. Zhu, J. Du, and R. H. Sun. 2019. Study on the ecological efficiency and its temporal dynamic evolution of China's cruise tourism from the perspective of carbon footprint. *Applied Ecology and Environmental Research* 17: 13001–13. [CrossRef]

Ye, Yan-Qiong, Guo-Jie Chen, and Fan Hong. 2003. Impacts of the "Grain for Green" project on rural communities in the Upper Min River Basin, Sichuan, China. *Mountain Research and Development* 23: 345–52.

Zhou, Haomiao, Zhihong Deng, Yuanqing Xia, and Mengyin Fu. 2016. A new sampling method in particle filter based on Pearson correlation coefficient. *Neurocomputing* 216: 208–15. [CrossRef]

economies

MDPI

Article

Radial Symmetry Does Not Preclude Condorcet Cycles If Different Voters Weight the Issues Differently

Richard F. Potthoff

Department of Political Science and Social Science Research Institute, Duke University, Box 90420, Durham, NC 27708, USA; potthoff@duke.edu

Abstract: Radial symmetry, by our definition, is a precise condition on continuous ideal-point distributions, rarely if ever found exactly in practice, that is similar to the classical 1967 symmetry condition of Plott but pertains to an infinite electorate; the bivariate normal distribution provides an example. A Condorcet cycle exists if the electorate prefers alternative X to Y, Y to Z, and Z to X. An alternative K is a Condorcet winner if there is no alternative that the electorate prefers to K. Lack of a Condorcet winner may engender turmoil. The nonexistence of a Condorcet winner implies that a Condorcet cycle exists. Radial symmetry precludes the existence of Condorcet cycles and thus guarantees a Condorcet winner; but this result assumes that all voters weight the dimensions alike. Our counterexamples show that a Condorcet cycle can arise, even under radial symmetry, if the weighting of issues varies across voters. This finding may be of more than theoretical value: It may suggest that in an empirical setting (without radial symmetry), a Condorcet cycle may be more frequent if voters differ as to how they weight the dimensions. We examine, for illustration based on two dimensions (left–right, linguistic), a Condorcet preference cycle in Finland's 1931 presidential election.

Keywords: Condorcet cycle; Condorcet winner; Condorcet paradox; multidimensional issue space; radial symmetry; spatial modeling

Citation: Potthoff, Richard F.. 2022. Radial Symmetry Does Not Preclude Condorcet Cycles If Different Voters Weight the Issues Differently. *Economies* 10: 166. https://doi.org/10.3390/economies10070166

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 9 April 2022
Accepted: 9 June 2022
Published: 13 July 2022

1. Introduction

An election or vote may breed discord if there is no Condorcet winner—if every alternative faces at least one competing alternative that the electorate (strictly) prefers. That quandary reflects a Condorcet cycle, where (more than half of) the electorate prefers alternative X to Y, Y to Z, and Z to X.

We define a symmetry condition, which we refer to as radial symmetry, that is similar to that in the classic work of Plott (1967) (and is a large-electorate analog thereof) but involves an infinite (rather than finite) number of voters along with a continuous distribution of voter ideal points. Similar to Plott's condition, it is highly restrictive.

It precludes a Condorcet cycle among any three alternatives (thus guaranteeing a Condorcet winner) if all voters weight the issues alike. However, such is no longer the case if voters differ as to the relative importance they attach to the issues, as we show using counterexamples with two dimensions (issues). This interesting theoretical finding may have practical implications: It may suggest that, correspondingly, Condorcet cycles are more likely empirically if voters weight the issues disparately. Dimensions that may be afforded different relative importance by different political parties or voters may include economic left–right, social–cultural, and others (see, e.g., Polk et al. 2017). Of course, the extent to which Condorcet cycles have real-world relevance in the first place has been questioned, as in Tullock (1967). Simpson (1969) amplifies Tullock's (1967) work and mentions Plott's (1967) condition.

Our work here has several distinguishing features. (i) Its model deals with differential issue weightings among voters, a facet of spatial modeling that is not often recognized in prior literature. (ii) Its model uses a continuous distribution of ideal points with an infinite number of voters, in contrast to the setup with a finite number of voters that is generally specified under the Plott (1967) model. This infinite-electorate model not only bears greater resemblance to the real world of voters in large public elections but also has certain advantages of lucidity. (iii) For the case where issue weightings are the same for all voters, the continuous distribution of ideal points with an infinite electorate leads to a simpler proof of the result—akin to the classical one under the Plott model—that radial symmetry precludes Condorcet cycles. (iv) For the case where issue weightings differ among voters, we focus on showing that radial symmetry does not preclude Condorcet cycles under the infinite-electorate setup but also find that that conclusion holds under the Plott finite-electorate model as well. This primary result of ours is expressed in the title of the paper. (v) In addition to their theoretical interest, our results may have the practical implications indicated in the preceding paragraph: They may suggest a hypothesis that cycles are more apt to occur with greater weighting differences among voters.

In what follows, Section 2 covers the details of our setting and concepts. For the case where voters' weightings are homogeneous and the electorate is infinite, propositions for radial symmetry established in Section 3 have simpler proofs and stronger results than comparable ones for a finite electorate with their standard symmetry conditions.

A limited discussion of earlier writings on differential weightings is in Section 4. This review of prior results covers a short note (Hoyer and Mayer 1975) that presents a finite-electorate example with radial symmetry where a Condorcet cycle occurs under differential weighting, though with an atypical set of utility functions.

For the case (under radial symmetry) where the voter weightings differ from one another, only one counterexample is needed to prove that Condorcet cycles are not precluded. Nonetheless, Section 5 provides several of them to illustrate different ways in which the cycles can occur for an infinite electorate; Section 6 covers the case of a finite electorate.

The degree of empirical prevalence of Condorcet cycles is examined in Section 7. In an empirical illustration based on two dimensions consisting of left–right and linguistic issues, Section 8 looks at a Condorcet preference cycle in Finland's 1931 presidential election. Section 9 summarizes.

2. Framework

We start by defining our framework, which is similar, but not identical, to that of Plott (1967; hereafter Pl); is even closer to that of Feld and Grofman (1987, especially Theorem 6; hereafter F-G) and Miller (2015, especially p. 172; hereafter Mi); and has still more common elements with Davis et al. (1972; hereafter DD&H). As we lay out our framework below, we indicate differences between it and Pl, F-G, Mi, and DD&H. (Note, though, that the second and third framework items below obviously do not apply to Section 6 below.)

Dimensions. We work with an issue space of two dimensions, but results are generally extendable to more than two.

Voters. We have an infinite number of voters (which closely approximates a large electorate). Each voter V has a (unique) ideal point, where V's utility is maximal. By contrast, Pl, F-G, and Mi use a finite number of voters, though they can be located in varying ways. DD&H allow for either a finite or an infinite electorate.

Distribution of voters' ideal points. We posit that this ideal-point distribution is continuous. It is discrete for Pl, F-G, and Mi (since their number of voters is finite). DD&H provide for it to be either discrete or continuous. Use of a continuum of voters in voting applications is also found in (e.g.) Caplin and Nalebuff (1988, p. 789).

Alternatives. Alternatives, or options, could be political candidates, legislative proposals, or something else. Hereafter, we will just refer to the alternatives as *candidates* but with the understanding that other possibilities are not precluded. Candidates' (unique) ideal points are located in the same issue space as those of voters. Our counterexamples use three candidates, though similar ones with more than three could easily be constructed.

Pl, F-G, and Mi do not consider triples of candidates (or of other alternatives) in the way that we do. DD&H have two examples that use three alternatives.

Utility function. We write P_V: (x_V, y_V) for the ideal point of voter V and P_G: (x_G, y_G) for the ideal point of candidate G. Our utility function is negative squared Euclidean distance for the case where all voters weight the two issues alike, so that the utility of voter V for candidate G is

$$U(V,\ G) = -(x_V - x_G)^2 - (y_V - y_G)^2\,. \tag{1}$$

For F-G, Mi, and DD&H, utility is also based on Euclidean distance, but for Pl it involves the utility gradient because the Pl treatment is more general.

Disparate weighting of issues. For the case where voters differ in their relative weightings of the two issues, we generalize (1) and use the utility function

$$U(V,\ G) = -w_{1V}(x_V - x_G)^2 - w_{2V}(y_V - y_G)^2\,. \tag{2}$$

Weight ratios w_{1V}:w_{2V} of 6:3, 2:1, and 1:$\frac{1}{2}$ (e.g.) all have the same effect. (Suitable weight ratios will vary depending on the relative scaling of x and y.) In general, the weight ratios w_{1V}:w_{2V} follow a distribution that can vary with (i.e., can be conditional on) the point P_V: (x_V, y_V). However, for our counterexamples later, except those in Section 6, the w_{1V}:w_{2V} weight ratios have the same distribution for the voters at every point (x_V, y_V) in the issue space (e.g., a distribution that, at each point, assigns 1:9 for half the voters and 9:1 for the other half). Accordingly, with the w_{1V}:w_{2V} distribution independent of the ideal point P_V, it is designated more simply as the w_1:w_2 distribution. (We can create each type of counterexample in Section 5 without a need for the w_{1V}:w_{2V} distribution to differ for different P_V.) Pl, F-G, Mi, and DD&H do not consider different weightings of the dimensions by different voters and so do not deal with anything like (2).

Radial symmetry. We define *radial symmetry* to be present in the distribution of voter ideal points if there exists a point M such that every line that passes through M has half the voters on each side of the line. (Because the distribution is continuous, points that are on the line itself can be disregarded.) Without loss of generality, we assume M to be the origin, (0, 0). Note that, with our framework, the definition of radial symmetry pertains only to the distribution of voters' ideal points and says nothing about a utility function (or about weighting of issues, or about the candidates). F-G and Mi do not use the term *radial symmetry* but, for a discrete distribution of voter ideal points, have a concept that can be deemed analogous to ours: There is a voter at M, and every line through M has the same number of voters on each side of M. The framework of Pl is also analogous but is more general and does involve utility. DD&H use concepts similar to ours.

Condorcet cycles. Under (1), our definitions that follow are consonant with standard ones and with those of Pl, F-G, Mi, and DD&H. Voter V at point P_V: (x_V, y_V) *prefers* candidate G at P_G: (x_G, y_G) to candidate H at P_H: (x_H, y_H) if and only if $U(V, G) > U(V, H)$. We write G►H to denote that more than half the electorate prefers G to H. (G►H if and only if more than half the electorate has greater utility for G than for H (i.e., is closer to P_G than to P_H, or, equivalently, is on the P_G side of the perpendicular bisector of the line segment joining P_G and P_H).) We define a *Condorcet cycle* to exist among three candidates A, B, and C if A►B, B►C, and C►A; or if B►A, C►B, and A►C. We then define *preclusion of Condorcet cycles* to mean that it is not possible to choose, from anywhere in the issue space, a triple of candidates A, B, and C that exhibits a Condorcet cycle. If, in a group of existing candidates, A is a(n existing) candidate such that G►A (A►G) for no other (existing) candidate G located elsewhere, then A is a *Condorcet winner* (*Condorcet loser*). Under (2) with the weight ratio following the same distribution throughout the space of voter ideal points, definitions akin to the foregoing apply.

Miscellaneous. We mention three further details. First, although we use straightforward definitions of Condorcet winner and Condorcet loser, it should be noted that (rarely) there can be more than one of either (e.g., two Condorcet winners, X and Y, will exist if the preference ranking is XYZ for half the electorate and YXZ for the other half). Second, although the absence of a Condorcet winner implies the presence of a Condorcet cycle, the converse is not true (e.g., B►C, C►D, and D►B, but A►B, A►C, and A►D). Third, we refrain from dealing with concepts such as the *core* (e.g., Saari 1997; Schofield 2008) that are closely related to Condorcet winners but are not essential to our development here.

3. Radial Symmetry as a Sufficient Condition to Preclude Condorcet Cycles under (1)

For (1) for our framework above, we state and prove three propositions.

Proposition 1. *Condorcet cycles are precluded if radial symmetry holds.*

Proof. For two candidates G at (x_G, y_G) and H at (x_H, y_H), define

$$D_G^2 = x_G^2 + y_G^2 \quad \text{and} \quad D_H^2 = x_H^2 + y_H^2 \,,$$

which are the respective squares of their distances from the origin (i.e., from M). We start by showing that G►H if and only if $D_G^2 < D_H^2$.

A voter V at (x_V, y_V) prefers G to H if and only if $U(V,G) > U(V,H)$, or if and only if

$$-\left(x_V - x_G\right)^2 - \left(y_V - y_G\right)^2 > -\left(x_V - x_H\right)^2 - \left(y_V - y_H\right)^2,$$

or if and only if

$$2(x_G - x_H)x_V + 2(y_G - y_H)y_V > D_G^2 - D_H^2\,. \tag{3}$$

The line that bounds the region (3) is parallel to the line that bounds

$$2(x_G - x_H)x_V + 2(y_G - y_H)y_V > 0\,. \tag{4}$$

However, the latter line passes through the origin and so, by the assumption of radial symmetry, has half the electorate on each side of it. The region defined by (3) (which is the set of voters V who prefer G to H) is larger than the region defined by (4) if and only if $D_G^2 < D_H^2$. Thus, G►H if and only if $D_G^2 < D_H^2$, since the latter region contains half the electorate.

To complete the proof, consider any three candidates A, B, and C whose respective squared distances from the origin are D_A^2, D_B^2, and D_C^2. It is, of course, not possible for these three values to exhibit intransitivity, so a Condorcet cycle among A, B, and C cannot occur. □

Although Proposition 1 and its proof are only for two dimensions, we remark that extension to more than two dimensions is immediate. Note also that it is easy to extend the proof to preclude any cyclicity among any group of more than three candidates.

Proposition 2. *If (given radial symmetry) any set of candidates includes a candidate K located at M (the origin), then K is a Condorcet winner.*

Proof. Follows at once from the foregoing, by noting that $D_K^2 = 0$. □

Proposition 2 can be considered an analog, for continuous ideal-point distributions, of the basic result of Pl, F-G, and Mi (for discrete distributions) that proves that a Condorcet winner, located at the origin, must exist under their symmetry conditions. Not only is Proposition 2 unusual in this setting in that it pertains to continuous rather than discrete ideal-point distributions, but also its proof has the benefit of being comparatively simple.

Proposition 3. *Radial symmetry guarantees a Condorcet winner even if there is no candidate located at the origin.*

Proof. This result, stronger than Proposition 2, follows from Proposition 1 because the impossibility of any Condorcet cycles implies the existence of a Condorcet winner. □

Curiously, as a matter of theoretical interest, and in contrast to the finite-electorate setup, Proposition 3 functions even if no *voter* is located at the origin. For example, in the distribution of voter ideal points there could be a doughnut hole centered at the origin.

The possible pedagogical value of the above results should not be overlooked. In particular, the proof of Proposition 1, under a continuous distribution of ideal points, is shorter and more straightforward than corresponding proofs under discrete distributions. Related results and proofs in DD&H, however, do apply to a continuous ideal-point distribution and an infinite electorate and are similar (though not identical) to ours above. Our development may be less comprehensive, but also more understandable, than that of DD&H.

We note that Propositions 1–3 still hold under a special case of (2) in which w_{1V}, w_{2V} are replaced by w_1, w_2 such that the weights are the same for all voters at a given point (as well as from one point to another) but can still differ between the two issues. This is just a transformation of scale: If we substitute $x = x_0/w_1^{1/2}$ and $y = y_0/w_2^{1/2}$ into the modified (2), then x_0 and y_0 satisfy (1), and radial symmetry still holds after the transformation.

4. Previous Work That Has Considered Unequal Weighting

Thus far, we have concentrated on the case where all voters weight the issues alike. For utility functions, whether with or without radial or other symmetry in the ideal-point distributions that apply, most authors do not consider cases where voters differ as to how they weight the dimensions. However, such differences in weighting are central to the present paper.

In earlier work, Davis and Hinich (1968, p. 68) treat a situation where some voters care only about a single issue (which may vary from one voter to another). Although Davis et al. (1970, p. 434) refer to but eschew a model with different weights for different voters, they do briefly mention that a voter may care about only one issue (p. 433), that farmers may care especially about farm price supports and petroleum-industry people about oil import quotas (p. 440), and that an assumption of equal weighting is not generally satisfied (p. 446). Hoyer and Mayer (1974) provide extensive comparisons of a model where all voters have the same weighting versus a model where voter weightings can differ. Hinich and Pollard (1981) use a model that allows voters to differ not only in their weightings of issues but also in their perceptions of candidates' positions on issues. Each candidate is located along a single "predictive dimension" (though more than one such dimension may be possible). Rabinowitz et al. (1982), Niemi and Bartels (1985), and Erikson and Romero (1990) consider weighting differences that are associated with differences in the saliences that voters attach to issues. Rivers (1988, p. 740) disapproves that voter heterogeneity in issue weighting is ignored "in common practice," before developing an approach that handles it. Although the utility function in the model of Adams (1997, p. 1254) allows for differences among voters as to how they weight the dimensions, that flexibility is not put to use. Adams and Adams (2000, p. 140) mention the possibility that "different policy dimensions matter to different voters" but judge that such a point does not affect their conclusions. Feld et al. (2014, pp. 480–81) briefly allude to a possible model "that would allow variation in the salience different voters attach to different issues."

Unlike the references just cited, Hoyer and Mayer (1975, p. 806) provide an example, for a finite electorate, where a Condorcet cycle occurs with radial symmetry and weighting differences among voters. That example has three candidates and nine voters. Though it is presented geometrically rather than algebraically, its utility functions do not all conform to (2) above. They do so for three of the nine voters. For the other six, however, the utility function is the same as (2) but with an added cross-product term of the form $-2w_{12V}(x_V - x_G)(y_V - y_G)$.

5. With an Infinite Electorate, Radial Symmetry, and Unlike Weighting, Cycles Can Occur

For an infinite electorate, we next provide counterexamples to establish that, even with (unlikely) radial symmetry of the distribution of voter ideal points, preclusion of Condorcet cycles is not guaranteed if voters differ (in their utility functions) as to the comparative importance they attach to the issues in the policy space. Even given the tight radial-symmetry restrictions, some critics might still find this result unsurprising. Here, we are proving the result by counterexample to provide broad insights, although other methods of proof might be used. Even though with equal weighting slight perturbances from radial symmetry are sufficient to create a Condorcet cycle, our counterexamples with unequal weighting (and radial symmetry) were not easy to construct, rest on thin margins for some candidate pairings, and (in most cases) place all or almost all of a given voter's weight on just one of the two issues.

We consider two bivariate (infinite-electorate) ideal-point distributions that observe radial symmetry. The first is a bivariate normal distribution with a mean of 0 and standard deviation of 10 for both x and y; we use a correlation coefficient of 0 (although, even if it is not 0, radial symmetry still holds). The second is a uniform (rectangular) distribution with probability density function $f(x, y) = 1/400$ over the square with vertices $(x, y) = (-10, 10)$, $(10, 10)$, $(-10, -10)$, and $(10, -10)$.

A single counterexample would suffice to show that Condorcet cycles are no longer precluded for voter ideal-point distributions with radial symmetry in an infinite electorate if voters differ as to how they judge the relative importance of the issues. We provide several counterexamples, however, in order to illustrate particular features.

Examples 1.1 and 1.2. The respective spatial locations of candidates A, B, and C are

$$P_A\text{: } (x_A, y_A) = (1, 7)\text{; } P_B\text{: } (x_B, y_B) = (3, -7)\text{; and } P_C\text{: } (x_C, y_C) = (-9, 5).$$

If all voters have the same utility function, as in (1), then A►B, B►C, and A►C since $D_A^2 = 50$, $D_B^2 = 58$, and $D_C^2 = 106$ (no cycle).

However, now suppose instead that (at each point in the issue space) half the electorate attaches sole importance to the issue on the x-axis whereas the other half cares only about the y-dimension. The utility functions (of voter V for candidate G) can then be represented by (2) with respective weight ratios w_1:w_2 of 1:0 and 0:1.

Example 1.1. *If the ideal points follow the uncorrelated bivariate normal distribution indicated above, then A►B, B►C, and C►A, so a Condorcet cycle exists. See Appendix A.1 for calculation details.*

Example 1.2. *Suppose now that the distribution of ideal points is uniform (as described above) rather than bivariate normal. Then a Condorcet cycle exists again, once more with A►B, B►C, and C►A. Here, the margins by which the electorate prefers A over B, B over C, and C over A are identical: 55% to 45% in all three cases. See Appendix A.1 for calculation details.*

In Examples 1.1–1.2, as well as below in Examples 3.1–3.2 and in parts of Example 4, each voter is concerned about *only one* of the two issues. Those examples were constructed in that way to provide ease of exposition and calculation.

It might be claimed, though, that all these counterexamples are problematic because the concern of voters for only one dimension renders the setting unidimensional rather than truly two-dimensional. Although we find it hard to grant that this is a valid objection, we nonetheless counter by pointing out that the proportion of the electorate with G►H (for two candidates G and H) can be a continuous—not discontinuous—function of the weight ratios. Thus, a change in the weight ratios from 1:0 and 0:1 to 1:ε and ε:1, for tiny enough ε, will not break the cycle. In fact, ε does not (in general) even need to be tiny, as evidenced by the next example.

Example 2. *The same cycle as in Examples 1.1–1.2 (A►B, B►C, C►A) also occurs if (e.g.) Example 1.2 is unchanged except that all voters care (somewhat) about both dimensions with weight ratios w_1:w_2 of 10:1 and 1:10 in place of 1:0 and 0:1. The (complicated) calculation details are in Appendix A.2.*

Examples 3.1 and 3.2. For candidates A, B, and C, respective locations are

$$P_A\text{: }(x_A, y_A) = (-3, 9)\text{; }P_B\text{: }(x_B, y_B) = (7, 7)\text{; and }P_C\text{: }(x_C, y_C) = (5, -9).$$

With the same utility function for all voters as in (1), no cycle can exist: A►B, B►C, and A►C, with $D_A^2 = 90$, $D_B^2 = 98$, and $D_C^2 = 106$. However, suppose now that, as in Examples 1.1–1.2, weightings differ across voters according to (2) with weight ratios w_1:w_2 of 1:0 for half the electorate and 0:1 for the other half.

Example 3.1. *If the ideal points follow the uncorrelated bivariate normal distribution, a Condorcet cycle occurs with B►A, C►B, and A►C. See Appendix A.3 for calculation details.*

Example 3.2. *If the ideal points follow the uniform rather than the bivariate normal distribution, a Condorcet cycle exists again, also with B►A, C►B, and A►C. See Appendix A.3 for calculation details.*

Note that, even though Examples 1.1–1.2 and 3.1–3.2 all have A►B, B►C, and A►C if all voters have the same utility function, with the change to 1:0 and 0:1 Examples 1.1–1.2 have the cycle A►B, B►C, and C►A whereas the cycle in Examples 3.1–3.2 is the opposite—B►A, C►B, and A►C. Thus, interestingly, only A►C is reversed in the former case, but both A►B and B►C are reversed in Examples 3.1–3.2—a salient difference between the two cases.

Example 4. *We extend Example 3.2, still with the same uniform distribution for the ideal points and the same three candidates, but now we suppose that the weight ratios w_1:w_2 for (2) are 1:0 for only 6% of the voters, 0:1 for another 6%, and 1:1 (equal weighting) for the remaining 88%. Even though an overwhelming proportion of the voters (88%) are identical with one another as to how they weight the issues, it turns out that the Condorcet cycle with B►A, C►B, and A►C nonetheless emerges again. See Appendix A.4 for calculation details. Therefore, here, even with radial symmetry, the generation of a cycle does not require a very large fraction of voters with non-conforming weightings.*

Obviously, where radial symmetry holds, many possible cases of unlike voter weightings of issues will *not* serve to trigger Condorcet cycles. The above counterexamples serve simply to establish the possibility of those cycles.

In all of those counterexamples, the distribution of the weight ratios w_{1V}:w_{2V} is the same (set at w_1:w_2) for voters at every point in the issue space (e.g., 10:1 for half the voters at each point and 1:10 for the other half). This allows for easier presentation but also shows that counterexamples need not be made more elaborate through use of different distributions at different points. Because of similar considerations, all of our chosen weight ratio distributions are discrete (in fact, mostly dichotomous) rather than continuous, although a possible continuous weight-ratio distribution is described briefly just before Table A1 in Appendix A.2 below.

6. Counterexamples for the Case of a Finite Electorate

Although this paper concentrates on establishing that Condorcet cycles are not precluded with differential weighting under radial symmetry if the number of voters is infinite, a comparable result holds for a finite number of voters. Thus, our final two examples demonstrate that, even under the customary symmetry conditions for a finite electorate, Condorcet cycles can occur if voters are disparate in their weightings of the issues based on the utility function (2). Similar to the examples in Section 5, which each have a common weight-ratio distribution at all points, Example 5 observes this condition (with a 7:1 ratio for one voter and 1:7 for the other) at each voter point except the origin. Example 6, though, has only one voter at each point, with weight-ratios that differ among points.

These two examples differ from the one in Hoyer and Mayer (1975, p. 806) in that both are based on the simple utility function (2). That earlier example uses a more complex utility function for some voters, as noted above in Section 4.

Example 5. *This counterexample has three candidates at the same locations as in Examples 1.1–1.2, and nine voters as follows:*

$V1$–$V4$ are at $(-6, 2)$, $(6, 2)$, $(-6, -2)$, $(6, -2)$;
$V5$–$V8$ are (respectively) at these same four locations;
$V9$ is at $(0, 0)$.
These nine voters satisfy the symmetry conditions of F-G, Mi, and Pl: A voter exists at the origin, and any line through the origin has the same number of voters on each side. Under (1), there is no Condorcet cycle because A►B, B►C, and A►C with 5-to-4 preference margins for each matchup.

Now let (2) apply with weight ratios of 7:1 for $V1$–$V4$, 1:7 for $V5$–$V8$, and 1:1 for $V9$. A Condorcet cycle—A►B, B►C, C►A, with 5-to-4 preference margins for each—then results. See Appendix A.5 for calculation details.

Example 6. *This example differs from Example 5 by having no more than one voter at the same point in the issue space—a condition that may sometimes be invoked but (unsurprisingly) does not avoid the possibility of a cycle, as Example 6 shows. Here, the three candidates have the same locations as in Example 5 (and Examples 1.1–1.2). The voters are five of the nine voters of Example 5: V5, V2, V3, V8, and V9. This set of five voters satisfies the radial-symmetry requirements. Again, there is no Condorcet cycle under (1), with a 3-to-2 margin for each of A►B, B►C, and A►C.*

However, under (2), with weight ratios as in Example 5 for the five voters, the cycle A►B, B►C, C►A results again in Example 6, with 3-to-2 margins this time. For calculation details, see Appendix A.6.

7. Empirical Frequency of Condorcet Cycles

The importance of Condorcet winners stems from the concept that, in a single-winner contest or election, the selectee or electee should be a Condorcet candidate (Condorcet winner)—assuming that one exists. Most voting theorists (e.g., Abramson et al. 2002; Dasgupta and Maskin 2004; Gehrlein 2006; Maskin and Sen 2018; Merrill 1988; Regenwetter et al. 2006) embrace this principle, though exceptions can be found (e.g., Saari 1995). The absence of a Condorcet winner reflects a Condorcet cycle.

Our results for radial symmetry may have empirical ramifications even though the symmetry condition itself would not be expected to exist (probably not even approximately) in practical situations. Thus, in broader, real-world milieus, it is reasonable to conjecture that Condorcet cycles—and the muddlement that they can create where no Condorcet winner emerges (e.g., Van Deemen 2014)—will be more frequent if voters weight the issues differently (rather than alike), in line with what happens where radial symmetry does exist.

The impact, though, may be greater or less depending both on the extent to which unlike versus like weightings can more often trigger Condorcet cycles and on the general frequency of conditions that are ripe for Condorcet cycles in the first place. As for the latter, judgments may differ as to how often Condorcet cycles occur empirically. Table 4 of Van Deemen (2014) shows that a Condorcet paradox (considered there to be the absence of a candidate against whom no rival is strictly preferred by the electorate) existed in 25, or 9.4%, of 265 elections or votes. Although that percentage is quite important and meaningful, an alternative percentage, calculated as the number of triples with a Condorcet cycle divided by total number of triples rather than as the number of elections with a Condorcet paradox divided by total number of elections, might be appreciably lower. It also appears that Condorcet paradoxes in the table occur more frequently in legislative votes or the like than in elections that involve political candidates or political parties; the former may have been winnowed more selectively, with votes perhaps likelier to reflect contrived outcomes rather than true preferences.

A few results do not appear in Table 4 of Van Deemen (2014). They include the 2009 mayoral election in Burlington, Vermont (USA), in which voters ranked five candidates and none of the $(5 \times 4 \times 3)/6 = 10$ candidate triples were cyclical (Laatu and Smith 2009; Olson 2009); a 1952 poll of 562 college students that asked for ranking of 10 candidates for U.S. president and showed no Condorcet cycles among the $(10 \times 9 \times 8)/6 = 120$ triples (Potthoff 1970); a 2019 poll of "1002 likely Democratic presidential primary voters" that requested rankings of 20 Democratic candidates for U.S. president and produced no cycles in any of the $(20 \times 19 \times 18)/6 = 2280$ triples (FairVote 2019); a similar 2020 poll of 825 Democratic voters covering eight presidential candidates that also produced no cycles, in any of the $(8 \times 7 \times 6)/6 = 56$ triples (FairVote 2020); and some of the items listed by Lagerspetz (2016, pp. 384–85). In any event, although the degree of empirical incidence of Condorcet cycles, and of the nonexistence of Condorcet winners, can be difficult to determine, their impact cannot be dismissed given their disruptive potential, and so any effect of differential issue weights on generation of greater cyclicity is deserving of attention.

8. The 1931 Election for President of Finland

Finding a good example of an actual election that is in tune with the work of this paper is challenging. The real-world example that we present here is necessarily imperfect. It does have a Condorcet cycle that is well-documented, which is a rarity. It has a two-dimensional issue space but does not have radial symmetry, of course. Numerical values associated with the voters and candidates have to be assigned somewhat arbitrarily. The distribution of voter ideal points cannot be portrayed as continuous, in part because the voters are electors representing political parties (rather than typical citizens) and are thus subject to party pressure for homogeneity. In contrast to our main counterexamples above, the weighting of the two issues is the same for all voters with a given ideal point but differs from one point to another. Although some features of our example are thus not in accord with our structures set forth previously, it may still provide some useful insight.

The example is from the 1931 presidential election in Finland, in which 300 electors, chosen through proportional representation and representing six political parties or factions thereof, voted among four candidates. The preference orderings of all 300 electors were generally known (Lagerspetz 2016, p. 397) and are shown, along with the number of electors, in Table 1 for each of the six party groups. The Swedish People's Party had no candidate and had two blocs (which we call Sw1 and Sw2). Otherwise, the four candidates and their parties were: Tanner, Social Democrats; Ståhlberg, Progressives; Kallio, Agrarians; and Svinhufvud, Conservatives. Tanner was the Condorcet loser. A preference cycle, however, encompassed the three at the top: Ståhlberg over Kallio, Kallio over Svinhufvud, and Svinhufvud over Ståhlberg, as shown on the right side of Table 1.

Lagerspetz (2016, pp. 393–94) identifies three applicable dimensions or issues: traditional left–right; linguistic—Swedish-speaking versus Finnish-speaking; and degree of acceptance of democracy. Because the third dimension seems to be closely aligned with the first, we treat it as combined with the first and consider just the left–right (x) and linguistic (y) issues. The latter concerns mainly a sharp discord between Kallio and the Swedish blocs: "the Agrarian candidate Kallio—a Finnish nationalist who, unlike most leading politicians, could not even speak Swedish—was totally unacceptable to the Swedish-speaking group" (Lagerspetz 2016, p. 396).

Table 1. For each party or faction, preference rankings among the four candidates.

Party Group (Party or Faction)	Number of Electors	Preference Ranking [a]	Head-to-Head Preference Matchups [a,b]					
			St vs. K		K vs. Sv		Sv vs. St	
			St	K	K	Sv	Sv	St
Social Democrats	90	T St K Sv	90	0	90	0	0	90
Progressives	52	St K Sv T	52	0	52	0	0	52
Agrarians	69	K Sv St T	0	69	69	0	69	0
Conservatives	64	Sv K St T	0	64	0	64	64	0
Swedish People's Party								
Bloc 1—Sw1	7	St Sv K T	7	0	0	7	0	7
Bloc 2—Sw2	18	Sv St K T	18	0	0	18	18	0
Total	300		167	133	211	89	151	149

[a] T = Tanner, St = Ståhlberg, K = Kallio, Sv = Svinhufvud. [b] Pairings of T against St, K, and Sv are not shown; T loses to each of the other three by 210 to 90.

Shown in Table 2, and also graphically in Figure 1, are our (x, y) spatial-location numerical assignments to the six party groups and the four candidates, selected so as to try to reflect the essence of the political situation for the two issues. For all 24 combinations of party groups with candidates, the table provides the weighted squared distance corresponding to (2) above, using our chosen weight ratios $w_{1V}{:}w_{2V}$ for the first to second issue. V now refers to a party group rather than an individual voter. There is one ratio for each party group V.

Table 2. Weighted squared distance from each party group to each candidate.

Party Group (Party or Faction), V	Weight Ratio, [a] $w_{1V}{:}w_{2V}$	Spatial Location, P_V: (x_V, y_V)	Weighted Squared Distance				Preference Ranking [b]
			Candidate, [b] G, and His Spatial Location, P_G: (x_G, y_G)				
			T	St	K	Sv	
			(−5, 0)	(−1, 0)	(1, 1)	(2, 0)	
Social Democrats	2:1	(−5, 0)	0	32	73	98	T St K Sv
Progressives	2:1	(−1, 0)	32	0	9	18	St K Sv T
Agrarians	1:1	(1, 1)	37	5	0	2	K Sv St T
Conservatives	2:1	(2, 0)	98	18	3	0	Sv K St T
Swedish People's Party							
Bloc 1—Sw1	1:2	(0, −1)	27	3	9	6	St Sv K T
Bloc 2—Sw2	1:2	(1, −1)	38	6	8	3	Sv St K T

[a] Same for all voters in the party group. [b] T = Tanner, St = Ståhlberg, K = Kallio, Sv = Svinhufvud.

Our weight ratios (Table 2) are taken as 2:1 except for the two Swedish blocs, which each receive 1:2 to reflect their strong concern for the linguistic dimension, and the Agrarians, with 1:1. As an illustration, the weighted squared distance between Sw1 and Svinhufvud is $1(0 - 2)^2 + 2(-1 - 0)^2 = 1{\cdot}4 + 2{\cdot}1 = 6$. From low to high, the four values for Sw1 are 3 (Ståhlberg), 6 (Svinhufvud), 9 (Kallio), and 27 (Tanner), thus producing the preference ordering on the right side of Table 2.

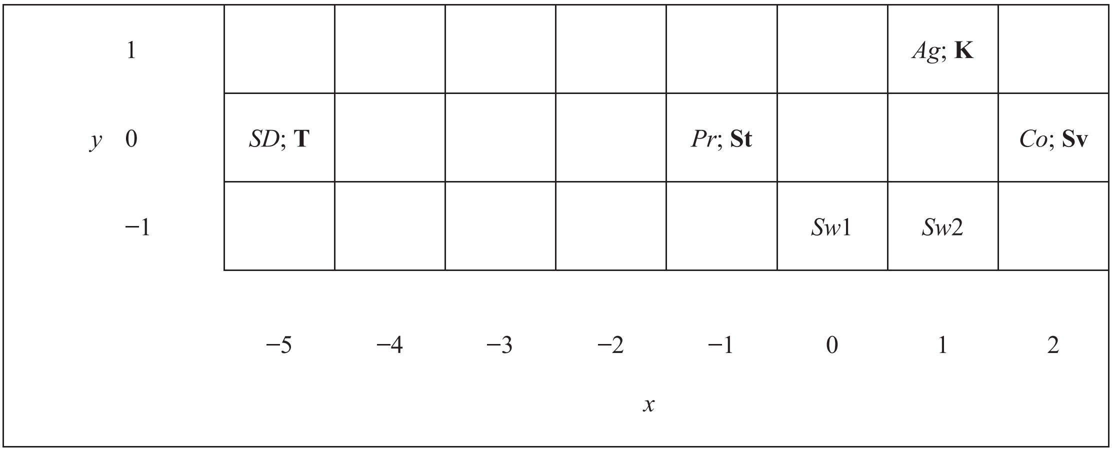

Figure 1. Two-dimensional issue space for party groups and candidates. Traditional left-right dimension is on the x-axis; linguistic dimension (Swedish versus Finnish) is on the y-axis. Party groups: SD = Social Democrats; Pr = Progressives; Ag = Agrarians; Co = Conservatives; $Sw1$, $Sw2$ = blocs 1, 2 of Swedish People's Party. Candidates: **T** = Tanner, **St** = Ståhlberg, **K** = Kallio, **Sv** = Svinhufvud.

All six of the preference orderings in Table 2—obtained using our (x, y) points and through the weight ratios w_{1V}:w_{2V} that differ by party group V—agree with the documented orderings in Table 1. They, thus, beget the same Condorcet cycle.

If for all six party groups w_{1V}:w_{2V} were set alike (to w_1:w_2) as either 2:1 or 1:1 (though not 1:2), one can show that no cycle would arise. See Appendix A.7 for details.

The 1931 Finnish presidential election provides a far-from-perfect illustration for our work. However, it does show how an election with a Condorcet cycle can be fitted by a model that has a two-dimensional issue space together with differential weights for the voters.

9. Summary

Different aspects of our work may be of significance or have interesting implications. Our work brings out some benefits of continuous (versus discrete) ideal-point distributions and infinite (versus finite) electorates. It calls attention to the topic, mostly neglected heretofore, of voters' differential weightings of issues. The findings on the failure of symmetry to preclude cycles under differential weighting may be of some theoretical interest, and our approach to proofs in Section 3 for the case of equal weighting and an infinite electorate may have some pedagogical advantages over alternative methods.

We showed that, even under the stringent condition of radial symmetry (which we define for an infinite electorate), Condorcet cycles can occur if voters differ on their weightings of issues (though not if they do not differ). It is reasonable to hypothesize that this theoretical result has a real-world counterpart: that, in empirical settings (necessarily without radial symmetry), the absence of Condorcet winners stemming from Condorcet cycles will occur more often where voters weight the issues heterogeneously. Such a hypothesis might be examined. If it holds, the question may arise, in some cases, as to whether measures to try to lessen voters' differences in issue weightings could (or should) be taken to try to reduce cycles and associated instability.

Funding: This research received no external funding.

Data Availability Statement: All data used in this paper are hypothetical and are shown in the paper itself.

Acknowledgments: The author thanks the three anonymous reviewers for their useful comments.

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

Appendix A

Appendix A.1. Details for Examples 1.1 and 1.2

A voter V in the first half of the electorate has preferences as follows:
V prefers A to B if $x_V < 2$ [since $\frac{1}{2}(x_A + x_B) = \frac{1}{2}(1 + 3) = 2$],
V prefers B to C if $x_V > -3$,
V prefers A to C if $x_V > -4$,
and conversely (e.g., V prefers B to A if $x_V > 2$). The preferences of a voter V who is in the second half are:
V prefers A to B if $y_V > 0$ [since $\frac{1}{2}(y_A + y_B) = 0$],
V prefers B to C if $y_V < -1$,
V prefers A to C if $y_V > 6$,
and conversely.

Example 1.1. *The ideal points follow the uncorrelated bivariate normal distribution (means of 0, standard deviations equal to 10). Let Φ(•) denote the cumulative distribution function of the standard (univariate) normal distribution (with mean 0 and variance 1). Then, the proportion of the electorate*

that prefers A to B is $[\Phi(0.2) + 1 - \Phi(0)]/2$,
$>\frac{1}{2}$ since $\Phi(0.2) > \Phi(0)$;
that prefers B to C is $[1 - \Phi(-0.3) + \Phi(-0.1)]/2$,
$>\frac{1}{2}$ since $\Phi(-0.1) > \Phi(-0.3)$; and
that prefers A to C is $[1 - \Phi(-0.4) + 1 - \Phi(0.6)]/2 = [\Phi(0.4) + 1 - \Phi(0.6)]/2$,
$<\frac{1}{2}$ since $\Phi(0.6) > \Phi(0.4)$.

Thus A►B, B►C, and C►A.

Example 1.2. *The distribution of ideal points is uniform rather than bivariate normal. Then the proportion of the electorate*
that prefers A to B is $\left[\frac{2-(-10)}{20} + \frac{10-0}{20}\right]/2 = 55\%$;
that prefers B to C is $\left[\frac{10-(-3)}{20} + \frac{-1-(-10)}{20}\right]/2 = 55\%$; and
that prefers A to C is $\left[\frac{10-(-4)}{20} + \frac{10-6}{20}\right]/2 = 45\%$.

Thus A►B, B►C, and C►A with 55%-to-45% preference margins in all three cases.

Appendix A.2. Details for Example 2

Table A1 shows that the cycle A►B, B►C, C►A results. To illustrate the calculations, we consider the bottom of the table. The last two lines show that A►C for 69.4% and 30% of the halves of the electorate with 10:1 and 1:10 weight ratios, respectively. Thus C►A overall, by 50.3% to 49.7%.

For the last line, with a 1:10 weight ratio, (2) in the main text yields

$$U(V,\ A) - U(V, C) = \left[-(x-1)^2 - 10(y-7)^2\right] - \left[-(x+9)^2 - 10(y-5)^2\right] = 0,$$

or

$$20(x + 2y - 8) = 0,$$

as the applicable bisector of $\overline{P_A P_C}$. This bisector intersects the line segment $\overline{P_A P_C}$ at $(x, y) = (-4, 6)$, and crosses the left side of the square at $y = 9$, the right side at $y = -1$, and the y-axis at $y = 4$. Because the average ordinate of the bisector is $y = 4$, the proportion of the group that prefers A to C is $\frac{10-4}{20} = 30\%$.

Remark. *The same cycle of A►B, B►C, C►A also emerges if (at each point in the issue space) w_1:w_2 follows a continuous distribution over the range from ($\frac{10}{11}$: $\frac{1}{11}$, $\frac{1}{11}$: $\frac{10}{11}$) to (1:0, 0:1)—for instance, half ($1 - \frac{1}{11}u$): $\frac{1}{11}u$ and half $\frac{1}{11}u$:($1 - \frac{1}{11}u$) with u distributed according to the uniform (rectangular) distribution on the interval [0, 1].*

Table A1. Details demonstrating the existence of the Condorcet cycle.

Candidate		$\overline{P_G P_H}$ Midpoint	Weight Ratio, w_1:w_2	Bisector of $\overline{P_G P_H}$: $U(V, G)-U(V, H)$=0, Using (2)	Value of						% of Electorate with G►H	
G at (x_G, y_G)	H at (x_H, y_H)				y When x =			x When y =				
					−10	0	10	−10	0	10	Each Half	Both
A, (1, 7)	B, (3, −7)	(2, 0)	10:1	$4(-10x + 7y + 20) = 0$				−5	2	9	(2 − (−10))/20 = 60%	55.07%
			1:10	$4(-x + 70y + 2) = 0$	−6/35	−1/35	4/35				(10 − (−1/35))/20 = 50.14%	
B, (3, −7)	C, (−9, 5)	(−3, −1)	10:1	$24(10x - y + 29) = 0$				−3.9	−2.9	−1.9	(10 − (−2.9))/20 = 64.5%	55.5%
			1:10	$24(x - 10y - 7) = 0$	−1.7	−0.7	0.3				(−0.7 − (−10))/20 = 46.5%	
A, (1, 7)	C, (−9, 5)	(−4, 6)	10:1	$4(50x + y + 194) = 0$				−3.68	−3.88	−4.08	(10 − (−3.88))/20 = 69.4%	49.7%
			1:10	$20(x + 2y - 8) = 0$	9	4	−1				(10 − 4)/20 = 30%	

Appendix A.3. Details for Examples 3.1 and 3.2

A voter V in the first half of the electorate has these preferences:
V prefers A to B if $x_V < 2$,
V prefers B to C if $x_V > 6$,
V prefers A to C if $x_V < 1$,
and conversely. The preferences of a voter V in the second half are:
V prefers A to B if $y_V > 8$,
V prefers B to C if $y_V > -1$,
V prefers A to C if $y_V > 0$,
and conversely.

Example 3.1. *For the uncorrelated bivariate normal distribution, the proportion of voters*

that prefers A to B is $[\Phi(0.2) + 1 - \Phi(0.8)]/2$,
$<\frac{1}{2}$ since $\Phi(0.8) > \Phi(0.2)$;
that prefers B to C is $[1 - \Phi(0.6) + 1 - \Phi(-0.1)]/2 = [1 - \Phi(0.6) + \Phi(0.1)]/2$,
$<\frac{1}{2}$ since $\Phi(0.6) > \Phi(0.1)$; and
that prefers A to C is $[\Phi(0.1) + 1 - \Phi(0)]/2$,
$>\frac{1}{2}$ since $\Phi(0.1) > \Phi(0)$.

Thus B►A, C►B, and A►C.

Example 3.2. *For the uniform distribution, the proportion of the electorate*

that prefers A to B is $\left[\frac{2-(-10)}{20} + \frac{10-8}{20}\right]/2 = 35\%$;
that prefers B to C is $\left[\frac{10-6}{20} + \frac{10-(-1)}{20}\right]/2 = 37.5\%$; and
that prefers A to C is $\left[\frac{1-(-10)}{20} + \frac{10-0}{20}\right]/2 = 52.5\%$.

Again, B►A, C►B, and A►C.

Appendix A.4. Details for Example 4

For the third group (the one with 88% of the voters):

The perpendicular bisector of the line segment $\overline{P_A P_B}$ intersects $\overline{P_A P_B}$ at $(x, y) = (2, 8)$ and crosses the top of the square ($y = 10$) at $x = 2.4$, the bottom at $x = -1.6$, and the x-axis at $x = 0.4$. The proportion of the group that prefers A to B is thus $\frac{0.4-(-10)}{20} = 52\%$.

The perpendicular bisector of $\overline{P_B P_C}$ crosses $\overline{P_B P_C}$ at $(x, y) = (6, -1)$, the left side of the square ($x = -10$) at $y = 1$, the right side at $y = -1.5$, and the y-axis at $y = -0.25$. Thus, the proportion of the group that prefers B to C is $\frac{10-(-0.25)}{20} = 51.25\%$.

The perpendicular bisector of $\overline{P_A P_C}$ crosses $\overline{P_A P_C}$ at $(x, y) = (1, 0)$, the left side of the square at $y = -4\frac{8}{9}$, the right side at $y = 4$, and the y-axis at $y = -\frac{4}{9}$. The proportion of the group that prefers A to C is, therefore, $\frac{10-\left(-\frac{4}{9}\right)}{20} = 52.22\%$.

Within the third group, of course, no cycle exists: A►B, B►C, and A►C.

The proportions for the first two groups (35%, 37.5%, 52.5%) are the ones obtained for Example 3.2 in Appendix A.3 just above. Thus, across all three groups, the proportion of the electorate

that prefers A to B is (0.12 × 35%) + (0.88 × 52%) = 49.96%;
that prefers B to C is (0.12 × 37.5%) + (0.88 × 51.25%) = 49.6%; and
that prefers A to C is (0.12 × 52.5%) + (0.88 × 52.22%) = 52.26%.

Therefore, for the entire electorate, the result is B►A, C►B, and A►C.

Appendix A.5. Details for Example 5

Table A2 (top part) establishes the Condorcet cycle when voters' weightings are disparate. The bottom part of the table confirms that no cycle exists when all voters are alike in their weightings of the issues.

Table A2. Weighted [a] squared distance from each voter to each candidate, and preference rankings and matchups.

		Weighted [a] Squared Distance				Preference Matchups					
	Weight Ratio	Candidate			Ranking	*A* vs. *B*		*B* vs. *C*		*C* vs. *A*	
Voter(s)		*A*, (1, 7)	*B*, (3, −7)	*C*, (−9, 5)		*A*	*B*	*B*	*C*	*C*	*A*
*V*1, (−6, 2)	7:1	7·49 + 25 = 368	7·81 + 81 = 648	7·9 + 9 = 72	*C A B*	1	0	0	1	1	0
*V*5, (−6, 2)	1:7	49 + 7·25 = 224	81·7 + 81 = 648	9 + 7·9 = 72	*C A B*	1	0	0	1	1	0
*V*2, (6, 2)	7:1	7·25 + 25 = 200	7·9 + 81 = 144	7·225 + 9 = 1584	*B A C*	0	1	1	0	0	1
*V*6, (6, 2)	1:7	25 + 7·25 = 200	9 + 7·81 = 576	225 + 7·9 = 288	*A C B*	1	0	0	1	0	1
*V*3, (−6, −2)	7:1	7·49 + 81 = 424	7·81 + 25 = 592	7·9 + 49 = 112	*C A B*	1	0	0	1	1	0
*V*7, (−6, −2)	1:7	49 + 7·81 = 616	81 + 7·25 = 256	9 + 7·49 = 352	*B C A*	0	1	1	0	1	0
*V*4, (6, −2)	7:1	7·25 + 81 = 256	7·9 + 25 = 88	7·225 + 49 = 1624	*B A C*	0	1	1	0	0	1
*V*8, (6, −2)	1:7	25 + 7·81 = 592	9 + 7·25 = 184	225 + 7·49 = 568	*B C A*	0	1	1	0	1	0
*V*9, (0, 0)	1:1	1 + 49 = 50	9 + 49 = 58	81 + 25 = 106	*A B C*	1	0	1	0	0	1
All						5	4	5	4	5	4
*V*1 and *V*5	1:1	49 + 25 = 74	81 + 81 = 162	9 + 9 = 18	*C A B*	2	0	0	2	2	0
*V*2 and *V*6	1:1	25 + 25 = 50	9 + 81 = 90	225 + 9 = 234	*A B C*	2	0	2	0	0	2
*V*3 and *V*7	1:1	49 + 81 = 130	81 + 25 = 106	9 + 49 = 58	*C B A*	0	2	0	2	2	0
*V*4 and *V*8	1:1	25 + 81 = 106	9 + 25 = 34	225 + 49 = 274	*B A C*	0	2	2	0	0	2
*V*9	1:1	1 + 49 = 50	9 + 49 = 58	81 + 25 = 106	*A B C*	1	0	1	0	0	1
All						5	4	5	4	4	5

[a] Squared distances at the bottom of the table are unweighted.

Appendix A.6. Details for Example 6

Table A3 (top part) establishes the Condorcet cycle when voters' weightings are disparate. The bottom part of the table confirms that no cycle exists when all voters are alike in their weightings of the issues.

Table A3. Weighted [a] squared distance from each voter to each candidate, and preference rankings and matchups.

		Weighted [a] Squared Distance				Preference Matchups					
	Weight Ratio	Candidate			Ranking	*A* vs. *B*		*B* vs. *C*		*C* vs. *A*	
Voter		*A*, (1, 7)	*B*, (3, −7)	*C*, (−9, 5)		*A*	*B*	*B*	*C*	*C*	*A*
*V*5, (−6, 2)	1:7	49 + 7·25 = 224	81·7 + 81 = 648	9 + 7·9 = 72	*C A B*	1	0	0	1	1	0
*V*2, (6, 2)	7:1	7·25 + 25 = 200	7·9 + 81 = 144	7·225 + 9 = 1584	*B A C*	0	1	1	0	0	1
*V*3, (−6, −2)	7:1	7·49 + 81 = 424	7·81 + 25 = 592	7·9 + 49 = 112	*C A B*	1	0	0	1	1	0
*V*8, (6, −2)	1:7	25 + 7·81 = 592	9 + 7·25 = 184	225 + 7·49 = 568	*B C A*	0	1	1	0	1	0
*V*9, (0, 0)	1:1	1 + 49 = 50	9 + 49 = 58	81 + 25 = 106	*A B C*	1	0	1	0	0	1
All						3	2	3	2	3	2
*V*5	1:1	49 + 25 = 74	81 + 81 = 162	9 + 9 = 18	*C A B*	1	0	0	1	1	0
*V*2	1:1	25 + 25 = 50	9 + 81 = 90	225 + 9 = 234	*A B C*	1	0	1	0	0	1
*V*3	1:1	49 + 81 = 130	81 + 25 = 106	9 + 49 = 58	*C B A*	0	1	0	1	1	0
*V*8	1:1	25 + 81 = 106	9 + 25 = 34	225 + 49 = 274	*B A C*	0	1	1	0	0	1
*V*9	1:1	1 + 49 = 50	9 + 49 = 58	81 + 25 = 106	*A B C*	1	0	1	0	0	1
All						3	2	3	2	2	3

[a] Squared distances at the bottom of the table are unweighted.

Appendix A.7. Details for Statement about 1931 Finnish Election

Parts (a), (b), and (c) of Table A4, for respective common weight ratios for all six party groups of 2:1, 1:1, and 1:2, establish that a Condorcet cycle still occurs only in the third case and not in the first two.

Table A4. Weighted squared distance from each party group to each candidate, plus preference rankings and matchups.

(a) Weight ratio is common at 2:1 within each party group and across all party groups.													
		Weighted Squared Distance											
V [a]	**Spatial Location of *V***	**Candidate, [b] *G*, and His Spatial Location, P_G: (x_G, y_G)**				**Preference Ranking [b]**	**No. of Electors**	**Head-to-Head Preference Matchups [b]**					
		T	**St**	**K**	**Sv**			**St vs. K**		**K vs. Sv**		**Sv vs. St**	
		(−5, 0)	**(−1, 0)**	**(1, 1)**	**(2, 0)**			**St**	**K**	**K**	**Sv**	**Sv**	**St**
SD	(−5, 0)	0	32	73	98	T St K Sv	90	90	0	90	0	0	90
Pr	(−1, 0)	32	0	9	18	St K Sv T	52	52	0	52	0	0	52
Ag	(1, 1)	73	9	0	3	K Sv St T	69	0	69	69	0	69	0
Co	(2, 0)	98	18	3	0	Sv K St T	64	0	64	0	64	64	0
Sw1	(0, −1)	51	3	6	9	St K Sv T	7	7	0	7	0	0	7
Sw2	(1, −1)	73	9	4	3	Sv K St T	18	0	18	0	18	18	0
All							300	149	151	218	82	151	149
(b) Weight ratio is common at 1:1 within each party group and across all party groups.													
		Weighted Squared Distance											
V [a]	**Spatial Location of *V***	**Candidate, [b] *G*, and His Spatial Location, P_G: (x_G, y_G)**				**Preference Ranking [b]**	**No. of Electors**	**Head-to-Head Preference Matchups [b]**					
		T	**St**	**K**	**Sv**			**St vs. K**		**K vs. Sv**		**Sv vs. St**	
		(−5, 0)	**(−1, 0)**	**(1, 1)**	**(2, 0)**			**St**	**K**	**K**	**Sv**	**Sv**	**St**
SD	(−5, 0)	0	16	37	49	T St K Sv	90	90	0	90	0	0	90
Pr	(−1, 0)	16	0	5	9	St K Sv T	52	52	0	52	0	0	52
Ag	(1, 1)	37	5	0	2	K Sv St T	69	0	69	69	0	69	0
Co	(2, 0)	49	9	2	0	Sv K St T	64	0	64	0	64	64	0
Sw1	(0, −1)	26	2	5	5	St K = Sv T	7	7	0	$3\frac{1}{2}$	$3\frac{1}{2}$	0	7
Sw2	(1, −1)	37	5	4	2	Sv K St T	18	0	18	0	18	18	0
All							300	149	151	$214\frac{1}{2}$	$85\frac{1}{2}$	151	149

Table A4. *Cont.*

(c) Weight ratio is common at 1:2 within each party group and across all party groups.														
		Weighted Squared Distance												
V [a]	Spatial Location of V	Candidate, [b] G, and His Spatial Location, P_G: (x_G, y_G)				Preference Ranking [b]	No. of Electors	Head-to-Head Preference Matchups [b]					
		T	St	K	Sv			St vs. K		K vs. Sv		Sv vs. St	
		(−5, 0)	(−1, 0)	(1, 1)	(2, 0)			St	K	K	Sv	Sv	St
SD	(−5, 0)	0	16	38	49	T St K Sv	90	90	0	90	0	0	90
Pr	(−1, 0)	16	0	6	9	St K Sv T	52	52	0	52	0	0	52
Ag	(1, 1)	38	6	0	3	K Sv St T	69	0	69	69	0	69	0
Co	(2, 0)	49	9	3	0	Sv K St T	64	0	64	0	64	64	0
Sw1	(0, −1)	27	3	9	6	St Sv K T	7	7	0	0	7	0	7
Sw2	(1, −1)	38	6	8	3	Sv St K T	18	18	0	0	18	18	0
All							300	167	133	211	89	151	149

[a] V is for party group: SD = Social Democrats; Pr = Progressives; Ag = Agrarians; Co = Conservatives; Sw1, Sw2 = blocs 1, 2 of Swedish People's Party. [b] Candidates: T = Tanner, St = Ståhlberg, K = Kallio, Sv = Svinhufvud.

References

Abramson, Paul R., John H. Aldrich, and David W. Rohde. 2002. *Change and Continuity in the 2000 Elections*. Washington, DC: CQ Press.

Adams, James. 1997. Condorcet efficiency and the behavioral model of the vote. *The Journal of Politics* 59: 1252–63. [CrossRef]

Adams, James F., and Ernest W. Adams. 2000. The geometry of voting cycles. *Journal of Theoretical Politics* 12: 131–53. [CrossRef]

Caplin, Andrew, and Barry Nalebuff. 1988. On 64% majority rule. *Econometrica* 56: 787–814. [CrossRef]

Dasgupta, Partha, and Eric Maskin. 2004. The fairest vote of all. *Scientific American* 290: 92–97. [CrossRef]

Davis, Otto A., and Melvin J. Hinich. 1968. On the power and importance of the mean preference in a mathematical model of democratic choice. *Public Choice* 5: 59–72. [CrossRef]

Davis, Otto A., Melvin J. Hinich, and Peter C. Ordeshook. 1970. An expository development of a mathematical model of the electoral process. *American Political Science Review* 64: 426–48. [CrossRef]

Davis, Otto A., Morris H. DeGroot, and Melvin J. Hinich. 1972. Social preference orderings and majority rule. *Econometrica* 40: 147–57. [CrossRef]

Erikson, Robert S., and David W. Romero. 1990. Candidate equilibrium and the behavioral model of the vote. *American Political Science Review* 84: 1103–26. [CrossRef]

FairVote. 2019. Democratic Primary 2020 Interactive—20 Candidates. Available online: https://www.fairvote.org/interact_with_the_data?utm_campaign=yougov_poll&utm_medium=email&utm_source=fairvote (accessed on 25 September 2019).

FairVote. 2020. Report on FairVote/SurveyUSA National Poll of the 2020 Democratic Presidential Field, Table 2, "Head-to-Head Matchups". Available online: https://fairvote.app.box.com/s/8syue8jxdikiueknwaoedpar9q9w6roj (accessed on 22 March 2020).

Feld, Scott L., and Bernard Grofman. 1987. Necessary and sufficient conditions for a majority winner in n-dimensional spatial voting games: An intuitive geometric approach. *American Journal of Political Science* 31: 709–28. [CrossRef]

Feld, Scott L., Samuel Merrill, III, and Bernard Grofman. 2014. Modeling the effects of changing issue salience in two-party competition. *Public Choice* 158: 465–82. [CrossRef]

Gehrlein, William V. 2006. *Condorcet's Paradox*. Berlin: Springer.

Hinich, Melvin J., and Walker Pollard. 1981. A new approach to the spatial theory of electoral competition. *American Journal of Political Science* 25: 323–41. [CrossRef]

Hoyer, R. W., and Lawrence S. Mayer. 1974. Comparing strategies in a spatial model of electoral competition. *American Journal of Political Science* 18: 501–23. [CrossRef]

Hoyer, Robert W., and Lawrence S. Mayer. 1975. Social preference orderings under majority rule. *Econometrica* 43: 803–6. [CrossRef]

Laatu, Juho, and Warren D. Smith. 2009. The Rank-Order Votes in the 2009 Burlington Mayoral Election. Available online: https://rangevoting.org/JLburl09.txt (accessed on 26 September 2019).

Lagerspetz, Eerik. 2016. *Social Choice and Democratic Values*. Cham: Springer.

Maskin, Eric, and Amartya Sen. 2018. Maine tries a better way to vote. *New York Times*, June 11, A21.

Merrill, Samuel, III. 1988. *Making Multicandidate Elections More Democratic*. Princeton: Princeton University Press.

Miller, Nicholas R. 2015. The spatial model of social choice and voting. In *Handbook of Social Choice and Voting*. Edited by Jac C. Heckelman and Nicholas R. Miller. Cheltenham: Edward Elgar Publishing, pp. 163–81.

Niemi, Richard G., and Larry M. Bartels. 1985. New measures of issue salience: An evaluation. *The Journal of Politics* 47: 1212–20. [CrossRef]

Olson, Brian. 2009. IRV Failure in the Real World. Available online: http://bolson.org/~{}bolson/2009/20090303_burlington_vt_mayor.html (accessed on 26 September 2019).

Plott, Charles R. 1967. A notion of equilibrium and its possibility under majority rule. *The American Economic Review* 57: 787–806.

Polk, Jonathan, Jan Rovny, Ryan Bakker, Erica Edwards, Liesbet Hooghe, Seth Jolly, Jelle Koedam, Filip Kostelka, Gary Marks, Gijs Schumacher, et al. 2017. Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data. *Research & Politics* 4 (January–March). [CrossRef]

Potthoff, Richard F. 1970. The problem of the three-way election. In *Essays in Probability and Statistics*. Edited by Raj Chandra Bose, Indra M. Chakravarti, Prasanta Chandra Mahalanobis, Calyampudi Radhakrishna Rao and Kempton J. C. Smith. Chapel Hill: University of North Carolina Press, pp. 603–20.

Rabinowitz, George, James W. Prothro, and William Jacoby. 1982. Salience as a factor in the impact of issues on candidate evaluation. *The Journal of Politics* 44: 41–63. [CrossRef]

Regenwetter, Michel, Bernard Grofman, A. A. J. Marley, and Ilia Tsetlin. 2006. *Behavioral Social Choice: Probabilistic Models, Statistical Inference, and Applications*. Cambridge: Cambridge University Press.

Rivers, Douglas. 1988. Heterogeneity in models of electoral choice. *American Journal of Political Science* 32: 737–57. [CrossRef]

Saari, Donald G. 1995. *Basic Geometry of Voting*. Berlin: Springer.

Saari, Donald G. 1997. The generic existence of a core for *q*-values. *Economic Theory* 9: 219–60.

Schofield, Norman. 2008. *The Spatial Model of Politics*. London: Routledge.

Simpson, Paul B. 1969. On defining areas of voter choice: Professor Tullock on stable voting. *The Quarterly Journal of Economics* 83: 478–90. [CrossRef]

Tullock, Gordon. 1967. The general irrelevance of the general impossibility theorem. *The Quarterly Journal of Economics* 81: 256–70. [CrossRef]

Van Deemen, Adrian. 2014. On the empirical relevance of Condorcet's paradox. *Public Choice* 158: 311–30. [CrossRef]

economies

Article

A Bibliometric Analysis of the Developments and Research Frontiers of Agent-Based Modelling in Economics

Ayesha Zehra *,† and Amena Urooj †

PIDE School of Economics, Pakistan Institute of Development Economics, Islamabad 44000, Pakistan; amna@pide.org.pk
* Correspondence: ayeshatuzzehra_15@pide.edu.pk
† These authors contributed equally to this work.

Abstract: Dynamic Stochastic General Equilibrium (DSGE) models are widely used as a tool for policy decision-making. These models lost their fame when they could not predict the crisis in 2008 and could not address policy problems afterward. Meanwhile, the Agent-Based Modelling (ABM) approach emerged as an alternative to DSGE models. Between 2000 and 2020, this study examined scholarly research on the topic of ABM in economics. The information is gathered using the SCOPUS database. Numerous bibliometric indicators are provided, including the total number of publications and citations. The study reveals that agent-based modelling in economics research has grown in recent years. The majority of active research occurs in countries such as the United States of America, and collaboration allows researchers to reach out to many more countries. ABM has the potential to be applied in a wide range of economic fields. ABM also necessitates research into its own development to be used to better understand economic phenomena.

Keywords: agent-based modelling; bibliometric analysis; network analysis; conceptual structure; social structure

PACS: B41; E10; D70; D50

Citation: Zehra, Ayesha, and Amena Urooj. 2022. A Bibliometric Analysis of the Developments and Research Frontiers of Agent-Based Modelling in Economics. *Economies* 10: 171. https://doi.org/10.3390/economies10070171

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 12 April 2022
Accepted: 13 June 2022
Published: 19 July 2022

1. Introduction

Modelling is defined as a mathematical and statistical way of reproducing events and their possible consequences due to policy decisions (Shanahan et al. 2016). In econometric modelling, models formed on a theoretical framework are constructed using one or many exogenous variables, identifying quantitative relationships which generate various responses. Dynamic Stochastic General Equilibrium (DSGE) models have ruled as a tool for policy decisions. Then the 2008 financial crisis became the downfall of this modelling technique because of "no response" to policy problems afterwards. These models cannot predict a crisis or any non-linear event. Meanwhile, the Agent-Based Modelling (ABM) approach emerged as an alternative to DSGE models. ABM emergence property allows foreseeing complex behaviours. The Great Recession made policymakers see the "economy as [a] complex evolving system" consisting of heterogeneous agents and non-equilibrium state continuously change the economy's structure.

Bibliometric analysis is a method to explore the information-rich environment on research activities and findings extracted through data from research publications in academic journals and their citations. Bibliometric indicators help investigate the knowledge structure of a particular field and its scope in the future. In the framework of research developments, questions like "where are we now?" and "where will we be in the future?" are answered by this form of analysis. There are two types of bibliometric analysis techniques: (1) performance analysis and (2) science mapping. In essence, performance analysis considers the contributions of research parts, whereas science mapping considers the relationships among them.

This article is dedicated to exploring the research development of agent-based modelling in economics using bibliometric analysis techniques. This study aims to use bibliometric analytic approaches to investigate the knowledge base of agent-based modelling in economics. We used both performance analysis metrices and scientific mapping methodologies to achieve the study's research aims. Also, ABM in economics is examined in terms of its conceptual and social structure.

2. What ABM Has and What DSGE Lacks

The primary instrument for generating policy judgments remains DSGE models. These models did not predict the financial crisis of 2008. These models' limitation in addressing many policy concerns is their data-driven approach and lack of macroeconomic data. These models' openness and transparency is a strength, but it also leaves them vulnerable to criticism. It is possible to draw attention to suspicious assumptions. Evidence of inconsistencies are readily apparent. It is possible to identify elements that are not included in the model (Christiano et al. 2018). DSGE modelling is a death tale that has been predicted. Joseph Stiglitz writes, "... much of the main parts of the DSGE model are defective—sufficiently seriously wrong that they do not give even a good starting point for creating a good macroeconomic model" (Stiglitz 2018). Vines and Wills want DSGE models to be able to achieve what they desire, which is to allow modellers to get a quick look at crucial issues. Central banks routinely use estimated DSGE models for forecasting and quantitative policy analysis. Estimating these models and interpreting the results to formulate policy are both difficult tasks (Schorfheide 2011). The issues and flaws of DSGE models are similar to those of generalised equilibrium (GE) models. Many researchers have discovered that the agents in DSGE models cannot be constrained so that their uniqueness and stability are preserved.

Furthermore, assuming individual rationality does not imply aggregate rationality. Because reactions to shocks or parameter changes may not resemble in aggregate, the representative agent assumption in these models is not reliable for policy analysis. In DSGE models, solving systems of equations can lead to another difficulty of identification, resulting in skewed estimates of some structural parameters and raising doubts about statistical significance. This modelling technique cannot predict infrequent economic crises, which is not surprising given that fat tail densities are approximated distributions of macroeconomic time series (Fagiolo et al. 2008), and Gaussian distributed shocks are a typical assumption in DSGE models. The assumption that Representative Agents (RAs) are rational prevents these models from addressing distributional issues because it implies that one: agents are fully aware of the economy; two: agents are capable of understanding and solving any problem they encounter without making mistakes; and three: agents are aware that all others follow the same pattern. These issues demonstrate that DSGE models are ill-equipped to solve policy concerns and cannot forecast future crises. The DSGE approach is so enthralled by its internal logic that it confuses the model's precision with the real one. (Caballero 2010).

ABM has evolved rapidly in economics over the previous two decades. Due to the following features of this modelling method;

- Bottom-up Perspective,
- Heterogeneity,
- A complex system approach that is always evolving,
- Non-linearity,
- Endogeneity vs direct interactions,
- Rationality with bounds,
- Agents' ability to learn,
- a market mechanism that is based on selection,

These models are computer simulations that use a top-down strategy to investigate developing dynamic patterns. Policies and the social behaviours that result from them act like a weather system constantly battered by storms and invasions. The ability to make large-scale modifications and crash systems are inherited. External disturbances throw the equilibrium condition off. ABM allows little effects like herding and fear driving bubbles and crashes to be amplified via feedback processes. Models are non-linear in mathematical terms, implying that the result may not be proportional to the cause. The capacity to represent emergent phenomena resulting from the interaction of each agent is a major advantage of this modelling technique. Emergent phenomena can have traits that are opposed to those of their constituents. Agent-based models are a natural way to describe a system of behavioural elements. ABM can explain that designing a virtual agent with a shopping basket is more natural than describing average effects using a synthetic basket density. The flexibility of these models allows for the addition of new agents and changes in behaviour, learning, evolution, and complexity by altering interaction rules.

3. Bibliometric Analysis: A Technique of Systematic Literature Review

One of the most important knowledge discovery methods is synthesising the results of earlier studies. The use of bibliometric analysis is growing in popularity (Zupic and Čater 2015). In a qualitative study of published research papers, journals, and books, the bibliometric technique has been employed (Ellegaard and Wallin 2015). It aids in the identification of frequently referenced authors and institutions, related publications, and the keywords most commonly used in a given study filed (Daim et al. 2006). Furthermore, bibliometric analysis can be used to assess the publication's popularity among specialists and verify the author's reputation (Ball and Tunger 2005). It also aids in literature review by leading the researcher to influential research works or publications, as well as objectively mapping the study field (Zupic and Čater 2015). (Donthu et al. 2021) discused in detail the methodology to conduct bibliometric analysis and concluded that bibliometric analysis can aid knowledge generation not just in business research but also in other sectors, thanks to a better comprehension of science.

Bibliometric approaches are used for a variety of purposes, including performance analysis and science mapping (Cobo et al. 2011). Performance analysis is used to assess individual, institutional, and individual research and publishing performance. A generic approach of domain analysis and visualization is science mapping. A scientific discipline, a field of research, or topic areas related to specific research topics can all be included in the scope of a science mapping study. In other words, an area of scientific knowledge expressed through an aggregated collection of intellectual contributions from members of a scientific community or more clearly defined specialty is the unit of analysis in science mapping (Chen 2017).

Citation analysis, co-citation analysis, bibliographic coupling, co-word analysis, and co-authorship analysis are some of the methodologies utilised in science mapping. Such methodologies are beneficial for illustrating the bibliometric and intellectual structure of a study field when combined with network analysis (Baker et al. 2020; Tunger and Eulerich 2018).

4. Study Design

The goal of this study was to fulfil two purposes. The first was to determine research trends over time, and the second was to investigate research content in order to assess the application of agent-based modelling techniques in various economic sectors. Hence we combined both bibliometric methodologies; performance analysis and science mapping to reach our objectives.

Research Objectives: The analysis is constructed in such a way as to achieve the following goals.

- Identifying the knowledge base of agent-based modelling and its intellectual structure particularly in economics.
- Examine the research front/conceptual structure of agent-based modelling in context of monetary policy.
- Exploring the social network structure of agent-based modelling in economics.

Research Design: Many methods are described in Section 3 of the bibliometric analysis. Citation and co-citation analysis (by author and journal), co-word analysis, and network analysis were among the bibliometric methodologies we considered.

Bibliometric Data Collection: The Scopus database offered a total of 1568 documents for examination. A sophisticated keyword selection was required for data extraction in order to provide a relevant set of data. The keywords chosen must match the following four criteria: high search volume, relevancy, high conversion value, and low competition. The search terms "agent-based modelling" AND ("DSGE" OR "monetary policy" OR "crisis" OR "central banks") were used to extract bibliometric data.

Inclusion Criteria: After filtering the data based on the inclusion and exclusion criteria, a total of 1568 data were gathered. Three inclusion rules were followed: (1) Articles in which one of the keywords appears in the title, abstract, or keywords (2) The publication date ranges from 2000 to 2020. (3) Journal articles, conference papers, and book chapters. If they met all inclusion criteria, English language abstracts were included in the bibliometric review.

Exclusion Criteria: All documents with a core subject of agent-based modelling but not relevant to the field of economics were left out of the analysis.

Methodology and Software: We used the Scopus dataset in the analysis and the approaches listed above to answer our research questions. R was chosen as software for both visuals and quantitative analysis.

5. Knowledge Base of ABM: Results and Findings

5.1. Productivity Assessment

After the financial crisis of 2008, the agent-based model became a widely researched topic. The crisis was not predicted by DSGE models, and they also did not respond to policy questions. Since the most publications were in 2020, ABM has become a new topic in its development. ABM is also the ideal tool for experimenting with different policy scenarios during a pandemic (Figure 1).

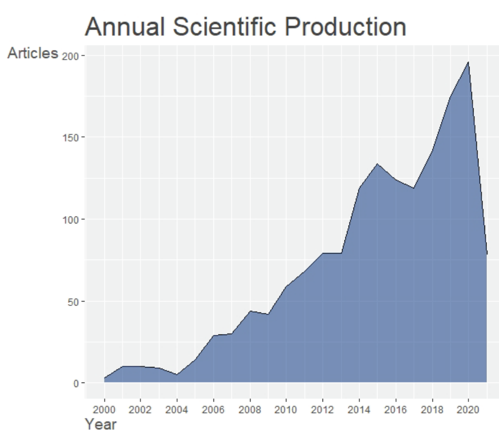

Figure 1. Annual scientific output on economics-related ABM: A Scopus database analysis (2000–2020).

Countries such as the United States have played a critical role in the field's continuous progress. The authors are working with researchers from the same country as well as from other ones. The top nations and authors working on the issue of agent-based modelling in economics are shown in Figure 2. According to the findings, American researchers are putting a greater emphasis on this modelling technique and examining its possibilities for solving difficult challenges. Researchers prefer to collaborate with researchers from their own nation rather than researchers from other countries, according to stacked bar charts.

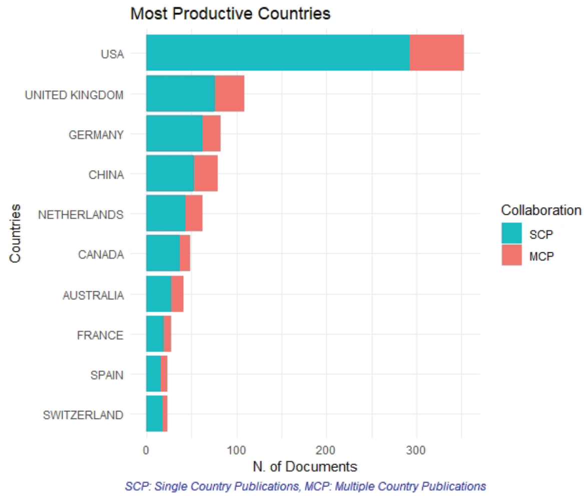

Figure 2. Ranking of countries based on scientific output on economics-related ABM (2000–2020).

Wilensky, being the most active contributor, also described how to use agent-based simulations to answer complex questions. His writings capture the thrill of re-creating social phenomena in computer simulations to better understand them (Wilensky and Rand 2015) (See Figure 3). In Figure 4, we can see that in which year the authors were most productive and long lived. There are authors from economics who were working on agent-based modelling before the financial crisis took place. But the focus of their research was not directed to answer the policy questions to overcome the aftershocks of crisis.

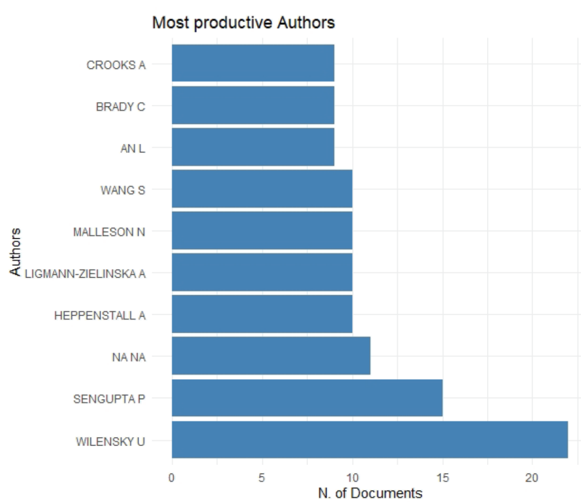

Figure 3. Ranking of authors based on scientific output on economics-related ABM (2000–2020).

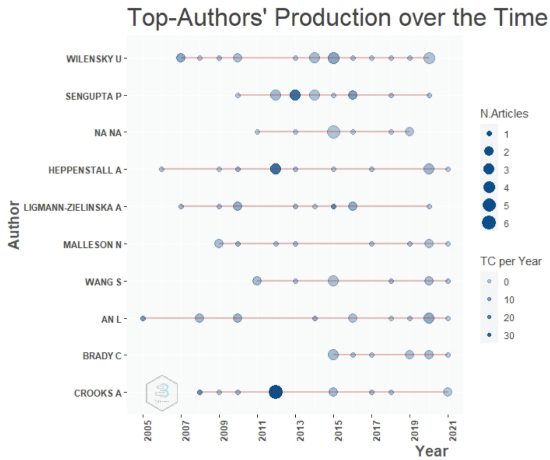

Figure 4. An overview of the author's output over the years (2000–2020).

Lotka's Law is one of the most fundamental bibliometric rules, and it deals with the frequency with which authors in a specific subject publish. The frequency of publishing by authors in a specific field is described by Lotka's law, presented as follows:

$$f(x) = \frac{\beta}{x^{\alpha}}$$

where $f(x)$ is the frequency of authors having x publications, and x is the positive integer, representing number of publications. The parameter estimates of Lotka's Law are $\beta = 2.82$ and $\alpha = 0.97$. According to the findings, there are 2833 writers with a single ABM publication in economics. There are approximately 70 authors with at least five published works. Over 20 documents were released by only one contributor. Theoretical and observed frequency are depicted graphically in Figure 5.

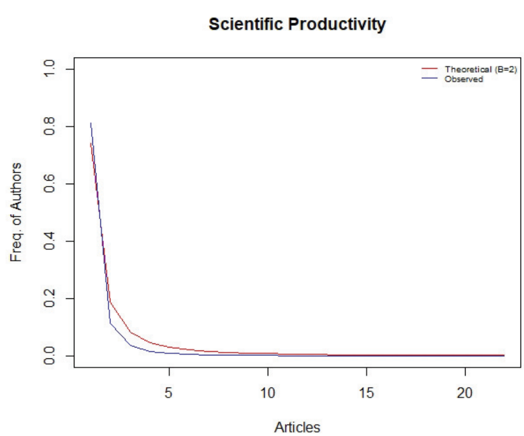

Figure 5. Lotka's law of scientific productivity from 2000–2020 (authors publishing on economics-related ABM).

A journal's impact on a specific study topic can be measured by its publications. It can be seen from Table 1 regarding ABM research in economics, Journal of Artificial Societies and Social Simulations (JASSS) is head and shoulders above in the league table. JASSS has the highest number of publications on the theme of agent-based modelling in Scopus index journals followed by Computer environment and urban system on the league table.

Table 1. Most Relevant Sources.

SR No	Sources	Publications
1	JASSS	210
2	COMPUTERS ENVIRONMENT AND URBAN SYSTEMS	42
3	SUSTAINABILITY (SWITZERLAND)	36
4	INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE	24
5	SOCIAL SCIENCE COMPUTER REVIEW	21
6	ENVIRONMENT AND PLANNING B: PLANNING AND DESIGN	20
7	TRANSPORTATION RESEARCH PART C: EMERGING TECHNOLOGIES	19
8	TRANSPORTATION RESEARCH PROCEDIA	19
9	JOURNAL OF INDUSTRIAL ECOLOGY	15
10	ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION	14

The sum of published documents on ABM in the top five journals is around 450 in ten years. As outlined in Table 2, authors are ranked via the Dominance Factor (DF). (Kumar and Kumar 2008) developed the formula as,

$$\text{DF} = \frac{\text{num of multi-authored publications of an author as first author (Nmf)}}{\text{total num of multi-authored publications (Nmt)}}.$$

The value of dominance factor indicates collaboration in the field. A value less than 0.5, reflects a good sign for collaboration. Authors who have published nine or more publications on the theme of agent-based modelling are selected and their dominance factor is calculated by using the above formula. Sengupta and Wilensky are top authors with respect to publication number, i.e., 15 and 22, but they rank 6th and 10th, respectively. If the authors' dominance factor values are less than 0.5, this is a good sign of collaboration.

Table 2. Dominance factor Ranking.

Rank	Author	Dominance Factor	Total Publications	Single Authored	Multi Authored	First Authored	Rank by Publications
1	AN L	0.67	9	0	9	6	6
2	MALLESON N	0.6	10	0	10	6	3
3	LIGMANN-ZIELINSKA A	0.57	10	3	7	4	3
4	CROOKS A	0.5	9	1	8	4	6
5	TANG W	0.5	9	1	8	4	6
6	SENGUPTA P	0.2	15	0	15	3	2
7	BRADY C	0.125	9	1	8	1	6
8	GILBERT N	0.11	9	0	9	1	6
9	HEPPENSTALL A	0.1	10	0	10	1	3
10	WILENSKY U	0.04	22	0	22	1	1

5.2. Importance Assessment

Table 3 is about author level metrics based on three indices. *H-index* measures the productivity as well as the impact of publication. The H-index is calculated as "author has H publications, and each publication has H or more citations". Whereas *g-index* is one variant of H-index which gives credits for highly cited authors in the data set. (Hirsch 2005), its inventor says: highly cited papers play a key role in the determination of H-index. The selected paper for the top h category are then dropped for the further determination of the H-index over time. This means that the H-index of the subsequent years are not influenced by the papers of the top category, even if the number of citations increases over time. Value of g-index is always either equal to or greater than H-index.

The *m-index* is another form of H-index. The M index in the Table 3 shows the comparison of authors within the field of agent-based modelling but with very different career lengths. The h-index is constrained by the fact that it is time-based and field-specific, and it ignores highly cited works. In the early identification of young researchers, bibliometrics that account for time, such as the m-index, should be evaluated, ideally in conjunction with critical peer review. The m-index has the most potential for identifying early-stage high-potential researchers. An m-index of 1 is normal, 1–2 is above average, and >2 is exceptional, according to a suggested rule of thumb for interpreting the index (Ndwandwe et al. 2021).

Table 3. h-index top authors.

Sr No.	Authors	h-Index	g-Index	m-Index
1	AN L	6	7	0.352941176
2	BRADY C	3	6	0.428571429
3	CROOKS A	7	7	0.5
4	HEPPENSTALL A	5	8	0.3125
5	LIGMANN-ZIELINSKA A	7	9	0.466666667
6	MALLESON N	6	7	0.461538462
7	SENGUPTA P	7	12	0.7
8	WANG S	5	9	0.454545455
9	WILENSKY U	8	16	0.533333333

We attempted to plot average citations and average total citations (see Figure 6) over time in order to study the most cited literature on the application of ABM in economics. We can readily see that total citations were much higher for work published prior to 2008, with a downward trend after that. On the other side, due to the pandemic, average citations climbed in 2020.

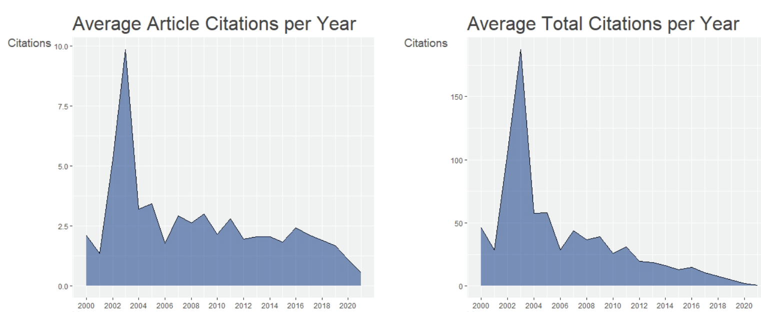

Figure 6. Citation analysis of published scientific documents.

5.3. Network Analysis

Figure 7 shows the social network maps of the co-occurance matrix, collaboration matrix and coupling.The size of the nodes reflects the frequency of keywords in each cluster. A larger size suggests a stronger citation burst and suggests high significance of the subfield. The conceptual structure captured through keyword co-occurrence indicates diversity within the research sub-fields.

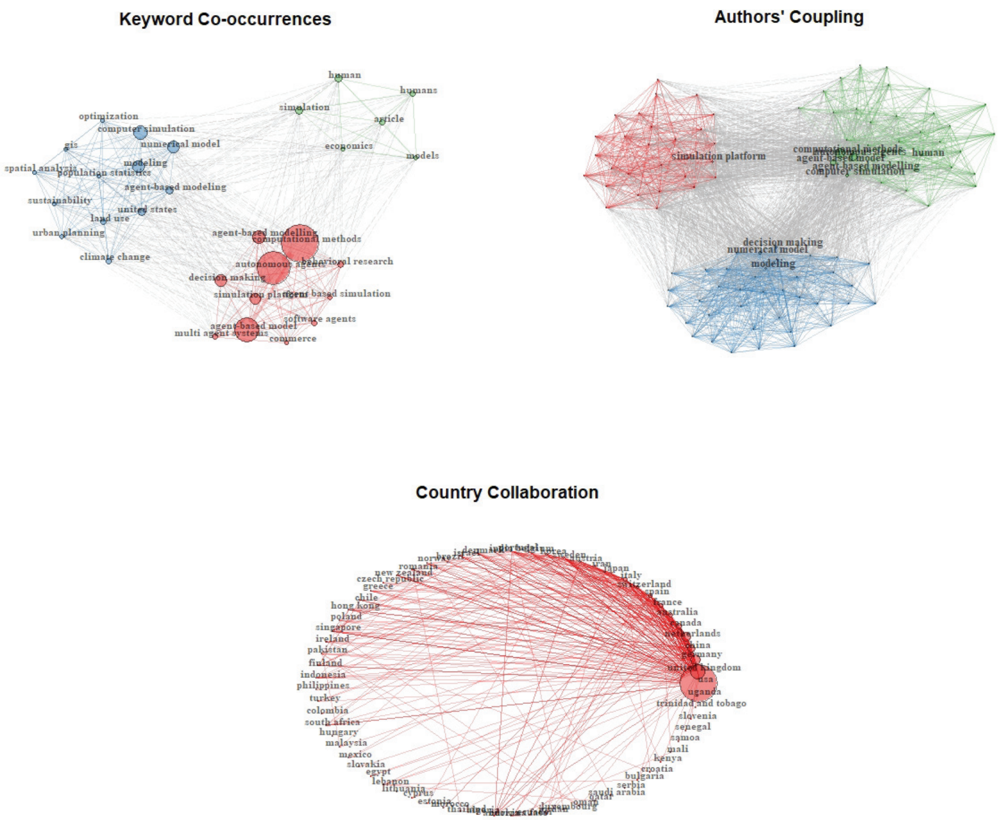

Figure 7. Network Analysis based on SCOPUS database.

The network indicates three clusters in Keyword Co-occurrence and Author's Coupling. Each cluster represents a theme/field of study in economics research using ABM. Three major themes emerge from the Keyword Co-occurrence network: human behavioral research, climate change and urban development, and the development of agent-based models that may be used to investigate various economic phenomena. The citing authors in the topic area that are mapped are known as author coupling, and these maps can be used to focus on research areas that are shared by many currently active authors. These are split into three groups in this instance. One of the primary cluster authors shares simulation platforms, while the other two work on computational simulations and numerical models together.

The expansion of international collaboration in ABM research was placed in a highly stratified way, resulting in a clear divide between the main contributing countries and many others who collaborated globally on a more occasional basis. The network is made up of a core that is dominated by research outputs from scholars in research-intensive countries, with numerous additional countries gravitating around that core—unsurprisingly, given the emphasis on English-language journals.

5.4. Research HotSpots

Conceptual structure of agent-based modelling in economics is shown in a thematic map (Figure 8). The map was constructed by using keywords with min word frequency of 250. Minimum cluster frequency per 1000 documents is 5 and number of labels to each cluster are 3. On the x-axis we have centrality, which measures degree of interaction of one network with other networks, while density on the y-axis is measure of internal strength of a network. Themes on the upper right quadrant, i.e., motor themes have well developed internal ties and are important for the structure of the research field. Motor skills have strong centrality and high density. Niche themes are of marginal importance in the research field as they have well developed internal ties but unimportant external ties. Emerging or disappearing themes with low density and low centrality are weakly developed and marginal. Whereas basic themes are of high importance for the field of ABM in economics, but these are not well developed.

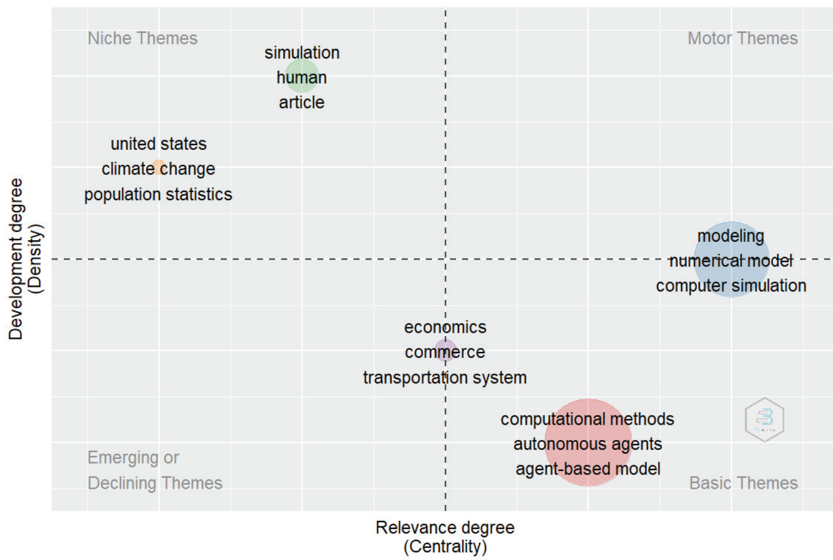

Figure 8. Thematic Map in Agent-based Modelling Research.

We discovered that research on agent-based modelling in economics can be classified in two ways, the first being the development of an agent-based model, and the second being the use of a developed agent-based model to investigate policy possibilities, based on Keyword Co-occurance analysis. The development of agent-based models employing economic theories, the application of diverse computational methods to calibrate these models, and the simulation of these models are all major topics in ABM currently. Although the largest bubble incorporates numerical models, computer simulations, and computational methods, these motor themes and basic themes both require a significant amount of research to get to a good measure of density and centrality.

6. Research Frontiers

There are several intriguing options for agent-based modelling research. ABM's versatility in applying to new issues has always been one of its best attributes. While certain classes of models have been established in fields like macroeconomics or financial markets, ABM has always been a transdisciplinary methodology that can be applied to problems involving a variety of rules, interactions, and behavioral phenomena (Steinbacher et al. 2021). ABMs can also be used to investigate issues that arise due to greater AI use, such as the societal impact of ranking algorithms and recommender systems and the potential reinforcement of social inequities and biases. ABMs can be used to build priors for machine learning algorithms in a semi-supervised manner in cases where the given data are noisy or biassed, reducing errors and preventing the amplification of distortions. Artificial agents can also be incorporated into large-scale simulators once they have been built based on the behavior of human subjects (Dosi et al. 2020). Such synergy between ABM and experimental technique is still in its infancy, but it represents an exciting avenue for future research, in our opinion. Further research on the estimation of ABMs is also required, as little is known about the benefits and drawbacks of various techniques. The majority of current models allow for the formulation and stochastic approximation of a likelihood function. As models become more complicated, such approximations will become increasingly difficult. In such instances, Approximation through Bayesian Computation methods and GMM/SMM should be considered (ABC). This approach (Sisson et al. 2007; Toni et al. 2009) employs measurements (moments) of the data other than the likelihood to approximate the posterior distribution of the parameters using a rejection sampling or Markov Chain Monte Carlo technique. While this approach has gained much traction in ecological ABMs (Csilléry et al. 2010), economic applications are still a work in progress.

7. Conclusions

During the 2008 financial crisis, the discussion began when economists began to investigate the possibility of agent-based modelling techniques for answering policy problems and performing "what-if" scenarios to aid policy decisions. This discussion about the future of agent-based modelling has yielded the desired result, with researchers currently working on developing agent-based models so that simulations based on simple rules can portray the complex economy. The European Central Bank is funding projects to construct agent-based economic models. The science of economics has long been in need of more robust methodologies that do not assume reasonable expectations and do not encourage optimism about the behaviours of agents. The economy appears to be a system, yet individual decisions cause the system's complex nonlinearity. Similarly, agent-based models recognise individual interactions and adapt in response to them.

The objective of the review analysis was to investigate the current developments in economics using this new modelling technique. According to the study's findings, research on the topic of agent-based modelling in economics is growing at a quick pace. Researchers experiment and publish their findings in research papers, books, and conference papers. Researchers are increasingly collaborating in order to improve the quality of their publications. The United States of America (USA), the United Kingdom (UK), Germany, and China are the most productive countries. The most prominent research on this topic has been published in the United States of America (USA), and American-based journals have taken the lead in publishing research in this field. This might be due to the increased formation trend of academic journals in the USA. The most often used keywords by the researchers (e.g., decision making, sustainability, and commerce) indicate the hotspots in ABM research. The main purpose of adopting different bibliometric analysis methodologies was to uncover research trends and the substance of published work. The knowledge structure and research trends were discovered through co-occurrence analysis. According to the findings, several well-established economic themes benefit from ABM techniques, but many require them. Researchers are not introducing central banks in an agent-based economy and conducting more monetary policy experiments. However, monetary policy had a considerable impact during and after the Great Recession. The channels of collaboration among scholars worldwide were uncovered through social network analysis. In the field of agent-based modelling, researchers promote cross-national and intra-national collaboration, which fosters the creation of new ideas. Although there has been much research into using agent-based modelling to comprehend the complexities of economic problems, underdeveloped countries like Pakistan have been slow to adopt this modelling technique. Even after a decade of economic disaster, economists discover agent-based modelling. This policy decision modelling technique is not widely used in economics, especially in monetary policy issues. Agent-based models, in addition to existing ones, could be critical instruments for assessing economic policy.

Author Contributions: A.Z. developed the proposed research model by reviewing relevant literature, analysing the data, and writing the paper, as well as revising the manuscript; A.U. contributed to the development of both research and practical implications in the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data for this study was collected from the SCOPUS database and is available upon request.

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

Abbreviations

The following abbreviations are used in this manuscript:

ABM	Agent-Based Modelling
DSGE	Dynamic Stochastic General Equilibrium

References

Baker, H. Kent, Satish Kumar, and Nitesh Pandey. 2020. A bibliometric analysis of managerial finance: A retrospective. *Managerial Finance* 46: 1495–517. [CrossRef]

Ball, Rafael, and Dirk Tunger. 2005. *Bibliometric Analysis Data, Facts and Basic Methodological Knowledge Bibliometrics for Scientists, Science Managers, Research Institutions and Universities*. Julich: Research Center Julich, vol. 12.

Caballero, Ricardo J. 2010. Macroeconomics after the crisis: Time to deal with the pretense-of-knowledge syndrome. *Journal of Economic Perspectives* 24: 85–102. [CrossRef]

Chen, Chaomei. 2017. Science mapping: A systematic review of the literature. *Journal of Data and Information Science* 2: 1–40. [CrossRef]

Christiano, Lawrence J., Martin S. Eichenbaum, and Mathias Trabandt. 2018. On dsge models. *Journal of Economic Perspectives* 32: 113–40. [CrossRef]

Cobo, Manuel J., Antonio Gabriel López-Herrera, Enrique Herrera-Viedma, and Francisco Herrera. 2011. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *Journal of Informetrics* 5: 146–66. [CrossRef]

Csilléry, Katalin, Michael G. B. Blum, Oscar E. Gaggiotti, and Olivier François. 2010. Approximate bayesian computation (abc) in practice. *Trends in Ecology & Evolution* 25: 410–8.

Daim, Tugrul U., Guillermo Rueda, Hilary Martin, and Pisek Gerdsri. 2006. Forecasting emerging technologies: Use of bibliometrics and patent analysis. *Technological Forecasting and Social Change* 73: 981–1012. [CrossRef]

Donthu, Naveen, Satish Kumar, Debmalya Mukherjee, Nitesh Pandey, and Weng Marc Lim. 2021. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research* 133: 285–96. [CrossRef]

Dosi, Giovanni, Mauro Napoletano, Andrea Roventini, Joseph E. Stiglitz, and Tania Treibich. 2020. Rational heuristics? expectations and behaviors in evolving economies with heterogeneous interacting agents. *Economic Inquiry* 58: 1487–516. [CrossRef]

Ellegaard, Ole, and Johan A. Wallin. 2015. The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics* 105: 1809–31. [CrossRef] [PubMed]

Fagiolo, Giorgio, Mauro Napoletano, and Andrea Roventini. 2008. Are output growth-rate distributions fat-tailed? some evidence from oecd countries. *Journal of Applied Econometrics* 23: 639–69. [CrossRef]

Hirsch, Jorge E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences* 102: 16569–72. [CrossRef] [PubMed]

Kumar, Sudhir, and Surendra Kumar. 2008. Collaboration in research productivity in oil seed research institutes of india. In *Proceedings of Fourth International Conference on Webometrics, Informetrics and Scientometrics*. Berlin: Institute for Library and Information: Humboldt-Universitat zu Berlin, vol. 28.

Ndwandwe, M., David Bishop, Robert Wise, and Reitze Rodseth. 2021. Bibliometrics to assess the productivity and impact of medical research. *South African Journal of Higher Education* 35: 224–36. [CrossRef]

Schorfheide, Frank. 2011. *Estimation and Evaluation of Dsge Models: Progress and Challenges*. Cambridge: Cambridge University Press.

Shanahan, Marian, Nagesh Shukla, Pascal Perez, Michael Farrell, and Alison Ritter. 2016. A systematic review of modelling approaches in economic evaluations of health interventions for drug and alcohol problems. *BMC Health Services Research* 16: 1–14.

Sisson, Scott A., Yanan Fan, and Mark M. Tanaka. 2007. Sequential monte carlo without likelihoods. *Proceedings of the National Academy of Sciences* 104: 1760–65. [CrossRef] [PubMed]

Steinbacher, Mitja, Matthias Raddant, Fariba Karimi, Eva Camacho Cuena, Simone Alfarano, Giulia Iori, and Thomas Lux. 2021. Advances in the agent-based modeling of economic and social behavior. *SN Business & Economics* 1: 1–24.

Stiglitz, Joseph E. 2018. Where modern macroeconomics went wrong. *Oxford Review of Economic Policy* 34: 70–106.

Toni, Tina, David Welch, Natalja Strelkowa, Andreas Ipsen, and Michael P. H. Stumpf. 2009. Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface* 6: 187–202. [CrossRef] [PubMed]

Tunger, Dirk, and Marc Eulerich. 2018. Bibliometric analysis of corporate governance research in german-speaking countries: Applying bibliometrics to business research using a custom-made database. *Scientometrics* 117: 2041–59. [CrossRef]

Wilensky, Uri, and William Rand. 2015. *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. Cambridge: MIT Press.

Zupic, Ivan, and Tomaž Čater. 2015. Bibliometric methods in management and organization. *Organizational Research Methods* 18: 429–72. [CrossRef]

economies

Article

Moderation Effects of Government Institutional Support, Active and Reactive Internationalization Behavior on Innovation Capability and Export Performance

Antonio Moreira [1,2,3,*], Eurico Navaia [1,4] and Cláudia Ribau [2,5]

1 Department of Economics, Management, Industrial Engineering and Tourism, University of Aveiro, 3810-193 Aveiro, Portugal; eunavaia@gmail.com
2 GOVCOPP—Research Unit on Governance, Competitiveness and Public Policies, University of Aveiro, 3810-193 Aveiro, Portugal; claudiapiresribau@gmail.com
3 INESCTEC—Institute for Systems and Computer Engineering, Technology and Science, 4200-465 Porto, Portugal
4 Faculdade de Ciências Sociais e Humanidades, Universidade Zambeze, Beira 369, Sofala, Mozambique
5 Instituto Superior de Contabilidade e Administração, University of Aveiro, 3810-193 Aveiro, Portugal
* Correspondence: amoreira@au.pt

Abstract: Although innovation capabilities are important drivers of export performance, few studies address how they influence export performance in the context of emerging economies. This paper evaluates the moderating effects of government institutional support and firms' active and reactive internationalization behaviors on the relationship between innovation capabilities and export performance. The sample analyzed is based on 250 Mozambican small and medium enterprises (SMEs). The results indicate that although innovation capabilities positively influence the export performance of Mozambican SMEs, the moderating effects of government institutional support and firms' active and reactive internationalization behaviors were not found to be statistically significant.

Keywords: SMEs; export performance; innovation capabilities; government institutional support; active internationalization behavior; reactive internationalization behavior; Mozambique

Citation: Moreira, Antonio, Eurico Navaia, and Cláudia Ribau. 2022. Moderation Effects of Government Institutional Support, Active and Reactive Internationalization Behavior on Innovation Capability and Export Performance. *Economies* 10: 177. https://doi.org/10.3390/economies10080177

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 30 June 2022
Accepted: 19 July 2022
Published: 22 July 2022

1. Introduction

Small and medium-sized enterprises (SMEs) represent over 96% of Mozambique's business population (Instituto Nacional de Estatística 2017), but their contribution in terms of gross domestic product (GDP) and employment is still relatively low. SMEs contribute 28% of GDP and 42% in formal employment, facing challenges at the business environment level, namely in terms of access to markets, access to finance, and coordination of support mechanisms (Ministério da Indústria e Comércio 2016). SMEs need to enhance their competitiveness to increase their economic contribution in emerging economies. Thus, product and process innovation can contribute to promoting structural change among developing countries and support the international competitiveness of SMEs (Ministério da Indústria e Comércio 2016).

It is important for SMEs to invest in their innovation capabilities (ICs), as they are considered important levers of product and process innovation (Blanchard 2020) and important sources of competitive advantage (Guan and Ma 2003; Knight and Kim 2009; Sen and Egelhoff 2000; Wang and Ahmed 2004), positively influencing the export performance of SMEs (Guan and Ma 2003; Oura et al. 2016; Ribau et al. 2017a). Based on the concept of dynamic capabilities, ICs are recognized as key drivers of business growth (Teece et al. 1997).

Government programs supporting internationalization play a primary role in the development of businesses, with positive consequences for national economies (Jalali 2012). These programs play an important role for SMEs, as they have some resource constraints, especially when competing in international markets with larger and more experienced companies. Thus, it is not uncommon for SMEs to seek to explore international markets, even though they have many limitations on resources and skills and little international experience (Freixanet 2012; Ayob and Freixanet 2014). Government support, in these situations, is crucial for some SMEs to overcome their limitations and be able to compete in international markets. Take, for example, export support programs (Comi and Resmini 2020; Malca et al. 2020). Likewise, given that SMEs have a multifaceted role and are involved in increasing internationalization processes, international support programs have been of added value in allowing SMEs to adjust their resources to contexts in order to progress to broader international markets (Francis and Collins-Dodd 2004; Mota et al. 2021).

While there is empirical research on the determinants of export performance (Guan and Ma 2003; Vicente et al. 2015; Oura et al. 2016; Ribau et al. 2017a), there are some studies addressing export performance in the context of emerging economy environments (Guan and Ma 2003; LiPuma et al. 2013; Malca et al. 2020; Oura et al. 2016). In addition, the importance of dynamic capabilities and ICs in the context of emerging economies has been overlooked as there are few studies addressing the relationship between ICs and export performance (Guan and Ma 2003; Krammer et al. 2018; Oura et al. 2016). Moreover, government support to firms is expected to create differential value in export performance (Freixanet 2012). Indeed, if firms need institutional support to increase their competitiveness in foreign markets (Francis and Collins-Dodd 2004; Krammer et al. 2018) in emerging economies firms, due to the lack of proper resources, which normally depend on institutional support to increase their competitiveness in foreign markets (LiPuma et al. 2013; Yi et al. 2013).

Another aspect that influences export performance is the type of active/reactive internationalization (Westhead et al. 2004; Ribau et al. 2017b). Proactive stimuli can result from SMEs' behavior and a deliberate search for market opportunities abroad, with the external environment being the source of these stimuli (Westhead et al. 2004). Moreover, firms embracing active internationalization strategies experience better export performance than those implementing reactive internationalization strategies (Ribau et al. 2017b). However, little is known on how active/reactive internationalization strategies influence export performance among SMEs from emerging countries.

Based on the two gaps above referred, this paper seeks to add theoretical value by analyzing the influence of both government institutional support and active/reactive internationalization strategies in the relationship between innovation capabilities and export performance, in emerging economies, particularly in Mozambique. For this reason, we raise the following research questions: How do ICs affect export performance in SMEs? What is the moderating effect of institutional factors (relationship with government) on the relationship between ICs and export performance? What is the moderating effect of SMEs' proactive and reactive internationalization strategies on the relationship between ICs and export performance?

The paper is divided into six sections. After this introduction, Section 2 reviews the relevant literature including hypothesis development, in which we examine the relationships between ICs and export performance, and the moderating effects of relationship with government and proactive/reactive behavior of SMEs. The research methodology and model are discussed in Section 3. Section 4 reveals the discussion of the most significant results, while Section 5 details the main findings. Section 6 presents the limitations and future lines of research.

2. Literature Review and Formulation of Hypotheses

2.1. Innovation Capabilities

There is a variety of perspectives regarding innovation capabilities (Olsson et al. 2010), with multifaceted and inconsistent constructs (Lawson and Samson 2001; Guan and Ma 2003; Oura et al. 2016; Perdomo-Ortiz et al. 2006; Ribau et al. 2017a) based on the context and the characteristics of the companies, with implications for the methodologies used, as it is necessary to adapt the scales to the methodology (Yi et al. 2013; Vicente et al. 2015;

Ahmad and Lee 2016). From this perspective, a diversity of concepts about innovation capabilities was also found (see Table 1). For example, Lawson and Samson (2001) define ICs as the ability to continuously transform knowledge and innovative ideas into new products, production processes, and systems for the benefit of the firm and stakeholders. In the view of Guan and Ma (2003), ICs are firms' assets related to internal and acquired experiences. Akman and Yilmaz (2008) define innovation capabilities as organizational culture, promotional activities, and abilities to perceive and cope appropriately with the external environment. Hogan et al. (2011) and Saunila (2016) build on Lawson and Samson's (2001) concept of mainstream and newstream capabilities. However, we will consider the scales presented by Guan and Ma (2003) and use a scale suitable for Mozambican small and SMEs, which has been previously tested in China, Brazil, and Portugal (Guan and Ma 2003; Oura et al. 2016; Ribau et al. 2017a).

Table 1. Definitions and dimensions of innovation capability.

Definition	Dimension	Author
It is the 'ability to continuously transform knowledge and ideas into new products, processes and systems for the benefit of the firm and its stakeholders.'	Vision and strategy; harnessing the competence base; organizational intelligence; creativity and idea management; organizational structure & systems; culture and climate; and management of technology.	Lawson and Samson (2001)
It is a special asset of a firm. It is tacit and non-modifiable, and it is correlated closely with interior experiences and experimental acquirement.	Learning capability; R&D capability; manufacturing capability; marketing capability; organizational capability; resource exploiting; and strategic capability.	Guan and Ma (2003)
Innovative capability is defined as a crucial element that facilitates the organizational culture, the distinctiveness of in-house promotional activities and the ability to understand and respond appropriately to the external environment.	Organizational culture; knowledge from different resources for product development activities; reflect changes in market conditions on products and processes; support of product and process innovation development; new ideas that come from customers, suppliers, to include into product development activities; and adaptation to environmental changes.	Akman and Yilmaz (2008)
A firm's ability, relative to its competitors, to apply the collective knowledge, skills, and resources to innovation activities related to new products, processes, services, or management, marketing or work organization systems, in order to create added value for the firm or its stakeholders.	Client-focused innovation capability; marketing-focused innovation capability; and technology-focused innovation capability.	Hogan et al. (2011)
Innovation capability is defined in this study as an internal capability aiming to describe the determinants affecting an organization's ability to achieve innovations continuously and add value for the organization and its stakeholders.	Participatory leadership culture; ideation and organizing structures; work climate and well-being; know-how development; regeneration; external knowledge; and individual activities.	Saunila (2016)

2.2. Export Performance

In the international literature, export behavior is described as the result of numerous variables (Bonaccorsi 1992). The measurement of export performance suffers from some conceptual, methodological, and practical limitations. Despite the large number of different measures of export performance, few have been used frequently, such as export intensity, export sales growth, export profitability, export market share and overall satisfaction, export performance, and export success (Sousa 2004), mainly as a result of the difficulty of obtaining data and firm secrecy.

Export performance was measured as the share of export sales over total sales (Yi et al. 2013), and has also been measured using financial (e.g., export sales and profits) and non-financial indicators, which include some strategy-based items (e.g., firms' export goals, satisfaction and perceived success) (Zou and Stan 1998; Ribau et al. 2017a).

Meanwhile, from a broader perspective, export performance can be divided into structural factors (size, age, management systems, technology, and R&D), firm management factors (export expectation, profitability, risk, costs, and experience), and incentives and obstacles in the internationalization process (Guan and Ma 2003). The scale used in this research is the result of an adaptation of several authors, namely Jantunen et al. (2005); Kuivalainen et al. (2007); Aulakh et al. (2000); Zou and Stan (1998).

2.3. Government Institutional Support

Government institutional support reflects the extent to which government institutions provide support to firms in order to reduce adverse market effects (Shu et al. 2019; Xin and Pearce 1996). This relationship is a key element of the institutional environment and shapes the relationship between innovation capabilities and export performance (Yi et al. 2013).

Yi et al. (2013) found that the relationship with the government has a positive and significant moderating effect on the relationship between ICs and export performance only in regions where the level of marketization is high. Conversely, Tian et al. (2019) concluded that firm–government relationships have a significant positive impact on firm innovation. Therefore, it is expected that the stronger the government–firm relationship, the greater the firm's innovation output, especially since public policy, often under the aegis of internationalization support programs, helps firms leverage their own resources.

According to Li and Atuahene-Gima (2001), government institutional support plays a significant role in increasing the effectiveness of firms' product innovation strategy. For Szeto and Kim (2018), government–firm relations can help firms access resources and improve their performance. Peng and Heath (1996) argued that government relations play a greater role in facilitating new initiatives, including exporting. Clearly, from a resource perspective, government institutional support is expected to help address the need for resources that many SMEs have (Mota et al. 2021).

The government–enterprise relationship ceases to make sense in regions where governments are corrupt (Qian 1996; Yi et al. 2013). This is what happens in developing and underdeveloped countries, which ultimately undermines the precious help that some SMEs need to bridge their internal resource needs, undermining the innovative development and competitive advantages of SMEs. Signs of good government functioning include lack of intervention, lower levels of regulation and bureaucracy, successful provision of public goods and services, and efficient spending (Porta et al. 1999). These attributes can provide services, resources, and other factors that help firms upgrade innovation capabilities and export their products to foreign markets (Yi et al. 2013).

2.4. Active/Reactive Internationalization Behavior

The internationalization of SMEs is a complex process that requires firms' active and reactive involvement and commitment (Ribau et al. 2017b). Active/reactive behaviors are related to endogenous or exogenous factors that affect firms' internationalization processes and their export performance (Bruyat and Julien 2001). Moreover, active stimuli can result from aggressive behavior by SMEs and a deliberate search for market opportunities abroad, the origin of these stimuli being the external environment (i.e., external proactive stimuli). Reactive motives may arise from within SMEs, but reflect involvement in international business as a reaction to certain internal conditions or events (i.e., reactive-internal stimuli). Alternatively, export motives may be the result of incidental circumstances or a response to environmental pressure (i.e., reactive-external stimuli) (Westhead et al. 2004).

More active firms tend to internationalize more quickly; in contrast, traditional firms tend to take a more ad hoc, reactive, and opportunistic approach to internationalization (Bell et al. 2003). On the other hand, SMEs with greater resources are more likely to actively pursue market opportunities. SMEs that benefit from the munificence of local resources may therefore be able to proactively seek customers in foreign markets (Westhead et al. 2004).

Mediation effects between entrepreneurial orientation and export performance suggest that active firms are not only better at innovating, but also their entrepreneurial orientation capabilities sustain better performance in international markets when compared to firms that react to external stimuli (Ribau et al. 2017a). Likewise, Ribau et al. (2017a) confirm that innovative skills are not as powerful in reactive SMEs as in active SMEs. Reactive SMEs neither generate nor depend on innovation to compete in international markets, while active firms implement their ICs to successfully compete and sustain activities in international markets; reactive firms not only lack these innovation capabilities, but investment in these ICs may divert their scarce resources to riskier activities.

2.5. Development of Hypotheses

There is evidence that ICs positively influence export performance. For example, Guan and Ma (2003) consider the role of seven dimensions (learning, manufacturing, R&D, marketing, organizational, resources exploitation, and strategic capabilities) and three firm characteristics (domestic market share, firm size, and productivity growth rate) in determining the performance of 213 Chinese manufacturing firms. The results indicate that export growth is closely related to total improvement in the dimensions of ICs (except for manufacturing capability) and productivity growth. Conversely, there was no evidence that export performance depends on firm size or domestic market share. Moreover, core innovation skills (a set of R&D, manufacturing, and marketing capabilities) do not lead to sustainable export growth. On the contrary, supplementary ICs (learning, organizational, resources exploitation, and strategic capabilities) allow not only the integration of all capabilities, core and supplementary, but also enable a firm to gain sustainable international competitiveness.

Ribau et al. (2017a) assessed the impact of internal ICs on the export performance of 147 Portuguese SMEs in the plastics industry, with the mediating role of entrepreneurial orientation (EO) based on firms' proactive or reactive behavior in the face of external stimuli. The results show that ICs have a positive impact on export performance. However, the mediation effects of EO suggest that proactive firms not only are better innovators, but also their EO competencies sustain better performance in international markets when compared to firms that react to external stimuli.

Oura et al. (2016) investigated the impact of innovation capacities and international experience on the export performance of Brazilian manufacturing SMEs. Conversely, the research indicated that international experience has a greater impact on export performance than innovation capacities. On the other hand, Vicente et al. (2015) identified important dimensions to build a scale to measure ICs in exporting firms. The results reveal that a construct formed by four dimensions (product development capability, innovativeness, strategic capability and technological capability) positively affects export performance. Thus, we present the following hypothesis:

Hypothesis 1 (H1). *Innovation capabilities have a direct positive effect on the export performance of SMEs.*

According to Yi et al. (2013), the government relationship has a positive moderating effect on the relationship between innovative capabilities and export performance. Government institutional support can offset the negative effects of market imperfections by reducing transaction costs and enhancing the role of innovative skills in export performance. Government institutional support functions as an important formal regulatory mechanism that remedies the adverse effects of institutional voids and helps organize and direct effective business operations (Stephan et al. 2015). Export support programs are a clear example of how public policies can support firms in their competitive development (Mota et al. 2021; Malca et al. 2020). We thus argue that the government relationship moderates the relationship between ICs and export performance. As such, it is possible to defend the following hypothesis:

Hypothesis 2 (H2). *Government institutional support positively moderates the relationship between innovation capabilities and the export performance of SMEs.*

SMEs' with proactive internationalization behavior, besides being better at managing their innovation processes, have entrepreneurial orientation capabilities that support better performance in international markets when compared to firms that react to external stimuli (Ribau et al. 2017a). On the other hand, proactive stimuli may result from SMEs' aggressive behavior and a deliberate search for market opportunities abroad (Westhead et al. 2004). Therefore, we argue that SMEs' active internationalization behavior positively moderates the relationship between innovation capabilities and export performance. Thus, it is possible to defend the following hypothesis:

Hypothesis 3a (H3a). *Proactive internationalization behavior positively moderates the relationship between SMEs' innovation capabilities and export performance.*

Reactive internationalization behaviors result from endogenous or exogenous factors that affect the firm's internationalization processes and export performance (Bruyat and Julien 2001; Westhead et al. 2004). Reactive internationalization behaviors can arise from within SMEs, but reflect involvement in international business as a reaction to certain external conditions or internal events (Westhead et al. 2004). Thus, we propose the following hypothesis:

Hypothesis 3b (H3b). *Reactive internationalization behavior positively moderates the relationship between ICs and export performance of SMEs.*

The proposed conceptual model is shown in Figure 1.

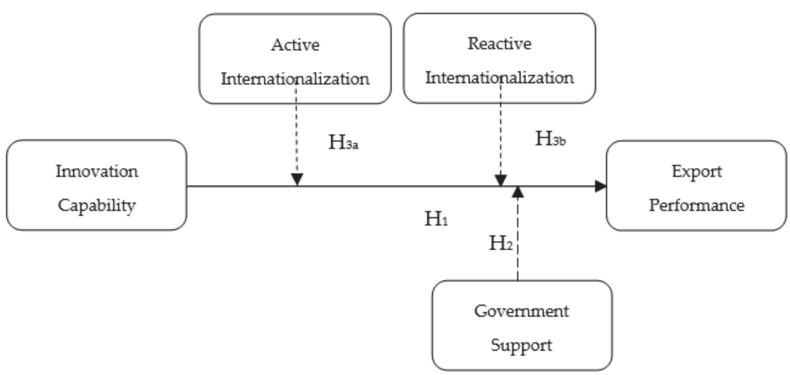

Figure 1. Proposed conceptual model.

3. Methodology

This study follows a quantitative methodology underpinned by the application of a questionnaire that was created as a result of a thorough literature review of different factors involving innovation capabilities, export performance, government support, and passive/active internationalization.

Data were collected using a questionnaire composed of scales adapted, validated, and published in previous research. We implemented a seven-point Likert-type scale in which respondents have the opportunity to agree or disagree and express the intensity of agreement (1 = Strongly disagree; 7 = Strongly agree). ICs were measured with a multidimensional scale developed by Guan and Ma (2003), the export performance scale was adapted from Jantunen et al. (2005); Kuivalainen et al. (2007); Aulakh et al. (2000); Zou and Stan (1998). Relationship with government was adapted from Shu et al. (2019), and Li and Atuahene-Gima (2001), and active/reactive internationalization was adapted from Westhead et al. (2004).

The questionnaire was applied in Portuguese. As such, Brislin's (1971) recommendations were followed to avoid misunderstandings when translating the questionnaire from English into Portuguese and back to English. At the beginning of the questionnaire, respondents were shown a small introduction of what the questionnaire was about. For operational reasons, the questionnaire was divided into five different sections—one for each topic addressed: innovation capabilities, government institutional support, active/passive internationalization, and export performance—and a final one containing respondents' data.

The questionnaire was subjected to a pre-test conducted with a convenience sample of eight individuals (three university professors and five managers) in order to check the organization and formatting of the questionnaire, the correct wording, how the respondents understood the questions and the response time required, and to eliminate typos. As a result of the pre-test, some changes were made in the terminology to facilitate the respondents' understanding. In addition, the number of items per variable was reduced to a minimum to keep the questionnaire to an adequate size. The final version of the questionnaire was made available online to companies via Google Drive LimeSurvey for 4 months, finishing in March 2020 before the outbreak of COVID-19 pandemic in Mozambique. Several rounds of emails were sent to the firms, complemented with telephone calls, to increase the response rate.

This research is grounded on the information obtained with the help of the National Institute of Statistics of Mozambique and with the support of the Investment and Export Promotion Agency (APIEX) of Mozambique, which provided a database of 400 exporting SMEs. Over the course of the survey, 305 responses were obtained. However, 55 questionnaires with incomplete answers were excluded, and a set of 250 questionnaires with complete answers was obtained, constituting 62.5% of the total sample. The characteristics of the sample are presented in Table 2. The sample size is considered suitable for data analysis using partial least squares structural equation modeling (PLS-SEM) (Hair et al. 2011).

Table 2. Characterization of respondents.

	n	%
No. employees		
5–49	168	67.2
50–100	82	32.8
Sector		
Agro-industry	48	19.2
Wood processing	89	35.6
Fishing products	67	26.8
Agricultural products	46	18.4
Respondent		
Owner	163	65.2
Manager	79	31.6
Others	8	3.2

Source: Own preparation.

As this research is based on quantitative approaches, it needs to be supported with reliability and validity analysis to ensure replicability and generalizability. While reliability is concerned with the consistency of measurements, validity is related to the extent to which the study reflects the social phenomena being studied. With reliable and valid measurement, it is possible to replicate the study (Wahyuni 2012).

The statistical analysis of the data was carried out using partial least squares structural equation models (PLS-SEM), using SmartPLS 3.2. This methodology was justified because its results are robust and because PLS-SEM supports linear regression equations that explain both linear and moderation effects when researchers seek to test and validate exploratory models (Henseler and Chin 2010).

4. Results

The evaluation of the models was based on reliability, convergent and discriminant validity. Tables 3–6 present the factor loadings of the items, which were obtained through bootstrapping with 5000 interactions, the average variance explained (AVE), and composite reliability (CR) for the different constructs under analysis. All items have loadings equal to or greater than the recommended minimum threshold of 0.7 (Götz et al. 2010) and items below this were removed.

Table 3. Loadings, AVE, CR and Cronbach's alpha of export performance.

Questionnaire Item	Loading	AVE	CR	Cronbach Alpha
Exporting has contributed to the sales growth of our firm	0.871	0.777	0.961	0.952
Exporting has improved our firm's market share	0.877			
Our export activity has made our firm more competitive	0.955			
Exporting has contributed to our Profitability	0.833			
Exporting has contributed to enter in new markets	0.831			
Exporting has contributed to improve international image	0.923			
Exporting improves the development of our know-how	0.874			

Source: Own preparation. Scale adapted from Jantunen et al. (2005); Kuivalainen et al. (2007); Aulakh et al. (2000); Zou and Stan (1998).

Table 4. Loadings, AVE, CR and Cronbach's alpha of government institutional support.

Questionnaire Item	Loading	AVE	CR	Cronbach Alpha
Government provides technology information and support	0.826	0.716	0.883	0.810
Government provides support to seek for financial resources	0.784			
Government provides with direct tax reduction and subsidy	0.923			

Source: Own preparation. Scale adapted from Shu et al. (2019); Li and Atuahene-Gima (2001).

Discriminant validity is shown in Table 7 using the Fornell–Lacker criterion. It is clear that the square root of AVE is larger than the correlation values of the two variables under analysis (Hair et al. 2011).

In order to test the four hypotheses put forward, seven different modes were tested, as shown in Table 8, considering export performance as the dependent variable. Model 1 tests the direct effect of innovation capabilities on export performance. Models 2, 4, and 6 test government institutional support, reactive internationalization behavior, and active internationalization behavior as antecedents of export performance, respectively. Models 3, 5, and 7 test the moderating effects of government institutional support, reactive internationalization behavior and active internationalization behavior on the relationship between innovation capabilities and export performance.

Table 5. Loadings, AVE, CR, and Cronbach's alpha of first order innovation capability construct.

Questionnaire Item	Loading	AVE	CR	Cronbach Alpha
Learning capability				
Monitoring technology development trends	0.717	0.679	0.913	0.880
Assimilating and absorbing ability	0.714			
Re-innovation ability facing international market	0.874			
Learning from past experiences and failings	0.910			
Cultivating and investing on learning consciousness	0.884			
Manufacturing capability				
Technological level of manufacturing equipment	0.821	0.747	0.936	0.915
Advanced manufacturing technology	0.864			
Equipment operating skill of personnel	0.918			
Production regulations and system	0.875			
Total quality management	0.841			
Marketing capability				
Understanding subdivided market	0.866	0.737	0.918	0.881
Monitoring the situation of market	0.853			
Controlling and managing distribution network	0.896			
Improving brand name and firm repute	0.816			
Organizational capability				
Adjusting organization structure to innovation projects	0.872	0.720	0.911	0.870
Centralizing resources on innovation activity quickly	0.858			
Adapting and responding to external environment	0.818			
Information flow and interconnection between departments	0.843			
R&D capability				
Building organization to collect various innovation ideas	0.691	0.702	0.921	0.894
Cross-functional project teamwork	0.895			
Facilitating communication among R&D personal	0.867			
Communication between R&D and marketing department	0.860			
Harmonizing product and process innovation	0.859			
Resource exploitation capability				
Attaching importance to human resources	0.924	0.740	0.895	0.822
Selecting key personnel in each functional department	0.861			
Making fully use of external technologies	0.790			
Strategic capability				
Understanding technological goals of top management	0.864	0.805	0.925	0.879
Entrepreneur spirit and intense innovation environnement	0.950			
Knowing industry's technological development trend	0.875			

Source: Own preparation. Scale adapted from Guan and Ma (2003); Ribau et al. (2017b).

Table 6. Loadings, AVE, CR, and Cronbach's alpha of first order innovation capability construct.

Questionnaire Item	Loading	AVE	CR	Cronbach Alpha
Reactive				
We follow or meet the actions of our competitors	0.861	0.650	0.847	0.734
To offset seasonal sales and reduce financial risks	0.805			
Declining profits in domestic market	0.748			
Active				
Part of the intrinsic growth objective of the firm	0.848	0.562	0.863	0.802
Export markets actively targeted by owner/manager	0.839			
Exporting seen as the easiest way to grow	0.778			
Excess capacity 'pushed' the business into exporting	0.674			
Public agencies with contacts with overseas clients	0.572			

Source: Own preparation. Scale adapted from Westhead et al. (2004).

Table 7. Discriminant validity.

Variables	Correlations										
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Export performance	**0.882**										
2. Government relationship	0.306	**0.846**									
3. Learning capability	0.514	0.198	**0.824**								
4. Manufacturing capability	0.373	0.192	0.802	**0.864**							
5. Marketing capability	0.417	0.098	0.653	0.793	**0.859**						
6. Organisational capability	0.572	0.260	0.720	0.809	0.823	**0.848**					
7. Proactive behavior	0.718	0.246	0.479	0.325	0.458	0.533	**0.749**				
8. R&D capability	0.277	0.122	0.733	0.576	0.519	0.508	0.307	**0.838**			
9. Reactive beahavior	0.603	0.078	0.554	0.410	0.435	0.504	0.543	0.309	**0.806**		
10. Resources exploitation capability	0.415	0.071	0.758	0.675	0.585	0.618	0.390	0.696	0.437	**0.860**	
11. Strategic capability	0.255	−0.062	0.534	0.366	0.308	0.405	0.256	0.640	0.247	0.416	**0.897**

Note: The values of the diagonal (in bold) are the square root of AVE.

Table 8. Summary of the regression analyses.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Innovation capability	0.501 (0.000)	0.458 (0.000)	0.450 (0.000)	0.252 (0.000)	0.182 (0.004)	0.195 (0.000)	0.146 (0.005)
Government relationship		0.225 (0.000)	0.229 (0.000)				
Innovation capability × Government relationship			−0.042 (0.464)				
Reactive behavior				0.421 (0.000)	0.389 (0.000)		
Innovation capability × Reactive behavior					−0.228 (0.000)		
Active behavior						0.626 (0.000)	0.599 (0.000)
Innovation capability × Active behavior							−0.129 (0.001)
R^2	0.251	0.296	0.298	0.482	0.482	0.547	0.569

Source: Own preparation. Dependent variable: Export performance.

The results presented in Table 8 for model 1 indicate that the correlation coefficient is high (β = 0.501) and explains 25.1% of the export performance in model 1, i.e., ($R^2 = 0.251$), thus confirming previous literature (Guan and Ma 2003; Vicente et al. 2015; Oura et al. 2016; Ribau et al. 2017a) and hypothesis 1.

Model 2 in Table 8 measures the effect of government institutional support as an antecedent of export performance. The results indicate that the explanation of model 2 rises from 25.1% ($R^2 = 0.251$) to 29.6% ($R^2 = 0.296$), i.e., government institutional support plays an important role (β = 0.225) as an antecedent of export performance (Peng and Heath 1996; Szeto and Kim 2018).

Model 3 presents the mediating relationship of government institutional support in the relationship between innovative capabilities and export performance. A stagnation of the indicators can be noted with regard to the explanation of the models as R^2 presents a marginal increase from 29.6% to 29.8%, as presented in Table 8. Although Yi et al. (2013) and Stephan et al. (2015) advocate the importance of the role of government institutional support, the result of this research confirms that the moderating effect is almost null (β = −0.042) and is not statistically significant. Therefore, contrary to that expected, the moderating effect of SMEs' relationship with the government is non-existent. According to Yi et al. (2013), this effect is only positive and significant in contexts where the level of marketing activities among firms is high. In Mozambique, the marketing dimension is considered to be incipient, despite several firms already defining their target customers and adjusting products/services in response to the market (Ministério da Indústria e Comércio 2016). Meanwhile, it is confirmed that Mozambican SMEs are supported by the government. This is done either through the Institute for the Promotion of Small and Medium Enterprises (IPEME) through Decree no. 47/2008, of 3 December, as the public entity that has the responsibility not only to ensure the implementation of the Strategy for the Development, promotion and dynamization of Micro, Small and Enterprises (MSMEs), or by the Agency for the Promotion of Investment and Exports (APIEX), created through Decree no. 60/2016, of 12 December, the objective of which is to promote and facilitate private, public investment and exports, in accordance with the objectives and goals of the government's economic policy. However, this effort is not significant as SMEs still face problems such as regulatory barriers, lack of financing, high tax burdens and costs of procedures, and poor access to international markets.

On the other hand, this relationship ceases to make sense in regions where governments are corrupt (Qian 1996; Yi et al. 2013). Corruption in developing countries is a hindrance to the growth of SMEs. In Mozambique, in particular, SMEs are the most confronted with bribes and other corrupt practices because they are less equipped to defend themselves and or turn to politicians (Ministério da Indústria e Comércio 2016). Clearly, if the relationship with the government is far from ideal, which can be justified by the lack of resources of Mozambican SMEs, as well as the inadequacy of support for SMEs that have difficulty competing in international markets, then government institutional support is lost.

The result of the moderating effect of the mediation of government institutional support tested in model 3 is presented in Figure 2 where it can be seen that as the innovation capability increases the marginal increase in export performance of firms with lower and higher relationship with the government is practically null.

Models 4 and 5 present the relationship between reactive internationalization behavior, export performance, and the mediating effect of reactive internationalization behavior on the relationship between ICs and export performance, respectively.

The results of model 4 indicate that innovation capabilities and reactive internationalization behavior explain 48.2% ($R^2 = 0.482$) of export performance, and there is a positive and statistically significant relationship between reactive behavior and export performance (β = 0.421), confirming what is postulated in the literature (Bruyat and Julien 2001; Westhead et al. 2004; Ribau et al. 2017a).

Model 5 shows the moderating effect of reactive behavior between innovation skills and export performance, where it is found that the R^2 did not increase relative to model 4

(R^2 = 0.482) and that the moderating effect is negative and statistically significant ($\beta = -0.228$). Thus, as shown in Figure 3, although firms with weak innovation capacity increase their export performance as they react to requests from international markets, firms that are truly innovative end up not benefiting from their export potential, confirming the results presented by Ribau et al. (2017a), as reactive SMEs do not rely on innovation to compete in international markets.

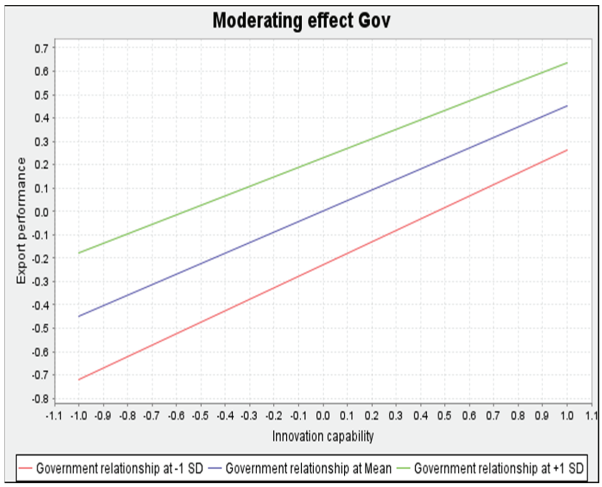

Figure 2. Graphic representation of the moderating effect of government institutional support on the relationship between innovation capabilities and export performance.

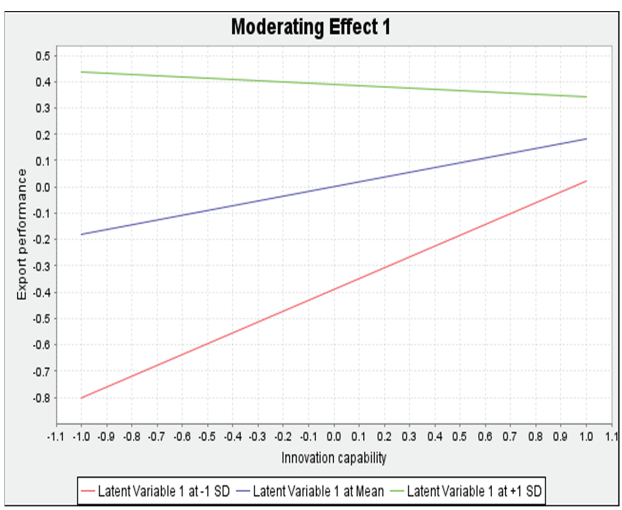

Figure 3. Graphic representation of the moderating effect of reactive internationalization strategy behavior on the relationship between innovation capabilities and export performance.

The results of model 6 indicate that the relationship between active internationalization behavior and export performance is positive and statistically significant ($\beta = 0.547$), so the higher the proactive internationalization behavior, the higher the export performance.

Therefore, the previous literature is confirmed (Bruyat and Julien 2001; Westhead et al. 2004; Ribau et al. 2017a).

The moderating effect of proactive behavior does not confirm the hypothesis 3b raised, since, although it is statistically significant, it has a negative effect on moderating the relationship between ICs and export performance (β = −0.129). Thus, as the innovation skills of Mozambican SMEs increase, as well as their proactive behavior, their export performance also increases. However, as shown in Figure 4, while firms with weak innovation skills increase their export performance by using active internationalization strategies, firms with higher innovation skills do not benefit from this active internationalization strategy.

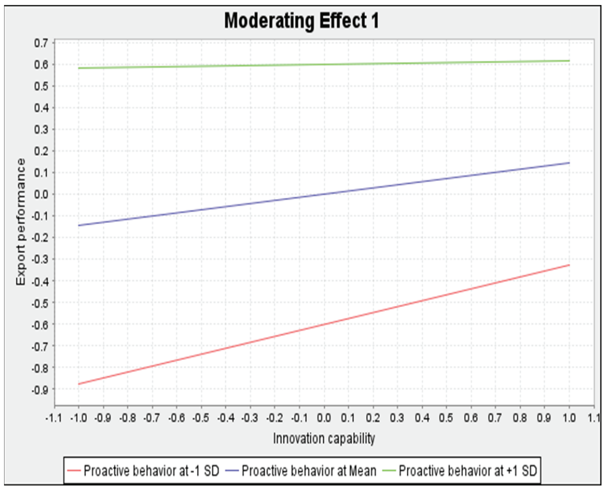

Figure 4. Graphic representation of the moderating effect of active internationalization strategy behavior on the relationship between innovation capabilities and export performance.

Table 8 presents the results of this research. The first hypothesis (H1) confirms that the direct effect between ICs and export performance is statistically significant (β = 0.501; $p < 0.000$). Thus, H1 is validated. Although the relationship between government institutional support and export performance is statistically significant, the moderating effect of government institutional support on the relationship between ICs and export performance was not shown to be statistically significant (β = −0.042; $p = 0.464$). Thus, the second hypothesis H2 cannot be validated. The third hypothesis (H3a) confirms that the moderating effect of reactive internationalization behavior between ICs and export performance is statistically significant, however, the sign is negative (β = −0.228; $p = 0.000$), thus rejecting hypothesis H3a. Finally, the fourth hypothesis (H3b) confirms that the moderating effect of proactive behavior between ICs and export performance is statistically significant, however, the sign is also negative (β = −0.129; $p = 0.001$), thus rejecting hypothesis H3b.

5. Discussion

The results of model 1 confirm the importance of ICs to Mozambican SMEs confirming previous studies (Guan and Ma 2003; Oura et al. 2016; Ribau et al. 2017a; Vicente et al. 2015). Thus, it is possible to state that exporting firms from emerging countries depend on their ICs to leverage their export performance. Another important aspect is that the moderating effect of government relationship, which is highly relevant in several contexts (Comi and Resmini 2020; Malca et al. 2020; Mota et al. 2021; Stephan et al. 2015; Yi et al. 2013), is not statistically significant, being close to zero (β = −0.042) in the Mozambican case. Thus, as presented in Figure 2, although export performance improves as innovation skills improve, the marginal difference in export performance is slightly lower as innovation skills improve for higher levels of government relationship. As such, the effect of the government relationship on the relationship between innovation skills and export performance remains virtually unchanged for increasing levels of government relationship.

As seen in Figures 3 and 4 and models 5 and 7 in Table 8, the moderating effect of active and reactive internationalization behaviors of Mozambican firms on the relationship between ICs and export performance is negative and statistically significant. Thus, it can be stated that as innovation skills increase, export performance increases marginally, though much more for firms with relatively modest active strategies than for firms with high levels of active internationalization. This may indicate that Mozambican exporting firms that are more active in international markets face problems penetrating international markets and do not improve their performance as ICs increase. Similarly, with passive internationalization, it is found that with increasing innovation skills, the performance of Mozambican firms decreases with higher levels of reactive exporting strategies, although export performance increases with increasing innovation skills for low levels of reactive exporting behavior. This clearly indicates that Mozambican SMEs may not be benefiting from their innovation skills when trying to push their products into international markets with reactive strategies. Likewise, it can be seen that Mozambican SMEs have difficulty in implementing active internationalization strategies, which may indicate that international competitiveness is not properly assured, despite investment and effort put into improving their innovation capabilities.

One issue is clear: Mozambican SMEs face clear resource constraints to competing in larger international markets, especially because investment in innovation is much more risky and is clearly hampered by the lack of government aid.

6. Conclusions

This research analyzes the moderating effect of government support and active/reactive internationalization of Mozambican SMEs in the relation between ICs and export performance. The target population of the study, 250 Mozambican SMEs, contributed to the literature on SME internationalization in the context of an emerging country. This study was undertaken on the premise that in the context where SMEs carry out their activities, the support that the government provides and active/reactive behavior can influence the relationship between ICs and export performance. Its originality stems from the fact that it deals with how ICs impact export performance in emerging countries, and the effect of government support and active/reactive internationalization strategies on the relationship between ICs and export performance. Both aspects are novel among emerging countries, namely in Mozambique.

Although most Mozambican SMEs, as in other emerging economies, particularly in sub-Saharan Africa, have structural problems, low productivity levels, little modern technology and, as a consequence, low product/service quality, Mozambican exporting SMEs have shown that the use of ICs leads to good export performance.

Paradoxically, government support to SMEs is nonexistent, as is the moderating effect of the relationship of SMEs with the government, i.e., as this effort is not significant, SMEs continue to face problems such as regulatory barriers, lack of funding, high tax burden and cost of procedures, and poor access to international markets. As such, Mozambican SMEs, despite improving their ICs, are not able to leverage their innovation improvements with the government programs available to them. This ends up hindering the normal growth of SMEs, especially in emerging countries whose domestic resources are very limited.

The lack of resources is a general characteristic of SMEs. In the case of emerging economies, this lack of resources is even more prevalent. Although Mozambican SMEs invest in their innovation skills, these are not having the desired effect on export performance because, on the one hand, the lack of resources often makes active internationalization strategies impossible and, on the other hand, government support is non-existent, which does not benefit the internationalization effort undertaken by many companies. It is not enough just to develop internal ICs, it is also necessary to operationalize proactive strategies so that SMEs can face international competitiveness, which requires clear and sustainable public policies and state support, able to make up for the companies' lack of pressing resources. However, the priority of governments in most emerging countries is oriented towards solving immediate socio-economic problems such as the percentage of the population with low survival rates and high corruption rates. In this context, SMEs aspiring to international business projects find themselves alone, relying on their own resources and capabilities, strangled by an unenthusiastic domestic market.

The proactive or reactive international strategy of Mozambican SMEs only benefits less innovative SMEs. In this context, it can be stated that Mozambican SMEs with high levels of innovation have some difficulty in competing internationally, not because of their lack of ICs, but because of the lack of resources, and perhaps because riskier investments are needed to compete within broader international horizons. This is possible only with effective public governmental business-support policies from the governments of emerging countries.

Clearly, at the government level, the challenge is clear: without governmental support, international business performance is compromised, especially because SMEs in emerging countries need to overcome the liability of newness that they face in international markets by increasing business risk, which ends up hindering active innovation development to implement internationalization strategies. On the other hand, the need for financial resources to adapt the firm's product portfolio may compromise the success in international competitive markets and relegate Mozambican and emerging countries' SMEs to the much less demanding domestic market, without solving the lack of competitiveness at the international level. Thus, the great challenge that many emerging economies have to overcome is to create clear government support for companies in order to support them to get to know and compete in international markets through support that increases their innovative capabilities.

At the business level, it is recommended that Mozambican SMEs, in particular, and those in emerging countries, in general, continue to invest in their ICs so that they can improve their internal processes and the development of new products and manufacturing skills, which will support them in marketing and implementing competitive strategies in wider markets. The development of ICs appears to be core to improving international competitiveness. They should also be more demanding of governmental institutions, demanding clear policies to support business competitiveness.

The main limitation of this study is that its content is based on the responses of 250 companies from a single emerging country, Mozambique, which could be complemented by samples from other emerging countries that could give a broader perspective. As it is a cross-sectional study, the intrinsic characteristics of a longitudinal research were not considered. The fact that only one informant per company was considered, and that it was not possible to compare several industrial sectors, may also be considered limitations. However, in the Mozambican context, this would be very difficult to implement. In addition to the limitations presented, future research should take into account the analysis of the context of the internationalization of Mozambican companies, especially their modes of entry and their main international competitors, as well as the qualifications of Mozambican managers and their degree of knowledge of the international context.

Author Contributions: Conceptualization, A.M., E.N. and C.R.; methodology, A.M., E.N. and C.R.; validation, A.M., E.N. and C.R.; formal analysis, A.M., E.N. and C.R.; investigation, E.N.; data curation, A.M. and E.N.; writing—original draft preparation, E.N.; writing—review and editing, A.M., E.N. and C.R.; supervision, A.M. and C.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

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

References

Ahmad, Abdi Yuya, and Keun Lee. 2016. Embodied technology transfer and learning by exporting in the Ethiopian manufacturing sector. *Innovation and Development* 6: 281–303. [CrossRef]

Akman, Gülşen, and Cengiz Yilmaz. 2008. Innovative capability, innovation strategy and market orientation: An empirical analysis in Turkish software industry. *International Journal of Innovation Management* 12: 69–111. [CrossRef]

Aulakh, Preet, Masaaki Kotabe, and Hildy Teegen. 2000. Export strategies and performance of firms from emerging economies: Evidence from Brazil, Chile, and Mexico. *Academy of Management Journal* 43: 342–61. [CrossRef]

Ayob, Abu, and Joan Freixanet. 2014. Insights into public export promotion programmes in an emerging economy: The case of Malaysian SMEs. *Evaluation and Programme Planning* 46: 38–46. [CrossRef] [PubMed]

Bell, Jim, Rod McNaughton, Stephen Young, and Dave Crick. 2003. Towards an integrative model of small firm internationalisation. *Journal of International Entrepreneurship* 1: 339–62. [CrossRef]

Blanchard, Kevin. 2020. Innovation and strategy: Does it make a difference! A linear study of micro & SMEs. *International Journal of Innovation Studies* 4: 105–15. [CrossRef]

Bonaccorsi, Andrea. 1992. On the relationship between firm size and export intensity. *Journal of International Business Studies* 23: 605–35. [CrossRef]

Brislin, Richard. 1971. Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology* 1: 185–216. [CrossRef]

Bruyat, Christin, and Pierre-Andre Julien. 2001. Defining the field of research in entrepreneurship. *Journal of Business Venturing* 16: 165–80. [CrossRef]

Comi, Simona, and Laura Resmini. 2020. Are export promotion programmes effective in promoting the internalization of SMEs? *Economia Política* 37: 547–81. [CrossRef]

Francis, June, and Colleen Collins-Dodd. 2004. Impact of export promotion programmes on firm competencies, strategies and performance: The case of Canadian high-technology SMEs. *International Marketing Review* 21: 474–95. [CrossRef]

Freixanet, Joan. 2012. Export promotion programmes: Their impact on companies' internationalization performance and competitiveness. *International Business Review* 21: 1065–86. [CrossRef]

Götz, Oliver, Kerstin Liehr-Gobbers, and Manfred Krafft. 2010. Evaluation of Structural Equation Models Using Partial Least Square (PLS) Approach. In *Handbook of Partial Least Squares*. Edited by Vincenzo Vinzi, Wynne Chin, Jorg Henseler and Huiwen Wang. Heidelberg: Springer, pp. 691–711. [CrossRef]

Guan, Jiancheng, and Nan Ma. 2003. Innovative capability and export performance of Chinese firms. *Technovation* 23: 737–47. [CrossRef]

Hair, Joe, Christian Ringle, and Marko Sarstedt. 2011. PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice* 19: 139–52. [CrossRef]

Henseler, Jorg, and Wynne Chin. 2010. A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling: A Multidisciplinary Journal* 17: 82–109. [CrossRef]

Hogan, Suellen, Geoffrey Soutar, Janet McColl-Kennedy, and Jilian Sweeney. 2011. Reconceptualizing professional service firm innovation capability: Scale development. *Industrial Marketing Management* 40: 1264–73. [CrossRef]

Instituto Nacional de Estatística. 2017. *Empresas em Moçambique: Resultados do Segundo Censo Nacional (2014–2015)*. Maputo: INE.

Jalali, Seyed. 2012. The effect of export promotion programmes on export performance: Evidence from Iranian food manufacturers. *International Journal of Business and Globalisation* 9: 122–33. [CrossRef]

Jantunen, Ari, Kaisu Puumalainen, Sami Saarenketo, and Kalevi Kylaheiko. 2005. Entrepreneurial orientation, innovation capabilities and export performance. *Journal of International Entrepreneurship* 3: 223–43. [CrossRef]

Knight, Gary, and Daekwan Kim. 2009. International business competence and the contemporary firm. *Journal of International Business Studies* 40: 255–73. [CrossRef]

Krammer, Sorin, Roger Strange, and Addisu Lashitew. 2018. The export performance of emerging economy firms: The influence of firm capabilities and institutional environments. *International Business Review* 27: 218–30. [CrossRef]

Kuivalainen, Oli, Sanna Sundqvist, and Per Servais. 2007. Firms' degree of born-globalness, international entrepreneurial orientation and export performance. *Journal of World Business* 42: 253–67. [CrossRef]

Lawson, Benn, and Danny Samson. 2001. Developing innovation capability in organisations: A dynamic capabilities approach. *International Journal of Innovation Management* 5: 377–400. [CrossRef]

Li, Haiyang, and Kwaku Atuahene-Gima. 2001. Product innovation strategy and the performance of new technology ventures in China. *Academy of Management Journal* 44: 1123–34. [CrossRef]

LiPuma, Joseph, Scott Newbert, and Jonathan Doh. 2013. The effect of institutional quality on firm export performance in emerging economies: A contingency model of firm age and size. *Small Business Economics* 40: 817–41. [CrossRef]

Malca, Oscar, Jesus Peña-Vinces, and Francisco Acedo. 2020. Export promotion programmes as export performance catalysts for SMEs: Insights from an emerging economy. *Small Business Economics* 55: 831–51. [CrossRef]

Ministério da Indústria e Comércio. 2016. *Pequenas e Médias Empresas em Moçambique. Situação, Perspectivas e Desafios*. Maputo: DNI-MIC.

Mota, Jorge, Antonio Moreira, and Alexandra Alves. 2021. Impact of export promotion programs on export performance. *Economies* 9: 127. [CrossRef]

Olsson, Annika, Carl Wadell, Per Odenrick, and Margareta Bergendahl. 2010. An action learning method for increased innovation capability in organisations. *Action Learning: Research and Practice* 7: 167–79. [CrossRef]
Oura, Mauricio, Silvia Zilber, and Evandro Lopes. 2016. Innovation capacity, international experience and export performance of SMEs in Brazil. *International Business Review* 25: 921–32. [CrossRef]
Peng, Mike, and Peggy Heath. 1996. The growth of the firm in planned economies in transition: Institution, and strategic choice. *Academy of Management Review* 21: 492–528. [CrossRef]
Perdomo-Ortiz, Jesus, Javier González-Benito, and Jesus Galende. 2006. Total quality management as a forerunner of business innovation capability. *Technovation* 26: 1170–85. [CrossRef]
Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. 1999. The quality of goverment. *Journal of Law, Economics, & Organization* 15: 222–79. [CrossRef]
Qian, Yingyi. 1996. Enterprise reform in China: Agency problems and political control. *Economics of Transition* 4: 427–47. [CrossRef]
Ribau, Claudia, Antonio Moreira, and Mario Raposo. 2017a. SMEs innovation capabilities and export performance: An entrepreneurial orientation view. *Journal of Business Economics and Management* 18: 920–34. [CrossRef]
Ribau, Claudia, Antonio Moreira, and Mario Raposo. 2017b. Export performance and the internationalisation of SMEs. *International Journal of Entrepreneurship and Small Business* 30: 214–40. [CrossRef]
Saunila, Minna. 2016. Performance measurement approach for innovation capabilty in SMEs. *International Journal of Productivity and Performance* 65: 162–76. [CrossRef]
Sen, Falguni, and William Egelhoff. 2000. Innovative capabilities of a firm and the use of technical alliances. *IEEE Transactions on Engineering Management* 47: 174–83. [CrossRef]
Shu, Chengli, Dirk De Clercq, Yunyue Zhou, and Cuijan Liu. 2019. Government institutional support, entrepreneurial orientation, strategic renewal, and firm performance in transitional China. *International Journal of Entrepreneurial Behaviour and Research* 25: 433–56. [CrossRef]
Sousa, Carlos. 2004. Export performance measurement: An evaluation of the empirical research in the literature. *Academy of Marketing Science Review* 9: 1–23.
Stephan, Ute, Lorraine Uhlaner, and Christofer Stride. 2015. Institutions and social entrepreneurship: The role of institutional voids, institutional support, and institutional configurations. *Journal of International Business Studies* 46: 308–31. [CrossRef]
Szeto, Maggie, and Young-Choon Kim. 2018. Costs and benefits of business-government relations: Empirical analysis of former-communist transition economies. *International Journal of Business and Society* 19: 219–32.
Teece, David, Gary Pisano, and Amy Shuen. 1997. Dynamic capabilties and strategic management. *Strategic Management Journal* 18: 509–33. [CrossRef]
Tian, Yuan, Yupei Wang, Xuemei Xie, Jie Jiao, and Hao Jiao. 2019. The impact of business-government relations on firms' innovation: Evidence from Chinese manufacturing industry. *Technological Forecasting & Social Change* 143: 1–8. [CrossRef]
Vicente, Margarida, Jose Luis Abrantes, and Mario Teixeira. 2015. Measuring innovation capabiliy in export firms: The INNOVSCALE. *International Marketing Review* 32: 29–51. [CrossRef]
Wahyuni, Dina. 2012. The research design maze: Understanding paradigms, cases, methods and methodologies. *Journal of Applied Management Accounting Research* 10: 69–80.
Wang, Catherine, and Pervaiz Ahmed. 2004. The development and validation of the organisational innovativeness construct using confirmatory factor analysis. *European Journal of Innovation Management* 7: 303–13. [CrossRef]
Westhead, Paul, Deniz Ucbasaran, and Martin Binks. 2004. Internationalization strategies selected by established rural and urban SMEs. *Journal of Small Business and Enterprise Development* 11: 8–22. [CrossRef]
Xin, Katherine, and Jone Pearce. 1996. Guanxi: Good connections as substitutes for institutional support. *Academy of Management Journal* 39: 1641–58. [CrossRef]
Yi, Jingtao, Chenqi Wang, and Mario Kafouros. 2013. The effects of innovative capabilities on exporting: Do institutional forces matter? *International Business Review* 22: 392–406. [CrossRef]
Zou, Shaoming, and Simona Stan. 1998. The determinants of export performance: A review of the empirical literature between 1987 and 1997. *International Marketing Review* 15: 333–56. [CrossRef]

economies

Review

A Systematic Literature Review of the Impact of Complexity Theory on Applied Economics

Philip Haynes [1,*] and David Alemna [2]

[1] School of Humanities and Social Science, University of Brighton, Brighton BN1 9PH, UK
[2] School of Area Studies, History, Politics and Literature, University of Portsmouth, Portsmouth PO1 3AS, UK
* Correspondence: p.haynes@brighton.ac.uk

Abstract: A systematic literature review is used to explore the relationship between complexity theory and economics. Broad search terms identify an unmanageable large number of hits. A more focused search strategy follows the PRISMA protocol and screens for Economics branded publications, and with key words for different applications of economics occurring in the abstract. This results in a distinct group of 247 publications. One hundred and twenty-two publications are excluded due to inclusion criteria or a lack of relevance. The remaining 113 are analysed for (1) use of complexity theory concepts, (2) types of methodology and methods, and (3) the applications for macro, meso, and micro issues. The publication with the greatest frequency of resulting articles is Complexity, closely followed by Ecological Economics. The highest annual citation ratio for a single article was 33.88. Complexity theory concepts included: non-linearity, system interactions, adaption, and resilience. Many developed a meso application, rather than solely focusing on macro or micro designs. Agent Based Models (ABMs) were popular, as were general systems models following the practice of the late system theorist, Donella Meadows. Applications were interdisciplinary and diverse, including world system models that linked macroeconomics to climate and sustainability, as contrast with micro and meso models trying to explain the complexity of agent-based behaviour on specific organisations or higher-level processes.

Keywords: complexity theory; economics; public policy; systematic review

Citation: Haynes, Philip, and David Alemna. 2022. A Systematic Literature Review of the Impact of Complexity Theory on Applied Economics. *Economies* 10: 192. https://doi.org/10.3390/economies10080192

Academic Editor: Sajid Anwar

Received: 30 June 2022
Accepted: 29 July 2022
Published: 8 August 2022

1. Introduction

Today's policy makers are faced with unprecedented challenges in tackling their immediate priorities of economic growth as well as how to approach other long-term issues suhch as climate change, energy security, and public health amongst others. For issues with many interdependent factors ('wicked problems'), it is difficult to determine drivers as multiple factors may produce similar or unidentical outcomes. For this reason, it is becoming more ubiquitous across the economic scholarship that understanding complexity offers a new science in which economic systems are understood as complex systems which cannot be judged using traditional linear analytical frameworks and methodologies. In light of these emerging complex policy challenges, advancements in economic conceptions have led to the development of 'non-orthodox' economic thinking with the labels heterodox and/or post-Keynesian economics (Lee and Lavoie 2012). Here, there is a growing scholarship on new ways of thinking that provide complementary, and alternative, perspectives to the equilibrium assumption of economic modelling. Within this economic paradigm lies the assumption that, amongst other things, economic systems are dynamic and oscillate between periods of stability and chaos, making them hard to predict.

Complexity theory is known to have had a growing influence on the broader social sciences in recent decades with increased citations that demonstrate this (Byrne and Callaghan 2013). Complexity theory was first developed in the physical sciences influencing the development of scientific concepts and methods for better understanding of unstable and difficult to predict systems such as meteorology (Lorenz 1963). Given the indeterminate nature of many social science phenomena, with novel behaviour and events resulting from a diversity of social interactions, many scholars soon saw the potential for complexity theory to assist in the explanation of the collective behaviours of societies and economies. Complex systems demonstrate a high level of uncertainty with low agreement between and across systems with regard to the causes of systemic pressures and the potential solutions to resolve such pressures (Bernardo and Smith 2009). This suggests an amount of irreducible uncertainty exists within the system (Sornette 2006). The uncertainty experienced within complex systems denotes non-linearity between cause and effect. This approach to thinking highlights properties that demonstrate features of complexity including, sensitivity to initial conditions and path dependency, emergence and self-organisation, feedback and feedback loops, and dynamic behaviours, as well as the interactions between these properties. Such an approach to economic modelling goes beyond the traditional orthodox approach where systems are seen to share identical patterns of behaviour, with interactions averaging each other out.

Castellani and Gerrits (2021) updated Map of the Complexity Sciences argues that Complexity Theory and Economics became increasingly linked from the 1990s onwards. The Santa Fe Institute (https://www.santafe.edu/ accessed on 20 July 2022) founded in 1984 was the first international scientific research institute dedicated to the study of complex adaptive systems. It succeeded in attracting leading scientists from across the world to consider important interdisciplinary science questions. Seminal academic leaders in this field included: Holland (1992), Kauffman (1993). Such scholars were ambitious in their desire to expand the new interdisciplinary scientific framework to cover the major social and economic challenges of the day. An economics program started in 1987, with much emphasis on the boundaries of the discipline, and the potential contribution of the interdisciplinary complexity science to economics. Fontana (2010), in a seminal historical summary of the impact of Santa Fe on economics, summarises three key impacts: dynamics, computational, and connectives. Complex dynamics is concerned with mathematical changes in economics with the developments to model chaos, sensitivity to initial conditions, and bifurcation. The focus here is on nonlinear approaches. Computational modelling primarily includes the development of agent-based modelling (ABM) allowing for a much more complex consideration of the behaviour of economic agents. Connective explanations are interested in the relational aspects of the economy such as positive and negative feedback and the operation of networks.

With this backdrop in mind, we use the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) to comprehend the impact of complexity theory on applied economics. Before commencing this, we use an initial exploration to identify the most cited and highly influential scholars who have affected the metatheoretical union of complexity theory with general economics. We then proceed to make an original and substantial contribution, through the use of PRISMA to identify where this fusion of complexity theory with economics is having the most applied influence. In particular, we identify impacts in the use of microeconomics, including in business, management and organisations, and in macroeconomics, incorporating also political and policy-based interventions. There are also meso applications that link micro and macro in innovative ways. The paper structure is as follows, the subsequent section explains the method of using both an indicative literature search and a more structured systematic approach. Next, we present our results with a discussion on the most cited relevant scholars and PRISMA findings highlighting methodological trends as well as the thematic application of specific complexity concepts across our reviewed documents. Finally, we provide some concluding remarks.

2. Research Method

To assess developments in the application of complexity theory in applied economics, we undertake a two-step approach to data collection and review. First, we use Google

Scholar as an initial search tool to explore the broad relationship between scholars and publications that link complexity theory with economics and to observe some quantitative citation evidence about the most important scholars and source material. Google Scholar is used for this indicative purpose because of its wide breadth of coverage, and relatively limited ability for the researcher to control and manipulate the search focus. Google Scholar only offers limited text search options (i.e., publication title, or text from the whole article) and uses automatic search algorithms to find what should be the most useful and relevant examples (Beel and Gipp 2009). The date of the search is 18 July 2022. The search term is: allintitle: complexity OR "complex systems" OR "complex adaptive systems" AND economics. It yields 523 references. We use this to construct an indicative summary of the major scholars who influence the use of complexity theory in economics.

Second, a focused systematic literature review is conducted to identify applied influences of complexity theory in economics research. As noted by Liberati et al. (2009), systematic reviews and meta-analyses are useful for summarising evidence in an accurate and reliable manner. The explicit use of systematic procedures to identify selected literature reduces bias, thereby providing reliable findings from which a researcher can draw conclusions and provide recommendations (Oxman and Guyatt 1993). This approach helps researchers keep up-to-date with topical developments while allowing readers to judge the quality of reporting, through the evidence-based rationale provided (Moher et al. 2016). In this study, our systematic review of relevant literature was undertaken in accordance with the reporting techniques outlined within the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol (PRISMA) (Moher et al. 2009). Initially adopted in the medical sciences, the PRISMA protocol offers a set of procedures for the collection and reporting of systematic reviews and meta-analysis. PRISMA consists of four-phases, that is (a) document identification; (b) screening, (c) eligibility and (d) inclusion (see Figure 1 below for a flow diagram). These steps are aimed at improving the reporting of systematic reviews and meta-analysis. As outlined by Liberati et al. (2009), the PRISMA guidelines require a researcher to: (i) explicitly outline the research objectives with areproducible methodology; (ii) undertake a systematic search to identify studies that meet the eligibility criteria; (iii) validate the included studies; and (iv) present a synthesis of the content, characteristics and findings of studies included.

In this systematic review, we utilised a combination of the Elsevier Scopus ('Scopus') and Web of Science (WoS) database to search for publications in selected journals—imposing some further restrictions. Scopus is Elsevier's largest citation and abstract database with peer-reviewed academic literature that cover the areas of social sciences, life sciences as well as health and physical sciences. Similar to Scopus, WoS is an academic citation and indexing database that provides access to journals covering the arts and humanities disciplines, sciences, and social science. The use of both databases offered a wide interdisciplinary coverage in the identification of specific research outputs. Both Scopus and WoS also provide filtering options for the researcher to control the search focus. To ensure a focus search scope, and to capture only relevant documents that fall within the theoretical parameters of the study, we generated the following search string to identify documents with the mention of "complexity theory" OR "complex systems" OR "complex adaptive systems" in the document keywords. From this search string, a combined total of 135,610 documents were identified across both databases. In order to filter for only relevant papers, we restricted our search to documents that included the initial search string in only the publication title, AND econ* OR complex* in the publication source title. This restriction allowed for the identification of publications in journals with a thematic focus on economics and complexity. Likewise, additional restrictions were placed to include policy OR management OR organization OR finan* in the abstract and further limiting this to include only publications in English[1]. These additional restrictions provided a sample of publications with 'real world' applications, rather than 'theoretical conceptualisations'. After removing duplicates, a subset of 242 documents were identified. The data from Scopus and WoS was then extracted as a .csv file (comma-separated values) for further screening.

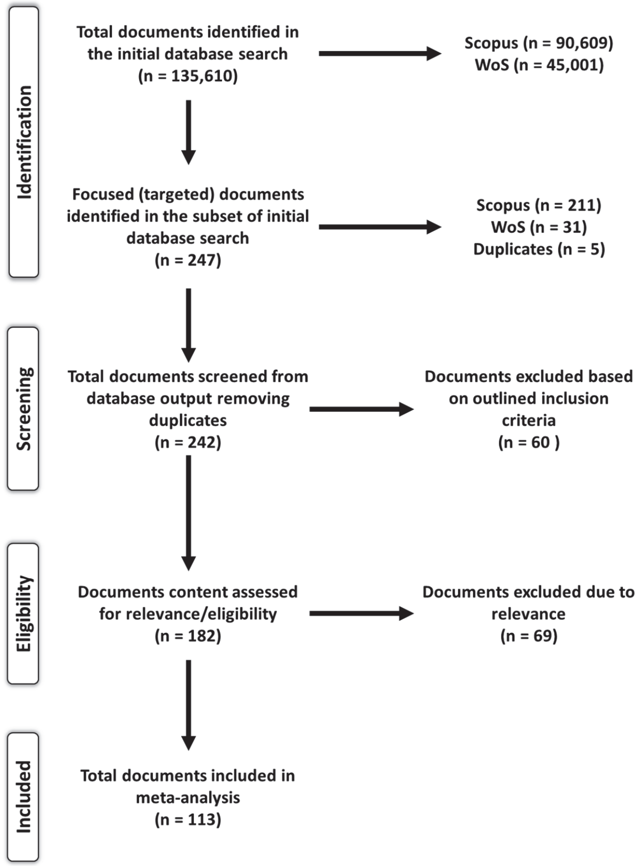

Figure 1. Literature Collection using the PRISMA Protocol.

Data screening and eligibility were two-part, first, we examined the abstracts of each document to detect and remove irrelevant literature. From this, an initial 60 documents were excluded from our analysis as the content of these publications are not openly assessible to the public domain. The second screening consisted of an examination of the full text of each document. At this stage, we developed an eligibility criterion based on the relevance of the publications (i.e., a direct focus on the application of complexity theory in the areas of applied economics, management, policy or finance). We also utilised publications impacts as both Scopus and WoS offer numeric values on the number of citations each publication has had. Lastly, the document type (i.e., Article, Review, etc.) were also consider in our criterion. From this, an additional 69 documents were removed from our analysis. Three publications were classified as Editorial, 55 documents were classified as either Conference Papers, Proceedings or Note, and the remaining 11 publications were either not contextually relevant or dated prior to 2019 with little research impact (no citations). Figure 1 (above) provides an overview of the document collection process using the PRISMA framework.

Table 1 below shows the top 15 frequently occurring publication titles for the documents included in our meta-analysis. Overall, *Complexity* has the most journal publications considered in our sample. This journal publishes studies that contribute to discussion on complex systems across a broad range of disciplines. For journals with a direct focus on economics, *Ecological Economics* features the most and is promoted as covering the situation of economics within ecology and the importance of ecological values to micro and macroeconomics. The *International Journal of Production Economics* deals with the interface of management and production, including manufacturing and engineering, and is marketed as an interdisciplinary journal. This has similarities with *Engineering Economics* and *Agricultural Economics*.

Table 1. Top 15 Journal Outputs: Frequency of Occurrence.

Journals	Total
Complexity	28
Ecological Economics	9
Ecological Complexity	7
International Journal of Production Economics	7
Construction Management and Economics	3
Engineering Economics	3
Handbook of Computational Economics	3
Journal of Economic Dynamics and Control	3
Journal of Physics: Complexity	3
Journal of Systems Science and Complexity	3
Cambridge Journal of Regions, Economy and Society	2
Complex Adaptive Systems Modeling	2
Complexity International	2
Journal of Economic Behavior and Organization	2
Journal of Economic Issues	2
Total Number of Journals included	*47*

3. Results and Discussion

3.1. Initial Exploration

One of the most cited is a book by Beinhocker (2006) entitled: The origin of wealth: evolution, complexity, and the radical remaking of economics with 2242 citations. Beinhocker is currently a professor at the University of Oxford. His book links economics with evolutionary biology and the thermodynamic laws of physics, therefore replicating some of the Santa Fe influence of seeing the natural sciences as important to economics. The core ideas of his book are summarised on page 97, Table 4.1. Economic systems are dynamic, nonlinear and far from equilibrium. Economic agents are diverse individuals with incomplete information who are subject to errors and biases. They adapt their behaviour. The economic interactions between agents can be partly understood through their changing networks of connections. Micro and macroeconomics are joined by the emergence of behaviours and interactions. The economic system evolves through differentiation, selection and amplification towards novelty and complexity. In the conclusion of his book, he makes a case for the linking of environmental issues and economics.

Arthur's (2013) Complexity Economics is relevant with 450 citations. Furthermore, his recent (2021) article in Nature Review Physics. This already has 85 hits. Arthur is documented as being one of the first economists to be substantially involved with the Santa Fe Institute and this involvement has continued for several decades (Fontana 2010). An examination of Arthur's own Google Scholar credentials reveal that he has 50,315 citations. Arthur challenges the basis of neoclassical economics. He rejects the dominant idea of an equilibrium where markets are clear to balance demand and supply. Instead, he argues that consumer and agent behaviour is diverse and evolving, leading to the emergence of new and novel aggregate outcomes. Therefore, economic interactions are not homogeneous but heterogeneous across a range of social networks. This often requires new and different mathematical approaches in economics. For policy makers, this means that they search for plausible patterns of interest that are limited in time and space rather than determined by universal laws. Policy makers face 'decision making under fundamental uncertainty' (Arthur 2021, p. 143).

In addition to an interest in the concept of 'emergence', Arthur is particularly influenced by several other concepts from complexity science such as: self-organisation (for explaining diversity within networks), power laws and long tails (for changing how policy makers and economists understand risk), and attractors (for explaining where specific empirical data patterns become important for a given time and space). With regard to methodology, Arthur focuses on consumer and agent behaviour in social networks and notes the importance of agent-based modelling as an excellent computation tool for modelling degrees of diversity in patterns of emergent behaviour in a given market context. Like almost all complexity theorists and practitioners in economics, he is committed to interdisciplinarity across the sciences and social sciences and is concerned if economics operates as a discipline in isolation from others.

Other notable substantial contributions identified in the Google Scholar search include Durlauf with two papers that are highly cited (Durlauf 2005, 2012) with 116 and 264 citations. His paper (2012) in Politics, Philosophy and Economics argues that complexity thinking adds value to contemporary economic modelling and analysis, but that it is not a theoretical paradigm shift, and he doubts the real benefits for public policy evaluation. He argues for a greater clarity about the mathematical tools that complexity theory provides for economic analysis. In the earlier paper (2005) published in The Economic Journal, he defines the empirical methods most used by complexity economists as: historical studies, power laws, and analyses of social interactions. He expresses scepticism about the extent to which the use of these methods validates the properties of complex systems.

Another scholar with substantial relevant citations is Rosser with 6660 citations on his Google Scholar author page. His most cited article (581 citations) of direct relevance is a paper in The Journal of Economic Perspectives (Rosser 1999). He argues that complexity economics have evolved from previous approaches examining cybernetic, catastrophic, and chaotic systems. Economic agents are dispersed and adapt their learning and novelty. Rationality becomes bounded. System simulation becomes an important method to understand complexity.

Rosser has also published with two other well cited authors (Holt et al. 2011). Their review of the state of the art of complexity and economics has 172 citations. Neither have Google Scholar author summary web pages, but both have other books and articles listed. For example, Colander's (2000) single authored book: Complexity and the history of economic thought, is cited 109 times.

Antonelli (2008, 2009) has two single authored papers both with substantial numbers of citations (363, in 2008, and 159, in 2009). He is professor of Economics at the University of Torino and has 12,842 citations on Google Scholar. His scholarship is specific to the economics of technological innovation. He explores and explains innovation as a path-dependent process rooted in the interdependence and interaction of a diversity of heterogeneous agents. He argues: location is important (relative to other agents), agent knowledge of others is always limited (so, none has complete knowledge), interaction is often localised, agents are creative and can deviate from given rules, but agents are also highly interdependent causing systemic phenomena.

It is important to conclude at this point that using Google Scholar in this way is exploratory and not as rigorous and focused as imposing systematic boundaries as used later in this article by applying the PRIMSA method. Nevertheless, it allows for illustration of some of the most important historical influences. The worst consequence is the exclusion of important publications that are very closely related to the topic of interest, but which use title labels that are different.

A good example, offered by one of the reviewers of this paper is when "complexity" is substituted with "evolutionary". Evolutionary economics is another subject having high impact on the discipline and with much overlap with complexity theory. A specific example is the work of Jason Potts. His Google Scholar author home page has 11,411 citations with several highly cited publications that include the keyword "evolutionary" in the title. On examination, the content overlaps with the conceptual domain of complexity economics. For example, his book: The New Evolutionary Microeconomics (2000) has 868 citations.

Similarly, peer reviewers of our article have pointed out that the eminent international scholar Doyne Farmer, Professor of Mathematics and Director of Complexity Economics at the University of Oxford, does not feature in our Google Scholar summary results, but he has highly cited articles that include the title keywords "chaos" and "chaotic". In total he has 41,215 citations and one of his most highly cited articles is relevant to complexity and economics: 'Predicting chaotic time series' (Farmer and Sidorowich 1987) published with Sidorowich. It is cited 2782 times. These two examples illustrate the limits of using Google Scholar to acquire an overview.

This first overview search with Google Scholar provides a coherent but imperfect sense of the theoretical and conceptual framework of complexity theory as applied to economics.

3.2. Systematic Literature Review

Given what was identified in the broad Google Scholar search about conceptual priorities of complexity theory for economics, our thematic analysis of the PRISMA selected articles focused first on identifying the key conceptual issues presented by each selected publication, and how these compared with each other. Next, we identify the main methodological frameworks used by each publication, placing them into groups of similarity and difference in this respect. Finally, we examined the application of the research and scholarship in the context of the traditional coverage of economics: Macro, Meso and Micro.

3.2.1. Complexity Themes

A central theme emerging across the studies included is the fact that complex systems demonstrate multiple properties (Cilliers 1998). For this reason, in our thematic analysis, we identify a central focus on specific complexity themes across the publications included. While there are overlaps, the majority of studies considered in our analysis discuss non-linearity (43), adaptation (16), system interactions (49) and resilience (5). Table 2 provides an overview of the dominant complexity properties discussed across the 113 publications.

Table 2. Emergent Complexity Properties.

Complexity Themes	Publications	Total
System Interactions	Aeeni and Saeedikiya (2019); Ahmad (2019); Albin and Foley (2001); Aouad and Bento (2019); Bianchi and Labory (2019); Brocal et al. (2019); Bruno et al. (2018); Budd et al. (2017); Chakraborti et al. (2021); Chikumbo et al. (2000); Çıdık and Phillips (2021); Coyne et al. (2021); Dong and Fisher (2019); Evans et al. (2017); Forbes and Xie (2018); Garmendia and Stagl (2010); Georgiev et al. (2015); Gimzauskiene and Kloviene (2010, 2011); Guo et al. (2021); Hartwell (2019); Jemmali (2022); Korotkikh and Korotkikh (2009); Kopp and Salecker (2020); Lamghari Elidrissi et al. (2020); Liu et al. (2021); Markose (2005); Marle (2020); Matesanz Gomez et al. (2017); Millhiser and Solow (2007); Mylek and Schirmer (2020); Naderpajouh and Hastak (2014); Oughton et al. (2018); Patrucco (2011); Phillips (2019); Qiu-Xiang et al. (2018); Ryan et al. (2007); Stuart et al. (2022); Tang and Gao (2014); Vallance (2016); Varga et al. (2016); Watson et al. (2011); Watson et al. (2011); Wheeler (2007); Wink et al. (2017); Xepapadeas (2010); Zhang et al. (2021b); Zheng and Chen (2012); Zhu et al. (2017);	49
Non-linearity	Aldhyani et al. (2021); Aymanns et al. (2018); Balint et al. (2017); Batabyal and Beladi (2011); Berg et al. (2002); Brunk and Hunter (2008); Chae (2012); Cioffi-Revilla (2014); Colander et al. (2010); Cooper (2011); Dai (2021); Dosi and Roventini (2017); Elsner (2017); Espejo (2018); Friedrich et al. (2021); Gaffeo and Tamborini (2011); Garmendia and Gamboa (2012); Gligor et al. (2022); González-Velasco et al. (2019); Green and Newth (2001); Hausner et al. (2021); Hommes (2006); Kirman (2010); Kukacka and Kristoufek (2020); Lee and Kim (2018); Li et al. (2020); Majeed and Shah (2015); May et al. (2011); Monasterolo et al. (2019); Mueller (2020); Oldham (2020); Raimbault (2019); Rammel et al. (2007); Rutkauskas et al. (2014); Shen (2021); Sitthiyot (2019); Stahel (2006); Stauffer et al. (2022); Sun and Zhong (2020); Tesfatsion (2006); Villani et al. (2018); Yaneer (2004); Zhang et al. (2021a);	43
Adaptation	Adamides and Pomonis (2009); Aldunate et al. (2005); Bento and Garotti (2019); Braz and de Mello (2022); Corbacioglu and Kapucu (2006); Garver (2019); Kim and Mackey (2014); Kukacka and Kristoufek (2021); Li et al. (2010); Maswana (2009); Matutinović (2001); Milne (2009); Sfa et al. (2020); Shobe (2020); Wiesner and Ladyman (2021); Zhang and Cui (2016)	16
Resilience	Darnhofer (2014); Fraccascia et al. (2018); Korhonen and Snäkin (2015); Plummer and Armitage (2007); Shachak and Boeken (2010);	5
Grand Total		113

In complexity thinking, systems exhibit non-linear effects and as such they behave in ways that the effects of inputs may not be proportional to outcomes (Beinhocker 2006). From this perspective, slight changes to conditions (initial or in the external environment) can result in larger unpredictable consequences (Turner and Baker 2019). In this realm, systems operate in an unpredictable manner (Hanseth and Lyytinen 2016), reacting disproportionately to their environment (Turner and Baker 2019). For publications that focus on non-linearity, these studies attempt to develop and conceptualise social reality from a Complex Adaptive Systems (CAS) viewpoint, although placing emphasis on the non-linear nature of these systems or their external environment. For example, Gligor et al. (2022) apply this perspective in their observation of gender differences in logistical innovations. Touching on other concepts such as emergence and multiple causality, these scholars highlight how diversity in innovation teams and workforce provides a deeper understanding to customer needs. Their research also shows how applying complexity appropriate methods such as QCA can provide insights that other mainstream ‘regression-based’ approaches may be unable to. Taking an evolutionary approach, Chae (2012) also applies complexity theory to demonstrate predictability, localization, and emergence in service innovation. Importantly, Chae (2012) notes that the environment of service innovation is multifaceted, and uncertain.

Monasterolo et al. (2019) argue that traditional economic and financial risk models do not offer the capacity needed to develop appropriate climate risks models and climate-alignment opportunities. For these authors, this is due to the constrains of ‘equilibrium conditions and linearity of impacts, as well as by representative agents and intertemporal optimization’ (Monasterolo et al. 2019, p. 177). Instead, they attempt to fill this gap by advocating the use of complexity appropriate methods, such as agent-based and network models, for effective alignment between national and global climate targets. Supporting this, Balint et al. (2017) also argue that decentralised economic models offer alternatives to equilibrium-based models in their assessment of non-linear effects. Batabyal and Beladi (2011) take a similar non-equilibrium view in their assessment of agricultural resilience. These scholars also note a need for a departure from equilibrium-based approaches. From these studies, it is evident that the influence of complexity theory has resulted in a different worldview. This particular set of complexity thinkers demonstrate ways to identify and tackle non-linearity across complex systems. The particular focus on the area of ecological economics highlights the need for more realist assessment of policy impacts within this area. Nevertheless, publications within this cluster are premised on the notion that the social world operates in in an unstable and non-predictable uncertain manner. This highlights the need for new ontological and methodological frameworks that transcends the reductionist paradigm.

Our thematic analysis also identified 49 publications (Table 2) that attempt to capture the interactions between and across systems from a multidimensional perspective. These publications provide demonstrations on how individual components of a system affect each other, and in some cases, influence actions. Xepapadeas (2010) advocates the need for adequate modelling that looks at spatial interactions induced by feedback. He finds that linear dynamics are not adequate illustrations of ecological systems. Gimzauskiene and Kloviene (2010, 2011) provide two publications that focus on the application of complexity theory in performance measurement systems. Here, they show an understanding of how systems interact with, and react to, the external environment. Applying complexity theory to the management of building construction projects, both Naderpajouh and Hastak (2014) and Çıdık and Phillips (2021) emphasise the importance of social interactions on risk management. Zhu et al. (2017) provide a unique demonstration of levers and hubs, that is, when a component of a system has disproportionate influence over the whole due to structure and connections. Here, they attempt to develop a model that can predict degrading components. Mylek and Schirmer (2020) also provide an application of complexity thinking in communication strategies. They develop an approach to the design of communication, with the intent to match the complexity of the information with the population. Zhang et al.

(2021b) show how institutional complexity can be applied with paradox theory to aid in efficient industrial change management, especially when faced with paradoxicalities.

In so far as systems operate in a non-static dynamic manner, and interact with each other, their components tend to adapt to, or learn from, changes to their environment. Sixteen publications (Table 2) also focus on adaptation, these studies provide a diversity of theoretical and contextual applications within which a system is seen to generate adaptive capabilities. Corbacioglu and Kapucu (2006), Li et al. (2010), Adamides and Pomonis (2009) and Zhang and Cui (2016) apply complexity theory to management practices and show how their attributes spontaneously adapt to changes in the environment. Here, adaptation may stem from organisational learning and self-adaptation in for instance, dynamic disaster environments (Corbacioglu and Kapucu 2006) or interactions within the external environment (Adamides and Pomonis 2009). Zhang and Cui (2016) attempt to quantitatively describe how a complex adaptive system highlightssystems self-adaptive to changing environments while Li et al. (2010) attempts to develop a multi-agent model that also factors in path dependency. Kim and Mackey (2014) and Garver (2019) also demonstrate how the environmental legal system can be viewed as adapting to its environment and suggest a systems-based assessment methodological viewpoint. Studies within this thematic cluster demonstrate how adaption may emerge when systems are at tipping points (Shobe 2020) or on the edge of chaos and uncertainty (Kim and Mackey 2014).

Finally, five articles (Table 2) major on the concept of resilience within the context of complexity theory. These publications take a more ecological perspective in their application of complexity theory. Darnhofer (2014) sees resilience as how complexity theory views the economic world as fundamentally unpredictable and actors and organisations must adapt to face this unpredictability. Korhonen and Snäkin (2015) examine resilience alongside efficiency and see resilience as achieved through diversity of resources. Plummer and Armitage (2007) and Shachak and Boeken (2010) take a non-equilibrium viewpoint in their development of an evaluation frameworks for ecological co-management. These authors argue that interactions do not always produce linear outcomes but are important for social-ecological resilience. Using a bibliometric analysis, Fraccascia et al. (2018) provide a comparison research study in ecological studies focusing on resilience. They show the multidisciplinary nature of resilience, especially in the fields of environmental science, ecology, and engineering.

3.2.2. Methodology and Method

Table 3 shows the overall comparison of the dominant methodology approach to research design across the 113 publications. The largest frequency is for those 45 publications that use a quantitative design. There is a variety of quantitative designs including: multi-agent and agent based modelling (Kukacka and Kristoufek 2020; Li et al. 2010; Hommes 2006; Tesfatsion 2006); Scenario Analysis (Korhonen and Snäkin 2015); Risk Assessment (Naderpajouh and Hastak 2014); Real Option Analysis (ROA) (Guo et al. 2021); statistical analysis of empirical data (Mylek and Schirmer 2020; Kijazi and Kant 2013; Gimzauskiene and Kloviene 2010, 2011); Intelligent Algorithms (Jemmali 2022); Power Law Distributions (Phillips 2019); modelling of live and empirical data (Zhu et al. 2017); and a NK model of fitness landscapes (Adamides and Pomonis 2009).

Table 3. Frequency of Overall Method in the Research Design.

Methodological Approach	Publications	Total
Quantitative	Adamides and Pomonis (2009); Aldhyani et al. (2021); Aymanns et al. (2018); Chakraborti et al. (2021); Chikumbo et al. (2000); Evans et al. (2017); Forbes and Xie (2018); Friedrich et al. (2021); Gimzauskiene and Kloviene (2010, 2011); González-Velasco et al. (2019); Guo et al. (2021); Hartwell (2019); Hommes (2006); Jemmali (2022); Kopp and Salecker (2020); Korhonen and Snäkin (2015); Korotkikh and Korotkikh (2009); Kukacka and Kristoufek (2021); Kukacka and Kristoufek (2020); Lamghari Elidrissi et al. (2020); Lee and Kim (2018); Li et al. (2010); Maswana (2009); Matesanz Gomez et al. (2017); Millhiser and Solow (2007); Mylek and Schirmer (2020); Naderpajouh and Hastak (2014); Oldham (2020); Phillips (2019); Qiu-Xiang et al. (2018); Rutkauskas et al. (2014); Shachak and Boeken (2010); Shen (2021); Sitthiyot (2019); Stauffer et al. (2022); Tang and Gao (2014); Tesfatsion (2006); Watson et al. (2011); Xepapadeas (2010); Zhang and Cui (2016); Zheng and Chen (2012); Zhu et al. (2017); Villani et al. (2018); Zhang et al. (2021a)	45
Case study and/or systems model	Ahmad (2019); Albin and Foley (2001); Aldunate et al. (2005); Aouad and Bento (2019); Batabyal and Beladi (2011); Bento and Garotti (2019); Berg et al. (2002); Bianchi and Labory (2019); Braz and de Mello (2022); Brunk and Hunter (2008); Budd et al. (2017); Chae (2012); Cioffi-Revilla (2014); Darnhofer (2014); Dong and Fisher (2019); Dosi and Roventini (2017); Espejo (2018); Garmendia and Stagl (2010); Garver (2019); Kim and Mackey (2014); Liu et al. (2021); Matutinović (2001); Oughton et al. (2018); Patrucco (2011); Plummer and Armitage (2007); Ryan et al. (2007); Sfa et al. (2020); Shobe (2020); Stuart et al. (2022); Vallance (2016); Varga et al. (2016); Watson et al. (2011); Markose (2005); Yaneer (2004); May et al. (2011); Marle (2020); Brocal et al. (2019); Wheeler (2007); Raimbault (2019); Li et al. (2020); Dai (2021); Sun and Zhong (2020)	42
Qualitative	Bruno et al. (2018); Çıdık and Phillips (2021); Coyne et al. (2021); Elsner (2017); Georgiev et al. (2015); Hausner et al. (2021); Kirman (2010); Milne (2009); Mueller (2020); Stahel (2006); Wiesner and Ladyman (2021)	11
Literature Review	Aeeni and Saeedikiya (2019); Balint et al. (2017); Colander et al. (2010); Cooper (2011); Gaffeo and Tamborini (2011); Green and Newth (2001); Majeed and Shah (2015); Monasterolo et al. (2019); Rammel et al. (2007)	9
Mixed Methods	Corbacioglu and Kapucu (2006); Garmendia and Gamboa (2012); Gligor et al. (2022); Wink et al. (2017)	4
Systematic Literature Review	Fraccascia et al. (2018); Zhang et al. (2021b)	2
Grand Total		113

These examples show the use of quantitative methods to model complexity are diverse, ranging from theoretical mathematical modelling of what a complex economic system might be like, to empirical based models that use historical or current data collections. The quantitative designs explore research questions both for macroeconomics and microeconomics. Furthermore, it is clear in the systematic review that complexity theorists often try to include aspects of the interaction of macro and microeconomics, and the interface between them. This mid-level interaction is referred to in this article as 'meso economics'. The quantitative microeconomics research that is identified includes applications for financial markets, manufacturing, production, engineering, construction, and environmental concerns.

There are four mixed methods publications in Table 3 and these include the mathematical approaches of Qualitative Comparative Analysis (QCA) (Gligor et al. 2022) and Cluster Analysis (Garmendia and Gamboa 2012). This is interesting given that these are case based methods widely advocated for exploring and explaining complexity in the political sciences and sociology (Rihoux and Ragin 2009). Case based methods are regarded as appropriate in these disciplines because of their ability to detect different causal configurations that evidence social complexity (Haynes 2018). Corbacioglu and Kapucu (2006) used mixed methods to compare the economic adaptation of communities in disasters in Turkey.

Forty-two publications included in our systematic review are best described methodologically as case studies and/or system models. This combined category is because the case studies about complexity theory are not mutually exclusive from system models, which also often use real world examples to embed their concepts. However, some of these articles did this case example embedding much more than others. Some of the system models were primarily case studies and then analysed as system models, while other articles set out much more to define approaches to system modelling and perhaps only included a limited and generalised real-world example. An example of theoretical system modelling is Plummer and Armitage's (2007) model of adaptive co-management of resources in a complex environment.

Garver's (2019) publication in Ecological Economics is recognisable as a form of economic discourse, and it takes an abstract theoretical approach to the ambitious topic of the global economic system. The article examines the economy in relation to the interventions of the law and governance and how these macro entities interact through leverage points and lock-ins. If there is a case study in this article, it is the global system. This system modelling approach has some similarities to Kim and Mackey's (2014 synthesis of international law as a complex adaptive system).

In contrast, the publications by Darnhofer (2014) and Braz and de Mello (2022) develop system models that are much more explicitly embedded in real world case studies. Darnhofer focuses on farm management as a definable system, but not a specific farm, or farming community in time and space. Braz and de Mello use the case study of a well specified supply chain economy in Brazil, using 'within' and 'across' case analysis. This case analysis clearly aids the explanation and conceptualisation of their theoretical model.

There were two previous systematic literature reviews discovered in our systematic review. Zhang et al. (2021b) explore complex 'paradoxes' in supply chain management, for example, improving inventory levels for operational flexibility and effectiveness, but whilst reducing inventory costs. Their article therefore sought out previous research on a very discrete subtopic on the periphery of mainstream complexity theory and was unlike the theoretical coverage of many (but not all) of the publications in our review. In other words, it is found that many of the articles attempted the opposite approach to Zhang et al. (2021b), preferring to offer ambitious and broad coverages across the metatheoretical landscape of complexity theory, rather than understanding a discrete sub-concept such as paradox. The narrow focus of their systematic review is a methodology strength in our opinion. Fraccascia et al. (2018) examine what they describe as 'state of the art' literature on complex systems and resilience and argue the literature is interdisciplinary but lacking in a shared understanding of a definition of resilience as a concept. They use a novel cross citation network analysis of the literature identified.

Nine other articles listed in our review were primarily unsystematic literature reviews. These publications deliberately use the research design of focusing on a controversial or seminal set, or single piece, of literature about the application of complexity to economics. The authors seek to add some original points to these arguments. Sometimes these designs are related to interdisciplinary theoretical areas. For example, Monasterolo et al. (2019) use some existing literature to argue for the need for a more robust approach to integrating macroeconomics models with an ecological perspective. Levanti (2018) looks at specific aspects of leading macroeconomic policy in complex socio-economic networks. While these two papers are in danger of presenting a rather esoteric contribution to the metatheoretical challenges of applying complexity theory, in contrast, Balint et al. (2017) present a well-structured critique of the literature on key areas of methodology. They focus on some literature covering the use of agent-based, network, and system dynamics models in ecological economics. While an obvious critique is that this is not done using a systematic method, the paper nevertheless gives a well-structured and robust account of methods that our own systematic review here also evidences as core territory to the application of complexity theory to economics. As a result, on page 262, Table 1, they provide a convincing summary of a comparison of system dynamics models and agent-based models with traditional equilibrium-based models. In their conclusion, they add weight to the prevailing direction of methodological changes in the sphere of applying complexity theory to economics:

' . . . agent-based models are increasingly considered as a prominent alternative to standard general equilibrium models which overlook many of the risks of climate change.' (Op cit, 262)

In another of the identified articles, Holt et al. (2011) draw primarily on existing writings rather than data and modelling to progress scholarship. They make an explicit and unapologetic contribution to an ongoing debate in the literature about whether complexity economics is really mainstream or heterodox. They argue that heterodox economics, especially when exploring scientific theories such as complexity, is not heterodox, but rather a necessary evolving of the mainstream discipline.

Finally, there are 11 articles in Table 3 using qualitative research methods. The definition of 'qualitative' here is broad and includes publications that discuss and argue conceptual issues, without being founded on specific literature or literature searches. For example, Kirman (2010) argues the weaknesses of General Equilibrium Theory and its impact on financial modelling and sees a complex adaptive systems approach as needed to better forecast major economic change. Coyne et al. (2021) look at conceptual issues and challenges in the interdisciplinary domain that overlaps public health with economics.

Çıdık and Phillips (2021) collect empirical qualitative data and analyse it for the journal Construction Management and Economics. They combine complexity theories with organisational approaches to reliable organisations to understand the impact of organisational culture on building safety. This is an alternative to taking a reductionist and quantitative approach to risk that assumes organisational stability over time. Twelve unstructured interviews are used to obtain evidence from experts. Social interaction is viewed as an important aspect in mitigating risk, this in addition to classical approaches to quantitative assessments of materials and costs associated with risks of combustibility.

3.2.3. Applications and Impact

Table 4 shows the frequencies of the type of application in economics contributed by the publications reviewed. This is on the basis of a division into the categories: macro, meso and micro.

Table 4. Summary of the Publication Applications.

Summary Type of Application	Publications	Total
Macro	Aymanns et al. (2018); Balint et al. (2017); Bianchi and Labory (2019); Brunk and Hunter (2008); Bruno et al. (2018); Chikumbo et al. (2000); Colander et al. (2010); Cooper (2011); Darnhofer (2014); Dosi and Roventini (2017); Elsner (2017); Espejo (2018); Evans et al. (2017); Forbes and Xie (2018); Friedrich et al. (2021); Gaffeo and Tamborini (2011); Garver (2019); González-Velasco et al. (2019); Green and Newth (2001); Kim and Mackey (2014); Kirman (2010); Kukacka and Kristoufek (2020); Lee and Kim (2018); Maswana (2009); Matesanz Gomez et al. (2017); Matutinović (2001); Milne (2009); Monasterolo et al. (2019); Mueller (2020); Plummer and Armitage (2007); Ryan et al. (2007); Sitthiyot (2019); Stauffer et al. (2022); Tang and Gao (2014); Wiesner and Ladyman (2021); Xepapadeas (2010); Zheng and Chen (2012); Watson et al. (2011); Markose (2005); Yaneer (2004); Villani et al. (2018); Fraccascia et al. (2018); Raimbault (2019); Li et al. (2020); Sun and Zhong (2020)	45
Meso	Adamides and Pomonis (2009); Aeeni and Saeedikiya (2019); Ahmad (2019); Batabyal and Beladi (2011); Braz and de Mello (2022); Budd et al. (2017); Chae (2012); Chakraborti et al. (2021); Cioffi-Revilla (2014); Corbacioglu and Kapucu (2006); Coyne et al. (2021); Garmendia and Gamboa (2012); Garmendia and Stagl (2010); Georgiev et al. (2015); Gimzauskiene and Kloviene (2010, 2011); Hartwell (2019); Hausner et al. (2021); Hommes (2006); Korhonen and Snäkin (2015); Korotkikh and Korotkikh (2009); Kukacka and Kristoufek (2021); Li et al. (2010); Millhiser and Solow (2007); Oldham (2020); Patrucco (2011); Phillips (2019); Qiu-Xiang et al. (2018); Rammel et al. (2007); Rutkauskas et al. (2014); Shachak and Boeken (2010); Shobe (2020); Stahel (2006); Stuart et al. (2022); Tesfatsion (2006); Wink et al. (2017); Zhang and Cui (2016); Zhang et al. (2021a); Dai (2021)	39
Micro	Albin and Foley (2001); Aldunate et al. (2005); Aouad and Bento (2019); Bento and Garotti (2019); Berg et al. (2002); Çıdık and Phillips (2021); Dong and Fisher (2019); Gligor et al. (2022); Guo et al. (2021); Jemmali (2022); Kopp and Salecker (2020); Lamghari Elidrissi et al. (2020); Liu et al. (2021); Majeed and Shah (2015); Mylek and Schirmer (2020); Naderpajouh and Hastak (2014); Oughton et al. (2018); Sfa et al. (2020); Shen (2021); Vallance (2016); Varga et al. (2016); Watson et al. (2011); Zhu et al. (2017); May et al. (2011); Marle (2020); Zhang et al. (2021b); Brocal et al. (2019); Aldhyani et al. (2021); Wheeler (2007)	29
Grand Total		113

Macroeconomics counts applications that are primarily directed at national and global economic issues. Microeconomic applications count applications that are primarily concerned with specific organisations and how individual agent behaviour contributes to collective phenomena. Meso economics is focused on the interaction of micro elements with macro elements and how the two dimensions influence and change each other. For example, from the publications considered in our systematic review, Li et al. (2010) see the disruption and uncertainties that external influences have on local organisational processes and systems.

The frequencies of the trio of macro, meso, micro groupings for the publications considered are relatively evenly distributed (Table 4). The importance of meso considerations (n = 39) shows that complex systems theoretical frameworks can be expected to lead researchers towards the interface between individual agents with their locality and the associated relational connections they have with regions, nations, and globalisation. Examples of this from the meso publications we review, include Shobe's (2020) critique of the difficulty with applying an optimal policy process to the decentralisation of environmental policy making due to the tightly connected and interlinking of organisations and devolved political geographies who have a stake in outcomes. Shobe sees studies of linked and adaptive complex systems as a key methodology for improving policy applications.

Zhang et al. (2021b) acknowledge the paradoxes and contradictions within patterns of agent interactions and how they make sense of the world. This is an aspect of complex outcomes that needs acknowledgment in applications to management practice. Korhonen and Snäkin's (2015) approach to modelling the Finnish energy system argues it is important to research the interdependent aspect of systems and their boundaries, thereby examining a specific city in relation to other municipalities with which it is regionally linked. Such a quantitative model evidences the importance of collaboration between cities and regions. Similarly, but applying a case study approach rather than a quantitative model, Braz and de Mello (2022) propose a meso complexity informed systems framework that includes management mechanisms, in addition to attributes of the internal and external environment.

Patrucco (2011) describes how changes in the network of the automobile production system in northern Italy is sustained by the dynamic interactions between firms. This modifies the behaviour of the key economic actors involved and promotes cross sector innovation and change with macro consequences. Furthermore, Gimzauskiene and Kloviene (2011), argue for an integrated meso approach to performance management that includes both internal and external influences for any individual organisation. Complexity, as uncertainty in the operating environment, becomes a key component of adaptive performance management (Gimzauskiene and Kloviene 2010). Likewise Batabyal and Beladi (2011) argue the multiple influences on range management in farming.

Garmendia and Gamboa's (2012) seek to model the many interests of different social groups towards sustainable natural resource management in northern Spain. They evidence that patterns of actor priorities can be grouped rather than being unrelated, but more importantly can give feedback into the dynamics of higher-level deliberation about social and economic change.

A model of service innovation developed by Chae (2012, p. 820) provides evidence of the meso dimensions that impact change in the service industry.

'Services arise and are emergent through recombination and/or reconfiguration of diverse resources and contexts from service provider, customer, and other economic actors. This recombinant/reconfiguring process, along with an effective balance of mutation and crossover, is a key for business growth and customer service experience.'

Garmendia and Stagl (2010) examine the interaction of participation about sustainability with the need to change social views and attitudes to ecology and economics. They conclude that there are uncertainties about participatory approaches to changing public attitudes towards the economics of sustainability.

These sorts of ambitious attempts at modelling the meso complexity and uncertainty of interactive agents and systems raises the issue about how useful such models can be for applied operational management applications, and whether the research outcomes offer only broad advice, such as the need for a good external view of economic and social change, and the ability to adapt policy and decisions rapidly in response.

Nevertheless, Adamides and Pomonis (2009) argue the emergence of new forms of organisational order from this complex and unpredictable range of meso influences. Rammel et al. (2007) assert the importance of a conceptual approach that understands non-linearity, non-equilibrium, and the resulting co-evolution of the system, if any progress is to be made in a research agenda that informs resource management. For Tesfatsion (2006) and Hommes (2006) ABM and its advancement through related dynamic methods is the research design of choice for progressing research on these meso-economic approaches. Phillips (2019) is concerned with the real-world example of bankers and financiers at the micro level misunderstanding risk with adverse consequences for macroeconomic policy, as in the Great Financial Crisis of 2008. Power Law distributions are seen as the research solution for getting better decision making that avoids such risks in the future.

Given that complexity focused approaches to solving meso-economic challenges are conceptually and methodologically ambitious, and offer limited insights to 'wicked' problems, it is not surprising that some applications identified in the selected publications reviewed still use either a macro or micro approach. Here, the system boundaries are restricted to either the global geopolitical economic system, often in the form of ecological economics, or the detail of production or performance within a single organisational system.

Of the 45 publications identified in Table 4 as having macro applications, five of these are published in the journal Ecological Economics. Monasterolo et al. (2019) argue for the changes in economic modelling necessary if countries are to hit global climate targets. Complexity science and evolutionary economics are seen as providing the fundamental framework for these changes. Garver (2019) argues for law and governance system changes in order to make global ecological improvements and identifies system leverage points to achieve change. Balint et al. (2017) cite both micro and macroeconomic literature, but produce largely macroeconomic conclusions about the importance of complexity economics models to provide knowledge on coalition formation, the macroeconomic impact of climate change, energy market dynamics, and the uptake of sustainable technologies. An earlier article by Plummer and Armitage (2007) provides an evaluative framework for the adoption of macro ecosystem and livelihood conditions, alongside the necessary governance changes. Matutinović (2001) hypothesises that socio-economic diversity is a prerequisite for social and ecological stability. At the core of these articles published in Ecological Economics is a meta world view of economics embedded with other global systems such as the available physical resources, climate dynamics and population demographics.

The other publications classified as being applied to macroeconomics in Table 4 include two in sources covering agricultural economics. Xepapadeas (2010) argues for a complexity approach to modelling that includes: nonlinear feedbacks, and spatial and temporal aspects. Darnhofer (2014) develops the concept of resilience as an alternative to equilibrium. This is a method for substantiating both complex system dynamics and the role of individual farms in managing macro social and ecological change. Farms are resilient to external changes by having a buffer capability, and an adaptive and transformative capability.

Other articles classified in Table 4 as having a macroeconomic approach include one about the importance of legal governance within a global economic system (Kim and Mackey 2014). Another by Holt et al. (2011) in Post Keynesian Economics argues that complexity economics needs a broad view of what is accepted within the economics discipline, rather than recognizing alternatives to classicism as heterodoxy. Furthermore, Brunk and Hunter (2008) offer an ecological approach to economic stagnation. They conclude that traditional macroeconomic policy approaches will exacerbate economic problems. Mueller (2020) concludes on the inevitable risk of policy failure for much macroeconomic intervention into complex global and national systems. Matesanz Gomez et al. (2017) argue the great financial crisis of 2008 resulted in increased macroeconomic system changes in Europe that challenge the commonly accepted notion of identifiable core and peripheral euro zone countries. Gaffeo and Tamborini (2011) examine the challenge of regulating macroeconomic finance in an age of globalisation and open capital markets. They see the usefulness of network theory and approaches but conclude that major questions remain about how to apply such idea to regulatory policy interventions. Milne (2009) examines the aims of macroprudential policy and concludes the most important aspect is to maintain the flow of finance through the economic system.

Twenty-nine publications are classified in Table 4 as being primarily micro in their complexity economics design. Jemmali (2022) creates a smart car parking system that enables a health service to vaccinate the most efficient number of people with a given period of time and resources. Gligor et al. (2022) research gender differences in logistics innovation. They identify different causal configurations for innovation with important gender differences across these configurations.

Several publications focus on novel approaches to understanding risk. Brocal et al. (2019) critique current models of risk management and propose a complex system of governance for better risk management. Çıdık and Phillips (2021) research professional opinions about high-risk buildings and conclude that collective culture as social interaction is an important aspect for reducing risk in addition to structured interventions for assessing physical conditions. Mylek and Schirmer (2020) measure the extent to which public actors have some degree of cognitive complexity as 'integrated complexity' in their ability to comprehend the socio-economic trade-offs, if policy interventions are to reduce wildfire risks. Another study that brings advances in complexity theory to agent perspectives and how they interact dynamically, is Naderpajouh and Hastak's (2014) concept of Interaction Analysis (IA). This forms a new type of risk assessment that estimates newly emerging risks in major construction projects.

Other microeconomic publications informed by complexity economics also examine uncertain outcomes from agent activities. Guo et al.'s (2021) model theorises how to optimise concessions to suppliers. This is during the unstable context of managing an ongoing public-private partnership contract. Kopp and Salecker (2020) look at how traders interact with their neighbours. Sellers' decisions about a buyer are often influenced by debt obligations and past interactions, including similar experiences of education, and living in close proximity. Kijazi and Kant (2013) theorise an approach to complex agent behaviour as 'socially rational' agents. These types of analysis of complex agents takes economic perspectives far beyond the concept of rational economic agents who are assumed to behave in similar ways. Nevertheless, agents are influenced by complex patterns of behaviour, often influenced by similar cross cutting social networks.

Not all the micro focused publications examine human interaction and behaviour. Zhu et al. (2017) analyse physical inputs and processes, rather than human actors, this in an engineering production process. They develop a quantitative model to measure the likely boundaries of component degradation (condition-based maintenance) within a complex multifaceted engineering process.

4. Conclusions

Figure 2 shows the frequency of publications from our systematic review over time. This highlights the growing applications of complexity theory in economic research. The trend for the publications included in our systematic review also provides some interesting suggestions on the influence of the global environment and the importance of the contextualization of such events. For instance, the increase in publications during 2010 reflects the impacts of the financial crisis and a moment when academic scholarship further questioned the use of traditional 'linear' economic ontologies. There is also an increasing trend in relevant publications during the last 15 years, with some fluctuations in the most recent period after 2018. This may result because of publications taking their time to get listed on databases.

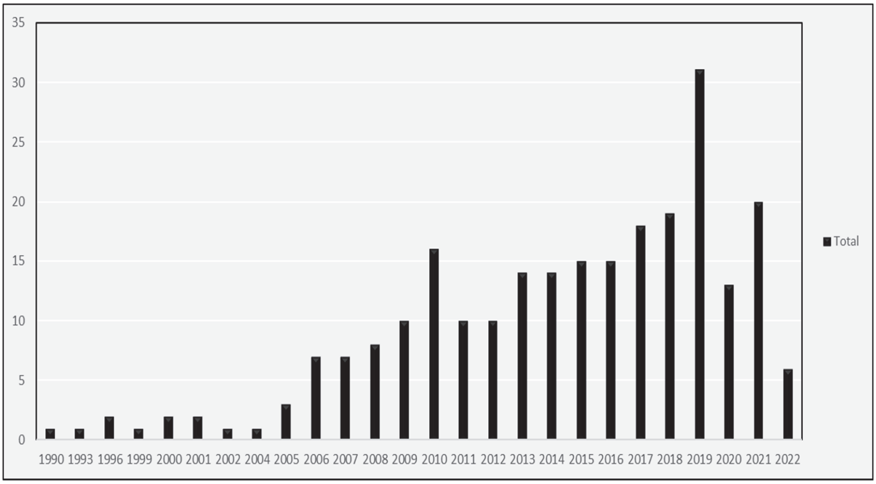

Figure 2. Frequency of Publications Over time.

Table 5 shows which of the reviewed publications have had the greatest impact through citations. The ratio index shows the annual ratio of citations per year, to avoid overly rating the impact of older articles. Hommes (2006) and Tesfatsion's (2006) contributions to the Handbook of Computational Economics have the highest ratios, an indication of the primacy that economics places on these kinds of contemporary quantitative models. Next, in Table 5 are a cluster of several authors publishing in the journal Ecological Economics, a journal that is clearly making an important impact in the integration of complexity theory with macro and meso system modelling that juxtaposes environmental and sustainability issues with the operations of governments and markets. This typifies the interdisciplinary approach of complexity economics. It is also an important observation given the current international concern about climate change and climate warming.

Table 5. Yearly ratio of citations—Top 15.

Rank	Author(s)	Year	Cited by	Ratio
1	Hommes	2006	576	33.88
2	Tesfatsion	2006	447	26.29
3	Plummer and Armitage	2007	346	21.63
4	Darnhofer	2014	161	17.89
5	Rammel, Stagl, and Wilfing	2007	228	14.25
6	Garmendia and Stagl	2010	175	13.46
7	Balint et al.,	2017	73	12.17
8	Fraccascia, Giannoccaro and Albino	2018	45	9.00
9	Kirman (2010)	2010	116	8.92
10	Coyne et al. (2021)	2021	16	8.00
11	Jemmali (2022)	2022	7	7.00
12	Garmendia and Gamboa (2012)	2012	72	6.55
13	Brocal et al. (2019)	2019	26	6.50
14	Zhang et al. (2021b)	2021	13	6.50
15	Kukacka and Kristoufek	2020	19	6.33

The interdisciplinary nature of complexity theory goes much wider than just this important single journal of Ecological Economics, and the different juxtapositions of other disciplines with economics is evidenced by the range of publication titles in Table 1. The nature of this interdisciplinarity is that it manifests itself especially in sub-disciplinary areas such as business and management studies, and ecology, rather than impacting the historical core of the foundation social science disciplines such as economics, sociology, and psychology. This leaves complexity economics open to the criticism that it is on the periphery of the single discipline. However, conversely the interdisciplinary approaches are contemporary and dealing with current real-world problems and applications. An example is Darnhofer's system model of farm management published in 2014 that deals with how farms can be resilient to ecological and economic change, and it has the fourth highest citation ratio in Table 5. Not all interdisciplinary complexity work is applied across the meso 'connect' and some highly cited scholars attempt an ambitious system world macro view especially when modelling global issues such as sustainability (Balint et al. 2017; Plummer and Armitage 2007).

In the publications reviewed there was a strong presence of ecology and climate change modelling that locate economics in a greater 'world/ecological/global system'. Table 1 evidences this with nine publications coming from Ecological Economics and seven from Ecological Complexity. (The later journal includes important global systems approaches but is less directly related to economics and the analysis of markets). Highly cited examples of the more economics focused approaches in Table 5 include Rammel et al. (2007) and Garmendia and Stagl (2010) and Garmendia and Gamboa (2012).

In terms of metatheory, complexity contributes a stinging critique of earlier economic approaches through its focus on unpredictability (Arthur 2013, 2021; Beinhocker 2006). Linked to this is a revision of economic agents who become not only consumers, but active actors politically, and motivated by a range of social aspects. Agents are therefore seen as heterogeneous and diverse, but with some consistent patterns of behaviour, often manifest in networks. They are not homogeneous. This has influenced the type of modelling that results (Hommes 2006; Tesfatsion 2006; Monasterolo et al. 2019; Balint et al. 2017).

This aspect of complexity, and the acknowledgement of multiple influences on dynamic economic systems, results in numerous publications having a strong ambition to develop meso models that have the potential to include both macroeconomics and microeconomics in some aspect. Here, there is a linking of the levels of analysis (Korhonen and Snäkin 2015; Garmendia and Stagl 2010; Chae 2012). However, this also results in a caution about how the theory is applied. Sometimes applications are speculative and based on general principles rather than offering prescriptive techniques. Good examples are the approaches towards the management of risk in sectors such as finance, engineering, and production (Chae 2012; Naderpajouh and Hastak 2014; Zhu et al. 2017).

Some important conclusions can be drawn about the use of methods in complexity economics. Agent based models and similar theoretical simulations of how a complex economy might behave are popular and often widely cited (Hommes 2006; Tesfatsion 2006). It is surprising that mixed methods and causal configurative methods do not feature more. The best example of such practices being a recent publication that includes Qualitative Comparative Analysis (Gligor et al. 2022). It is the experience of the authors of this systematic literature review that these types of contemporary complexity appropriate methods are used much more in political science, public policy and sociology. Furthermore, in the UK, they have made an important impact on the work of the HM Treasury (Bicket et al. 2020).

The use of general systems models in the tradition of the late Donella Meadows (2008) continue to have a wide use and impact in complexity economics, especially when concurrent with ecological issues (for example, Matutinović 2001; Garver 2019). This is again evidenced from the important impact of the journal Ecological Economics. System models that combine economics with environmental issues also span a wide range of publications including: Cambridge Journal of Regions, Economy and Society, Ecological Complexity, Forest Policy and Economics and The Economic Review of Agricultural Economics.

Overall, there are two key sets of publications that our systematic literature review exposed about complexity economics. On the one hand, there is the scholarship that explicitly addresses recognisable aspects of the contemporary agenda of the economics discipline. We have tended to focus on this literature in the examples used in the thematic analysis. The second area of literature is more implicit in its juxtaposition of complexity and economics, it being primarily concerned with the scientific development of complexity theory. In this second field of publications, the ambitious and wide coverage of complexity science across many disciplines, results in the continuing development of the broad theoretical perspective, but where the application to the working practice of economics is often weak. In our review we have chosen to highlight more the best working examples of the application of complexity science that in our opinion are having the greatest impact on the application and practice of economics. It is our argument that this continues to be an important development and one that will continue to change the nature of economics and its applications. The literature review provided gives evidence for our argument.

Author Contributions: Both authors have contributed equally to the article overall, especially with regard to thematic analysis. D.A. led on the PRISMA literature review, P.H. on the introduction and indicative Google Scholar exploration. All authors have read and agreed to the published version of the manuscript.

Funding: We are grateful to our employing institutions for their support: University of Brighton and University of Portsmouth. David Alemna's current post is funded by the Economic and Social Research Council [Project Reference: ES/W005743/1], as a South Coast Doctoral Training Partnership Postdoctoral Fellowship.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: It is possible to reproduce the systematic literature review used in this article, as the search terms are included in full in the text.

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

Note

[1] The resulting search string: (KEY ("complexity theory" OR "complex system" OR "complex adaptive system") AND SRCTITLE (econ* OR complex*) AND ABS (policy OR management OR organization OR finan*)) AND (LIMIT-TO (LANGUAGE, "English")).

References

Adamides, Emmanuel D., and Nikolaos Pomonis. 2009. The Co-Evolution of Product, Production and Supply Chain Decisions, and the Emer-Gence of Manufacturing Strategy. *International Journal of Production Economics* 121: 301–12. [CrossRef]

Aeeni, Zeynab, and Mehrzad Saeedikiya. 2019. Complexity Theory in the Advancement of Entrepreneurship Ecosystem Research: Future Research Directions. In *Eurasian Business Perspectives*. Cham: Springer International Publishing, pp. 19–37.

Ahmad, Syed Amaar. 2019. Urbanization, Energy Consumption and Entropy of Metropolises. *Complex Systems* 28: 287–312. [CrossRef]

Albin, Peter S., and Duncan K. Foley. 2001. The Co-Evolution of Cooperation and Complexity in a Multi-Player, Local-Interaction Pris-Oners' Dilemma. *Complexity* 6: 54–63. [CrossRef]

Aldhyani, Theyazn H. H., Manish R. Joshi, Shahab A. AlMaaytah, Ahmed Abdullah Alqarni, and Nizar Alsharif. 2021. Using Sequence Mining to Predict Com-Plex Systems: A Case Study in Influenza Epidemics. *Complexity* 2021: 9929013. [CrossRef]

Aldunate, Roberto G., Feniosky Pena-Mora, and Gene E. Robinson. 2005. Collaborative Distributed Decision Making for Large Scale Disaster Re-Lief Operations: Drawing Analogies from Robust Natural Systems. *Complexity* 11: 28–38. [CrossRef]

Antonelli, C. 2008. *Localised Technological Change: Towards the Economics of Complexity*. London: Routledge.

Antonelli, Cristiano. 2009. The Economics of Innovation: From the Classical Legacies to the Economics of Complexity. *Economics of Innovation and New Technology* 18: 611–46. [CrossRef]

Aouad, Jennifer, and Fabio Bento. 2019. A Complexity Perspective on Parent–Teacher Collaboration in Special Education: Narratives from the Field in Lebanon. *Journal of Open Innovation Technology Market and Complexity* 6: 4. [CrossRef]

Arthur, W. Brian. 2013. *Complexity Economics*. Oxford: Oxford University Press.

Arthur, W. Brian. 2021. Foundations of Complexity Economics. *Nature Reviews. Physics* 3: 136–45. [CrossRef] [PubMed]

Aymanns, Christoph, J. Doyne Farmer, Alissa M. Kleinnijenhuis, and Thom Wetzer. 2018. Models of Financial Stability and Their Application in Stress Tests ✶. In *Handbook of Computational Economics*. Amsterdam: Elsevier, pp. 329–91.

Balint, Tomas, Francesco Lamperti, Antoine Mandel, Mauro Napoletano, Andrea Roventini, and Alessandro Sapio. 2017. Complexity and the Economics of Cli-Mate Change: A Survey and a Look Forward. *Ecological Economics* 138: 252–65. [CrossRef]

Batabyal, Amitrajeet A., and Hamid Beladi. 2011. Ecological, Economic, and New Synthetic Perspectives in Range Management: An Interpretative Essay. *International Review of Environmental and Resource Economics* 5: 147–98. [CrossRef]

Beel, Jöran, and Bela Gipp. 2009. Google Scholar's Ranking Algorithm: An Introductory Overview. Paper presented at 12th International Conference on Scientometrics and Informetrics (ISSI'09), Rio de Janeiro, Brazil, July 14–17; vol. 1, pp. 230–41.

Beinhocker, Eric D. 2006. *The Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*. Boston: Harvard Business Press.

Bento, Fabio, and Luciano Garotti. 2019. Resilience beyond Formal Structures: A Network Perspective towards the Challenges of an Aging Workforce in the Oil and Gas Industry. *Journal of Open Innovation Technology Market and Complexity* 5: 15. [CrossRef]

Berg, Johannes, Matteo Marsili, Aldo Rustichini, and Riccardo Zecchina. 2002. Are Financial Markets Efficient? Phase Transition in the Aggregation of Information. *Complexity* 8: 20–23. [CrossRef]

Bernardo, José M., and Adrian F. M. Smith. 2009. *Bayesian Theory*. Hoboken: John Wiley and Sons, vol. 405.

Bianchi, Patrizio, and Sandrine Labory. 2019. Regional Industrial Policy for the Manufacturing Revolution: Enabling Conditions for Complex Transformations. *Cambridge Journal of Regions Economy and Society* 12: 233–49. [CrossRef]

Bicket, Martha, Ian Christie, Nigel Gilbert, Dione Hills, Alexandra Penn, and Helen Wilkinson. 2020. *Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation*. London: HM Treas.

Braz, Antonio Carlos, and Adriana Marotti de Mello. 2022. Circular Economy Supply Network Management: A Complex Adaptive System. *International Journal of Production Economics* 243: 108317. [CrossRef]

Brocal, Francisco, Cristina González, Dragan Komljenovic, Polinpapilinho F. Katina, and Miguel A. Sebastián. 2019. Emerging Risk Management in Industry 4.0: An Approach to Improve Organizational and Human Performance in the Complex Systems. *Complexity* 2019: 2089763. [CrossRef]

Brunk, Gregory G., and Kennith G. Hunter. 2008. An Ecological Perspective on Interest Groups and Economic Stagnation. *The Journal of Socio-Economics* 37: 194–212. [CrossRef]

Bruno, Bruna, Marisa Faggini, and Anna Parziale. 2018. Innovation Policies: Strategy of Growth in a Complex Perspective. In *Social Dynamics in a Systems Perspective*. Cham: Springer International Publishing, pp. 65–84.

Budd, Leslie, Alessandro Sancino, Michela Pagani, Ómar Kristmundsson, Borut Roncevic, and Michael Steiner. 2017. Sport as a Complex Adaptive System for Place-Based Leadership: Comparing Five European Cities with Different Administrative and Socio-Cultural Traditions. *Local Economy* 32: 316–35. [CrossRef]

Byrne, David, and Gill Callaghan. 2013. *Complexity Theory and the Social Sciences: The State of the Art*. London: Routledge.

Castellani, Brian, and Lasse Gerrits. 2021. Map of Complexity Sciences. Available online: https://www.art-sciencefactory.com/complexity-map_feb09.html (accessed on 25 July 2022).

Chae, Bongsug K. 2012. An Evolutionary Framework for Service Innovation: Insights of Complexity Theory for Service Science. *International Journal of Production Economics* 135: 813–22. [CrossRef]

Chakraborti, Anirban, Hrishidev, Kiran Sharma, and Hirdesh K. Pharasi. 2021. Phase Separation and Scaling in Correlation Structures of Financial Markets. *Journal of Physics: Complexity* 2: 015002. [CrossRef]

Chikumbo, Oliver, Roger Bradbury, and Stuart Davey. 2000. Large-Scale Ecosystem Management as a Complex Systems Problem: Mul-Ti-Objective Optimisation with Spatial Constraints. *Applied Complexity-From Neural Nets to Managed Landscapes* 8: 124–55.

Çıdık, Mustafa Selçuk, and Stephen Phillips. 2021. Buildings as Complex Systems: The Impact of Organisational Culture on Building Safety. *Construction Management and Economics* 39: 972–87. [CrossRef]

Cilliers, Paul. 1998. *Complexity and Postmodernism: Understanding Complex Systems*. London: Routledge.

Cioffi-Revilla, Claudio. 2014. Seeing It Coming: A Complexity Approach to Disasters and Humanitarian Crises. *Complexity* 19: 95–108. [CrossRef]

Colander, David. 2000. *Complexity and the History of Economic Thought*. London: Routledge.

Colander, David, Richard P. F. Holt, and J. Barkley Rosser. 2010. How to Win Friends and (Possibly) Influence Mainstream Economists. *Journal of Post Keynesian Economics* 32: 397–408. [CrossRef]

Cooper, Melinda. 2011. Complexity Theory after the Financial Crisis: The Death of Neoliberalism or the Triumph of Hayek? *Journal of Cultural Economy* 4: 371–85. [CrossRef]

Corbacioglu, Sitki, and Naim Kapucu. 2006. Organisational Learning and Selfadaptation in Dynamic DisasterEnvironments. *Disas-Ters* 30: 212–33. [CrossRef] [PubMed]

Coyne, Christopher J., Thomas K. Duncan, and Abigail R. Hall. 2021. The Political Economy of State Responses to Infectious Disease. *Southern Economic Journal* 87: 1119–37. [CrossRef]

Dai, Wensheng. 2021. Development and Supervision of Robo-Advisors under Digital Financial Inclusion in Complex Systems. *Complexity* 2021: 6666089. [CrossRef]

Darnhofer, Ika. 2014. Resilience and Why It Matters for Farm Management. *European Review of Agricultural Economics* 41: 461–84. [CrossRef]

Dong, Xiaoli, and Stuart G. Fisher. 2019. Ecosystem Spatial Self-Organization: Free Order for Nothing? *Ecological Complexity* 38: 24–30. [CrossRef]

Dosi, Giovanni, and Andrea Roventini. 2017. Agent-Based Macroeconomics and Classical Political Economy: Some Italian Roots. *Italian Economic Journal* 3: 261–83. [CrossRef]

Durlauf, Steven N. 2005. Complexity and Empirical Economics. *Economic Journal (London, England)* 115: F225–F243. [CrossRef]

Durlauf, Steven N. 2012. Complexity, Economics, and Public Policy. *Politics Philosophy & Economics* 11: 45–75. [CrossRef]

Elidrissi, Hajar Lamghari, Nait-Sidi-Moh Ahmed, and Tajer Abdelouahed. 2020. Knapsack Problem-Based Control Approach for Traffic Signal Management at Urban Intersections: Increasing Smooth Traffic Flows and Reducing Environmental Impact. *Ecological Complexity* 44: 100878. [CrossRef]

Elsner, Wolfram. 2017. Complexity Economics as Heterodoxy: Theory and Policy. *Journal of Economic Issues* 51: 939–78. [CrossRef]
Espejo, Raul. 2018. *Social Dynamics in a Systems Perspective*. Cham: Springer, pp. 121–35.
Evans, John, Neil Allan, and Neil Cantle. 2017. A New Insight into the World Economic Forum Global Risks. *Economic Papers A Journal of Applied Economics and Policy* 36: 185–97. [CrossRef]
Farmer, J. Doyne, and John J. Sidorowich. 1987. Predicting chaotic time series. *Physical review letters* 59: 845. [CrossRef] [PubMed]
Fontana, Magda. 2010. The Santa Fe Perspective on Economics: Emerging Patterns in The Science of Complexity. *History of Eco-Nomic Ideas* XVIII: 167–94.
Forbes, Dolores Jane, and Zhixiao Xie. 2018. Identifying Process Scales in the Indian River Lagoon, Florida Using Wavelet Transform Analysis of Dissolved Oxygen. *Ecological Complexity* 36: 149–67. [CrossRef]
Fraccascia, Luca, Ilaria Giannoccaro, and Vito Albino. 2018. Resilience of Complex Systems: State of the Art and Directions for Future Research. *Complexity* 2018: 3421529. [CrossRef]
Friedrich, Jan, Joachim Peinke, Alain Pumir, and Rainer Grauer. 2021. Explicit Construction of Joint Multipoint Statistics in Complex Systems. *Journal of Physics: Complexity* 2: 045006. [CrossRef]
Gaffeo, Edoardo, and Roberto Tamborini. 2011. If the Financial System Is Complex, How Can We Regulate It? *International Journal of Political Economy* 40: 79–97. [CrossRef]
Garmendia, Eneko, and Gonzalo Gamboa. 2012. Weighting Social Preferences in Participatory Multi-Criteria Evaluations: A Case Study on Sustainable Natural Resource Management. *Ecological Economics: The Journal of the International Society for Ecological Economics* 84: 110–20. [CrossRef]
Garmendia, Eneko, and Sigrid Stagl. 2010. Public Participation for Sustainability and Social Learning: Concepts and Lessons from Three Case Studies in Europe. *Ecological Economics: The Journal of the International Society for Ecological Economics* 69: 1712–22. [CrossRef]
Garver, Geoffrey. 2019. A Systems-Based Tool for Transitioning to Law for a Mutually Enhancing Human-Earth Relationship. *Ecological Economics: The Journal of the International Society for Ecological Economics* 157: 165–74. [CrossRef]
Georgiev, Georgi Yordanov, Kaitlin Henry, Timothy Bates, Erin Gombos, Alexander Casey, Michael Daly, Amrit Vinod, and Hyunseung Lee. 2015. Mechanism of Organization Increase in Complex Systems. *Complexity* 21: 18–28. [CrossRef]
Gimzauskiene, Edita, and Lina Kloviene. 2010. Research of the Performance Measurement System: Environmental Perspective. *Inžinerinė Ekonomika* 21: 180–86.
Gimzauskiene, Edita, and Lina Kloviene. 2011. Performance Measurement System: Towards an Institutional Theory. *Engineering Economics* 22. [CrossRef]
Gligor, David, Ivan Russo, and Michael J. Maloni. 2022. Understanding Gender Differences in Logistics Innovation: A Complexity Theory Perspective. *International Journal of Production Economics* 246: 108420. [CrossRef]
González-Velasco, Carmen, Marcos González-Fernández, and José-Luis Fanjul-Suárez. 2019. Does Innovative Effort Matter for Corporate Performance in Spanish Companies in a Context of Financial Crisis? A Fuzzy-Set QCA Approach. *Empirical Economics* 56: 1707–27. [CrossRef]
Green, David G., and David Newth. 2001. Grand Challenges in Complexity and Informatics. *Complexity International* 8: 36.
Guo, Kai, Limao Zhang, and Tao Wang. 2021. Concession Period Optimisation in Complex Projects under Uncertainty: A Public–Private Partnership Perspective. *Construction Management and Economics* 39: 156–72. [CrossRef]
Hanseth, Ole, and Kalle Lyytinen. 2016. Design Theory for Dynamic Complexity in Information Infrastructures: The Case of Build-Ing Internet. In *Enacting Research Methods in Information Systems*. Cham: Palgrave Macmillan, pp. 104–42.
Hartwell, Christopher A. 2019. Short Waves in Hungary, 1923 and 1946: Persistence, Chaos, and (Lack of) Control. *Journal of Economic Behavior & Organization* 163: 532–50. [CrossRef]
Hausner, Jerzy, Michał Możdżeń, and Marek Oramus. 2021. Fragility of Social and Economic Systems and the Role of 'Modality'. *Journal of Economic Issues* 55: 1118–38. [CrossRef]
Haynes, P. 2018. *Social Synthesis: Finding Dynamic Patterns in Complex Social Systems*. London: Routledge.
Holland, John H. 1992. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Cambridge: MIT Press.
Holt, Richard P. F., J. Barkley Rosser Jr., and David Colander. 2011. The Complexity Era in Economics. *Review of Political Economy* 23: 357–69. [CrossRef]
Hommes, Cars H. 2006. Chapter 23 Heterogeneous Agent Models in Economics and Finance. In *Handbook of Computational Economics*. Amsterdam: Elsevier, pp. 1109–86.
Jemmali, Mahdi. 2022. Intelligent Algorithms and Complex System for a Smart Parking for Vaccine Delivery Center of COVID-19. *Complex & Intelligent Systems* 8: 597–609. [CrossRef]
Kauffman, Stuart. A. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.
Kijazi, Martin Herbert, and Shashi Kant. 2013. Complexity Theory and Forest Resource Economics. In *Post-Faustmann Forest Resource Economics*. Dordrecht: Springer, pp. 41–70.
Kim, Rakhyun E., and Brendan Mackey. 2014. International Environmental Law as a Complex Adaptive System. *International Environmental Agreements Politics Law and Economics* 14: 5–24. [CrossRef]
Kirman, Alan. 2010. The Economic Crisis Is a Crisis for Economic Theory. *CESifo Economic Studies* 56: 498–535. [CrossRef]

Kopp, Thomas, and Jan Salecker. 2020. How Traders Influence Their Neighbours: Modelling Social Evolutionary Processes and Peer Ef-Fects in Agricultural Trade Networks. *Journal of Economic Dynamics and Control* 117: 103944. [CrossRef]

Korhonen, Jouni, and Juha-Pekka Snäkin. 2015. Quantifying the Relationship of Resilience and Eco-Efficiency in Complex Adaptive Energy Systems. *Ecological Economics: The Journal of the International Society for Ecological Economics* 120: 83–92. [CrossRef]

Korotkikh, Victor, and Galina Korotkikh. 2009. On Irreducible Description of Complex Systems. *Complexity* 14: 40–46. [CrossRef]

Kukacka, Jiri, and Ladislav Kristoufek. 2020. Do 'Complex'Financial Models Really Lead to Complex Dynamics? Agent-Based Models and Multifractality. *Journal of Economic Dynamics and Control* 113: 103855. [CrossRef]

Kukacka, Jiri, and Ladislav Kristoufek. 2021. Does Parameterization Affect the Complexity of Agent-Based Models? *Journal of Economic Behavior & Organization* 192: 324–56. [CrossRef]

Lee, Frederic S., and Marc Lavoie, eds. 2012. *Defense of Post-Keynesian and Heterodox Economics: Responses to Their Critics*. London: Routledge.

Levanti, Gabriella. 2018. Governing complex strategic networks: Emergence versus enabling effects. In *Governing Business Systems*. Cham: Springer, pp. 49–65.

Lee, Jaesung, and Dae-Won Kim. 2018. Scalable Multilabel Learning Based on Feature and Label Dimensionality Reduction. *Complexity* 2018: 6292143. [CrossRef]

Li, Gang, Hongjiao Yang, Linyan Sun, Ping Ji, and Lei Feng. 2010. The Evolutionary Complexity of Complex Adaptive Supply Networks: A Simulation and Case Study. *International Journal of Production Economics* 124: 310–30. [CrossRef]

Li, Wenrui, Menggang Li, Yiduo Mei, Ting Li, and Fang Wang. 2020. A Big Data Analytics Approach for Dynamic Feedback Warning for Complex Systems. *Complexity* 2020: 7652496. [CrossRef]

Liberati, Alessandro, Douglas G. Altman, Jennifer Tetzlaff, Cynthia Mulrow, Peter C. Gøtzsche, John P. A. Ioannidis, Mike Clarke, P. J. Devereaux, Jos Kleijnen, and David Moher. 2009. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *Journal of Clinical Epidemiology* 62: e1–e34. [CrossRef] [PubMed]

Liu, Jieling, Franz W. Gatzweiler, and Manasi Kumar. 2021. An Evolutionary Complex Systems Perspective on Urban Health. *Socio-Economic Planning Sciences* 75: 100815. [CrossRef]

Lorenz, Edward N. 1963. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences* 20: 130–41. [CrossRef]

Majeed, Abdul, and Munam Ali Shah. 2015. Energy Efficiency in Big Data Complex Systems: A Comprehensive Survey of Modern Energy Saving Techniques. *Complex Adaptive Systems Modeling* 3. [CrossRef]

Markose, Sheri M. 2005. Computability and Evolutionary Complexity: Markets as Complex Adaptive Systems (CAS). *Economic Journal (London, England)* 115: F159–F192. [CrossRef]

Marle, Franck. 2020. An Assistance to Project Risk Management Based on Complex Systems Theory and Agile Project Management. *Complexity* 2020: 3739129. [CrossRef]

Maswana, Jean-Claude. 2009. A Contribution to the Empirics of Finance-Growth Nexus in China: A Complex System Perspective. *Global Economic Review* 38: 29–47. [CrossRef]

Matesanz Gomez, David, Hernan J. Ferrari, Benno Torgler, and Guillermo J. Ortega. 2017. Synchronization and Diversity in Business Cycles: A Network Analysis of the European Union. *Applied Economics* 49: 972–86. [CrossRef]

Matutinović, Igor. 2001. The Aspects and the Role of Diversity in Socioeconomic Systems: An Evolutionary Perspective. *Ecological Economics: The Journal of the International Society for Ecological Economics* 39: 239–56. [CrossRef]

May, Christopher J., Michelle Burgard, and Imran Abbasi. 2011. Teaching Complexity Theory through Student Construction of a Course Wiki: The Self-Organization of a Scale-Free Network. *Complexity* 16: 41–48. [CrossRef]

Meadows, D. 2008. *Thinking in Systems: A Primer*. London: Earthscan.

Millhiser, William P., and Daniel Solow. 2007. How Large Should a Complex System Be? An Application in Organizational Teams. *Complexity* 12: 54–70. [CrossRef]

Milne, Alistair. 2009. Macroprudential Policy: What Can It Achieve? *Oxford Review of Economic Policy* 25: 608–29. [CrossRef]

Moher, David, Alessandro Liberati, Jennifer Tetzlaff, Douglas G. Altman, and PRISMA Group. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine* 151: 264–69. [CrossRef] [PubMed]

Moher, David, Lesley Stewart, and Paul Shekelle. 2016. Implementing PRISMA-P: Recommendations for Prospective Authors. *Systematic Reviews* 5: 15. [CrossRef] [PubMed]

Monasterolo, Irene, Andrea Roventini, and Tim J. Foxon. 2019. Uncertainty of Climate Policies and Implications for Economics and Finance: An Evolutionary Economics Approach. *Ecological Economics: The Journal of the International Society for Ecological Economics* 163: 177–82. [CrossRef]

Mueller, Bernardo. 2020. Why Public Policies Fail: Policymaking under Complexity. *EconomiA* 21: 311–23. [CrossRef]

Mylek, Melinda R., and Jacki Schirmer. 2020. Understanding Acceptability of Fuel Management to Reduce Wildfire Risk: Informing Communication through Understanding Complexity of Thinking. *Forest Policy and Economics* 113: 102120. [CrossRef]

Naderpajouh, Nader, and Makarand Hastak. 2014. Quantitative Analysis of Policies for Governance of Emergent Dynamics in Complex Construction Projects. *Construction Management and Economics* 32: 1222–37. [CrossRef]

Oldham, Matthew. 2020. Quantifying the Concerns of Dimon and Buffett with Data and Computation. *Journal of Economic Dynamics & Control* 113: 103864. [CrossRef]

Oughton, Edward J., Will Usher, Peter Tyler, and Jim W. Hall. 2018. Infrastructure as a Complex Adaptive System. *Complexity* 2018: 3427826. [CrossRef]

Oxman, Andrew D., and Gordon H. Guyatt. 1993. The Science of Reviewing Research. *Annals of the New York Academy of Sciences* 703: 125–33; discussion 133–4. [CrossRef] [PubMed]

Patrucco, Pier Paolo. 2011. Changing Network Structure in the Organization of Knowledge: The Innovation Platform in the Evidence of the Automobile System in Turin. *Economics of Innovation and New Technology* 20: 477–93. [CrossRef]

Phillips, E. 2019. Nassim Taleb heads international banking's first Grey/Black Swan Committee. *The Quarterly Review of Economics and Finance* 72: 117–22. [CrossRef]

Plummer, Ryan, and Derek Armitage. 2007. A Resilience-Based Framework for Evaluating Adaptive Co-Management: Linking Ecology, Economics and Society in a Complex World. *Ecological Economics: The Journal of the International Society for Ecological Economics* 61: 62–74. [CrossRef]

Qiu-Xiang, Li, Zhang Yu-Hao, and Huang Yi-Min. 2018. The Complexity Analysis in Dual-Channel Supply Chain Based on Fairness Concern and Different Business Objectives. *Complexity* 2018: 4752765. [CrossRef]

Raimbault, Juste. 2019. Second-Order Control of Complex Systems with Correlated Synthetic Data. *Complex Adaptive Systems Modeling* 7. [CrossRef]

Rammel, Christian, Sigrid Stagl, and Harald Wilfing. 2007. Managing Complex Adaptive Systems-a Co-Evolutionary Perspective on Natural Resource Management. *Ecological Economics* 63: 9–21. [CrossRef]

Rihoux, Benoît, and Charles C. Ragin. 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks: Sage Publication.

Rosser, J. Barkley, Jr. 1999. On the Complexities of Complex Economic Dynamics. *The Journal of Economic Perspectives: A Journal of the American Economic Association* 13: 169–92. [CrossRef]

Rutkauskas, Aleksandras Vytautas, Viktorija Stasytyte, and Edvard Michnevic. 2014. Universally Sustainable Development Strategy for a Small Country: A Systemic Decision. *Engineering Economics* 25. [CrossRef]

Ryan, Justin G., John A. Ludwig, and Clive A. Mcalpine. 2007. Complex Adaptive Landscapes (CAL): A Conceptual Framework of Mul-Ti-Functional, Non-Linear Ecohydrological Feedback Systems. *Ecological Complexity* 4: 113–27. [CrossRef]

Sfa, Fatima Ezahra, Mohamed Nemiche, and Hamza Rayd. 2020. A Generic Macroscopic Cellular Automata Model for Land Use Change: The Case of the Drâa Valley. *Ecological Complexity* 43: 100851. [CrossRef]

Shachak, Moshe, and Bertrand R. Boeken. 2010. Patterns of Biotic Community Organization and Reorganization: A Conceptual Framework and a Case Study. *Ecological Complexity* 7: 433–45. [CrossRef]

Shen, Yanan. 2021. Modeling and Research on Human Capital Accumulation Complex System of High-Tech Enterprises Based on Big Data. *Complexity* 2021: 6635228. [CrossRef]

Shobe, William. 2020. Emerging Issues in Decentralized Resource Governance: Environmental Federalism, Spillovers, and Linked Socio-Ecological Systems. *Annual Review of Resource Economics* 12: 259–79. [CrossRef]

Sitthiyot, Thitithep. 2019. On Rank-Size Distribution of Local Government Debt. *SSRN Electronic Journal* 37: 45–60. [CrossRef]

Sornette, Didier. 2006. *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools*. Berlin: Springer Science and Business Media.

Stahel, Andri W. 2006. Complexity, Oikonomía and Political Economy. *Ecological Complexity* 3: 369–81. [CrossRef]

Stauffer, Maxime, Isaak Mengesha, Konrad Seifert, Igor Krawczuk, Jens Fischer, and Giovanna Di Marzo Serugendo. 2022. A Computational Turn in Policy Process Studies: Coevolving Network Dynamics of Policy Change. *Complexity* 2022: 8210732. [CrossRef]

Stuart, Morgan, Andrew Placona, Gabe Vece, Kelsi Lindblad, Saikou Diallo, and Bob Carrico. 2022. Perspective on Macroscale Complexity in the National Transplant System. *Complexity* 2022: 3221885. [CrossRef]

Sun, Hui, and Xiong Zhong. 2020. Impact of Financial R&D Resource Allocation Efficiency Based on VR Technology and Machine Learning in Complex Systems on Total Factor Productivity. *Complexity* 2020: 6679846. [CrossRef]

Tang, Bo, and Jianzhong Gao. 2014. System Simulation and Reliability Assessment of Chinese Carbon Sequestration Market. *Journal of Systems Science and Complexity* 27: 760–76. [CrossRef]

Tesfatsion, Leigh. 2006. Chapter 16 Agent-Based Computational Economics: A Constructive Approach to Economic Theory. In *Handbook of Computational Economics*. Amsterdam: Elsevier, pp. 831–80.

Turner, John R., and Rose M. Baker. 2019. Complexity Theory: An Overview with Potential Applications for the Social Sciences. *Systems* 7: 4. [CrossRef]

Vallance, Paul. 2016. Universities, Public Research, and Evolutionary Economic Geography. *Economic Geography* 92: 355–77. [CrossRef]

Varga, Liz, Marguerite Robinson, and Peter Allen. 2016. Multiutility Service Companies: A Complex Systems Model of Increasing Resource Efficiency. *Complexity* 21: 23–33. [CrossRef]

Villani, Marco, Laura Sani, Riccardo Pecori, Michele Amoretti, Andrea Roli, Monica Mordonini, Roberto Serra, and Stefano Cagnoni. 2018. An Iterative Infor-Mation-Theoretic Approach to the Detection of Structures in Complex Systems. *Complexity* 2018: 3687839. [CrossRef]

Watson, Richard A., C. L. Buckley, and Rob Mills. 2011. Optimization in 'Self-Modeling' Complex Adaptive Systems. *Complexity* 16: 17–26. [CrossRef]

Wheeler, Thomas J. 2007. Analysis, Modeling, Emergence & Integration in Complex Systems: A Modeling and Integration Framework & System Biology. *Complexity* 13: 60–75. [CrossRef]

Wiesner, Karoline, and James Ladyman. 2021. Complex Systems Are Always Correlated but Rarely Information Processing. *Journal of Physics: Complexity* 2: 045015. [CrossRef]

Wink, Rüdiger, Laura Kirchner, Florian Koch, and Daniel Speda. 2017. Agency and Forms of Path Development along Transformation Processes in German Cities. *Cambridge Journal of Regions Economy and Society* 10: 471–90. [CrossRef]

Xepapadeas, Anastasios. 2010. Modeling Complex Systems: Modeling Complex Systems. *Agricultural Economics (Amsterdam, Netherlands)* 41: 181–91. [CrossRef]

Yaneer, Bar-Yam. 2004. Multiscale variety in complex systems. *Complexity* 9: 37–45.

Zhang, Ge, Weijie Wang, and Yikai Liang. 2021a. Understanding the Complex Adoption Behavior of Cloud Services by SMEs Based on Complexity Theory: A Fuzzy Sets Qualitative Comparative Analysis (FsQCA). *Complexity* 2021: 5591446. [CrossRef]

Zhang, Jiayuan, Mehmet G. Yalcin, and Douglas N. Hales. 2021b. Elements of Paradoxes in Supply Chain Management Literature: A Systematic Literature Review. *International Journal of Production Economics* 232: 107928. [CrossRef]

Zhang, Meng, and Jinchuan Cui. 2016. A Quantitative Description of Complex Adaptive System: The Self-Adaptive Mechanism of the Ma-Terial Purchasing Management System towards the Changing Environment. *Journal of Systems Science and Complexity* 29: 151–70. [CrossRef]

Zheng, Xiaolian, and Ben M. Chen. 2012. Modeling and Forecasting of Stock Markets under a System Adaptation Framework. *Journal of Systems Science and Complexity* 25: 641–74. [CrossRef]

Zhu, Qiushi, Peng Hao, Timmermans Bas, and van Houtum Geert-Jan. 2017. A Condition-Based Maintenance Model for a Single Compo-Nent in a System with Scheduled and Unscheduled Downs. *International Journal of Production Economics* 193: 365–80. [CrossRef]

economies

MDPI

Case Report

Domestic vs. External Economic Sectors and the Political Process: Insights from Greece

George Petrakos [1], Kostas Rontos [2], Luca Salvati [3,*], Chara Vavoura [4] and Ioannis Vavouras [1]

[1] Department of Public Administration, Panteion University, 136 Syngrou Avenue, EL-17671 Athens, Greece
[2] Department of Sociology, University of the Aegean, University Hill, EL-81100 Mytilene, Greece
[3] Department of Methods and Models for Economics, Territory and Finance, Sapienza University of Rome, Via del Castro Laurenziano 9, I-00161 Rome, Italy
[4] Department of Economics, National and Kapodistrian University of Athens, 1 Sofokleous Str., EL-10559 Athens, Greece
* Correspondence: luca.salvati@uniroma1.it

Abstract: Building on the well-established relationship between economic dynamics and political processes, we focus on the most important element of the political process, namely, general (or national) elections, and look into their effects on public finance and total economic output. In this vein, the present study has three objectives: (i) to investigate political budget cycles in Greece during the period known as the 'Third Hellenic Republic' (in Greek, 'Metapolitefsi', hereafter THR) since 1974; (ii) to assess whether national elections affect total economic activity in a stabilizing or destabilizing way; and (iii) to examine the possible effects of the external sector of the economy on the budget balance. The empirical findings of our analysis document how the Greek economy was characterized by sharp political budget cycles in correspondence with the THR, exerting a destabilizing effect on the total output of the economy. Performances of the external sector of the economy have significantly affected budget balances in Greece.

Keywords: political budget cycles; electoral cycles; economic destabilization; international trade

JEL Classification: D72; E62; H62

Citation: Petrakos, George, Kostas Rontos, Luca Salvati, Chara Vavoura, and Ioannis Vavouras. 2022. Domestic vs. External Economic Sectors and the Political Process: Insights from Greece. *Economies* 10: 198. https://doi.org/10.3390/economies10080198

Academic Editor: Franklin G. Mixon

Received: 30 March 2022
Accepted: 10 August 2022
Published: 17 August 2022

1. Introduction

Over the last decade, Greece has repeatedly faced bankruptcy risk as a result of its public debt crisis. Bankruptcy is still plausible nowadays, since the Greek government's consolidated gross debt has surpassed 200% of GDP[1], topping all European Union (EU) member states and amounting to more than twice the EU average. In this study, we document how the political process has been instrumental in the Greek debt crisis—mainly through the emergence of severe political budget cycles—and we explore the generating mechanism of these cycles and their impact on the stability of the aggregate economy and international trade.

Political budget cycles (or electoral fiscal cycles) occur when governments in pre-election periods pursue expansionary fiscal policies (e.g., public spending increases, tax cuts, budget deficit increases) followed by restrictive fiscal policies in post-election periods. In fact, excessive public spending and tax reductions are basic mechanisms generating political budget cycles. Each one is associated with different social welfare implications (Alesina 1987). Reducing taxes tends to be of practical concern to relatively wealthier people, while policies leveraging public spending—usually materializing through increases in transfer payments—favor the less well-off (Alesina 1988). Earlier studies support the hypothesis that budget cycles are stronger on the expenditure side (e.g., Rogoff 1990; Alt and Rose 2007), although there is some evidence on the existence of budget cycles on the revenue side (e.g., Katsimi and Sarantides 2012). The final outcome of this political and economic behavior is the creation of budget cycles which are neither due to exogenous shocks nor due to policy failures. Instead, they are deliberately generated in the context of governments' opportunistic efforts to maximize their re-election chances (Philips 2016). The success of this practice hinges on the voters' myopic perspective and the ability of incumbent politicians to exploit their temporary information advantage[2].

Political budget cycles have been the subject of empirical research documenting how the ability of governments to create political-economic cycles becomes limited as (i) the level of socioeconomic development rises (Block 2002; Shi and Svensson 2006; Klomp and de Haan 2013), (ii) the quality of institutions improves, and (iii) the transparency of its political process increases (Persson and Tabellini 2005; De Haan and Klomp 2013; Veiga et al. 2017). In other words, this happens when the quality and duration (or maturity) of democracy in a given country increase (Shi and Svensson 2003; Brender and Drazen 2005; De Haan 2013). In particular, the quality of institutions, and especially the existence of effective mechanisms controlling the financial decisions of governments (checks, balances, and fiscal rules), seem to be the most important means of limiting the extent of political budget cycles, as institutions shape the choices of the electorate and determine the incentives as well as the opportunities for incumbents to resort to these cycles (Chang 2008; Streb and Torrens 2013; Gootjes et al. 2021)[3].

Therefore, the main conclusion of the existing works is that political budget cycles are more evident in developing economies (Brender and Drazen 2005), being negligible (or very limited) in wealthier countries (Andrikopoulos et al. 2004; Shi and Svensson 2006; Mandon and Cazals 2019). Our analysis focuses on Greece, a European country representing a remarkable deviation from the aforementioned rule. Chortareas et al. (2018) have documented the occurrence of opportunistic budgetary policies in this country, despite mixed evidence on their effectiveness in maximizing re-election prospects (Brender and Drazen 2008; Aidt et al. 2011; De Haan and Klomp 2013). In a more recent work, we have documented the significance, direction and size of political budget cycles in Greece (Petrakos et al. 2021a). This process is justified with the limited quality of the country's institutions (Afonso et al. 2015) and lack of internal/external controls (Trantidis 2016), as well as with the fact that its governments often pursue clientele policies (Mitsopoulos and Pelagidis 2011) and resort to populism (Christodoulakis 2019).

This paper serves three main objectives. The first is to investigate the existence of political budget cycles in Greece in the Third Hellenic Republic (in Greek, 'Metapolitefsi', hereafter THR) since 1974. For this purpose, we specified econometric models in which the actual budget balance (as the percent share of Gross Domestic product, GDP) is used as the dependent variable, with the aim of quantifying the effects of the electoral cycle on the formation of budget balances in Greece. The second objective is to examine whether general (or parliamentary) elections affect a country's GDP, and whether the effects are stabilizing or destabilizing. More specifically, the question is whether general elections have statistically significant (stabilizing or destabilizing) effects on GDP. The third objective is to examine the possible impact of the external sector of the economy (i.e., external trade) on budget deficits. Based on these objectives, the structure of the paper is organized as follows. Section 2 introduces empirical data and presents the empirical methodology. Section 3 illustrates our findings in full detail and discusses the main results of this study, and Section 4 concludes the paper with some remarks for future research.

2. Methodology

2.1. Study Area

Considering Greece, a Southern European country, as the study area, our empirical analysis tests the intensity of political economic cycles and, particularly, political budget cycles over the THR since 1974, and more specifically, during a time interval encompassing more than four decades (1974–2020). This relatively long time horizon is considered representative of sequential economic downturns (Salvia et al. 2020), from the democratic recovery after the colonel's dictatorship (1967–1973) in the hand of the center-right 'Nea

Dimokratia' party of K. Karamanlis and G. Rallis (1974–1981), to intense economic development under the government of the socialist party of A. Papandreou (1981–1989). After a relatively short dominance of a center-right government with K. Mitsotakis (1990–1993), the socialist party of A. Papandreou won the 1993 elections, taking the lead continuously up to 2004 with three governments of K. Simitis (1996–2004), after the premature death of A. Papandreou (1996). A moderate economic growth—mostly characterized by rising expenditure in public infrastructure expanding the public debt significantly—was the main trait of this political phase. The subsequent center-right government of K. Karamanlis (2004–2009) consolidated the intense economic expansion following the 2004 Olympic Games celebrated in Athens, introducing some reforms and trying to manage the early signs of the imminent financial crisis (Vinci et al. 2022). The short center-left government of G. Papandreou (2009–2010) was unable to manage the drastic impact of the recession on Greek society (Tomao et al. 2021), and a more evident instability characterized the political life of the country for some years (2011–2012). Alternated center-right (A. Samaras) and center-left (A. Tsipras) governments was representative of the 2010s political course of Greece. While reducing the impact of the crisis, the center-right government—answering the pressing requests of the European Commission and International Monetary Fund—was forced to cut salaries and pensions (Salvati 2016), and introduced a sort of austerity regime in public expenditures causing urban poverty, among others (Gkartzios 2013; Rontos et al. 2016; Panori et al. 2019). The subsequent center-left government tried to alleviate the economic consequences of Troika's austerity, while fueling social spending, public infrastructure, and tourism (Salvati 2019).

2.2. Logical Framework and Scope of the Study

The existence of political budget cycles has been documented for the Greek economy, with specific reference to a shorter time period, namely, 1980–2017 (Petrakos et al. 2021a). We have also demonstrated how these cycles have been generated primarily on the expenditure side and not on the revenue side (Petrakos et al. 2021c). The present analysis differs substantially from the existing research in many aspects. For the first time in the literature, the effects of the external sector of the economy (exports, imports, trade balance) were examined as explanatory variables of the public budget deficit in the context of political budget cycles and prove to be significant. These outcomes support the 'twin deficit hypothesis' (Miller and Russek 1989; Cavallo 2005; Corsetti and Müller 2006; Kim and Roubini 2008; Algieri 2013; Badinger et al. 2017; Afonso et al. 2022). Relating budget deficits to trade deficits, these outcomes had important policy implications for economic competitiveness (e.g., Rontos et al. 2016; Di Feliciantonio et al. 2018; Benassi et al. 2022). Moreover, we explicitly examined the effects of unemployment on the evolution of public budget deficits, regarded as a significant policy issue (Salvati and Benassi 2021). We studied a relatively long time series, incorporating predictors that may evaluate the influence of several election rounds from 1974 to 2019.

Our second goal was to ascertain the possible existence of a relationship between election years and total real output (hereafter 'real GDP'). In other words, we investigated whether the rate of change of total real GDP (hereafter 'DTYGR1') during the election years differs from the corresponding rate of change of non-election years (Petrakos et al. 2021b). Our final objective is to examine the possible effects of the external trade (exports of goods and services, imports of goods and services, and balance of goods and services) on the state budget balance, testing in this way the hypothesis of a direct linkage of budget balances with economic competitiveness at large.

2.3. Variables and Indicators

To achieve our objectives, we constructed five econometric models using the following dependent (1) and explanatory (2–10) variables:

(1) The actual budget balance (ABB) of a general government as a percent of GDP, as defined and measured by Eurostat. The minus (−) sign corresponds to a deficit, while the

plus (+) sign corresponds to a surplus. We consider the actual budget balance as a percent share of GDP (instead of computing absolute monetary terms) for three main reasons: because (i) percent metrics provide a more reliable indicator of the relative magnitude of the actual budget balance; (ii) percentages remove the long-term effect of inflation on fiscal aggregates; (iii) the main fiscal policy condition of the EU member states, in order to avoid entering the Excessive Deficit Procedure (EDP) of the Stability and Growth Pact, is not to exceed 3% of their GDP.

(2) The one-year lag of the actual budget balance (ABB-1) of a general government as a percent of GDP, since the public balance might be 'compounded' in the sense that the budget deficit of the previous year might affect, to some extent, the deficit of the current year. In effect, the one-year lag of the dependent variable is used to control for the autoregressive, AR(1), component of the actual budget balance (Petrakos et al. 2021a).

(3) The two-year lag of the actual budget balance (ABB-2) of a general government as a percent of GDP, since not only the budget deficit of the previous year but also the deficit of the year before might affect the deficit of the current year. In effect, the two year lag of the dependent variable is used to capture the autoregressive, AR(2), component of the actual budget balance. We also considered the impact of a three-year lag of the actual budget, but our preliminary analysis suggests that this variable is largely insignificant.

(4) The growth rate of real total GDP (TYGR), as estimated by the World Bank.

(5) The change in the growth rate of total real GDP (DTYGR1), i.e., the difference between the growth rate of total real GDP of the current year and the previous one, i.e., $\mathrm{DTYGR1_i} = \mathrm{TYGR1_i} - \mathrm{TYGR1_{i-1}}$. This variable is introduced as it might not be the rates of change in total real GDP that cause pressures for expansionary fiscal policy as much as it is the variations (or fluctuations) of these rates from year to year.

(6) The unemployment rate (UNR) as it is measured by Eurostat and the Hellenic Statistical Authority (ELSTAT).

(7) Exports of goods and services as a percent of GDP (EXPO) as they are estimated by the World Bank (World Development Indicators).

(8) Imports of goods and services as a percent of GDP (IMPO) as they are estimated by the World Bank (World Development Indicators).

(9) Balance of goods and services as a percent of GDP (BAGS), estimated as the difference between exports of goods and services and imports of goods and services as a percent of GDP (BAGS = EXPO − IMPO).

(10) Election (ELE), a binary variable taking the value of 1 for (general) election years and 0 otherwise.

2.4. Statistical Analysis

We have specified five linear regression models with the actual budget balance as the dependent variable. In what follows (Section 3), results of these models were presented and commented in the order they are listed in Table 1, reporting the regression coefficients of each model in separated tables. All models were checked against the basic assumptions of Ordinary Least Square (OLS) model fitting, i.e., testing for linearity, homoscedasticity, independence, and normality assumptions (Ciommi et al. 2019). Results of statistical checks were presented in a graphical form (Lamonica et al. 2020).

Table 1. Specification of econometric models used in our analysis (Acronyms in Section 2.3).

Model	Dependent Variable	Predictors
1	Actual Budget Balance	ABB-1, TYGR, UNR, ELE
2	Actual Budget Balance	ABB-1, TYGR, EXPO, IMPO, ELE
3	Actual Budget Balance	ABB-1, ABB-2, TYGR, EXPO, IMPO, ELE
4	Actual Budget Balance	ABB-1, ABB-2, TYGR, BAGS, ELE
5	Actual Budget Balance	ABB-1, ABB-2, TYGR, BAGS, UNR, ELE

3. Results and Discussion

3.1. Outcomes of Model 1

Model 1 provided the simplest specification of the variability in the actual budget balances of Greece (1974–2020) using only four variables. Since the lagged dependent variable is included in the explanatory variables, residuals are independently distributed (Salvati 2022). Plotting the residuals versus the fit, we verify that the linearity and homoscedasticity assumptions hold (Figure 1). Finally, a QQ plot and histogram of the residuals reveal no serious departures from normality (Ciommi et al. 2018). ABB-1 and UNR are significant at $\alpha = 0.01$, ELE is significant at $\alpha = 0.05$, and TYGR is significant at $\alpha = 0.1$. Low Variance Inflation Factors (VIF) values indicate no evidence of multicollinearity (Table 2). The proportion of explained variance is relatively high (adjusted-$R^2 = 0.70$).

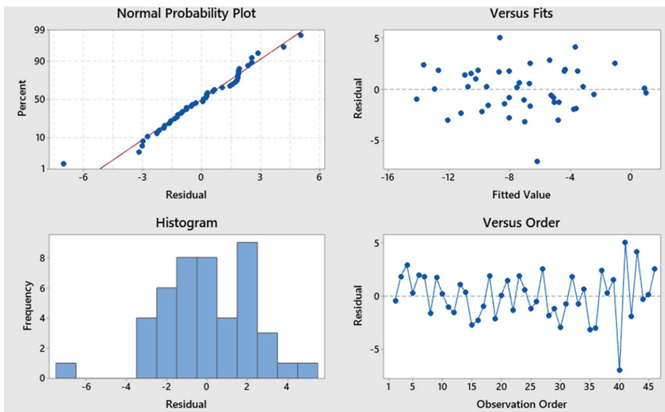

Figure 1. Residual plots for Model 1 with Actual Budget Balance as dependent variable.

Table 2. Regression coefficients of Model 1 (see Table 1 and Section 2 for model's description).

Variable	Coefficient	Standard Error	*t*-Value	*p*-Value	VIF
Constant	−3.18	1.170	−2.71	0.010	-
Budget Balance, previous year	0.75	0.098	8.02	0.000	1.22
GDP growth	0.22	0.122	1.81	0.078	1.52
Unemployment	0.16	0.057	2.83	0.007	1.26
Election year	−1.84	0.755	−2.44	0.019	1.02

3.2. Outcomes of Model 2

EXPO and IMPO were introduced in Model 2 together with the predictors already considered in Model 1. Residual plots in Figure 2 document how results of Model 2 fully adhere to linearity, homoscedasticity, independence, and normality assumptions. ABB-1, TYGR, and EXPO are significant at $\alpha = 0.01$, while ELE and IMPO are significant at $\alpha = 0.05$. VIF values indicate evidence of a residual multi-collinearity in Model 2, mainly attributable to IMPO and EXPO variables (Table 3). The proportion of explained variance is rather high (adjusted-$R^2 = 0.74$).

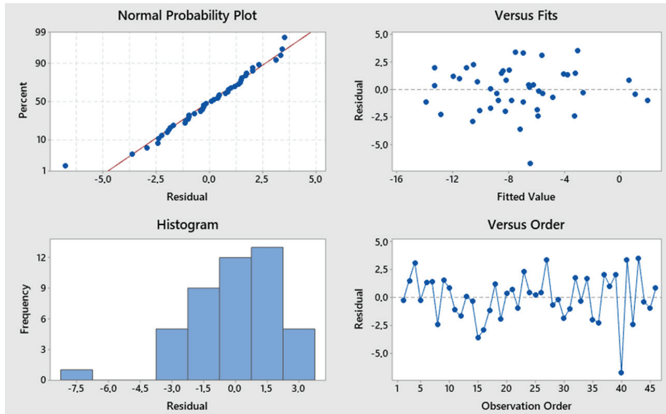

Figure 2. Residual plots for Model 2 with Actual Budget Balance as the dependent variable.

Table 3. Regression coefficients of Model 2 (see Table 1 and Section 2 for the model description).

Variable	Coefficient	Standard Error	*t*-Value	*p*-Value	VIF
Constant	−3.710	2.360	−1.57	0.125	-
Budget Balance, previous year	0.531	0.108	4.93	0.000	1.86
GDP growth	0.335	0.122	2.74	0.009	1.77
Exports	0.444	0.122	3.63	0.001	5.97
Imports	−0.312	0.140	−2.24	0.031	4.66
Election year	−1.733	0.708	−2.45	0.019	1.03

3.3. Outcomes of Model 3

The lagged variable ABB-2 was introduced as a predictor in Model 3 and accounted for a particularly high proportion of the explained variance (adjusted-R^2 = 0.75). The residual plots in Figure 3 indicate no deviations from the basic regression assumptions. EXPO was significant at α = 0.01, ELE, IMPO, TYGR, and ABB-1 were all significant at α = 0.05, and ABB-2 was significant at α = 0.1 (Table 4). A weak multi-collinearity issue (VIF > 5) was a distinctive characteristic of the estimation of Model 3, in line with what has been observed for the results of Model 2.

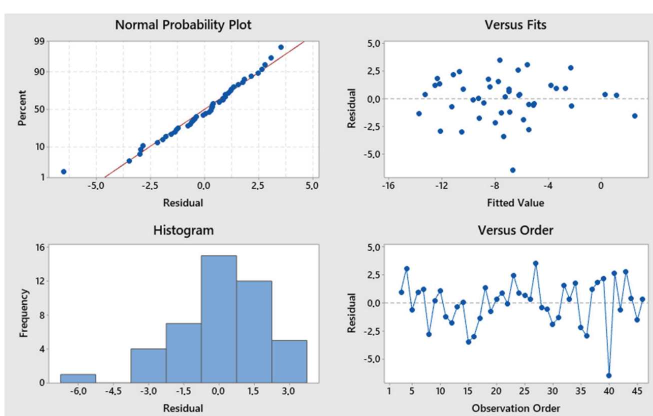

Figure 3. Residual plots for Model 3 with Actual Budget Balance as the dependent variable.

Table 4. Regression coefficients of Model 3 (see Table 1 and Section 2 for the model description).

Variable	Coefficient	Standard Error	*t*-Value	*p*-Value	VIF
Constant	−3.100	2.410	−1.28	0.207	-
Budget Balance, previous year	0.365	0.144	2.53	0.016	3.29
Budget Balance, two years before	0.228	0.130	1.76	0.087	2.52
GDP growth	0.313	0.121	2.58	0.014	1.71
Exports	0.458	0.121	3.78	0.001	5.88
Imports	−0.326	0.138	−2.36	0.023	4.51
Election year	−1.767	0.701	−2.52	0.016	1.03

3.4. Outcomes of Model 4

To cope with the weak multi-collinearity observed in Model 3, the variables IMPO and EXPO were replaced with their difference, named BAGS, in Model 4. The residual plots in Figure 4 indicate that the basic assumptions were not violated with this specification. The proportion of explained variance remained satisfactory (adjusted-R^2 = 0.73). ABB-1 and BAGS are significant at α = 0.01, ELE is significant at α = 0.05, TYGR is marginally significant at α = 0.05, and ABB-2 is marginally significant at α = 0.1 (Table 5). Low VIF values indicate no evidence of multicollinearity.

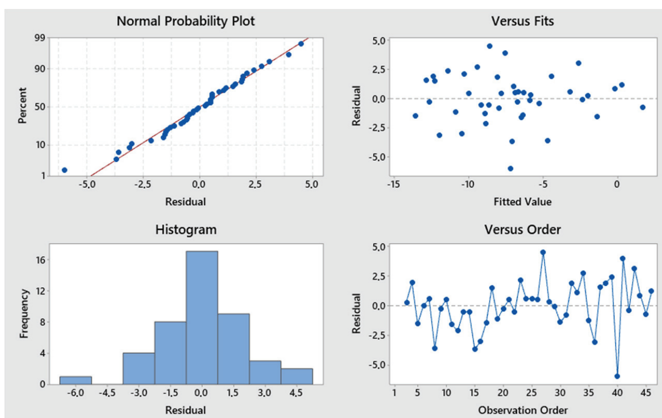

Figure 4. Residual plots for Model 4 with Actual Budget Balance as the dependent variable.

Table 5. Regression coefficients of Model 4 (see Table 1 and Section 2 for the model description).

Variable	Coefficient	Standard Error	*t*-Value	*p*-Value	VIF
Constant	1.003	0.979	1.02	0.312	-
Budget Balance, previous year	0.426	0.145	2.93	0.006	3.12
Budget Balance, two years before	0.224	0.134	1.68	0.102	2.52
GDP growth	0.236	0.118	2.01	0.052	1.51
Balance of goods and services	0.449	0.125	3.60	0.001	1.52
Election year	−1.724	0.723	−2.39	0.022	1.03

3.5. Outcomes of Model 5

Unemployment rate (UNR) was introduced as an additional predictor in Model 5. The proportion of explained variance was satisfactory (adjusted-R^2 = 0.75). Plots of residuals (Figure 5) indicate no violation of the basic econometric assumptions for Model 5. ABB-1 is significant at α = 0.01, ELE, ABB-2, TYGR, BAGS, are all significant at α = 0.05, and UNR is significant at α = 0.1. Low VIF values indicate no evidence of multicollinearity (Table 6).

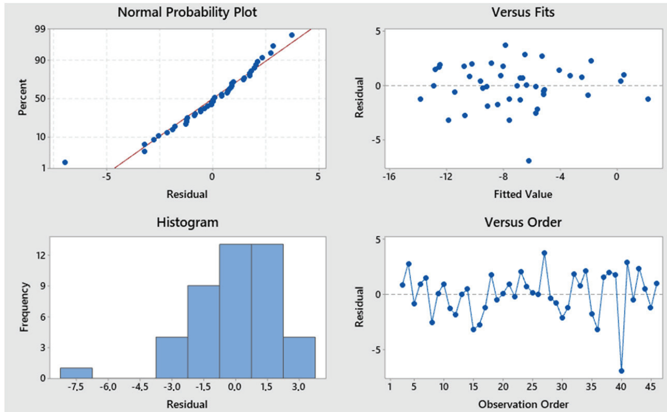

Figure 5. Residual plots for Model 5 with Actual Budget Balance as the dependent variable.

Table 6. Regression coefficients of Model 5 (see Table 1 and Section 2 for the model description).

Variable	Coefficient	Standard Error	*t*-Value	*p*-Value	VIF
Constant	−0.920	1.450	−0.63	0.532	-
Budget Balance, previous year	0.417	0.141	2.95	0.005	3.12
Budget Balance, two years before	0.290	0.136	2.14	0.039	2.73
GDP growth	0.274	0.117	2.35	0.024	1.56
Balance of goods and services	0.304	0.147	2.06	0.046	2.23
Unemployment rate	0.117	0.067	1.75	0.089	1.96
Election year	−1.744	0.704	−2.48	0.018	1.03

In order to investigate possible differences in the mean growth rate (%) of total (real) GDP between the two ELE groups (i.e., election and non-election years), the null hypothesis H_0: $\mu_1 - \mu_2 = 0$ was tested against the alternative H_1: $\mu_1 - \mu_2 \neq 0$, where μ_1 = E(DTYGR1 | ELE = 0) and μ_2 = E(DTYGR1 | ELE = 1). H_0 is rejected at $\alpha = 0.05$, and the differences in the means of GDP growth are significant. The results of our analysis were presented in Table 7 and Figure 6.

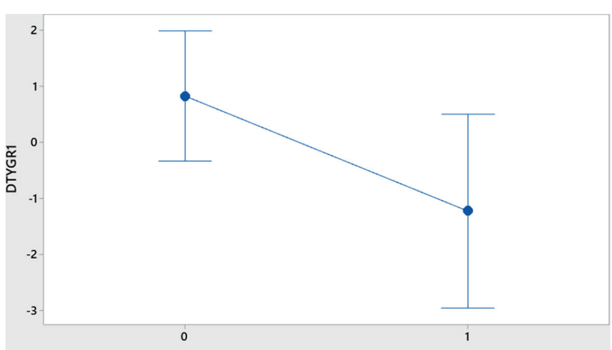

Figure 6. Interval Plot (95% Confidence Interval around the mean) of changes over time in the growth rate (%) of total (real) Gross Domestic Product (DTYGR1) against election year (1) or non-election year (0); pooled standard deviation (3.20) was used to calculate the intervals.

Table 7. Descriptive statistics of the growth rate (%) of total (real) Gross Domestic Product (DTYGR1) by election year.

Year	No. Years	Mean	Standard Deviation	95% Confidence Interval
No election	31	0.823	3.620	(−0.338; 1.983)
Election	14	−1.229	1.929	(−2.956; 0.499)

3.6. Discussing the Outcomes of Econometric Models (1 to 5)

The models illustrated above have a good fit as the adjusted R^2 was systematically above 0.7. In all of them, the actual budget balances (which correspond to public deficits in the case of Greece) were strongly correlated to their one-year lagged values (the coefficient of ABB-1 is 0.75 and 0.53, respectively, in Models 1 and 2). The impact of ABB-1 was reduced slightly when considering the two-year lagged variable ABB-2 (ABB-1 coefficient ranged between 0.37 in Model 3 and 0.43 in Model 4, while ABB-2 coefficient ranged between 0.22 in Model 4 and 0.29 in Model 5). Hence, we conclude that public deficits were persistent in Greece, adjusting slowly over longer periods of time that could last up to two years.

Regarding the impact of the external sector of the economy (EXPO, IMPO, BAGS), exports of goods and services tend to reduce the budget deficit, while imports seem to increase it (Models 2 and 3). Moreover, the balance of goods and services (BAGS = EXPO − IMPO) tends to reduce the budget deficit (Models 4 and 5). This result is a strong indication for the presence of 'twin deficits' in the case of the Greek economy. In particular, according to the 'twin deficits hypothesis', large and growing budget deficits are reflected in the widening of the current account deficit, leading to declining economic policy credibility, macroeconomic imbalances, and a slowdown in economic growth. Although making use of different approaches, these findings are in line with earlier literature testing this hypothesis for Greece (Vamvoukas 1999; Pantelidis et al. 2009; Kalou and Paleologou 2012; Magazzino 2012; Piersanti 2000; Forte and Magazzino 2013; Panousis and Koukouritakis 2020; Katrakilidis and Trachanas 2011; Paparas et al. 2016; Kosteletou 2013). Few additional studies ended up with more mixed, and sometimes contrasting, results (Algieri 2013; Papadogonas and Stournaras 2006).

The results of all econometric models presented in our study demonstrate how the growth rate of real GDP positively affects the actual budget balance, reducing budget deficits. The coefficient of GDP growth rate varied between 0.335 (Model 2) and 0.221 (Model 1). This finding implies that during times of economic slowdown, Greek governments tend to respond with expansionary fiscal policies. Unemployment also affected the dependent variable in a positive way, although with a relatively low coefficient (0.162 in Model 1 and 0.117 in Model 5). This correlation appears counterintuitive at first (*sensu* Rontos et al. 2019). However, economic dynamics in the 2010s may justify this outcome in the case of Greece, since the country was under budgetary supervision by the European Institutions ('Troika memoranda'). In that decade, budget deficits were strictly controlled, and unemployment skyrocketed to unprecedented levels, fueling income inequalities and rising social segregation, especially in urban areas (Gavalas et al. 2014; Di Feliciantonio and Salvati 2015; Panori et al. 2019; Rontos et al. 2016; Salvati 2016, 2018).

Election years (ELE) were found statistically significant in all econometric models. In other words, the political cycle (general parliamentary elections) in Greece was demonstrated to significantly affect the actual budget balance as a percent share of GDP. In all models, the ELE coefficient was higher than −1.5, ranging between −1.72 (Model 4) to −1.77 (Model 3). That is, in the years of general (or national) elections, budget deficits increased by more than 1.5% of GDP. More precisely, the increase was as high as 1.84% of GDP according to Model 1, 1.73%, 1.77%, 1.72% and 1.74% of GDP, respectively, according to Models 2, 3, 4, and 5. This effect is disproportionately high in comparison with other developed economies, where the effects of budget cycles have been estimated to be well below 1% of GDP (Andrikopoulos et al. 2004) and, in most cases, insignificant (Mandon and Cazals 2019).

We finally investigated the relationship between general elections (ELE) and changes over time in the growth rate of real GDP (DTYGR1) between 1974 and 2019. Descriptive statistics indicate that in non-election years, economic growth rates amounted to 0.82%, on average. On the contrary, in election years, the value of economic output decreased at a rate of 1.23% on average. This evidence suggests how the electoral cycle in Greece has serious destabilizing effects on total real output.

4. Concluding Remarks

The empirical results of our study demonstrate how severe political budget cycles have characterized Greece since 1974. These persistent political budget (or fiscal) cycles have contributed to the country's public debt. Moreover, the political budget cycles in Greece have played a destabilizing or pro-cyclical role, in terms of their effects on the country's real GDP. This is mainly due to the fact that the cycles materialize through an 'unproductive' public expenditures' increase, especially a rise in social transfers, during the election years (Petrakos et al. 2021d). It is therefore crucial to restrict political budget cycles, basically to improve the country's public finances and to stabilize the Greek economy—considering that the country had lost 30% of its real GDP between 2007 and 2020. Moreover, at the end of 2021, the gross debt of the general government in Greece had skyrocketed to 209% of its GDP, a magnitude more than two times higher than the corresponding average of the Eurozone (102% of GDP).

In addition, the existence of large political budget cycles suggests that certain socio-political characteristics in Greece—such as low-quality political institutions and insufficient control mechanisms or ineffective checks and balances—strongly resemble those of developing countries. These characteristics are basically the result of low-quality governance. Consequently, the effects of the electoral cycles cannot be ignored by any systematic scrutiny of the reforms needed to boost economic growth and to reduce budget deficits and public debt. Short-term performances of the external sector of the economy, namely, the balance of goods and services, were finally documented to affect the budget balance in Greece. Reduced external sector deficits and, more generally, a reduction in the balance of payments deficits, may contribute to the handling of budget deficits. Since the improvement of the trade balance is the result of the rising competitiveness of a given economic system, this should be a long-term objective of economic policy pursued through the implementation of specific strategies. Such issues are arguably among the most serious structural problems of the Greek economy and, as such, call for an effective and immediate policy response.

Author Contributions: Conceptualization, G.P. and I.V.; methodology, C.V. and K.R.; software, C.V. and L.S.; validation, I.V. and G.P.; formal analysis, C.V. and K.R.; investigation, L.S.; resources, I.V.; data curation, K.R.; writing—original draft preparation, I.V. and C.V.; writing—review and editing, K.R. and L.S.; visualization, G.P. and C.V.; supervision, I.V.; project administration, I.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

Notes

1. See Eurostat, online data code: GOV_10DD_EDPT1, update: 22 February 2022.
2. See Downs (1957); Nordhaus (1975); Rogoff and Sibert (1988); Rogoff (1990); Alesina and Perotti (1995); Persson and Tabellini (2000); Eslava (2011); De Haan (2013); Bonfiglioli and Gancia (2013). More recent contributions include the empirical studies by Dubois (2016); Mandon and Cazals (2019); Bohn (2019); Garcia and Hayo (2021); Gootjes et al. (2021).
3. Another factor that may influence the occurrence of political budget cycles is uncertainty regarding the outcome of the forthcoming elections, given that if governments are very confident of their re-election possibilities, they have limited motivation or no motivation at all to resort to them (Schultz 1995; Alt and Rose 2007; Hanusch and Magleby 2014; Eibl and Lynge-Mangueira 2017).

References

Afonso, António, Florence Huart, João Tovar Jalles, and Piotr Stanek. 2022. Twin deficits revisited: A role for fiscal institutions? *Journal of International Money and Finance* 121: 102506. [CrossRef]

Afonso, António, Sotirios Zartaloudis, and Yannis Papadopoulos. 2015. How party linkages shape austerity politics: Clientelism and fiscal adjustment in Greece and Portugal during the eurozone crisis. *Journal of European Public Policy* 22: 315–34. [CrossRef]

Aidt, Toke S., Francisco José Veiga, and Linda Gonçalves Veiga. 2011. Election results and opportunistic policies: A new test of the rational political business cycle model. *Public Choice* 148: 21–44. [CrossRef]

Alesina, Alberto. 1987. Macroeconomic policy in a two-party system as a repeated game. *Quarterly Journal of Economics* 102: 651–78. [CrossRef]

Alesina, Alberto. 1988. Credibility and policy convergence in a two-party system with rational voters. *American Economic Review* 78: 796–805.

Alesina, Alberto, and Roberto Perotti. 1995. The political economy of budget deficits. *International Monetary Fund Staff Papers* 42: 1–31. [CrossRef]

Algieri, Bernardina. 2013. An empirical analysis of the nexus between external balance and government budget balance: The case of the GIIPS countries. *Economic Systems* 37: 233–53. [CrossRef]

Alt, James E., and Shanna S. Rose. 2007. Context-conditional political business cycles. In *Oxford Handbook of Comparative Politics*. Edited by Carles Boix and Susan C. Stokes. Oxford: Oxford University Press, pp. 845–67.

Andrikopoulos, Andreas, Ioannis Loizides, and Kyprianos Prodromidis. 2004. Fiscal policy and political business cycles in the EU. *European Journal of Political Economy* 20: 125–52. [CrossRef]

Badinger, Harald, Aurélien Fichet de Clairfontaine, and Wolf Heinrich Reuter. 2017. Fiscal rules and twin deficits: The link between fiscal and external balances. *World Economy* 40: 21–35. [CrossRef]

Benassi, Federico, Alessia Naccarato, Ricardo Iglesias-Pascual, Luca Salvati, and Salvatore Strozza. 2022. Measuring residential segregation in multi-ethnic and unequal European cities. *International Migration*. [CrossRef]

Block, Steven. A. 2002. Political business cycles, democratization, and economic reform: The case of Africa. *Journal of Development Economics* 67: 205–28. [CrossRef]

Bohn, Frank. 2019. Political budget cycles, incumbency advantage, and propaganda. *Economics and Politics* 31: 43–70. [CrossRef]

Bonfiglioli, Alessandra, and Gino Gancia. 2013. Uncertainty, electoral incentives and political myopia. *The Economic Journal* 123: 373–400. [CrossRef]

Brender, Adi, and Allan Drazen. 2005. Political budget cycles in new versus established democracies. *Journal of Monetary Economics* 52: 1271–95. [CrossRef]

Brender, Adi, and Allan Drazen. 2008. How do budget deficits and economic growth affect reelection prospects? Evidence from a large panel of countries. *American Economic Review* 98: 2203–20. [CrossRef]

Cavallo, Michele. 2005. *Understanding the Twin Deficits: New Approaches, New Results*. FRBSF Economic Letter. San Francisco: Federal Reserve Bank of San Francisco.

Chang, Eric C. C. 2008. Electoral incentives and budgetary spending: Rethinking the role of political institutions. *The Journal of Politics* 70: 1086–97. [CrossRef]

Chortareas, Georgios, Vassilis E. Logothetis, and Andreas A. Papandreou. 2018. Elections and opportunistic budgetary policies in Greece. *Managerial and Decision Economics* 40: 854–62. [CrossRef]

Christodoulakis, Nicos. 2019. *Greece and the Euro Area: The Cost of Weak Institutions*. CGK Working Paper No. 2019-05. Medford: The Fletcher School, Tufts University, May.

Ciommi, Mariateresa, Francesco M. Chelli, and Luca Salvati. 2019. Integrating parametric and non-parametric multivariate analysis of urban growth and commuting patterns in a European metropolitan area. *Quality & Quantity* 53: 957–79.

Ciommi, Mariateresa, Francesco M. Chelli, Margherita Carlucci, and Luca Salvati. 2018. Urban growth and demographic dynamics in southern Europe: Toward a new statistical approach to regional science. *Sustainability* 10: 2765. [CrossRef]

Corsetti, Giancarlo, and Gernot J. Müller. 2006. Twin deficits: Squaring theory, evidence and common sense. *Economic Policy* 21: 598–638. [CrossRef]

De Haan, Jakob. 2013. Democracy, elections and government budget deficits. *German Economic Review* 15: 131–42. [CrossRef]

De Haan, Jakob, and Jeroen Klomp. 2013. Political budget cycles and election outcomes. *Public Choice* 157: 245–67.

Di Feliciantonio, Cesare, and Luca Salvati. 2015. 'Southern' Alternatives of Urban Diffusion: Investigating Settlement Characteristics and Socio-Economic Patterns in Three Mediterranean Regions. *Tijdschrift voor Economische en Sociale Geografie* 106: 453–70. [CrossRef]

Di Feliciantonio, Cesare, Luca Salvati, Efthymia Sarantakou, and Kostas Rontos. 2018. Class diversification, economic growth and urban sprawl: Evidences from a pre-crisis European city. *Quality & Quantity* 52: 1501–22.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.

Dubois, Eric. 2016. Political business cycles 40 years after Nordhaus. *Public Choice* 166: 235–59. [CrossRef]

Eibl, Ferdinand, and Halfdan Lynge-Mangueira. 2017. Electoral confidence and political budget cycles in non-OECD countries. *Studies in Comparative International Development* 52: 45–63. [CrossRef]

Eslava, Marcela. 2011. The political economy of fiscal deficits: A survey. *Journal of Economic Surveys* 25: 645–73. [CrossRef]

Forte, Francesco, and Cosimo Magazzino. 2013. Twin deficits in the European countries. *International Advances in Economic Research* 19: 289–310. [CrossRef]
Garcia, Israel, and Bernd Hayo. 2021. Political budget cycles revisited: Testing the signaling process. *European Journal of Political Economy* 69: 1–18. [CrossRef]
Gavalas, Vasilis S., Kostas Rontos, and Luca Salvati. 2014. Who becomes an unwed mother in Greece? Sociodemographic and geographical aspects of an emerging phenomenon. *Population, Space and Place* 20: 250–63. [CrossRef]
Gkartzios, Menelaos. 2013. 'Leaving Athens': Narratives of counterurbanisation in times of crisis. *Journal of Rural Studies* 32: 158–67. [CrossRef]
Gootjes, Bram, Jakob de Haan, and Richard Jong-A-Pin. 2021. Do fiscal rules constrain political budget cycles? *Public Choice* 188: 1–30. [CrossRef]
Hanusch, Marek, and Daniel B. Magleby. 2014. Popularity, polarization, and political budget cycles. *Public Choice* 159: 457–67. [CrossRef]
Kalou, Sofia, and Suzanna-Maria Paleologou. 2012. The twin deficits hypothesis: Revisiting an EMU country. *Journal of Policy Modeling* 34: 230–41. [CrossRef]
Katrakilidis, Constantinos P., and Emmanouil Trachanas. 2011. Has the accession of Greece in the EU influenced the dynamics of the country's "twin deficits"? An empirical investigation. *European Research Studies Journal* 14: 45–54.
Katsimi, Margarita, and Vassilis Sarantides. 2012. Do elections affect the composition of fiscal policy in developed, established democracies? *Public Choice* 151: 352–62. [CrossRef]
Kim, Soyoung, and Nouriel Roubini. 2008. Twin deficit or twin divergence? Fiscal policy, current account, and real exchange rate in the U.S. *Journal of International Economics* 74: 362–83. [CrossRef]
Klomp, Jeroen, and Jakob de Haan. 2013. Do political budget cycles really exist? *Applied Economics* 45: 329–41. [CrossRef]
Kosteletou, Nikolina E. 2013. Financial integration, euro and the twin deficits of southern eurozone countries. *Panoeconomicus* 60: 161–78. [CrossRef]
Lamonica, Giuseppe R., Maria C. Recchioni, Francesco M. Chelli, and Luca Salvati. 2020. The efficiency of the cross-entropy method when estimating the technical coefficients of input–output tables. *Spatial Economic Analysis* 15: 62–91. [CrossRef]
Magazzino, Cosimo. 2012. Fiscal policy, consumption and current account in the European countries. *Economics Bulletin* 32: 1330–44.
Mandon, Pierre, and Antoine Cazals. 2019. Political budget cycles: Manipulation by leaders versus manipulation by researchers? Evidence from a meta-regression analysis. *Journal of Economic Surveys* 33: 274–308. [CrossRef]
Miller, Stephen M., and Frank S. Russek. 1989. Are the twin deficits really related? *Contemporary Economic Policy* 7: 91–115. [CrossRef]
Mitsopoulos, Michaele, and Theodore Pelagidis. 2011. *Understanding the Crisis in Greece: From Boom to Bust*. Basingstoke: Palgrave Macmillan.
Nordhaus, William D. 1975. The political business cycle. *Review of Economic Studies* 42: 169–90. [CrossRef]
Panori, Anastasia, Yannis Psycharis, and Dimitris Ballas. 2019. Spatial segregation and migration in the city of Athens: Investigating the evolution of urban socio-spatial immigrant structures. *Population, Space and Place* 25: e2209. [CrossRef]
Panousis, Konstantinos P., and Minoas Koukouritakis. 2020. Twin deficits: Evidence from Portugal, Italy, Spain and Greece. *Intereconomics* 55: 332–38. [CrossRef]
Pantelidis, Panagiotis, Emmanouil Trachanas, and Athanasios L. Athanasenas. 2009. On the dynamics of the Greek twin deficits: Empirical evidence over the period 1960–2007. *International Journal of Economic Sciences and Applied Research* 2: 9–32.
Papadogonas, Theodore, and Yannis Stournaras. 2006. Twin deficits and financial integration in EU member-states. *Journal of Policy Modeling* 28: 595–602. [CrossRef]
Paparas, Dimitrios, Christian Richter, and Hairong Mu. 2016. An econometric analysis of the twin deficits hypothesis in Greece during the period 1960–2014. *Applied Economics Quarterly* 62: 341–60. [CrossRef]
Persson, Torsten, and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge: The MIT Press.
Persson, Torsten, and Guido Tabellini. 2005. *The Economic Effects of Constitutions*. Cambridge: The MIT Press.
Petrakos, George, Kostantinos Rontos, Chara Vavoura, and Ioannis Vavouras. 2021a. The mechanism of political budget cycles in Greece. Paper presented at 2020 International Conference on Applied Economics (ICOAE 2020): Advances in Longitudinal Data Methods in Applied Economic Research, Virtual, July 2–3; Springer Proceedings in Business and Economics. Edited by Nicholas Tsounis and Aspasia Vlachei. Berlin and Heidelberg: Springer, pp. 123–33.
Petrakos, George, Konstantinos Rontos, Chara Vavoura, and Ioannis Vavouras. 2021b. The destabilizing effects of political budget cycles: The case of Greece. Paper presented at the International Conference on Applied Economics (ICOAE): Advances in Quantitative Economic Research, Heraklion Crete, Greece, June 30; Springer Proceedings in Business and Economics. Berlin: Springer.
Petrakos, George, Konstantinos Rontos, Chara Vavoura, Luca Salvati, and Ioannis Vavouras. 2021c. Political budget cycles and the effects of the Excessive Deficit Procedure: The case of Greece. *Regional Statistics* 11: 1–20. [CrossRef]
Petrakos, George, Konstantinos Rontos, Chara Vavoura, Luca Salvati, and Ioannis Vavouras. 2021d. Toward a political budget cycle? Unveiling long-term latent paths in Greece. *Quality and Quantity*. [CrossRef]
Philips, Andrew Q. 2016. Seeing the forest through the trees: A meta-analysis of political budget cycles. *Public Choice* 168: 313–41. [CrossRef]
Piersanti, Giovanni. 2000. Current account dynamics and expected future budget deficits: Some international evidence. *Journal of international Money and Finance* 19: 255–71. [CrossRef]

Rogoff, Kenneth. 1990. Equilibrium political budget cycles. *American Economic Review* 80: 21–36.

Rogoff, Kenneth, and Anne Sibert. 1988. Elections and macro-economic policy cycles. *Review of Economic Studies* 55: 1–16. [CrossRef]

Rontos, Kostas, Efstathios Grigoriadis, Adele Sateriano, Maria Syrmali, Ioannis Vavouras, and Luca Salvati. 2016. Lost in protest, found in segregation: Divided cities in the light of the 2015 "Οχι" referendum in Greece. *City, Culture and Society* 7: 139–48. [CrossRef]

Rontos, Kostas, Ioannis Vavouras, Maria Teresa Ciommi, and Luca Salvati. 2019. Two faces of the same coin? A comparative, global approach to corruption and socioeconomic development. *Quality & Quantity* 53: 1875–94.

Salvati, Luca. 2016. The dark side of the crisis: Disparities in per capita income (2000–12) and the urban-rural gradient in Greece. *Tijdschrift voor Economische en Sociale Geografie* 107: 628–41. [CrossRef]

Salvati, Luca. 2018. Population growth and the economic crisis: Understanding latent patterns of change in Greece, 2002–16. *Letters in Spatial and Resource Sciences* 11: 105–26. [CrossRef]

Salvati, Luca. 2019. Bridging the divide: Demographic dynamics and urban–rural polarities during economic expansion and recession in Greece. *Population, Space and Place* 25: e2267. [CrossRef]

Salvati, Luca. 2022. Exploring long-term urban cycles with multivariate time-series analysis. *Environment and Planning B: Urban Analytics and City Science* 49: 1212–27. [CrossRef]

Salvati, Luca, and Federico Benassi. 2021. Rise (and decline) of European migrants in Greece: Exploring spatial determinants of residential mobility (1988–2017), with special focus on older ages. *Journal of International Migration and Integration* 22: 599–613. [CrossRef]

Salvia, Rosanna, Gianluca Egidi, Luca Salvati, Jesús Rodrigo-Comino, and Giovanni Quaranta. 2020. In-between 'Smart'urban growth and 'Sluggish'rural development? reframing population dynamics in Greece, 1940–2019. *Sustainability* 12: 6165. [CrossRef]

Schultz, Kenneth A. 1995. The politics of the political business cycle. *British Journal of Political Science* 25: 79–99. [CrossRef]

Shi, Min, and Jakob Svensson. 2003. Political budget cycles: A review of recent developments. *Nordic Journal of Political Economy* 29: 69–76.

Shi, Min, and Jakob Svensson. 2006. Political budget cycles: Do they differ across countries and why? *Journal of Public Economics* 90: 1367–89. [CrossRef]

Streb, Jorge M., and Gustavo Torrens. 2013. Making rules credible: Divided government and political budget cycles. *Public Choice* 156: 703–22. [CrossRef]

Tomao, Antonio, Giovanni Quaranta, Rosanna Salvia, Sabato Vinci, and Luca Salvati. 2021. Revisiting the 'southern mood'? Post-crisis Mediterranean urbanities between economic downturns and land-use change. *Land Use Policy* 111: 105740. [CrossRef]

Trantidis, Aris. 2016. *Clientelism and Economic Policy: Greece and the Crisis*. Advances in European Politics (Book 126). London: Routledge.

Vamvoukas, George A. 1999. The twin deficits phenomenon: Evidence from Greece. *Applied Economics* 31: 1093–100. [CrossRef]

Veiga, Francisco José, Linda Gonçalves Veiga, and Atsuyoshi Morozumi. 2017. Political budget cycles and media freedom. *Electoral Studies* 45: 88–99. [CrossRef]

Vinci, Sabato, Francesca Bartolacci, Rosanna Salvia, and Luca Salvati. 2022. Housing markets, the great crisis, and metropolitan gradients: Insights from Greece, 2000–14. *Socio-Economic Planning Sciences* 80: 101171. [CrossRef]

economies

Article

The Impact of Environmental Uncertainty on Accounting Information Relevance and Performance: A Contingency Approach

Rui Pires [1] and Maria-Ceu G. Alves [2,*]

[1] Departamento de Ciências Empresariais e Jurídicas, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal
[2] NECE-UBI Research Unit in Business Sciences, University of Beira Interior, 6201-001 Covilhã, Portugal
* Correspondence: mceu@ubi.pt

Abstract: This paper examines the association between environmental uncertainty, accounting information relevance, and organizational performance. From a contingency approach, this paper attempts to contribute to a stream of research that investigates the relationship between accounting information relevance and organizational performance. The presence of environmental uncertainty in this relationship has not been fully established. This paper contributes to this area by suggesting a framework to study and explain this connection. An online questionnaire-based survey was conducted, which produced 119 valid responses (a response rate of 23%) from large manufacturing companies operating in Portugal. The results suggest that in contexts of environmental uncertainty, the relevance of non-financial information increases. However, the relevance of financial information continues to outstrip that of non-financial information. The results also suggest that financial information and non-financial information are complementary, and not substitutes, and can be used simultaneously in different situations. These results have several implications for professionals involved in decision-making activities. It offers findings which are potentially useful for both theory and practice. The study addresses an identified gap in the literature and adds to the existing body of work analyzing the association between environmental uncertainty, accounting information relevance for decision-making purposes, and organizational performance.

Keywords: environmental uncertainty; financial information; non-financial information; organizational performance; contingency theory; survey; Portugal

Citation: Pires, Rui, and Maria-Ceu G. Alves. 2022. The Impact of Environmental Uncertainty on Accounting Information Relevance and Performance: A Contingency Approach. *Economies* 10: 211. https://doi.org/10.3390/economies10090211

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 1 July 2022
Accepted: 30 August 2022
Published: 2 September 2022

1. Introduction

In recent decades the organizational environment has changed, being more complex, turbulent, and unpredictable than in the past, posing new challenges to managers as they must make decisions in uncertain environments (Al-Mawali and Am 2016; Baines and Langfield-Smith 2003; Chenhall 2003; Chenhall and Langfield-Smith 1998a; Chong 1996; Otley 2016). Recently, in dealing with the unprecedented situation of the COVID-19 pandemic, decision makers have faced many types of environmental uncertainty; for example, uncertainty about the reliability of information flows (Lodge and Boin 2020), irrational reactions of financial markets, extreme volatility of the economy, etc. Making decisions in such an environment where information is incomplete and there are no correct and clear answers is unprecedented in its degree of uncertainty (Aon 2020). Nevertheless, decision makers must obtain proper information to make strategic decisions that would influence organizational competitive advantages and performance (Adeniran and Obembe 2020; Oyewo 2021).

Although several empirical studies have shown the influence of perceived environmental uncertainty on decision-making, few have examined the role of accounting information in explaining this association (Abu-Rahma and Jaleel 2019). In this context, accounting information becomes relevant, namely non-financial, external, and future-oriented information, to support organizational change and decision-making. For an organization to survive and operate with success, it is critical that managers have access to useful and timely information so they can act and make the best decisions (Al-Mawali and Am 2016; Alves 2017; Oyewo 2021; Thuan et al. 2022; Visedsun and Terdpaopong 2021).

In this study, accounting information is defined as the formalized financial and non-financial information (Massicotte and Henri 2021) provided on a regular basis for decision-making purposes. However, a major problem that managers face is the need for credible information to assist them in the decision-making process (Frazer 2020). Additionally, despite non-financial information having gained growing relevance (Czaja-Cieszyńska et al. 2021), it does not substitute for financial information, which is considered of little relevance in uncertain decision-making contexts, because it is too aggregated and available too late (Chenhall and Langfield-Smith 1998b; Johnson and Kaplan 1987), even if in some cases, financial information continues to have the confidence of managers in decision-making (Bhimani and Langfield-Smith 2007; Chow and Van der Stede 2006; Hyvönen 2007). Non-financial information is, therefore, complementary to financial information and these two types of information can be used simultaneously in different situations or for different purposes (Chenhall and Langfield-Smith 2007; Chow and Van der Stede 2006; Lau and Sholihin 2005; Massicotte and Henri 2021; Monteiro et al. 2021; Visedsun and Terdpaopong 2021).

Following the contingency approach, several authors argue that the fit between internal factors (e.g., organizational structure, management accounting systems and characteristics of accounting information) and external factors (e.g., environmental uncertainty and intensity of competition) leads to a better management and organizational performance (Al-Mawali and Am 2016; Baines and Langfield-Smith 2003; Boulianne 2007; Chenhall 2003; Chong 1996; Löfsten and Lindelöf 2005; Oyewo 2022; Turner et al. 2017). For example, some studies suggest that giving greater weight to non-financial and external information in a context of environmental uncertainty improves organizational performance (Al-Mawali and Am 2016; Hoque 2004, 2005; Hoque and James 2000). Thus, to promote organizational performance, it is critical to adjust the accounting information relevance to the level of environmental uncertainty.

In fact, some empirical studies report a direct and positive effect of environmental uncertainty on accounting information relevance/use for decision-making purposes (e.g., Al-Mawali and Am 2016; Cescon et al. 2019; Latan et al. 2018; Oyewo 2021, 2022). Furthermore, previous literature also shows a direct and positive effect of accounting information or management accounting practices on performance (e.g., Adeniran and Obembe 2020; Baines and Langfield-Smith 2003; Cadez and Guilding 2008; Hoque and James 2000; Latan et al. 2018; Turner et al. 2017; Visedsun and Terdpaopong 2021). Concerning moderation analysis, Hoque (2005) showed that environmental uncertainty moderates the relationship between the use of some non-financial performance measures for performance evaluation and organizational performance. Al-Mawali and Am (2016) concluded that environmental uncertainty moderates the relationship between customer accounting information use and organizational performance. However, to the best of our knowledge, no recent study examines, specifically, the association between environmental uncertainty, accounting information relevance, and organizational performance, or investigates the influence of environmental uncertainty on the relationship between non-financial information relevance for decision-making purposes and organizational performance.

Therefore, the main purpose of this paper is to analyze the association between environmental uncertainty, accounting information relevance in decision-making, and organizational performance. Furthermore, we intend to analyze how environmental uncertainty moderates the relationship between non-financial information relevance and organizational performance. In this way, this paper aims to answer to the following research questions: What is the association between environmental uncertainty, accounting information relevance, and organizational performance? How does environmental uncertainty influence the relationship between non-financial information relevance and organizational performance?

This study provides additional insights regarding the relevance of financial and non-financial information for decision-making purposes under uncertainty contexts and its influence on organizational performance, thus contributing to the literature in this research field. Hence, it extends the knowledge on which is the most suitable accounting information for certain situations. Moreover, our study also documents the moderating role of environmental uncertainty on the relationship between non-financial information relevance for decision-making purposes and organizational performance.

The remainder of the paper is structured as follows. In Section 2, we present a summary of the relevant literature and develop our research hypotheses. The research design of this study, including the sample, data collection, and variables measurement, is detailed in Section 3. Section 4 presents and discusses the results regarding the association between the degree of environmental uncertainty, accounting information relevance, and organizational performance. Finally, Section 5 provides the main conclusions, theoretical and practical implications, and limitations of this research, outlining, also, future research opportunities.

2. Theoretical Framework and Research Hypotheses

Organizational environment considers the set of physical and social factors external to organizations that influence their internal characteristics and are therefore considered in the decision-making process (Haldma and Lääts 2002; Löfsten and Lindelöf 2005). Factors such as globalization of operations, increased competition, technological changes, the demand for continuous improvement, and new demands on social and environmental responsibility cause constant changes that lead to increased environmental uncertainty (Latan et al. 2018; Löfsten and Lindelöf 2005; Mia and Clarke 1999; Newkirk and Lederer 2006; Otley 2016). Environmental uncertainty represents a challenge for every organization today and is related to the lack of information and the speed of information, which limit actions (Latan et al. 2018). Given its influence on organizational structure and systems, environmental uncertainty is one of all the variables used in the pioneer contingent studies that has gained the widest attention in management accounting research (Otley 2016).

Recently, in dealing with the unprecedented situation of the COVID-19 pandemic, many decision makers have faced environmental uncertainty, about the reliability of information flows (Lodge and Boin 2020), irrational reactions of financial markets, extreme volatility of the economy, etc. Making decisions in such an environment is unprecedented in its degree of environmental uncertainty (Aon 2020). In these contexts, the information needed by managers in decision-making increases (Baines and Langfield-Smith 2003; Chenhall and Langfield-Smith 1998a, 1998b; Chong 1996; Latan et al. 2018), given the increased unpredictability of future events (Chenhall and Morris 1986). In these conditions of high uncertainty, sophisticated accounting information can help managers enhance decision-making, providing some alternatives and solutions (Latan et al. 2018). On the other hand, environmental uncertainty makes it more necessary for managers to resort to strategic planning (Baines and Langfield-Smith 2003; Newkirk and Lederer 2006) and thus it encourages the introduction of action plans to respond to threats and opportunities (Mia and Clarke 1999).

Managers that face greater environmental uncertainty attach a greater utility to non-financial information, because they consider it more suitable (Boulianne 2007; Chenhall and Morris 1986; Hoque 2005; Hoque and James 2000; Lal and Hassel 1998; Monteiro et al. 2022). Traditional accounting information, mainly on financial and internal events, is too aggregated and inappropriate (Chenhall and Langfield-Smith 1998b; Johnson and Kaplan 1987) and is therefore inadequate and unhelpful in contexts of environmental uncertainty (Hayes 1977). Another limitation of this information is that it does not reflect the efficiency of organizations concerned with intangible factors such as quality, continuous improvement, and customer satisfaction (Baines and Langfield-Smith 2003). In these cases, decision makers need timely, non-financial, and external accounting information (e.g., on markets, customers, and competitors) (Afifa and Saleh 2021; Boulianne 2007; Chenhall and Morris 1986). This is the most relevant information for making better decisions (Baines and

Langfield-Smith 2003), as well as to focus management control on the strategic uncertainties of the organization (Vaivio 1999). It is within this context that the first research hypothesis is formulated.

Hypothesis 1. *The relevance attributed to non-financial information for decision-making is greater when environmental uncertainty is higher.*

As previously mentioned, traditional management accounting information, mostly internal and financial information, has lost some relevance for decision makers in the current circumstances of the organizational environment (Johnson and Kaplan 1987; Ma et al. 2022). Under such circumstances, non-financial, external, and timely accounting information has won the confidence of managers because it alerts them to new situations, allowing a better understanding and control of costs, and enables the achievement of competitive advantages (Chenhall and Langfield-Smith 1998a; Oyewo 2022; Vaivio 1999). Quantitative non-financial information, in particular, helps to attain organizational alignment by integrating the horizontal and vertical dimensions of performance (Bertolotti et al. 2019). Moreover, non-financial information gains relevance for assessing performance in various areas of the organization, such as processes and operations, human resources, customers, and corporate strategy (Baines and Langfield-Smith 2003; Bhimani and Langfield-Smith 2007; Chenhall and Langfield-Smith 2007). Managers know that good performance in these areas leads to good non-financial performance, which increases an organization's financial performance (Baines and Langfield-Smith 2003; Turner et al. 2017). For instance, Turner et al. (2017) have shown that customer performance positively influences financial performance.

However, non-financial information does not replace financial information, and in many cases financial information is more relevant than non-financial information (Bhimani and Langfield-Smith 2007; Chow and Van der Stede 2006; Hyvönen 2007; Massicotte and Henri 2021). According to Bhimani and Langfield-Smith (2007), in strategy development both financial and non-financial information is used, while in strategy implementation greater emphasis is placed on financial information. In turn, Massicotte and Henri (2021) report the use of financial and non-financial information to oversee strategy implementation. Therefore, several authors consider that the main purpose of non-financial information is to complement financial information (Chenhall and Langfield-Smith 2007; Chow and Van der Stede 2006; Lau and Sholihin 2005), which is not sufficient for decision-making purposes (Monteiro et al. 2021). Additionally, for companies' performance measurement, financial measures are usually used, while non-financial measures such as customer satisfaction and loyalty, and employee satisfaction, cannot be ignored (Visedsun and Terdpaopong 2021). Based on the literature review, the following two hypotheses are proposed.

Hypothesis 2. *Managers assign greater relevance to financial information for decision-making than to non-financial information.*

Hypothesis 3. *Managers believe that financial information and non-financial information are complementary.*

Contingency research in management accounting considers that several internal and external contingent factors influence the characteristics of management accounting systems and accounting information required for decision-making (e.g., Abdel-Kader and Luther 2008; Cescon et al. 2019; Chenhall and Morris 1986; Chong and Chong 1997; Haldma and Lääts 2002; Hoque 2005; Hoque and James 2000; Mia and Clarke 1999; Oyewo 2022; Turner et al. 2017). In fact, according to Otley (2016), contingency research seeks to discover when specific accounting information and management accounting practices might be most appropriate for organizations in their specific circumstances. Furthermore, several studies also conclude that the fit between contingent factors such as environmental uncertainty, technology, business strategy, and organizational structure, management accounting systems, and accounting information required for decision-making improves organizational performance (e.g., Al-Mawali and Am 2016; Baines and Langfield-Smith 2003; Boulianne 2007; Cadez and Guilding 2008; Chong 1996; Hoque 2005; Hoque and James 2000; Löfsten and Lindelöf 2005; Mia and Clarke 1999; Oyewo 2022; Turner et al. 2017). For instance, Al-Mawali and Am (2016) show that the fit between environmental uncertainty and customer accounting information use improves organizational performance. In contexts of environmental uncertainty, if higher relevance is attributed to non-financial information, organizational performance will also be greater (Hoque 2004, 2005; Hoque and James 2000). Quality non-financial information contributes to decision-making success which, in turn, is relevant for business success (Monteiro et al. 2022). Therefore, the following research hypothesis is formulated:

Hypothesis 4. *The match between environmental uncertainty and non-financial information relevance improves organizational performance.*

3. Research Design

The goals of this section are threefold: (i) to present the data collection instrument, (ii) to describe the sample and procedure of data collection, and (iii) to define the variables considered in the construction of the survey questionnaire.

3.1. Data Collection

A quantitative research approach was adopted to test the developed hypotheses. For the data collection, an online questionnaire-based survey was chosen, primarily because it could reach many respondents (Abdel-Kader and Luther 2008), in addition to enabling the respondent to answer how and where he/she wants, without feeling the pressure to respond immediately (Gillham 2008). The questionnaire also allowed us to collect the data needed to test the developed hypotheses in the previous section (Gillham 2008). Furthermore, with a questionnaire survey, it is possible to collect sufficient data to allow the generalization of results to the population being analyzed, which allows advances in contingency research that comes from the ascertainment of general patterns (Chenhall 2003).

As this instrument is not without limitations, to minimize them, the recommendations of Hill and Hill (2008) were followed. The questions were organized into sections to create a common thread between them and a pre-test was conducted. We sent to experts (i.e., three academics and five practitioners) a draft of the questionnaire for review and recommendation. The purposes of the pre-test were threefold: (i) to assess the adequacy of our questionnaire design, (ii) to verify whether the items captured the relevant dimensions of our variables, and (iii) to verify whether practitioners found the questionnaire items understandable and plausible. The feedback led to minor changes in the structure of the questionnaire and the wording of some individual items.

To implement the survey, we chose to send it by electronic mail, as in previous studies (e.g., Cadez and Guilding 2008; Cescon et al. 2019; Holm et al. 2016; Hyvönen 2007; Latan et al. 2018; Lau and Sholihin 2005; Massicotte and Henri 2021; Monteiro et al. 2022). A free software program available on the internet was used for this purpose. It does not require programming skills and is inexpensive to use (Fleming and Bowden 2009; Ganassali 2008). Compared with conventional questionnaire-based surveys (i.e., mail, telephone, and on-site), online questionnaire-based surveys have two main advantages: (i) low-cost administration, enabling large sample sizes, and (ii) speed and accuracy of data collection (Fleming and Bowden 2009). Moreover, Loomis and Paterson (2018) stress that there are no significant differences between the two data collection modes regarding response rate, item nonresponse, and nature of the data, which validates this approach.

3.2. Sample and Procedure

After finalizing the questionnaire, the target population of the study was selected. For this, we asked the National Statistics Institute (NSI) a list of the top 500 manufacturing companies, according to turnover, operating in Portugal. The choice of the larger manufacturing companies is justified by the fact that several studies state that larger companies have greater information needs for planning, control, and decision-making purposes (Abdel-Kader and Luther 2008; Alves 2010; Chenhall 2003; Chenhall and Langfield-Smith 1998a; Haldma and Lääts 2002). In addition, most of the studies reviewed had as targets the larger manufacturing companies (e.g., Baines and Langfield-Smith 2003; Cescon et al. 2019; Hoque and James 2000; Oyewo 2021; Visedsun and Terdpaopong 2021). Therefore, to be able to compare our results with those of these studies, it was considered important that the target population had similar characteristics.

Since we needed email addresses to send the questionnaire and collect the data, we updated the information provided by the NSI through a telephone call made to all companies. This led to the exclusion of nine companies[1]. After that, the questionnaire was sent to the person in charge of the financial department of 491 large manufacturing companies operating in Portugal that are the subjects of this study. Three months after the first mailing, a second mailing of questionnaires was carried out. Aware that the number of contacts, the persistence of the researcher, and the personalization of those contacts significantly affect the response rate to surveys (Ganassali 2008), along with the third mailing of questionnaires, we started making phone calls inviting financial managers to participate in the study. At the end of this process, 119 questionnaires were received. As there was a need to exclude five questionnaires because they were incorrectly completed (Gillham 2008), 114 usable responses were considered, which corresponds to a response rate of 23%. This response rate is above or in line with those of previous studies on management accounting (Baines and Langfield-Smith 2003; Cescon et al. 2019; Holm et al. 2016; Latan et al. 2018; Monteiro et al. 2022; Visedsun and Terdpaopong 2021).

Regarding the characteristics of the firms surveyed (turnover and number of employees) it was verified that 49% of the firms had turnover exceeding EUR 55 M and 57% had more than 250 employees (Table 1). Regarding the respondents, it appears that 50% were in an administrative/financial position.

Table 1. Characteristics of sample and respondents.

Description	Quantity (%)
Size (turnover EUR):	
Less than 25 M	13 (11.40%)
25 M to 35 M	21 (18.42%)
35 M to 55 M	24 (21.05%)
55 M to 90 M	22 (19.30%)
More than 90 M	34 (29.82%)
Size (number of employees):	
Less than 100	15 (13.16%)
100 to 249	34 (29.82%)
250 to 500	38 (33.33%)
More than 500	27 (23.68%)
Position of respondents:	
Managing/financial director	57 (50.00%)
Management controller	22 (19.30%)
Accountant	21 (18.42%)
Administrator/manager	7 (6.14%)
Other	7 (6.14%)

To examine possible differences between respondents and non-responding companies concerning turnover and number of employees, like Drury and Tayles (2006) and Guilding et al. (2000) we used the nonparametric test of Mann-Whitney. Regarding the turnover, some differences were found in central tendency. Through the analysis of the median, it was found that firms with higher turnover were also the ones that answered the questionnaire.

3.3. Variables Measurement

Some instruments developed in previous studies were used or adapted to measure the variables needed to test the research hypotheses developed in this study. Besides ensuring a greater reliability and validity of the information gathered, using instruments tested in other studies allows a more accurate comparison of results (Chenhall 2003). In this context, we will describe how the variables were operationalized and the instruments used to measure them.

The environmental uncertainty variable was measured based on an instrument developed by Teo and King (1997) and used later by Newkirk and Lederer (2006). Some adjustments were made in this instrument resulting from the studies of Chenhall (2003) and Löfsten and Lindelöf (2005). We added two items presented by Löfsten and Lindelöf (2005) (the intensity of research and development; and the legal, political, and economic constraints), and an item suggested by Chenhall (2003) (the requirements in terms of social and environmental responsibility) that recognizes social and environmental issues as relevant sources of environmental uncertainty (Latan et al. 2018). Respondents were requested to indicate their perception about the predictability of 15 items using a Likert-type scale ranging from 1 (strongly disagree) to 5 (strongly agree). Rating on these items was averaged to determine the environmental uncertainty index. Prior studies have used a similar approach to measure environmental uncertainty (e.g., Hoque 2005; Oyewo 2022).

In order to measure accounting information relevance (financial and non-financial), we used an instrument based on a set of items identified in the literature (Baines and Langfield-Smith 2003; Chow and Van der Stede 2006; Vaivio 1999). Comprising 25 items, this instrument considers nine items to measure financial information relevance (indicators of costs, results, profitability, and return on investment) and 16 items to measure non-financial information relevance (indicators related to processes and operations, employees, and suppliers and customers). Respondents were asked to indicate the relevance attributed to these items of financial and non-financial information to decision-making purposes using a Likert-type scale ranging from 1 (no impact) to 5 (very relevant). The respective scores were additively combined and averaged to derive indexes for financial information relevance and non-financial information relevance. A similar approach was used by Baines and Langfield-Smith (2003).

To measure the organizational performance variable, an instrument previously used by Hoque and James (2000) and modified by Cadez and Guilding (2008) was adopted. This instrument, which was used in several studies (e.g., Abernethy and Lillis 1995; Baines and Langfield-Smith 2003; Chenhall and Langfield-Smith 1998b; Oyewo 2022), consists in asking respondents to assess the organization's performance compared to that of competitors. This instrument considers seven[2] financial and non-financial dimensions and uses a Likert-type scale ranging from 1 (much worse than competitors) to 5 (much better than competitors). Rating on these dimensions was averaged to determine the organizational performance index for each company. Prior studies have used a similar approach to measure organizational performance (e.g., Cadez and Guilding 2008; Oyewo 2022).

Statistical analysis of the data collected was performed using the SPSS program, as in previous studies (e.g., Hoque 2005; Monteiro et al. 2021; Monteiro et al. 2022). We carried out univariate, bivariate, and multivariate analyses. In the next section, we proceed to the presentation and discussion of results.

4. Results and Discussion

The goals of this section are to describe the variables and analyze the association between environmental uncertainty, financial and non-financial information relevance in decision-making, and organizational performance. In addition, we analyze how the interaction between environmental uncertainty and non-financial information relevance for decision-making purposes influences organizational performance. That is, in this section we present a descriptive analysis of the variables and test the hypotheses developed in Section 2.

4.1. Descriptive Analysis

The reliability and descriptive statistics of the variables under study are presented in Table 2. As shown, and except for environmental uncertainty, the variables were established with all the items considered in the questionnaire. For the creation of the environmental uncertainty variable, 2 of the 15 items were excluded (the shortage of skilled labor and the shortage of materials) since Cronbach's alpha improves with this exclusion. We used Cronbach's alpha to assess the reliability, or internal consistency, and it appears that all the items had a good or very good internal consistency (Marôco 2021; Pestana and Gageiro 2014).

Table 2. Reliability and descriptive statistics of variables.

Variable	Number of Items Used	Cronbach's Alpha	Mean (n = 114)	Standard Deviation
Environmental uncertainty (EU)	13	0.79	3.31	0.53
Financial information (FI)	9	0.87	4.24	0.55
Non-financial information (NFI)	16	0.93	4.04	0.62
Process and operations (PO)	5	0.87	4.13	0.68
Employees (E)	5	0.90	3.75	0.76
External environment (EE)	6	0.88	4.21	0.67
Organizational performance (OP)	7	0.86	3.60	0.50

Regarding environmental uncertainty, the perception of respondents indicated that the level of uncertainty was moderate (3.31) and resulted essentially from the intense competition on prices, quality, and product differentiation, and the high level of social and environmental responsibility. Thus, it seems that factors related to intense competition drive environmental uncertainty (Baines and Langfield-Smith 2003; Mia and Clarke 1999). The same was true regarding the requirements in terms of social and environmental responsibility, as advocated by Chenhall (2003) and Latan et al. (2018), which were perceived by respondents as very high. In fact, most Portuguese companies operate in the highly competitive environment of the European Union (Monteiro et al. 2021). In addition, social and environmental issues have been also one of the biggest priorities of the European Union.

Concerning the relevance attributed to financial and non-financial information for decision-making purposes, the results indicate that it was significant (high or very high). The financial indicators considered most relevant were the overall results of the organization, production costs, and sales costs. In turn, the most relevant non-financial indicators were related to customer satisfaction, loyalty, and complaints, all of which are external indicators. Curiously, non-financial indicators related to employees were deemed the least relevant for decision-making.

Respondents' perceptions of organizational performance when compared to competitors' performance are also shown in Table 2. The results obtained reveal that respondents considered their organization's performance to be moderately higher than that of competitors.

4.2. Hypotheses Testing

In this section we demonstrate how the hypotheses formulated in the second section of the paper were tested, using bivariate and multivariate analyses. An objective of this study was to examine the nature and strength of the association between key variables. This was achieved using correlation analysis, in particular Spearman correlations. Here, when the coefficients are closer to 1 it means a strong positive association between two variables; in contrast, coefficients closer to zero imply weak association. A summary of the results obtained is presented in Table 3.

Table 3. Spearman correlations.

Variable	EU	FI	NFI	PO	E	EE	OP
Environmental uncertainty (EU)	1						
Financial information (FI)	0.09	1					
Non-financial information (NFI)	0.21 *	0.76 **	1				
Process and operations (PO)	0.12	0.73 **	0.85 **	1			
Employees (E)	0.23 *	0.63 **	0.90 **	0.68 *	1		
External environment (EE)	0.21 *	0.72 *	0.90 **	0.69 *	0.73 **	1	
Organizational performance (OP)	−0.06	−0.08	−0.07	−0.18	−0.07	0.02	1

**, * Significant at 1 and 5% (two-tailed), respectively.

Our analysis begins by examining the strength of the association between variables using the Spearman correlation coefficient (Table 3). It appears that there was a positive association, although weak, between environmental uncertainty and non-financial information relevance, related to employees and the external environment. These results are statistically significant (p-value < 0.05), suggesting that with increasing environmental uncertainty greater emphasis should be placed on non-financial information in decision-making. It should be noted, however, that this association between environmental uncertainty and non-financial information relevance was not evident when it comes to information related to processes and operations. Still, the results allow supporting Hypothesis 1. These results are consistent with those of other studies, conducted in other countries (Boulianne 2007; Chenhall and Morris 1986; Chong and Chong 1997; Hoque 2005; Hoque and James 2000; Lal and Hassel 1998). Therefore, we can state that the relevance attributed to non-financial information for decision-making purposes is greater when environmental uncertainty is higher. In these contexts, managers need sophisticated accounting information to enhance decision-making quality (Latan et al. 2018) and non-financial information, in particular, contributes to decision-making success (Monteiro et al. 2022).

No statistically significant association was evidenced between environmental uncertainty and financial information relevance in decision-making. Regarding the relevance attributed to accounting information, the results (Table 3) indicate a strong and positive association between the relevance attributed to financial information and the relevance attributed to non-financial information in decision-making. Thus, it appears that when greater relevance was attributed to financial information in decision-making it was also attributed to non-financial information, whether related to processes and operations and employees (p-value < 0.01), non-financial information connected with customers and suppliers (p-value < 0.05), or vice versa. As argued by Chenhall and Langfield-Smith (2007), non-financial information complements financial information but does not replace it, confirming Hypothesis 3. These findings are in line with the results obtained by Lau and Sholihin (2005) and Chow and Van der Stede (2006). Financial information is not sufficient for decision-making purposes (Monteiro et al. 2021). In several situations, both financial and non-financial information is used (Bhimani and Langfield-Smith 2007; Massicotte and Henri 2021). Therefore, we can state that financial information and non-financial information are complementary.

To compare the relevance given to financial information and non-financial information, both overall and partial (processes and operations, employees and outside), we used the nonparametric Wilcoxon test, which allows comparing measures of central tendency of two variables (Marôco 2021; Pestana and Gageiro 2014). Previous studies have used the Wilcoxon test for similar analyses (e.g., Holm et al. 2016). The results obtained and summarized in Table 4 show that there are statistically significant differences (p-value < 0.01) between the relevance attributed to financial information and non-financial information in decision-making. Through the analysis of the measure of central tendency of these variables it was found that the median of the relevance attributed to financial information (4.17) is above the median of relevance attributed to non-financial information (4.06). However, despite this difference between the relevance attributed to financial information and non-financial information related to processes and operations and employees, no statistically significant differences between the relevance attributed to financial information and non-financial information related to the external environment could be demonstrated. In this context, as verified by Bhimani and Langfield-Smith (2007), there are situations in which financial information is considered more relevant to decision-making than non-financial information. In this study, these situations fell into the contexts of decision-making associated with processes and operations, and employees. These results confirm Hypothesis 2 and are consistent with those obtained by other authors (Chow and Van der Stede 2006; Hyvönen 2007). Financial information remains relevant for managers (Hyvönen 2007). Therefore, we can state that managers assign greater relevance to financial information for decision-making than to non-financial information in decisions related to internal processes, operations, and employees.

Table 4. Wilcoxon tests: relevance of financial and non-financial information.

Description	Financial and Non-Financial	Financial and Process	Financial and Employees	Financial and External Environment
Z	−4.87	−2.54	−7.02	−0.12
Asymp. Sig. (two-tailed)	0.00	0.01	0.00	0.90

To test Hypothesis 4, and thus assess the adjustment between the environmental uncertainty and the relevance assigned to non-financial information, as in previous studies (e.g., Hoque 2005; Hyvönen 2007; Oyewo 2022), we used the linear regression model that allows estimating the direct effects and interaction of independent variables on the dependent variable. Following Afifa and Saleh (2021) and Oyewo (2022), before performing the regression analysis, multicollinearity between the predictor variables considered in each regression model was inspected using correlation analysis. Table 3 shows that all Spearman correlations were at the acceptance level (low correlation levels) and, thus, there was no problem with multicollinearity. Consequently, a regression model was run using the SPSS program. According to Gerdin and Greve (2008, p. 1003) "one frequently used technique for testing the existence of a significant difference in regression coefficients is the moderate regression analysis (MRA)".

MRA has usually the following format:

$$Y = \alpha_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + e$$

where Y represents organizational performance (dependent variable); X_1 represents environmental uncertainty (independent variable); X_2 represents the non-financial information relevance (moderator); $X_1 X_2$ is the moderating effect that X_2 has on the relationship between X_1 and Y; α_0 represents a constant; and e is the error variable (Gerdin and Greve 2008).

The results (Table 5) show that there was no effect of environmental uncertainty in the relevance attributed to non-financial information and in the interaction between these variables on the dependent variable related to organizational performance. This is because the linear regression model and the effects (direct and interactive) of independent variables on organizational performance were not statistically significant.

Table 5. Results of regression.

Variable	Coefficient	Standard Error	*t*-Value	*p*-Value
Constant (α_0)	6.02	1.93	3.12	0.00
Environmental uncertainty (X_1)	−0.71	0.59	−1.21	0.23
Non-financial information (X_2)	−0.59	0.48	−1.24	0.22
X_1X_2	9.173	0.14	1.21	0.23
$R^2 = 0.01$; $F_{3,110} = 0.53$; $p = 0.66$				

In order to verify a possible partial effect of non-financial information and environmental uncertainty on the organizational performance, we used the linear regression for each of the components of non-financial information (processes and operations, employees, and external environment). The results presented in Table 6 confirm that there was no direct effect of the relevance attributed to non-financial information, nor any effect resulting from the interaction with environmental uncertainty on the organizational performance. The models presented and the effects of independent variables on organizational performance are not statistically significant.

Table 6. Additional regression analysis (partial effects).

Variable	Coefficient	Standard Error	*t*-Value	*p*-Value
Constant (α_0)	3.78	1.99	1.89	0.06
Environmental uncertainty (X_1)	0.08	0.61	0.14	0.89
Process and Operations (X_2)	−0.05	0.48	−0.09	0.92
X_1X_2	−0.02	0.15	−0.13	0.90
$R^2 = 0.02$; $F_{3,110} = 0.80$; $p = 0.49$				
Constant (α_0)	5.77	1.42	4.07	0.00
Environmental uncertainty (X_1)	−0.67	0.44	−1.55	0.13
Employees (X_2)	−0.56	0.37	−1.51	0.13
X_1X_2	0.17	0.11	1.54	0.13
$R^2 = 0.02$; $F_{3,110} = 0.81$; $p = 0.49$				
Constant (α_0)	5.96	1.67	3.57	0.00
Environmental uncertainty (X_1)	−0.78	0.52	−1.51	0.13
External environment (X_2)	−0.54	0.40	−1.37	0.17
X_1X_2	0.18	0.12	1.48	0.14
$R^2 = 0.02$; $F_{3,110} = 0.81$; $p = 0.49$				

As there was no direct effect from environmental uncertainty and the relevance assigned to non-financial information, not even one resulting from the interaction between them on organizational performance, Hypothesis 4 is not confirmed. In this sense, the results are surprising and contradict the results obtained by Hoque and James (2000), Hoque (2005), and Al-Mawali and Am (2016), who concluded that organizational performance is influenced by the interplay between environmental uncertainty and the relevance assigned to non-financial information for decision-making or the use of customer accounting information. These differences can occur because other aspects are not being considered, which may also influence environmental uncertainty, non-financial information relevance, and organizational performance. Therefore, additional research should be conducted to investigate this association.

5. Conclusions

In recent years, the organizational environment has changed significantly because of the globalization of business, increased competition, rapidly changing technologies and demands for continuous improvement, and more recently as a result of the pandemic crisis and the recent war in Ukraine, all of which have increased uncertainty for companies. In these circumstances, managers need not only more accounting information, but also more timely accounting information to make better decisions and thus achieve their goals. That is why non-financial information becomes relevant in decision-making, and also because it is available sooner and a good non-financial performance may translate into good financial performance.

From the literature review, four hypotheses were formulated aiming to analyze the association between environmental uncertainty, (financial and non-financial) accounting information relevance in decision-making, and organizational performance. It appears that the organizational environment of large manufacturing companies operating in Portugal has a moderate uncertainty, resulting mainly from intense competition and high demands for social and environmental responsibility. In situations of higher environmental uncertainty, greater emphasis is given to non-financial information for decision-making, particularly to the non-financial information related to employees and customers, and supplier assessment. However, in general, managers continue to attach greater relevance to financial information than to non-financial information related to processes, operations, and employees. The same is not true for non-financial information related to the external environment. Our results suggest that in the evaluation of customers and suppliers, non-financial information is of similar relevance to managers as financial information. Furthermore, when greater relevance is assigned to financial information, greater emphasis is also given to non-financial information. Thus, we conclude that financial information and non-financial information are complementary.

Regarding the impact of environmental uncertainty and the relevance attributed to non-financial information on organizational performance, contrary to the results of Hoque and James (2000), Hoque (2005), and Al-Mawali and Am (2016), the data collected did not reveal statistically significant results. Thus, there is a need to understand what additional factors (e.g., contingent factors such as competitiveness, strategy, technology, organizational structure, and organizational and national culture) can influence these variables and have led to different findings.

The implications of this paper are relevant not only for researchers but also for practitioners (e.g., accounting professionals and managers). From the theoretical point of view, the results of this study contribute to theory by validating the findings of previous contingency studies regarding the influence of environmental uncertainty on non-financial information relevance for decision-making purposes. In line with the contingency perspective, non-financial information is more relevant under specific circumstances, that is, uncertainty contexts. Nevertheless, this information should be combined with financial information, given their complementary nature. The results regarding the influence of environmental uncertainty on the relationship between non-financial information relevance and organizational performance, however, challenge findings of previous studies. Thus, further research is needed to examine this relationship, considering also what additional factors can influence this relationship and justify the different results achieved.

From the practical point of view, the findings of this research reveal the need for accounting professionals and managers to provide and use non-financial information in decision-making, namely in uncertainty contexts. In these contexts, given their complementary nature, not only non-financial information should be provided and used, but also financial information. These findings support the implementation of accounting practices, in particular strategic management accounting practices, which collect and provide financial and non-financial information. Some of these strategic management accounting practices are particularly devoted to providing non-financial information on customers, which is considered relevant to decision-making purposes. Therefore, organizations regarding the external environment as uncertain have a greater tendency to use strategic management accounting practices to survive intense competition (Oyewo 2022).

This study has some limitations that should be considered when interpreting the results. A main limitation relates to the instrument used for data collection, the questionnaire-based survey, as it restricts the number of questions, prevents the placement of new questions to clarify certain situations detected, does not allow the respondent to aid in the interpretation of questions, and is not always filled in by the most suitable person. Another limitation arises from the fact it was a "cross section" study. Studies with contingency approaches have been the target of some criticism (Chapman 1997; Tillema 2005). Longitudinal studies using qualitative methods, such as case studies, may help to overcome these criticisms, since these approaches consider the specific context of the organizations not captured by questionnaire-based surveys (Otley 2016). However, "surveys" are needed to make generalizations and thus build a coherent whole (Chenhall 2003; Tillema 2005).

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

Funding: This research was supported by National Funds through the FCT (Portuguese Foundation for Science and Technology), I.P., within the scope of the project Ref. UIDB/04630/2020.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Notes

1 One company was excluded because it ceased activity. Another company has been excluded for failing to engage in the transformation. Two companies were excluded because they had ceased to form two new companies (already part of the list supplied by the INE). Finally, five companies were excluded because we could not reach them (via phone and email).

2 We adopted the seven dimensions used by Cadez and Guilding (2008).

References

Adeniran, Adetayo Olaniyi, and Oluwabukunmi Eunice Obembe. 2020. The significance of strategic management accounting on the performance of transport businesses in Nigeria. *Insights Into Regional Development* 2: 5. [CrossRef]

Afifa, Malik Muneer Abu, and Isam Saleh. 2021. Management accounting systems effectiveness, perceived environmental uncertainty and enterprise risk management: Evidence from Jordan. *Journal of Accounting & Organizational Change* 17: 704–27. [CrossRef]

Al-Mawali, Hamzah, and Tri-Dung L. Am. 2016. Customer accounting and environmental uncertainty: Sequential explanatory study. *International Review of Management and Marketing* 6: 532–43.

Alves, Maria do Céu F. Gaspar. 2010. Management accounting and information technology – Some empirical evidence. In *Studies in Managerial and Financial Accounting*. Edited by Marc J. Epstein, Jean-François Manzoni and Antonio Davila. Bingley: Emerald Group Publishing Limited, vol. 20, pp. 429–55. [CrossRef]

Alves, Maria do Céu Gaspar. 2017. Perspectives on accounting information use: A multiple case study. *International Business Management* 11: 1316–27. [CrossRef]

Aon. 2020. Decision Making in Complex & Volatile Times: Keys to Managing COVID-19. Available online: https://www.aon.com/getmedia/86577ac3-1ce2-43eb-8323-1a1191488355/Decision-Making-in-Complex-and-Volatile-Times-Keys-to-Managing-COVID-19-2020-05-06.aspx (accessed on 12 April 2022).

Abu-Rahma, Ali, and Bushra Jaleel. 2019. Perceived uncertainty and use of environmental information in decision making. *International Journal of Organizational Analysis* 27: 690–711. [CrossRef]

Abdel-Kader, Magdy, and Robert Luther. 2008. The impact of firm characteristics on management accounting: A UK-based empirical analysis. *The British Accounting Review* 40: 2–27. [CrossRef]

Abernethy, Margaret A., and Anne M. Lillis. 1995. The impact of manufacturing flexibility on management control system design. *Accounting, Organizations and Society* 20: 241–58. [CrossRef]

Baines, Annette, and Kim Langfield-Smith. 2003. Antecedents to management accounting change: A structural equation approach. *Accounting, Organizations and Society* 28: 675–98. [CrossRef]

Bertolotti, Fabiola, Diego Maria Macrì, and Matteo Vignoli. 2019. Strategic alignment matrix: Supporting management coordination in complex organizations. *Journal of Accounting & Organizational Change* 15: 557–79. [CrossRef]

Bhimani, Alnoor, and Kim Langfield-Smith. 2007. Structure, formality and the importance of financial and non-financial information in strategy development and implementation. *Management Accounting Research* 18: 3–31. [CrossRef]

Boulianne, Emilio. 2007. Revisiting fit between AIS design and performance with the analyzer strategic-type. *International Journal of Accounting Information Systems* 8: 1–16. [CrossRef]
Cadez, Simon, and Chris Guilding. 2008. An exploratory investigation of an integrated contingency model of strategic management accounting. *Accounting, Organizations and Society* 33: 836–63. [CrossRef]
Cescon, Franco, Antonio Costantini, and Luca Grassetti. 2019. Strategic choices and strategic management accounting in large manufacturing firms. *Journal of Management and Governance* 23: 605–36. [CrossRef]
Chapman, Christopher S. 1997. Reflections on a contingent view of accounting. *Accounting, Organizations and Society* 22: 189–205. [CrossRef]
Chenhall, Robert H. 2003. Management control systems design within its organizational context: Findings from contingency-based research and directions for the future. *Accounting, Organizations and Society* 28: 127–68. [CrossRef]
Chenhall, Robert H., and Kim Langfield-Smith. 1998a. Adoption and benefits of management accounting practices: An Australian study. *Management Accounting Research* 9: 1–19. [CrossRef]
Chenhall, Robert H., and Kim Langfield-Smith. 1998b. The relationship between strategic priorities, management techniques and management accounting: An empirical investigation using a systems approach. *Accounting, Organizations and Society* 23: 243–64. [CrossRef]
Chenhall, Robert H., and Kim Langfield-Smith. 2007. Multiple perspectives of performance measures. *European Management Journal* 25: 266–82. [CrossRef]
Chenhall, Robert H., and Deigan Morris. 1986. The impact of structure, environment, and interdependence on the perceived usefulness of management accounting systems. *The Accounting Review* 61: 16–35.
Chong, Vincent K. 1996. Management accounting systems, task uncertainty and managerial performance: A research note. *Accounting, Organizations and Society* 21: 415–21. [CrossRef]
Chong, Vincent K., and Kar Ming Chong. 1997. Strategic choices, environmental uncertainty and SBU performance: A note on the intervening role of management accounting systems. *Accounting and Business Research* 27: 268–76. [CrossRef]
Chow, Chee W., and Wim A. Van der Stede. 2006. The use and usefulness of nonfinancial performance measures. *Management Accounting Quarterly* 7: 1–8.
Czaja-Cieszyńska, Hanna, Dominika Kordela, and Beata Zyznarska-Dworczak. 2021. How to make corporate social disclosures comparable? *Entrepreneurship and Sustainability Issues* 9: 268–88. [CrossRef]
Drury, Colin, and Mike Tayles. 2006. Profitability analysis in UK organizations: An exploratory study. *The British Accounting Review* 38: 405–25. [CrossRef]
Fleming, Christopher M., and Mark Bowden. 2009. Web-based surveys as an alternative to traditional mail methods. *Journal of Environmental Management* 90: 284–92. [CrossRef]
Frazer, Linval. 2020. Does internal control improve the attestation function and by extension assurance services? A practical approach. *Journal of Accounting and Finance* 20: 28–38. [CrossRef]
Ganassali, Stéphane. 2008. The influence of the design of web survey questionnaires on the quality of responses. *Survey Research Methods* 2: 21–32. [CrossRef]
Gerdin, Jonas, and Jan Greve. 2008. The appropriateness of statistical methods for testing contingency hypotheses in management accounting research. *Accounting, Organizations and Society* 33: 995–1009. [CrossRef]
Gillham, Bill. 2008. *Developing a Questionnaire*. London: Continuum.
Guilding, Chris, Karen S. Cravens, and Mike Tayles. 2000. An international comparison of strategic management accounting practices. *Management Accounting Research* 11: 113–35. [CrossRef]
Haldma, Toomas, and Kertu Lääts. 2002. Contingencies influencing the management accounting practices of Estonian manufacturing companies. *Management Accounting Research* 13: 379–400. [CrossRef]
Hayes, David C. 1977. The contingency theory of managerial accounting. *The Accounting Review* 52: 22–39.
Hill, Manuela Magalhães, and Andrew Hill. 2008. *Investigação por questionário*. Lisboa: Edições Sílabo.
Holm, Morten, Vanktesh Kumar, and Thomas Plenborg. 2016. An investigation of customer accounting systems as a source of competitive advantage. *Advances in Accounting, incorporating Advances in International Accounting* 32: 18–30. [CrossRef]
Hoque, Zahirul. 2004. A contingency model of the association between strategy, environmental uncertainty and performance measurement: Impact on organizational performance. *International Business Review* 13: 485–502. [CrossRef]
Hoque, Zahirul. 2005. Linking environmental uncertainty to non-financial performance measures and performance: A research note. *The British Accounting Review* 37: 471–81. [CrossRef]
Hoque, Zahirul, and Wendy James. 2000. Linking balanced scorecard measures to size and market factors: Impact on organizational performance. *Journal of Management Accounting Research* 12: 1–17. [CrossRef]
Hyvönen, Johanna. 2007. Strategy, performance measurement techniques and information technology of the firm and their links to organizational performance. *Management Accounting Research* 18: 343–66. [CrossRef]
Johnson, H. Thomas, and Robert S. Kaplan. 1987. The rise and fall of management accounting. *Management Accounting* 68: 22–30. [CrossRef]
Lal, Mohan, and Lars Hassel. 1998. The joint impact of environmental uncertainty and tolerance of ambiguity on top managers' perceptions of the usefulness of non-conventional management accounting information. *Scandinavian Journal of Management* 14: 259–71. [CrossRef]

Latan, Hengky, Charbel Jose Chiappetta Jabbour, Ana Beatriz Lopes de Sousa Jabbour, Samuel Fosso Wamba, and Muhammad Shahbaz. 2018. Effects of environmental strategy, environmental uncertainty and top management's commitment on corporate environmental performance: The role of environmental accounting. *Journal of Cleaner Production* 180: 297–306. [CrossRef]

Lau, Chong M., and Mahfud Sholihin. 2005. Financial and nonfinancial performance measures: How do they affect job satisfaction? *The British Accounting Review* 37: 389–413. [CrossRef]

Lodge, Martin, and Arjen Boin. 2020. COVID-19 as the Ultimate Leadership Challenge: Making Critical Decisions without Enough Data. Available online: https://blogs.lse.ac.uk/politicsandpolicy/covid-19-as-the-ultimate-leadership-challenge/ (accessed on 12 April 2022).

Löfsten, Hans, and Peter Lindelöf. 2005. Environmental hostility, strategic orientation and importance of management accounting—An empirical analysis of new technology-based firms. *Technovation* 25: 725–38. [CrossRef]

Loomis, David K., and Shona Paterson. 2018. A comparison of data collection methods: Mail versus online surveys. *Journal of Leisure Research* 49: 133–49. [CrossRef]

Ma, Lindong, Xihui Chen, Jiawen Zhou, and Luigi Aldieri. 2022. Strategic management accounting in small and medium-sized enterprises in emerging countries and markets: A case study from China. *Economies* 10: 74. [CrossRef]

Marôco, João. 2021. *Análise Estatística Com o SPSS Statistics*. Lisboa: Edições Sílabo.

Massicotte, Steeve, and Jean-François Henri. 2021. The use of management accounting information by boards of directors to oversee strategy implementation. *The British Accounting Review* 53: 100953. [CrossRef]

Monteiro, Albertina Paula, Catarina Cêpeda, Amélia Silva, Eduardo Leite, and Élvio Camacho. 2021. The role of accounting information in decision-making and economic performance: The Portuguese accountants' perspective. *Entrepreneurship and Sustainability Issues* 9: 30. [CrossRef]

Monteiro, Albertina Paula, Joana Vale, Eduardo Leite, Marcin Lis, and Joanna Kurowska-Pysz. 2022. The impact of information systems and non-financial information on company success. *International Journal of Accounting Information Systems* 45: 100557. [CrossRef]

Mia, Lokman, and Brian Clarke. 1999. Market competition, management accounting systems and business unit performance. *Management Accounting Research* 10: 137–58. [CrossRef]

Newkirk, Henry E., and Albert L. Lederer. 2006. The effectiveness of strategic information systems planning under environment uncertainty. *Information & Management* 43: 481–501. [CrossRef]

Otley, David. 2016. The contingency theory of management accounting and control: 1980–2014. *Management Accounting Research* 31: 45–62. [CrossRef]

Oyewo, Babajide. 2021. Do innovation attributes really drive the diffusion of management accounting innovations? Examination of factors determining usage intensity of strategic management accounting. *Journal of Applied Accounting Research* 22: 507–38. [CrossRef]

Oyewo, Babajide. 2022. Contextual factors moderating the impact of strategic management accounting on competitive advantage. *Journal of Applied Accounting Research* ahead-of-print: 1–29. [CrossRef]

Pestana, Maria Helena, and João Nunes Gageiro. 2014. *Análise de Dados para Ciências Sociais—A Complementaridade do SPSS*. Lisboa: Edições Sílabo.

Teo, Thompson S. H., and William R. King. 1997. Integration between business planning and information systems planning: An evolutionary-contingency perspective. *Journal of Management Information Systems* 14: 185–214. [CrossRef]

Tillema, Sandra. 2005. Towards an integrated contingency framework for MAS sophistications—Case studies on the scope of accounting instruments in Dutch power and gas companies. *Management Accounting Research* 16: 101–29. [CrossRef]

Thuan, Pham Quoc, Nguyen Vinh Khuong, Nguyen Duong Cam Anh, Nguyen Thi Xuan Hanh, Vo Huynh Anh Thi, Tieu Ngoc Bao Tram, and Chu Gia Han. 2022. The determinants of the usage of accounting information systems toward operational efficiency in industrial revolution 4.0: Evidence from an emerging economy. *Economies* 10: 83. [CrossRef]

Turner, Michael J., Sean A. Way, Demian Hodari, and Wiarda Wiettman. 2017. Hotel property performance: The role of strategic management accounting. *International Journal of Hospitality Management* 63: 33–43. [CrossRef]

Vaivio, Juhani. 1999. Exploring a 'non-financial' management accounting change. *Management Accounting Research* 10: 409–37. [CrossRef]

Visedsun, Nimnual, and Kanitsorn Terdpaopong. 2021. The effects of the strategy and goal on business performance as mediated by management accounting systems. *Economies* 9: 149. [CrossRef]

economies

MDPI

Article

The Impact of Remittances on Saving Behaviour and Expenditure Patterns in Vietnam

Thanh Xuan Hua [1], Roselinde Kessels [2,3] and Guido Erreygers [3,4,*]

[1] Department of Economics, College of Economics, Can Tho University, Can Tho City 94000, Vietnam
[2] Department of Data Analytics and Digitalization, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands
[3] Department of Economics, University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium
[4] Centre for Health Policy, University of Melbourne, 207 Bouverie Street, Carlton, VIC 3010, Australia
* Correspondence: guido.erreygers@uantwerpen.be

Abstract: We examine the effects of receiving remittances on household saving behaviour and expenditure patterns in Vietnam. We consider the amount of saving, the saving rate, and the share of expenditure, as well as per capita expenditure on education, health, assets, house repairs, food, non-food, and utilities. We apply propensity score matching to data from the Vietnam Household Living Standard Survey (VHLSS) of 2012. We find that remittances have a positive impact on household savings and increase both the amount of saving and the saving rate. As far as expenditure patterns are concerned, our results indicate that receiving households spend more on health, assets, and house repairs, and less on food. This finding suggests that households tend to use remittances productively, with receiving households increasing their investments in human and physical capital. For the economy as a whole, remittances can create more opportunities for the development of services provided by banks, financial institutions, hospitals and healthcare centres, and give incentives to the production of building materials and tangible assets.

Keywords: remittances; household saving behaviour; household expenditure patterns; propensity score matching; Vietnam

Citation: Hua, Thanh Xuan, Roselinde Kessels, and Guido Erreygers. 2022. The Impact of Remittances on Saving Behaviour and Expenditure Patterns in Vietnam. *Economies* 10: 223. https://doi.org/10.3390/economies10090223

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 18 July 2022
Accepted: 29 August 2022
Published: 14 September 2022

1. Introduction

Remittances are a common source of income for households in many developing countries (McKenzie and Sasin 2007), and Vietnam is no exception (Nguyen 2008). The flow of external remittances into the country increased significantly between 2000 and 2021. According to the World Bank's *Migration and Development Brief 36*, in 2021, Vietnam received an inflow of remittances estimated at 18 billion USD, equal to 4.9% of GDP (Ratha et al. 2022, p. 31). In addition, Vietnam's rapid urbanisation and the migration of labourers from rural to urban areas have led to an increasing trend in internal remittances (World Bank 2012). In spite of the fact that both external and internal remittances have become more prevalent in the country, there are only a few papers that have investigated the effects of remittances on income and welfare (Nguyen 2008; Nguyen and Mont 2012; Nguyen et al. 2017; Nguyen and Vu 2018). This paper contributes to the literature on the effects of remittances in Vietnam by studying how they affect saving behaviour and expenditure patterns.

In general, quite a few papers have examined the effect of remittances on savings and expenditure patterns (see, e.g., Castaldo and Reilly 2007; Ang et al. 2009; Tabuga 2010; Clément 2011; Haider et al. 2016; Quartey et al. 2019; Opiniano 2021; Salahuddin et al. 2022). Both the life-cycle theory of Modigliani and Brumberg (1954) and the permanent income hypothesis proposed by Friedman (1957) have been influential with regard to the identification of possible relations among remittances, expenditures, and savings. According to the life-cycle theory, how much households consume and save depends on the total income they receive rather than on the specific sources from which they derive their income. This means that remittances are fungible and used like income from other sources, such as wages. The implication is that remittances, like any other income, can influence both household consumption and household saving (Haider et al. 2016; Nguyen and Vu 2018). Likewise, remittances can be used both for investment expenditures (such as education, health, and housing) and consumption expenditures (such as food, non-food, and utilities), as shown by Castaldo and Reilly (2007), Adams et al. (2008), Ang et al. (2009), and Tabuga (2010). In this view, there would be no significant differences between the expenditure patterns of households that receive remittances and those that do not.

According to the permanent income hypothesis, however, remittances can be either a type of permanent income or a type of transitory income. In the first case, remittance income is perceived as stable over time, and households tend to use it for consumption. Therefore, remittances are predominantly used for consumption expenditures rather than for savings and investments. For example, Clément (2011) found that, in Tajikistan, households receiving remittances allocated a higher share of expenditure to food and utilities and a lower share to housing and investment than households not receiving remittances, concluding that remittances are not used productively. Likewise, Zhu et al. (2014) found that Chinese rural households tend to use their remittances for consumption rather than for investment, concluding that, in rural areas of China, remittances should be considered as permanent income. Their findings were confirmed by Démurger and Wang (2016), who stressed that remittances could be detrimental to sustaining investment in human capital for rural families in China. By contrast, if remittances are treated as a form of transitory income with unexpected, accidental occurrence, then they will be used mainly for saving and investment due to the zero propensity to consume from this income component. In this way, remittances would have a productive use and influence the growth and development capabilities of households (Yang 2008; Randazzo and Piracha 2019). For example, Adams and Cuecuecha (2010) found that, in Guatemala, households receiving external remittances used these productively on two forms of investment expenditure: education and housing. Moreover, these households also spent less on food than households not receiving remittances. The productive use of remittances for education was also found in the studies of Cardona-Sosa and Medina (2006) in Colombia, Yang (2008) in the Philippines, and Randazzo and Piracha (2019) in Senegal. In addition, Taylor and Mora (2006) found that, in Mexico, remittances increased expenditures on education, health, and housing rather than those on consumption goods. Similar findings were obtained by Ponce et al. (2011) for Ecuador, and by Berloffa and Giunti (2019) for Peru. Using a dataset of 141 countries, Ait Benhamou and Cassin (2021) found that remittances tended to increase investment in education at the expense of investment in physical capital.

As far as Vietnam is concerned, the literature about the impact of remittances has focused mainly on income and expenditure. Most studies have relied on the Vietnamese Household and Living Standards Survey (VHLSS) datasets and defined remittances as receipts of households from other people, such as migrant members, relatives, and friends. Using the VHLSS 2002 and 2004 datasets, Nguyen (2008) found that, as far as external remittances are concerned, the impact on income was much higher than the impact on consumption expenditures, which suggests that a large proportion of these remittances were used for saving and investment. For internal remittances, by contrast, the effect on consumption expenditures was only slightly smaller than the effect on income.

The effect of external remittances on investment expenditures was confirmed by Nguyen and Mont (2012), who used the updated VHLSS 2006 and 2008 datasets. Households receiving remittances were likely to invest these in housing, land, debt payments, and saving, rather than to increase their consumption. In addition, Nguyen and Vu (2018) examined the patterns and impact of migration and remittances on household welfare in Vietnam using data from the VHLSS 2010 and 2012 datasets. They found that remittances help households increase per capita income and per capita expenditure, with the effect of remittances on expenditure being smaller than the effect on income. They concluded that households receiving remittances use these not only for consumption but also for saving.

Lastly, Nguyen et al. (2017) conducted a study to investigate the effect of remittances on the expenditures of internal migrant households in rural areas of three provinces (Ha Tinh, Thua Thien Hue, and Dak Lak) for the years 2007, 2008, and 2010, using data from a non-VHLSS survey. Remittances were defined as household receipts from migrant members who had moved to urban areas outside the original province for at least 1 month. The authors showed that migration with remittances had a positive effect on housing and other non-food expenditures, while migration without remittances had a positive effect on food, healthcare, and other non-food expenditures, but a negative one on education expenditures.

This paper focuses on the impact of remittances on the saving behaviour and expenditure patterns of Vietnamese households. More specifically, our research question was the following: when it comes to saving, investment, and consumption, do households that receive remittances differ from households that do not receive remittances? Using the propensity score matching (PSM) approach proposed by Rosenbaum and Rubin (1983), we compare households that receive remittances to households that do not, as well as explore how the former use their remittances. As far as saving behaviour is concerned, we look at the saving amount and the saving rate. Regarding expenditure patterns, we consider both the shares and the per capita expenditures on education, health, assets, house repairs, food, non-food, and utilities. We treat the first four categories as household investment expenditures, and the last three as household consumption expenditures. We also explore whether remittances increase household income, by comparing the incomes of receiving and non-receiving households.

2. Methodology

2.1. Methods

Two approaches have been dominant in the empirical research on the impact of remittances on saving behaviour and expenditure patterns: the Working–Leser Engel curve regression approach (Working 1943; Leser 1963) and the PSM approach (Rosenbaum and Rubin 1983). The first relies on the hypothesis of household utility maximisation to construct a basic model to estimate the shares of expenditure as a function of the logarithm of total household expenditure. The basic model has then been extended to include other variables assumed to affect the shares, such as household characteristics (Deaton 2019). In this way, empirical researchers have added a dummy variable to study the impact of remittances on household expenditure patterns (Taylor and Mora 2006; Castaldo and Reilly 2007; Tabuga 2010). The Working–Leser Engel curve regression model then looks as follows:

$$w_{ij} = \alpha_j + \beta_j \ln(z_i) + \gamma_j x_i + \theta_j REMIT_i + \nu_{ij}, \tag{1}$$

where w_{ij} is the share of expenditure of good j by household i, z_i is the total expenditure of household i, x_i is the vector of household characteristics of household i, $REMIT_i$ is a binary variable indicating whether household i receives remittance $(REMIT_i = 1)$ or not $(REMIT_i = 0)$, and ν_{ij} the error term.

The coefficient θ_j in Equation (1), considered as the difference in the shares of expenditure on good j by a receiving and non-receiving household, is often estimated by OLS regression. However, since the remittance variable $REMIT_i$ is also influenced by the household characteristics x_i, this can lead to an endogeneity problem. Unobserved variables may affect both the household expenditure pattern and its remittance status. Theoretically, this is a major problem that needs to be solved. If not, the estimated impact of household remittances on expenditure patterns will be biased (Deaton et al. 1989).

One popular method to avoid endogeneity in this approach is to use instrumental variables. With aggregate data, Aggarwal et al. (2011) suggested using per capita GDP and the unemployment rate as instrumental variables in studying the impact of remittances on financial development. With microdata, however, it is difficult to identify a suitable instrumental variable for remittances (Randazzo and Piracha 2019). Furthermore, McKenzie and Sasin (2007) argued that it is difficult to determine a valid instrumental variable that strongly correlates with the receipt of remittances but has no direct influence on the household expenditure pattern. Using invalid instruments can result in an even larger bias in impact estimates (Nguyen and Mont 2012; Randazzo and Piracha 2019).

Instead, the PSM approach proposed by Rosenbaum and Rubin (1983) performs well when it comes to estimating the effect of remittances on expenditure patterns (Caliendo and Kopeinig 2008; McKenzie et al. 2010; Clément 2011; Li 2012; Randazzo and Piracha 2019). That is why we apply this alternative method in our paper. In general, PSM has been applied to estimate causal treatment effects in various fields of study. The basic idea of PSM is to compare and match households in the treated group with those in the non-treated group in terms of similar observable characteristics. In other words, the causal effect of the treatment is measured by the difference in outcomes between the treated and non-treated groups that have similar observable characteristics. In this way, selection bias between treated and non-treated households can be reduced (Clément 2011). Usually, this approach consists of six steps as shown in an application by Caliendo and Kopeinig (2008).

Step 1. We begin by constructing a logit/probit model to estimate the propensity score, i.e., the probability that a household receives remittances, as a function of a set of household characteristics. According to Caliendo and Kopeinig (2008), the choice between logit and probit is not critical since both usually yield similar results in the case of a binary dependent variable. Following previous papers (Caliendo and Kopeinig 2008; McKenzie et al. 2010; Clément 2011; Li 2012), we use a logit regression,

$$logit(P(REMIT_i = 1)) = \ln\left(\frac{P(REMIT_i = 1)}{1 - P(REMIT_i = 1)}\right) = \beta x_i + \varepsilon_i, \tag{2}$$

where $P(REMIT_i = 1)$ is the probability of receiving remittances for household i with observed covariates x_i, β is the effect of x_i on (the logit of) the probability of receiving remittances, and ε_i is the error term.

The observed covariates to estimate the propensity score using Equation (2) should be chosen on the basis of relevant theories, institutional settings, and previous empirical studies, and they should have a simultaneous impact on the treatment (receiving remittances) and on the potential outcomes (expenditure patterns and saving behaviour) to attain a reliable result (Heckman et al. 1997; Caliendo and Kopeinig 2008; Li 2012). We followed Clément (2011), Randazzo and Piracha (2019), and other remittance studies applying the PSM approach by not including income as a covariate to estimate the propensity score, for two reasons. First, according to the authors of these studies, household income does not have an impact on receiving remittances; thus, its addition in the logit regression would not be meaningful. Second, theories of consumption state that household characteristics influence household income. Hence, including income as a covariate together with these characteristics could cause an endogeneity bias.

Step 2. We determine the region of common support by comparing the range of the propensity scores of households receiving remittances (the treated group) to that of households not receiving remittances (the non-treated group). The determination is based on the minima, maxima, and density of the propensity scores in the treated and non-treated groups (Caliendo and Kopeinig 2008). The purpose is to ensure that treated households can be matched with some non-treated households having a similar propensity score. Any treated household which has a propensity score lying outside the region of common support is dropped, since we cannot find any non-treated household with which it can be matched (Caliendo and Kopeinig 2008).

Step 3. We use PSM estimators to match each observation in the treated group with one or more observations in the non-treated group. Rosenbaum and Rubin (1983) constructed several PSM estimators, which differ in the way the neighbourhood for each treated observation is defined and in the weights that are assigned to the neighbours. Each estimator presents advantages and drawbacks in terms of the quality and quantity of the matches (Caliendo and Kopeinig 2008). Asymptotically, all PSM estimators should yield the same results; in practice, however, various matching estimators should be implemented to compare the results and check the robustness of the findings (Caliendo and Kopeinig 2008; Garrido et al. 2014; Randazzo and Piracha 2019). Following previous empirical studies, we consider the *k*-nearest neighbour (*k*NN) estimator (with k = 5), the radius calliper estimator (with caliper r = 0.001), and the kernel estimator.

The *k*NN estimator matches each treated household with the k closest non-treated households in terms of propensity score. We applied both k = 5 and k = 10, but report the results for k = 5 only, since the results for k = 10 are similar. Using the *k*NN estimator, bad matches can occur if there are treated households for which the nearest non-treated households lie relatively far away. This problem can be avoided by applying the radius calliper estimator. The calliper is the maximum propensity score difference that can be allowed. A calliper fixed at 0.001 means that each treated household must be matched with non-treated households with a propensity score that differs at most 0.001 from the propensity score of the treated household. For both of these estimators, each treated household is matched with only a few households in the non-treated group. As a result, non-treated households not matched with any treated household are excluded from the matched sample. By contrast, typical of the kernel estimator is that it matches each treated household with a weighted average of all households in the non-treated group.

Step 4. It is crucial to check whether the observed covariates and propensity score distributions of the treated and non-treated groups are balanced after matching. If these distributions are not balanced or equivalent after matching, the results of the PSM approach could be misleading and biased, and the propensity scores estimated by the selected covariates need to be re-examined (Rosenbaum and Rubin 1983; Caliendo and Kopeinig 2008; Garrido et al. 2014). Various tests have been proposed to check the balancing property of the observed covariates and propensity score distributions after matching (Rosenbaum and Rubin 1985; Sianesi 2004; Ho et al. 2007; Austin 2009). More details on the tests we apply can be found in the Supplementary Materials.

Step 5. If the matched sample is sufficiently balanced, the effect caused by the treatment can be determined by the average treatment effect on the treated (*ATT*), which is defined as the difference between the expected outcomes with and without treatment for households in the treatment group. In this paper, the *ATT* is the effect of remittances on saving behaviour and expenditure patterns, defined as follows:

$$ATT = E(Y_{i1}|REMIT_i = 1) - E(Y_{i0}|REMIT_i = 1), \tag{3}$$

where $E(Y_{i1}|REMIT_i = 1)$ and $E(Y_{i0}|REMIT_i = 1)$ are the expected outcomes with and without treatment for households in the treatment group. In PSM, the expected outcomes without treatment for households in the treatment group, $E(Y_{i0}|REMIT_i = 1)$, are simply the expected outcomes for households in the non-treated group after matching, $E(Y_{i0}|REMIT_i = 0)$. Hence, we obtain

$$ATT = E(Y_{i1}|REMIT_i = 1) - E(Y_{i0}|REMIT_i = 0). \tag{4}$$

Step 6. We apply the bounding approach proposed by Rosenbaum (2002) (see the Supplementary Materials for more details) to test whether our results are robust. In PSM, unobserved covariates are assumed not to influence the *ATT*. If there are any unobserved covariates that affect the treatment and the outcomes simultaneously, a hidden bias might occur. Rosenbaum's test allows us to examine whether the *ATT* is sensitive to the influence of unobserved covariates under varying degrees of assumed hidden bias (Becker and Caliendo 2007).

As mentioned before, the PSM method performs well in situations where we cannot find suitable instrumental variables to avoid the endogeneity problems which often occur when studying the effect of remittances on household expenditure patterns and saving behaviour (Caliendo and Kopeinig 2008; McKenzie et al. 2010; Li 2012; Démurger and Wang 2016; Randazzo and Piracha 2019). An additional advantage of the PSM method is that it allows us to reduce the sources of bias in observational data (Heckman et al. 1998). Since non-matched households are excluded, the bias from non-overlapping observations is reduced. Moreover, the technique allows us to reweigh the non-treated households so as to obtain equivalent distributions for the treated and non-treated households, which diminishes the bias due to the difference in density weighting between these two groups. In the next section, we apply the PSM method making use of the *psmatch2* module in Stata (Leuven and Sianesi 2018) and the statistical software package JMP (SAS Institute Inc., Cary, NC, USA).

2.2. Data

We use the VHLSS survey of the year 2012, which collected data on 9399 households. We deleted six households with household heads younger than 18 years. Next, we excluded three households which answered 'do not remember' to the question whether they received remittances, 608 households with missing educational information, and three households for which we could not determine the saving rate. Overall, our final sample consists of 8778 households of which 2174 households (24.73%) received remittances, and 6604 households (75.23%) did not. Regarding the sources of remittances, 2071 received internal remittances, and 159 households received external remittances. Because the sample of households receiving external remittances was small, we did not analyse the impact of internal and external remittances separately. Instead, we focused on total remittances.

In the literature, remittances have been defined in at least two different ways. Remittances can be conceived broadly as the sum of what a household receives from migrant members, relatives, friends, and neighbours (Castaldo and Reilly 2007; Clément 2011). Most studies of remittances in Vietnam have defined remittances in this way, by using available information of household receipts in the VHLSS dataset (Nguyen et al. 2008; Nguyen and Mont 2012; Nguyen and Vu 2018). A more narrow definition of remittances comes from the literature on the new economics of labour migration (Stark and Bloom 1985). Here, remittances are limited to what migrant members send to their families. This provides information on both remittances and migration, which has been used in numerous studies (Tabuga 2010; Démurger and Wang 2016; Nguyen et al. 2017; Randazzo and Piracha 2019).

In this study, we adopt the more narrow definition of remittances. One reason for this is that migration within Vietnam, as well as to other countries, has increased rapidly in recent years (Junge et al. 2015; Nguyen et al. 2017; Luong 2018; Nguyen and Vu 2018). Vietnam's 2009 census showed that 8.5% of the population represented inter- and intra-provincial migrants, and the government expects that this percentage will continue to rise (World Bank 2016). Furthermore, around four million people of Vietnamese descent are living abroad (Ministry of Foreign Affairs of Vietnam 2012). Another reason is that most of the current literature on remittances in Vietnam is based on the broad definition.

We use the VHLSS 2012 dataset because it contains a special module on migration with extensive data on both migrants and how much they send home—information which is missing in earlier and later waves of the VHLSS survey. This also explains why the VHLSS 2012 dataset has been used in other studies of migration in Vietnam, such as the volume on rural–urban migration edited by Liu and Meng (2019). In accordance with the data of the special module, we define migrant members as people who have left their households, but are still considered as important to the household in terms of either filial responsibility or financial contributions. We also take our remittance data from this module. It should be noted, however, that these data do not coincide perfectly with the remittance data in other modules of the survey. Since the survey calculates a household's total income on the basis of the remittance data of the other modules rather than on the data of the special migration module, we adjusted the income data of households. More precisely, we replaced the remittance data used by the survey to calculate total income using the remittance data from the migration module. In this way, we obtain what we call the adjusted income of households, which we use as an outcome variable.

We use two sets of outcomes for saving behaviour and expenditure patterns. Regarding saving behaviour, we study the saving amount (expressed in 1000 Vietnamese dong, VND) and the saving rate. With respect to expenditure patterns, we analyse the share of expenditure and per capita expenditure (1000 VND) on various categories, including education, health, assets, house repairs, food, non-food, and utilities. A description of the expenditure categories can be found in Table 1.

Table 1. Description of expenditure categories.

Category	Description
Education	All education expenses of the household members, including tuition fees, contributions to class, school, uniforms, books, study instruments (paper, pen, etc.), coaching sessions, and others (such as exam fees, travel, rent, and student body insurance).
Health	All expenses for healthcare and health checks, such as doctor fees, lab fees, hospitalisation, prescription, travel, and insurance fees.
Assets	All expenditures on house equipment, such as bikes, motorbike, car, boat, phone, air conditioner, and washing machine.
House repairs	All costs for repairing and maintaining the house.
Food	All expenditures on food and drink.
Non-food	All expenditures on non-food consumption categories.
Utilities	All expenditures on water, electricity, and waste.

Following previous empirical studies on remittances in developing countries, we include the following numerical covariates in the logit model: household size, number of members with a high-school degree or above, age of the household head and its squared mean-centred term (to observe a nonlinear relationship), number of elderly members over 70 years old, number of children below 6 years old, and number of children between 6 and 14 years old. We also explore the effect of the following categorical covariates: living area of the household (urban/rural), ethnicity of the household head (Kinh/minorities), marital status of the household head (married/otherwise), and the six regions of Vietnam (Red River Delta, Midlands and Northern Mountainous Areas, Northern and Coastal Central Region, Central Highlands, South-Eastern Area, and Mekong River Delta). We specify this region covariate by means of five dummy variables in the logit model, with Mekong River Delta as the base region group.

The characteristics of the households with and without remittances can be found in Table 2. Households receiving remittances tended to have a smaller size, lower educational levels, more elderly members, and fewer children than those not receiving remittances. In addition, households receiving remittances were more likely to be rural and of Kinh ethnicity, and to have an older and nonmarried household head. Furthermore, there were some differences in the regional distribution of both types of households; for example, the South-Eastern Area was characterised by a relatively small proportion of households receiving remittances.

Descriptive statistics of adjusted income, remittances, saving amount and rate, and expenditures (both shares and per capita amounts) for the whole sample, as well as for families with and without remittances, are shown in Table 3. For households receiving remittances, these constituted about 15% of their adjusted income. Nevertheless, their adjusted income, total expenditure, and per capita expenditure were on average lower than those of households who did not receive remittances. However, receiving households appeared to save more, as shown by a somewhat higher saving amount and saving rate. Lastly, concerning expenditure patterns, receiving households tended to spend more on health, assets, and house repairs, while non-receiving families tended to spend more on education, food, non-food, and utilities.

Table 2. Descriptive statistics of the sample characteristics.

	Whole Sample		Households with Remittances		Households without Remittances	
	(N = 8778 and W = 21,870,190)		(N = 2174 and W = 5,427,473)		(N = 6604 and W = 16,442,717)	
Numerical variables	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Household size	3.836	1.516	3.478	1.686	3.954	1.437
Number of members with at least high-school degree	0.881	1.128	0.738	1.017	0.928	1.158
Age of the household head	49.736	13.817	58.116	11.748	46.969	13.324
Number of elderly members	0.210	0.498	0.329	0.595	0.171	0.454
Number of children 0–5 years	0.342	0.589	0.244	0.512	0.375	0.609
Number of children 6–14 years	0.539	0.760	0.283	0.597	0.623	0.789
Categorical variables (%)						
Living area household						
Urban	30.79		20.99		34.02	
Rural	69.21		79.01		65.98	
Ethnicity household head						
Kinh	87.95		91.37		86.81	
Minor ethnicity	12.05		8.63		13.19	
Marital status household head						
Married	81.93		78.24		83.14	
Otherwise	18.07		21.76		16.86	
Region of household living						
Red River Delta	24.92		28.36		23.78	
Midlands and Northern Mountainous Areas	12.54		12.39		12.59	
Northern and Coastal Central Region	22.39		26.96		20.88	
Central Highlands	5.17		2.92		5.91	
South-Eastern Area	16.81		8.75		19.47	
Mekong River Delta	18.17		20.62		17.37	

Note: N = actual sample size; W = total sample weight.

Table 3. Descriptive statistics for adjusted income, remittances, saving behaviour, and expenditure with sample weights.

	Whole Sample		Households with Remittances		Households without Remittances	
Numerical Variables	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Adjusted income (1000 VND)	99,406.460	99,510.700	89,412.370	80,718.890	102,705.400	104,774.100
Remittances (1000 VND)	3259.227	15,572.290	13,133.170	29,116.130		
Saving amount (1000 VND)	22,027.620	71,713.220	22,227.710	67,708.120	21,961.570	72,991.840
Saving rate	−0.005	1.456	0.059	0.642	−0.026	1.638
Total expenditure (1000 VND)	77,378.850	61,546.320	67,184.660	56,730.740	80,743.780	62,696.420
Share of education expenditure (%)	0.044	0.069	0.034	0.064	0.048	0.070
Share of health expenditure (%)	0.048	0.079	0.064	0.096	0.043	0.072
Share of assets expenditure (%)	0.040	0.089	0.041	0.095	0.040	0.087
Share of house repairs (%)	0.012	0.056	0.015	0.066	0.010	0.052
Share of food expenditure (%)	0.542	0.129	0.536	0.133	0.544	0.128
Share of non-food expenditure (%)	0.280	0.096	0.278	0.101	0.280	0.095
Share of utilities (%)	0.034	0.033	0.031	0.025	0.035	0.035
Per capita expenditure (PCE) (1000 VND)	21,329.770	17,054.220	20,383.920	16,176.260	21,641.980	17,324.120
PCE for education (1000 VND)	1086.043	4559.845	718.995	1839.826	1207.199	5145.896
PCE for health (1000 VND)	1072.264	2672.007	1420.606	3260.805	957.283	2436.128
PCE for assets (1000 VND)	1223.634	6611.444	1426.143	10,424.190	1156.789	4718.759
PCE for house repairs (1000 VND)	3582.293	2255.531	537.692	3030.480	298.992	1929.331
PCE for food (1000 VND)	10,704.140	6671.854	10,007.310	5234.823	10,934.150	7067.761
PCE for non-food (1000 VND)	6111.310	6252.114	5649.425	4550.661	6263.771	6713.151
PCE for utilities (1000 VND)	99,406.460	99,510.700	623.750	787.180	823.795	1278.468

Note: VND = Vietnamese dong.

3. Results

We proceed in three steps. First, we present the results of the propensity score regression. Next, we derive the region of common support and compare the density distributions before and after matching. In the third and final step, we analyse to what extent receiving remittances influences saving behaviour and expenditure patterns.

3.1. Results of the Estimated Propensity Score by Logit Regression

Table 4 presents two models for the estimation of the propensity score based on the selected household characteristics. In an initial model, we included all observed covariates. However, the effects of the number of elderly members and of the number of children below 6 years old on the propensity score turned out to be insignificant at the 5% level. Therefore, we excluded these two variables from the final model, to only show the impact of the significant characteristics.

Table 4. Initial and final logit model for propensity score estimation.

Variable	Coefficients Initial Model (Chi-Square Value)	Coefficients Final Model (Chi-Square Value)
Age of the household head	0.101 *** (677.99)	0.101 *** (700.42)
Squared age of the household head (mean-centred)	−0.003 *** (208.09)	−0.002 *** (252.88)
South-Eastern Area	−0.625 *** (31.04)	−0.627 *** (31.28)
Central Highlands	−0.529 *** (13.05)	−0.527 *** (12.95)
Northern and Coastal Central Region	0.162 ** (3.88)	0.165 ** (4.00)
Midlands and Northern Mountainous Areas	0.169 * (2.71)	0.178 * (3.00)
Red River Delta	0.004 (0.00)	0.015 (0.03)
Urban	−0.491 *** (53.99)	−0.488 *** (53.49)
Number of members with at least high-school degree	−0.193 *** (41.81)	−0.200 *** (45.94)
Married household head	0.389 *** (25.70)	0.388 *** (26.70)
Number of children 6–14 years	−0.249 *** (21.25)	−0.273 *** (28.48)
Household size	−0.097 *** (12.80)	−0.074 *** (12.08)
Kinh household head	0.287 *** (8.17)	0.294 *** (8.61)
Number of elderly members	0.116 (2.64)	/
Number of children 0–5 years	0.079 (1.41)	/
Constant	−5.791 *** (533.53)	−5.857 *** (559.30)
Pseudo R^2	0.189	0.189
−Log-likelihood (full model–constant model)	928 *** (1856.23)	926 *** (1852.70)
Observations, N	8778	8778

Note: *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The reference household of the model is a household in a rural area of the Mekong River Delta, where the household head is not married and of minor ethnicity.

The importance of each explanatory variable was estimated by means of the log-worth, defined as $-\log_{10}$(p-value of the F-test) (Kessels and Erreygers 2019). In Figure 1, the horizontal bar graph depicts the log-worth values of the variables relative to the most important variable, age of the household head, for which the log-worth values of its main and squared term combined are normalised to 100%. The age of the household head and the squared age, thus, had the greatest effect on the propensity score, followed by region, urban living area, education, older children, marital status of the household head, household size, and ethnicity of the household head.

Source	Logworth	PValue
Age of the household head	241.650	0.00000
Squared age of the household head (mean-centred)	66.219	0.00000
Region of household living	14.786	0.00000
Urban living area household	12.896	0.00000
Number of members with at least high-school degree	11.177	0.00000
Number of children 6–14 years	7.217	0.00000
Married household head	6.755	0.00000
Household size	3.328	0.00047
Kinh household head	2.511	0.00309

Figure 1. Importance of the explanatory variables to the propensity score estimates obtained from logit regression.

The overall goodness of fit, as measured by the pseudo R^2, indicates that the observed characteristics explain 18.9% of the propensity score. There is no threshold for this number in the PSM approach. Most of the previous papers set it around 10%, e.g., 10% and 8% in the research of Clément (2011) and Démurger and Wang (2016), respectively. Hence, the explanatory power of the logit model for the estimated propensity score in our study seems satisfactory. In addition, to check whether there is any indication of missing variables in our logit regression, we performed a lack-of-fit test. The test result (p = 0.9997) suggests that there is no immediate evidence of missing variables. Note that the PSM approach does not aim to maximise the fit of the model, but uses the propensity score as a balancing mechanism (Randazzo and Piracha 2019).

The final model in Table 4 suggests the following results:

- The probability of receiving remittances increases steeply with the age of the household head until the age of 70 is reached, after which the probability decreases.
- There are significant differences in the probability of receiving remittances between households in different regions. According to the coefficients of the different regions, the probability of households receiving remittances is the highest for the Midlands and Northern Mountainous Areas, followed by the Northern and Coastal Central Region, the Red River Delta, the Mekong River Delta, the Central Highlands, and the South-Eastern Area. However, the difference in the probability of receiving remittances of households in the Mekong River Delta and households in the Red River Delta is not significant.
- Rural households have a higher probability of receiving remittances than urban households.
- As far as education is concerned, the probability of receiving remittances depends negatively on the number of well-educated members. Migrants from a well-educated household could have less strong motives to send remittances to support their home families.
- With regard to the marital status of household head, married household heads have a higher probability of receiving remittances than the others.
- Older children negatively affect the probability of receiving remittances. Migrant members could have less responsibility to support their home families in cases where these consist of older children. As in Hua and Erreygers (2020), we considered older children as belonging to the household labour force, and not as dependent members, as in other empirical papers. This result confirms the role of older children as labourers in households.
- The probability of receiving remittances depends negatively on household size. This implies that small families tend to receive remittances more often than larger families.
- Lastly, the effect of the ethnicity covariate reveals that the Kinh have a higher probability of receiving remittances than other ethnic groups. This result supports the conclusion of Nguyen and Vu (2018) and Coxhead et al. (2019), who found that people from minor ethnicities were less likely to migrate than Kinh people.

3.2. Defining the Common Support for Propensity Scores of Treated and Non-Treated Groups

Table 5 contains the minima and maxima of the estimated propensity scores for treated and non-treated households. The range of propensity scores for both groups is largely overlapping with a region of common support ranging from 0.0017 to 0.7048. Any treated household with a propensity score outside this range was excluded.

Table 5. Descriptive statistics of propensity score estimates from the final logit model.

Household Group	*N*	Mean	Std. Dev.	Min	Max
With remittances	2174	0.3965	0.1662	0.0017	0.7063
Without remittances	6604	0.1987	0.1732	0.0007	0.7048

The top left panel of Figure 2 plots the densities of the propensity scores of the treated and non-treated households before matching. The density plot for the treated households shows an inverted U-shaped distribution, while, for the non-treated households, the distribution is right-skewed. Hence, there is a substantial difference between the two distributions before matching. After matching, the distributions are nearly equivalent, as shown in the remaining panels of Figure 2.

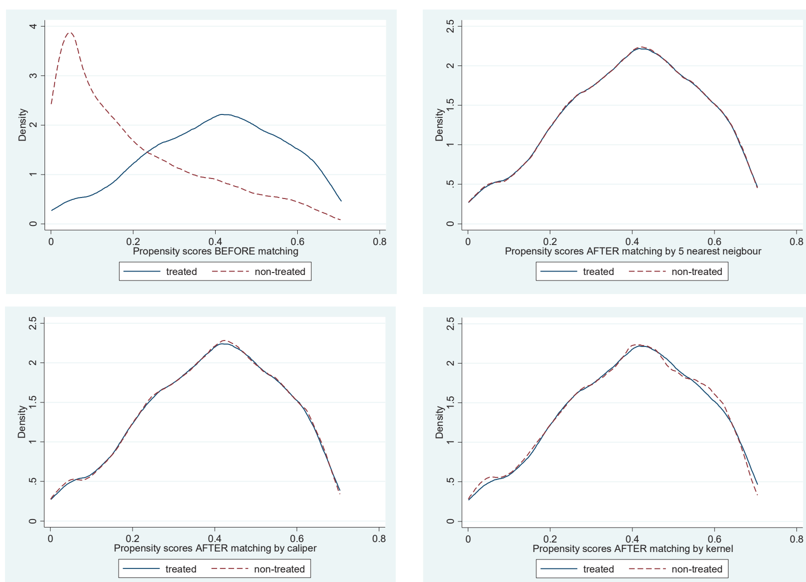

Figure 2. Propensity score distribution before and after matching.

3.3. Impact of Remittances on Saving Behaviour and Expenditure Patterns

3.3.1. Effect of Remittances on Saving, Adjusted Income, and Total Expenditure

The *ATT* estimates presented in Table 6 are the differences in saving amount and rate, adjusted income, and total expenditure between households with and without remittances. As far as the saving amount and rate before matching are concerned, we find that the effect of remittances is only significant for the saving rate. After matching, however, the results of all matching estimators show that the differences in both saving amount and rate between treated and non-treated households are significant. Mainly, households with remittances tend to save more in level (between 4.6 and 4.9 million VND), as well as in rate (between 16.8% and 20.7%) than those without remittances. These results imply that remittances impact saving behaviour positively. Moreover, the significant results for all matching estimators confirm the robustness of our findings.

Table 6. *ATT* estimates of the effect of remittances on saving, adjusted income and total expenditure.

	Treated	Non-Treated	Difference	S.E.	*t*-Statistic
Saving amount					
Unmatched	21,322.069	18,775.257	2546.813	1680.144	1.52
5NN	21,353.926	16,441.609	4912.318	2000.714	2.46 **
Radius	21,362.624	16,712.992	4649.633	2044.478	2.27 **
Kernel	21,353.926	16,537.695	4816.232	1932.336	2.49 **
Saving rate					
Unmatched	0.047	−0.049	0.096	0.039	2.44 **
5NN	0.048	−0.125	0.173	0.055	3.17 ***
Radius	0.048	−0.119	0.168	0.041	4.14 ***
Kernel	0.048	−0.159	0.207	0.037	5.66 ***
Adjusted income					
Unmatched	89,125.092	96,554.327	−7429.236	2291.376	−3.24 ***
5NN	89,214.019	83,375.030	5838.989	2617.316	2.23 **
Radius	89,509.807	84,803.818	4705.989	2651.937	1.77 *
Kernel	89,214.019	83,836.196	5377.824	2483.378	2.17 **
Total expenditure					
Unmatched	67,803.022	77,779.071	−9976.049	1486.577	−6.71 ***
5NN	67,860.093	66,933.422	926.670	1753.725	0.53
Radius	68,147.183	68,090.826	56.356	1780.121	0.03
Kernel	67,860.093	67,298.501	561.592	1681.281	0.33

Note: *, **, and *** indicate significance at the 10% level ($t > 1.645$), 5% level ($t > 1.96$), and 1% level ($t > 2.575$), respectively.

The high savings of remittance-receiving households could be caused by their higher income. Table 6 shows that the adjusted income of non-receiving households is significantly higher both before and after matching (between 4.7 and 5.8 million VND after matching). This finding supports the evidence of Nguyen and Vu (2018) that remittances help reduce poverty for the receiving households. By contrast, as Table 6 also shows, the difference in total expenditure between remittance receiving and non-receiving households is insignificant after matching.

3.3.2. Effect of Remittances on Household Expenditure Patterns

With respect to shares of expenditure, Table 7 shows a significant difference in the shares of expenditure on health, assets, house repairs, and food between households with and without remittances after matching. Receiving households spend significantly more on health (between 0.9% and 1.0% of their total expenditure), assets (between 0.6% and 0.7%), and house repairs (between 0.3% and 0.4%), but less on food (between 2.0% and 2.2%).

As far as per capita expenditure is concerned, we observe in Table 8 that receiving households have a higher per capita expenditure than non-receiving households, al-though the difference is not significant according to the radius estimator. Nguyen and Vu (2018) also found that remittances help receiving households increase per capita consumption. Furthermore, Table 8 shows significantly higher spending on health (between 228 and 243 thousand VND), assets (between 496 and 537 thousand VND), and house repairs (between 215 and 236 thousand VND) by receiving households compared with non-receiving households. These differences in expenditure patterns provide evidence that remittances tend to be used productively for human and physical capital investment, since spending on health, assets, and house repairs can be considered as investment expenditures (Taylor and Mora 2006; de Brauw and Rozelle 2008).

Table 7. *ATT* estimates for the impact of remittances on the share of expenditure.

	Treated	Non-Treated	Difference	S.E.	*t*-Statistic
Education					
Unmatched	0.033	0.046	−0.013	0.002	−7.92 ***
5NN	0.033	0.033	0.000	0.002	0.22
Radius	0.033	0.033	0.000	0.002	0.08
Kernel	0.033	0.033	0.000	0.002	0.05
Health					
Unmatched	0.065	0.043	0.022	0.002	11.17 ***
5NN	0.065	0.055	0.010	0.003	3.50 ***
Radius	0.065	0.056	0.010	0.003	3.68 ***
Kernel	0.065	0.056	0.009	0.003	3.49 ***
Assets					
Unmatched	0.043	0.043	0.000	0.002	0.03
5NN	0.043	0.036	0.006	0.003	2.24 **
Radius	0.043	0.037	0.006	0.003	2.04 **
Kernel	0.043	0.036	0.007	0.003	2.47 **
House repairs					
Unmatched	0.015	0.011	0.005	0.001	3.36 ***
5NN	0.015	0.012	0.003	0.002	1.80 *
Radius	0.015	0.011	0.004	0.002	2.21 **
Kernel	0.015	0.012	0.004	0.002	2.19 **
Food					
Unmatched	0.537	0.548	−0.012	0.003	−3.63 ***
5NN	0.536	0.559	−0.022	0.004	−5.34 ***
Radius	0.536	0.555	−0.020	0.004	−4.93 ***
Kernel	0.536	0.559	−0.022	0.004	−5.78 ***
Non-food					
Unmatched	0.278	0.278	0.000	0.002	0.09
5NN	0.278	0.275	0.003	0.003	1.07
Radius	0.278	0.277	0.001	0.003	0.43
Kernel	0.278	0.274	0.004	0.003	1.30
Utilities					
Unmatched	0.030	0.032	−0.002	0.001	−2.93 ***
5NN	0.030	0.031	−0.001	0.001	−1.32
Radius	0.030	0.031	−0.001	0.001	−1.44
Kernel	0.030	0.031	−0.001	0.001	−1.37

Note: *, **, and *** indicate significance at the 10% level ($t > 1.645$), 5% level ($t > 1.96$), and 1% level ($t > 2.575$), respectively.

The higher spending on health by receiving households is consistent with the results of other studies on the role of remittances in developing countries such as Taylor and Mora (2006), Ponce et al. (2011), Wen and Lin (2012), and Berloffa and Giunti (2019). This remittance-induced expenditure on health confirms the altruism motive of migrant members who send money home to care for other family members. This also creates more opportunities for the development of hospitals and healthcare centres in the local economy. Furthermore, expenditures on assets and house repairs can improve the quality of life of the households, e.g., by expanding family business facilities for informal self-employed workers (Adams and Cuecuecha 2010). This is especially relevant for Vietnam, where there are large numbers of informal self-employed workers and small-scale family businesses.

Table 8. *ATT* estimates for the impact of remittances on per capita expenditure.

	Treated	Non-Treated	Difference	S.E.	*t*-Statistic
Expenditure per capita					
Unmatched	20,331.619	20,633.493	−301.874	413.994	−0.73
5NN	20,341.403	19,483.273	858.130	477.386	1.80 *
Radius	20,368.527	19,768.648	599.879	498.593	1.20
Kernel	20,341.403	19,563.238	778.165	472.459	1.65 *
Education					
Unmatched	700.610	1129.769	−429.158	112.147	−3.83 ***
5NN	701.579	737.116	−35.538	64.935	−0.55
Radius	708.431	798.690	−90.258	115.047	−0.78
Kernel	701.579	782.200	−80.622	103.507	−0.78
Health					
Unmatched	1448.926	935.685	513.241	68.120	7.53 ***
5NN	1449.837	1207.229	242.608	95.503	2.54 **
Radius	1457.375	1218.209	239.166	90.199	2.65 ***
Kernel	1449.837	1221.525	228.311	86.381	2.64 ***
Assets					
Unmatched	1488.071	1220.869	267.201	175.100	1.53
5NN	1489.008	952.164	536.843	247.270	2.17 **
Radius	1491.888	995.499	496.390	258.983	1.92 **
Kernel	1489.008	969.065	519.942	251.939	2.06 **
House repairs					
Unmatched	536.503	307.776	228.727	57.022	4.01 ***
5NN	537.245	321.867	215.378	79.441	2.71 ***
Radius	519.236	303.151	216.086	74.544	2.90 ***
Kernel	537.245	301.387	235.858	75.382	3.13 ***
Food					
Unmatched	9956.491	10,489.693	−533.203	159.427	−3.34 ***
5NN	9958.977	10,202.106	−243.129	184.490	−1.32
Radius	9973.769	10,244.294	−270.525	180.117	−1.50
Kernel	9958.977	10,184.127	−225.150	169.063	−1.33
Non-food					
Unmatched	5614.585	5847.380	−232.795	143.296	−1.62
5NN	5618.003	5455.790	162.212	163.256	0.99
Radius	5629.198	5592.111	37.086	161.619	0.23
Kernel	5618.003	5500.074	117.929	150.432	0.78
Utilities					
Unmatched	586.433	702.320	−115.888	24.676	−4.70 ***
5NN	586.756	607.001	−20.245	23.182	−0.87
Radius	588.630	616.696	−28.065	25.842	−1.09
Kernel	586.756	604.858	−18.103	24.316	−0.74

Note: *, **, and *** indicate significance at the 10% level ($t > 1.645$), 5% level ($t > 1.96$), and 1% level ($t > 2.575$), respectively.

As far as food is concerned, we find that remittance-receiving households have a lower expenditure share compared with non-receiving households. Nevertheless, per capita expenditures on food are not significantly different between the two groups. This suggests that remittances do not increase the demand for food.

In summary, we observe that remittances not only increase household saving, but also have a positive effect on investment categories such as human and physical capital investment. Our findings are consistent with those in the studies of Nguyen (2008), and Nguyen and Mont (2012). These results reveal that remittances tend to be used productively and, therefore, have a positive effect on the growth and development of households and the economy in Vietnam.

4. Discussion

In this section, we discuss the quality of our matching estimators and the sensitivity of our results. Since some of the material is quite technical, we relegated most of it to the Supplementary Materials.

4.1. Assessing the Quality of the Matching

As explained in more detail in Part I of the Supplementary Materials, two types of tests can be conducted to check the quality of the matching process. The first type looks at the difference in the distribution of the individual covariates between the treated and the control group after matching. We consider three specific tests and report the results in Tables S1–S3 of the Supplementary Materials. First, we use a *t*-test to check whether there is a difference in the means of a covariate between households with and without remittances after matching. Since the *t*-tests after matching are all insignificant, this is evidence of covariate balance between the treated and the control group. Second, we calculate the absolute standardised bias (ASB) for each covariate between households with and without remittances after matching. The bias percentages for the different estimators, which are all below 5%, with a maximum of 4.1% for the kernel matching estimator, indicate sufficient covariate balance. Third, for the continuous covariates, we report the ratios of the variance in the treated group to that in the control group. After matching, these ratios are closer to unity compared with the ratios in the unmatched sample, providing further evidence of proper balancing.

The second type re-estimates the propensity score by logit regression for the matched sample and compares the joint explanatory power of all covariates to that of the initial regression from the unmatched sample. We report the results for the three matching estimators in Table S4 of the Supplementary Materials. First, while the pseudo R^2 value is 18.9% before matching, it is only 0.1% after matching, indicating that the observed covariates can explain very little of the propensity score in the matched sample. Accordingly, the new regressions of the matched sample reveal that the likelihood ratio (LR) tests on the joint significance of all covariates are insignificant. This confirms that the distributions of the propensity scores for the treated and non-treated households after matching are balanced, as shown in Figure 2. In addition, we consider the mean and median of the absolute standardised difference or bias in the covariates across the treated and non-treated households. The means and medians for the unmatched sample (before matching) are rather high (25% and 18%), while, for the matched sample, they are low (between 1.3% and 1.6%). The kernel estimator has the lowest mean and median (1.5% and 1.3%), thus reducing the bias the most. This is line with Garrido et al. (2014) who found that the kernel estimator had the lowest bias in their empirical study.

In summary, the results of all the tests confirm that the balancing property is satisfied. Thus, the distributions of households with and without remittances after matching are equivalent. Therefore, the PSM approach can be applied to estimate the impact of remittances on saving behaviour and expenditure patterns.

4.2. Analysing Sensitivity

Since the *ATT* estimates for the outcomes obtained from PSM are based on the observed covariates only, other unobserved covariates are assumed not to impact these outcomes. However, a sensitivity test is needed to investigate whether the average treatment effects are sensitive to the influence of unobserved covariates or hidden biases. As described in Part II of the Supplementary Materials, the bounding approach proposed by Rosenbaum (2002) allows us to analyse the sensitivity of the *ATT* estimates. In this approach, the sensitivity of the *ATT* to hidden bias is indicated using the critical value of the odds ratio. A higher critical value indicates the greater extent to which an unobserved confounder would have to alter the odds of receiving remittances to completely determine the *ATT*. If this critical value is relatively low, i.e., smaller than 2, the *ATT* is likely sensitive to hidden bias (Clément 2011; Li 2012).

Table S5 of the Supplementary Materials shows that the critical values producing a 95% confidence interval including zero fluctuate around 1 to 1.2 for the saving amount and adjusted income for the three matching estimators. This implies that households with the same observed characteristics can differ in their odds of receiving remittances by as much as 20% before the confidence interval on the *ATT* starts including zero. That is to say, an unobserved characteristic would have to increase the odds ratio by at most 20% before it could bias the *ATT*. Saving amount and adjusted income are, therefore, susceptible to hidden bias. Nevertheless, this finding does not mean we have to reject the impact of remittances on the saving level and adjusted income of households. It means that, if there were an unobserved characteristic, the result could be different. In other words, it is considered a 'worst-case scenario' that could happen if an unobserved covariate caused the odds ratio of treatment assignment to differ between the treated and non-treated groups (Clément 2011; Li 2012). In this research, we defined remittances as receipts of households from migrant members, but other factors affecting income of those members could have an impact on household remittances. Due to lack of information in the dataset, we could not observe these covariates in our study. For the saving rate, the critical odds ratio is at least equal to 2, highlighting that the impact of remittances on the household saving rate is robust to the presence of unobserved characteristics.

Concerning the shares of expenditure, Table S6 of the Supplementary Materials reveals that the *ATT* of house repairs is robust to hidden bias since the critical odds ratios are larger than 2 for all matching estimators. However, there is no robustness guarantee for the other outcomes, health, assets, and food, as indicated by critical odds ratios smaller than 2. Regarding per capita expenditures, Table S7 shows that the *ATT* estimates of assets and house repairs are robust, while the effect of health is likely sensitive to hidden bias.

The critical odds ratios of saving behaviour and expenditure patterns in our research are of the same magnitude as those identified in other studies (Clément 2011; Li 2012). Hence, the impact of remittances in our study has the same degree of sensitivity to unobservables as in other studies.

5. Conclusions

This study was an econometric analysis to analyse the impact of remittances on the saving and expenditure behaviour of Vietnamese households. We applied the PSM approach proposed by Rosenbaum and Rubin (1983) to the VHLSS 2012 dataset, which allowed us to avoid the endogeneity problem when investigating the impact of remittances using the Working–Leser Engel method. PSM has been widely used in the study of causal treatment effects. The impact of remittances on saving behaviour and expenditure patterns is determined by estimating the average treatment effect on the treated (*ATT*) for the outcomes.

We found that households with remittances tend to have a higher saving amount and rate than those without, while expenditures per capita of the two household groups did not differ significantly. As far as expenditure patterns are concerned, remittances were used productively in human and physical capital investment. We also observed that households receiving remittances had a significantly lower expenditure share on food. This implies that remittances stimulated investment rather than consumption. Our findings are consistent with those of Nguyen (2008), as well as Nguyen and Mont (2012), who studied the impact of external and internal remittances separately. Our study adds to the existing literature on the effect of remittances by considering both external and internal remittances together.

Households receiving remittances were unlikely to consider their remittances as expected and stable income. This suggests that remittances were treated as transitory income, in accordance with the permanent income hypothesis. Since remittances were used mainly for investment in human and physical capital, they created more opportunities for the development of services provided by banks, financial institutions, hospitals, and healthcare centres, also in addition to providing incentives for the production of building materials and tangible assets. Nevertheless, consistent with the study of Nguyen et al. (2017), we found that remittances did not influence education. Therefore, we agree with their suggestion that other capital sources should be considered for improving education, especially in rural areas in Vietnam. More generally, as Yoshino et al. (2020) found, the importance of remittances for investment in human and physical capital tends to diminish as countries become richer. Remittances are gradually replaced by other forms of capital inflows, such as foreign direct investment (FDI). To what extent this applies to a country like Vietnam, which has grown quickly from the group of low-income countries to that of middle-income countries, remains to be seen. As long as the flow of remittances is relatively large, policymakers should encourage the productive use of these resources, e.g., by improving the efficiency of the banking system. In the long run, however, they must take into account that the role of remittances will probably become smaller and smaller.

There are still some limitations to our research. Firstly, we were not able to investigate the impact of external and internal remittances separately, due to the small size of foreign receipts (only 159 households). Secondly, the matching method is based on the propensity score estimated from the observed covariates. The *ATT* does not reflect the possible effect of unobserved factors. Even when the impact of the observed covariates is significant, there might still be some variables that remain unexplored. Lastly, while the PSM approach allows us to test whether remittances have an impact on saving behaviour and expenditure patterns, it does not reveal the extent of the effect. This means that other methods should be applied to estimate the magnitude of the effect.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/economies10090223/s1: Additional explanation and Tables S1–S7 on the tests of the matching quality (Part I) and the robustness (Part II).

Author Contributions: Conceptualization, T.X.H.; methodology, T.X.H. and R.K.; validation, T.X.H. and R.K.; data curation, T.X.H.; writing—original draft preparation, T.X.H.; writing—review and editing, R.K. and G.E.; supervision, R.K. and G.E. All authors read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: This study used data from the Vietnam Household Living Standard Survey (VHLSS) for the year 2012. This dataset was collected by the General Statistics Office of Vietnam (https://www.gso.gov.vn/en/homepage/) (accessed on 23 December 2019).

Acknowledgments: Thanh Xuan Hua acknowledges the support of the Ministry of Education and Training in Vietnam which allowed her to work on this paper as part of her PhD research. Roselinde Kessels acknowledges the JMP Division of SAS Institute for financial support. We are grateful to the reviewers of this journal for their comments and suggestions.

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

References

Aakvik, Arild. 2001. Bounding a matching estimator: The case of a Norwegian training program. *Oxford Bulletin of Economics and Statistics* 63: 115–43.

Adams, Richard H., Jr., and Alfredo Cuecuecha. 2010. Remittances, household expenditure and investment in Guatemala. *World Development* 38: 1626–41. [CrossRef]

Adams, Richard H., Jr., Alfredo Cuecuecha, and John Page. 2008. *Remittances, Consumption and Investment in Ghana*. Policy Research Working Paper No. 4515. Washington, DC: World Bank.

Aggarwal, Reena, Asli Demirgüç-Kunt, and Maria Soledad Martínez Pería. 2011. Do remittances promote financial development? *Journal of Development Economics* 96: 255–64. [CrossRef]

Ait Benhamou, Zouhair, and Lesly Cassin. 2021. The impact of remittances on savings, capital and economic growth in small emerging countries. *Economic Modelling* 94: 789–803. [CrossRef]

Ang, Alvin P., Guntur Sugiyarto, and Shikha Jha. 2009. *Remittances and Household Behavior in the Philippines*. Asian Development Bank Economics Working Paper No. 188. Manila: Asian Development Bank. Available online: https://www.adb.org/publications/remittances-and-household-behavior-philippines (accessed on 15 July 2022).

Austin, Peter C. 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine* 28: 3083–107. [CrossRef] [PubMed]

Becker, Sascha O., and Marco Caliendo. 2007. Sensitivity analysis for average treatment effects. *The Stata Journal* 7: 71–83. [CrossRef]

Berloffa, Gabriella, and Sara Giunti. 2019. Remittances and healthcare expenditure: Human capital investment or responses to shocks? Evidence from Peru. *Review of Development Economics* 23: 1540–61. [CrossRef]

Caliendo, Marco, and Sabine Kopeinig. 2008. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22: 31–72. [CrossRef]

Cardona-Sosa, Lina, and Carlos Medina. 2006. *Migration as a Safety Net and Effects of Remittances on Household Consumption: The Case of Colombia*. Borradores de Economía No. 414. Bogotá: Banco de la Republica de Colombia.

Castaldo, Adriana, and Barry Reilly. 2007. Do migrant remittances affect the consumption patterns of Albanian households? *South-Eastern Europe Journal of Economics* 5: 25–44.

Clément, Matthieu. 2011. Remittances and household expenditure patterns in Tajikistan: A propensity score matching analysis. *Asian Development Review* 28: 58–87. [CrossRef]

Coxhead, Ian, Viet Cuong Nguyen, and Hoang Linh Vu. 2019. Internal migration in Vietnam: 2002–2012. In *Rural-Urban Migration in Vietnam*. Edited by Amy Y. C. Liu and Xin Meng. Cham: Springer, pp. 67–96.

Deaton, Angus S. 2019. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington, DC: World Bank.

Deaton, Angus S., Javier Ruiz-Castillo, and Duncan Thomas. 1989. The influence of household composition on household expenditure patterns: Theory and Spanish evidence. *Journal of Political Economy* 97: 179–200. [CrossRef]

de Brauw, Alan, and Scott Rozelle. 2008. Migration and household investment in rural China. *China Economic Review* 19: 320–35. [CrossRef]

Démurger, Sylvie, and Xiaoqian Wang. 2016. Remittances and expenditure patterns of the left behinds in rural China. *China Economic Review* 37: 177–90. [CrossRef]

Esquivel, Gerardo, and Alejandra Huerta-Pineda. 2007. Remittances and poverty in Mexico: A propensity score matching approach. *Integration and Trade* 27: 45–71.

Friedman, Milton. 1957. *A Theory of the Consumption Function*. Princeton: National Bureau of Economic Research, Inc.

Gangl, Markus. 2004. RBOUNDS: Stata Module to Perform Rosenbaum Sensitivity Analysis for Average Treatment Effects on the Treated. Statistical Software Components S438301, Boston College Department of Economics. Available online: https://ideas.repec.org/c/boc/bocode/s438301.html (accessed on 28 August 2022).

Garrido, Melissa M., Amy S. Kelley, Julia Paris, Katherine Roza, Diane E. Meier, R. Sean Morrison, and Melissa D. Aldridge. 2014. Methods for constructing and assessing propensity scores. *Health Services Research* 49: 1701–20. [CrossRef]

Haider, Mohammed Ziaul, Tanbir Hossain, and Ohidul Islam Siddiqui. 2016. Impact of remittance on consumption and savings behavior in rural areas of Bangladesh. *Journal of Business* 1: 25–34. [CrossRef]

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies* 64: 605–54. [CrossRef]

Heckman, James J., Hidehiko Ichimura, Jeffrey Smith, and Petra E. Todd. 1998. Characterizing selection bias using experimental data. *Econometrica* 66: 1017–98. [CrossRef]

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15: 199–236. [CrossRef]

Hua, Thanh Xuan, and Guido Erreygers. 2020. Applying quantile regression to determine the effects of household characteristics on household saving rates in Vietnam. *Journal of Asian Business and Economic Studies* 27: 175–93. [CrossRef]

Jimenez-Soto, Eliana V., and Richard P. C. Brown. 2012. Assessing the poverty impacts of migrants' remittances using propensity score matching: The case of Tonga. *Economic Record* 88: 425–39. [CrossRef]

Junge, Vera, Javier Revilla Diez, and Ludwig Schätzl. 2015. Determinants and consequences of internal return migration in Thailand and Vietnam. *World Development* 71: 94–106. [CrossRef]

Kessels, Roselinde, and Guido Erreygers. 2019. A direct regression approach to decomposing socioeconomic inequality of health. *Health Economics* 28: 884–905. [CrossRef]

Leser, Conrad E. V. 1963. Forms of Engel functions. *Econometrica* 31: 694–703. [CrossRef]

Leuven, Edwin, and Barbara Sianesi. 2018. PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. Available online: https://EconPapers.repec.org/RePEc:boc:bocode:s432001 (accessed on 15 July 2022).

Li, Mingxiang. 2012. Using the propensity score method to estimate causal effects: A review and practical guide. *Organizational Research Methods* 16: 188–226. [CrossRef]

Liu, Amy Y. C., and Xin Meng, eds. 2019. *Rural-Urban Migration in Vietnam*. Cham: Springer.

Luong, Hy V. 2018. The changing configuration of rural-urban migration and remittance flows in Vietnam. *Sojourn: Journal of Social Issues in Southeast Asia* 33: 602–46. [CrossRef]

McKenzie, David, and Marcin J. Sasin. 2007. *Migration, Remittances, Poverty, and Human Capital: Conceptual and Empirical Challenges*. Policy Research Working Paper No. 4272. Washington, DC: World Bank.

McKenzie, David, Steven Stillman, and John Gibson. 2010. How important is selection? Experimental vs. non-experimental measures of the income gains from migration. *Journal of the European Economic Association* 8: 913–45. [CrossRef]

Ministry of Foreign Affairs of Vietnam. 2012. *Review of Vietnamese Migration Abroad*; Ha Noi: Consular Department, Ministry of Foreign Affairs of Vietnam.

Modigliani, Franco, and Richard Brumberg. 1954. Utility analysis and the consumption function: An interpretation of cross-section data. In *Post-Keynesian Economics*. Edited by Kenneth K. Kurihara. New Brunswick: Rutgers University Press, pp. 388–436.

Nguyen, Duc Loc, Ulrike Grote, and Trung Thanh Nguyen. 2017. Migration and rural household expenditures: A case study from Vietnam. *Economic Analysis and Policy* 56: 163–75. [CrossRef]

Nguyen, Thu Phuong, Ngo Thi Minh Tam Tran, Thi Nguyet Nguyen, and Remco Oostendorp. 2008. *Determinants and Impacts of Migration in Vietnam*. Working Papers 2008/01. Ha Noi: Development and Policies Research Center (DEPOCEN).

Nguyen, Viet Cuong. 2008. Impacts of international and internal remittances on household welfare: Evidence from Viet Nam. *Asia-Pacific Development Journal* 16: 59–92. [CrossRef]

Nguyen, Viet Cuong, and Daniel Mont. 2012. Economic impacts of international migration and remittances on household welfare in Vietnam. *International Journal of Development Issues* 11: 144–63.

Nguyen, Viet Cuong, and Huang Linh Vu. 2018. The impact of migration and remittances on household welfare: Evidence from Vietnam. *Journal of International Migration and Integration* 19: 945–63.

Opiniano, Jeremaiah M. 2021. Remittances and the financial capabilities of migrant households in the Philippines. *Asian and Pacific Migration Journal* 30: 370–85. [CrossRef]

Ponce, Juan, Iliana Olivié, and Mercedes Onofa. 2011. The role of international remittances in health outcomes in Ecuador: Prevention and response to shocks. *International Migration Review* 45: 727–45. [CrossRef]

Quartey, Peter, Charles Ackah, and Monica Puoma Lambon-Quayefio. 2019. Inter-linkages between remittance and savings in Ghana. *International Journal of Social Economics* 46: 152–66. [CrossRef]

Randazzo, Teresa, and Matloob Piracha. 2019. Remittances and household expenditure behaviour: Evidence from Senegal. *Economic Modelling* 79: 141–53. [CrossRef]

Ratha, Dilip, Eung Ju Kim, Sonia Plaza, Elliott J. Riordan, and Vandana Chandra. 2022. *Migration and Development Brief 36: A War in a Pandemic: Implications of the Russian Invasion of Ukraine and the COVID-19 Crisis on Global Governance of Migration and Remittance Flows*. Washington, DC: KNOMAD-World Bank.

Rosenbaum, Paul R. 2002. *Observational Studies*. New York: Springer.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55. [CrossRef]

Rosenbaum, Paul R., and Donald B. Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 39: 33–38.

Salahuddin, Sarah, Muhammad Mehedi Masud, and Kian Teng Kwek. 2022. Remittances and households' savings behaviour in Bangladesh. *Society and Business Review* 17: 120–40. [CrossRef]

Sianesi, Barbara. 2004. An evaluation of the Swedish system of active labor market programs in the 1990s. *Review of Economics and Statistics* 86: 133–55. [CrossRef]

Stark, Oded, and David E. Bloom. 1985. The new economics of labor migration. *American Economic Review* 75: 173–78.

Tabuga, Aubrey D. 2010. International remittances and family expenditure patterns: The Philippines' case. *Philippine Journal of Development* 35: 103–23.

Taylor, J. Edward, and Jorge Mora. 2006. *Does Migration Reshape Expenditures in Rural Households? Evidence from Mexico*. Policy Research Working Paper No. 3842. Washington, DC: World Bank.

Wen, Ming, and Danhua Lin. 2012. Child development in rural China: Children left behind by their migrant parents and children of nonmigrant families. *Child Development* 83: 120–36. [CrossRef]

World Bank. 2012. *Well Begun, Not Yet Done: Vietnam's Remarkable Progress on Poverty Reduction and the Emerging Challenges*. Hanoi: World Bank.

World Bank. 2016. *Transforming Vietnamese Agriculture: Gaining More from Less*. Washington, DC: World Bank.

Working, Holbrook. 1943. Statistical laws of family expenditure. *Journal of the American Statistical Association* 38: 43–56. [CrossRef]

Yang, Dean. 2008. International migration, remittances and household investment: Evidence from Phillipine migrants' exchange rate shocks. *Economic Journal* 118: 591–630. [CrossRef]

Yoshino, Naoyuki, Farhad Taghizadeh-Hesary, and Miyu Otsuka. 2020. Determinants of international remittance inflow in Asia-Pacific middle-income countries. *Economic Analysis and Policy* 68: 29–43. [CrossRef]

Zhu, Yu, Zhongmin Wu, Liquan Peng, and Laiyun Sheng. 2014. Where did all the remittances go? Understanding the impact of remittances on consumption patterns in rural China. *Applied Economics* 46: 1312–22. [CrossRef]

economies

Article

Which Demographic Quintile Benefits from Public Health Expenditure in Nigeria: A Marginal Benefit Analysis

Kanayo K. Ogujiuba

School of Development Studies, University of Mpumalanga, Mbombela 1200, South Africa; kanayo.ogujiuba@ump.ac.za

Abstract: Policymakers concur that social investments are crucial, and that inequality must be decreased to accomplish long-term poverty reduction. Nigeria, one of the 20 poorest countries in the world, has a severely unequal society at the moment, with over 80% of the people living in deep, severe, and pervasive poverty, with an estimated 5% of the population possessing 85% of the country's resources. This article's focus is on how benefits are dispersed among various demographic groups. Previous data collection does not reflect the present realities of this topic. For this analysis, in southeast Nigeria, data sets from government agencies and for-profit service providers were utilized. The benefits of distinct quintiles were estimated using a marginal benefit incidence analysis. The results show that governmental spending in Nigeria is not pro-poor and that the country's southeast governments supported spending for the wealthy rather than the poor. The results show, among other things, that investment in health is not well directed; benefits from primary education and primary healthcare appear to be disproportionately dispersed to the upper class in the states studied, as they are throughout Nigeria. The paper serves as an example of the value of benefit incidence analysis (BIA). This article recommends effective targeted discretionary spending to lower systemic poverty and inequality. If education and health spending were more pro-poor, better education and health outcomes, strong governance, high per capita income, and wider access to information would all be more likely.

Keywords: health; expenditure; poverty

JEL Classification: A11; A13; G31; G32; G38

Citation: Ogujiuba, Kanayo K. 2022. Which Demographic Quintile Benefits from Public Health Expenditure in Nigeria: A Marginal Benefit Analysis. *Economies* 10: 253. https://doi.org/10.3390/economies10100253

Academic Editors: Ralf Fendel, Sajid Anwar and Robert Czudaj

Received: 29 June 2022
Accepted: 19 September 2022
Published: 12 October 2022

1. Introduction

Lack of resources can hurt relationships, health, life expectancy, education, and other factors (Victoria 2018). However, global sustainable development goals are supporting initiatives to fight poverty and enhance living conditions. Despite being Africa's top oil producer, Nigeria still has trouble turning its resources into better living circumstances. According to recent estimations, Nigeria has emerged as the archetype for high-income poverty in Africa. This is especially important given the 2014 GDP rebasing, which elevated it to Africa's largest economy, and the mismanagement of oil revenues since the 1970s. When multiple perspectives on the nature of poverty in Nigeria are applied, there is evidence that socially excluded people are impacted (Victoria 2018). However, compared to other African and Asian nations with comparable histories, Nigeria's development paradox stands out. Despite receiving over $300 million in total oil revenue from 1975 to 2015, Nigeria's current per capita income is still meager in real terms and lower than the prediction. Adejuwon and Adekunle estimate that approximately 70% of Nigerians live in poverty (2012). In Nigeria, poverty is widespread, severe, and enduring.

Nigeria is currently one of the world's 20 poorest countries, with roughly half of the population controlling 5% of the nation's resources. This raises significant questions regarding what transpired with the social investments, and, on the other hand, who benefited from them. Furthermore, Dauda (2017) observes that Nigeria's poverty pattern differs from that of many other countries in that, despite appearing to be making economic progress, poverty is still increasing, with the North-West and North-East zones leading the poverty index. On the other hand, prosperity has reduced poverty in developing nations in Europe, North America, and Asia. This confirms the widely accepted view that there is no direct correlation between poverty, development, and advancement.

Agu and Caliari (2014) assert that the government, including Nigeria, is responsible for a sizeable percentage of the economic activity in most rising nations. They contend that determining who gets what, when, and how is an important responsibility for the government at all levels. However, experts have long disagreed on the connection between government spending and economic growth. However, governments serve two main purposes. These functions encompass social, public, and security benefits. Despite the lack of an established framework on the mechanisms, the relationship between poverty reduction and public spending has been a topic of discussion over the past 20 years (Aigbokhan 2008; Al-Yousif 2000). The relationship between demographic dividends, inequality, poverty, and public health spending is part of society's larger picture. Injustice and a lack of resources generate poor health outcomes since the impoverished receive few significant benefits. Data show that residents of developing and transitional nations have less access to healthcare services than citizens of affluent nations. Additionally, access to healthcare is limited for the impoverished in their home countries. Furthermore, it is said that in developing nations, access to healthcare is frequently hampered by a lack of information and financial means.

Research indicates that there is a wide range of access to and utilization of government services by Nigerian families (Eboh and Diejomaoh 2010). Typically, higher-income groups benefit more from social investment by the government. Ben-Shahar asserts that the poor are frequently left unprotected from the detrimental effects of budget cuts; these cuts are frequently brought on by a framework's inability to predict revenue, such as a decline in the price of oil globally, which could result in a reduction in distributions to socioeconomic areas (Brian et al. 2001). The Nigerian health system, which includes tertiary facilities, is in disarray, according to data from numerous studies, which has fueled health tourism to other nations with better health facilities. The wealthy classes are still supported of despite this. In addition, Nigeria's productive sectors of the economy have declined significantly over the last two decades, and there is currently concern that the poor will pass on their poverty to their children, as poverty has become hereditary in the country.

Throughout most of Sub-Saharan Africa, the poor are especially vulnerable. Poorer households, for example, face crippling healthcare costs, thus allocating more healthcare discretionary non-food spending has become imperative. Long-term, business analysts and demographers have concluded that social interests in health and education should be prioritized to achieve equitable gains across demographic groups and long-term poverty reduction (Ogujiuba and Mngometulu 2022). Furthermore, contrary to earlier advancement assumptions that financial disparity promotes poverty reduction and growth, Cornia et al. (2004) assert that while inequality persists, no discernible progress toward sustainable development can be made. This contradiction draws attention to the well-being framework's apparent value in promoting economic development and eliminating poverty. However, rather than focusing on who receives what, when, and how, a large portion of the current debate among experts on the relationship between public spending and demographic dynamics has focused on determining the ideal population size and its implications for expectations for everyday comforts.

In 2015, 368 million of the world's 736 million extremely poor lived in just five countries, accounting for half of the total. India, Nigeria, the Democratic Republic of the Congo, Ethiopia, and Bangladesh have the highest number of extremely poor people (in descending order). They are also the most populous countries in South Asia and Sub-Saharan Africa, which account for 85 percent (629 million) of the world's poor. As a result, large reductions in poverty in these five countries will be critical if the global target of reducing extreme poverty (those living on less than $1.90 per day) to less than 3% by 2030 is to be met (Roy and Divyanshi 2019). Nonetheless, it is undeniable that most efforts made by densely populated nations to offer a respectable standard of living in terms of social services are continually stymied by rapid population growth without commensurate economic growth. The benefit incidence among the various quintiles hence becomes crucial in general.

Howbeit, interests in health, education, and social issues were highlighted by the Poverty Reduction Strategies Programs of the World Bank during the 1990s. They contend that boosting the creation of human capital lowers poverty. According to the World Bank (2021), human capital improves people's human capacities and productivity, allowing them to alleviate poverty and raise income through better chances. This simply suggests that poverty is reduced positively when expenditure is placed on social issues, and adversely when spending is concentrated on loss financing, economic, and community services. Policymakers must comprehend the distributional effects of such expenditure since the Nigerian government has resorted to using discretionary funds to mitigate the effects of poverty over time due to economic hardship. There are various perspectives and studies on the relationship between public sector spending and social development, even though most scholars believe that there are situations where less government spending is beneficial to social development and others where the opposite is true. As a result, there is a two-way relationship between social progress and government investment. This is because higher growth leads to better results, and higher sectoral outcomes, on the other hand, complement expenditures on social infrastructure. As a result, understanding the links between poverty reduction and redistribution impacts has become not just necessary but also imperative.

Nigeria, a third-world African country, is known as the world's poverty capital. It has surpassed India in terms of the number of people living in extreme poverty. Approximately 86.9 million people, or roughly 50% of the total population, live in extreme poverty. Despite its smaller size, both geographically and in terms of population, the country is failing to reduce poverty rates. This is due in part to the mismanagement of the oil industry and the presence of corruption. In addition, the country is experiencing a "population boom", which will make managing poverty rates more difficult. However, the Nigerian government has started several programs to aid those who are poor, but it is obvious that these efforts have not been effective enough, largely to poor targeting of resources. The country's programs are not successfully reducing the rates of poverty because of the high levels of corruption, unemployment, and inequality. The fact that Nigeria is the world's poorest nation influences the entire world in addition to Nigeria. Thus, Nigeria falls short of the United Nations ambition to end world poverty by 2050. A bigger plan to adopt a sustainable development framework, reducing the income gap between the wealthy and the poor, and efficient resource allocation, is essential. Ogundipe and Nurudeen (2013) assert that such a strategy would significantly lower poverty if social investments, notably in health and education, were made with care. Due to inadequate healthcare, the poor are vulnerable. Quality healthcare services affect attitudes toward work and society and beliefs and value systems. They also have an impact on precise knowledge and the development of a broad cognitive skill set. Lack of access to healthcare makes the poverty of the poor even worse.

With a growing emphasis on the importance of pro-poor health financing, it is becoming increasingly important to be able to monitor the impact of policy and strategy on poor healthcare consumers. While National Health Accounts are an important tool for establishing the level, sources, and allocation of financial resources within the health system, they provide little information about who benefits from the expenditure. Benefit incidence studies look at how well governments use their limited resources to meet the needs of the poor. They provide a revealing analysis of how, for example, groups divided by income or gender use primary and hospital services differently in rural and urban settings. The objective of this article is to demonstrate the effects of spending on a variety of demographic groupings. Because of the close links between health sector policy goals relating to equity of access, determining who benefits from public subsidies is critical for policymakers. The distribution of the benefits of social programs has a significant impact on progress toward the SDGs' general poverty reduction targets. Thus, periodic benefit incidence studies could be a useful performance indicator for the health sector (for possible incorporation into Sector Wide Approaches and Poverty Reduction Strategy Papers). Governments can be held accountable for their success in allocating public resources, but they cannot be held accountable for health-related improvements. This article provides context and assists in the correction of existing flaws, allowing for the more effective implementation of future resource distribution methods for reducing poverty and economic imbalance in Nigeria.

2. Literature Review

There is an alternative path that the causal relationship between healthcare benefits and absenteeism takes. Leo (2014) claims that Nigeria's social investment was already among the lowest in the world before the 2010 GDP rebase, but it dropped even lower after that. For instance, public health spending represented 2% of GDP before rebasing but just 1% following it. In most impoverished countries, medical care is routinely postponed, which hurts people's well-being, income, and out-of-pocket spending over the long run. Income, poverty, and education have a negative correlation, according to a prior study (Dabalen et al. 2013). Before the Boko Haram insurgency in northern Nigeria, 10 million Nigerian children (42 percent), according to Zachary (2014) and Bourne (2015), were not enrolled in school (2015). In emerging nations, fiscal policy formulation and implementation are challenging. Contrary to established economies, developing nations lack a de facto progressive tax system and efficient tax administration to change the distribution of income after taxes (Alesina 1999; Zee 1999; Atkinson 1999; Chu et al. 2000; Tanzi et al. 1999). In a similar vein, many nations lack the administrative capacity and the resources necessary to carry out cash transfer programs that can affect post-transfer income, consumption, or other welfare metrics (Tanzi 1998; Chu et al. 2000; Bourguignon et al. 2008). Since in-kind transfers often consist of social services such as education, healthcare, and social safety net programs, governments of developing countries prefer to distribute resources in this way. While many sorts of government spending are important for individual well-being, social services are frequently viewed as the most important for enhancing the population's long-term earning capacity, especially for the poor.

The utilization of government inpatient and delivery services in India is pro-poor, according to the findings of Bowser et al. (2019). When gross and net benefits are considered, however, services become more equal and less pro-poor. Gross benefits are nearly equal for all services when measured using state-level unit costs. Although there are some pro-poor gross benefit trends for national outpatient services, the findings also show that national gross benefit equality conceals a significant disparity across Indian states. While several Indian states have pro-poor outpatient gross benefits, few have pro-poor inpatient and delivery services. Net benefits, which consider both unit costs for each service and out-of-pocket (OOP) expenses, follow a similar pattern. In addition, those who use public facilities spend considerable OOP to supplement government services. On the other hand, according to Peter et al. (2017), based on available international costing norms, India does not finance primary healthcare services at a level sufficient to provide a comprehensive package of services to its citizens. While states bear most of the financial burden for healthcare, the federal government provides targeted grants for primary care to supplement state spending, particularly in poorer states. Furthermore, the weak and disadvantaged members of society may experience uncomfortable conditions and higher degrees of poverty as a result of the lack of social benefits (transfers).

Castro-Leal et al. (1999) looked at health and education spending using comparative benefit incidence analysis. Government health expenditures for the wealthiest 20% of the population were roughly 2.5 times greater than for the poorest 20%, according to their data. The wealthiest 20% of the population in five of the seven countries had more financial improvements than the bottom 20%. Overall, the wealthiest 20% received around 1.5 times the amount spent on primary care as the lowest 20%. The analysis indicated that public spending is reversing in all the countries studied. Moreover, the bounds of traditional benefit-incidence were extended in research in Nepal, Sri Lanka, and Bangladesh studied by Rannan-Eliya et al. (2001). To examine the total fairness of healthcare finance rather than just the government's share, this included both private and public healthcare spending. While Sri Lanka's health and financial situation has improved, Bangladesh's has worsened. The distributional effects of health expenditures in India and its major states were also evaluated by Ajay et al. (2000) using BIA, and they found that the wealthy benefit more from health spending than the poor. Compared to hospital treatment, the financial benefits of primary and outpatient care were divided more fairly. Pro-rich favoritism was more prevalent in rural than urban areas and poorer than affluent states. However, the findings of Gomanee et al. (2005) demonstrate the need for new techniques in battling poverty rates, as social service spending is not as successful as it should be in reducing poverty due to inadequate targeting strategies. This has sparked a lot of debate over how effective the Nigerian government's targeting strategy is.

MBIA (marginal benefit incidence analysis) is a well-known approach for assessing healthcare and education spending distributions in connection to socioeconomic welfare distributions (Bowser et al. 2019). Instead of comparing descriptive statistics by stratified variables, it integrates the distribution of benefits throughout the population into a single number that is nearly comparable to the Gini coefficient and may be used to compare results over time and space. This characteristic has led to the method's usage in a variety of non-industrialized nations, such as Vietnam, Pakistan, Jordan, and Nigeria, to give quantitative proof of how medical care administration and expenditures are dispersed to various population segments based on their socioeconomic position (Bowser et al. 2019). Despite inherent subjective limitations and challenges to quantifying concepts, marginal benefit incidence analysis has assumed a central role in most policy evaluations. This method (MBA) is frequently employed to discover which demographic groups benefit from government spending and to investigate the distributional impact of government spending (cash or in-kind). The main premise is that government spending benefits should be dispersed evenly, with those in the lowest quintile benefiting more than those in the richest quintile. The method can therefore be used to assess how pro-poor the government's fiscal policies are. In other words, the Keynesian model proposes that government investment be utilized to close skill gaps in the market, raise proficiency, and guarantee a fair distribution of economic advantages (Van-de-Walle 1995).

The World Bank, Demery et al. (1995), Castro-Leal (1996), Sahn and Younger (1998), Van-de-Walle (1995), and others have all utilized this method to examine low- and middle-income countries. Furthermore, according to Reinikka (2002), the use of BIA is most beneficial because there is minimal evidence of the effect of public sector investment on development indicators, which is the case in Nigeria. However, classifying the receivers of government healthcare spending is the same as evaluating how well healthcare initiatives are working to combat poverty and inequality. Nonetheless, most subsequent benefit incidence investigations were started by (Selowsky 1979). An estimate of the distributional advantages of government spending is produced by their research.

3. Data and Methods

The numbers for this study were compiled using data from the Nigerian Living Standard Survey 2018 report as compiled by the World Bank (2021) microdata. The survey targeted both city and rural families at the same time, and it included roughly 19,000 families. The survey included a wide range of topics (including social and financial advice), and the data included information on a family's total spending. The 2018/19 NLSS questionnaire covers all the demographic indicators, including those related to education, health, labor, spending on food and non-food items, non-farm enterprises, household assets, and durable goods, access to safety nets, housing conditions, economic shocks, exposure to crime, and farm production indicators. The survey data includes information on resources and their accessibility at the third level of government and is broken down by state, geographical region, and orientation (male/female). The 2018 HNLSS was a development of the 2010 review and an expanded adaptation of it. The sampling frame for the 2018/19 NLSS was the national master sample created by the NBS, referred to as the NISH2 (Nigeria Integrated Survey of Households 2). The enumeration areas (EAs) for the 2006 Nigeria Census Housing and Population conducted by the National Population Commission were used to create this master sample (NPopC). The NISH2 was developed by the NBS as a framework for surveys covering issues at the state level. For surveys containing LGA-level domains, NBS created a separate master sample from which the NISH2 EAs were produced (the "LGA master sample").

This analysis is premised on the data on income, spending, health, and education in each of the five states in the southeast states of Nigeria. BIA identifies the recipients of public goods and services. Additionally, using the information on unit costs, shows how benefits affect various populations. Expenditure on health, for example, can be formally written as:

$$X_j \equiv \sum_{i=1}^{3} E_{ij}\frac{S_i}{E_i} \equiv \sum_{i=1}^{3} \frac{E_{ij}}{E_i} S_i \tag{1}$$

This is computed by multiplying the main health facility unit cost by the number of secondary accesses multiplied by the secondary health unit cost, plus the number of tertiary accesses multiplied by the tertiary unit cost. The number of secondary healthcare consumers is then multiplied by the secondary healthcare unit cost, which is then multiplied by the secondary healthcare unit cost. The number of tertiary healthcare users is multiplied by the unit cost of supplying tertiary healthcare to obtain the result. According to Amakom (2012), X_j is the total amount of social assistance (wellness) cash that benefits group j1 (j is the economic group, and for the sake of this article, all families were divided into five quintiles based on their economic status—from the lowest to the highest income group). The subscript i denotes the level of social assistance (in Nigeria, medical care is divided into primary, secondary, and tertiary, so i = 1 to 3), and the subscript S and E denote the government social sector (health) subsidy appropriation and the number of people expected to benefit from the wellbeing office for the (wellbeing area), respectively (wellbeing area). This is determined by increasing the unit cost of an essential wellbeing office by the quantity of primary access duplicated by the optional unit cost of auxiliary wellbeing, in addition to the quantity of tertiary access duplicated by the tertiary unit cost.

As per Amakom (2012), the advantageous occurrence of absolute well-being attributed to the group is not entirely settled by "the number of clients of essential medical services from the gathering" (Epj). The quantity of buyers of optional medical services is then duplicated by the unit cost of giving auxiliary medical care, which is then increased by the unit cost of giving auxiliary medical services. The outcome is determined by duplicating the amount of tertiary medical care clients by the unit cost of conveying tertiary medical services.

The percentage of total health spending attributable to group (X_j) is then calculated by:

$$X_j \equiv \sum_{i=1}^{3} \frac{E_{ij}}{E_i}\left\{\frac{S_i}{S}\right\} \equiv \sum_{i=1}^{3} e_{ij}s_i \tag{2}$$

Two important determinants are used in Equation (2):

1. The e_{ij}s are the group's percentages of overall service usage (number of people who visit a health facility in the health sector), indicating household behavior.
2. Government behavior is reflected in the s_i, which is the share of public spending allocated to different types of services.

This study employed an approach where the monetary worth of the benefits people receives from utilizing public services is not based on their behavior. Instead, the unit cost of the service, or its value, was distributed equally among all users of the services in the form of rewards. Instead, rather than determining the precise worth of services provided by the government, this research concentrates on the distribution of service recipients and the benefits of counterfactual reciprocity of expenditures (Heltberg et al. 2003). The possibility that a group will benefit from a government subsidy or investment is known as this. Ajwad and Wodon (2001) and Lanjouw and Ravallion (1999) calculated the distribution of additions to public service access rates at the margin using a single cross-sectional data set. The spread of new access in lagging regions is predicted to follow the pattern observed in areas with greater access rates; hence, both studies used this assumption to anticipate the development of access through time. However, the methods used by Ajwad and Wodon (2001) and Lanjouw and Ravallion (1999) to rate persons are dissimilar. In addition, while Ajwad and Wodon categorized individuals based on their position in the local income distribution, Lanjouw and Ravallion classed individuals as affluent or poor based on their place in the national income distribution. This assumes that the works of Lanjouw and Ravallion (1999) and Ajwad and Wodon (2001) disagree. The details are as follows:

1. How is endogeneity bias handled in the marginal benefit incidence analysis calculation? Both authors used the overall access rate means to derive the access rate in each quintile. Ajwad and Wodon use the leave-out mean as their right-hand side variable to eliminate endogeneity. Except for the quintile in which the relapse is complete, they regressed against the mean of the admission rates across all quintiles. Lanjouw and Ravallion, on the other hand, used an instrumental method to instrument the real mean, using the leave-out mean.
2. Ajwad and Wodon used marginal benefit incidence analysis to constrain their estimates. Although Lanjouw and Ravallion disagree, they believe that removing the restriction will slant the numbers downward.

Lanjouw and Ravallion also describe the following econometric procedure, which has been used in other studies (Ajwad and Wodon 2001; Kamgnia-Dia et al. 2008):

$$\rho_{i,j,q} = \alpha_q + \beta\rho_k + \mu_q \tag{3}$$

where i refers to a small geographic unit, k to a bigger one, and q to the welfare quantile. The program engagement rate for the partition and quantile is the left-hand variable. The regressor is the percentage of people who participate in programs in the division's region. The marginal effect of increasing people's program participation rates in a particular location and quantile is q. Lanjouw and Ravallion (1999), as the percentage of a specific quintile population who partakes in a program sponsored by the government, defined the average participation rate. Each quantile's regressor is run separately. Furthermore, because ijk is included in k, the estimation has an upward bias. Lanjouw and Ravallion, as previously stated, settled this issue by instrumenting k with the left-out mean. The thinking behind the estimate is that by recognizing contrasts in cooperation, it will be feasible to see how better coverage affects the participation of various demographic groupings. Assuming q is greater than one, an overall expansion in inclusion is related to a lopsidedly critical expansion in support of that division and quantile. Nonetheless, a key postulate in our model is that the political process is consistent across all places and drives the link between program preference and prevalence.

MBIA, on the other hand, is the consequence of government and household efforts achieving "equilibrium". It does not refer to a paradigm that guides government or household activity. On the other hand, studies of demand functions for public services (e.g., Younger 1999) address this gap, but they are few. It also contrasts the costs and benefits. The cost of providing public services is used by BIA as a measure of the value ascribed to such services, and it is assumed that the costs of provision are a decent approximation to the benefit that users attach to such services. MBIA does not, in most cases, cover the total cost of providing public services (e.g., tax administration), including both monetary and nonmonetary costs. At a given point in time, MBIA captures the best benefit incidence of government spending. MBIA is required to provide a dynamic picture of incidence over time for different years. Behavioral models, on the other hand, can better capture dynamic advantages from government spending than BIA. Benefit incidence estimates frequently represent the average incidence. This means that BIA does not often provide information on who gains from an increase or decrease in government spending, which are crucial problems for policymakers; for a study of marginal incidence, see Younger (2002). More MBIA should be carried out more frequently for emerging countries and transition economies, despite the several inherent limitations of MBIA that have been recognized, even when BIA is absent in many poor nations. When creating and implementing future policy interventions, it is essential to establish a benchmark benefit incidence pattern with the understanding that better methodologies for incidence evaluation should be employed when resources are available.

Analytical Technique (Marginal Benefit Analysis)

Summarily, the steps employed in this article to compute marginal benefit analysis for the southeast states of Nigeria are as follows:

1. The population was divided into equal-sized sections according to welfare standards. This made it possible to split the population into quintiles. More information was broken down by states, locales, and gender.
2. Identification of the households that received government assistance (education and health). The Nigerian Living Standard Survey from 2018 served as the basis for this (NLSS). Additional data was acquired from hospitals visited and schools.
3. The State Universal Basic Education Board (SUBEB) provided primary school information; the Post-Primary School Management Board (PPSMB) provided secondary school information; and the NCCoE, National Board for Technical Education (NBTE), and National Universities Commission provided tertiary school information (NUC).
4. Information on primary healthcare was obtained from the Primary Health Care Development Agency (PHDA) of the southeast states, information on secondary healthcare from the Hospital Management Boards (HMB) of states, and information on tertiary healthcare from the Ministry of Health within the states. See Figure 1 for a map of Nigeria showing the area for analysis
5. Data sets were matched while considering any potential biases in household data brought on by survey design, questionnaire format, and sample size. Using 2018 NLSS household data, we ranked persons in the southeast states based on household consumption per capita, and an exception was given to those who received benefits.
6. The estimated cost of providing a service was determined by dividing government expenditure on the service by the total number of users of the service, and then dividing the estimated cost of providing quality service by the estimated price of providing the service to arrive at the average benefit from government expenditure on the service, using the methodology of Amakom (2012).
7. Finally, a two-stage least square method was used to calculate the distributional spread of benefits across quintiles.

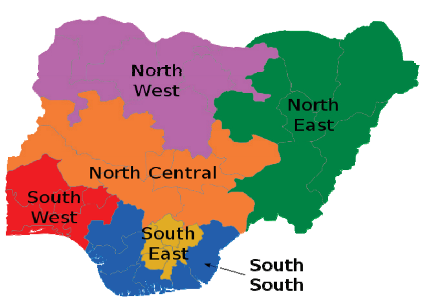

Figure 1. Map of Nigeria showing regions (spatial context for Southeast).

4. Results and Discussion

Tables 1–3 below show the projected benefits and marginal odds of using public primary, secondary, and tertiary healthcare services. Regressing each quintile's participation rate against the average participation rate yielded these results. The tables show the increase in subsidy incidence per capita for each quintile because of a one Naira increase in aggregate primary healthcare spending. According to Table 2, only three of the five SE states (Anambra, Ebonyi, and Enugu) have a pro-poor goal in terms of primary healthcare. Quintiles (1, 2, and 3) for the states demonstrated that the allocations helped the neediest people the most. They were given a government benefit that was worth more than the N1 they spent. SDG 3: Good Health and well-being correspond to this. This increase, however, will not result in a better average because the other states are still significantly below the planned aim.

Beneficiaries of Health Expenditure across Quintiles in Southeast Nigeria

Below Table 1 shows the spread across quintiles for primary health care amongst the SE States in Nigeria.

Table 1. Primary healthcare (benefit spread using 2018/19 HNLSS).

States	Abia	Anambra	Ebonyi	Enugu	Imo
Quintile 1	0.964	1.385	1.113	1.453	0.988
T—Stat	2.515	2.335	2.449	2.262	1.558
Quintile 2	0.997	1.520	1.063	1.480	1.088
T—Stat	1.644	2.521	2.277	1.734	1.785
Quintile 3	1.112	1.393	1.067	1.327	1.063
T—Stat	2.559	3.060	2.446	1.555	2.466
Quintile 4	1.014	0.514	0.957	0.449	1.039
T—Stat	1.875	1.768	1.788	2.150	1.908
Quintile 5	0.923	0.200	0.822	0.300	0.837
T—Stat	6.619	1.997	5.992	3.908	8.282
Total	5.010	5.012	5.021	5.008	5.015

Source: Author's Computations.

Table 2. Secondary healthcare (benefit spread using 2018/19 HNLSS).

States	Abia	Anambra	Ebonyi	Enugu	Imo
Quintile 1	1.015	1.068	0.964	1.074	0.995
T—Stat	3.747	3.156	2.246	3.622	3.889
Quintile 2	0.927	1.117	1.016	1.076	1.015
T—Stat	6.721	2.458	7.400	2.142	3.556
Quintile 3	1.115	0.979	1.015	1.065	1.023
T—Stat	4.832	2.152	2.720	2.663	5.956
Quintile 4	0.879	0.973	0.891	0.993	1.057
T—Stat	1.445	1.591	1.958	1.163	1.727
Quintile 5	1.066	0.863	1.117	0.793	0.913
T—Stat	3.887	3.519	2.847	2.649	3.483
Total	5.002	5.000	5.003	5.001	5.002

Source: Author's Computations.

N/b:

1. Based on the 2018 HNLSS, Tables 1–3 show the instrumental variables estimate of the quintile-specific service rate regression coefficient on the average rate for the southeast region.
2. The instrument for the real mean is the leave-out mean area service rate.
3. The t-ratios are the numbers in parentheses.
4. Quintile 1 (very poor); quintile 2 (poor); quintile 3 (moderate); quintile 4 (rich); and quintile 5 (extremely wealthy).

Table 3. Tertiary healthcare (benefit spread using 2018/19 HNLSS).

States	Abia	Anambra	Ebonyi	Enugu	Imo
Quintile 1	0.633	0.700	0.670	0.760	0.803
T—Stat	2.246	2.934	1.705	2.450	3.177
Quintile 2	0.814	0.850	0.788	0.882	0.699
T—Stat	3.208	1.776	2.012	1.499	2.803
Quintile 3	1.082	0.974	1.019	1.070	0.975
T—Stat	4.525	2.167	2.456	2.886	3.850
Quintile 4	1.207	1.132	1.113	1.067	1.186
T—Stat	4.648	2.360	2.829	1.996	4.485
Quintile 5	1.264	1.344	1.411	1.223	1.338
T—Stat	4.613	3.926	3.258	4.165	3.615
Total	5.001	5.001	5.002	5.001	5.001

Source: Author's Computations.

The affluent were substantially subsidized by the other three states, which increased poverty and weakened the nation's healthcare system. Despite some targeting displayed by Ebonyi state, an overwhelming amount of money was given to the wealthy. This cash might have gone toward aiding the underprivileged and impoverished people in our society. Quintile 5 received funding from the states of Anambra and Enugu of 3.9 percent and 5.9 percent, respectively, on a scale of 100 percent, while the states of Ebonyi, Imo, and Abia provided 26.4 percent, 16.7 percent, and 18.4 percent, respectively, in support of the same group. This scenario assumes that anyone without access to basic healthcare will at the very least experience poverty. The affluent were extensively supported by the other three states, which led to a rise in poverty and a deterioration of the nation's healthcare infrastructure. Even though Ebonyi state showed some targeting, a disproportionate amount of money was given to the wealthy. The less fortunate people in society may have benefited from the use of these funds. Quintile 5 was financed by the states of Anambra and Enugu, with 3.9 percent and 5.9 percent, respectively, on a scale of 100 percent, while the states of Ebonyi, Imo, and Abia supported the same group with 26.4 percent, 16.7 percent, and 18.4 percent, respectively. According to the hypothetical situation, those without access to primary healthcare will, at the very least, experience poverty.

The predicted gains and marginal probability from utilizing public secondary and tertiary healthcare are shown in Tables 2 and 3 and were determined by regressing the participation rates of each quintile against the overall participation rate. The estimated figures in the table show the increase in the distribution of subsidies per capita for each quintile because of an increase in total healthcare spending in Naira. For the demographic group in quintile 1, three states in the southeast had an incidence of more than one at the secondary healthcare level. However, the quintile 4 and quintile 5 demographic groupings benefited the most from state governments' discretionary spending. Table 3 demonstrates that quintiles 1, 2, and 3 were more effectively targeted than quintiles 4 and 5. More than 60% of the resources were obtained by them. However, quintile 5 only received about 20% of the funding, which is anti-poor and goes against SDG 1's No Poverty aim. An identical pattern was discovered at the tertiary level.

Table 4 below demonstrates that quintiles 4 and 5 had high significant values throughout the five states. This group received around half of the public resources. However, it is believed that this trend may be found across all of Nigeria's states. The conceptual framework demonstrates how the poor are shut out of the development process. Targeting was insufficient for quintiles 1 and 2. In this case, the SDG 10 goal of reducing gaps was not achieved because the disadvantaged groups (quintiles 1 and 2) did not benefit at this level. They had about 155 resources in total. This condition is an abnormality in development. Additionally, tertiary healthcare is designed to help the poor who cannot afford specialized care, but in this case, that is not the case. In Nigeria's five eastern states, the extremely wealthy and wealthy fared better than the less fortunate. This circumstance will simply serve to exacerbate the system's growing inequity. Foster et al. (2002) agree with our results.

Furthermore, previous research in Ghana, Malawi, Mozambique, Tanzania, and Uganda revealed a lopsided benefit structure that favored the wealthy over the poor.

Table 4. Access to the healthcare system (poor versus rich): In the SE states, a summary of health benefits is available (quintiles 2, 4, and 5).

Abia State			
Quintiles	*Primary System*	*Secondary System*	*Tertiary System*
Poor	19.9	18.5	16.3
Rich	38.6	38.9	49.4
Benefit/Loss	**−18.7**	**−20.4**	**−33.1**
Anambra State			
Poor	30.3	22.3	17
Rich	14.2	36.7	49.5
Benefit/Loss	**16.1**	**−14.4**	**−32.5**
Ebonyi State			
Poor	21.2	20.3	15.8
Rich	35.4	40.1	50.4
Benefit/Loss	**−14.2**	**−19.8**	**−34.6**
Enugu State			
Poor	29.6	21.5	17.6
Rich	14.8	35.7	45.8
Benefit/Loss	**14.8**	**−14.2**	**−28.2**
Imo State			
Poor	21.7	20.3	13.9
Rich	37.4	39.4	50.5
Benefit/Loss	**−15.7**	**−19.1**	**−36.6**

Source: Author compilation; Underlying data were derived from estimates sourced from the National Bureau of Statistics (NBS) and Harmonized Nigeria Living Standard Survey (HNLSS). Notes: Figures derived from Annexure A; Benefits refers to Coefficient/Total × 100; Rich—Quintile 4+5; Poor—Quintile 2.

Because they cannot afford to pay for private medical care, the poor and very poor make up a larger share of the population in developing countries and are in urgent need of medical care. The state's inability to lower child mortality, improve maternal health, combat malaria, and treat other diseases, as outlined in the Sustainable Development Goals, may be evidenced by lower life expectancies, an increase in the prevalence of chronic diseases, and poor health conditions for the poor and very poor (SDGs). As a result, poor people with poor health access have higher rates of poverty, lower incomes, and reduced productivity. The fact that impoverished people have low earnings, which makes it difficult for them to pay for treatment, remains the driving force behind initiatives to improve their access to quality healthcare.

Compared to the wealthy in Abia state, the poor and very poor have less access to healthcare services. However, the poor people of Anambra have little access to medical facilities. Study findings indicate that wealthy people profit from secondary and tertiary health programs available in the state, whereas the very poor receive fewer healthcare advantages than wealthy people. The very poor are given precedence when it comes to access to primary healthcare, but they still need more access than the wealthy, who can pay for their care. Due to the outcomes, the poor and very poor are at risk of going even further into poverty if a large portion of their population is unemployed or underproductive because of inadequate access to healthcare. This goes against the SDGs, which call for the eradication of extreme poverty and hunger while simultaneously protecting the environment. On the one hand, the wealthy may be vulnerable to the spread of disease or illness, requiring additional spending and possibly a sacrifice in the meager comforts enjoyed by the poor and severely poor to protect their area. The government will need to increase recurrent spending to safeguard the state due to the chronic under-provision of healthcare services for the poor. The gap between healthcare benefits for the very poor and poor in Ebonyi state has widened. Healthcare remains prohibitively expensive for low-income citizens of the state since the poor and the very poor have limited access to it.

In addition to the prior discussion of the affordability issue, the lack of the restricted benefit could hinder social development because many individuals are ill and cannot obtain the right diagnosis or treatment because they lack access to qualified medical personnel. For the state to encourage and maintain growth, it is pertinent to provide healthcare access for the poor and the very poor. However, several factors that contribute to the core poor's lack of access to healthcare, including a lack of skilled personnel and funds, insufficient logistics, a lack of a maintenance culture, and a high level of leadership turnover has been identified in extant literature and studies. However, the repercussions could be detrimental to the state's ability to achieve inclusive growth.

The expansion of healthcare facilities in Enugu state has not directly led to higher incomes for the underprivileged. The three healthcare systems' distribution of benefits to citizens, particularly the primary, secondary, and tertiary systems, continues to favor the wealthy. This implies that the state's most vulnerable residents cannot afford its healthcare system. The rural health system's provision of primary healthcare services aids the poor and very poor, yet it is still insufficient for those with low incomes and a high risk of requiring medical attention. Increased access is necessary to promote pro-poor growth because these people have limited awareness of affirmative action and healthcare. Increased access to healthcare services for these groups is justified economically in terms of revenue production and stable policies that provide a smooth and well-functioning environment that encourages entrepreneurship and creative thinking. Only productive citizens, that is, those who are physically and mentally well will provide more revenue to the government. If the poor and very poor have better access to healthcare, they will be more productive and contribute their fair share to economic advancement. Even though the findings showed increased access to primary healthcare services for all categories, which supports MDGs 4, 5, and 6, Enugu state could perhaps increase the benefit of secondary and tertiary healthcare services, which specifically benefit the core poor, to achieve these goals.

Additionally, findings indicate that Imo state's extremely poor residents are benefiting less from healthcare services. As a result, those who fall into the extremely poor and poor groups are more likely to experience early death and low quality of life. Imo state is still far from achieving the Sustainable Development Goals due to the unequal distribution of healthcare services. that nonetheless, in order to provide the core poor with sustainable healthcare services, the state must manage the resources allotted, involve the population of the poor, encourage education, and employ skilled health staff. Figures 2–6 shown below typify healthcare services in Abia, Anambra, Ebonyi, Enugu and Imo States in Nigeria.

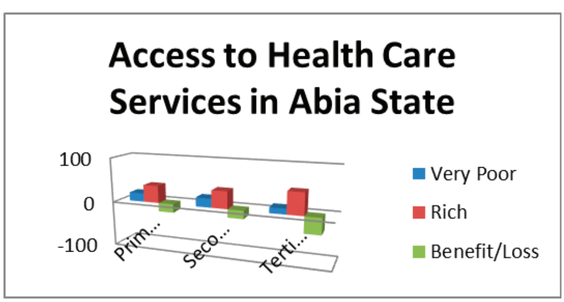

Figure 2. Access to healthcare services (Abia state). Source: Graphed by the author; Underlying data were sourced from the Central Bank of Nigeria (CBN) and the National Bureau of Statistics (NBS).

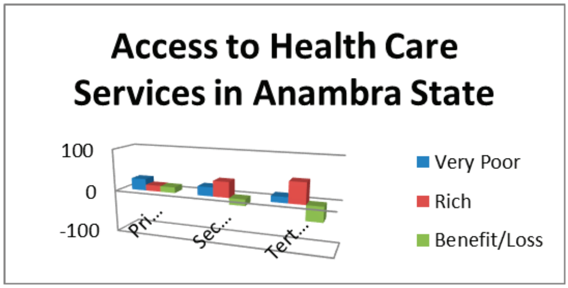

Figure 3. Access to healthcare services (Anambra state). Source: Graphed by the author; Underlying data were sourced from the Central Bank of Nigeria (CBN) and the National Bureau of Statistics (NBS).

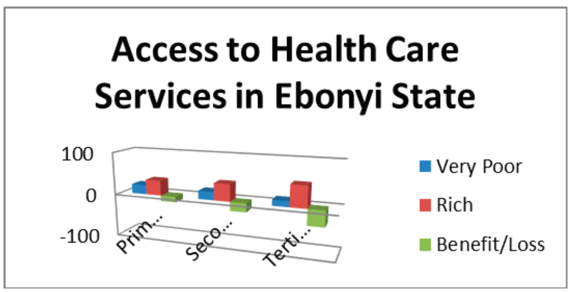

Figure 4. Access to healthcare services (Ebonyi state). Source: Graphed by the author; Underlying data were sourced from the Central Bank of Nigeria (CBN) and the National Bureau of Statistics (NBS).

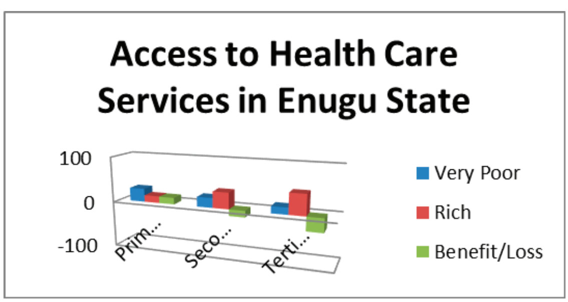

Figure 5. Access to healthcare services (Enugu state). Source: Graphed by the author; Underlying data were sourced from the Central Bank of Nigeria (CBN) and the National Bureau of Statistics (NBS).

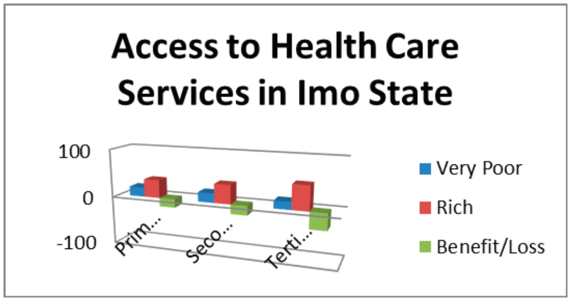

Figure 6. Access to healthcare services (Imo state). Source: Graphed by the author; Underlying data were sourced from the Central Bank of Nigeria (CBN) and the National Bureau of Statistics (NBS).

Amakom and Uju (2012) investigated the relationship between a family's health out-of-pocket spending and its level of poverty. Their findings imply that substantial out-of-pocket healthcare expenses have helped numerous families escape poverty. The foregoing conclusions seem to be supported by WDI statistical data. The Nigerian Medical Association (NMA) estimated in 2012 that roughly 5000 Nigerians travel to India and other countries monthly for clinical therapy, indicating that the nation spends between $1 billion and $2 billion on medical tourism annually. This means that out-of-pocket health spending made up 95.34 percent of all private health spending in 2018. In the southeast region of Nigeria, healthcare spending is wholly out of control. Families in need do not receive aid, but wealthier people receive substantial subsidies. The Nigerian Living Standards Survey from 2004 formed the basis for Amakom's (2011) study of Nigeria, which is supported by our findings. Using the benefit incidence technique, he examined public spending initiatives aimed at alleviating poverty and inequality at all levels of the health system (BIA). Primary healthcare did not appear to be specifically pro-poor, despite secondary healthcare's inconsistent outcomes.

Furthermore, results from other studies (Olamide et al. 2022) suggest the use of bilateral links among countries in the reduction of poverty within a zone. This could be achieved by leveraging the benefits of ICT's poverty-reducing impacts, economic growth, financial development, and trade openness. As applicable in other advanced and emerging economies, the digital competence of Nigeria needs to be synchronized for effective service delivery to the most vulnerable groups.

5. Conclusions

Improvement in the health of a country's population, providing financial risk protection, and citizen satisfaction are the three goals of health systems. The availability of good-quality and relevant data and evidence on the resources allocated and used to finance and deliver health services is important for developing and implementing strategies to meet those goals. Systematic health resource tracking can contribute to this effort. The methodological framework of this article is more concerned with a broad philosophy than with a policy framework for extending social services. The effect must be within the political restrictions set by each group's cost, benefits, and political clout, regardless of whether procedures are approved, according to the research. The demographic quintiles gained from discretionary government spending were analyzed in this article. This article employed a modified method in which behavior data is not used to regulate the value of money regarding the benefits a person receives from using government amenities. Fairly, all persons who used the services received the same monetary worth of benefits, which is the worth of the unit cost of delivering the service. This study focused on the distribution of service beneficiaries rather than determining the exact value of government-sponsored services to recipients. The binary strategy given by Sahn and Younger (1998) was introduced because of the government's deficient data structure and plan estimates. As a result, technicalities, such as the necessity to estimate unit subsidies, which are not included in the first and second formulae, were avoided. The focus was solely on whether a service is used, with consumers of communal services being counted and receiving one benefit while non-consumers receive zero. Furthermore, how the health budget is split among the population for health facilities is determined by the utilization of government resources. This is known as current accounting, although the results may not accurately reflect the distribution of changes across all groups or quintiles.

Empowerment in Nigeria could be enhanced by improving social investment facilities. An increase in assets and capabilities is defined as empowerment. Quintiles 1 and 2 will be able to fully participate, bargain, and hold accountable authorities that have an impact on their lives because of this. For Nigeria to meet the task of invigorating sweeping economic growth and social change, spending options should be created considering proof-based examination and a survey of who obtains what, when, and how. To sum up, measures for lessening poverty and inequality in Nigeria should consolidate distributive components in addition to a high level of inclusion. Several indicators are indicating that those living in poverty in the southeast region of Nigeria do not receive the same level of attention and treatment as others who have a higher income. Some people might not be able to receive any medical care at all because of inefficient discretionary targeting of healthcare resources. This implies that the poor are less likely to benefit from the health protection provided to others and that most children living in impoverished households do not receive the necessary vaccinations against avoidable diseases. According to the aforementioned, the most efficient strategy for all Nigerian states to tackle poverty in the nation would be to focus all social sector resources on households in quintiles 1 (very poor) and 2 (poor) throughout the entire nation. This would help to alleviate poverty by increasing the productivity of able-bodied men and women.

The foregoing findings present a dilemma to policymakers: if existing users of public education and health services are not the policymakers' intended beneficiaries, a pertinent question arises as to what policies should be implemented to modify the observed benefit incidence and, as a result, enhance social spending targeting. To make public services, at least, progressive, policymakers in nations with high poverty levels must make every effort to skew the incidence of public social spending in favor of the poor and boost the use of public services by the poor.

Recommendations

To make progress in this area, it is necessary to address several issues, including governance, gender bias, the location of public healthcare and education services, pro-urban bias, and the search for alternative in-kind transfer modalities in addition to joint public financing and public provision of healthcare and education services. This article recommends that policymakers in Nigeria should, amongst others, educate households and communities and make them more aware of how to effectively receive healthcare in their numerous jurisdictions to have the greatest positive influence on their socioeconomic situation. Furthermore, a proper targeting mechanism for safety net measures must be developed and is critical, as it would mitigate the current detrimental effects of health and education spending, especially on the poor. One of these possibilities is better targeting of healthcare and education spending among the poor. In addition, income redistribution can still be accomplished through subsidies rather than direct consumer transfers, if persons with exceptional needs are adequately targeted. The Nigerian government should invest more in social services and enhance access to basic healthcare to reach SGD targets, concentrating on outcomes rather than outputs.

To ensure an equitable distribution of resources, the income-sharing formula between Nigeria's three levels of government needs to be revised. Nonetheless, future studies could focus on calculating the precise value of government-sponsored services to its users. An important methodological lesson from this work is that future MBIA studies should pay more attention to recording incidence statistics and other data breakdowns (such as by location, gender, and ethnicity), and the required auxiliary IDs. To make it simple to assess the degree of progressivity of social spending, they should also provide information on income or consumption distribution.

Funding: This research received no external funding.

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

References

Agu, Chukwuma, and Aldo Caliari. 2014. Economic diversification and macroeconomic policies: Is Africa's growth enough? Paper presented at the CSAE Conference 2014, Oxford, UK, March 23–25.

Aigbokhan, Ben. 2008. Growth, Inequality, and Poverty in Nigeria, United Nations Economic Commission for Africa—UNECA Economic Commission for Africa, ACGS/MPAMS Discussion Paper No.3. Available online: www.uneca.org/sites/default/files/.../growthinequalitypoverty.pdf (accessed on 12 January 2013).

Ajay, Mahal, Jarunejaya Singh, Farzana Afridi, Lamba Vikram, Anil Gumber, and V. Selvaraju. 2000. *Who Benefits from Public Health Spending in India? Results of a Benefit-Incidence Analysis for India*. Washington, DC: World Bank.

Ajwad, Mohamed, and Quentin Wodon. 2001. *Do Governments Maximize Access Rates to Public Services Across Areas?* Washington, DC: World Bank, Latin America Poverty Group.

Alesina, Alberto. 1999. Too Large and too Small Governments. In *Economic Policy and Equity*. Washington, DC: IMF.

Al-Yousif, Y. 2000. Does Government Expenditure Inhibit or Promote Economic Growth: Some Empirical Evidence from Saudi Arabia. *Indian Economic Journal* 48: 92–96.

Amakom, Uzochukwu. 2011. *Public Spending and Poverty Reduction in Nigeria: A Benefit*. Incidence Analysis in Education and Health, Final Report for Publication Submitted to the African Economic Research Consortium (AERC). Nairobi: African Economic Research Consortium (AERC).

Amakom, Uzochukwu. 2012. Public Spending on Education and Healthcare in Nigeria: Who Benefits and Why. *International Journal of Business and Management* 7: 48–59. [CrossRef]

Amakom, Uzochukwu, and Ezenekwe Uju. 2012. Implications of households catastrophic pocket out-of-pocket (OOP) healthcare spending in Nigeria. *Journal of Research in Economics and International Finance* 1: 136–40.

Atkinson, Anthony. 1999. Equity Issues in a Globalizing World: The Experience of OECD Countries. In *Economic Policy and Equity*. Washington, DC: IMF.

Bourguignon, François, Maurizio Bussolo, and Luiz A. Pereira da Silva. 2008. *The Impact of Macroeconomic Policies on Poverty and Income Distribution: Macro-Micro Evaluation Techniques and Tools*. Washington, DC: World Bank and Palgrave Macmillan.

Bourne, Richard. 2015. *Nigeria: A New History of a Turbulent Century*. London: Zed Books.

Bowser, Diana, B. Patenaude, M. Bhawalkar, Denizhan Duran, Peter Berman, Bowser Diana, B. Patenaude, M. Bhawalkar, Denizhan Duran, and Peter Berman. 2019. Benefit incidence analysis in public health facilities in India: Utilization and benefits at the national and state levels. *International Journal for Equity in Health* 18: 13. [CrossRef] [PubMed]

Brian, Ames, Brown Ward, Devarajan Shanta, and Izquierdo Alejandro. 2001. Macroeconomic Policy and Poverty Reduction. The IMF's Poverty Reduction and Growth Facility: A Factsheet. Available online: http://www.imf.org/external/pubs/ft/exrp/macropol/eng/ (accessed on 15 February 2022).

Castro-Leal, Florencia. 1996. *The Impact of Public Health Spending on Poverty and Inequality in South Africa*. PSP Discussion Paper Series 102; Washington, DC: World Bank, Poverty, and Social Policy Department.

Castro-Leal, Florencia, Julia Dayton, Lionel Demery, and Kalpana Mehra. 1999. Public Social Spending in Africa: Do the Poor Benefit? *World Bank Research Observer* 14: 49–72. [CrossRef]

Chu, Ke-young, Hamid Davoodi, and Sanjeev Gupta. 2000. *Income Distribution and Tax, and Government Social Spending Policies in Developing Countries*. Working Paper 214. Helsinki: UNU World Institute for Development Economic Research.

Cornia, Giovanni Andrea, Tony Addison, and Sampsa Kiiski. 2004. Income Distribution Changes and Their Impact in the Post-Second World War Period. In *Inequality, Growth, and Poverty in an Era of Liberalisation and Globalization*. Oxford: Oxford University Press.

Dabalen, Andrew, Vasco Molini, and Rose Mungai. 2013. *Nigeria: Where Has All the Growth? A Policy Note*. Washington, DC: The World Bank. Available online: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2014/07/21/000470435_20140721115331/Rendered/PDF/789080REVISED00Box0377361B00PUBLIC0.pdf (accessed on 20 June 2022).

Dauda, Rasaki Stephen. 2017. Poverty and Economic Growth in Nigeria: Issues and Policies. *Journal of Poverty* 21: 61–79. [CrossRef]

Demery, Lionel, Shiyan Chao, Rene Bernier, and Kalpana Mehra. 1995. *The Incidence of Social Spending in Ghana*. PSP Discussion Paper Series 82; Washington, DC: World Bank, Poverty, and Social Policy Department.

Eboh, Eric, and Ito Diejomaoh. 2010. Local Governments in Nigeria: Relevance and Effectiveness in Poverty Reduction and Economic Development. *Journal of Economic and Sustainable Development* 1: 12–28.

Foster, Mick, Fozzard Adrian, Felix Naschold, and Tim Conway. 2002. *How, When and Why does Poverty Get Budget Priority: Poverty Reduction Strategy and Public Expenditure in Five African Countries*. Working Paper 168. London: Overseas Development Institute.

Gomanee, Karuna, Sourafel Girma, and Oliver Morrissey. 2005. Aid, Public Spending and Human Welfare: Evidence from Quantile Regressions. *Journal of International Development* 17: 299–309. [CrossRef]

Heltberg, Rasmus, Kenneth Simler, and Finn Tarp. 2003. Public Spending and Poverty in Mozambique. Available online: https://econpapers.repec.org/paper/fprfcndbr/167.htm (accessed on 5 December 2021).

Kamgnia-Dia, Bernadette, Simon LeunkeuWangun, Christophe Tatsinkou, and Josephine Afor. 2008. *Bénéfices acquis et ciblage des pauvres dans les dépenses publiques de santé et d'éducation au Cameroun PEP-PMMA Working paper May*. Quebec: Universite Laval.

Lanjouw, Peter, and Martin Ravallion. 1999. Benefit Incidence, Public Spending Reforms, and the Timing of Program Capture. *The World Bank Economic Review* 13: 257–73. [CrossRef]

Leo, Ben. 2014. Beyond the Headlines: 3 under-Reported Stories about Nigeria's 90% GDP Boost. Center for Global Development Blog, 8. Available online: http://www.cgdev.org/blog/beyond-headlines-3-under-reported-stories-about-nigerias-90-gdp-boost (accessed on 5 December 2021).

Ogujiuba, Kanayo, and Ntombifuthi Mngometulu. 2022. Does Social Investment Influence Poverty and Economic Growth in South Africa: A Cointegration Analysis? *Economies* 10: 226. [CrossRef]

Ogundipe, Mushay, and Lawal Nurudeen. 2013. Health, Poverty Reduction and Human Capital Development in Nigeria (1980–2011). *International Journal of Management Sciences and Business Research*. 2. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2715899 (accessed on 15 February 2022).

Olamide, Ebenezer, Kanayo K. Ogujiuba, Andrew Maredza, and Phetole Semosa. 2022. Poverty, ICT and Economic Growth in SADC Region: A Panel Cointegration Evaluation. *Sustainability* 14: 9091. [CrossRef]

Peter, Berman, Bhawalkar Manjiri, and Jha Rajesh. 2017. *A Report of the Resource Tracking and Management Project*. Boston: Harvard T.H. Chan School of Public Health Boston.

Rannan-Eliya, Ravi, Pande Badri, James Killingworth, and Aparna Somanathan. 2001. *Equity in Financing and Delivery of Health Services in Bangladesh*. Nepal, Sri Lanka and Washington, DC: World Bank.

Reinikka, R. 2002. *Public Expenditure Tracking and Service Delivery Surveys: A Review of Design and Implementation Issues*. PEAM Course, May 23, DECRG. Washington, DC: The World Bank.

Roy, Katayama, and Wadhwa Divyanshi. 2019. Half of the World's Poor Live in just 5 Countries. World Bank Blogs. Available online: https://blogs.worldbank.org/opendata/half-world-s-Poor-live-just-5-countries (accessed on 5 December 2021).

Sahn, David, and Stephen Younger. 1998. *Fiscal Incidence in Africa: Microeconomic Evidence*. AERC Working Papers CR-2-5. Nairobi: AERC.

Selowsky, Marcelo. 1979. *Who Benefits from Government Expenditure? A Case Study of Colombia*. New York: Oxford University Press.

Tanzi, Vito. 1998. *Fundamental Determinants of Inequality and the Role of Government*. Washington, DC: IMF.

Tanzi, Vito, Ke-young Chu, and Sanjeev Gupta, eds. 1999. *Economic Policy and Equity*. Washington, DC: IMF.

Van-de-Walle, Dominique. 1995. *Public Spending and the Poor: What We Know, What We Need to Know*. Research Working Paper No. 1476. Washington, DC: World Bank.

Victoria, Danaan. 2018. Analyzing poverty in Nigeria through theoretical lenses. *Journal of Sustainable Development* 11: 20–31.

World Bank. 2021. Available online: https://microdata.worldbank.org/index.php/catalog/3827 (accessed on 5 December 2021).

Younger, Stephen. 1999. The Relative Progressivity of Social Services in Ecuador. *Public Finance Review* 27: 310–52. [CrossRef]

Younger, Stephen. 2002. *Benefits on the Margin: Observations on Marginal Benefit Incidence*. USAID Cooperative Agreement AOT-0546-A-00-3178-00. Washington, DC: USAID.

Zachary, Wright. 2014. Poverty in Nigeria Increasing Despite GDP Growth. The Borgen Project. Available online: http://borgenproject.org/poverty-in-nigeria/ (accessed on 5 December 2021).

Zee, Howel. 1999. Inequality and Optimal Redistributive Tax and Transfer Policies. *Public Finance Review* 32: 359–81. [CrossRef]

economies

Article

Prices and Taxes in a Ramsey Climate Policy Model under Heterogeneous Beliefs and Ambiguity

Peter von zur Muehlen †

Federal Reserve Board, Washington, DC 20551, USA; pmuehlen@verizon.net
† Retired.

Abstract: In a Ramsey policy regime, heterogeneity in beliefs about the potential costs of climate change is shown to produce policy ambiguities that alter carbon prices and taxation. Three sources of ambiguity are considered: (i) the private sector is skeptical, with beliefs that are unknown to the government, (ii) private agents have pessimistic doubts about the model, or (iii) the policy authority itself does not trust the extant scientific climate model and fears the worst. These three sources of ambiguity give rise to four potential belief regimes characterized by differentials between the government's and the private sector's inter-temporal rates of substitutions, with implications for the prices of carbon and capital, framed in terms of distorted Arrow–Debreu pricing theory that establishes an equivalence between the optimal carbon tax and the permit price of an underlying asset—the government-imposed limit on emissions in economies with cap and trade. This paper shows that in most instances, skeptical beliefs and resulting ambiguities justify higher carbon taxes and lower capital taxes to offset the private sector's increased myopia compared with rational expectations. Conversely, ambiguities created by worst-case fears in either the private sector or in government tend produce forces in the opposite direction.

Keywords: Knightian uncertainty; multiplier preferences; Ramsey planner; social planner; carbon tax; capital tax; ambiguity premium; dynamic stochastic integrated general equilibrium (DSIGE); robust Arrow–Debreu asset prices

JEL Classification: C61; C73; D80; E62; H21; H23; H63; Q32; Q38; Q54

Citation: von zur Muehlen, Peter. 2022. Prices and Taxes in a Ramsey Climate Policy Model under Heterogeneous Beliefs and Ambiguity. *Economies* 10: 257. https://doi.org/10.3390/economies10100257

Academic Editor: Ralf Fendel

Received: 2 July 2022
Accepted: 8 October 2022
Published: 17 October 2022

1. Introduction

"The trouble with the world is that the stupid are cocksure and the intelligent are full of doubt." Bertrand Russell

That global warming poses an existential threat to the future of humanity and the planet is now a universally accepted scientific fact. Yet significant segments of society still harbor doubt and skepticism about what scientists tell us, with some, including a previous US Administration, either minimizing the threat or denying it altogether, and others fearing the worst.[1]

In the minds of many, doubt or skepticism is justified by uncertainties surrounding climate science, including the extent of terrestrial carbon uptake, mankind's role in this, the relationship between carbon and temperature, and, ultimately any claimed economic damages.[2] The nature and causes of resistance to accepting bad news and environmental threats have been much discussed in the literature (see Meyer and Kunreuther (2017), Kunreuther et al. (1978), Kaufmann et al. (2017)). At the simplest level, disbelief may be motivated by myopic economic self interest, such as a refusal to contemplate the possibly enormous costs of carbon abatement, and an unwillingness to accept significant taxation on fossil energy (see Lifton (2017)). Alternatively, the much publicized threats of pending climate-related disasters have likely contributed to a fear in some sectors of society that the worst of climate catastrophes will ensue. For the analyst, some humility may be in order, because, as Millner et al. (2010) have noted, "our knowledge of the climate system is not of sufficiently high quality to be described with unique probability distributions, and that formal frameworks that account for aversion to ambiguity are normatively legitimate". In a recent paper, Pindyck (2017) warned that excessive reliance on Integrated Assessment Models (IAMs), first introduced by Nordhaus (1993), "create a perception of knowledge and precision that is illusory, possibly fooling policy-makers into thinking that the forecasts that models generate have some kind of scientific legitimacy. The same goes for any illusion that the probability distributions underling uncertainty can even be known". It is this last thought that mostly motivates the present paper.

For purposes of exposition, I shall give specific meaning to the terms *skepticism* and *doubt*. Skeptics are said to be individuals who distort their probability assessments in favor of different outcomes than predicted by science. Because such beliefs are arbitrary, they are also, in principle, unknown to the government, thereby creating an ambiguity for the policy maker (see Hansen and Sargent (2012)). By doubt I shall mean the kind of pessimism regarding the trustworthiness of a given model that Hansen and Sargent (2012) have described as emanating from a belief that nature is likely to produce worst-case outcomes. Although the literature has focused solely on ambiguities inhabiting the policy authority, both, government and private agents, may be pessimistic regarding the climate model. Pessimists in this paper make decisions by playing a game against an imagined malevolent opponent. A powerful and rational motive for pessimism is the prospect of tipping points that Dietz et al. (2021) have described as the most important source of uncertainty, one capable of throwing off all modeling descriptions and justifying preparedness for the worst.[3] In this paper, all pessimistic players use Hansen and Sargent's (2008) min-max strategies to compute their policies, meaning they are revealed to the authority, assumed to be a Ramsey planner, in the form of additional implementability constraints under which such a planner must operate.

The preceding sources of ambiguity give rise to four potential belief regimes:

1. Unknown private beliefs
   (a) Political planner: The government strategically assumes private beliefs to be true. Its ignorance of private beliefs produces ambiguity.
   (b) Paternalistic planner: The government trusts and adheres to the true model. However, the presence of unknown private beliefs creates ambiguity for policy.
   (c) Pessimistic planner: Having doubts about the model and facing unknown private beliefs, the planner confronts two sources of ambiguity.
2. Known pessimistic private beliefs
   (a) Pessimistic planner: The planner's own doubts doubts about the model and the private sector's pessimistic doubts that constrain the Ramsey planner's policy produce two sources of ambiguity.

The framework for studying policy with ambiguity is based on work by Hansen and Sargent (2005, 2007, 2008), a general rationale for pursuing robust climate policy under deep uncertainty having been provided by Workman et al. (2021). In their theory, deviations of private-sector beliefs from some approximating or reference distribution—the true scientific distribution—are represented as martingale multiplier distortions M, later defined as likelihood ratios having well defined properties. A justification for this approach to modeling ambiguity is based on a theorem by Strzalecki (2011), who axiomatized the robust control criterion of multiplier preferences introduced by Hansen and Sargent (2001), relating them to other classes of preferences studied in decision theory, in particular, the variational preferences introduced by Maccheroni et al. (2006), and proving them to be equivalent to multiplier distortions of probabilities.

As indicated earlier, the study of ambiguity in the context of climate policy is not without precedent, but discussions have generally been limited to cases when only the planner has doubts about the approximating climate-economic model. Such doubts include concerns about potential mis-specification of alternative models and ambiguity over how much weight to assign to each of these models, while agents themselves are usually assumed to have rational beliefs (see Millner et al. (2012), Brock and Durlauf (2015), Cai et al. (2013), Cai and Lontzek (2019), Anderson et al. (2013), Berger et al. (2016), Li et al. (2016), Lemoine and Traeger (2016), Rezai and van der Ploeg (2017), and Barnett et al. (2020)). Hennlock's (2009) is no exception in that, although he attributes deep uncertainty to the consumer, the government, being a direct extension of the consumer, remains the consumer's sole agent, so that, in effect, it is the planner who is modeled as having doubts about the model. In a bit of a twist to the approaches taken by other researchers, Rezai and van der Ploeg (2017) studied the implications of adopting max-min, max-max, and min-max regret policies when the planner faces alternative models ranging from science-based paradigms to denialist imaginings, concluding that max-min or min-max regret climate policies that rely on a non-skeptic view of global warming lead to a substantial and moderate amount of caution, respectively, while max-max policies produce policies that do not match the beliefs of climate skeptics. Later, Rezai and van der Ploeg (2019) applied a version of *Pascal's wager* and asked: what would an agnostic but rational planner—one who does not know or care which model is correct—do when faced with some probability that the approximating model, adhered to by so-called deniers, is false? Their conclusions are briefly described in Section 13.

The literature on optimal climate policy has generally followed the tradition of welfare analysis based on expected utility maximization within the framework of an integrated climate assessment model, with government defined as a social planner (vid. Golosov et al. 2014) seeking to maximize the expected welfare of society unconstrained by private decisions or market outcomes. An alternative is to assume that the government is a Ramsey planner, likewise seeking to maximize consumer welfare, akin to the authority introduced by Chari et al. (1994) to study optimal dynamic capital taxation, but under constraints imposed by market equilibrium.[4] This paper studies both versions of government, in which each type of planner must acknowledge the possibly distorted beliefs held by the private sector.

In its essentials, the description of the economy here follows the recent literature on optimal carbon taxation, foremost among them Nordhaus (2008), Acemoglu et al. (2012), von Below (2012), Golosov et al.'s (2014), van der Ploeg and Withagen (2014), and Belfiori (2017, 2018). The analytical framework is a by now familiar dynamic stochastic integrated general equilibrium (DSIGE) model similar to those in Anderson et al. (2013) and Golosov et al. (2014), which in turn are based on RICE—Regional Dynamic Integrated Model of Climate and the Economy—developed by Nordhaus (1993, 2008, 2007).

The government's fiscal policy tools include bond finance and taxes on carbon and capital. I include a tax on capital because within a Ramsey planning framework, capital and Pigouvian carbon taxation are tightly linked: (1) the government's stochastic discount factor for the return to capital and for the expected social cost of carbon (SCC) is the same, and (2) the SCC and taxes on capital and carbon are influenced by the same shadow price.[5]

In essence, this paper will show how the planner's ignorance of the model or of skeptical private beliefs creates an endogenous gap between the government's and the household's discount factors, leading to a gap in their respective Arrow–Debreu pricing of carbon and capital that contributes to an *ambiguity premium* over the standard certainty-equivalent formulation of the expected social cost of carbon.[6] This finding is related to recent papers by von Below (2012), Barrage (2018), and Belfiori (2017), who proved in different contexts that the optimal tax on capital is negative, and the optimal tax on carbon is higher than the standard Pigou rate, if the government's subjective discount rate is *exogenously* lower than the public sector's.[7] The underlying reason is that climate change decreases the returns to capital, so that individuals, who are too impatient from a social point of view, i.e., skeptical, do not save enough without a capital subsidy and, at the same time, burn too much fossil fuel unless the latter is taxed sufficiently. The optimal policy response is therefore to tax capital less and to tax carbon more.[8] In the reverse case, a pessimistic public motivated to over-invest and to under-utilize carbon may justify lower carbon and higher capital taxation, unless the government is also pessimistic. In this paper, any disparity in discount factors between the government and the private sector is endogenously driven, in this instance by heterogeneity in beliefs, fear of mis-specification, and ambiguity.

The next section provides formal definitions of pessimism and skepticism as understood in this paper, using multiplier preferences introduced by Hansen and Sargent (2001, 2005, 2008). Section 3 derives the Euler conditions for a consumer who may be skeptical (Section 3.5) or pessimistic (Section 3.4), as defined in Section 2. Section 4 presents a three-factor production function subject to damaging climate-related total productivity shocks in a model of a firm renting capital and purchasing energy from the household. Section 5 uses results from Sections 3.3, 3.6, and 4 to derive two versions of Hotelling's rule. Subject to constraints derived in Section 6, Section 7 presents the Ramsey planner's Euler conditions for the three belief regimes under study. Section 8 derives the possibly distorted equilibrium prices of carbon damage and capital. In anticipation of the main results, Section 9 comments on this paper's methodology and approach, which contain some innovations. Section 10 derives the expected social cost of carbon, including an ambiguity premium that governments in all four policy/belief regimes will implement. Section 11 establishes the conditions under which the planner may or may not impose an additional ambiguity-related carbon tax premium over the social cost of carbon. Section 12 presents a number of conditions under which a planner may or may not raise the subsidy rate on capital, where it will become apparent that such conditions mirror those that drive results for the carbon tax premium. Finally, before the paper's conclusion, Section 14 describes a reverse feedback from taxation to beliefs whereby a planner facing an economy with pessimistic agents is able to manipulate debt and taxes to affect pessimistic beliefs.

Throughout, references to state-conditioned distorted (including robust) Arrow–Debreu prices reflect the basic theme in this paper: that a Ramsey planner's Pigouvian tax policy under ambiguity is able to implement allocations via equilibrium pricing of an underlying asset with unknown returns that have an equivalence in a cap and trade economy. In this respect, this paper is most closely related to Barnett et al. (2020), who use asset pricing methods not only to impute market valuations but also to ascertain social valuations, as this paper intends. As in Barnett et al. (2020), the asset prices in this paper can be viewed as equivalent to shadow prices of the expected discounted values of stochastic processes impinging on the economy.

Climate policy that is motivated by a planner's own deep or Knightian uncertainty is an interesting and important topic and has been fairly exhaustively treated in the literature. Furthermore, as has also been shown elsewhere in the literature, the effects of ambiguities originating from the planner turn out not always to be clear-cut, depending on specific features related to preferences and returns in the economy. More importantly, as maintained in this paper, an equally or possibly more urgent issue for policy must be the role of beliefs held in the private sector, because they affect consumption and investment decisions, where heterogeneity between private and government beliefs will surely impact policy. That public acceptance of climate science and its policy prescriptions are not unanimous is uncontroversial and has been well documented. However, little attention has been paid to its implications, particularly its effects on ambiguity in optimal climate policy. By altering inter-temporal rates of substitution or pricing kernels that determine consumption decisions and household wealth, belief distortions in the private sector impact a planner's implementability constraints and become more salient by the addition of ambiguities that arise when those beliefs are unknowable to the government. Whatever deep uncertainty may or may not already inhabit the mind of a planner, a welfare maximizing policy authority would be remiss in ignoring the effects of private-sector belief heterogeneity and associated ambiguities on its own policy decisions.

This paper then is the first systematic attempt to analyze the policy implications of ambiguities arising from belief heterogeneity in the private sector regarding the nature of anthropomorphic climate change. Significantly, this paper distinguishes between mere skepticism producing myopic behavior in the economy and true doubt as manifested by worst-case fears leading to increased foresight. The analytic choice to investigate the implications of belief heterogeneity and associated ambiguities by evaluating the planner's Euler conditions follows the example of Anderson et al. (2013). However, the particular approach leading to a role for second-order moments as factors in optimal policy is an innovation of this paper. The covariances between multiplier distortions and variables in the economy that arise as relevant to policy will be shown to allow more detailed descriptions of the effects of ambiguity on the social cost of carbon and on carbon and capital taxation. As will become clear later, ambiguities resulting from belief heterogeneity and distortions in the private sector produce ambiguity-related premiums on the social cost of carbon and, separately, on the carbon tax and on capital subsidies.

2. Multiplier Preferences

The representative consumer and the government share a reference probability model, given a joint density $\pi(x^t)$ of the history of shocks $x^t = x_0, \cdots, x_t$, where, as detailed later in Section 4 (see Equation (21)), x_t is a climate-induced damage shock to the production economy at time t. The consumer and the government do not necessarily agree that π is the true probability. The government may doubt the model fearing the worst, and the private sector may either have similar doubts or be skeptical in some arbitrary way. Either may then choose an alternative model via some distortion of π in a manner described by Hansen and Sargent (2001, 2005, 2008), who invoke the Radon-Nikodym theorem to express any alternative model as as a non-negative measurable mapping $\hat{\pi}(x^t) = M_t(x^t)\pi(x^t)$, with $\mathbb{E}_t M_t = 1$, where, since uncertainty is realized in $t = 0$, $M_0 = 1$.[9] Being the unconditional likelihood ratio $M_t(x^t) = \frac{\hat{\pi}_t(x^t)}{\pi_t(x^t)}$ of an alternative density to $\hat{\pi}_t(x^t)$, M_t, is a martingale with respect to the reference model π, $\mathbb{E}_t M_{t+1} = M_t$, with the interpretation of a change in measure. The *distorted expectation* of x_{t+1}, given history x^t, is[10]

$$\tilde{\mathbb{E}}_t[x_{t+1}|x^t] = \mathbb{E}_t[\frac{M_{t+1}}{M_t}x_{t+1}].$$

As described in more detail below, disbelief may take two forms, either as skepticism or as pessimism, the latter being a manifestation of worst-case fears. Importantly, throughout, any alternative model $\hat{\pi}(x^t)$ is assumed to be absolutely continuous with respect to the reference model $\pi(x^t)$.[11]

Later, it will be convenient to decompose M_t by defining the conditional likelihood ratio $m_t(x^{t+1}) \equiv \frac{M(x^{t+1})}{M(x^t)}$, such that $\mathbb{E}_t m_{t+1} = 1$.

2.1. *Skepticism (Random Belief Distortion)*

As indicated in the introduction, for the purpose of this paper, skepticism refers to an arbitrary rejection of the extant approximating model $\pi(x^t)$ in favor of some other model $M_t\pi(x^t)$, where M_t is a random variable with properties previously set out, including the assumption of absolute continuity with respect to the true distribution $\pi_t(x^t)$, which means that households can be skeptics but not outright climate change deniers.

2.2. *Pessimism (Ambiguity Aversion)*

Pessimism or model doubt refers to a worst-case belief distortion $\hat{\pi}_t(x^t) = M_t\pi_t(x^t)$ derived from the consumer's having solved a min-max problem shown later.

Following Hansen and Sargent (2008), define discounted *relative entropy* conditional on date zero information as the distance $v_0(\hat{\pi}_t, \pi_t)$ between $\hat{\pi}_t$ and π_t associated with M_t over time-t information and over an infinite horizon as

$$\begin{aligned} v_0(\hat{\pi}_t, \pi_t) &= (1-\beta)\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t M_t \log M_t \\ &= \beta\mathbb{E}_0[\sum_{t=0}^{\infty} \beta^t M_t \mathbb{E}_t \frac{M_{t+1}}{M_t}(\log M_{t+1} - \log M_t)], \\ &= \beta\mathbb{E}_0[\sum_{t=0}^{\infty} \beta^t M_t \mathbb{E}_t m_{t+1} \log m_{t+1}], \end{aligned}$$

where β is the subjective discount factor. With this definition, an agent's ambiguity about π is represented by a set of joint densities $\{M_t\}_{t=0}^{\infty}$ satisfying the constraint,

$$\beta\mathbb{E}_0 M_t \mathbb{E}_t[m_{t+1} \log m_{t+1}] \le \eta, \tag{1}$$

where $\eta > 0$. In the following, a pessimistic consumer will choose a consumption plan subject to the constraint in (1).[12]

3. Households

3.1. CRRA Preferences

Consumers derive utility from consumption c_t, given a constant-elasticity preference function $u(c_t)$,[13]

$$\begin{aligned} u(c_t) &= \frac{c_t^{1-\gamma}}{1-\gamma}, \; 0 < \gamma < 1, \\ &= \log c_t; \;\; \gamma \to 1, \end{aligned}$$

with elasticities $\epsilon_{cc} = -u_{cc}c/u_c = \gamma$ (implying constant relative risk aversion).

If $\gamma = 1$ (logarithmic preferences), $\epsilon_{cc} = 1$. For the record, the constant intertemporal elasticity of substitution for consumption is $1/\gamma$. I will assume that $\gamma \le 1$, which accords with much of the literature on long-run risk—the kind this paper is most concerned with.[14]

3.2. The Household's Budget Constraint

The household owns three assets: (1) the stock of depreciating capital with a net yield $[(1-\tau_t^k(x^t))(r_t(x^t) - \delta)]k_t(x^{t-1})$, where r_t is the real rental rate on capital $k_t(x^{t-1})$ left over from last period, and τ_t^k is the tax on capital; (2) a government bond $b(x^t)$, defined as an Arrow–Debreu security promising one unit of consumption in period $t+1$, if the state is x_{t+1} and zero otherwise, and (3) the resource Q_t of fossil fuels from which it draws E_t units every period, according to the law of motion[15]

$$Q_{t+1} = Q_t - E_t, \tag{2}$$

which, by the assumption of exhaustibility, implies

$$\sum_{t=0}^{\infty} E_t \le Q_0. \tag{3}$$

The household sells E_t to the firm at a price p_t^e with after-tax revenue $(p_t^e(x^t) - \tau_t^e(x^t))E_t(x^t)$, where τ_t^e is an excise (carbon) tax per unit of energy.

The household receives income from (1) inelastically supplied labor H_t at the competitive wage $w_t = 1$, (2) rent from capital, (3) revenues from the sale of fossil energy, and (4) a lump-sum transfer from the government g_t. It spends its resources on consumption c_t, new capital k_{t+1}, and the purchase of a new Arrow security b_{t+1}, trading at the state-contingent price $\hat{p}_{t+1}(x_{t+1}|x^t)$ to be defined later. Summarizing, the household's one-period budget constraint is

$$\begin{aligned} 0 &\leq H_t(x^t) + b_t(x^t) + g_t(x^t) + R_t^k(x^t)k_t(x^{t-1}) \\ &+ (p_t^e - \tau_t^e)[Q_t(x^t) - Q_{t+1}(x^t)] - c_t(x^t) - k_{t+1}(x^t) - \mathbb{E}_t\hat{p}_{t+1}b_{t+1}(x^{t+1}) \\ &\equiv \mathcal{L}_t(x^t), \end{aligned} \tag{4}$$

where $R_t^k(x^t) = 1 + (1 - \tau_t^k)(r_t(x^t) - \delta)$ is the after-tax gross return to capital.

3.3. The Consumer's Maximization Problem

This section studies the two types of consumer introduced earlier: those who are skeptical of the model and form some arbitrary belief distortion M to the true distribution π according to Section 2.1, and those who are pessimistic and play a game against a malevolent force to determine a worst-case value for M following Section 2.2. It is convenient to set this problem up for the latter and then show the former to be a special case.

3.4. Pessimistic Consumer

The general framework for the consumer's problem is a game against some malevolent force representing extreme uncertainty about the model. Given the resource constraint in (2) and the budget constraint (4), the representative consumer solves the Lagrangian

$$\max_{\{c_t, E_t, b_{t+1}, k_{t+1}, Q_{t+1}\}} \min_{\{M_t, m_{t+1}\}} \quad \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t M_t \left(u(c_t) + \lambda_t^c \mathcal{L}_t - \frac{\beta}{\sigma^c} m_{t+1} \log m_{t+1} \right),$$

subject to $M_{t+1} = m_{t+1}M_t$, $M_0 = 1$, $\mathbb{E}_t m_{t+1} = 1$, where λ_t^c is a time-varying Lagrangian multiplier associated with the household budget constraint (4), and $0 \geq -\frac{1}{\sigma^c} < \infty$ denotes a Lagrangian shadow cost for the penalty of deviating from rational expectations represented by the approximating distribution π, also known as the Kullback-Leibler distance (see Kullback and Leibler 1951) between the two probability measures π and $\hat{\pi}$; so $0 > \sigma^c > -\infty$ may be considered the consumer's *parameter of ambiguity aversion*.[16] As in all of this paper, expectations $\mathbb{E}_t$ are taken over the measure π.

The preceding criterion has the Bellman recursion

$$\begin{aligned} \mathcal{U}_t(k_t, b_t, Q_t) &= \max_{c_t, E_t, k_{t+1}, b_{t+1}, Q_{t+1}} \min_{m_{t+1}} u(c_t) + \lambda_t^c \mathcal{L}_t(x^t) \\ &+ \beta \mathbb{E}_t \left(m_{t+1} \mathcal{U}_{t+1}(k_{t+1}, b_{t+1}, Q_{t+1}) - \frac{\beta}{\sigma^c} m_{t+1} \log m_{t+1} \right) \\ &+ \varphi_t^c (Q_t - E_t - Q_{t+1}), \end{aligned} \tag{5}$$

where φ_t^c is the Lagrangian shadow price associated with the resource constrains (2).

3.4.1. Inner Minimization

The optimal conditional likelihood ratio that minimizes (5) has the familiar exponentially twisting form[17]

$$m_{t+1}^c = \frac{e^{\sigma^c \mathcal{U}_{t+1}}}{\mathbb{E}_t e^{\sigma^c \mathcal{U}_{t+1}}},$$

or, equivalently,

$$M_{t+1}^c = \frac{e^{\sigma^c \mathcal{U}_{t+1}}}{\mathbb{E}_t e^{\sigma^c \mathcal{U}_{t+1}}} M_t^c. \tag{6}$$

The worst-case martingale distortion in (6) is *pessimistic* in that it attaches higher probabilities to histories with low continuation utilities and lower probabilities to histories with high continuation utilities. Notably, because of its dependence on continuation values, it is also a function of future distortions m_{t+j}^c and future decisions by both the consumer and the government that, as shown later, determine equilibrium prices in the economy.

3.4.2. Outer Maximization with Implied Risk-Sensitive Recursion

Substituting m_{t+1}^c from (6) into (5) produces the *risk-sensitive* recursion[18]

$$\mathcal{U}(k_t, b_t, Q_t) \quad = \quad \max_{c_t, E_t, k_{t+1}, b_{t+1}, Q_{t+1}} u(c_t) + \frac{\beta}{\sigma^c} \mathbb{E}_t \log e^{\sigma^c \mathcal{U}(k_{t+1}, b_{t+1}, Q_{t+1})} + \lambda_t^c \mathcal{L}_t. \tag{7}$$

The first-order conditions (FONCs) for the flow variables $\{c_t, E_t\}$ are

$$\lambda_t^c \quad = \quad u_{c_t}, \tag{8}$$
$$(p_t^e - \tau_t^e)\lambda_t^c \quad = \quad \varphi_t^c. \tag{9}$$

The envelope conditions are,

$$\mathcal{U}_{k_t} \quad = \quad \lambda_t^c R_t^k = u_{c_t} R_t^k, \tag{10}$$
$$\mathcal{U}_{b_t} \quad = \quad \lambda_t^c = u_{c_t}, \tag{11}$$
$$\mathcal{U}_{Q_t} \quad = \quad \varphi_t^c. \tag{12}$$

In conjunction with these envelope conditions, the first-order conditions for the stock variables b_{t+1}, k_{t+1}, and Q_{t+1} are, respectively,

$$\hat{p}_{t+1} \quad = \quad \beta m_{t+1}^c \frac{u_{c_{t+1}}}{u_{c_t}}, \tag{13}$$
$$1 \quad = \quad \beta \mathbb{E}_t m_{t+1}^c \frac{u_{c_{t+1}}}{u_{c_t}} R_{t+1}^k = \mathbb{E}_t \hat{p}_{t+1} R_{t+1}^k, \tag{14}$$
$$\varphi_t^c \quad = \quad \beta \mathbb{E}_t m_{t+1}^c \varphi_{t+1}^c. \tag{15}$$

Note that $\hat{p}_{t+1}$ is the one-period worst-case equilibrium price of a state-contingent claim as shown later in (18). According to the preceding conditions, the price of such a claim is determined by continuation utilities and the random climate cost shock x_{t+1}. This is information that the government can exploit to determine optimal fiscal policy.

3.5. Skeptical Consumer

Let M^s denote a skeptical consumer's belief distortion. As stated earlier, skepticism represents an arbitrary random distortion $\hat{\pi}_t \equiv M_t^s \pi_t$ of the true distribution π_t. From the point of view of society (or the government), M_{t+1}^s is an unknowable exogenous variable. However, this household is sure of its beliefs and evaluates (5) by disabling the penalty on belief distortions and letting $\sigma^c \to 0$.

The Euler equations are the same as before, except that $m = m^s$ is random and unknown to the authorities, and the equilibrium price $\hat{p}$ of a state-contingent claim based on skepticism as defined here is unrelated to the consumer's continuation utility. Additionally, since m_{t+1}^s is random, it is unrelated to x_{t+1}.

3.6. Arrow–Debreu Prices under Belief Distortions

The belief distortions of probabilities, ranging from skepticism to deep uncertainty, treated in this paper, add an important dimension to stochastic discounting. This section shows how to construct their corresponding asset-price measures.

For preferences $u(c_t)$ distorted by a martingale process M_{t+j} (either M_{t+j}^s or M_{t+j}^c), the j-period-ahead stochastic discount factor (MSDF) is,

$$\hat{\varrho}_{t+j,t} \quad = \quad \beta^j \frac{M_{t+j}}{M_t} \frac{u_{c_{t+j}}}{u_{c_t}} \equiv \frac{M_{t+j}}{M_t} \varrho_{t+j,t}, \tag{16}$$

where $\varrho_{t+j,t}$ is the pricing kernel under rational expectations. When $j = 1$, this becomes the familiar one-period-ahead stochastic discount factor SDF.

Let $\hat{q}_{t+j}(x^{t+j})$ be the j-period numeraire

$$\hat{q}_{t+j}(x^{t+j}) \equiv \beta^j M_{t+j}\pi_{t+j}(x^{t+j})\frac{u_{c_{t+j}}(x^{t+j})}{u_{c_0}(x^0)}, \quad \hat{q}_0 = 1, M_0 = 1, \tag{17}$$

and define

$$\begin{aligned}\hat{p}_{t+j,t}(x_{t+j}|x^t) &\equiv \frac{\hat{q}_{t+j}(x^{t+j})}{\hat{q}_t(x^t)} = \beta^j\left(\prod_{i=1}^{j} m_{t+i}\frac{u_c(x^{t+i})}{u_c(x^t)}\right)\pi_{t+j}(x^{t+j}) \\ &= \frac{M_{t+j}}{M_t}\varrho_{t+j,t}\pi_{t+j}(x^{t+j})\end{aligned}$$

as the market's distorted $t+j$ equilibrium price in (13) of an Arrow–Debreu security in terms of consumption at history x^t, or equivalently,

$$\hat{p}_{t+j,t}(x_{t+j}|x^t) = \hat{\varrho}_{t+j,t}\pi_{t+j}(x^t), \tag{18}$$

which, for $j = 1$, also corresponds to the first-order condition for capital in (14). For future reference, denote the undistorted rational expectations price by

$$p_{t+j,t}(x_{t+j}|x^t) = \varrho_{t+j,t}\pi_{t+j}(x^t). \tag{19}$$

4. Firms

The economy's output is produced by a continuum of atomistic firms with a Cobb-Douglas production technology:

$$Y_t = (1 - D_t(T_t - T_0))F(k_t, H_t, E_t, Q_t) = (1 - D_t(Q_t))k_t{}^{\alpha}E_t^{\nu}H_t^{1-\alpha-\nu}, \tag{20}$$

where k_t is the stock of capital, H_t is hours of labor, E_t is the flow of fossil energy, and Q_t is the remaining stock of carbon energy in the ground. As explained presently, $D_t(Q_t)$ is a damage function measuring the proportion of GDP lost due to the change in average global temperatures $T_t - T_0$ since the beginning of the Industrial Revolution at $t = 0$.

Dietz and Venmans (2019) observe that "climate has delivered two important and related insights. First, global warming appears to be approximately linearly proportional to cumulative emissions of carbon dioxide. Second, the temperature response to an emission of CO_2 appears to be approximately instantaneous and then constant as a function of time". This conclusion accords with Matthews et al. (2012), and Collins et al. (2013), who earlier defined the *Transient Climate Response to Cumulative Carbon Emissions* (TCRE)[19]

$$\lambda_t^{TC} = \frac{T_t - T_0}{Q_0 - Q_t},$$

where $0 \leq Q_0 - Q_t = \sum_{i=0}^{t-1} E_i$ is accumulated carbon emissions since the beginning of the Industrial Revolution.[20] The TCRE parameter is generally assumed to be a stochastic variable, due to uncertainties surrounding climate modeling.

The preceding motivates a damage function having the following exponential form

$$D_t(Q_t) = e^{\zeta_t(T_t - T_0)} = e^{\zeta_t\lambda_t^{TC}(Q_0 - Q_t)} \equiv e^{x_t(Q_0 - Q_t)}, \quad \frac{\partial D_t}{\partial Q_t} = -x_t e^{x_t(Q_0 - Q_t)} < 0, \tag{21}$$

where ζ_t is a stochastic parameter that translates the damaging effects of temperature changes into units of GDP, and $x_t = \zeta_t\lambda_t^{TC}$ combines the damage parameter with the TCRE parameter λ_t^{TC}.[21] As posited in Section 2, I shall consider x_t to be a random variable with either known distribution $\pi_t(x_t)$ or unknown distribution giving rise to ambiguities described earlier. Confining the source of economic damage to a single catch-all variable follows the practice of a number of authors, including Li et al. (2016), who create a single source of model uncertainty with a stochastic variable γ that reduces their end-of-period capital stock. Others who have used similar exponential formulations include Golosov et al.'s (2014), and Anderson et al. (2013).[22] To the extent that there exist possibilities of catastrophic tipping points that would likely upend all calculations and planning, the underlying distribution of x_t may be taken as Knightian and unknowable by either the private sector or the government or both. Tipping points, analyzed by Lemoine and Traeger (2016), and Cai et al. (2013), are abrupt nonlinear climate disruptions that pose a potentially existential threat to humanity in ways that may override concerns with belief and skepticism. A further source of extreme uncertainty is *polar amplification* analyzed by Brock and Xepapadeas (2017).

The typical firm rents capital k_t from consumers and buys energy E_t from households in order to maximize an expected infinite stream of profits. Hours of labor H_t are supplied inelastically at the going wage 1—a standard assumption in this literature (see Golosov et al. (2014) and Li et al. (2016)).

$$\begin{aligned} & \max_{k_t,H_t,E_t} \mathbb{E}_o \sum_{t=0}^{\infty} \hat{q}(x_t)[Y_t - r_t k_t - H_t - p_t^e E_t], \\ = & \max_{k_t,H_t,E_t} \mathbb{E}_o \sum_{t=0}^{\infty} \hat{q}_t(x_t)\Big((1 - e^{-x_t(Q_t - Q_0)}){k_t}^{\alpha} E_t^{\nu} H_t^{1-\alpha-\nu} - r_t k_t - H_t - p_t^e E_t\Big), \end{aligned}$$

where $\hat{q}(x_t|x_0)$ is the belief-distorted and possibly robust numeraire defined earlier in (17). The first-order conditions with respect to $\{k_t, H_t, E_t\}$ are

$$r_t = Y_{k_t} = \alpha \frac{Y_t}{k_t}, \tag{22}$$

$$1 = Y_{H_t} = (1 - \alpha - \nu)\frac{Y_t}{H_t}, \tag{23}$$

$$p_t^e = Y_{E_t} = \nu \frac{Y_t}{E_t}. \tag{24}$$

5. Two Versions of Hotelling's Rule

Denote by R_{t+1}^e the rate of return to energy, net of the carbon tax τ_t^c,

$$R_{t+1}^e = \frac{p_{t+1}^e - \tau_{t+1}^e}{p_t^e - \tau_t^e}. \tag{25}$$

Additionally, since, for either $m = m^s$ or $m = m^c$, (9) and (15) imply

$$p_t^e - \tau_t^e = \beta \mathbb{E}_t m_{t+1} \frac{u_{c_{t+1}}}{u_{c_t}} (p_{t+1}^e - \tau_{t+1}^e), \tag{26}$$

it follows that

$$\mathbb{E}_t \hat{p}_{t+1} R_{t+1}^e = 1, \tag{27}$$

where expectations are taken with respect to $\pi(x^t)$, as before. Substituting (24) for period t and $t+1$ in (26) and using the result $p^e = Y_E$ from (24), yields the firm's dynamic rule for optimal energy use, given taxes,

$$Y_{E_t} - \tau_t^e = \mathbb{E}_t \hat{\varrho}_{t+1}(Y_{E_{t+1}} - \tau_{t+1}^e), \tag{28}$$

which is a version of Hotelling's rule that will become relevant later for determining the social cost of carbon.

If $m_{t+1} = 1$, the preceding equation confirms Hotelling's original formula that with zero or constant taxes, energy consumption falls over time at the subjective rate of discount. If consumers are climate-skeptic, a multiplier $m_{t+1}^s < 1$ effectively lowers the private-sector's discount factor for future benefits of fossil fuel in favor of current returns, leading to increased current consumption relative to rational expectations. If the consumer is pessimistic with $m_{t+1} = m^c_{t+1}$, energy expenditures are delayed, following the same argument.

Hotelling's (1931) original rule states that the price of an exhaustible resource net of extraction costs should rise at the rate of interest, which, on average, is above the rate of real GDP growth. By Equation (14), the t-period Arrow–Debreu price of a claim on a unit of capital is

$$P^k_t \quad = \quad \mathbb{E}_t \sum_{j=0}^{\infty} \hat{p}_{t+j,t} R^k_{t+j}, \tag{29}$$

while Equation (27) implies that the t-period Arrow–Debreu price of a claim on a unit of carbon energy is

$$P^e_t \quad = \quad \mathbb{E}_t \sum_{j=0}^{\infty} \hat{p}_{t+j,t} R^e_{t+j}. \tag{30}$$

From Equations (14) and (27) follows a version of Hotelling's rule in terms of returns:

$$Version\ 1: \qquad R^e_{t+1} = R^k_{t+1}, \ \forall t, \tag{31}$$

while Equations (29) and (30) restate Hotelling's rule in terms of Arrow–Debreu contingent prices:

$$Version\ 2: \qquad P^e_{t+1} = P^k_{t+1}, \ \forall t. \tag{32}$$

Note that in each case, returns and prices are distorted by the consumer's beliefs, be they skeptical m^s or pessimistic m^c.

These two versions of Hotelling's rule (31) and (32) constitute binding no-arbitrage conditions that require prices and returns to capital and fossil fuel to be equal: in equilibrium, the return to fossil fuel left in the ground for one more period equals the return to the next unit of capital. This rule illuminates two important features of a competitive market for exhaustible energy when there is uncertainty and skepticism regarding the underlying model: (i) the pricing of energy resources continues to obey the laws of asset markets requiring equality of returns to all activities, including capital, bonds, and energy stores, but (ii) the market now uses a martingale-distorted and possibly robust stochastic discount factor to evaluate the expected future returns to all assets.

Remark 1. *Hotelling's rule assures efficient allocation but not necessarily socially optimal outcomes if it fails to internalize costs to society created by private economic activity. Later, Section 11 shows how the government can remedy this failure with a public version of Hotelling's rule that includes a social accounting of all costs.*

The next few sections discuss how a Ramsey planner implements competitive equilibrium depending on assumptions about heterogeneity in beliefs, the degree of doubt about the model by either the government or the public, and by how ignorant the planner is about private beliefs.

6. The Ramsey Planner's Constraints

6.1. National Income Identity and the Government Budget Constraint

In all belief regimes considered here, a Ramsey planner commits to policy in period 0 by choosing a competitive equilibrium that maximizes the consumer's expected utility over time. This means the government chooses allocations that satisfy the natural resource constraint (2) and (3) and the national income resource constraint that output exhausts consumption plus investment plus government spending, which, to keep things simple, is assumed to be entirely devoted to a lump-sum transfer g_t,

$$Y_t = c_t + k_{t+1} - (1-\delta)k_t + g_t. \tag{33}$$

The government's budget constraint requires that spending and the redemption of bonds from the preceding period be covered by tax receipts, lump-sum transfers, and new issuance of bonds:

$$b_t \quad = \quad \mathbb{E}_t \hat{p}_{t+1} b_{t+1} - g_t + \tau_t^e E_t + \tau_t^k (r_t - \delta) k_t.$$

When $b_t < 0$, the government is a lender. Solved forward, the preceding equation becomes the government's dynamic budget constraint,

$$b_t \leq \mathbb{E}_t \sum_{j=0}^{\infty} \hat{q}_{t+j} [\tau_{t+j}^k (r_{t+j} - \delta) k_{t+j} - \tau_{t+j}^e (Q_{t+j+1} - Q_{t+j}) - g_{t+j}]. \tag{34}$$

Note the added term involving receipts from the carbon tax.

6.2. The Ramsey Planner's Implementability Constraints

In solving for optimal taxation in the case of a Ramsey planner, I use a so-called primal approach due to Chamley (1986) that searches directly for allocations by solving the government's problem subject to an *implementability* constraint. A starting point is the household's dynamic budget constraint (4), which, when solved forward for b_t, utilizing the no-arbitrage condition (14) and a no-Ponzi game condition, yields the intertemporal budget constraint[23]

$$b_t \quad \geq \quad \mathbb{E}_t \lim_{T\to\infty} \sum_{j=0}^{T-1} \hat{p}_{t+j} [c_{t+j} - H_{t+j} - g_{t+j}]. \tag{35}$$

Let $\mathcal{W}_t$ denote household wealth in period t, composed of government bonds, the after-tax equity value of physical capital, and fossil fuels still in the ground, valued at current after-tax energy prices:

$$\mathcal{W}_t \quad \equiv \quad b_t + R_t^k k_t + (p_t^e - \tau_t^e) Q_t. \tag{36}$$

Appendix B shows that

$$\begin{aligned} \mathcal{W}_t \quad &\geq \quad c_t - H_t - g_t + \beta \mathbb{E}_t m_{t+1} \frac{u_{c_{t+1}}}{u_{c_t}} \mathcal{W}_{t+1} \\ &= \quad c_t - H_t - g_t + \mathbb{E}_t \hat{p}_{t+1} \mathcal{W}_{t+1}, \end{aligned} \tag{37}$$

where $\hat{p}_{t+j,t}$ is the market's previously defined distorted $t+j$ equilibrium price of an Arrow–Debreu security in terms of consumption at history x^t. Solved forward, household wealth is

$$\mathcal{W}_t \quad \geq \quad \mathbb{E}_t \sum_{j=0}^{\infty} \hat{p}_{t+j,t} [c_{t+j} - H_{t+j} - g_{t+j}]. \tag{38}$$

For later use, it is convenient to define the marginal-utility-of-consumption-scaled market value of wealth $\Upsilon_t = u_{c_t} \mathcal{W}_t$, so that the Ramsey planner's implementability constraint (37) becomes, equivalently,

$$\Upsilon_t \quad \geq \quad \Omega_t + \mathbb{E}_t m_{t+1} \Upsilon_{t+1}, \tag{39}$$

where

$$\Omega_t \quad = \quad u_{c_t}[c_t - H_t - g_t], \tag{40}$$

with derivative,

$$\begin{aligned}\Omega_{c_t} &= \frac{u_{cc_t}}{u_{c_t}}(c_t - H_t - g_t)u_{c_t} + u_{c_t} \\ &= [1 - \gamma c_t^{-1}(c_t - H_t - g_t)]u_{c_t} = [1 - \gamma + \gamma\frac{H_t + g_t}{c_t}]u_{c_t} \geq 0,\end{aligned} \tag{41}$$

since, from Section 3.1, $\frac{u_{cc_t}}{u_{c_t}} = -\gamma/c_t$. Note that if the constraint (39) is nonbinding, the government is the social planner widely treated in the literature.

Solved forward,

$$Y_t = \mathbb{E}_t u_{c,t} \sum_{j=0}^{\infty} \hat{p}_{t+j}(c_{t+j} - H_{t+j} - g_{t+j}) = u_{ct} b_t, \tag{42}$$

is the government's surplus valued in terms of the marginal utility of consumption.

Finally, by its very definition, a Ramsey planner heeds all equilibrium constraints imposed by competitive markets. In particular, when consumers are pessimistic with belief distortions defined in terms of their continuation values, where $m_t = m^c_{t+1} = \frac{e^{\sigma^c \mathcal{U}_{t+1}}}{\mathbb{E}_t e^{\sigma^c \mathcal{U}_{t+1}}}$, the planner faces two additional implementability constraints that come from (6) and (7):

$$M^c_{t+1} \quad = \quad \frac{e^{\sigma^c \mathcal{U}_{t+1}}}{\mathbb{E}_t e^{\sigma^c \mathcal{U}_{t+1}}} M^c_t, \tag{43}$$

$$\mathcal{U}_t \quad = \quad u(c_t) + \frac{\beta}{\sigma^c} \log \mathbb{E}_t e^{\sigma^c \mathcal{U}_{t+1}}. \tag{44}$$

7. Ramsey Planning in Four Belief Regimes

Subject to the fossil resource constraint (2) and (3), the national income identity (33), and the implementability constraint (39), the taxing Ramsey authority chooses $\{c, E, k, Q, Y\}$ to maximize society's expected welfare, and $\{N, n\}$ to minimize discounted *relative entropy* defined in (1),

$$\max_{c,E,k,Q,Y} \min_{N,n,} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t N_t(x^t) \Big[u(c_t) - \frac{\beta}{\sigma}(\mathbb{E}_t n_{t+1} \log n_{t+1}) \Big], \tag{45}$$

subject to $N_{t+1} = n_{t+1} N_t$, $N_0 = 1$, $\mathbb{E}_t n_{t+1} = 1$, and $0 > \sigma > -\infty$, where N_t is the government's martingale multiplier, equivalent to M_t defined earlier for the consumer, and $\sigma < 0$ is the planner's *parameter of ambiguity aversion*.

Exploiting a linear homogeneity property such that $V(Y_t, k_t, Q_t, N_t) = N_t \mathcal{V}_t(Y_t, k_t, Q_t)$, the problem may be cast as the recursive (Bellman) Lagrangian

$$\begin{aligned}\mathcal{V}_t(Y_t, k_t, Q_t) \quad &= \quad \max_{c_t, E_t, k_{t+1}, Q_{t+1} Y_{t+1}} \min_{n_{t+1}} u(c_t) - \frac{\beta}{\sigma} \mathbb{E}_t n_{t+1} \log n_{t+1} \\ &+ \quad \beta \mathbb{E}_t n_{t+1} \mathcal{V}_{t+1}(Y_{t+1}, k_{t+1}, Q_{t+1}) \\ &+ \quad \Phi_t[\Omega_t + \mathbb{E}_t m_{t+1} Y_{t+1} - Y_t] + \mathcal{M}_t,\end{aligned} \tag{46}$$

where

$$\mathcal{M}_t \quad = \quad \begin{pmatrix} \lambda_t[(1 - e^{x_t(Q_0 - Q_t)})F(k_t, H_t, E_t) + (1-\delta)k_t - c_t - g_t - k_{t+1}] \\ \varphi_t[Q_t - E_t - Q_{t+1}] \\ \upsilon[Q_0 - \sum_{i=0}^{\infty} E_i] \end{pmatrix},$$

is a group of Lagrangian constraints and λ_t, Φ_t, φ_t are non-negative Lagrangian co-state variables, and v is constant.[24]

The preceding problem constitutes a Stackelberg game between the government as the leader and the private sector as the follower. Embedded in this game is another game with a minimizing opponent to represent worst-case outcomes. This latter game, framed as *inner minimization*, may be played by either government or consumers, as in Section 3.4, or even both, depending on respective attitudes toward extreme risk. The solution to this subgame implies a plan for *outer maximization* of an indirect value function, typically a *risk sensitivity recursion*, such as the one introduced in Section 3.4.2.

7.1. Political Planner

In this example, consumers are assumed to have arbitrarily distorted—skeptical—beliefs about the approximating scientific model π, as described in Section 2.1. Importantly, the government is ignorant of these beliefs. The reason for its ignorance is fundamental: as noted by Hansen and Sargent (2012), the government's ambiguity arises because the possible space of models that are unknown to the fiscal planner but known to the consumer is so vast that it is impossible to infer the private sector's probability model from finite data, providing a motive to construct robust climate and fiscal policies by solving a so-called multiplier problem that protects against worst-case belief distortions. As Hansen and Sargent (2012) put it, the government's ignorance of private beliefs is akin to a set or cloud of probability distributions over events x centered on the reference or approximating density π constrained by a discounted relative entropy set of probability distributions reflecting the unknown beliefs of the private sector.[25] Its ignorance notwithstanding, this political planner acts under the assumption that the unknown beliefs are true.

With this in mind, the recursion (46) becomes

$$\begin{aligned} \mathcal{V}_t(Y_t, k_t, Q_t) &= \max_{c_t, E_t, k_{t+1}, Q_{t+1} Y_{t+1}} \min_{n_{t+1}} u(c_t) - \frac{\beta}{\sigma}\mathbb{E}_t n_{t+1} \log n_{t+1} \\ &+ \beta \mathbb{E}_t n_{t+1} \mathcal{V}_{t+1}(Y_{t+1}, k_{t+1}, Q_{t+1}) \\ &+ \Phi_t[\Omega_t + \mathbb{E}_t n_{t+1} Y_{t+1} - Y_t] + \mathcal{M}_t. \end{aligned} \tag{47}$$

The political government's acceptance of private-sector beliefs M_t as true means that m_{t+1}, which multiplies Y_{t+1} in the implementability constraint, is set equal to n_{t+1}—the authority adopts the consumer's distortion as its own.[26] Ambiguity is activated with the entropy constraint penalizing deviations of distorted beliefs from true beliefs (1), where $-\frac{1}{\sigma}$ is a Lagrangian multiplier, and $\sigma < 0$ measures the planner's ambiguity aversion. As σ moves toward $-\infty$, the government's preference for robustness rises. As σ approaches a break-down $\underline{\sigma} < 0$ from above, the government's concern about distortions to expectations is maximal. Conversely, as σ approaches zero, the government's preference for robustness diminishes, until, in the limit the government fully adopts the approximating reference model as true and not subject to doubt.

7.1.1. Inner Game with Nature

Ambiguity for this planner produces an *ex post* worst-case probability model with distorted *ex post homogeneity* in beliefs between the government and the private sector.

Minimization of (47) with respect to n_{t+1} leads to

$$n_{t+1} = \frac{e^{\sigma[\mathcal{V}_{t+1} + \Phi_t Y_{t+1}]}}{\mathbb{E}_t e^{\sigma[\mathcal{V}_{t+1} + \Phi_t Y_{t+1}]}} \equiv n^{PO}_{t+1}, \tag{48}$$

with limits $\lim_{\sigma \downarrow -\infty} n^{PO}_{t+1} \to 0$, and $\lim_{\sigma \uparrow 0} n^{PO}_{t+1} \to 1$, indicating that the conditional likelihood ratio n^{PO} is inversely related to the intensity of the planner's doubts about private beliefs, approaching 1 as doubt ceases.

7.1.2. Outer Maximization with Implied Risk-Sensitive Recursion

The continuation value $\mathcal{V}_{t+1}$ in the exponent of the formula for n^{PO} comes from the planner's distrust of the reference model itself, while continuation wealth Υ_{t+1} points to the planner's ignorance of private-sector beliefs. Formula (48), indicates that, given $\sigma < 0$, ambiguity leads this robust planner to put more probability weight on histories with low continuation values $\mathcal{V}_{t+1}$ and Υ_{t+1}, and lower probabilities on histories with high continuation values.

Substituting (48) for both n_{t+1} and m_{t+1} in (47) produces a variation on Hansen and Sargent's (1995) discounted risk-sensitive recursion, one that is augmented with household wealth Υ, a forward-looking entity valued at the shadow price Φ, in the exponent,[27]

$$\begin{aligned}\mathcal{V}(\Upsilon_t, k_t, Q_t) &= \max_{c_t, E_t, k_{t+1}, Q_{t+1}, \Upsilon_{t+1}} u(c_t) + \Phi_t[\Omega_t - \Upsilon_t] + \mathcal{M}_t \\ &+ \frac{\beta}{\sigma} \log \mathbb{E}_t e^{\sigma(\mathcal{V}_{t+1} + \Phi_t \Upsilon_{t+1})}. \end{aligned} \tag{49}$$

The distortion n^{PO} that attains the minimum of the right side of (49) tilts the x_{t+1} distribution exponentially toward lower continuation values via multiplication of $\pi(x_{t+1})$ by n^{PO}_{t+1} in (48).

The first-order conditions (FONCs) for $\{c_t, E_t\}$ for an interior maximum are: $\forall t \geq 1$,

Flows

$$\begin{aligned} c_t: &\quad \lambda_t = u_c + \Omega_{c_t}\Phi_t, & (50)\\ E_t: &\quad Y_{E_t}\lambda_t = \varphi_t + v, & (51)\end{aligned}$$

where $Y_{E_t} = \nu\frac{Y_t}{E_t}$, is the marginal product of fossil energy based on the Cobb-Douglas assumption.

Envelope conditions

$$\begin{aligned} k_t: &\quad \mathcal{V}_{k_t} = (1 - \delta + Y_{k_t})\lambda_t, & (52)\\ Q_t: &\quad \mathcal{V}_{Q_t} = \varphi_t + \lambda_t x_t D_t F_t, & (53)\\ \Upsilon_t: &\quad \mathcal{V}_{\Upsilon_t} = -\Phi_t. & (54)\end{aligned}$$

Note that $\lambda \geq 0$ and $\varphi \geq 0$ imply $V_k \geq 0$ and $\mathcal{V}_{Q_t} \geq 0$, respectively. Further, $\Phi \geq 0$ implies $V_\Upsilon \leq 0$.

Stocks

$$\begin{aligned} &k_{t+1}: \\ &\quad 1 = \frac{\beta}{\lambda_t}\mathbb{E}_t n^{PO}_{t+1}\mathcal{V}_{k_{t+1}} = \beta\mathbb{E}_t n^{PO}_{t+1}\frac{\lambda_{t+1}}{\lambda_t}(1 - \delta + Y_{k_{t+1}}), & (55)\\ &Q_{t+1}: \\ &\quad \varphi_t = \beta\mathbb{E}_t n^{PO}_{t+1}\mathcal{V}_{Q_{t+1}} = \beta\mathbb{E}_t n^{PO}_{t+1}(\varphi_{t+1} + \lambda_{t+1}x_{t+1}D_{t+1}F_{t+1}), & (56)\\ &\Upsilon_{t+1}: \\ &\quad \Phi_t = -\mathcal{V}_{\Upsilon_{t+1}} = \Phi_{t+1} = \bar{\Phi}, & (57)\end{aligned}$$

where the second equality in each equation follows from the envelope conditions (52)–(54), and (56), giving a marginal utility valuation of the benefit of fossil energy net of climate damages. Given (57), the Lagrangian multiplier Φ_t is a constant and, by the Kuhn-Tucker condition, is zero if the wealth constraint is nonbinding, indicating that the planner is a social and not a Ramsey planner.

Note that because both the private sector's and the planner's probabilities are twisted by a martingale multiplier, the ambiguity for this planner effectively delivers *ex post* a model of endogenously distorted *homogeneous* beliefs. Hansen and Sargent (2012) emphasize that n^{PO} is not intended to "solve" an impossible inference problem, and being the planner's *cautious* inference about unknown private beliefs, should be viewed as merely a device to construct a robust Ramsey policy. If the planner were to solve the private sector's Euler equations using the minimizing n^{PO} in order to derive its *ex post* decision rules for consumption, labor, and energy, it would not necessarily end up reproducing their observed values, meaning that private beliefs cannot be reverse-engineered from such observations.

7.1.3. Paternalistic Planner

Woodford (2010) originally introduced a monetary authority facing a type of ambiguity described here: while trusting the reference model π, the government is ignorant about the private sector's distorted beliefs m^s. It expresses its ambiguity by setting $\sigma < 0$. Trust in its own model means that the planner sets n_{t+1} multiplying continuation value $\mathcal{V}_{t+1}$ equal to unity.[28]

7.1.4. Inner Game with Nature

The minimizing value of n_{t+1} is

$$n_{t+1} = \frac{e^{\sigma[\Phi_t Y_{t+1}]}}{\mathbb{E}_t e^{\sigma[\Phi_t Y_{t+1}]}} \equiv n^{PA}_{t+1}, \tag{58}$$

with limits $\lim_{\sigma\downarrow-\infty} n^{PA}_{t+1} \to 0$, and $\lim_{\sigma\uparrow 0} n^{PA}_{t+1} \to 1$, indicating that the conditional likelihood ratio n^{PA} is inversely related to the intensity of the planner's doubt about private beliefs, approaching 1 as doubt ceases.

An important distinction between the political planner and the paternalistic planner with ambiguity is that here, the planner's worst-case distortion of beliefs n^{PA} is solely determined by continuation values of wealth, since the only sources of ambiguity are private-sector beliefs that distort the consumer's expectation of future household wealth. Formula (58) instructs us that, via n^{PA}, a robust paternalistic planner assigns greater probability weights to histories with low continuation values of wealth, weighted with marginal utility of consumption in (39).

7.1.5. Outer Maximization with Implied Risk-Sensitive Recursion

Substitution of the formula for n^{PA}_{t+1} in (46) implies another variation on Hansen and Sargent's (1995) discounted risk-sensitive recursion,

$$\begin{aligned} \mathcal{V}(Y_t, k_t, Q_t) &= \max_{c_t, E_t, k_{t+1}, Q_{t+1}, Y_{t+1}} \{u(c_t) + \Phi_t[\Omega_t - Y_t] + \frac{\beta}{\sigma}\mathbb{E}_t \log e^{\sigma\Phi_t Y_{t+1}} + \mathcal{M}_t \\ &+ \beta\mathbb{E}_t \mathcal{V}_{t+1}(Y_{t+1}, k_{t+1}, Q_{t+1})\}, \end{aligned}$$

which differs from the recursion (49) for a political planner in that the exponent does not include the continuation value $\mathcal{V}_{t+1}$, since, here, risk sensitivity does not apply to the planner's own trusted model.

The FONCs for $\{c, E\}$ are previously given by (50) and (51). The envelope conditions are also the same as before. However, for choosing $\{k_{t+1}, Q_{t+1}, Y_{t+1}\}$, the government must solve

$$1 = \mathbb{E}_t \beta \frac{\lambda_{t+1}}{\lambda_t}(1 - \delta + Y_{k_{t+1}}), \tag{59}$$

$$\varphi_t = \beta\mathbb{E}_t \mathcal{V}_{Q_{t+1}} = \beta\mathbb{E}_t(\varphi_{t+1} + \lambda_{t+1} x_{t+1} D_{t+1} F_{t+1}), \tag{60}$$

$$\Phi_t = -\mathcal{V}_{Y_{t+1}} = \Phi_{t+1} = n^{PA}_{t+1}\Phi_t. \tag{61}$$

From (61) follows that Φ_t is a martingale: $\mathbb{E}_t\Phi_{t+1} = \mathbb{E}_t n^{PA}_{t+1}\Phi_t = \Phi_t$, unless the wealth constraint $0 \leq [\Omega_t + \mathbb{E}_t n^{PA}_{t+1} Y_{t+1} - Y_t]$ is non-binding, in which case $\Phi_t = 0$, and the government reverts to a social planner.

7.2. Pessimistic Planner, Skeptical Consumer

This section treats a variation of the paternalistic planner in Section 7.1.3, where, instead of trusting the approximating model π, the authority has pessimistic doubts about it, meaning it now faces two kinds of ambiguity: one that derives from its ignorance of private beliefs, indexed by M, and the other stemming from its own doubts about the model, indexed by N. Accordingly, the planner minimizes with respect to both m and n, given two Lagrange penalty functions. For the sake of simplicity, I will assume that in the following recursion, a single risk sensitivity σ applies to both kinds of ambiguities:

$$
\begin{aligned}
\mathcal{V}_t(\Upsilon_t, k_t, Q_t) &= \max_{c_t, E_t, k_{t+1}, Q_{t+1} \Upsilon_{t+1}} \min_{n_{t+1}, m_{t+1}} u(c_t) \\
&- \frac{\beta}{\sigma}\mathbb{E}_t(n_{t+1}\log n_{t+1} + m_{t+1}\log m_{t+1}) \\
&+ \beta\mathbb{E}_t n_{t+1}\mathcal{V}_{t+1}(\Upsilon_{t+1}, k_{t+1}, Q_{t+1}) \\
&+ \Phi_t[\Omega_t + \mathbb{E}_t m_{t+1}\Upsilon_{t+1} - \Upsilon_t] + \mathcal{M}_t.
\end{aligned}
$$

7.2.1. Inner Minimization

Minimization with respect to n and m produces the following worst-case multipliers

$$
n^p_{t+1} = \frac{e^{\sigma\mathcal{V}_{t+1}}}{\mathbb{E}_t e^{\sigma\mathcal{V}_{t+1}}}, \tag{62}
$$

$$
m^p_{t+1} = \frac{e^{\sigma\Phi_t\Upsilon_{t+1}}}{\mathbb{E}_t e^{\sigma\Phi_t\Upsilon_{t+1}}}. \tag{63}
$$

7.2.2. Outer Maximization with Implied Risk-Sensitive Recursion

Substitution as before yields the risk-sensitive recursion

$$
\begin{aligned}
\mathcal{V}(\Upsilon_t, k_t, Q_t) &= \max_{c_t, H_t, E_t, k_{t+1}, Q_{t+1}, \Upsilon_{t+1}} u(c_t, H_t) + \Phi_t[\Omega_t - \Upsilon_t] + \mathcal{M}_t \\
&+ \frac{\beta}{\sigma}(\log\mathbb{E}_t e^{\sigma\mathcal{V}_{t+1}} + \log\mathbb{E}_t e^{\sigma\Phi_t\Upsilon_{t+1}}).
\end{aligned} \tag{64}
$$

The first-order conditions for $\{k_{t+1}, Q_{t+1}, \Upsilon_{t+1}\}$ are

$$
1 = \frac{\beta}{\lambda_t}\mathbb{E}_t n^p_{t+1}\mathcal{V}_{k_{t+1}} = \beta\mathbb{E}_t n^p_{t+1}\frac{\lambda_{t+1}}{\lambda_t}(1 - \delta + \Upsilon_{k_{t+1}}), \tag{65}
$$

$$
\varphi_t = \beta\mathbb{E}_t n^p_{t+1}\mathcal{V}_{Q_{t+1}} = \beta\mathbb{E}_t n^p_{t+1}(\varphi_{t+1} + \lambda_{t+1}x_{t+1}D_{t+1}F_{t+1}) \tag{66}
$$

$$
\Phi_t = -\mathcal{V}_{\Upsilon_{t+1}} = \Phi_{t+1} = \frac{m^p_{t+1}}{n^p_{t+1}}\Phi_t. \tag{67}
$$

Those for $\{c, H, E\}$ remain the same as before.

A special case:

1. The consumer has rational expectations ($m^p = 1$). As noted earlier, this case has been widely treated in papers on robust climate policy.

7.3. Pessimistic Planner, Pessimistic Consumer

This section treats a variation on the preceding belief regime by replacing its skeptical consumers with the pessimistic consumers from Section 3.4. Doing so will require adding two more implementability constraints: (1) the law of motion for the households' worst-case beliefs M^c_t in (43), because the authority needs to keep track of its evolution, and (2) the consumer's risk-sensitive utility recursion (44), because increments to the worst-case likelihood ratio M_t^c are determined by that household's utility $\mathcal{U}_t$.[29] For this policy maker, the Bellman recursion is,

$$\begin{aligned}\mathcal{V}_t(\Upsilon_t, k_t, Q_t) &= \max_{c_t, E_t, k_{t+1}, Q_{t+1}\Upsilon_{t+1}, M_t^c, \mathcal{U}_t} \min_{n_{t+1},} u(c_t) - \frac{\beta}{\sigma}\mathbb{E}_t(n_{t+1}\log n_{t+1}) \\ &+ \beta\mathbb{E}_t n_{t+1}\mathcal{V}_{t+1}(\Upsilon_{t+1}, k_{t+1}, Q_{t+1}) \\ &+ \Phi_t[\Omega_t + \beta\mathbb{E}_t\frac{M_{t+1}^c}{M_t^c}\Upsilon_{t+1} - \Upsilon_t] + \mathcal{M}_t \\ &+ \beta\mathbb{E}_t\mu_{t+1}\left[\frac{e^{\sigma^c\mathcal{U}_{t+1}}}{\mathbb{E}_t e^{\sigma^c\mathcal{U}_{t+1}}}M_t^c - M_{t+1}^c\right] \\ &+ \varepsilon_t\left[u(c_t, H_t) + \frac{\beta}{\sigma^c}\log\mathbb{E}_t e^{\sigma^c\mathcal{U}_{t+1}} - \mathcal{U}_t\right] + \mathcal{M}_t,\end{aligned}$$

where μ_{t+1} and ε_t are the Lagrangian shadow prices for the law of motion for M_{t+1}^c and the consumer's worst-case utility $\mathcal{U}_t$, respectively.

Outer Maximization with Implied Risk-Sensitive Recursion

With m^c computed by the consumer in Section 3.4,

$$m_{t+1}^c = \frac{M_{t+1}^c}{M_t^c} = \frac{e^{\sigma_n\mathcal{U}_{t+1}}}{\mathbb{E}_t e^{\sigma_n\mathcal{U}_{t+1}}},$$

and n^p the worst-case multiplier chosen by this planner,

$$n_{t+1}^p = \frac{e^{\sigma_n\mathcal{V}_{t+1}}}{\mathbb{E}_t e^{\sigma_n\mathcal{V}_{t+1}}},$$

the risk-sensitive recursion to be solved under dual ambiguities is,

$$\begin{aligned}\mathcal{V}(\Upsilon_t, k_t, Q_t) &= \max_{c_t, E_t, M_t^c, \mathcal{U}_t k_{t+1}, Q_{t+1}, \Upsilon_{t+1}} u(c_t) + \Phi_t[\Omega_t + \beta\mathbb{E}_t\frac{M_{t+1}^c}{M_t^c}\Upsilon_{t+1} - \Upsilon_t] \\ &+ \frac{\beta}{\sigma}\log\mathbb{E}_t e^{\sigma\mathcal{V}_{t+1}} + \beta\mathbb{E}_t\mu_{t+1}\left[\frac{e^{\sigma^c\mathcal{U}_{t+1}}}{\mathbb{E}_t e^{\sigma^c\mathcal{U}_{t+1}}}M_t^c - M_{t+1}^c\right] \\ &+ \varepsilon_t\left[u(c_t, H_t) + \frac{\beta}{\sigma^c}\log\mathbb{E}_t e^{\sigma^c\mathcal{U}_{t+1}} - \mathcal{U}_t\right] + \mathcal{M}_t.\end{aligned}$$

The first-order condition for c changes a little from before and becomes

$$c_t: \quad \lambda_t = [1 + \varepsilon_t]u_c + \Omega_{c_t}\Phi_t, \tag{68}$$

while the condition for E remains (51). The first-order conditions for $\{k_{t+1}, Q_{t+1}, \Upsilon_{t+1}\}$ are

$$1 = \beta\mathbb{E}_t n_{t+1}^p\frac{\lambda_{t+1}}{\lambda_t}(1 - \delta + \Upsilon_{k_{t+1}}), \tag{69}$$

$$\varphi_t = \beta\mathbb{E}_t n_{t+1}^p\mathcal{V}_{Q_{t+1}} = \beta\mathbb{E}_t n_{t+1}^p(\varphi_{t+1} + \lambda_{t+1}x_{t+1}D_{t+1}F_{t+1}), \tag{70}$$

$$\Phi_{t+1} = \frac{m_{t+1}^c}{n_{t+1}^p}\Phi_t, \tag{71}$$

$$0 \leq \Phi_t[\Omega_t + \mathbb{E}_t m_{t+1}^c\Upsilon_{t+1} - \Upsilon_t]. \tag{72}$$

From (71), Φ_t is the submartingale:

$$\mathbb{E}_t\Phi_{t+1} \geq \frac{\mathbb{E}_t m_{t+1}^c}{\mathbb{E}_t n_{t+1}^p}\Phi_t = \Phi_t, \tag{73}$$

unless the wealth constraint is nonbinding, when, by the Kuhn-Tucker condition (72), $\Phi_t = 0$. Finally, given endogenous belief distortions M^c in the private sector, the first-order conditions with respect to M_t^c and $\mathcal{U}_t$ are

$$M_t^c: \quad \mu_t - \frac{\Phi_{t-1}}{M_{t-1}^c}\Upsilon_t = \beta \mathbb{E}_t m_{t+1}^c \left(\mu_{t+1} - \frac{\Phi_t}{M_t^c}\Upsilon_{t+1} \right), \tag{74}$$

$$\mathcal{U}_t: \quad \varepsilon_t = \sigma^c m_t^c M_{t-1}^c (\mu_t - \mathbb{E}_{t-1} m_t^c \mu_t) + m_t^c \varepsilon_{t-1}. \tag{75}$$

The forward solution of (74) implies that the shadow value of increasing m_t^c is proportional to the value of debt (to the consumer) in units of the marginal utility of consumption,

$$\mu_t = \frac{\Phi_{t-1}}{M_{t-1}^c}\Upsilon_t = \frac{\Phi_{t-1}}{M_{t-1}^c} u_{c,t} b_t, \tag{76}$$

obtained by using (42). Substituting this in (75) yields

$$\varepsilon_t = \sigma^c \Phi_{t-1} m_t^c u_{ct} (b_t - \mathbb{E}_{t-1} m_t^c b_t) + m_t^c \varepsilon_{t-1}, \tag{77}$$

where the term in parentheses is the innovation in government debt, with positive surprises producing a negative shock to the pessimistic likelihood ratio.

8. The Equilibrium Price of Capital

The consumption Euler condition (68) implies the discount factor,

$$\begin{aligned} \varrho_{t+j,t}^* &\equiv \beta \tfrac{\lambda_{t+j}}{\lambda_t} \\ &= \beta \tfrac{(1+\varepsilon_{t+j}) u_{c_{t+j}} + \Omega_{c_{t+j}} \Phi_{t+j}}{(1+\varepsilon_t) u_{c_t} + \Omega_{c_t} \Phi_t} \\ &= \beta \tfrac{u_{c_{t+j}}}{u_{c_t}} \tfrac{1+\varepsilon_{t+j} + [1-\gamma+\gamma\vartheta_{t+1}]\Phi_{t+j}}{1+\varepsilon_t + [1-\gamma+\gamma\vartheta_t]\Phi_t} \\ &\equiv \Psi(\varepsilon_{t+j}, n_{t+j}^*) \varrho_{t+j,t}, \end{aligned} \tag{78}$$

where

$$\Psi(\varepsilon_{t+j}, n_{t+j}^*) = \frac{1 + \varepsilon_{t+j} + [1 - \gamma + \gamma\vartheta_{t+j}] n_{t+j}^* \Phi_{t+j-1}}{1 + \varepsilon_t + [1 - \gamma + \gamma\vartheta_t]\Phi_t}, \tag{79}$$

$\vartheta_t = \frac{H_t + g_t}{c_t}$ is the inverse of the average propensity to consume, namely the ratio of wage income plus the lump-sum rebate to consumption, and n_{t+j}^*, associated with the planner's implementability constraint on household wealth, varies according to policy regime as shown in Table 1. The shadow prices μ and ε are zero, except when the private sector has pessimistic beliefs $m = m^c$.

For the Ramsey plans derived previously, the conditions for capital k imply the distorted discount factor,

$$\varrho_{t+j,t}^{**} \equiv n_{t+j}^{**} \beta \frac{\lambda_{t+j}}{\lambda_t} = n_{t+j}^{**} \Psi(\varepsilon_{t+j}, n_{t+j}^*) \varrho_{t+j,t}, \tag{80}$$

where, like n_{t+j}^*, n_{t+j}^{**} varies by belief regime as shown in Table 1.[30] Note that from (78) and (80), the discount factor $\varrho_{t+j,t}^{**}$ is an n^{**}-distorted version of the previous consumption discount factor $\varrho_{t+j,t}^*$. The corresponding $t + j$ distorted equilibrium price of capital is

$$\hat{p}_{t+j,t}^{**} = n_{t+j}^{**} \hat{p}_{t+j,t}^*. \tag{81}$$

Table 1. Equilibrium Arrow–Debreu prices in alternative belief regimes.

	m	n^*	n^{**}	$\hat{p}$	$\hat{p}^*$	$\hat{p}^{**}$
Skeptical consumers						
Political planner	m^s	1	n^{PO}	$m^s p$	$\Psi(0,1)p$	$n^{PO}\Psi(0,1)p$
Paternalistic planner	m^s	n^{PA}	1	$m^s p$	$\Psi(0,n^{PA})p$	$\Psi(0,n^{PA})p$
Pessimistic planner	m^s	$\frac{m^p}{n^p}$	n^p	$m^s p$	$\Psi(0,\frac{m^p}{n^p})p$	$n^p\Psi(0,\frac{m^p}{n^p})p$
Pessimistic consumers						
Pessimistic planner	m^c	$\frac{m^c}{n^p}$	n^p	$m^c p$	$\Psi(\varepsilon,\frac{m^c}{n^p})p$	$n^p\Psi(\varepsilon,\frac{m^c}{n^p})p$

9. A Comment about Methodology

In the extant literature, the typical approach to finding analytically tractable solutions to the kind of maximin dynamic programming problems posed in Section 7 is to form their Isaacs-Bellman-Flemming equations that involve guessing and verifying functional forms as well as specifying detailed assumptions regarding preferences and probability distributions. See for example Hennlock (2009) and Li et al. (2016). As a bit of an exception, Anderson et al. (2013) do provide some analytic insights based on their problem's first-order conditions, especially regarding the importance of the role of deep uncertainty in measuring the TCRE parameter defined in Section 4. However, they leave more detailed conclusions to an evaluation of numerical solutions of their stochastic finite-horizon robust optimization problems.

An innovation of this paper is to derive specific and detailed formulas for the social cost of carbon and both carbon and capital taxes through evaluations of the Euler conditions derived from optimization. As will become apparent in the next three sections, each case will reveal important roles for certain second-order moments in the economy, specifically for a number of covariances between measures of the martingale belief distortions and a variety of economic variables, including net returns to fossil energy and capital that arise as a consequence of evaluating the expectations of products of random variables.

Given some key assumptions and a number of results stated as lemmas in Section 13.1, it becomes possible to sign these covariances, allowing us to determine with fair accuracy the likely signs of ambiguity premiums that must be added to the social cost of carbon, the carbon tax, and any capital subsidy in the various belief regimes analyzed in this paper.

10. The Social Cost of Carbon

The generic form of the first-order condition for carbon stores Q_t, based (alternatively) on (56), (61), (66) and (70), is

$$\varphi_t \quad = \quad \beta\mathbb{E}_t n^{**}_{t+1}(\varphi_{t+1} + \lambda_{t+1}x_{t+1}D_{t+1}F_{t+1}). \tag{82}$$

Defining the marginal-utility scaled shadow price $\varpi_t = \varphi_t/\lambda_t$, this becomes

$$\begin{aligned}\varpi_t \quad &= \quad \mathbb{E}_t n^{**}_{t+1}\beta\frac{\lambda_{t+1}}{\lambda_t}(\varpi_{t+1} + x_{t+1}D_{t+1}F_{t+1}) \\ &\equiv \quad \mathbb{E}_t\varrho^{**}_{t+1}[\varpi_{t+1} + x_{t+1}D_{t+1}F_{t+1}].\end{aligned} \tag{83}$$

Formula (83) gives a recursion for worst-case climate-caused damages from the point of view of a planner who may or may not be facing ambiguity, depending on n^{**}. The *social cost of carbon* is its forward solution,

$$\varpi_t = \mathbb{E}_t \lim_{T\to\infty}\sum_{j=0}^{T}(\prod_{i=0}^{t} p^{**}_{t+i,t})x_{t+j}D_{t+j}F_{t+j}. \tag{84}$$

Formula (84) reveals that damages are priced at $p^{**}_{t+j,t} = \varrho^{**}_{t+j,t}\pi_{t+j}$, i.e., the current t-period robust Arrow–Debreu price of capital in Formula (81), demonstrating an equivalence between capital and climate damages. In essence, accumulated carbon emissions constitute a negative asset that, in a competitive economy, is optimally priced like any other asset.[31]

It is convenient to define the instantaneous undistorted social cost of carbon,

$$\Lambda_{t+1} \equiv \varrho_{t+1}[\varpi_{t+1} + x_{t+1}D_{t+1}F_{t+1}], \tag{85}$$

and also

$$f(\Phi_t, \varepsilon_t) \equiv \frac{1}{1+\varepsilon_t + \Phi_t[1-\gamma+\gamma\vartheta_t]} \leq f(\Phi_t, 0) < f(0,0) = 1, \tag{86}$$

where $\vartheta_t = \frac{H_t + g_t}{c_t}\ \forall t$. Then use (78) and (79) to write the worst-case social cost of carbon (83) as a distortion of Λ_{t+1}

$$\begin{aligned} \varpi_t &= \mathbb{E}_t n^{**}_{t+1}\Psi(\varepsilon_{t+1}, n^*_{t+1})\Lambda_{t+1} \\ &= f(\Phi_t, \varepsilon_t)\mathbb{E}_t\left[n^{**}_{t+1}\Lambda_{t+1}\left(1+\varepsilon_{t+1} + [1-\gamma+\gamma\vartheta_{t+1}]n^*_{t+1}\Phi_t\right)\right]. \end{aligned} \tag{87}$$

Notably, the SCC is the sum of two parts.

The first part is

$$I: \quad f(\Phi_t, \varepsilon_t)(1+\varepsilon_{t+1})\mathbb{E}_t\left[n^{**}_{t+1}\Lambda_{t+1}\right],$$

which discounts future damages valued at the planner's n^{**}-distorted price of damages. Based on Table 1, n^{**} is n^{PO} for the political planner, 1 for the paternalistic planner, and n^p for the pessimistic planner facing either climate-skeptic or climate-pessimistic consumers.

The second part is

$$II: \quad f(\Phi_t, \varepsilon_t)\mathbb{E}_t\left[n^{**}_{t+1}n^*_{t+1}\Lambda_{t+1}(1-\gamma+\gamma\vartheta_{t+1})\Phi_t\right],$$

which augments damages in part I with the value of the planner's implementability constraint (if binding) at the n^*-distorted price of the net benefits of fossil energy. Based on Table 1, the combined ambiguity distortion $n^* \times n^{**}$ is n^{PO} for the political planner, n^{PA} for the paternalistic planner, m^p for the pessimistic planner facing climate-skeptical consumers, and m^c for the pessimistic planner facing climate-pessimistic consumers.

In the special instance when beliefs are rational and homogeneous, and the Ramsey planner's implementability constraint (39) is not binding ($\Phi_t = 0$), the social cost of carbon is the expected value of Λ_{t+1} (see for example Golosov et al. (2014)),

$$\varpi^{SP}_t = \mathbb{E}_t\varrho_{t+1}[\varpi_{t+1} + x_{t+1}D_{t+1}F_{t+1}] = \mathbb{E}_t\Lambda_{t+1} \equiv \Lambda^s_t. \tag{88}$$

11. The Carbon Tax

The conventional formula for the optimal carbon tax, as derived in Golosov et al. (2014), is

$$Social\ \ Planner's\ \ CarbonTax: \qquad \tau^{e-SP}_t = \varpi^{SP}_t = \Lambda^s_t. \tag{89}$$

This section shows that when beliefs are heterogeneous, the preceding formula for the carbon tax is inadequate.

The Ramsey planner's first-order condition (51) with respect to energy E_t is

$$Y_{E_t}\lambda_t - \varphi_t = v,$$

where Y is defined in (20). Advancing one period and taking discounted expectations, the preceding expression implies,

$$\mathbb{E}_t\beta\left(Y_{E_{t+1}}\lambda_{t+1} - \varphi_{t+1}\right) = v.$$

Together, these two expressions imply

$$\mathbb{E}_t \frac{\beta \frac{\lambda_{t+1}}{\lambda_t}(Y_{E_{t+1}} - \varphi_{t+1}/\lambda_{t+1})}{Y_{E_t} - \varphi_t/\lambda_t} \quad = \quad \mathbb{E}_t \frac{\varrho^*_{t+1}(Y_{E_{t+1}} - \varpi_{t+1})}{Y_{E_t} - \varpi_t} = 1. \tag{90}$$

In Formula (90), $\varrho^*_{t+1,t}$ is the discount factor a Ramsey planner applies to future expected excess returns to energy over the social cost of carbon—which I shall call the *Net Social Benefit of Carbon*. The corresponding $t+j$ equilibrium price of fossil energy is

$$\hat{p}^*_{t+j,t} = \varrho^*_{t+j,t}\pi_{t+j}(x_{t+j}|x^t). \tag{91}$$

Market equilibrium requires that returns to all assets and activities be equal:

$$\mathbb{E}_t\left[\frac{\varrho^*_{t+1}(Y_{E_{t+1}} - \varpi_{t+1})}{Y_{E_t} - \varpi_t}\right] \quad = \quad \mathbb{E}_t\varrho^{**}_{t+1}(1 - \delta + Y_{k_{t+1}}) = 1, \tag{92}$$

where the right-hand side comes from Euler Equation (55) (or (60), (65), (69)).

Formula (92) is a version of Hotelling's welfare-optimal rule and posits that, in equilibrium, the expected socially optimal pre-tax gross return on the extant stock of fossil fuel remaining in the ground, which society discounts with stochastic discount factor ϱ^*, is equal to the expected pre-tax gross return to capital in place, discounted using the government's SDF ϱ^{**}. By comparison, Hotelling's rule, when derived for laissez-faire in (31), posits an equivalence between the private after-tax return to fossil fuel in the ground and the private after-tax gross return to the extant stock of capital, discounted using the private sector's distorted SDF $\hat{\varrho}_{t+1}$. Since both ϱ^*_{t+1} and $\hat{\varrho}_{t+1}$ may reflect ambiguous beliefs, formula (92) is potentially a robust version of Hotelling's rule. Note that the expected sign of the numerator is positive (or negative) in period $t+1$ if its sign is positive (or negative) in period t.

Market equilibrium for the price of fossil energy $p^e = Y_E$ was previously derived as obeying the difference Equation (28),

$$Y_{E_t} - \tau^e_t \quad = \quad \mathbb{E}_t\hat{\varrho}_{t+1}(Y_{E_{t+1}} - \tau^e_{t+1}), \tag{93}$$

while, from (90), a socially optimal energy price must obey the rule

$$Y_{E_t} - \varpi_t \quad = \quad \mathbb{E}_t\varrho^*_{t+1}(Y_{E_{t+1}} - \varpi_{t+1}).$$

Since, in equilibrium, both expressions must hold, subtract the first from the second equation to eliminate Y_{E_t}:

$$\begin{aligned}
\tau^e_t - \varpi_t \quad &= \quad \mathbb{E}_t\varrho^*_{t+1}[Y_{E_{t+1}} - \varpi_{t+1}] - \mathbb{E}_t[\hat{\varrho}_{t+1}(Y_{E_{t+1}} - \tau^e_{t+1})] \\
&+ \quad E_t\hat{\varrho}_{t+1}\varpi_{t+1} - E_t\hat{\varrho}_{t+1}\varpi_{t+1} \\
&= \quad \mathbb{E}_t\hat{\varrho}_{t+1}[\tau^e_{t+1} - \varpi_{t+1}] + \mathbb{E}_t\varrho^*_{t+1}[Y_{E_{t+1}} - \varpi_{t+1}] - \mathbb{E}_t\hat{\varrho}_{t+1}(Y_{E_{t+1}} - \varpi_{t+1}).
\end{aligned}$$

Adding and subtracting $E_t\hat{\varrho}_{t+1}\varpi_{t+1}$ yields

$$\tau^e_t - \varpi_t \quad = \quad \mathbb{E}_t\hat{\varrho}_{t+1}[\tau^e_{t+1} - \varpi_{t+1}] + \mathbb{E}_tZ_{t+1}, \tag{94}$$

where

$$Z_{t+1} \quad = \quad (\varrho^*_{t+1} - \hat{\varrho}_{t+1})(Y_{E_{t+1}} - \varpi_{t+1}),$$

is the difference between the government's discounted *Net Social Benefit* of fossil fuel and the private sector's discounted *Net Social Benefit*, where, from (90), $Y_{E_{t+j}} \geq \varpi_{t+j}$. Intuitively, the carbon tax exceeds the social cost of carbon ϖ in every period if $\forall j > 0$, the social discount factor is higher than the private discount factor, $\varrho^*_{t+j} \geq \hat{\varrho}_{t+j}$. So if, in every period, the private sector is more myopic, i.e., less patient with respect to returns to fossil fuels than the planner, the government will add a premium to the carbon tax above the expected social cost of carbon ϖ. Exactly how and by how much, is determined as follows.

Lemma 1 (The carbon tax). *The carbon tax is the sum of two terms: (1) the expected social cost of carbon ϖ_t and (2) a premium χ_t:*

$$\hat{\tau}^e_t \quad = \quad \varpi_t + \chi_t, \tag{95}$$

where, utilizing (17),

$$\begin{aligned}
\chi_t \quad &\equiv \quad \mathbb{E}_t \sum_{j=0}^{\infty} \prod_{i=0}^{j} \hat{\varrho}_{t+i} Z_{t+j} = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j \left(\prod_{i=1}^{j} m_{t+i} \frac{u_c(x^{t+i})}{u_c(x^t)} \right) Z_{t+j} \\
&= \quad \mathbb{E}_t \sum_{j=0}^{\infty} \frac{M_{t+j}}{M_t} \varrho_{t+j,t} Z_{t+j} = \mathbb{E}_t \sum_{j=0}^{\infty} \hat{p}_{t+j,t} (\varrho^*_{t+j} - \hat{\varrho}_{t+j})(Y_{E_{t+j}} - \varpi_{t+j}) \qquad (96) \\
&= \quad \mathbb{E}_t \sum_{j=0}^{\infty} \hat{p}_{t+j,t} Z_{t+j}.
\end{aligned}$$

Proof. Solve the difference Equation (94). □

The first component of $\hat{\tau}^e$ is the Social Cost of Carbon previously derived that may or may not already contain an ambiguity-related premium. The second component adds a further ambiguity premium depending on belief regime. In essence, a government uses this formula to impose a premium on (or grant a discount toward) the carbon tax if cumulative *expected* differentially discounted *net private benefits* of fossil fuel are positive (or negative), where the sign and size of the premium (or discount) χ_t depends on the signs and sizes of all future Z_{t+j} priced at $\hat{p}_{t+j,t}$ that need to be determined. Note that in Formula (96), $\hat{p}_{t+j,t}(Y_{E_{t+j}} - \varpi_{t+j})$ is the (possibly belief distorted) market value of the excess of private returns of fossil fuels over their social cost in period $t+j$. The premium χ_t is the expected sum of all such terms, each multiplied by the difference between the government's and the private sector's discount factor. In general, the premium is positive if the public is myopic compared with the planner: $\varrho^*_{t+j} - \hat{\varrho}_{t+j} > 0$ in all periods. Intuitively, a skeptical consumer who disbelieves the seriousness of climate change, will tend to use more energy than warranted from society's point of view because it myopically undervalues future climate costs. By contrast, a pessimistic private sector would opt to use less. As will become apparent later, this calculus is modified, if the authority itself has ambiguity about private beliefs.

An asset-pricing interpretation of (95) is that the optimal carbon tax is the sum of two possibly robust asset prices: (1) expected cumulative fossil fuel damages per unit of carbon valued at prices p^{**}_{t+j} in (84), and (2) cumulative net private benefits per unit of carbon over the social cost of carbon, valued at $\hat{p}_{t+j}$ and weighted by the difference in discount factors $\varrho^*_{t+j} - \hat{\varrho}_{t+j}$. The interpretation of $\hat{\tau}^e_t$ as the potentially robust price of an underlying asset—the government-imposed cap on emissions—extends a result by Belfiori (2017), who also derived an equivalence between the optimal carbon tax and the optimal price of traded carbon permits in an economy with a cap and trade.

The case of a social planner under rational expectations, in which $\Phi = 0$ and beliefs are undistorted and homogeneous, is a suitable benchmark for comparison. However, it is exceptional in that in all other policy/belief regimes, the government may either impose a premium or give a concession. The premium (or concession) in Formula (96) is the expected value of the sum of products of random variables over all $t+j$, so the sign is not obvious from inspection, although it can be deciphered via decomposition into covariance components derived in Appendix D, which distinguishes between two cases for each of the possible policy regimes when either (1) the wealth constraint in (39) is binding ($\Phi_t > 0$), implying Ramsey planning, or (2) it is not ($\Phi_t = 0$), implying social planning.

To pursue this, define a new variable Ξ^e_{t+j} as the weighted proportional difference between the two stochastic discount factors ϱ^*_{t+j} and $\hat{\varrho}_{t+j}$:

$$\Xi^e_{t+j} \equiv \varsigma^e_{t+j} \frac{\varrho^*_{t+j} - \hat{\varrho}_{t+j}}{\hat{\varrho}_{t+j}} = \varsigma^e_{t+j}\left(\frac{\varrho^*_{t+j}}{\hat{\varrho}_{t+j}} - 1\right), \tag{97}$$

where

$$\varsigma^e_{t+j} \equiv \frac{\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j}),}{\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j}),} \Longrightarrow \mathbb{E}_t\varsigma^e_{t+j} = 1, \tag{98}$$

is the j-th period's normalized discounted excess return to fossil energy over the social cost of carbon. Utilizing (78), Ξ^e_{t+j} is

$$\Xi^e_{t+j} = \varsigma^e_{t+j}(\frac{\Psi(\varepsilon_{t+j}, n^*_{t+j})}{m_{t+j}} - 1), \tag{99}$$

where $\Psi(\varepsilon_{t+j}, n^*_{t+j})$ is defined in (79). Notice that Ξ^e depends on the government's belief multiplier n^* which varies according to policy regime, as shown in Table 1, and on the private sector's belief distortion $m_{t+1} = m^c_{t+1}$ when consumers are pessimistic and $m_{t+1} = m^s_{t+1}$ when they are skeptical. Importantly, whereas m^s is an exogenous martingale process, m^c is the pessimistic consumer's worst-case multiplier that depends on continuation utility as derived in (6). From the preceding, using (78) and (79), the expected value of Ξ^e_{t+j} is

$$\begin{aligned} \mathbb{E}_t\Xi^e_{t+j} &= \mathbb{E}_t\frac{\varrho^*_{t+j} - \hat{\varrho}_{t+j}}{\hat{\varrho}_{t+j}}\varsigma^e_{t+j} \\ &= \mathbb{E}_t\left[\left(\frac{1}{m_{t+j}}\frac{1 + \varepsilon_{t+j} + [1 - \gamma + \gamma\vartheta_{t+j}]n^*_{t+j}\Phi_{t+j-1}}{1 + \varepsilon_t + [1 - \gamma + \gamma\vartheta_t]\Phi_t} - 1\right)\varsigma^e_{t+j}\right]. \end{aligned} \tag{100}$$

Note that (100) gives the difference between two non-centered covariances with the normalized discounted excess return to fossil energy, one involving the government's and the other the private sector's stochastic discount factor.

Lemma 2 is key to determining the sign of the carbon tax premium. Based on (95),

Lemma 2.

$$sign\ \chi_t \ = \ sign[\hat{\tau}^e_t - \varpi_t] \ = sign\ \mathbb{E}_t\Xi^e_{t+j}\ \forall j. \tag{101}$$

Proof. First,

$$\mathbb{E}_t Z_{t+j} \gtreqless 0, \quad if\ \ \mathbb{E}_t\Xi^e_{t+j} \gtreqless \ 0,\ j = 1, \cdots \infty.$$

To show this, multiply and divide Z_{t+j} by $\hat{\varrho}_{t+j}\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j})$:

$$\begin{aligned} Z_{t+j} &= Z_{t+j}\frac{\hat{\varrho}_{t+j}}{\hat{\varrho}_{t+j}}\frac{\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j})}{\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j})} \\ &= \left[\frac{\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j})}{\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j})}\frac{\varrho^*_{t+j} - \hat{\varrho}_{t+j}}{\hat{\varrho}_{t+j}}\right]\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j}) \\ &= \left[\varsigma^e_{t+j}\frac{\varrho^*_{t+j} - \hat{\varrho}_{t+j}}{\hat{\varrho}_{t+j}}\right]\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j}) \\ &= \Xi^e_{t+j}\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j}), \qquad (102) \\ \Longrightarrow \mathbb{E}_t Z_{t+j} &= \mathbb{E}_t\Xi^e_{t+j}\mathbb{E}_t\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j}), \end{aligned}$$

where, from (90), $Y_{E_{t+j}} \geq \varpi_{t+j}$. It follows that $\mathbb{E}_t Z_{t+j}$ and $\mathbb{E}_t \Xi^e_{t+j}$ have the same sign. Finally, with ϖ_t defined in (84), the lemma follows from Formulas (96) and (102). □

In the absence of any known time-dependent anomalies, there is no reason to believe that the sign of $\mathbb{E}_t \Xi^e_{t+j}$ is different at different times, justifying

Assumption 1. *sign* $\mathbb{E}_t \sum_{j=0}^{\infty} \Xi^e_{t+j} =$ *sign* $\mathbb{E}_t \Xi^e_{t+j} \;\; \forall j$

Since in (96), the sign of χ_t depends on the sign of all its future expected elements, $\mathbb{E}_t \hat{p}_{t,t+j} Z_{t+j}$, we require

Lemma 3. *If* $\mathbb{E}_t \Xi^e_{t+j} \gtreqless 0$ *for any* $j \geq 0$*, then* $\chi_t \gtreqless 0$ *for any* $j \geq 0$*.*

Proof. First, in (96),

$$\mathbb{E}_t \hat{p}_{t+j,t} Z_{t+j} \gtreqless 0, \quad if \quad \forall j \geq 1, \quad \mathbb{E}_t \hat{p}_{t+j,t} \Xi^e_{t+j} \gtreqless \; 0.$$

The lemma follows by applying Assumption 1, and (96) and (102). □

A suitable formula for the carbon tax in terms of the Ξ^e_{t+j} is found by substituting (96) and (102) into (95):

$$\hat{\tau}^e_t = \varpi_t + \mathbb{E}_t \sum_{j=0}^{\infty} \hat{p}_{t+j,t} \Xi^e_{t+j} \hat{\varrho}_{t+j} (Y_{E_{t+j}} - \varpi_{t+j}). \tag{103}$$

The essence of the preceding is that if we know the sign of Ξ^e_{t+j} for any j, then we shall know the sign of any premium χ_t over the standard carbon tax. The implications of this for the different belief regimes are stated in several propositions in Section 13.

12. An *Ex Ante* Tax on Capital

While $p^e_t(x^t)$ and $\tau^e_t(x^t)$ are uniquely determined in (24) and (95), respectively, it is easy to demonstrate, following Chamley (1986), Zhu (1992), and Chari et al. (1994), that a state-by-state capital income tax is not uniquely determined because an implementable allocation $\{b, k, Q\}$ that uniquely determines household wealth $\mathcal{W}$ in (36) can be obtained by a multiplicity of capital tax and bond policies $\{\tau^k, b\}$ at prices $\{\hat{p}, r\}$.[32] However, across states of nature, the history-dependent value of tax payments is fully determined. So with this in mind, define the *effective* or *ex ante* tax rate on capital as the ratio of the prices of two "assets", one yielding a stream of tax revenues and the other yielding a stream of gross before-tax capital returns, conditional on history x^t, where the latter is defined by

$$\bar{\tau}^k_{t+1}(x^t) \quad \equiv \quad \frac{\mathbb{E}_t \hat{p}_{t+1,t}(x_{t+1}, x^t) \tau^k_{t+1}(x^{t+1})(Y_{k_{t+1}}(x^{t+1}) - \delta)}{\mathbb{E}_t \hat{p}_{t+1,t}(x_{t+1}, x^t)(Y_{k_{t+1}}(x^{t+1}) - \delta))}. \tag{104}$$

In this formula, the *ex ante* tax rate $\bar{\tau}^k_{t+1}(x^t)$ is the ratio of expected tax revenues to expected gross returns from capital, conditional on history x^t and valued at possibly distorted Arrow–Debreu prices $\hat{p}$.

With the use of (14), (104) is re-written as

$$\begin{aligned} \bar{\tau}^k_{t+1}(x^t) \quad &\equiv \quad \frac{\mathbb{E}_t \hat{p}_{t+1,t}(x_{t+1}, x^t)[1 - \delta + Y_{k_{t+1}}(x^{t+1})] - 1}{\mathbb{E}_t \hat{p}_{t+1,t}(x_{t+1}, x^t)(Y_{k_{t+1}}(x^{t+1}) - \delta)}, \\ &= \quad \frac{\mathbb{E}_t \hat{\varrho}_{t+1}[1 - \delta + Y_{k_{t+1}}] - 1}{\mathbb{E}_t \hat{\varrho}_{t+1}(Y_{k_{t+1}} - \delta)}. \end{aligned} \tag{105}$$

Based on consumption-Euler Equations (50) or (68),

$$\begin{aligned} 1 &= \mathbb{E}_t n^{**}_{t+1}\beta\frac{\lambda_{t+1}}{\lambda_t}(1+Y_{k_{t+1}}-\delta) \\ &= \mathbb{E}_t n^{**}_{t+1}\beta\frac{[1+\varepsilon_{t+1}]u_{c_{t+1}}+\Omega_{c_{t+1}}\Phi_{t+1}}{[1+\varepsilon_t]u_{c_t}+\Omega_{c_t}\Phi_t}(1+Y_{k_{t+1}}-\delta) \\ &= \mathbb{E}_t n^{**}_{t+1}\beta\frac{1+\varepsilon_{t+1}+[1-\gamma+\gamma\vartheta_{t+1}]n^{*}_{t+1}\Phi_{t+j-1}}{1+\varepsilon_t+[1-\gamma+\gamma\vartheta_t]\Phi_t}(1+Y_{k_{t+1}}-\delta), \end{aligned} \tag{106}$$

the government discounts *gross* returns to capital in period $t+1$ using ϱ^{**}, derived in (79), which is also used to discount future damages resulting from fossil use in (84), as shown earlier, where the distorting factor n^{**}_{t+1}varies according to belief regime as detailed in Table 1.

Substituting (107) into (105), an alternative expression for $\bar{\tau}^k_t$ is

$$\bar{\tau}^k_{t+1} = \frac{\mathbb{E}_t[(\hat{\varrho}_{t+1}-\varrho^{**}_{t+1})(1-\delta+Y_{k_{t+1}}]}{\mathbb{E}_t[\hat{\varrho}_{t+1}(Y_{k_{t+1}}-\delta)]}. \tag{107}$$

Noteworthy in this expression is that it echoes the expression in (96) determining the sign of the carbon tax premium, likewise depending on the difference between the private sector's discount factor and that of the government, $\hat{\varrho}_{t+1}-\varrho^{**}_{t+1}$. Simply put, the numerator in (107) is the difference between two non-centered covariances between the stochastic pre-tax return $1-\delta+Y_{k_{t+1}}$ and the private sector's and the government's SDF, respectively. In general terms, the *ex ante* tax on capital is negative if the private sector is more myopic, i.e., less patient with respect to returns to capital than the government, specifically, if the covariance of private returns with the private stochastic discount factor is less than the corresponding covariance with the planner's stochastic discount factor. Conversely, if private returns are better correlated with the government's stochastic discount factor, or are equal to it, then the average tax is positive or zero. Earlier literature by Chamley (1986), Judd (1985), and Atkeson et al. (1999), concludes that in a deterministic economy with time-additive preferences, the optimal *ex ante* tax on capital is zero, except possibly at time 0. Subsequently, Chari et al. (1994) and Zhu (1992) found that in a stochastic economy with homogeneous beliefs, the tax is positive (or negative) if a weighted average of the change in period elasticities of the utility function is positive (or negative). Below, I shall show how a divergence in beliefs between the private sector and government produces a gap between the government's and the private sector's discount factors that contribute additional motives to subsidize or penalize capital, depending on the source of ambiguity.

Now mimic expression (100) and define the weighted proportional difference between the private and the government's stochastic discount factors,

$$\Xi^k_{t+1} \equiv \varsigma^k_{t+1}\frac{\hat{\varrho}_{t+1}-\varrho^{**}_{t+1}}{\hat{\varrho}_{t+1}} = \varsigma^k_{t+1}(1-\frac{\varrho^{**}_{t+1}}{\hat{\varrho}_{t+1}}) = \varsigma^k_{t+1}(1-\frac{n^{**}_{t+1}}{m_{t+1}}\Psi(\varepsilon_{t+1},n^{*}_{t+1})), \tag{108}$$

where

$$\varsigma^k_{t+1} \equiv \frac{\hat{\varrho}_{t+1}[1-\delta+Y_{k_{t+1}}]}{\mathbb{E}_t[\hat{\varrho}_{t+1}(1-\delta+Y_{k_{t+1}})]}, \tag{109}$$

is the normalized discounted gross return to capital, such that $\mathbb{E}_t\varsigma^k_{t+1}=1$, and where $\Psi(\varepsilon_{t+1},n^{*}_{t+1})$ is defined in (79). The next lemma proves that the sign of the capital tax equals the sign of $E_t\Xi^k_{t+1}$:

Lemma 4.

$$\bar{\tau}^k_{t+1} \lesseqgtr 0, \;\; if \;\; \mathbb{E}_t\Xi^k_{t+1} \lesseqgtr 0. \tag{110}$$

Proof. Use (109) in (108) and write,

$$\begin{aligned}\mathbb{E}_t\Xi^k_{t+1} &\equiv \mathbb{E}_t[\varsigma^k_{t+1}\frac{\hat{\varrho}_{t+1}-\varrho^{**}_{t+1}}{\hat{\varrho}_{t+1}}] \\ &= \mathbb{E}_t[\frac{\hat{\varrho}_{t+1}[1-\delta+Y_{k_{t+1}}]}{E_t[\hat{\varrho}_{t+1}(1-\delta+Y_{k_{t+1}})]}\frac{\hat{\varrho}_{t+1}-\varrho^{**}_{t+1}}{\hat{\varrho}_{t+1}}] \\ &= \frac{\mathbb{E}_t[\hat{\varrho}_{t+1}(Y_{k_{t+1}}-\delta)]}{\mathbb{E}_t[\hat{\varrho}_{t+1}(1-\delta+Y_{k_{t+1}})]}\bar{\tau}^k_{t+1},\end{aligned} \tag{111}$$

where the last term is obtained by multiplying and dividing by $\mathbb{E}_t[\hat{\varrho}_{t+1}(Y_{k_{t+1}}-\delta)]$ and using (107). The proof follows by observing that the term multiplying $\bar{\tau}^k_{t+1}$ is positive. □

For later use, re-write (111) as

$$\begin{aligned}\mathbb{E}_t\Xi^k_{t+1} &= \mathbb{E}_t\varsigma^k_{t+1}\left(1-\frac{n^{**}_{t+1}}{m_{t+1}}\frac{1+\varepsilon_{t+1}+[1-\gamma+\gamma\vartheta_{t+1}]n^*_{t+1}\Phi_t}{1+\varepsilon_t+[1-\gamma+\gamma\vartheta_t]\Phi_t}\right), \\ &= \mathbb{E}_t\varsigma^k_{t+1}\left(1-f(\Phi_t,\varepsilon_t)\frac{n^{**}_{t+1}}{m_{t+1}}(1+\varepsilon_{t+1}+[1-\gamma+\gamma\vartheta_{t+1}]n^*_{t+1}\Phi_t\right).\end{aligned} \tag{112}$$

13. Policy Implications of Belief Heterogeneity and Ambiguity

13.1. Preliminary Results

Having derived formulas for the expected social cost of carbon, and carbon as well as capital taxation, we are now prepared to consider the policy implications for the various belief regimes, with particular focus on any premia or surcharges that may arise as a consequence of belief distortions and ambiguities. The results are summarized in six propositions. As will become apparent, the sign and size of surcharges on the carbon tax and those of the *ex ante* tax on capital will depend on the signs of several key second-order moments, namely covariances involving measures related to the state of the economy, ς^k and ς^e, the average propensity to consume out of disposable income ϑ, and measures of the intensity of robustness concerns reflected in the endogenous worst-case belief distortions m^c, m^p, n^p, n^{POL}, n^{PAT} that arise from evaluating Formulas (87), (100), and (112) containing the expectations of products and ratios of random variables. The resulting expressions involve sums of expectations of their individual components and their covariances, known to be non-trivial because, as shown later, the assumption of endogeneity means that all involved variables are functions of the fundamental climate shock process x_t and are therefore related to each other. The manner of deconstructing optimal tax policy applied in this paper is novel in the literature, which has mostly had to rely on simulations to quantify policy effects, and should be considered a unique contribution of this paper.

The signs of the covariances needed to evaluate expectations and to prove later propositions are not always clear from intuition and need to be formally derived and stated in the form of several lemmas.

Assumption 2. $\frac{dc_{t+1}}{dx_{t+1}} \leq 0$.

It is generally assumed that consumption declines with global warming (see Frankhouser and Tol 2005; Weitzman 2009).

Lemma 5.

$$\frac{d\hat{\varrho}_{t+j}}{dx_{t+j}} > 0.$$

Proof.

$$\begin{aligned}
\frac{d\hat{Q}_{t+j}}{dx_{t+j}} &= \beta M_{t+j} \frac{d\frac{u_{c_{t+j}}}{u_{c_t}}}{dx_{t+j}} + \beta \frac{u_{c_{t+j}}}{u_{c_t}} \frac{dM_{t+j}}{dx_{t+j}} \\
&= -\beta M_{t+j} \left(\frac{c_t}{c_{t+j}} \right)^{\gamma} \left(\gamma c_{t+j}^{-1} \frac{dc_{t+j}}{dx_{t+j}} \right) \\
&+ \beta \left(\frac{c_t}{c_{t+j}} \right)^{\gamma} \frac{dM_{t+j}}{dx_{t+j}} > 0,
\end{aligned}$$

where first terms in each line are positive by Assumption 2, and the second lines are also positive since $\frac{dM_{t+j}}{dx_{t+j}} > 0$ if $M = M^c$, and zero otherwise. □

Assumption 3. $\frac{dk_{t+1}}{dx_{t+1}} < 0$.

Li et al. (2016) assume that capital utilization declines upon a climate cost shock. Frankhouser and Tol (2005, p. 5) observe that "the overall effect of climate change on the accumulation of capital is in principle ambiguous", but that "it seems safe to speculate that the capital accumulation effect will probably be negative". More recent evidence that a significant portion (50%) of total GDP losses can be attributed to disincentives to invest capital is less ambiguous (see Willner et al. 2021).

Assumption 4. $\frac{dH_{t+1}}{dx_{t+1}} \leq 0$.

Labor input likely declines as a result of a climate cost shock since, as documented in Dasgupta et al. (2021); Kjellstrom (2014); Kjellstrom et al. (2009); Somanathan et al. (2018), productive labor is lost as a consequence of warming, justifying

Assumption 5. $\frac{dE_{t+1}}{dx_{t+1}} \simeq 0$.

Wilbanks et al. (2008) find that climate warming reduces energy use and production, as verified by the BEA (BEA 2019), although, as an Environmental Protection Agency web page, deleted by the Trump Administration in 2017 but saved and still available on (EPA 2017), reports, climate warming may lead to partially offsetting heating and cooling demands.

Lemma 6.

$$\begin{aligned}
\frac{dY_{z_{i,t}}}{dx_t} &= -(Q_0 - Q_t) D_t F_{z_{i,t}} + \sum_{j=1}^{n} \frac{\partial Y_{z_{i,t}}}{\partial z_{j,t}} \frac{dz_{j,t}}{dx_t}, \\
&\approx -(Q_0 - Q_t) D_t F_{z_{i,t}} < 0,
\end{aligned} \tag{113}$$

for all factors of production, k, E, and H.

Proof. The signs of $\frac{dz_{j,t}}{dx_t}$ are given in Assumptions 3–5, and $\frac{\partial Y_{z_{i,t}}}{\partial z_{j,t}} > 0$. Recent literature (Dasgupta et al. 2021; Njuki et al. 2020; Zhang et al. 2017) suggests that the indicated sum is likely negative but small, leaving as main driver of climate cost shocks their effects on total factor productivity, here basically represented by the damage function. □

Assumption 6. $\frac{db_{t+1}}{dx_{t+1}} > 0$.

The effect of a climate shock on debt b_t can plausibly be said to be positive. In their study of Columbia and Peru, Maldonado and Gallagher (2022) provide some evidence that climate shocks significantly affect public debt trajectories towards significantly higher levels and, in some cases, raise probabilities of increasing debt during climate stress. Nor do developed economies in Europe seem to be immune from this effect. For example, Zenios (2022) combined projections from the IMF World Economic Outlook with simulations of versions of an IAM model obtained from Emmerling et al. (2016) and Gazzotti et al. (2021) to show that climate shocks will raise the sovereign debt-to-GDP ratio in Italy and Cyprus over time.

Lemma 7.

$$\frac{dm^c_{t+1}}{dx_{t+1}} \geq 0.$$

Proof. From (6),

$$\frac{dm^c_{t+1}}{dx_{t+1}} = \frac{\partial m^c_{t+1}}{\partial \mathcal{U}_{t+1}} \frac{d\mathcal{U}_{t+1}}{dx_{t+1}} \quad = \quad \sigma^c m^c_{t+1}(1 - m^c_{t+1}) \frac{d\mathcal{U}_{t+1}}{dx_{t+1}}.$$

The term multiplying $\frac{d\mathcal{U}_{t+1}}{dx_{t+1}}$ is negative, indicating that decreases in $\mathcal{U}_{t+1}$ raise m^p_{t+1} toward 1. The envelope conditions (10)–(12) imply

$$\begin{aligned}
\frac{d\mathcal{U}_{t+1}}{dx_{t+1}} &= \mathcal{U}_{k_{t+1}} \frac{dk_{t+1}}{dx_{t+1}} + \mathcal{U}_{b_{t+1}} \frac{db_{t+1}}{dx_{t+1}} + \mathcal{U}_{Q_{t+1}} \frac{dQ_{t+1}}{dx_{t+1}} \\
&= u_{c_{t+1}}[R^k_{t+1} \frac{dk_{t+1}}{dx_{t+1}} + \frac{db_{t+1}}{dx_{t+1}}] \leq 0,
\end{aligned}$$

if the negative effect on capital outweighs the presumably positive effect on debt. Otherwise, the result follows from $\sigma^c < 0$ and $m^c \leq 1$.[33] □

Lemma 8.

$$\frac{dn^{PO}_{t+1}}{dx_{t+1}} \quad > \quad 0.$$

Proof. Evaluate

$$\begin{aligned}
\frac{dn^{PO}_{t+1}}{dx_{t+1}} &= \frac{\partial n^{PO}_{t+1}}{\partial \mathcal{V}_{t+1}} \frac{d\mathcal{V}_{t+1}}{dx_{t+1}} + \frac{\partial n^{PO}_{t+1}}{\partial \mathcal{V}_{t+1}} \frac{d\Upsilon_{t+1}}{dx_{t+1}} \\
&= \sigma n^{PO}_{t+1}(1 - n^{PO}_{t+1}) \left(\mathcal{V}_{k_{t+1}} \frac{dk_{t+1}}{dx_{t+1}} + \mathcal{V}_{\Upsilon_{t+1}} \frac{d\Upsilon_{t+1}}{dx_{t+1}} + \bar{\Phi} \frac{d\Upsilon_{t+1}}{dx_{t+1}} \right) \\
&= \sigma n^{PO}_{t+1}(1 - n^{PO}_{t+1}) \left((1 - \delta + \Upsilon_{k_{t+1}})(u_{c_{t+1}} + \Omega_{c_{t+1}} \bar{\Phi}) \frac{dk_{t+1}}{dx_{t+1}} \right) \\
&+ \sigma n^{PO}_{t+1}(1 - n^{PO}_{t+1}) \left(\Phi_t \frac{d\Upsilon_{t+1}}{dx_{t+1}} - \Phi_{t+1} \frac{d\Upsilon_{t+1}}{dx_{t+1}} \right) \\
&= \sigma n^{PO}_{t+1}(1 - n^{PO}_{t+1}) \left((1 - \delta + \Upsilon_{k_{t+1}})(u_{c_{t+1}} + \Omega_{c_{t+1}} \bar{\Phi}) \frac{dk_{t+1}}{dx_{t+1}} \right) > 0,
\end{aligned}$$

which follows from $\sigma < 0$, $n^{PO} \leq 1$, envelope conditions (52) and (54), Euler condition (50) with $\xi = 0$, from Assumption 3.[34], and from the previous result that $\Omega_{c_{t+1}} > 0$.[35] □

The preceding result accords with the properties $0 < n_{t+1}^{PO} \leq 1$, $\lim_{\sigma\uparrow 0} n_{t+1}^{PO} \to 1$, $\lim_{\sigma\downarrow -\infty} n_{t+1}^{PO} \to 0$, and the partial derivative

$$\frac{\partial n_{t+1}^{PO}}{\partial \sigma} = (\mathcal{V}_{t+1} + \Phi_t \mathrm{Y}_{t+1})(1 - n_{t+1}^{PO})n_{t+1}^{PO} \geq 0. \tag{114}$$

As one might expect intuitively, increased ambiguity aversion ($\sigma \downarrow$) reduces n^{PO} and drives it towards zero, a result that echoes those in Gilboa and Schmeidler (1989) and Millner et al. (2012).

Lemma 9.

$$\frac{dn_{t+1}^{PA}}{dx_{t+1}} > 0.$$

Proof. From (39) and results (41),

$$\begin{aligned} \frac{dn_{t+1}^{PA}}{dx_{t+1}} &= \frac{\partial n_{t+1}^{PA}}{\partial \mathrm{Y}_{t+1}} \frac{d\mathrm{Y}_{t+1}}{dx_{t+1}} = \sigma \Phi_t n_{t+1}^{PA}(1 - n_{t+1}^{PA}) \frac{d\mathrm{Y}_{t+1}}{dx_{t+1}} \\ &= \sigma \Phi_t n_{t+1}^{PA}(1 - n_{t+1}^{PA}) \left(\Omega_{c_{t+1}} \frac{dc_{t+1}}{dx_{t+1}} \right) \geq 0. \end{aligned}$$

Given $\sigma < 0$ and Assumption 2, this implies a positive correlation between n^{PA} and x. □

By the preceding arguments,

Lemma 10.

$$\frac{\partial n_{t+1}^{p}}{\partial x_{t+1}} \geq 0.$$

Proof. The proof follows argument similar to the proof of Lemma 10. □

The preceding discussion leads up to the following two lemmas:

Lemma 11. *Given definition (109),*

$$\frac{d\varsigma_{t+1}^{k}}{dx_{t+1}} \gtreqless 0.$$

Proof. Based on Lemma 5 and the result in (113),

$$\begin{aligned} \frac{d\varsigma_{t+1}^{k}}{dx_{t+1}} &= \frac{d\hat{\varrho}_{t+1}(1 - \delta + \mathrm{Y}_{k_{t+1}})}{dx_{t+1}} \\ &= (1 - \delta + \mathrm{Y}_{k_{t+1}}) \frac{d\hat{\varrho}_{t+1}}{dx_{t+1}} + \hat{\varrho}_{t+1} \frac{d\mathrm{Y}_{k_{t+1}}}{dx_{t+1}} > 0. \end{aligned}$$

□

Lemma 12.

$$\frac{d\varsigma_{t+j}^{e}}{dx_{t+j}} > 0 \quad \forall j > 0.$$

Proof. Use Lemma 5 and the result in (113) and note that, from Formula (98), the sign of $\frac{d\varsigma^e_{t+j}}{dx_{t+j}}$ is the same as the sign of

$$\frac{d\hat{\varrho}_{t+j}(Y_{E_{t+j}} - \varpi_{t+j})}{dx_{t+j}} = (Y_{E_{t+j}} - \varpi_{t+j})\frac{d\hat{\varrho}_{t+j}}{dx_{t+j}} + \hat{\varrho}_{t+j}\frac{dY_{E_{t+j}}}{dx_{t+j}} > 0.$$

□

Lemma 13. $\frac{d\Lambda_{t+j}}{dx_{t+j}} \geq 0$.

Proof. This follows from its definition in (85). □

Assumption 7.

$$\frac{d\vartheta_{t+j}}{dx_{t+j}} = \frac{d\frac{H_t+g_t}{c_t}}{dx_{t+j}} \approx 0.$$

Since climate lowers both income and consumption, the net effect on ϑ is likely negligible, justifying an assumption that ϑ is not strongly negatively correlated with Λ, ς^k, ς^e, m^c, n^{PO}, m^c, n^{PA}, and n^p, all of which are positively correlated with x.

Under the assumption that climate shocks are the sole stochastic process x driving the economy, the next proposition reflects the fact that variables that are functions of x must be correlated with each other. The signs of the covariances listed therein follow from the preceding lemmas.

Lemma 14. *Signs of key covariances*

(i) The signs of $cov(\Lambda, \vartheta)$, $cov(\varsigma^e, \vartheta)$, $cov(\varsigma^k, \vartheta)$, $cov(n^{PO}, \Lambda)$, $cov(n^{PO}, \Lambda\vartheta)$, $cov(n^{PA}, \Lambda)$, $cov(n^{PA}, \Lambda\vartheta)$, $cov(n^{PA}, \vartheta\varsigma^e)$, $cov(n^{PA}, \varsigma^k)$, $cov(n^{PA}, \vartheta\varsigma^k)$, $cov(m^{PA}, \Lambda)$, $cov(m^{PA}, \vartheta\Lambda)$, $cov(n^p, \Lambda)$, $cov(n^p, \varsigma^k)$, $cov(n^p, m^c)$, $cov(m^p, \varsigma^k)$, $cov(m^p, \Lambda)$, $cov(m^p, \vartheta\Lambda)$, $cov(m^p\varsigma^k, \vartheta)$, $cov(m^c m^p, \Lambda)$, $cov(m^c, \Lambda)$, *and* $cov(m^c, \vartheta\Lambda)$ *are positive.*

(ii) The signs of $cov(\varsigma^e, \frac{1}{m^c})$, $cov(n^p, \frac{1}{m^c})$, $cov(\varsigma^e, \frac{1}{n^p})$, *and* $cov(\vartheta\varsigma^e, \frac{1}{n^p})$ *are negative.*

(iii) The signs of $cov(\varsigma^k, \frac{n^p}{m^c})$, $cov(\varsigma^e, \frac{n^p}{m^c})$, $cov(\varsigma^e, \frac{m^p}{n^p})$, *and* $cov(\varsigma^e\vartheta, \frac{m^p}{n^p})$ *are indeterminate.*

The following propositions distinguish two situations: (i) when the implementability (the marginal-utility-of-consumption value of household wealth) constraint is binding, ($\Phi > 0$), and (ii) when it is not, ($\Phi = 0$). In the latter case, all policies revert to those of an unconstrained social planner.[36] To set a baseline for comparison, the first proposition establishes results that obtain with homogeneous beliefs under rational expectations RE.

13.2. Homogeneous Beliefs

Proposition 1 (Benchmark rational expectations). *If (i) beliefs are homogeneous and rational, and (ii) labor income plus tax rebates as a proportion of consumption do not decrease over time in all periods, then*

1. *The social cost of carbon exceeds the standard formulation* Λ^s_t *derived in (88) for a social planner, unless* $\bar{\Phi} = 0$*,*

$$\begin{aligned} \varpi^{RE}_t &= \Lambda^s_t + \gamma\bar{\Phi}f(\bar{\Phi},0)[(\mathbb{E}_t\vartheta_{t+1} - \vartheta_t)\Lambda^s_t + cov(\Lambda,\vartheta)] \geq \Lambda^s_t, \;\; \bar{\Phi} > 0, \; \forall t, \\ &= \Lambda^s_t, \;\; \Phi_t = 0 \; \forall t. \end{aligned}$$

2. *The carbon tax premium is positive, unless* $\Phi = 0$*,*

$$\begin{aligned} \mathbb{E}_t\Xi^{e-RE}_{t+j} &= \gamma\bar{\Phi}f(\bar{\Phi},0)\left[\mathbb{E}_t\vartheta_{t+j} - \vartheta_t + cov(\varsigma^e,\vartheta)\right] > 0, \;\; \bar{\Phi} > 0, \; \forall t \\ &= 0, \;\; \bar{\Phi} = 0 \; \forall t. \end{aligned}$$

3. *The ex ante rate on capital is negative, unless $\Phi = 0$,*

$$\begin{aligned} \mathbb{E}_t \Xi_{t+1}^{k-RE} &= -\gamma \bar{\Phi} f(\bar{\Phi}, 0) \Big[\mathbb{E}_t \vartheta_{t+1} - \vartheta_t + cov(\varsigma^k, \vartheta) \Big] < 0, \;\; \bar{\Phi} > 0, \forall t \\ &= 0, \;\; \bar{\Phi} = 0 \; \forall t. \end{aligned}$$

This proposition sets up a background and baseline for comparison with the conclusions for the remaining belief regimes. It also establishes an important distinction between social and Ramsey planning. For a social planner, defined as one for whom the implementability constraint (39) is not binding ($\Phi = 0$)—nor invoked in most of the literature—the policy settings coincide with the formulas derived in Golosov et al. (2014) for the optimal carbon tax and the social cost of carbon, and with the optimal tax on capital derived by Zhu (1992), Chari et al. (1994), and Atkeson et al. (1999). By contrast, a true Ramsey planner, being mindful of preferences and household budgets, adjusts all calculations by factors involving the inter-temporal rate of substitution and the average propensity to consume.

13.3. Heterogeneous Beliefs: Skeptical Consumers

The next proposition establishes that even without ambiguity, skepticism alters optimal policy.

Proposition 2 (No ambiguity). *By producing incentives to spend more on fossil energy and less on capital than is socially optimal, the mere presence of skepticism is an inducement to the Ramsey planner but not a social planner to raise the carbon tax and to subsidize capital, where*

1. *Social cost of carbon is higher than under RE, unless $\Phi = 0$:*

$$\begin{aligned} \varpi_t &= \varpi_t^{RE} > \Lambda_t^s \;\; \Phi_t > 0, \;\; \forall t, \\ &= \Lambda_t^s \;\; \Phi_t = 0; \end{aligned}$$

2. *The carbon tax is higher than under RE, unless $\Phi = 0$:*

$$\begin{aligned} \mathbb{E}_t \Xi_{t+j}^e &\geq \mathbb{E}_t \Xi_{t+j}^{e-RE} \geq 0, \;\; \Phi_t > 0 \;\; \forall t \\ &= 0, \;\; \Phi_t = 0; \end{aligned}$$

3. *Ex ante capital tax rate is less than under RE, unless $\Phi = 0$:*

$$\begin{aligned} \mathbb{E}_t \Xi_{t+1}^k &\leq \mathbb{E}_t \Xi_{t+1}^{k-RE} \leq 0, \;\; \Phi_t > 0 \;\; \forall t \\ &= 0, \;\; \Phi_t = 0. \end{aligned}$$

This proposition highlights an intriguing point: having little faith in climate science, climate skeptics might naturally want to pay a lower carbon tax. Yet, with manifestly poetic justice, the very consequence of skepticism by itself is an increase in both the social cost of carbon and the carbon tax, and a decrease in the tax on capital.

The next two propositions provide results for two regimes in which the Ramsey planner faces a climate-skeptical public whose beliefs are not known and are, indeed unknowable. The first regime is political in that the planner believes the unknown beliefs of the private sector to be true. In the second regime, the planner is paternalistic in that it believes the science model to be true.

Proposition 3 (Political planner). *Ignorance of private beliefs believed to be true leads to the following policy alterations:*

1. *The social cost of carbon contains an ambiguity premium for both Ramsey and social plans:*

$$
\begin{aligned}
\varpi_t &= \varpi_t^{RE} + \bar{\Phi} f(\bar{\Phi}, 0)\Big[(1-\gamma)cov(n^{PO}, \Lambda) + \gamma cov(n^{PO}, \Lambda\vartheta)\Big] \\
&+ f(\bar{\Phi}, 0)cov(n^{PO}, \Lambda) \geq \varpi_t^{RE}, \;\; \forall t, \\
&= \Lambda_t^s + cov(n^{PO}, \Lambda) \geq \Lambda_t^s, \; \bar{\Phi} = 0;
\end{aligned}
$$

2. *The carbon tax is higher than under RE, unless $\Phi = 0$:*

$$
\begin{aligned}
\mathbb{E}_t \Xi_{t+j}^{e-PO} &\geq \mathbb{E}_t \Xi_{t+j}^{e-RE} \geq 0, \;\; \Phi_t > 0 \; \forall t, \\
&= 0, \;\; \Phi_t = 0;
\end{aligned}
$$

3. *Ex ante capital tax is lower than under RE, unless $\Phi = 0$:*

$$
\begin{aligned}
\mathbb{E}_t \Xi_{t+1}^{k-PO} &\leq \mathbb{E}_t \Xi_{t+1}^{k-RE} \leq 0, \; \bar{\Phi} > 0 \;\; \forall t \\
&= 0, \;\; \bar{\Phi} = 0.
\end{aligned}
$$

The preceding proposition establishes that a political planner's ignorance about private beliefs, even if held to be correct, justifies a positive ambiguity premium for the social cost of carbon, activated by correlations between the planner's worst-case belief multiplier n^{PO} and the social cost of carbon that would apply if rational expectations prevailed. In addition, unless the government is a social planner, ambiguity raises the carbon tax and lowers the capital tax.

Proposition 4 (Paternalistic planner). *Ignorance of private beliefs that the planner also believes to be false leads to the following policy alterations:*

1. *The social cost of carbon contains an ambiguity premium, unless $\Phi = 0$:*

$$
\begin{aligned}
\varpi_t &= \varpi_t^{RE} + \Phi_t f(\Phi_t, 0)\Big[(1-\gamma)cov(n^{PA}, \Lambda) + \gamma cov(n^{PA}, \Lambda\vartheta)\Big] \\
&\geq \varpi_t^{RE} \;\; \Phi_t > 0, ; \forall t, \\
&= \Lambda_t^s, \; \Phi_t = 0;
\end{aligned}
$$

2. *The carbon tax contains an ambiguity premium, unless $\Phi = 0$:*

$$
\begin{aligned}
\mathbb{E}_t \Xi_{t+j}^{e-PA} &\geq \mathbb{E}_t \Xi_{t+j}^{e-RE} + \gamma \Phi_t f(\Phi_t, 0) cov(n^{PA}, \varsigma^e \vartheta) \geq \mathbb{E}_t \Xi_{t+1}^{e-RE} \\
&= \geq 0, \;\; \Phi_t > 0 \;\; \forall t, \\
&= 0, \;\; \Phi_t = 0;
\end{aligned}
$$

3. *The ex ante capital tax contains an ambiguity subsidy, unless $\Phi = 0$:*

$$
\begin{aligned}
\mathbb{E}_t \Xi_{t+1}^{k-PA} &\leq \mathbb{E}_t \Xi_{t+1}^{k-RE} - \Phi_t f(\Phi_t, 0)\Big[1 + \gamma cov(n^{PA}, \varsigma^k) + \gamma cov(n^{PA}, \varsigma^k \vartheta)\Big] \\
&\leq \mathbb{E}_t \Xi_{t+1}^{k-RE} \leq 0, \;\; \Phi_t > 0 \;\; \forall t, \\
&= 0, \;\; \Phi_t = 0.
\end{aligned}
$$

The main distinction between a political and a paternalistic planner, as defined in this paper, is that for the latter, an ambiguity premium for the social cost of carbon applies only if the government is a Ramsey planner and not a social planner. For both types of Ramsey planner, a positive ambiguity premium on the social cost of carbon is optimal because the planner's worst-case martingale belief distortion correlates with the certainty-equivalent version of the social cost of carbon.

The next proposition introduces the possibility of the government itself having doubts about the model.

Proposition 5 (Pessimistic planner—skeptical consumer). *Ignorance of private beliefs that the government believes to be false, in combination with a planner's doubts about the model, leads to the following policy alterations:*

1. *The social cost of carbon contains an ambiguity premium in both Ramsey and social plans:*

$$\begin{aligned}\varpi_t &\geq \varpi_t^{RE} + f(\Phi_t, 0)cov(n^p, \Lambda) \\ &+ f(\Phi_t, 0)\Phi_t[(1-\gamma)cov(m^p, \Lambda) + \gamma cov(m^p, \vartheta\Lambda)] \geq \varpi_t^{RE}, \;\; \Phi_t > 0 \\ &\geq \varpi_t^{RE} + cov(n^p, \Lambda) \geq \varpi_t^{RE}, \;\; \Phi_t = 0 \;\; \forall t.\end{aligned}$$

2. *The carbon tax contains a positive or negative ambiguity premium, unless $\Phi = 0$:*

$$\begin{aligned}\mathbb{E}_t\Xi_{t+j}^e &\geq \mathbb{E}_t\Xi_{t+j}^{e-RE} + (1-\gamma)\Phi_t f(\Phi_t, 0)\left[cov(m^p, \frac{1}{n^p}) + cov(\varsigma^e, \frac{m^p}{n^p})\right] \\ &+ \gamma\Phi_t f(\Phi_t, 0)\left[cov(\vartheta, \varsigma^e) + cov(m^p, \frac{1}{n^p}) + cov(\varsigma^e\vartheta, \frac{m^p}{n^p})\right] \\ &\gtreqless \Xi_{t+j}^{e-RE}, \;\; \Phi_t > 0 \;\; \forall t \\ &= 0, \;\; \Phi_t = 0.\end{aligned}$$

3. *The ex ante capital tax is a subsidy in both Ramsey and social plans:*

$$\begin{aligned}\mathbb{E}_t\Xi_{t+1}^{k-R-s} &\leq \mathbb{E}_t\Xi_{t+1}^{k-RE} - f(\Phi_t, 0)[cov(\varsigma^k, n^p) + (1-\gamma)cov(\varsigma^k, m^p)] \\ &- \gamma\Phi_t f(\Phi_t, 0)cov(\vartheta, \varsigma^k m^p)] \leq \mathbb{E}_t\Xi_{t+1}^{k-RE} \leq 0, \;\; \Phi_t \geq 0, \\ &= -cov(\varsigma^k, n^p) - (1-\gamma)cov(\varsigma^k, m^p) \leq 0, \;\; \Phi_t = 0.\end{aligned}$$

Consistent with Propositions 3 and 4, the marginal contribution of an increased correlation between the government's ignorance of skeptical private beliefs $m^p\pi$ with the normalized discounted excess return to fossil energy over the social cost of carbon ς^e or the discounted gross return to capital ς^k, is to raise both the social cost of carbon and the carbon tax, and to lower the *ex ante tax* on capital. The marginal contribution of correlations involving the planner's own model doubts is to raise the social cost of carbon and to lower both the carbon tax and the *ex ante* tax on capital.

In the next and final proposition, the government remains pessimistic, but consumers are pessimistic rather than skeptical.

13.4. Heterogeneous Beliefs: Pessimistic Consumers

Proposition 6 (Pessimistic planner—pessimistic consumer). *The combination of pessimism in the private sector and the government's pessimism has the following implications:*

1. *The social cost of carbon contains an ambiguity premium in both Ramsey and social plans:*

$$\begin{aligned}\varpi_t &\geq [1 - f(\Phi_t, \varepsilon_t)\varepsilon_t]\varpi_t^{RE} + f(\Phi_t, \varepsilon_t)[\varepsilon_t + cov(n^p, \Lambda)] \\ &+ f(\Phi_t, \varepsilon_t)\Phi_t[(1-\gamma)cov(m^c, \Lambda) + \gamma cov(m^c, \vartheta\Lambda)] \\ &\geq \varpi_t^{RE}, \;\; \Phi_t > 0, \\ &\geq \frac{1}{1+\varepsilon_t}[\varpi_t^{RE} + \varepsilon_t + cov(n^p, \Lambda)] > \Lambda_t^s \;\; \Phi_t = 0, \;\; \forall t.\end{aligned}$$

2. *If $\Phi > 0$, the premium on the carbon tax may be positive or negative, and is negative otherwise:*

$$\begin{aligned}
\mathbb{E}_t\Xi^e_{t+j} &\geq \mathbb{E}_t\Xi^{e-RE}_{t+j} \\
&- f(\Phi_t,\varepsilon_t)f(\Phi_t,0)[1+\Phi_t(1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1}+\gamma cov(\varsigma^e,\vartheta))]\varepsilon_t \\
&+ f(\Phi_t,\varepsilon_t)\left[\varepsilon_t + cov(\varsigma^e,\frac{1}{m^c}) + \Phi_t[(1-\gamma)cov(\varsigma^e,\frac{1}{n^p}) + \gamma cov(\vartheta\varsigma^e,\frac{1}{n^p})]\right] \\
&\lesseqqgtr \mathbb{E}_t\Xi^{e-RE}_{t+j}, \quad \Phi_t > 0, \quad \forall t \\
&= \frac{1}{1+\varepsilon_t}cov(\varsigma^e,\frac{1}{m^c}) < 0, \quad \Phi_t = 0. \quad \forall t
\end{aligned}$$

3. *The ex ante capital tax rate may be positive or negative:*

$$\begin{aligned}
\mathbb{E}_t\Xi^{k-R-p}_{t+1} &\leq \mathbb{E}_t\Xi^{k-RE}_{t+1} \\
&- f(\Phi_t,\varepsilon_t)f(\Phi_t,0)\left[1+\Phi_t\left(1-\gamma+\gamma[\mathbb{E}_t\vartheta_{t+1}+cov(\varsigma^k,\vartheta)]\right)\right]\varepsilon_t \\
&- f(\Phi_t,\varepsilon_t)\left[cov(\varsigma^k,\frac{n^p}{m^c}) + cov(\frac{1}{m^c},n^p) + (1+\varepsilon_t)cov(\varsigma^k,n^p)\right] \\
&\lesseqqgtr \mathbb{E}_t\Xi^{k-RE}_{t+1}, \quad \Phi_t > 0 \\
&= -\frac{1}{1+\varepsilon_t}\left[cov(\varsigma^k,\frac{n^p}{m^c}) + cov(\frac{1}{m^c},n^p)\right] \\
&- cov(\varsigma^k,n^p) \lesseqqgtr 0, \quad \Phi_t = 0.
\end{aligned}$$

As in Proposition 5, which concerned skeptical beliefs, in this belief regime, the marginal effect of private-sector pessimism m^c is to raise the social cost of carbon. However, the effect on carbon and capital taxation is the opposite, producing a reduction in the carbon tax and an increase in the tax on capital. The intuition is that increased consumer doubts about the climate model tend to motivate carbon consumption below the socially optimal level and to increase capital spending above its socially optimal level. The net effect of an increase in ϖ_t and a decrease in $\mathbb{E}_t\Xi^e_{t+j}$ may or may not in the end produce a lower carbon tax itself, because from (95) is the sum of two terms: $\varpi_t + \chi_t$, where χ is the sum of terms that contain expected future values of Ξ^e_{t+j}.

Proposition 6 echoes Proposition 5 in that the marginal contribution of any correlation between the planner's own model doubts n^p and asset returns represented by ς^k and ς^e is to likewise raise the social cost of carbon and to lower both the carbon tax and the *ex ante* tax on capital. This becomes apparent if the role of private beliefs is de-activated by setting $m^p \equiv 1$ or $m^c \equiv 1$, respectively, leading to

$$\varpi_t = \varpi^{RE}_t + f(\Phi_t,0)cov(n^p,\Lambda), \tag{115}$$

$$\begin{aligned}
\mathbb{E}_t\Xi^e_{t+j} &\geq \mathbb{E}_t\Xi^{e-RE}_{t+j} \\
&+ f(\Phi_t,0)\Phi_t\left[(1-\gamma)cov(\varsigma^e,\frac{1}{n^p}) + \gamma cov(\vartheta\varsigma^e,\frac{1}{n^p})\right],
\end{aligned} \tag{116}$$

$$\mathbb{E}_t\Xi^{k-R-s}_{t+1} \leq \mathbb{E}_t\Xi^{k-RE}_{t+1} - f(\Phi_t,0)cov(\varsigma^k,n^p). \tag{117}$$

The belief regimes in the three preceding formulas, representing policies stripped of any effects due to belief distortions in the private sector—skeptical or pessimistic—are most closely related to the extant literature on robust climate policy and so serve best for comparisons, to which I now turn.

In a policy regime most similar to that treated in Proposition 5, with m^p set equal to 1 but differing in some details, Hennlock (2009) attributes doubts about the model not to a policy authority *per se* but to a utility-maximizing representative consumer representing society who computes robust feedback rules that, as in (115), generate an ambiguity premium on the expected social cost of carbon, so that with nonlinear damage, policy becomes more responsive to changes in climate. Li et al. (2016) study a dynamic optimization problem that is nearly identical to the planning models in Section 7, but under the assumption that the government is a social planner ($\Phi = 0$ in the present paper) and not a Ramsey planner. They find that even a relatively small increase in the concern about model uncertainty can cause a significant drop in optimal energy extraction and a rise in the socially optimal carbon tax, which in this paper would also correspond to the result in (115). Likewise, Cai and Lontzek (2019) find that with empirically plausible parameterizations of Epstein-Zin preferences to represent attitudes towards risk, the uncertainty associated with anthropogenic climate change implies carbon taxes much higher than associated with deterministic models, while Lemoine and Traeger (2016) conclude that a government's aversion to Knightian uncertainty in the face of an ambiguous tipping point increases the optimal tax on carbon dioxide emissions, but only by a small amount. As with Li et al. (2016), since the derived tax effects come via changes in the SCC, those two conclusions also best correspond to formula (115). Heal and Millner (2013) do not consider taxation *per se*, but find that the value of abatement (that would presumably include carbon taxation) rises as ambiguity aversion increases. Rezai and van der Ploeg (2017) take a somewhat different approach to modeling ambiguity and consider a so-called agnostic policy authority—essentially a government having ambiguity about the approximating climate model—as facing potential models ranging from denialist to scientific. If such a government pursues max-max policies, it imposes higher carbon taxes as a precaution. Later, with the help of Nordhaus's (1993) DICE model to simulate carbon taxation, Rezai and van der Ploeg (2019) broaden their earlier results and consider an agnostic planner who adopts Pascal's Wager (Pascal's (1670))[37] with the question: what would such an agnostic but rational planner—one who does not know or care which model is correct but who wishes to avoid the worst—do when faced with some probability that the approximating model, adhered to by so-called deniers, is false? Their conclusion is that the *hedge-your-bet* optimal carbon tax is quite close to the optimal tax derived in a non-denialist scientific setting, even if the probability of the model being false is a mere ten percent. Further, when ambiguity about whether scientists or deniers are correct rises, as represented by a parameter of constant relative ambiguity aversion, the optimal carbon pricing policy moves ever closer to the science-based policy.

Finally, Anderson et al. (2013), who study a model similar in spirit to the models in this paper, except for an additional robustness channel capable of affecting growth, find that a planner's increased deep uncertainty about the model can result in either a decrease or an increase in the optimal carbon tax, depending on other factors, such as market features and social preferences. In the present paper, the conclusions are driven by similar forces involving preferences and market features, but in the form of second-order moments represented by the covariances between worst-case martingale belief multipliers m and relevant market features, such as net asset-returns to fossil energy and capital.

Table 2 summarizes the preceding propositions. A principal leitmotif of this paper is belief heterogeneity, a main driver of ambiguities in all regimes studied here. Even if (as a mental, albeit unrealistic exercise in Proposition 2) one were to assume away ambiguity, the very presence of heterogeneity in beliefs between the private sector and the government, leading to a positive spread between their respective discount factors, is sufficient to alter the policies of a Ramsey planner, though not those of a social planner. As a consequence, the government increases its estimate of the social cost of carbon and adds a premium to the carbon tax, while raising the capital subsidy rate.

A government with its own doubts about the model is compelled to raise the social cost of carbon and acquires further motives to either raise or lower taxes, depending on the case. However, in all instances, the planner will raise the social cost of carbon. The political and paternalistic planners of Propositions 3 and 4 face ambiguity because of their ignorance of arbitrary private beliefs held by consumers who regard the scientific model with skepticism and discount the future at relatively higher rates. They therefore use relatively more fossil energy and invest in relatively less capital than is socially optimal. To encourage socially optimal choices, the government taxes carbon and subsidizes capital.

The pessimistic planners in Propositions 5 and 6 are identical except with respect to the kind of beliefs they face. The former operates under ambiguities that result from doubts about the model itself and from its ignorance of private beliefs. The latter government confronts a single ambiguity caused by its own doubts about the model, but being constrained by consumers who have known pessimistic beliefs, this government effectively manages two kinds of ambiguities, its own and those of the public. In both cases, the government's doubts about the model, indexed by n^p, provide motives to raise the social cost of carbon and the capital subsidy but to lower the carbon tax. A pessimistic government's ignorance of the beliefs of skeptical consumers in Proposition 5 motivates policies that mimic those of the political and paternalistic planners in Propositions 3 and 4.

Pessimistic consumers discount the future less than do skeptical or rational consumers and therefore use less carbon energy and invest more in capital relative to socially optimal rates and quantities. So the presence of private-sector pessimism is an inducement for the government to lower carbon taxes and to raise the capital tax. Given that the pessimistic government's own doubts produce motives in the opposite direction, the net effect of private and government doubts can be ambiguous.

Table 2. Propositions 1–6.

	Φ	ϖ_t	χ_t	$\bar{\tau}_t^k$
I. Homogeneous Beliefs (RE)				
Ramsey	>0	$\geq\Lambda^s$	$\chi^{RE}\geq 0$	$\tau^{k-RE}\leq 0$
Social	0	Λ^s	0	0
II. Heterogeneous Beliefs				
1. **Skeptical consumers**				
Ambiguity absent				
Ramsey	>0	$\geq\Lambda^s$	$\geq\chi^{RE}$	$\leq\tau^{k-RE}$
Social	0	Λ^s	≥ 0	≤ 0
Ambiguity present				
Political planner				
Ramsey	>0	$\geq\Lambda^s$	$\geq\chi^{RE}$	$\leq\tau^{k-RE}$
Social	0	$\geq\Lambda^s$	≥ 0	≤ 0
Paternalistic planner				
Ramsey	>0	$\geq\Lambda^s$	$\geq\chi_t^{RE}$	$\leq\tau^{k-RE}$
Social	0	$\geq\Lambda^s$	≥ 0	≤ 0
Pessimistic planner				
Ramsey	>0	$\geq\Lambda^s$	$\gtreqless\chi^{RE}$	$\lesseqqgtr\tau^{k-RE}$
Social	0	$\geq\Lambda^s$	$\geq\chi^{RE}$	$\leq\tau^{k-RE}$
Effect of planner's ambiguity				
Ramsey	>0	$\uparrow$	$\uparrow$	$\downarrow$
Social	0	$\uparrow$	$\uparrow$	$\downarrow$
Effect of consumer's ambiguity				
Ramsey	>0	$\uparrow$	$\downarrow$	$\downarrow$
Social	0	$\uparrow$	$\downarrow$	$\downarrow$
2. **Pessimistic consumers**				
Pessimistic planner				
Ramsey	>0	$\geq\Lambda^s$	$\gtreqless\chi^{RE}$	$\lesseqqgtr\tau^{k-RE}$
Social	0	$\geq\Lambda^s$	$\gtreqless\chi_t^{RE}$	$\lesseqqgtr\tau^{k-RE}$
Effect of planner's ambiguity				
Ramsey	>0	$\uparrow$	$\uparrow\downarrow$	$\uparrow\downarrow$
Social	0	$\uparrow$	$\uparrow\downarrow$	$\uparrow\downarrow$
Effect of consumer's ambiguity				
Ramsey	>0	$\uparrow$	$\uparrow\downarrow$	$\uparrow\downarrow$
Social	0	NA	$\downarrow$	$\uparrow\downarrow$

14. A Feedback from Taxes to Consumers' Pessimistic Beliefs

When faced with pessimistic consumers, the policies of a Ramsey planner ($\Phi > 0$) present an instance of a possible two-way feedback between carbon tax policy and private-sector pessimism via the debt channel.[38] For example, consider a positive surprise in debt $b_t > \mathbb{E}_{t-1} m_t^c b_t$. From (77), the shadow value to the taxing authority of the consumer's utility ε_t must drop, implying a drop in the consumer's worst-case utility $\mathcal{U}_t$, hence an increased pessimism via (6).

The tax implications follow from formulas (79), (86), (87), (95), (100) and (112), which show that the social cost of carbon and the carbon tax are decreasing functions of ε_t, and that the *ex ante tax* on capital is an increasing function of ε_t, so that a t-period surprise increase in debt that causes a drop in ε_t and an increase in consumer pessimism should be associated with an increase in the SCC and the carbon tax, and a drop in the tax on capital.

The taxing authority is motivated by two considerations. On the one hand, for purely fiscal reasons, it wants to raise taxes in those states against which it is cheaper to issue debt. It therefore raises tax rates in high-debt states caused by climate shocks, and conversely lowers taxes in low-debt states when climate is calmer. The government also has a goal of setting the prices of carbon and capital in ways that are optimal for society. It turns out that these twin goals coincide: by manipulating debt and taxes to raise (or reduce) the household's utility and turning it more (or less) pessimistic and less (or more) willing to use fossil energy but more (or less) willing to spend on capital, the government manages both, an optimal allocation and a need to smooth debt and taxes over time.

15. Conclusions

This paper is about the climate policy implications of belief heterogeneity and ambiguity in a dynamic market economy governed by a benign welfare maximizing authority, here referred to as a Ramsey planner. To keep this paper focused, many important features relevant for practical policy, such as non-linearities in the mechanisms from the burning of fossil fuels to climate change, are set aside, as are some macroeconomic issues, such as the implications of either exogenous or directed endogenous technological change on growth.[39] Nor do I account for substitution among available types of fossil energy inputs, such as oil, gas, and coal, and green energy from sun, water, and even nuclear power. Finally, following many examples in the literature, to keep the presentation manageable, all sources of uncertainty have been combined into a single "climate-cost shock" variable representing all CO_2-related economic climate damages, including those related to the dynamics of CO_2, productivity shocks other than from climate, and cost shocks from alternative sources of energy, including renewable energy.

This paper, has, in the main, sought to adhere to the spirit of its antecedents, including Stern (2007), Nordhaus (2008), Acemoglu et al. (2012), von Below (2012), van der Ploeg and Withagen (2014), Golosov et al. (2014), Belfiori (2017), Rezai and van der Ploeg (2017), Barrage (2018), and Cai and Lontzek (2019). Golosov et al. (2014), whose model I consider a benchmark for comparisons, proved that the optimal carbon tax (expressed as a proportion of GDP) depends solely on the social cost of carbon, the carbon persistence parameter, and a discount factor. While this is true here as well when beliefs are homogeneous and rational, this paper's value added is an accounting of how belief distortions about the underlying model will alter that discount factor, depending on type of ambiguity.

The principal insight in this paper is that dissonance in beliefs, expressed as skepticism or doubt about forces governing economic outcomes arising from climate cost shocks, produce ambiguity and create wedges between the private sector's discount factor and the government's discount factor and therefore between the way in which consumers and society differentially price two assets with uncertain pay-offs: one from physical capital, whose returns are positive, and the other from atmospheric carbon accumulation, whose returns are negative.

Ambiguity that is due to the government's ignorance of optimistically distorted private beliefs raises the expected social cost of carbon and the carbon tax, which is offset by a lowering of capital taxes. This conclusion carries a certain political irony: a public that confidently denies or minimizes the fact of climate change may actually see its carbon taxes increased.

Ambiguity that arises as a consequence of neither government nor private sector trusting the scientific model has mixed effects depending on the relative degree of pessimism in either sector. So, while any doubts held by either the planner or the private sector produce an ambiguity premium to the social cost of carbon, their differential tax effects are mixed. A combination of government doubt about the model and consumer skepticism motivates lowering the tax on capital, but has ambiguous effects on any carbon tax premium. The combination of fear of model misspecification on the part of both the government and the private sector justifies an ambiguity premium for the social cost of carbon, but opposing effects of some key correlations in the economy leave the net effect on carbon tax premiums ambiguous. The source of this disparity of outcomes derives from two channels: (i) an indirect positive effect through the social cost of carbon and (ii) a direct positive *or* negative effect on the tax itself, depending on the source of ambiguity.

In all cases discussed in this paper, the perspective of Arrow–Debreu asset pricing theory illuminates an equivalence between Pigouvian carbon taxation and optimal pricing of carbon permits in a cap-and-trade economy. This correspondence extends to conditions of Knightian uncertainty, when derived asset prices must meet the test of robustness.

Being theoretic, this paper begs the question of just how applicable the findings herein might be to the real world. In the absence of counter-factual history, the best approach is stochastic simulations of a properly calibrated DSGIE climate-economy model governed by alternative Ramsey regimes discussed here. Alternatively, solving Isaacs-Bellman-Flemming equations associated with the belief regimes in Section 7 as Hennlock (2009) and Li et al. (2016) have done, also holds promise.

Finally, a word of caution. This paper is normative in that it posits not what is but what should be, based on the ideal of welfare maximization implemented by a benign authority heedful of consumers' preferences and budgets. In the reality, such a Ramsey planner may be mere fiction, even in nominal democracies, such as the United States, where an overriding authority, the Supreme Court, has recently ruled that the Executive has no authority to implement carbon policies without explicit and detailed instructions by a legislature that has shown little inclination to address the approaching climate catastrophe. In the real world, policy regimes may even turn rogue: in 2017, the Envronmental Protection Agency under the previous US Administration scrubbed its website of all references to climate. The EPA's website EPA (2017), cited earlier, is available only because it was copied and preserved on another website.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: A previous version of the paper, entitled "Confronting Climate Skepticism or Fear: Ramsey Carbon Taxation and Fiscal Policy under Ambiguity", was presented at the Troxler Symposium, "Economic Analysis, Environmental Uncertainties, and Policy Implications" at East Carolina University, 8 November 2019. I would like to thank all participants for comments and insights.

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

Appendix A. The Household's Intertemporal Budget Constraint

Solve (4) forward for b_t, starting with $t = 0$:

$$\begin{aligned}
b_t \geq & \sum_{t=0}^{T-1} \left(\Pi_{j=0}^{t-1} \sum_{x^{j+1}} \hat{p}_{j+1}(x_{j+1}) \right) [c_t(x^t) - H_t - g_t] \\
& + \left(\Pi_{j=0}^{T-1} \sum_{x^{j+1}} \hat{p}_{j+1} x_{j+1,0} \right) \left[\sum_{x^{T-1}} k_T + \sum_{x^T} \hat{p}_{T+1} x_{T+1} b_{T+1} \right] \\
& + \left(\Pi_{j=0}^{T-1} \sum_{x^{j+1}} \hat{p}_{j+1} x_{j+1} \right) \sum_{x^T} [p_T^e - \tau_T^e] Q_{T-1}(x^T) \\
& + \sum_{t=1}^{T-1} \left(\Pi_{j=0}^{t-1} \sum_{x^{j+1}} \hat{p}_{j+1} x_{j+1} \right) \left[1 - \sum_{x^t} \hat{p}_{t+1} x_{t+1} R_t^k \right] k_t \\
& + \sum_{t=1}^{T-1} \left(\Pi_{j=0}^{t-1} \sum_{x^{j+1}} \hat{p}_{j+1} x_{j+1} \right) \left[1 - \sum_{x^t} \hat{p}_{t+1} x_{t+1} R_t^k \right] [p_t^e - \tau_t^e] Q_t, \qquad \text{(A1)}
\end{aligned}$$

where $p_0 = 1$, and Hotelling's rule (31) ($\frac{p_{t+1}^e - \tau_{t+1}^e}{p_t^e - \tau_t^e} = R_{t+1}^k$) is used in the last term. By the no-arbitrage condition (14), the last two lines are zero. The third line vanishes under the assumption of resource exhaustibility, where $\lim_{T\to\infty} Q_{T-1} = 0$. The second line disappears as a consequence of the household's no-Ponzi game condition

$$\lim_{T\to\infty} \sum_{x^T} \hat{q}_T(x^T)[k_{T+1}(x^T) + \sum_{x_T} \hat{p}_{T+1}(x_{T+1}|x^T) b_{T+1}(x^T)] = 0. \qquad \text{(A2)}$$

This leaves the first line to be evaluated. Note, from (17), the t-step ahead pricing kernel,

$$\begin{aligned}
\hat{q}_t(x^t, x_0) & \equiv \Pi_{j=0}^{t-1} \hat{p}_{j+1,0}(x^j) \\
& = \Pi_{j=0}^{t-1} \beta m_{j+1}(x^{j+1}) \pi_j(x^j) \frac{u_{c_{j+1}}(x^{j+1})}{u_{c_j}(x^j)} \\
& = \beta^t M_t(x^t) \pi_t(x^t) \frac{u_{c_t}(x^t)}{u_{c_0}(x^0)}, \quad q_0 = 1, M_0 = 1, \qquad \text{(A3)}
\end{aligned}$$

is the price of an Arrow–Debreu contract written at $t = 0$. Substituting the last line in (A3) into the term in parentheses in the first line of (A1) then yields an expression for the household's intertemporal budget constraint shown in (35).

Appendix B. Derivation of Implementability Constraint (37)

From (36), the expectation of $\mathcal{W}_{t+1}$ with respect to the distorted measure $m_{t+1}\pi_{t+1}$, valued at the price of an Arrow–Debreu security b, is

$$\begin{aligned}
& \sum_{x_{t+1}} \hat{p}_{t+1}(x_{t+1}|x^t) \mathcal{W}_{t+1}(x^{t+1}) = \\
& \sum_{x_{t+1}} \hat{p}_{t+1}(x_{t+1}|x^t)[b_{t+1}(x^{t+1}) + R_{t+1}^k(x^{t+1}) k_{t+1}(x^t)] \\
+ & \sum_{x_{t+1}} \hat{p}_{t+1}(x_{t+1}|x^t)(p_{t+1}^e(x^{t+1}) - \tau_{t+1}^e(x^{t+1})) Q_{t+1}(x^t) \\
= & \sum_{x_{t+1}} \hat{p}_{t+1}(x_{t+1}|x^t)[b_{t+1}(x^{t+1}) + R_{t+1}^k(x^{t+1}) k_{t+1}(x^t)] \\
+ & [p_t^e(x^t) - \tau_t^e(x^t)] \sum_{x_{t+1}} \hat{p}_{t+1}(x_{t+1}|x^t) R_{t+1}^k(x^{t+1}) Q_{t+1}(x^t),
\end{aligned}$$

where (25) and Hotelling's rule (31) are used to get the last line. Application of no-arbitrage condition (14) further simplifies the expression for household wealth:

$$\begin{aligned}\sum_{x_{t+1}} \hat{p}_{t+1}(x_{t+1}|x^t)\mathcal{W}_{t+1}(x^{t+1}) &= \sum_{x_{t+1}}\Big[\hat{p}_{t+1}(x_{t+1}|x^t)b_{t+1}(x^{t+1}) + k_{t+1}(x^t)\Big] \\ &+ [p^e_t(x^t) - \tau^e_t(x^t)]Q_{t+1}(x^t), \end{aligned} \tag{A4}$$

indicating that the consumer's expected wealth in the following period includes the unused store of fossil fuel valued at its current after-tax price in addition to the capital stock and bonds carried over into the next period. Using the household's period budget constraint (4) and (A4), household wealth defined in (36) is, equivalently,

$$\begin{aligned}\mathcal{W}_t(x^t) &= b_t(x^t) + R^k_t k_t(x^{t-1}) + (p^e_t(x^t) - \tau^e_t(x^t))Q_t(x^{t-1}) \\ &= c_t(x^t) - H_t(x^t) - g_t(x^t) + k_{t+1}(x^t) \\ &+ \sum_{x^{t+1}}\Big[\hat{p}_{t+1}(x_{t+1}|x^t)b_{t+1}(x^{t+1}) + (p^e_t(x^t) - \tau^e_t(x^t))Q_{t+1}(x^t)\Big] \\ &\geq c_t(x^t) - H_t(x^t) - g_t(x^t) + \sum_{x_{t+1}} \hat{p}_{t+1}(x_{t+1}|x^t)\mathcal{W}_{t+1}(x^{t+1}) \\ &= c_t(x^t) - H_t(x^t) - g_t(x^t) + \sum_{x_{t+1}} \hat{\varrho}_{t+1}\pi_{t+1}\mathcal{W}_{t+1}(x^{t+1}) \\ &= c_t(x^t) - H_t(x^t) - g_t(x^t) + \beta\sum_{x_{t+1}} m_{t+1}\pi_{t+1}\frac{u_{c_{t+1}}}{u_{c_t}}\mathcal{W}_{t+1}(x^{t+1}). \end{aligned} \tag{A5}$$

Appendix C. The Social Cost of Carbon

This appendix evaluates the expectation (83) for the four belief regimes:

$$\varpi_t = f(\Phi_t, \varepsilon_t)\mathbb{E}_t\big[n^{**}_{t+1}\Lambda_{t+1}\big(1 + \varepsilon_{t+1} + [1-\gamma+\gamma\vartheta_{t+1}]n^*_{t+1}\Phi_t\big)\big],$$

where $\Lambda_{t+1} = \varrho_{t+1}[(1-\rho)\varpi_{t+1} + x_{t+1}Y_{t+1}]$, $\mathbb{E}_t\Lambda_{t+1} = \Lambda^s_t$, $f(\Phi_t, \varepsilon_t) = \frac{1}{1+\varepsilon_t+[1-\gamma+\gamma\vartheta_t]\Phi_t}$ $\leq f(\Phi_t, 0) < f(0,0) = 1$, $\vartheta_t = \frac{H_t+g_t}{c_t}$ $\forall t$, and $\varepsilon_{t+1} > 0$ only if consumers are pessimistic.[40]

A. Homogeneous beliefs

RE solution ($n^* = 1; n^{**} = 1, \varepsilon = 0, \Phi_t = \bar{\Phi}$)

$$\begin{aligned}\varpi^{RE}_t &= f(\bar{\Phi},0)\mathbb{E}_t[\Lambda_{t+1}(1 + [1-\gamma+\gamma\vartheta_{t+1}]\bar{\Phi})] \\ &= f(\bar{\Phi},0)\mathbb{E}_t[1 + (1-\gamma)\bar{\Phi}]\mathbb{E}_t\Lambda_{t+1} + \gamma\bar{\Phi}f(\bar{\Phi},0)\mathbb{E}_t\Lambda_{t+1}\vartheta_{t+1} \\ &= f(\bar{\Phi},0)[(1 + \bar{\Phi}[1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1}])\Lambda^s_t + \bar{\Phi}\gamma cov(\Lambda,\vartheta)], \\ &= \Lambda^s_t + \gamma\bar{\Phi}f(\bar{\Phi},0)[(\mathbb{E}_t\vartheta_{t+1} - \vartheta_t)\Lambda^s_t + cov(\Lambda,\vartheta)], \quad \bar{\Phi} > 0, \ \forall t, \\ &= \Lambda^s_t, \quad \bar{\Phi} = 0 \ \ \forall t. \end{aligned} \tag{A6}$$

B. Heterogeneous beliefs

1. **Skeptical private sector**
   (a) **Planner has no ambiguity about distorted private beliefs** m^s ($n^* = m^s, n^{**} = 1, \varepsilon = 0$)

$$\begin{aligned}\varpi^o_t &= f(\Phi_t,0)\mathbb{E}_t\big[\Lambda_{t+1}\big(1 + [(1-\gamma)+\gamma\vartheta_{t+1}]m^o_{t+1}\Phi_t\big)\big] \\ &= f(\Phi_t,0)\mathbb{E}_t\Lambda_{t+1} + (1-\gamma)\Phi_t f(\Phi_t,0)\mathbb{E}_t\Lambda_{t+1} + \gamma\Phi_t f(\Phi_t,0)\mathbb{E}_t\Lambda_{t+1}\vartheta_{t+1} \\ &= f(\Phi_t,0)[(1 + \Phi_t(1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1})\Lambda^s_t + \gamma\Phi_t cov(\Lambda,\vartheta)] \\ &= \varpi^{RE}_t, \quad \Phi_t > 0 \\ &= \Lambda^s_t, \quad \Phi_t = 0 \ \forall t. \end{aligned} \tag{A7}$$

   (b) **Political planner** ($n^* = 1; n^{**} = n^{PO}, \varepsilon = 0$)

$$
\begin{aligned}
\varpi_t &= f(\bar{\Phi},0)\mathbb{E}_t\Big[n^{PO}_{t+1}\Lambda_{t+1}(1+[1-\gamma+\gamma\vartheta_{t+1}]\bar{\Phi})\Big] \\
&= f(\bar{\Phi},0)(1+(1-\gamma)\bar{\Phi})\mathbb{E}_t n^{PO}_{t+1}\Lambda_{t+1} + \gamma\bar{\Phi}f(\bar{\Phi},0)\mathbb{E}_t n^{PO}_{t+1}\Lambda_{t+1}\vartheta_{t+1} \\
&= f(\bar{\Phi},0)[(1+\bar{\Phi}[1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1}])\Lambda^s_t + \gamma\bar{\Phi}cov(\Lambda,\vartheta)] + f(\bar{\Phi},0)cov(n^{PO},\Lambda) \\
&+ \bar{\Phi}f(\bar{\Phi},0)\Big[(1-\gamma)cov(n^{PO},\Lambda)+\gamma cov(n^{PO},\Lambda\vartheta)\Big] \\
&= \varpi^{RE}_t + \bar{\Phi}f(\bar{\Phi},0)\Big[(1-\gamma)cov(n^{PO},\Lambda)+\gamma cov(n^{PO},\Lambda\vartheta)\Big] \\
&+ f(\bar{\Phi},0)cov(n^{PO},\Lambda) \geq \varpi^{RE}_t,\ \bar{\Phi}>0,\ \forall t \\
&= \Lambda^s_t + cov(n^{PO},\Lambda) \geq \Lambda^s_t,\ \bar{\Phi}=0.
\end{aligned} \tag{A8}
$$

(c) **Paternalistic planner** $(n^* = n^{PA}, n^{**} = 1, \varepsilon = 0)$

$$
\begin{aligned}
\varpi_t &= f(\Phi_t,0)\mathbb{E}_t\Big[\Lambda_{t+1}\Big(1+[1-\gamma+\gamma\vartheta_{t+1}]n^{PA}_{t+1}\Phi_t\Big)\Big] \\
&= f(\Phi_t,0)(\Lambda^s_t + (1-\gamma)\Phi_t\mathbb{E}_t n^{PA}_{t+1}\Lambda_{t+1} + \gamma\Phi_t f(\Phi_t,0)\mathbb{E}_t n^{PA}_{t+1}\Lambda_{t+1}\vartheta_{t+1} \\
&= f(\Phi_t,0)[(1+\Phi_t[1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1}])\Lambda^s_t + \gamma\Phi_t cov(\Lambda,\vartheta)] + f(\Phi_t,0)cov(n^{PA},\Lambda) \\
&+ \Phi_t f(\Phi_t,0)\Big[(1-\gamma)cov(n^{PA},\Lambda)+\gamma cov(n^{PA},\Lambda\vartheta)\Big] \\
&= \varpi^{RE}_t + \Phi_t f(\Phi_t,0)\Big[(1-\gamma)cov(n^{PA},\Lambda)+\gamma cov(n^{PA},\Lambda\vartheta)\Big] \geq \varpi^{RE}_t,\ \Phi_t>0,\ \forall t \\
&= \Lambda^s_t,\ \Phi_t = 0\ \forall t.
\end{aligned} \tag{A9}
$$

(d) **Pessimistic planner** $(n^* = \frac{m^p}{n^p}, n^{**} = n^p, \varepsilon = 0.$

$$
\begin{aligned}
\varpi_t &= f(\Phi_t,0)\mathbb{E}_t\left[n^p_{t+1}\Lambda_{t+1}\left(1+[1-\gamma+\gamma\vartheta_{t+1}]\frac{m^p_{t+1}}{n^p_{t+1}}\Phi_t\right)\right] \\
&= f(\Phi_t,0)[\Lambda^s_t + cov(n^p,\Lambda)] + (1-\gamma)\Phi_t f(\Phi_t,0)[\Lambda^s_t + cov(m^p,\Lambda)] \\
&+ \gamma\Phi_t f(\Phi_t,0)[\Lambda^s_t\mathbb{E}_t\vartheta_{t+1} + cov(\Lambda,\vartheta) + cov(m^p,\vartheta\Lambda)] \\
&= f(\Phi_t,0)[(1+\Phi_t[1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1})])\Lambda^s_t + \gamma\Phi_t cov(\Lambda,\vartheta)) + f(\Phi_t,0)cov(n^p,\Lambda) \\
&+ f(\Phi_t,0)\Phi_t[(1-\gamma)cov(m^p,\Lambda)+\gamma cov(m^p,\vartheta\Lambda)]. \\
&= \varpi^{RE}_t + f(\Phi_t,0)[cov(n^p,\Lambda) + \Phi_t[(1-\gamma)cov(m^p,\Lambda)+\gamma cov(m^p,\vartheta\Lambda)],\ \Phi_t>0,\ \forall t, \\
&= \Lambda^s_t + cov(n^p,\Lambda) \geq \Lambda^s_t,\ \Phi_t = 0,\ \ \forall t.
\end{aligned} \tag{A10}
$$

2. **Pessimistic consumers**

(a) **Pessimistic planner** $(n^* = \frac{m^c}{n^p}, n^{**} = n^p, \varepsilon \neq 0)$[41]

$$
\begin{aligned}
\varpi_t &= f(\Phi_t,\varepsilon_t)\mathbb{E}_t\left[n^p_{t+1}\Lambda_{t+1}\left(1+\varepsilon_{t+1}+[1-\gamma+\gamma\vartheta_{t+1}]\frac{m^c_{t+1}}{n^p_{t+1}}\Phi_t\right)\right] \\
&= f(\Phi_t,\varepsilon_t)[\mathbb{E}_t n^p_{t+1}\Lambda_{t+1} + \mathbb{E}_t n^p_{t+1}\Lambda_{t+1}\varepsilon_{t+1}] \\
&+ f(\Phi_t,\varepsilon_t)[(1-\gamma)\Phi_t\mathbb{E}_t m^c_{t+1}\Lambda_{t+1} + \gamma\Phi_t\mathbb{E}_t m^c_{t+1}\Lambda_{t+1}\vartheta_{t+1}] \\
&\geq f(\Phi_t,\varepsilon_t)[\Lambda^s_t + cov(n^p,\Lambda) + \varepsilon_t] \\
&+ (1-\gamma)\Phi_t f(\Phi_t,\varepsilon_t)[\Lambda^s_t + cov(m^c,\Lambda)] \\
&+ \gamma\Phi_t f(\Phi_t,\varepsilon_t)[\Lambda^s_t\mathbb{E}_t\vartheta_{t+1} + cov(\Lambda,\vartheta) + cov(m^c,\vartheta\Lambda)] \\
&= f(\Phi_t,\varepsilon_t)[(1+\Phi_t[1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1}])\Lambda^s_t + \gamma\Phi_t cov(\Lambda,\vartheta)] \\
&+ f(\Phi_t,\varepsilon)[\varepsilon_t + cov(n^p,\Lambda)] \\
&+ f(\Phi_t,\varepsilon_t)\Phi_t[(1-\gamma)cov(m^c,\Lambda)+\gamma cov(m^c,\vartheta\Lambda)].
\end{aligned}
$$

Adding and subtracting

$$\varpi_t^{RE} = f(\Phi_t, 0)[(1 + \Phi_t[1 - \gamma + \gamma \mathbb{E}_t \vartheta_{t+1}])\Lambda_t^s + \gamma \Phi_t cov(\Lambda, \vartheta)],$$

leads to,

$$\begin{aligned} \varpi_t &\geq [1 - f(\Phi_t, \varepsilon_t)\varepsilon_t]\varpi_t^{RE} + f(\Phi_t, \varepsilon_t)[\varepsilon_t + cov(n^p, \Lambda)] \\ &+ f(\Phi_t, \varepsilon_t)\Phi_t[(1-\gamma)cov(m^c, \Lambda) + \gamma cov(m^c, \vartheta\Lambda)] \\ &\geq \varpi_t^{RE}, \ \ \Phi_t > 0, \\ &\geq \Lambda_t^* + cov(n^p, \Lambda) > \Lambda_t^s \ \ \Phi_t = 0, \varepsilon_t = 0, \ \ \forall t. \end{aligned} \tag{A11}$$

Appendix D. Sign of Carbon Tax Premium χ

For the decompositions below, recall that the skeptic's belief multiplier m^s is assumed to be an arbitrary, independent random variable, uncorrelated with anything, where $\mathbb{E}_t m^s_{t+j} = 1$. This last property is shared by all martingale multipliers in this paper, so $\mathbb{E}_t n^{PO}_{t+j} = 1$ and $\mathbb{E}_t n^{PA}_{t+j} = 1$, as well. Where applicable, the derivations below apply Jensen's inequality to a convex function, such that, for all versions of m displayed in Table 1, $\mathbb{E}_t \frac{1}{m_{t+1}} \geq \frac{1}{\mathbb{E}_t m_{t+1}} = 1$. Also used, where necessary, is the law of iterated expectations.[42]

From (100),

$$\mathbb{E}_t \Xi^e_{t+j} = \mathbb{E}_t \left[\left(f(\Phi_t, \varepsilon_t) \frac{1}{m_{t+j}} (1 + \varepsilon_{t+j} + [1 - \gamma + \gamma \vartheta_{t+j}] n^*_{t+j} \Phi_{t+j-1}) - 1 \right) \varsigma^e_{t+j} \right],$$

where $\vartheta_t = \frac{H_t + g_t}{c_t}$, $\Phi_{t+j} = n^*_{t+j}\Phi_{t+j-1}$, n^* is defined in Table 1, $f(\Phi_t, \varepsilon_t)$ is defined as before, and $m_{t+j} = m^s_{t+j}$ or $m_{t+j} = m^c_{t+j}$.[43]

A. Homogeneous beliefs

RE solution ($m = 1$, $n^* = 1$, $\varepsilon = 0$, $\Phi_t = \bar{\Phi}$, $\varepsilon = 0$)[44]

$$\begin{aligned} \mathbb{E}_t \Xi^{e-RE}_{t+j} &= \mathbb{E}_t \varsigma^e_{t+j} \big[f(\bar{\Phi}, 0)\big(1 + [1 - \gamma + \gamma \vartheta_{t+j}]\bar{\Phi}\big) - 1 \big] \\ &= f(\bar{\Phi}, 0)\mathbb{E}_t \varsigma^e_{t+j}[1 + (1-\gamma)\bar{\Phi}] + \gamma \bar{\Phi} f(\bar{\Phi}, 0)\mathbb{E}_t \varsigma^e_{t+j}\vartheta_{t+j} - \mathbb{E}_t \varsigma^e_{t+j} \\ &= f(\bar{\Phi}, 0)[1 + \bar{\Phi}(1 - \gamma + \gamma \mathbb{E}_t \vartheta_{t+1} + \gamma cov(\varsigma^e, \vartheta))] - 1 \\ &= \bar{\Phi} f(\bar{\Phi}, 0)\gamma \big[\mathbb{E}_t \vartheta_{t+j} - \vartheta_t + cov(\varsigma^e, \vartheta)\big] > 0, \ \ \bar{\Phi} > 0, \\ &= 0, \ \ \bar{\Phi} = 0 \ \forall t. \end{aligned} \tag{A12}$$

The preceding utilizes $\mathbb{E}_t \varsigma^e_{t+1} = 1$, and

$$\begin{aligned} f(\Phi_t, 0)[1 + (1 - \gamma + \gamma \mathbb{E}_t \vartheta_{t+j})\Phi_t] - 1 &= \frac{1 + (1 - \gamma + \gamma \mathbb{E}_t \vartheta_{t+j})\Phi_t}{1 + (1 - \gamma + \gamma \vartheta_t]\Phi_t} - 1 \\ &= \gamma \Phi_t f(\Phi_t, 0)(\mathbb{E}_t \vartheta_{t+j} - \vartheta_t). \end{aligned}$$

B. Heterogeneous beliefs

1. **Skeptical private sector**

(a) **No ambiguity** ($m = m^s$; $n^* = 1 \rightarrow \Phi_t = \bar{\Phi}_t$, $\varepsilon = 0$)[45]

$$\begin{aligned}
\mathbb{E}_t\Xi^e_{t+j} &= \mathbb{E}_t\varsigma^e_{t+j}\left[\frac{f(\bar{\Phi},0)}{m^s_{t+j}}\left(1+[1-\gamma+\gamma\vartheta_{t+j}]\bar{\Phi}\right)-1\right] \\
&= f(\bar{\Phi},0)\mathbb{E}_t\varsigma^e_{t+j}\frac{1}{m^s_{t+j}}[1+(1-\gamma)\bar{\Phi}] \\
&+ \gamma f(\bar{\Phi},0)\mathbb{E}_t\varsigma^e_{t+j}\frac{1}{m^s_{t+j}}\vartheta_{t+j}\bar{\Phi}-\mathbb{E}_t\varsigma^e_{t+j} \\
&\geq f(\bar{\Phi},0)\frac{1}{\mathbb{E}_t m^s_{t+j}}\mathbb{E}_t[\varsigma^e_{t+j}(1+(1-\gamma)\bar{\Phi})] \\
&+ \gamma\bar{\Phi}f(\bar{\Phi},0)\frac{1}{\mathbb{E}_t m^s_{t+j}}\left(\mathbb{E}_t\varsigma^e_{t+j}\vartheta_{t+j}+cov(\varsigma^e,\vartheta)\right)-\mathbb{E}_t\varsigma^e \\
&= f(\bar{\Phi},0)\left[1+\bar{\Phi}\left(1-\gamma+\gamma[\mathbb{E}_t\vartheta_{t+j}+cov(\varsigma^e,\vartheta)]\right)\right]-1 \\
&= \gamma\bar{\Phi}f(\bar{\Phi},0)\left[\mathbb{E}_t\vartheta_{t+j}-\vartheta_t+cov(\varsigma^e,\vartheta)\right]>0 \\
&= \mathbb{E}_t\Xi^{e-RE}_{t+j}>0,\ \ \bar{\Phi}>0, \\
&= 0,\ \ \bar{\Phi}=0,\ \forall t.
\end{aligned} \tag{A13}$$

(b) **Political planner** ($m = m^s$, $n^* = 1$, $\varepsilon = 0$).

$$\begin{aligned}
\mathbb{E}_t\Xi^{e-PO}_{t+j} &= \mathbb{E}_t\varsigma^e_{t+j}\left[\frac{f(\bar{\Phi},0)}{m^s_{t+j}}\left(1+[1-\gamma+\gamma\vartheta_{t+j}]\bar{\Phi}\right)-1\right] \\
&= f(\bar{\Phi},0)\mathbb{E}_t\frac{1}{m^s_{t+j}}\mathbb{E}_t\varsigma^e_{t+j}[1+(1-\gamma)\bar{\Phi}] \\
&+ \gamma f(\bar{\Phi},0)\mathbb{E}_t\frac{1}{m^s_{t+j}}\varsigma^e_{t+j}\vartheta_{t+j}\bar{\Phi}-\mathbb{E}_t\varsigma^e_{t+j} \\
&\geq f(\bar{\Phi},0)\left[1+\bar{\Phi}\left(1-\gamma+\gamma[\mathbb{E}_t\vartheta_{t+j}+cov(\varsigma^e,\vartheta)]\right)\right]-1 \\
&= \bar{\Phi}f(\bar{\Phi},0)\gamma\left[\mathbb{E}_t\vartheta_{t+j}-\vartheta_t+cov(\varsigma^e,\vartheta)\right]>0, \\
&= \mathbb{E}_t\Xi^{e-RE}_{t+j},\ \ \bar{\Phi}>0, \\
&= 0,\ \ \bar{\Phi}=0,\ \forall t.
\end{aligned} \tag{A14}$$

(c) **Paternalistic planner** ($m = m^s$, $n^* = n^{PA}$; $\Phi_{t+j} = n^{PA}_{t+j}\Phi_{t+j-1}$, $\varepsilon = 0$)

$$\begin{aligned}
\mathbb{E}_t \Xi^{e-PA}_{t+j} &= \mathbb{E}_t \varsigma^e_{t+j} \left[\frac{f(\Phi_t, 0)}{m^s_{t+j}} \Big(1 + [1 - \gamma + \gamma \vartheta_{t+j}] n^{PA}_{t+j} \Phi_{t+j-1} \Big) - 1 \right] \\
&= f(\Phi_t, 0) \mathbb{E}_t \varsigma^e_{t+j} \frac{1}{m^s_{t+j}} [1 + (1-\gamma) n^{PA}_{t+j} \Phi_{t+j-1}] \\
&+ \gamma f(\Phi_t, 0) \mathbb{E}_t \frac{1}{m^s_{t+j}} \varsigma^e_{t+j} \vartheta_{t+j} n^{PA}_{t+j} \Phi_{t+j-1} - \mathbb{E}_t \varsigma^e_{t+j} \\
&\geq f(\Phi_t, 0) \frac{\mathbb{E}_t \varsigma^e_{t+j}}{\mathbb{E}_t m^s_{t+j}} + (1-\gamma) \Phi_t f(\Phi_t, 0) \frac{\mathbb{E}_t \varsigma^e_{t+j}}{\mathbb{E}_t m^s_{t+j}} \\
&+ \gamma \Phi_t f(\Phi_t, 0) \left(\frac{\mathbb{E}_t \varsigma^e_{t+j}}{\mathbb{E}_t m^s_{t+j}} \mathbb{E}_t \vartheta_{t+j} + cov(n^{PA}, \varsigma^e \vartheta) + cov(\varsigma^e, \vartheta) \right) - \mathbb{E}_t \varsigma^e_{t+j} \\
&= f(\Phi_t, 0) \big[1 + \Phi_t \big(1 - \gamma + \gamma [\mathbb{E}_t \vartheta_{t+j} + cov(\varsigma^e, \vartheta)] \big) \big] - 1 \\
&+ \gamma \Phi_t f(\Phi_t, 0) cov(n^{PA}, \varsigma^e \vartheta) \\
&= \mathbb{E}_t \Xi^{e-RE}_{t+j} + \gamma \Phi_t f(\Phi_t, 0) cov(n^{PA}, \varsigma^e \vartheta) > 0, \;\; \Phi_t > 0, \\
&\geq 0, \;\; \Phi_t = 0, \; \forall t.
\end{aligned} \tag{A15}$$

(d) **Pessimistic planner** ($m = m^s$, $n^* = \frac{m^p}{n^p}$; $\Phi_{t+j} = \frac{m^p_{t+j}}{n^p_{t+j}} \Phi_{t+j-1}$, $\varepsilon = 0$)[46]

$$\begin{aligned}
\mathbb{E}_t \Xi^{e}_{t+j} &= \mathbb{E}_t \varsigma^e_{t+j} \left[f(\Phi_t, 0) \frac{1}{m^s_{t+j}} \left(1 + [1 - \gamma + \gamma \vartheta_{t+j}] \frac{m^p_{t+1}}{n^p_{t+1}} \Phi_{t+j-1} \right) - 1 \right] \\
&= f(\Phi_t, 0) \mathbb{E}_t \left[\frac{\varsigma^e_{t+j}}{m^s_{t+j}} \left(1 + [1 - \gamma + \gamma \vartheta_{t+j}] \frac{m^p_{t+j}}{n^p_{t+j}} \Phi_{t+j-1} \right) \right] - 1 \\
&= f(\Phi_t, 0) \left[\mathbb{E}_t \left(\frac{\varsigma^e_{t+j}}{m^s_{t+j}} \right) + (1-\gamma) \Phi_t \left(\mathbb{E}_t \frac{\varsigma^e_{t+j} m^p_{t+j}}{m^s_{t+j} n^p_{t+j}} \right) + \gamma \Phi_t \left(\mathbb{E}_t \frac{\varsigma^e_{t+j} \vartheta_{t+j} m^p_{t+j}}{m^s_{t+j} n^p_{t+j}} \right) \right] - 1 \\
&\geq f(\Phi_t, 0) [1 + \Phi_t (1 - \gamma + \gamma \mathbb{E}_t \vartheta_{t+1} + \gamma cov(\varsigma^e, \vartheta))] - 1 \\
&+ \Phi_t f(\Phi_t, 0)(1-\gamma) \left[cov(m^p, \frac{1}{n^p}) + cov(\varsigma^e, \frac{m^p}{n^p}) \right] \\
&+ \Phi_t f(\Phi_t, 0) \gamma \left[cov(\vartheta, \varsigma^e) + cov(m^p, \frac{1}{n^p}) + cov(\varsigma^e \vartheta, \frac{m^p}{n^p}) \right] \\
&= \mathbb{E}_t \Xi^{e-RE}_{t+j} + (1-\gamma) \Phi_t f(\Phi_t, 0) \left[cov(m^p, \frac{1}{n^p}) + cov(\varsigma^e, \frac{m^p}{n^p}) \right] \\
&+ \gamma \Phi_t f(\Phi_t, 0) \left[cov(\vartheta, \varsigma^e) + cov(m^p, \frac{1}{n^p}) + cov(\varsigma^e \vartheta, \frac{m^p}{n^p}) \right] \gtreqless 0; \;\; \Phi_t \geq 0, \\
&= \mathbb{E}_t \Xi^{e-RE}_{t+j} > 0, \;\; \Phi_t = 0.
\end{aligned} \tag{A16}$$

2. **Pessimistic consumers**

(a) **Pessimistic planner** ($m = m^c, n^* = \frac{m^c}{n^p}; \Phi_{t+j} = \frac{m^c_{t+j}}{n^p_{t+j}}\Phi_{t+j-1}, \varepsilon \neq 0$)[47],

$$\begin{aligned}
\mathbb{E}_t\Xi^e_{t+j} &= \mathbb{E}_t\varsigma^e_{t+j}\left[f(\Phi_t,\varepsilon_t)\frac{1}{m^c_{t+j}}\left(1+\varepsilon_{t+j}+[1-\gamma+\gamma\vartheta_{t+j}]\frac{m^c_{t+1}}{n^p_{t+1}}\Phi_{t+j-1}\right)-1\right] \\
&= f(\Phi_t,\varepsilon_t)\mathbb{E}_t\left[\frac{\varsigma^e_{t+j}}{m^c_{t+j}}\left(1+\varepsilon_{t+j}+[1-\gamma+\gamma\vartheta_{t+j}]\frac{m^c_{t+j}}{n^p_{t+j}}\Phi_{t+j-1}\right)\right]-1 \\
&\geq f(\Phi_t,\varepsilon_t)\left[\frac{\mathbb{E}_t\varsigma^e_{t+j}}{\mathbb{E}_t m^c_{t+j}}+\varepsilon_t+cov(\varsigma^e,\frac{1}{m^c})\right]-1 \\
&+ (1-\gamma)\Phi_t f(\Phi_t,\varepsilon_t)\left[\left(\frac{\mathbb{E}_t\varsigma^e_{t+j}}{\mathbb{E}_t n^p_{t+j}}\right)+cov(\varsigma^e,\frac{1}{n^p})\right] \\
&+ \gamma\Phi_t f(\Phi_t,\varepsilon_t)\left[\left(\frac{\mathbb{E}_t\varsigma^e_{t+j}\vartheta_{t+j}}{\mathbb{E}_t n^p_{t+j}}\right)+cov(\varsigma^e\vartheta,\frac{1}{n^p})\right] \\
&= f(\Phi_t,\varepsilon_t)\left[1+\varepsilon_t+cov(\varsigma^e,\frac{1}{m^c})\right]-1 \\
&+ (1-\gamma)\Phi_t f(\Phi_t,\varepsilon_t)\left[1+cov(\varsigma^e,\frac{1}{n^p})\right] \\
&+ \gamma\Phi_t f(\Phi_t,\varepsilon_t)\left[\mathbb{E}_t\vartheta_{t+j}+cov(\frac{1}{n^p},\varsigma^e\vartheta)+cov(\varsigma^e,\vartheta)\right] \\
&= f(\Phi_t,\varepsilon_t)[1+\Phi_t(1-\gamma+\gamma(\mathbb{E}_t\vartheta_{t+j}+cov(\varsigma^e,\vartheta)))]-1 \\
&+ f(\Phi_t,\varepsilon_t)\left[\varepsilon_t+cov(\varsigma^e,\frac{1}{m^c})\right] \\
&+ \Phi_t f(\Phi_t,\varepsilon_t)\left[(1-\gamma)cov(\varsigma^e,\frac{1}{n^p})+\gamma cov(\vartheta\varsigma^e,\frac{1}{n^p})\right].
\end{aligned}$$

Adding and subtracting

$$f(\Phi_t,0)[1+\Phi_t(1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1}+\gamma cov(\varsigma^e,\vartheta))],$$

leads to

$$\begin{aligned}
\mathbb{E}_t\Xi^e_{t+j} &= \mathbb{E}_t\Xi^{e-RE}_{t+j} - f(\Phi_t,\varepsilon_t)f(\Phi_t,0)[1+\Phi_t(1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1}+\gamma cov(\varsigma^e,\vartheta))]\varepsilon_t \\
&+ f(\Phi_t,\varepsilon_t)\left[\varepsilon_t+cov(\varsigma^e,\frac{1}{m^c})\right] \\
&+ \Phi_t f(\Phi_t,\varepsilon_t)\left[(1-\gamma)cov(\varsigma^e,\frac{1}{n^p})+\gamma cov(\vartheta\varsigma^e,\frac{1}{n^p})\right] \lesseqqgtr 0,\;\; \Phi_t>0,\;\; \forall t \\
&= cov(\varsigma^e,\frac{1}{m^c})<0,\;\; \Phi_t=0,\;\; \forall t,
\end{aligned} \tag{A17}$$

where, from (75) and (77),

$$
\begin{aligned}
\mathbb{E}_t \frac{\varsigma^e_{t+1}}{m^c_{t+j}} \varepsilon_{t+1} &= \sigma^c M^c_t \left(\mathbb{E}_t \varsigma^e_{t+1} \mu_{t+1} - \mathbb{E}_t \varsigma^e_{t+1} \mathbb{E}_t m^c_{t+1} \mu_{t+1} \right) + \mathbb{E}_t \frac{\varsigma^e_{t+1}}{m^c_{t+1}} \varepsilon_t \\
&\geq \sigma^c \Phi_t \left[\mathbb{E}_t (1 - m^c_{t+1}) \mathbb{E}_t u_{c,t+1} b_{t+1} + cov(\varsigma^e, bu_c) - cov(m^c, bu_c) \right] \\
&+ [1 + cov(\varsigma^e, \frac{1}{m^c})] \varepsilon_t \\
&= \sigma^c \Phi_t [cov(\varsigma^e, bu_c) - cov(m^c, bu_c)] + [1 + cov(\varsigma^e, \frac{1}{m^c})] \varepsilon_t \geq \varepsilon_t.
\end{aligned}
$$

Appendix E. Sign of Ex Ante Capital Tax

This appendix evaluates the expression (112) for each of the four belief regimes,

$$
\mathbb{E}_t \Xi^k_{t+1} = \mathbb{E}_t \varsigma^k_{t+1} \left(1 - f(\Phi_t, \varepsilon_t) \frac{n^{**}_{t+1}}{m_{t+1}} (1 + \varepsilon_{t+1} + [1 - \gamma + \gamma \vartheta_{t+1}] n^*_{t+1} \Phi_t \right).
$$

A. Homogeneous beliefs
RE solution ($m = 1, n^* = 1, n^{**} = 1, \varepsilon = 0$)

$$
\begin{aligned}
\mathbb{E}_t \Xi^{k-RE}_{t+1} &= \mathbb{E}_t \left[\varsigma^k_{t+1} (1 - f(\Phi_t, 0)[1 + \Phi_t [1 - \gamma + \gamma \vartheta_{t+1}]]) \right] \\
&= \mathbb{E}_t \varsigma^k_{t+1} - f(\Phi_t, 0)[1 + (1-\gamma)\Phi_t] \mathbb{E}_t \varsigma^k_{t+1} + \gamma \Phi_t f(\Phi_t, 0) \mathbb{E}_t \left[\varsigma^k_{t+1} \vartheta_{t+1} \right] \\
&= 1 - f(\Phi_t, 0) \left[1 + \Phi_t \left(1 - \gamma + \gamma [\mathbb{E}_t \vartheta_{t+1} + cov(\varsigma^k, \vartheta)] \right) \right] \\
&= -\gamma f(\Phi_t, 0) \Phi_t \left[\mathbb{E}_t \vartheta_{t+1} - \vartheta_t + cov(\varsigma^k, \vartheta) \right] < 0, \quad \Phi_t > 0, \quad \forall t, \\
&= 0, \quad \Phi_t = 0 \quad \forall t.
\end{aligned} \tag{A18}
$$

B. Heterogeneous beliefs

1. **Skeptical private sector**
   (a) **No ambiguity** ($m = m^s, n^* = 1, n^{**} = 1, \varepsilon = 0$).

$$
\begin{aligned}
\mathbb{E}_t \Xi^k_{t+1} &= \mathbb{E}_t \left[\varsigma^k_{t+1} \left(1 - \frac{f(\Phi_t, 0)}{m^s_{t+1}} (1 + [(1-\gamma) + \gamma \vartheta_{t+1}] \Phi_t) \right) \right] \\
&= \mathbb{E}_t \varsigma^k_{t+1} - f(\Phi_t, 0) \mathbb{E}_t \varsigma^k_{t+1} \left[\frac{1}{m^s_{t+1}} (1 + [(1-\gamma) + \gamma \vartheta_{t+1})] \Phi_t \right] \\
&\leq 1 - f(\Phi_t, 0) \frac{\mathbb{E}_t \varsigma^k_{t+1}}{\mathbb{E}_t m^s_{t+1}} - (1-\gamma) \Phi_t f(\Phi_t, 0) \mathbb{E}_t \varsigma^k_{t+1} \\
&- \gamma \Phi_t f(\Phi_t, 0) \mathbb{E}_t \varsigma^k_{t+1} \mathbb{E}_t \vartheta_{t+1} - \gamma \Phi_t f(\Phi_t, 0) cov(\varsigma^k, \vartheta) \\
&= 1 - f(\Phi_t, 0) \left[1 + \Phi_t \left(1 - \gamma + \gamma ([\mathbb{E}_t \vartheta_{t+1} + cov(\varsigma^k, \vartheta)] \right) \right] \\
&= -\gamma \Phi_t f(\Phi_t, 0) \left[\mathbb{E}_t \vartheta_{t+1} - \vartheta_t + cov(\varsigma^k, \vartheta) \right] \\
&= \mathbb{E}_t \Xi^{k-RE}_{t+1} < 0, \quad \Phi_t > 0, \ \forall t \\
&= 0, \quad \Phi_t = 0 \ \forall t.
\end{aligned} \tag{A19}
$$

   (b) **Political planner** ($m = m^s, n^* = 1, n^{**} = n^{PO}, \varepsilon = 0$).

$$\begin{aligned}
\mathbb{E}_t\Xi_{t+1}^{k-PO} &= \mathbb{E}_t\left[\varsigma_{t+1}^k\left(1-f(\bar{\Phi},0)\frac{n_{t+1}^{PO}}{m_{t+1}^s}[1+[1-\gamma+\gamma\vartheta_{t+1}]\bar{\Phi}]\right)\right] \\
&= \mathbb{E}_t\varsigma_{t+1}^k - f(\bar{\Phi},0)\left[1+(1-\gamma)\bar{\Phi}\mathbb{E}_t\varsigma_{t+1}^k\frac{n_{t+1}^{PO}}{m_{t+1}^s}+\gamma\bar{\Phi}\mathbb{E}_t\varsigma_{t+1}^k\frac{n_{t+1}^{PO}}{m_{t+1}^s}\vartheta_{t+1}\right] \\
&\leq 1-f(\bar{\Phi},0)\Big[1+\bar{\Phi}\Big(1-\gamma+\gamma[\mathbb{E}_t\vartheta_{t+1}+cov(\varsigma^k,\vartheta)\Big]\Big) \\
&- \gamma\bar{\Phi}f(\bar{\Phi},0)cov(n^{PO},\varsigma^k\vartheta)-(1-\gamma)\bar{\Phi}f(\bar{\Phi},0)cov(n^{PO},\varsigma^k) \\
&= 1-f(\bar{\Phi},0)\Big[1+\Big(1-\gamma+\gamma\mathbb{E}_t\vartheta_{t+1}+\gamma cov(\varsigma^k,\vartheta)\Big)\bar{\Phi}\Big] \\
&- \bar{\Phi}f(\bar{\Phi},0)\Big[(1-\gamma)cov(n^{PO},\varsigma^k)+\gamma cov(n^{PO},\varsigma^k\vartheta)\Big] \\
&= \mathbb{E}_t\Xi_{t+1}^{k-RE}-\bar{\Phi}f(\bar{\Phi},0)\Big[(1-\gamma)cov(n^{PO},\varsigma^k)+\gamma cov(n^{PO},\varsigma^k\vartheta)\Big] \leq \mathbb{E}_t\Xi_{t+1}^{k-RE}<0, \bar{\Phi}>0, \\
&= 0,\ \ \bar{\Phi}=0.
\end{aligned} \tag{A20}$$

(c) **Paternalistic planner** ($m=m^s$, $n^*=n^{PA}$, $n^{**}=1$, $\varepsilon=0$).

$$\begin{aligned}
\mathbb{E}_t\Xi_{t+1}^{k-PA} &= \mathbb{E}_t\varsigma_{t+1}^k - f(\Phi_t,0)\mathbb{E}_t\frac{\varsigma_{t+1}^k}{m_{t+1}^s}\Big[(1+[(1-\gamma)+\gamma\vartheta_{t+1}]n_{t+1}^{PA}\Phi_t\Big] \\
&\leq 1-f(\Phi_t,0)\left[\frac{\mathbb{E}_t\varsigma_{t+1}^k}{\mathbb{E}_t m_{t+1}^s}+\Phi_t(1-\gamma)\mathbb{E}_t\varsigma_{t+1}^k\frac{\mathbb{E}_t n_{t+1}^{PA}}{\mathbb{E}_t m_{t+1}^s}\right] \\
&- \gamma\Phi_t f(\Phi_t,0)\left[\mathbb{E}_t\varsigma_{t+1}^k\frac{\mathbb{E}_t n_{t+1}^{PA}}{\mathbb{E}_t m_{t+1}^s}\mathbb{E}_t\vartheta_{t+1}+cov(n^{PA},\varsigma^k\vartheta)+cov(\varsigma^k,\vartheta)\right] \\
&- (1-\gamma)\Phi_t f(\Phi_t,0)cov(n^{PA},\varsigma^k) \\
&= 1-f(\Phi_t,0)\Big[1+\Phi_t\Big(1-\gamma+\gamma[\mathbb{E}_t\vartheta_{t+1}+\gamma cov(\varsigma^k,\vartheta)]\Big)\Big] \\
&- \Phi_t f(\Phi_t,0)\Big[(1-\gamma)cov(n^{PA},\varsigma^k)+\gamma cov(n^{PA},\varsigma^k\vartheta)\Big] \\
&= \mathbb{E}_t\Xi_{t+1}^{k-RE}-\Phi_t f(\Phi_t,0)\Big[(1-\gamma)cov(n^{PA},\varsigma^k)+\gamma cov(n^{PA},\varsigma^k\vartheta)\Big] \leq \mathbb{E}_t\Xi_{t+1}^{k-RE}<0, \Phi_t>0,\ \forall t \\
&= 0,\ \ \Phi_t=0,\ \forall t.
\end{aligned} \tag{A21}$$

(d) **Pessimistic planner** ($m=m^s$, $n^*=\frac{m^p}{n^p}$, $n^{**}=n^p$, $\varepsilon=0$)

$$\begin{aligned}
\mathbb{E}_t\Xi_{t+1}^{k-R-s} &= \mathbb{E}_t\left[\varsigma_{t+1}^k\left(1-f(\Phi_t,0)\frac{n_{t+1}^p}{m_{t+1}^s}\left[1+[1-\gamma+\gamma\vartheta_{t+1}]\frac{m_{t+1}^p}{n_{t+1}^p}\Phi_t\right]\right)\right] \\
&= \mathbb{E}_t\varsigma_{t+1}^k - f(\Phi_t,0)\left[\mathbb{E}_t\frac{n_{t+1}^p}{m_{t+1}^s}\varsigma_{t+1}^k+\Phi_t\mathbb{E}_t(1-\gamma+\gamma\vartheta_{t+1})\frac{m_{t+1}^p\varsigma_{t+1}^k}{m_{t+1}^s}\right] \\
&\leq 1-f(\Phi_t,0)\Big[1+cov(\varsigma^k,n^p)+(1-\gamma)\Phi_t(1+cov(\varsigma^k,m^p))\Big] \\
&- \gamma\Phi_t f(\Phi_t,0)\Big[\mathbb{E}_t\vartheta_{t+1}+cov(\varsigma^k,\vartheta)+cov(\vartheta,m^p\varsigma^k)\Big] \\
&= 1-f(\Phi_t,0)\Big[1+\Phi_t\Big(1-\gamma+\gamma[\mathbb{E}_t\vartheta_{t+1}+cov(\varsigma^k,\vartheta)]\Big)\Big] \\
&- f(\Phi_t,0)[cov(\varsigma^k,n^p)+(1-\gamma)\Phi_t cov(\varsigma^k,m^p)]-\gamma\Phi_t f(\Phi_t,0)cov(\vartheta,\varsigma^e m^p)] \\
&= \mathbb{E}_t\Xi_{t+1}^{k-RE}-f(\Phi_t,0)[cov(\varsigma^k,n^p)+(1-\gamma)cov(\varsigma^k,m^p)] \\
&- \gamma\Phi_t f(\Phi_t,0)cov(\vartheta,\varsigma^e m^p)\leq 0,\ \ \Phi_t\geq 0 \\
&= \mathbb{E}_t\Xi_{t+1}^{k-RE}-cov(\varsigma^k,n^p)-(1-\gamma)cov(\varsigma^k,m^p)\leq 0,\ \ \Phi_t=0.
\end{aligned} \tag{A22}$$

2. **Pessimistic consumers**
   (a) **Pessimistic planner** ($m = m^c$, $n^* = \frac{m^c}{n^p}$, $n^{**} = n^p$, $\varepsilon \neq 0$)

$$
\begin{aligned}
\mathbb{E}_t \Xi_{t+1}^{k-R-p} &= \mathbb{E}_t \left[\varsigma_{t+1}^k \left(1 - f(\Phi_t, \varepsilon_t) \frac{n_{t+1}^p}{m_{t+1}^c} \left[1 + \varepsilon_{t+1} + [1 - \gamma + \gamma \vartheta_{t+1}] \frac{m_{t+1}^c}{n_{t+1}^p} \Phi_t \right] \right) \right] \\
&= \mathbb{E}_t \varsigma_{t+1}^k - f(\Phi_t, \varepsilon_t) \left[\mathbb{E}_t \frac{n_{t+1}^p}{m_{t+1}^c} \varsigma_{t+1}^k (1 + \varepsilon_{t+1}) + \Phi_t \mathbb{E}_t (1 - \gamma + \gamma \vartheta_{t+1}) \varsigma_{t+1}^k \right] \\
&\leq 1 - f(\Phi_t, \varepsilon_t) \left[\frac{\mathbb{E}_t n_{t+1}^p}{\mathbb{E}_t m_{t+1}^c} \mathbb{E}_t \varsigma_{t+1}^k + \Phi_t (1 - \gamma + \gamma \mathbb{E}_t \vartheta_{t+1}) \mathbb{E}_t \varsigma_{t+1}^k \right] \\
&- f(\Phi_t, \varepsilon_t) \left[\gamma \Phi_t cov(\varsigma^k, \vartheta) + cov(\varsigma^k, \frac{n^p}{m^c}) + cov(\frac{1}{m^c}, n^p) - (1 + \varepsilon_t) cov(\varsigma^k, n^p) \right] \\
&= 1 - f(\Phi_t, \varepsilon_t) \left[1 + \Phi_t (1 - \gamma + \gamma (\mathbb{E}_t \vartheta_{t+1} + cov(\varsigma^k, \vartheta)) \right] \\
&- f(\Phi_t, \varepsilon_t) \left[cov(\varsigma^k, \frac{n^p}{m^c}) + cov(\frac{1}{m^c}, n^p) - (1 + \varepsilon_t) cov(\varsigma^k, n^p) \right].
\end{aligned}
$$

Adding and subtracting

$$
f(\Phi_t, 0) \left[1 + \Phi_t \left(1 - \gamma + \gamma [\mathbb{E}_t \vartheta_{t+1} + cov(\varsigma^k, \vartheta)] \right) \right],
$$

leads to

$$
\begin{aligned}
\mathbb{E}_t \Xi_{t+1}^{k-R-p} &\leq \mathbb{E}_t \Xi_{t+1}^{k-RE} - f(\Phi_t, \varepsilon_t) f(\Phi_t, 0) \left[1 + \Phi_t \left(1 - \gamma + \gamma [\mathbb{E}_t \vartheta_{t+1} + cov(\varsigma^k, \vartheta)] \right) \right] \varepsilon_t \\
&- f(\Phi_t, \varepsilon_t) \left[cov(\varsigma^k, \frac{n^p}{m^c}) + cov(\frac{1}{m^c}, n^p) + (1 + \varepsilon_t) cov(\varsigma^k, n^p) \right] \lesseqgtr 0, \quad \Phi_t > 0 \\
&= - \left[cov(\varsigma^k, \frac{n^p}{m^c}) + cov(\frac{1}{m^c}, n^p) + (1 + \varepsilon_t) cov(\varsigma^k, n^p) \right] \lesseqgtr 0, \quad \Phi_t = 0,
\end{aligned} \tag{A23}
$$

where, from (75) and (77),

$$
\begin{aligned}
\mathbb{E}_t \frac{\varsigma_{t+1}^k n_{t+1}^p}{m_{t+1}^c} \varepsilon_{t+1} &= \sigma^c M_t^c \left(\mathbb{E}_t \varsigma_{t+1}^k n_{t+1}^p \mu_{t+1} - \mathbb{E}_t \varsigma_{t+1}^k n_{t+1}^p \mathbb{E}_t m_{t+1}^c \mu_{t+1} \right) \\
&+ \mathbb{E}_t \frac{\varsigma_{t+1}^k n_{t+1}^p}{m_{t+1}^c} \varepsilon_t \\
&= \sigma^c \Phi_t \mathbb{E}_t n_{t+1}^p \varsigma_{t+1}^k \mathbb{E}_t [1 - m_{t+1}^c] u_{c,t+1} b_{t+1} \\
&+ \sigma^c \Phi_t [cov(\varsigma^k n^p, bu_c) - cov(m^c, bu_c)] \\
&+ [1 + cov(\varsigma^k, \frac{n^p}{m^c})] \varepsilon_t \\
&= \sigma^c \Phi_t \left[cov(\varsigma^k, bu_c) - cov(m^c, bu_c) \right] \\
&+ (1 + \varepsilon_t) cov(\varsigma^k, \frac{n^p}{m^c}) \geq \varepsilon_t.
\end{aligned}
$$

Notes

1 The Yale Climate Opinion Maps 2020 (Marlon et al. 2020) is encouraging in that 72 percent of respondents said global warming is real and a threat to humanity and the planet, with 57 percent believing it to be human caused.

2 Uncertainty in climate modeling has been intensively treated in the literature; see Pindyck (2007, 2013b), Roe and Baker (2007), Weitzman (2007, 2009, 2013), Heal and Millner (2013), and lately Barnett et al. (2020).

3 Tipping points include shutdown of the Atlantic Meridional Overturning Circulation, West Antarctic ice sheet disintegration, Amazon rainforest dieback, West African monsoon shift, permafrost and methane hydrates, coral reef die-off, Indian monsoon shift, Greenland ice sheet disintegration, boreal forest shift, and permafrost and methane hydrates. Dietz et al. (2021) note that such "climate tipping points are subject to considerable scientific uncertainty in relation to their size, probability, and how they interact with each other. Their economic impacts are even more uncertain, and consequently, these are often ignored or given a highly stylized treatment that fails to accurately represent geophysical dynamics and is nearly impossible to calibrate. As a result, tipping points are only weakly reflected in the policy advice economists give on climate change, typically by way of caveats and contextualization, rather than an integral part of the modeling that gives rise to estimates of the social cost of carbon (SCC) and other economic metrics of interest.

4 Barrage (2020) is perhaps the first to study optimal capital and carbon taxation in Ramsey planning framework.

5 This is not the first paper to dwell on a symmetry between carbon and capital taxation arising from a difference between social and private discount factors. Barrage's (2020) shows that the very logic leading to an optimal zero *ex ante* tax on capital derived in (Atkeson et al. 1999; Chari et al. 1994; Zhu 1992) implies a *positive* Pigou tax on carbon.

6 In a recent paper, Dietz and Niehörster (2019), estimate ambiguity loads—i.e., the extra insurance premiums due to ambiguity—and show how these depend on the insurer's attitude to ambiguity.

7 A treatment of the proper social discount rate for far-away and potentially catastrophic consequences of today's actions is beyond the scope of this paper. For discussions, see Weitzman (2009, 2013), Pindyck (2013a), and Traeger (2014).

8 In Belfiori's (2015) inter-generational model of altruism, the difference in discount factors is endogenous because each generation of households living in an infinite sequence of generations assigns a positive weight to the welfare of future generations, causing the government's discount factor to rise above that of households and leading to a carbon tax that exceeds the standard Pigou tax.

9 An axiomatization of multiplier preferences can be found in Strzalecki (2011).

10 Throughout, the expectations operator $\mathbb{E}_t y(x_{t+j})$ denotes the mathematical expectations of some function $y(x_{t+j})$, $j \geq 1$, with respect to the probability density $\pi(x_{t+j}|x^t)$: $\mathbb{E}_t y(x_{t+j}) = \sum_{x^{t+j}|x^t} y(x_{t+j})\pi(x_{t+j}|x^t)$.

11 This means one cannot assign positive probability to events as functions of x_t that have probability measure zero under the distribution of the approximating model or alternatively, the distorted and the approximating distributions are, at the very least, in agreement about which events have zero probability and which events are certain.

12 Previous systematic treatments of decision making under ambiguity include Klibanoff et al. (2005, 2009), and Traeger (2014).

13 Following examples by Anderson et al. (2013), Golosov et al. (2014), Nordhaus (2008), Li et al. (2016), and others, I omit leisure (or hours worked) to keep things simple. I also leave out any utility effects of environmental quality caused by climate change, since no conclusions relevant for this paper would be affected by their inclusion. See however, Barrage (2020).

14 See Hennlock (2009). While $\gamma < 1$ gets at early resolution of uncertainty, as noted by Bansal and Yaron (2004), the assumption of CRRA misses certain aspects of risk aversion that may create pricing issues.

15 For analytical convenience that will not affect conclusions, I will assume that energy production is costless, as is approximately true for oil (see Golosov et al. 2014). Accordingly, it is reasonable to fold the fossil energy extraction/production sector into the household, by making the consumer be the owner of the resource Q and the seller of energy E_t. This assumption differs from the literature (see Golosov et al. 2014; Nordhaus 1993), where, instead, the consumer owns the energy producing firm and receives its profits. However, ownership of the firm is merely a financial veil for direct ownership of the resource itself, hence a distinction without a difference.

16 The implied relative risk aversion is $CRR = 1 - \frac{\sigma^c}{1-\beta}$ that rises as $\sigma^c \to -\infty$. and $0 > \sigma^c > -\infty$. See Epstein and Zin (1991).

17 The formula for m^c can be derived as a special case of Epstein and Zin (1989) and Weil (1990) preferences

$$\mathcal{U}_t = [(1-\beta)u_t^{1-\rho} + \beta(\mathbb{E}_t\mathcal{U}_{t+1}^{1-\gamma})^{\frac{1-\rho}{1-\gamma}}]^{\frac{1}{1-\rho}},$$

where $\frac{1}{\rho}$ is the intertemporal elasticity of substitution. For $\gamma > 0$ and $\rho = 1$, this has the interpretation of a *risk-sensitive* recursion, such that in (6), $\sigma^c \equiv (1-\beta)(1-\gamma)$. With $\gamma = \rho$, preferences reduce to standard time-additive expected utility.

18 Readers will recognize Hansen and Sargent's (1995) discounted risk-sensitive recursion related to recursive preferences introduced earlier by Epstein and Zin (1989) and Weil (1990). Risk-sensitivity as a concept was introduced by Jacobson (1973) and later familiarized by Whittle (1981).

19 Excellent expositions of the physics of climate change for economists are Hsiang and Kopp (2018) and Traeger's (2018) description of the complex sets of channels involved in the transmission from carbon to temperature change.

20 Barnett et al. (2020) also adopt this measure of climate response.

21 Modeling damages resulting from climate change as negative total productivity shocks to the economy follows common practice, as for example in Golosov et al. (2014). The damaging effects of increased temperature on productivity in the world economy have been documented by Burke et al. (2015). Donadelli et al. (2017) used vector autoregressions to show that with a 50-year horizon, a one-standard deviation temperature shock lowers both cumulative output and labor productivity growth by 1.4

percentage points. Based on their model, they further show that temperature risk is associated with non-negligible welfare costs of 18.4% of the agent's lifetime utility that grow exponentially with the size of the impact of temperature on TFP.

22 Belfiori (2017) models climate damage as a reduction in household utility, while Barrage (2020) considers reductions in both production and utility.

23 For the derivation see Appendix A.

24 Henceforth, the n multiplier refers to the government's belief, and m to the consumer's belief distortion.

25 In this paper, I do not consider learning by either planner or private agents. An approach to doing so is available in Tetlow and von zur Muehlen (2009) who study robust monetary policy using *structured singular value* analysis when agents have misspecified models but are engaged in learning under the handicap that their learning of the reduced form of the economy is subject to potentially destabilizing parameter perturbations.

26 Hansen and Sargent's (2012) refer to this as *Type I ambiguity*.

27 The first application of risk-sensitive decision making to economic policy is van der Ploeg (1984). An early treatment of Knightian uncertainty in economics is von zur Muehlen (1982). In his analytical climate economy model, Traeger (2018) posits risk-sensitive preferences attributable to the planner representing the consumer.

28 Hansen and Sargent's (2012) refer to this as *Type III ambiguity*.

29 This belief regime is related to the robust fiscal policy model in Karantounias (2020). A fifth possible belief regime, wherein the planner has doubts but the private sector trusts the extant model, is not treated here having been widely discussed in the referenced literature.

30 In Table 1, p is defined by (19).

31 Current estimates of the social cost of carbon emissions are around USD75 per ton of carbon. Some consider this a gross underestimate and suggest the number is closer to USD220 per ton. (See Moore and Diaz (2015) and Than (2015)).

32 Initial wealth $\mathcal{W}_0$ is a function of k_0 and S_0 and the initial tax rates τ_0^k and τ_0^e. It is well known that since initial capital is supplied inelastically, the government has an incentive to raise the initial capital tax as high as possible. Likewise, with S_0 given, there is nothing (i.e., no welfare criterion) to prevent the planner from expropriating the energy sector by setting $\tau_0^e = 1$. As is conventional, I fix both τ_0^k and τ_0^e at 0.

33 By definition, next period's stock $Q_{t+1} = Q_t - E_t$ is given, i.e., determined by previous E_t, hence not influenced by x_{t+1}.

34 A positive warming shock reduces both consumption and productivity and therefore the value of future welfare $\mathcal{V}_{t+1}$.

35 Wilbanks et al. (2008) present empirical evidence that the net effect of declining consumption on wealth is negative.

36 Proofs are provided in Appendixes C–E, and utilize the preceding assumptions and lemmas, particularly the results that n^p, n^{POL}, n^{PAT}, m^p, ς^k, ς^e, and Λ, being positively correlated with x, are positively correlated with with each other.

37 Pascal argued that a rational person should live as though God exists and seek to believe in God. If God does not exist, such a person will have only a finite loss (some pleasures, luxury, etc.), whereas if God does exist, he stands to receive infinite gains (as represented by eternity in Heaven) and avoid infinite losses (an eternity in Hell).

38 This section is inspired by Karantounias (2013) who studied a fiscal authority's ability to manage pessimistic expectations.

39 See for example,Traeger (2018). Other features not addressed here include (i) *Arrhenius' Greenhouse Law* describing radiative forcing that connects carbon with temperature change, as described in Hassler et al. (2016), (ii) *tipping points* analyzed by Lemoine and Traeger (2016) and (Cai et al. 2013), describing abrupt nonlinear climate changes that pose a potentially existential threat to humanity in ways that may override concerns with belief and skepticism, and (iii) *polar amplification* analyzed by Brock and Xepapadeas (2017).

40 The derivations in this and the next two Appendices utilize these facts: Let $\{x, y, z\}$ represent four random, possibly related variables, and a a non-random variable or constant. Then

$$\begin{aligned} \mathbb{E}[ax \times y] &= a\mathbb{E}x\mathbb{E}y + acov(x,y) \\ \mathbb{E}[ax \times y \times z] &= a\mathbb{E}x\mathbb{E}y\mathbb{E}z + acov(x,yz) + a\mathbb{E}xcov(y,z). \end{aligned}$$

41 The derivation uses (77)

$$\begin{aligned} \mathbb{E}_t n^p_{t+1}\Lambda_{t+1}\varepsilon_{t+1} &= \sigma^c\Phi_t\mathbb{E}_t[n^p_{t+1}\Lambda_{t+1}u_{c_{t+1}}][\mathbb{E}_t m^c_{t+1}b_{t+1} - \mathbb{E}_t m^c_{t+1}\mathbb{E}_t m^c_{t+1}b_{t+1}] + E_t n^p_{t+1}\Lambda_{t+1}m^c_{t+1}\varepsilon_t \\ &+ \sigma^c\Phi_t[cov(n^p u_c\Lambda, m^c b) - cov(n^p u_c\Lambda, m^c)] \\ &= \sigma^c\Phi_t[cov(n^p u_c\Lambda, m^c b) - cov(n^p u_c\Lambda, m^c)] + [\Lambda_t^* + cov(n^p m^c, \Lambda]\varepsilon_t \\ &> \varepsilon_t. \end{aligned}$$

since, with the exception of marginal utility u_c, m^c, n^p, Λ, and b are positively related to x and therefore with each other.

42 By the law of iterated expectations $\mathbb{E}_t\mathbb{E}_{t+1}\cdots\mathbb{E}_{t+j-1}[\Xi^e_{t+j}] = \mathbb{E}_t[\Xi^e_{t+j}]$.

43 Where called for, the proofs use the result that for any martingale process m, $\mathbb{E}m = 1$, if

$$\Phi_{t+j} = m_{t+j}\Phi_{t+j-1},$$

then,

$$\mathbb{E}_t\Phi_{t+j} = \mathbb{E}_t m_{t+j}\Phi_{t+j-1} = \mathbb{E}_t \prod_{i=1}^{j} \frac{M_{t+i}}{M_{t+i-1}}\Phi_t = \mathbb{E}_t \frac{M_{t+j}}{M_t}\Phi_t = \frac{M_t}{M_t}\Phi_t = \Phi_t.$$

44 Constancy of Φ follows from (71).

45 In the following, m^s is purely random and independent of any other variable in the economy. Likewise, Φ_{t+j} is independent, where from footnote 42, $\mathbb{E}_t m^s_{t+j}\Phi_{t+j-1} = \Phi_t$, and $\mathbb{E}_t\varsigma^e_{t+j} = 1$.

46 In the following, I use the facts that $\mathbb{E}m = 1$ for each m, and also that $\mathbb{E}\varsigma^e = 1$. In addition, note that

$$\begin{aligned}
\mathbb{E}_t\Phi_{t+j} &= \mathbb{E}_t \frac{m^p_{t+j}}{n^p_{t+j}}\Phi_{t+j-1} = \mathbb{E}_t\left(\frac{m^p_{t+j}}{m^p_{t+j-1}} / \frac{N^p_{t+j}}{N^p_{t+j-1}}\right)\Phi_{t+j-1} = \mathbb{E}_t \prod_{i=1}^{j}\left(\frac{m^p_{t+i}}{m^c_{t+i-1}} / \frac{N^p_{t+i}}{N^p_{t+i-1}}\right)\Phi_t \\
&\geq \left(\frac{\mathbb{E}_t m^p_{t+j}}{\mathbb{E}_t N^p_{t+j}} / \frac{m^p_t}{N^p_t}\right)\Phi_t = \left(\frac{m^p_t}{m^p_t} / \frac{N^p_t}{N^p_t}\right)\Phi_t = \Phi_t,
\end{aligned}$$

since, by the law of iterated expectations, $\mathbb{E}_t m^p_{t+j} = m^p_t$, and likewise for N^p_t.

47 In the following, I use

$$\begin{aligned}
\mathbb{E}_t\Phi_{t+j} &= \mathbb{E}_t \frac{m^c_{t+j}}{n^p_{t+j}}\Phi_{t+j-1} = \mathbb{E}_t\left(\frac{m^c_{t+j}}{m^c_{t+j-1}} / \frac{N^p_{t+j}}{N^p_{t+j-1}}\right)\Phi_{t+j-1} = \mathbb{E}_t \prod_{i=1}^{j}\left(\frac{m^c_{t+i}}{m^c_{t+i-1}} / \frac{N^p_{t+i}}{N^p_{t+i-1}}\right)\Phi_t \\
&\geq \left(\frac{\mathbb{E}_t m^c_{t+j}}{\mathbb{E}_t N^p_{t+j}} / \frac{m^c_t}{N^p_t}\right)\Phi_t = \left(\frac{m^c_t}{m^c_t} / \frac{N^p_t}{N^p_t}\right)\Phi_t = \Phi_t,
\end{aligned}$$

since, by the law of iterated expectations, $\mathbb{E}_t m^c_{t+j} = m^c_t$, and likewise for N^p_t.

References

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012. The environment and directed technical change. *American Economic Review* 102: 131–66. [CrossRef] [PubMed]

Anderson, Evan W., William Brock, Lars Peter Hansen, and Alan H. Sanstad. 2013. Robust analytical and computational explorations of coupled economic-climate models with carbon-climate response. *SSRN Electronic Journal* 1–43. [CrossRef]

Atkeson, Andrew, Varadarajan V. Chari, and Patrick J. Kehoe. 1999. Taxing capital income: A bad idea. *Federal Reserve Bank of Minneapolis Quarterly Review* 23: 3–17. Available online: https://ideas.repec.org/a/fip/fedmqr/y1999isump3-17nv.23no.3.html (accessed on 28 September 2022).

Bansal, Ravi, and Amir Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59: 1481–509. [CrossRef]

Barnett, Michael, William Brock, and Lars Peter Hansen. 2020. Pricing uncertainty induced by climate change. *Review of Financial Studies* 33: 1024–66. [CrossRef]

Barrage, Lint. 2018. Be careful what you calibrate for: Social discounting in general equilibrium. *Journal of Public Economics* 160: 33–49. [CrossRef]

Barrage, Lint. 2020. Optimal dynamic carbon taxes in a climate-economy model with distortionary fiscal policy. *Review of Economic Studies* 87: 1–39. [CrossRef]

BEA. 2019. *Table 1.1.6. Real Gross Domestic Product, Chained Dollars*; Suitland: Bureau of Economic Analysis, U.S. Department of Commerce. Available online: https://apps.bea.gov/iTable/iTable.cfm? (accessed on 28 September 2022).

Belfiori, Maria Elisa. 2015. Time consistent climate policies. In *Political Economy And Instruments of Environmental Politics*. Edited by Friedrich Schneider, Andrea Kollmann and Johannes Reichl. Cambridge: MIT Press, Chapter 12.

Belfiori, Maria Elisa. 2017. Carbon trading, carbon taxes and social discounting. *European Economic Review* 96: 1–17. [CrossRef]

Belfiori, Maria Elisa. 2018. Climate change and intergenerational equity: Revisiting the uniform taxation principle on carbon energy inputs. *Energy Policy* 21: 292–99. [CrossRef]

Berger, Loïc, Johannes Emmerling, and Massimo Tavoni. 2016. Managing catastrophic climate risks under model uncertainty aversion. *Managemant Science* 57: 1–17. [CrossRef]

Brock, William A., and Anastasios Xepapadeas. 2017. Climate change policy under polar amplification. *European Economic Review* 94: 263–82. [CrossRef]

Brock, William A., and Steven N. Durlauf. 2015. On sturdy policy evaluation. *The Journal of Legal Studies* 44: 263–82. [CrossRef]

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel. 2015. Global non-linear effect of temperature on economic production. *Nature* 52: 235–39. [CrossRef] [PubMed]

Cai, Yongyang, Kenneth L. Judd, and Thomas S. Lontzek. 2013. *The Social Cost of Stochastic and Irreversible Climate Change*. Technical report, NBER Working Paper 18704. Cambridge: NBER. [CrossRef]

Cai, Yongyang, and Thomas S. Lontzek. 2019. The social cost of carbon with economic and climate risks. *Journal of Political Economy* 127: 2684–734. [CrossRef]

Chamley, Christophe. 1986. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica* 54: 607–22. [CrossRef]

Chari, V. V., Larence J. Christiano, and Patrick J. Kehoe. 1994. Optimal fiscal policy in a business cycle model. *The Journal of Political Economy* 102: 617–52. Available online: https://jstor.org/stable/2138759 (accessed on 28 September 2022). [CrossRef]

Collins, Matthew, Reto Knutti, Julie Arblaster, Jean-Louis Dufresne, Thierry Fichefet, Pierre Friedlingstein, Xuejie Gao, William J. Gutowski, Tim Johns, Gerhard Krinner, and et al. 2013. Long-term climate change: Projections, commitments and irreversibility. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by Thomas F. Stocker, Dahe Qin, Gian-Kasper Plattner, Melinda M. B. Tignor, Simon K. Allen, Judith Boschung, Alexander Nauels, Yu Xia, Vincent Bex and Pauline M. Midgley. Chapter Long-term Climate Change: Projections, Commitments and Irreversibility. Cambridge and New York: Cambridge University Press. Available online: https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter12_FINAL.pdf (accessed on 28 September 2022).

Dasgupta, Shouro, Nicole van Maanen, Franziska Piontek, Christian Otto, and Carl-Friedrich Schleussner. 2021. Effects of climate change on combined labour productivity and supply: An empirical, multi-model study. *The Lancet* 5: E455–65. [CrossRef]

Dietz, Simon, and Falk Niehörster. 2019. *Pricing Ambiguity in Catastrophe Risk Insurance*. Technical Report 354. Washington, DC: Resources for the Future.

Dietz, Simon, and Frank Venmans. 2019. Cumulative carbon emissions and economic policy: In search of general principles. *Journal of Environmnetal Economics and Management* 96: 108–29. [CrossRef]

Dietz, Simon, James Rising, Thomas Stoerk, and Gernot Wagner. 2021. Economic impacts of tipping points in the climate system. *Proceedings of the National Academy of Sciences of the United States of America* 118: e2103081118. [CrossRef] [PubMed]

Donadelli, Michael, Marcus Jüppner, Max Riedel, and Christian Schlag. 2017. Temperature shocks and welfare costs. *Journal of Economic Dynamics and Control* 82: 331–55. [CrossRef]

Emmerling, Johannes, Laurent Drouet, Lara Aleluia Reis, Michela Bevione, Loic Berger, Valentina Bosetti, Samuel Carrara, Enrica De Cian, Gauthier De MaereD'Aertrycke, Tom Longden, and et al. 2016. *Witch 2016 Model—Documentation and Implementation of the Shared Socioeconomic Pathways*. Technical Report 2016.42. Milan: Fondazione Eni Enrico Mattei. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2800970 (accessed on 28 September 2022).

EPA, US Government. 2017. Climate Impacts on Energy. Available online: https://19january2017snapshot.epa.gov/climate-impacts/climate-impacts-energy_.html (accessed on 28 September 2022)

Epstein, Larry G., and Stanley E. Zin. 1989. Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57: 937–69. [CrossRef]

Epstein, Larry G., and Stanley E. Zin. 1991. Substitution, risk aversion and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Ecnomy* 99: 263–86. [CrossRef]

Frankhouser, Samuael, and Richard S. J. Tol. 2005. On climate change and economic growth. *Resource and Energy Economics* 77: 1–17. [CrossRef]

Gazzotti, Paolo, Johannes Emmerling, Andrea Castelletti, Kaj-Ivar van der Wijst, Andries Hof, and Massimo Tavoni. 2021. Persistent inequality in economically optimal climate policies. *Nature Communications* 12: 3421. Available online: https://www.nature.com/articles/s41467-021-23613-y (accessed on 28 September 2022). [CrossRef]

Gilboa, Itzhak, and David Schmeidler. 1989. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18: 141–53. [CrossRef]

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski. 2014. Optimal taxes on fossil fuel in general equilibrium. *Econometrica* 82: 41–88. [CrossRef]

Hansen, Lars Peter, and Thomas J. Sargent. 1995. Discounted linear exponential quadratic gaussian control. *IEEE Transactions on Automatic Control* 40: 60–66. [CrossRef]

Hansen, Lars Peter, and Thomas J. Sargent. 2001. Robust control and model uncertainty. *American Economic Review* 91: 60–68. Available online: https://www.jstor.org/stable/2677734 (accessed on 28 September 2022). [CrossRef]

Hansen, Lars Peter, and Thomas J. Sargent. 2005. Robust estimation and control under commitment. *Journal of Economic Theory* 124: 258–301. [CrossRef]

Hansen, Lars Peter, and Thomas J. Sargent. 2007. Recursive robust estimation without commitment. *Journal of Economic Theory* 136: 1–27. [CrossRef]

Hansen, Lars Peter, and Thomas J. Sargent. 2008. *Robustness*. Princeton: Princeton University Press.

Hansen, Lars Peter, and Thomas J. Sargent. 2012. Three types of ambiguity. *Journal of Monetary Economics* 59: 422–45. [CrossRef]

Hassler, John, Per Krusell, and Anthony Smith. 2016. Environmental Macroeconomics. In *Handbook of Macroeconomics*. Edited by John B. Taylor and Harald Uhlig. Amsterdam: Elsevier, pp. 1893–2008.

Heal, Geoffrey, and Antony Millner. 2013. *Uncertainty and Decision in Climate Change Economics*. Technical Report 128. London: Centre for Climate Change Economics and Policy, University of Leeds. Available online: https://ideas.repec.org/p/lsg/lsgwps/wp108.html (accessed on 28 September 2022).

Hennlock, Magnus. 2009. *Robust Control in Global Warming Management: An Analytical Dynamic Integrated Assessment*. Technical Report 09-19. Washington, DC: Resources for the Future.

Hotelling, Harald. 1931. The economics of exhaustible resources. *Journal of Political Economy* 39: 137–75. Available online: https://www.jstor.org/stable/1822328 (accessed on 28 September 2022). [CrossRef]

Hsiang, Solomon, and Robert E. Kopp. 2018. An economists guide to climate change science. *Journalof Economic Perspectives* 32: 3–32. [CrossRef]

Jacobson, David H. 1973. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Transactions on Automatic Control* 18: 124–31. [CrossRef]

Judd, Kenneth L. 1985. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics* 28: 59–83. [CrossRef]

Karantounias, Anastasios. 2020. *Doubts about the Model and Optimal Policy*. Technical Report 2012-12. Atlanta: Federal Reserve Bank of Atlanta. [CrossRef]

Karantounias, Anastasios G. 2013. Managing pessimistic expectations and fiscal policy. *Theoretical Economics* 8: 193–231. [CrossRef]

Kaufmann, Robert K., Michael L. Mann, Sucharita Gopal, Jackie A. Liederman, Peter D. Howe, Felix Pretis, Xiaojing Tang, and Michelle Gilmore. 2017. Spatial heterogeneity of climate change as an experiential basis for skepticism. *Proceedings of the National Acadamy of Science of the United States of America* 114: 67–71. [CrossRef]

Kjellstrom, Tord. 2014. Productivity Losses Ignored in Economic Analysis of Climate Change. United Nations University. Available online: https://unu.edu/publications/articles/productivity-losses-ignored-in-economic-analysis-of-climate-change.html (accessed on 28 September 2022).

Kjellstrom, Tord, R. Sari Kovats, Simon J. Lloyd, Tom Holt, and Richard S. J. Tol. 2009. The direct impact of climate change on regional labor productivity. *Arch Environ Occup Health* 64: 217–27. [CrossRef]

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. 2005. A smooth model of decision making under ambiguity. *Econometrica* 73: 1849–92. [CrossRef]

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. 2009. Recursive smooth ambiguity preferences. *Journal of Economic Theory* 144: 930–76. [CrossRef]

Kullback, Solomon, and Richard A. Leibler. 1951. On information and sufficiency. *Annals of Mathematical Statistics* 22: 79–86. [CrossRef]

Kunreuther, Howard, Ralph Ginsberg, Louis Miller, Philip Sagi, Paul Slovic, Bradley Borkan, and Norman Katz. 1978. *Disaster Insurance Protection: Public Policy Lessons*. New York: Wiley.

Lemoine, Derek, and Christian P. Traeger. 2016. Ambiguous tipping points. *Journal of Economic Behavior and Organization* 132: 5–18. [CrossRef]

Lifton, Robert J. 2017. *The Climate Swerve*. New York: The New Press. Available online: https://thenewpress.com/books/climate-swerve (accessed on 28 September 2022).

Li, Xin, Borghan Narajabad, and Ted Temzelides. 2016. Robust dynamic energy use and climate change. *Quantitative Economics* 7: 821–57. [CrossRef]

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini. 2006. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica* 74: 1468–98. [CrossRef]

Maldonado, Franco, and Kevin P. Gallagher. 2022. *Climate Change and IMF Sustainability Analysis*. Technical report. Washington, DC: The Task Force on Climate, International Monetary Fund. Available online: https://www.bu.edu/gdp/files/2022/02/TF_WP_005_FIN.pdf (accessed on 28 September 2022).

Marlon, Jennifer, Peter Howe, Matto Mildenberger, Anthony Leiserowitz, and Xinran Wang. 2020. Yale climate opinion maps u.s. 2020. *Yale Program on Climate Change*. Available online: https://climatecommunication.yale.edu/visualizations-data/ycom-us/ (accessed on 28 September 2022).

Matthews, H. Damon, Susan Solomon, and Raymond Pierrehumbert. 2012. Cumulative carbon as a policy framework for achieving climate stabilization. *Philosophical Transactions of the Royal Society A* 370: 4365–79. [CrossRef]

Meyer, Robert, and Howard Kunreuther. 2017. *The Ostrich Paradox*. Philadelphia: Wharton Digital Press. Available online: https://wsp.wharton.upenn.edu/book/ostrich-paradox/ (accessed on 28 September 2022).

Millner, Anthony, Simon Dietz, and Geoffrey Heal. 2010. *Ambiguity and Climate Policy*. Technical Report 16050. Cambridge: National Bureau of Economic Research. Available online: https://www.nber.org/system/files/working_papers/w16050/w16050.pdf (accessed on 28 September 2022).

Millner, Antony, Geoffrey Heal, and Simon Dietz. 2012. Scientific ambiguity and climate policy. *Environmental and Resource Economics* 55: 21–46. [CrossRef]

Moore, Frances C., and Delavane B. Diaz. 2015. Temperature impacts on economic growth warrant stringent mitigation policy. *Nature Climate Change* 5: 127–31. [CrossRef]

Njuki, Eric, Boris E. Bravo-Ureta, and Víctor E. Cabrera. 2020. Climatic effects and total factor productivity: Econometric evidence for wisconsin dairy farms. *European Review of Agricultural Economicss* 47: 1276–301. [CrossRef]

Nordhaus, William D. 1993. Rolling the 'dice': An optimal transition path for controlling greenhouse gasses. *Resource and Energy Economics* 15: 27–50. [CrossRef]

Nordhaus, William D. 2007. *Estimates of the Social Cost of Carbon: Background and Results from the Rice-2011 Model*. Technical Report 1826. Cowles Foundation Discussion Paper. New Haven and London: Yale University Press. [CrossRef]

Nordhaus, William D. 2008. *A Question of Balance: Weighting the Options on Global Warming Policies*. New Haven and London: Yale University Press. Available online: https://www.jstor.org/stable/j.ctt1npzkh (accessed on 28 September 2022).

Pascal, Blaise. 1670. *Pensees*. London: W F Trotter.

Pindyck, Robert. 2007. Uncertainty in environmental economics. *Review of Environmental Economics and Policy* 1: 45–65. Available online: https://ssrn.com/abstract=949761 (accessed on 28 September 2022). [CrossRef]

Pindyck, Robert S. 2013a. Climate change policy: What do the models tell us? *Journal of Economic Literature* 51: 860–72. Available online: https://www.researchgate.net/publication/275930356_Climate_Change_Policy_What_Do_the_Models_Tell_Us (accessed on 28 September 2022). [CrossRef]

Pindyck, Robert S. 2013b. The climate policy dilemma. *Review of Environmental Economics and Policy* 7: 219–37. Available online: https://ideas.repec.org/a/oup/renvpo/v7y2013i2p219-237.html (accessed on 28 September 2022). [CrossRef]

Pindyck, Robert S. 2017. The use and misuse of models for climate policy. *Review of Environmental Economics and Policy* 11: 100–14. Available online: https://ssrn.com/abstract=2596426 (accessed on 28 September 2022). [CrossRef]

Rezai, Armon, and Rick van der Ploeg. 2017. Climate policies under climate model uncertainty: Max-min and min-max regret. *Energy Economics* 68: 4–16. [CrossRef]

Rezai, Armon, and Rick van der Ploeg. 2019. The agnostic's response to climate deniers: Price carbon! *European Economic Review* 111: 70–84. [CrossRef]

Roe, Gerard H., and Marcia B. Baker. 2007. Why is climate sensitivity so unpredictable? *Science* 318: 629–32. [CrossRef]

Somanathan, E., Rohini Somanathan, Anant Sudarshan, and Meenu Tewari. 2018. The Impact of Temperature on Productivity and Labor Supply: Evidence from Indian Manufacturing. Energy Policy Institute, University of Chicago. Available online: https://bfi.uchicago.edu/working-paper/the-impact-of-temperature-on-productivity-and-labor-supply-evidence-from-indian-manufacturing/ (accessed on 28 September 2022).

Stern, Lord Nicholas Stern. 2007. *The Economics of Climate Change: The Stern Review*. Cambridge: Cambridge University Press. Available online: https://www.lse.ac.uk/granthaminstitute/publication/the-economics-of-climate-change-the-stern-review/ (accessed on 28 September 2022).

Strzalecki, Tomasz. 2011. Axiomatic foundation of multiplier preferences. *Econometrica* 79: 47–43. [CrossRef]

Tetlow, Robert J., and Peter von zur Muehlen. 2009. Robustifying learnability. *Journal of Economic Dynamics and Control* 33: 296–316. [CrossRef]

Than, Ker. 2015. Estimated Social Cost of Climate Change Not Be Accurate, Stanford Scientists Say. *Stanford News*. Available online: http://news.stanford.edu/2015/01/12/emissions-social-costs-011215/ (accessed on 28 September 2022).

Traeger, Christian P. 2014. Why uncertainty matters—Discounting under intertemporal risk aversion and ambiguity. *Economic Theory* 56: 627–64. [CrossRef]

Traeger, Christian P. 2018. *Ace—Analytic Climate Economy (with Temperature and Uncertainty)*. Technical Report. Available online: https://ssrn.com/abstract=3307622 (accessed on 28 September 2022).

van der Ploeg, Frederick. 1984. Economic policy rules for risk-sensitive decision making. *Zeitschrift für Nationalökonomie* 44: 207–35. [CrossRef]

van der Ploeg, Frederick, and Cees Withagen. 2014. Growth, renewables, and the optimal carbon tax. *International Economic Review* 55: 283–311. [CrossRef]

von Below, David. 2012. Optimal Carbon Taxes with Social and Private Discounting. Unpublished manuscript. Available online: https://ethz.ch/content/dam/ethz/special-interest/mtec/cer-eth/resource-econ-dam/documents/research/sured/sured-2012/SURED-12_170_vonBelow.pdf (accessed on 28 September 2022).

von zur Muehlen, Peter. 1982. Activist vs. Non-Activist Monetary Policy: Optimal Rules under Extreme Uncertainty. Available online: https://ssrn.com/abstract=2735777 (accessed on 28 September 2022).

Weil, Philippe. 1990. Non-expected utility in macroeconomics. *Quarterly Journal Economics CV* 1: 29–42. [CrossRef]

Weitzman, Martin L. 2007. A review of the *Stern Review* on the economics of climate change. *Journal of Economic Literature* 45: 703–24. [CrossRef]

Weitzman, Martin L. 2009. On modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics* 91: 1–19. [CrossRef]

Weitzman, Martin L. 2013. Tail-hedge discounting and the social cost of carbon. *Journal of Economic Literature* 51: 873–82. [CrossRef]

Whittle, Peter. 1981. Risk sensitive linear quadratic gaussian control. *Advances in Applied Probability* 13: 764–77. [CrossRef]

Wilbanks, Thomas J., Vatsal Bhatt, Daniel E. Bilello, Stanley R. Bull, James Ekmann, William C. Horak, Y. Joe Huang, Mark D. Levine, Michael J. Sale, David K. Schmalzer, and et al. 2008. *Effects of Climate Change on Energy Production and Use in the United States*. Washington, DC: US Climate Change Assessment Program: Synthesis and Assessment Product 4.5, National Technology and Science Council, US Climate Change Science Program, US Department of Energy. Available online: https://digitalcommons.unl.edu/usdoepub/12/ (accessed on 28 September 2022).

Willner, Sven N., Nicole Glanemann, and Anders Levermann. 2021. Investment incentive reduced by climate damages can be restored by optimal policy. *Nature Communications* 12: 3245. [CrossRef]

Woodford, Michael. 2010. Robustly optimal monetary policy with near-rational expectations. *American Economic Review* 100: 274–303. [CrossRef]

Workman, Mark, Geoff Darch, Kate Dooley, Guy Lomax, James Maltby, and Hector Pollitt. 2021. Climate policy decision making in contexts of deep uncertainty - from optimisation to robustness. *Environmental Science and Policy* 120: 127–37. [CrossRef]

Zenios, Stavros A. 2022. The risks from climate change to sovereign debt. *Climatic Change* 172: 1. [CrossRef]

Zhang, Peng, Olivier Deschenes, Kyle C. Meng, and Junjie Zhang. 2017. *Temperature Effects on Productivity and Factor Reallocation: Evidence from a Half Million Chinese Manufacturing Plants*. Technical Report IZA DP No. 11132. Bonn: IZA Institute of Labor Economics. [CrossRef]

Zhu, Xiaodong. 1992. Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory* 58: 250–89. [CrossRef]

economies

Article

The Global Textile and Apparel Value Chain: From Mexico–US–China Linkages to a Global Approach

Óscar Rodil-Marzábal [1,*], Ana Laura Gómez Pérez [2] and Hugo Campos-Romero [1]

[1] ICEDE Research Group, Departamento de Economía Aplicada, Universidade de Santiago de Compostela, Av. do Burgo, s/n, 15782 Santiago de Compostela, A Coruña, Spain
[2] Instituto Politécnico Nacional (IPN), Escuela Superior de Economía (ESE), Av. Plan de Agua Prieta 66, Plutarco Elías Calles, Miguel Hidalgo, 11350 Ciudad de México, CDMX, Mexico
* Correspondence: oscar.rodil@usc.es

Abstract: The aim of this paper is to analyze the participation in the global textile and apparel value chain with special attention, first, to the case of three dynamic and interrelated economies (Mexico, the United States, and China); and second, to a general approach to a larger sample of countries through the analysis of trade in value added. From the descriptive analysis, a high domestic share in each country's exports is found. However, China is the leading exporter in the industry, accounting for around a third of the domestic value added in the global textile final demand. An econometric estimation has also been carried out to observe the effects of tariffs, FDI, and labor costs on the total and backward participation in the textile GVC. In this case, the sample has been extended to 39 developed and 22 developing countries. The results reveal tariff protection as a determinant of the degree of participation of the sector, especially when backward participation and developing economies are considered. However, FDI and labor costs only show the expected results in the case of developing countries. This may be due to the different tasks performed by developing economies (primarily manufacturing) versus developed economies (branding, design) within the sector's value chain.

Keywords: global value chain; textile and apparel; international trade; developed and developing countries; forward and backward participation; explanatory factors

Citation: Rodil-Marzábal, Óscar, Ana Laura Gómez Pérez, and Hugo Campos-Romero. 2022. The Global Textile and Apparel Value Chain: From Mexico–US–China Linkages to a Global Approach. *Economies* 10: 258. https://doi.org/10.3390/economies10100258

Academic Editor: Sajid Anwar

Received: 30 June 2022
Accepted: 12 October 2022
Published: 18 October 2022

1. Introduction

The textile and clothing industry is of great importance worldwide. In the case of countries such as Mexico or the United States, it is a relevant activity currently boosted by the advantages of the United States–Mexico–Canada Agreement (USMCA). The significant increase in the demand for textile products has opened the door to complex production processes that met international quality standards and paved the way to produce complete packages. This stimulates the formation of linkages between members of the production chain and provides opportunities for domestic producers. This process has led to the emergence of Global Value Chains (GVCs), a complex phenomenon reflecting the importance of global production linkages for access to new technologies, training, and innovation (Morrison et al. 2008).

The concept of GVCs provides a better understanding of the value creation process and helps understand how this value is captured, held, and leveraged in all industries. The GVC approach offers a global view of the World's industries from two perspectives: governance and upgrading. The former focuses mainly on leading companies and how their supply chains are organized on a global scale. At the same time, the latter involves the strategies that countries, regions, companies, and other actors use to maintain or improve their positions in the global value chain (Gereffi and Lee 2016). From this perspective, the case of the textile and apparel industry can be seen as a clear example of the strategic use of GVCs in a competitive and dynamic business world.

The theoretical definition of GVCs covers the full range of activities required to bring a good or service to the final consumer, from the acquisition of raw materials to delivery to the final consumer (Antrás 2020; Del Prete et al. 2017; Rodil 2017). In this sense, in the context of international fragmentation and the dynamism of production in the textile sector, labor seems to be a crucial factor, especially in manufacturing tasks. However, the progressive cheapening of products means that no country can forever maintain its comparative advantage in producing labor-intensive garments as its economy industrializes and advances (Lu 2018). Furthermore, these GVCs have expanded due to liberalization, the rise of ICTs, and lower transport costs. This has allowed the management of multiple geographically dispersed tasks in a value chain (Baldwin 2016). Thus, the GVC concept covers all value chain stages following an Input–Output structure. It is also defined by a governance structure, which refers to the power relations between the participating firms, and an institutional context, which refers to the local, national, and international political conditions that affect the different stages of the value chain (Gereffi and Fernandez-Stark 2016).

For decades, developing countries have imported parts and components from countries with more advanced technology, although usually only for the assembly of goods sold locally, forming part of a global network (Taglioni and Winkler 2014). However, several developing countries have managed to move up the chain to more advanced and higher value-added tasks (Pahl and Timmer 2020).

Trade in the supply chain is determined by international differences in production and unbundling costs, while technology determines how the different stages of production are linked (Amador and Cabral 2014). For example, a key part of China's success that has allowed it to achieve economies of scale and scope in GVCs is the constant interaction with various nations for the acquisition of inputs and technology to reduce production costs (Gereffi 2019). Thus, GVCs for developing countries are a fast path to industrialization, as internationally fragmented production allows them to join existing supply chains instead of building them, by sophisticating their goods and expanding their product range (Raei et al. 2019).

An essential factor for insertion in GVCs is industrial competitiveness, which is increasingly defined by international production networks (fragmented and spatially dispersed) and less by national borders (Ponte et al. 2019). In this sense, FDI also plays a central role, representing an opportunity for insertion in GVCs for developing countries. However, according to the WTO (2014), not all countries succeed in joining GVCs. Only those whose production is close enough to the global standards of quality and efficiency succeed. Knowledge and technology transfers, usually fostered by FDI and trade openness, tend to trigger the initial integration.

As a key global player, China has shown a trend as the World's leading exporter of manufactured goods and the largest importer of many raw materials, contributing to its status as an important country in the GVCs (Gereffi 2019). Moreover, the increase in Chinese trade in GVCs has been associated with significant changes in wages and employment in China's trading-partner countries (Robertson et al. 2020). Therefore, the dynamics of GVCs depend on the direction of current trade flows (Durand and Milberg 2020). Regardless of the specific type of GVC, the fragmentation of production results in a greater international division of labor and higher specialization gains exploited by the textile industry (Antrás 2020).

Traditionally, the textile sector has been seen as an ideal way for developing countries to enter GVCs. Although markets have become more complex and competitive, the work done by Whitfield et al. (2021) shows that it is still possible to promote industrialization through trade in textiles. This is due to the potential of this activity to generate intra-sectoral networks and generate industrial upgrading trajectories, initially based on a labor cost advantage. Moreover, successful upgrading processes can lead to greater resilience of companies to external shocks (Choksy et al. 2022).

The process of value creation in different countries generates a comparative advantage and a new division of labor, produces new sources for the flow of trade, and increases the level of innovation during the production process, where the main sources of value added are the industries. Therefore, according to Rodil (2017), measuring trade in value added is a fundamental tool for analyzing international trade in this fragmented context. This methodology is based on the decomposition of gross trade into value-added flows that capture the way and intensity in which international productive fragmentation affects the participating countries. Likewise for Banga (2014), domestic and foreign value added is created during manufacturing, so value-added exports will differ from gross exports and can be estimated by subtracting foreign value added.

Value-added trade is a series of measures that provide a better understanding of production networks and supply chains through statistical data. Thus, for this measure of trade, several indicators assess the participation of countries within the GVC: the backward participation index, which indicates the share of foreign value added as a percentage of gross exports; the forward participation index, which indicates the share of domestic value added embodied in foreign exports as a percentage of gross exports; and the total participation index, which is the sum of former.

This paper aims to analyze the participation of countries in textile and apparel GVC with special attention, first, to the case of three dynamic and interrelated economies (Mexico, the United States, and China); and second, extending the analysis to a larger sample of countries in the textile and apparel GVC through trade in value-added approach. The first part focuses on the changing role of the three selected economies, on their performance as value-added suppliers of the final global demand for textile products, and, especially, on verifying the rise of Chinese leadership in this global industry. Meanwhile, the second part includes an econometric analysis with panel data (61 countries, 24 years: 1995–2018) of some relevant factors explaining this GVC participation.

The main source of data is the TiVA database (December 2021 edition) provided by the Organization for Economic Co-operation and Development (OECD), which provides information on trade in value-added for 66 economies and 45 industrial sectors, covering the period of 1995–2018. Such information can be used, among others, to analyze the integration of economies into GVCs, as well as the country of origin of the value-added embodied in gross trade flows and final demand. Other databases used are UNCTAD for data on FDI flows, and the WTO for data on textile tariff rates.

The remainder of the paper is structured as follows. Section 2 describes the paper's methodology, highlighting the usefulness of trade in value-added approach for analyzing country participation in GVCs. Section 3 presents and discusses the empirical results, analyzing the participation of Mexico, the United States, and China in GVC from a general (all industries) and sectoral (textiles and apparel) perspective. It also analyzes the contribution of these countries as value-added suppliers to the World's final demand for textile products and adopts an extended econometric analysis with panel data (61 countries, 24 years: 1995–2018) to explore relevant factors explaining the participation of countries in this GVC. Finally, Section 4 presents the conclusions of the paper.

2. Data and Methodology

The empirical study of GVC participation has a growing number of works analyzing the role played by various explanatory factors (among others, Rahman and Zhao 2013; Arrighetti et al. 2014; Stehrer and Stöllinger 2015; Kowalski et al. 2015; Jona-Lasinio et al. 2016; Vrh 2018). However, the analysis of GVCs from a macroeconomic perspective usually follows the work of Koopman et al. (2014). Their methodology decomposes a country's gross exports into nine components of trade, providing several indicators. These include forward (export-linked) and backward (import-linked) participation indices, the sum of which is considered an indicator of total GVC participation (see Appendix A for the corresponding OECD TiVA indicators). This methodology allows for the tracing of each country's value-added flows to its final consumption destination.

The local supply of intermediate products is one of the main direct export channels attracting FDI, and specialization in the early stages is associated with the production of local inputs obtained by foreign investors (Amendolagine et al. 2017). Hence, one aspect to be considered as a possible explanatory factor for participation in GVCs is the degree of tariff protection, as this factor acts as a barrier to trade flows, among which trade in intermediate products associated with the GVC linkages is becoming increasingly essential. Thus, it is interesting to verify if there is a negative relationship between the level of tariff protection and participation in GVC.

As Yi (2003) points out, vertical specialization may have enhanced the reduction in tariff rates. Through this strategy, characteristic of GVCs, countries specialize in certain stages of a product's value chain. As a result, a slight reduction in tariff rates has multiple multiplier effects on trade growth. Conversely, increasing tariff rates can reduce trade in GVCs as parts and components pass multiple times across different national borders (OECD 2013).

Among the explanatory factors of GVC participation, FDI stands out as a determining element when analyzing the insertion of countries in the framework of international productive fragmentation. In this regard, various studies (Stehrer and Stöllinger 2015; Kowalski et al. 2015) point to a positive relationship between inward FDI stock and participation in GVC. However, no conclusive results can be found in the literature on the role played by outward FDI stock. Therefore, studying the relationship between FDI and GVC participation is interesting. In general, it is assumed that there is a positive relationship between them. This hypothesis is based on the role of multinational companies as major actors in GVCs.

Another explanatory factor of GVC participation is the labor cost, since labor has traditionally been a critical factor, especially in manufacturing or assembly tasks, usually offshored to developing countries. However, the progressive cheapening of global products has led to an unstable competitive framework (Lu 2018), and the explanatory relevance of this factor may sometimes be unclear. Hence, it is also interesting to analyze the influence of labor costs on countries' participation in GVCs.

Based on these assumptions, an econometric model has been estimated using panel data. This empirical analysis considers a group of 61 countries at different development levels, observed for 24 years (1995–2018). The division of the 61 countries into two development groups is based on the World Bank's most recent criteria (2021–2022). Countries classified as "high income" have been considered developed countries. All other cases have been included in the group of developing countries. This division divides the sample into two groups of 39 and 22 countries, respectively (see Appendix B). The general model to be estimated is specified as follows:

$$\gamma_{it} = \beta_0 + \beta_1 TARIF_{it} + \beta_2 FDI_{it} + \beta_3 LABC_{it} + \varepsilon_{it} \tag{1}$$

where i refers to the country and t refers to the period. Two dependent variables have been considered for estimation: total participation in GVCs (TPART), expressed as a percentage of gross exports, and backward participation (BPART), also expressed as a percentage of gross exports.

A total of four independent variables have been selected. The two first regressors are TARIF1 and TARIF2, which refer to the average tariff on textile raw materials and the main textile products, respectively. TARIF1 refers to 51 and 52 textile raw material groups and TARIF2 refers to 61 and 62 textile product groups, according to HS classification. Due to multicollinearity problems between both variables, two different models are considered: Model I, including only TARIF1 as the tariff variable, and Model II, including only TARIF2.

The other two independent variables considered are FDI, which refers to the inward foreign direct investment stock, expressed as a percentage of GDP, and LABC, which is the labor cost, expressed as a percentage of value-added. Except for FDI, all variables refer to the textile sector (ISIC Rev.4 codes 13, 14, and 15). FDI is obtained from UNCTAD, labor cost and GVC share variables are obtained from TiVA (OECD 2021), and TARIF data is obtained from WTO.

Therefore, the two considered models are as follows, where the expression relating to the dependent variable (PART) is a generic expression of the GVC participation, which can refer indistinctly to total participation (TPART) or backward participation (BPART):

$$\begin{array}{l} \text{Model I}: \ PART_{it} = \beta_0 + \beta_1 TARIF1_{it} + \beta_2 FDI_{it} + \beta_3 LABC_{it} + \varepsilon_{it} \\ \text{Model II}: \ PART_{it} = \beta_0 + \beta_1 TARIF2_{it} + \beta_2 FDI_{it} + \beta_3 LABC_{it} + \varepsilon_{it} \end{array} \quad (2)$$

The consideration of backward participation in GVCs (BPART) is due to its relevance for most developing economies, which, in the textile sector, tend to take on manufacturing tasks of lower value added, relative to other tasks, such as garment design and conception.

The reason for using panel data is motivated by the suspicion that participation in GVC is influenced by unobservable factors that correlate with observed variables, such as the factors mentioned above. Therefore, it is assumed that the panel techniques contribute to obtaining consistent estimates of the effect of the variables observed, offering greater possibilities at the time of facing the usual problems in this type of empirical approach.

The joint significance of differing group means and Breusch-Pagan statistic tests point to a panel data structure. The Hausman statistic test points to a fixed effects model. One of the immediate implications of this is that the error term ε_{it}, in Equations (1) and (2), is now broken down into two different effects: a specific country effect (m_i) and the remaining error (v_{it}). A relevant advantage of this econometric technique is that it allows us to obtain unbiased estimators.

3. Results

3.1. The Participation of Mexico, the United States, and China in GVC: General Perspective

International trade allows economies to integrate and increase their participation in GVC trade flows, so that activities along a value chain can be carried out by FDI or outsourcing (Kowalski et al. 2015). For example, in apparel, China has been the most dynamic exporter worldwide in clothing since 2001, when it joined the World Trade Organization (WTO).

It adopted a position that has not favored Mexico, due to the displacement of U.S. imports from that country with such a growth that they surpassed Mexican imports (Pino 2020). As a result, in 2003, Mexico dropped from first to second place as a textile and clothing supplier because of its dependence on the economic ups and downs of the United States (Rodríguez and Fernández 2006).

Trade through GVCs offers opportunities to developing countries, especially smaller ones, for global integration, changing the nature of competitiveness (Pathikonda and Farole 2017). This is because much of the labor-intensive production moved to the World developing in the last wave of globalization, with textiles being highly tradable products (Lund et al. 2019).

As can be seen in this paper, the analysis of the participation of Mexico, the United States, and China in the global textile and apparel value chain presents relevant changes in the last decades. Lu (2013) points out that one of the reasons for these changes is that a country's apparel industry gradually upgrades following the path of Cut, Make and Trim (CMT), Original Equipment Manufacturing (OEM), Original Design Manufacturing (ODM), or Original Brand Manufacturing (OBM). In the case of Mexico, for example, the textile industry has been transforming by assuming mainly assembly tasks (e.g., cutting and sewing) and abandoning a series of risk- and knowledge-intensive coordination and design tasks (Pipkin and Fuentes 2017).

The comparative analysis of the participation of Mexico, the United States, and China in the GVC in the period of 1995–2018 shows different participation levels, insertion patterns, and trends. In this sense, Figure 1 provides a clear picture of the different patterns observed. The first difference refers to the total participation rate, with high participation in Mexico in 2018 (46.4% of gross exports), compared to China (36.6%) and the United States (35.6%) but at the same time with a strong predominance of backward linkages in Mexico (35.9% of gross exports), compared to China (17.2%) and the United States (9.5%).

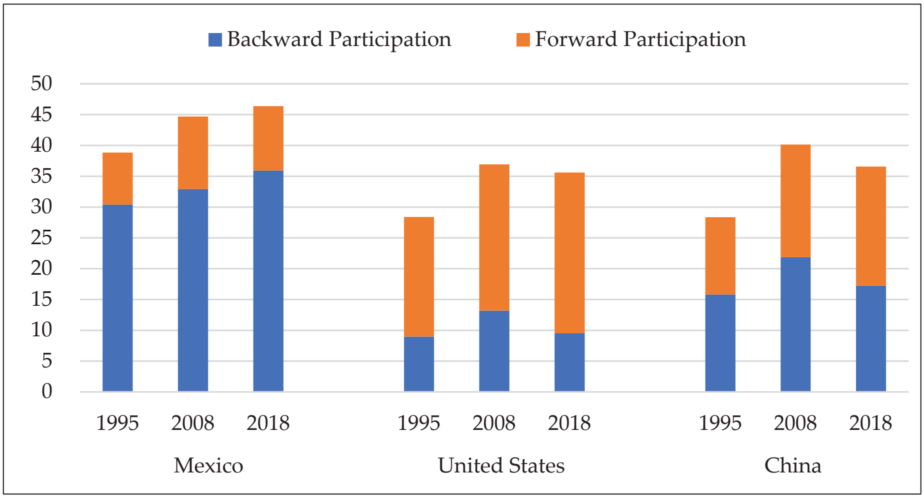

Figure 1. GVC participation index (%). Total participation (all sectors) 1995–2018. Source: Authors based on TiVA (OECD 2021).

However, this gap in the level of total participation differs significantly from that observed more than a decade earlier (2008), when Mexico started from a higher level (44.7%) than China (40.2%) and the United States (36.9%). Furthermore, another difference is given by the opposite trends observed in GVC participation; that is, Mexico's participation increased by more than seven percentage points during the study period, while the U.S. and China decreased their participation in GVC between 2008 and 2018 (even though their participation rates are higher in 2018 than in 1995).

The observation of the predominant type of production linkage is fundamental since this analysis is given by decomposing the total participation in its two components: backward and forward participation. Thus, the predominance of China's forward participation (12.6% of gross exports) in 1995 increased by more than six percentage points by 2018 (19.3%); in comparison, Mexico increases by two percentage points from 1995 to 2018 (from 8.5% to 10.5%) while the U.S. shows an increase of more than six percentage points (from 19.4% to 26.1%).

China's rapid growth has made it a major player in virtually all goods produced in GVCs, accounting for 20% of global gross output (Lund et al. 2019), which was initially due to cheap Chinese labor due to low wages (Gereffi and Memedovic 2003). In this sense, one of the causal factors contributing to the reduction in costs and the increase in production rates has been the supply of cheap Chinese labor, which brings low wages (Gereffi and Memedovic 2003).

However, the trend observed for Mexico reveals that its foreign trade operates more as a carrier of value added originating in other countries than as a channeler of domestic value added to later stages of production in the framework of international fragmentation of production. This is, to some extent, a direct consequence of China's productive strategy of gradually substituting foreign value added for domestic value added (Rodil 2017).

Therefore, the reduction in the intensity of participation in GVCs is due to the deepening of the domestic division of labor and the lengthening of domestic value chains (Li et al. 2019). In this sense, the GVC participation trends in Mexico, the United States, and China offer an interesting perspective on their behavior in the GVCs of developed and developing countries, highlighting the case of Mexico's backward linkages that are characteristic of a manufacturing country.

This predominance of backward participation could be associated with countries' participation in production stages close to the final demand, for example, in the case of assembly tasks (assembly line). However, it is also important to note that countries can show high rates of backward participation by doing non-manufacturing activities that generate high value-added, related to marketing and distribution (Rodil and Gómez 2021).

3.2. Participation of Mexico, the United States, and China in GVC of Textiles and Apparel: Sectoral Perspective

3.2.1. GVC Participation in the Textile and Apparel Industry

The analysis of the GVC participation at a sectoral level, focusing on the textile and apparel industry (Figure 2), provides a different picture from the general perspective (Figure 1). Firstly, the level of participation is higher when textile GVC is considered. Secondly, the trends at the sectoral level change for Mexico and China, with decreasing and increasing levels of participation, respectively, from 2008 to 2018. Thirdly, the sectoral approach shows a clearer pattern of participation by country, with a deeper forward GVC participation for China, the United States, and Mexico confirming the predominance of backward insertion.

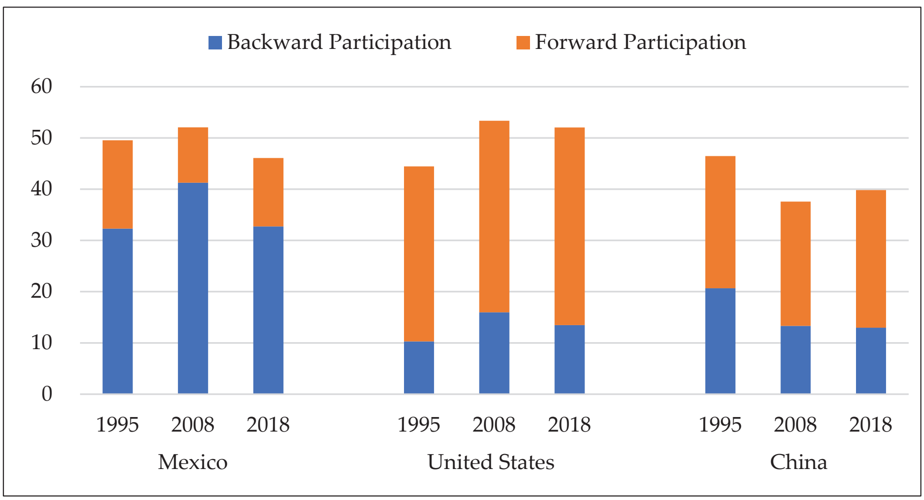

Figure 2. GVC participation index (%). Sectoral participation (textiles and apparel) 1995–2018. Source: Authors based on TiVA (OECD 2021).

One of the reasons for the dispute over the North American market between Mexico and China is due to the competition generated by U.S. manufacturing exports. There are two reasons for this rivalry. On the one hand, Mexico is strategically close to the U.S., geographically. On the other hand, China has the advantage of scale, being the world's largest exporter of manufactured goods, especially consumer goods such as textiles and electronics (Gereffi and Luo 2015).

In 2008, China's share of world textile exports was already 26.1%, while the United States and Mexico accounted for only 5% and 0.8%, respectively (Gracia-Hernández 2011). With a national strategy of considering the Fiber–Textile–Clothing Chain (CFTV) as one of the 12 priority branches of the textile industry, Mexico increased its export levels in cotton textile fibers from 7.2% in 2001 to 21.8% in 2010, showing its competitive potential (Vázquez et al. 2015).

Moreover, Mexico's privileged position of sharing more than 3000 km of border with the United States and the existence of free trade agreements account for its progressive trade liberalization, which results in the existence of cheap labor, more than that of all countries, except Asia (Montón 2015).

These aspects provide, in a way, a guarantee for Mexico to position itself in the U.S. market as a continuous development of its insertion in the more sophisticated GVCs. However, its direct participation in the fiber (yarn) and apparel (garment) links has unleashed a continuous cause of tension between China and Mexico (Chen and Goodman 2018). In this sense, some authors (Chen and Goodman 2018) propose that China and Mexico should develop a strategic partnership focused on cooperation by actively seeking business opportunities between them and being more understanding rather than showing their competitiveness with each other.

Therefore, it is interesting to analyze the textile sector's participation in the GVC through the study of value-added trade. The purpose of this analysis is to verify the main research question underlying this paper: is China consolidating its leadership as the leading supplier of value added to the global textile and apparel value chain developing an increasingly important role as a value creator in this global industry? To address this question, changes affecting the origin of value added embodied in textile exports from Mexico, the United States, and China are analyzed below.

3.2.2. Origin of the Value Added of Mexico's Textile and Apparel Exports

Although the origin of the value added embodied in Mexico's textile exports is mostly domestic, it is worth noting that 33% of the total value added exported in 2018 came from abroad, with the United States (14.4%) and China (6.6%) standing out in this share of value added (Figure 3).

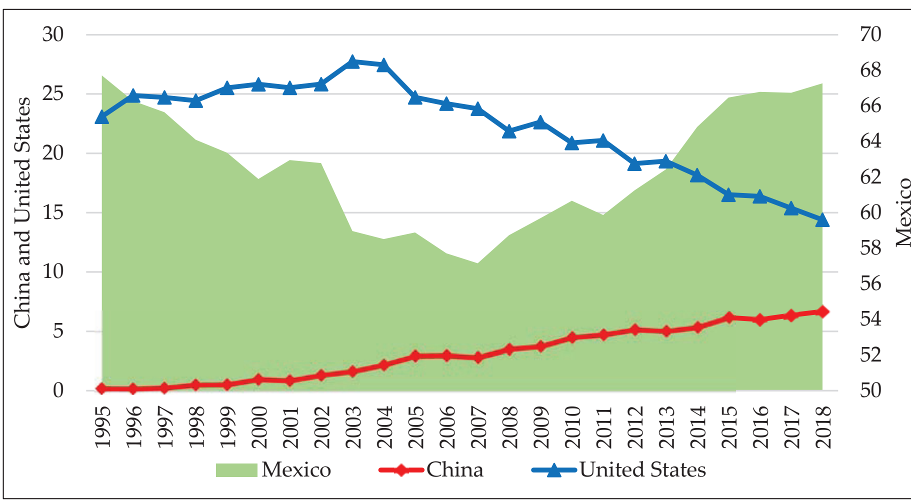

Figure 3. Origin of value added in Mexico's textile and apparel gross exports (%). 1995–2018. Source: Authors based on TiVA (OECD 2021).

In any case, what is most striking are the contrasting trends observed in the participation of the United States and China during the 1995–2018 period. In this regard, the U.S. share decreased by more than eight percentage points, while China's share increased by more than 6 percentage points during the same period, more than doubling its initial share. This increase in added value could imply a faster upgrading of activities performed in GVCs and the deepening of intra-product specialization brought about by the recovery of cross-country, production-sharing activities (Li et al. 2019).

Likewise, this increase could also be due to the increased identification of GVC conditions ranging from sourcing cheap labor inputs and basic assembly activities with cheap and unskilled labor, to more advanced forms of value production, such as the full package strategy (Fernández and Gereffi 2019).

The importance of the U.S., in terms of value-added incorporated in Mexican textile exports, stems from the territorial proximity and the productive historical interconnection forged between the two economies over time. A significant proportion of the manufactured products exported by Mexico are made up of dynamic products for world trade, such as textiles and clothing (Fuji et al. 2005).

The implication of the Mexican textile industry in GVCs has been attributed to high transportation costs that fragmented the domestic market and generated a geographically dispersed industry (Gómez-Galvarriato 1999). Likewise, the growth of garment production in China affected world markets by offering the possibility of a greater quantity of textile products being produced through a fragmented process (Robertson et al. 2020).

However, the textile industry in Mexico has obeyed the geographical proximity of the United States, which translates into lower transportation costs, as well as the ease of supplying foreign plants with machinery, components, and materials in general, in addition to the fact that specialized labor represents greater agility when required in Mexico (Hansen 2020).

Therefore, some of the factors that potentially explain the increase in China's share as a source of value-added embodied in Mexican textile exports are: (1) China's lower labor costs compared to Mexico; (2) China's exploitation of economies of scale through investments in infrastructure and transportation logistics that accelerate the commercialization of its exported products (Gereffi and Luo 2015); (3) China's coherent and multidimensional scaling strategy for diversifying its industrial composition which adds high value-added activities (Frederick and Gereffi 2011); and (4) the use of FDI to promote continuous learning in industries as well as leveraging domestic market knowledge.

3.2.3. Origin of the Value Added of the United States Textile and Apparel Exports

U.S. textile and apparel exports show a somewhat irregular variation from 1995 to 2018, decreasing almost three percentage points (from 89.7% to 86.5%). The participation of foreign companies, the introduction of new technologies, the continuous training of employees, and the growing competitiveness and innovation are causes that did not foresee an ideal integration of the textile-garment production (or supply) chain. This lets the full package strategy be carried out through selective alliances with leading Mexican companies.

Therefore, the small- and medium-sized companies that make up most of the Mexico sector have been affected. As a result, Mexico's participation in international trade has decreased (Arroyo and Cárcamo 2010). For example, Figure 4 shows a half percentage point share of Mexico in 1995 in the value added embodied in U.S. textile and apparel exports, which barely increased to 0.76% in 2018.

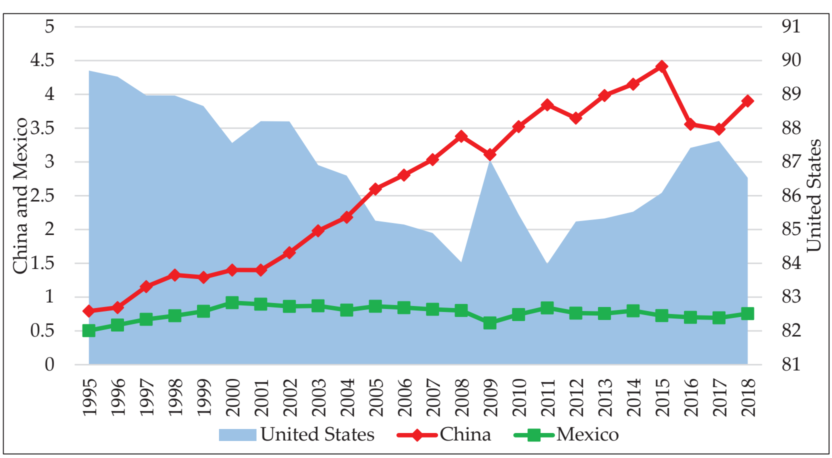

Figure 4. Origin of value added in the United States' textile and apparel gross exports (%). 1995–2018. Source: Authors based on TiVA (OECD 2021).

As for China's share, it shows an increase of more than three percentage points from 1995 to 2018 (0.79% to 3.90%), generating a considerable increase in the level of Chinese world exports through its share of the World export market in less than four decades (Gómez Chiñas and García 2017).

Moreover, China's upward trend contrasts both with the fall of domestic (U.S.) value added in the period of 1995–2018, as well as Mexico. This trend has caused the relative importance of China as the source of value-added embodied in U.S. textile exports to increase almost fivefold in 1995.

3.2.4. Origin of the Value Added of China's Textile and Apparel Exports

The origin of the added value of China's textile exports is due to the efficiency of the full package strategy in Asia, as it creates competitive advantages that highlight the manufacturing of textile products. This generates around 50% of the final product costs which makes Mexico's participation less than one percentage point and almost constant from 1995 to 2018, as seen in Figure 5 (Castro-Gonzáles and Mathews 2013).

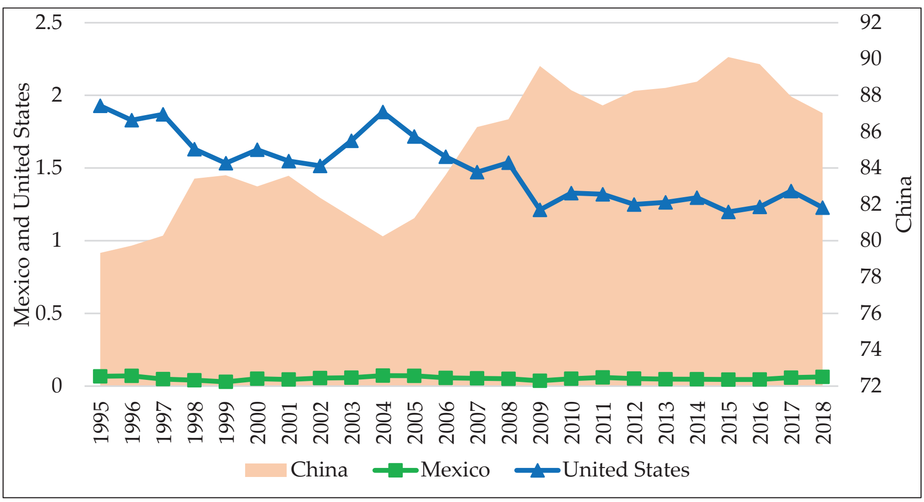

Figure 5. Origin of value added in China's textile and apparel gross exports (%). 1995–2018. Source: Authors based on TiVA (OECD 2021).

This behavior contrasts with the level and evolution of the domestic (Chinese) value-added share of Chinese textile exports, as it has a steady growth from 1995 to 1999 (79.3–83.6%) and increased just over four percentage points from 2000 to 2018 (82.9–87.0%).

Therefore, the United States has a descending trend in the origin of the added value of Chinese exports from 1995 to 2018. This predominance of China is not only due to the strategy of the complete package but also to the use of the strategy of devaluation of the currency against the dollar that has a positive effect on its textile exports whose purpose is to primarily affect the participation of the United States (Castro-Gonzáles and Mathews 2013).

However, there may be an increase in production costs in China, which would cause garment producers to have opportunities abroad to relocate their plants (Inomata and Taglioni 2019). Nevertheless, China has built a strong domestic market with a complete and independent manufacturing system, and through "Made in China" has been active in the global market internationalizing and inserting themselves into global supply chains (Cohen and Lee 2020; Ma et al. 2018).

3.3. The Contribution of Mexico, the United States, and China to the Final Global Demand for Textiles and Apparel

The previous section analyzed the evolution of the relative importance of Mexico, the United States, and China as the origin of the value added incorporated in their respective gross exports of textiles. However, for a better understanding of its implications, it is also interesting to study their importance in terms of their contribution to the final global demand for textile products.

This analysis shows the greater importance of China with a rising trend between 1995 and 2015 (from 6.2% to 38.3%), declining since then and contrasting notably with the declining share of the United States and Mexico, both with a share below 5% of global demand in 2018 (Figure 6).

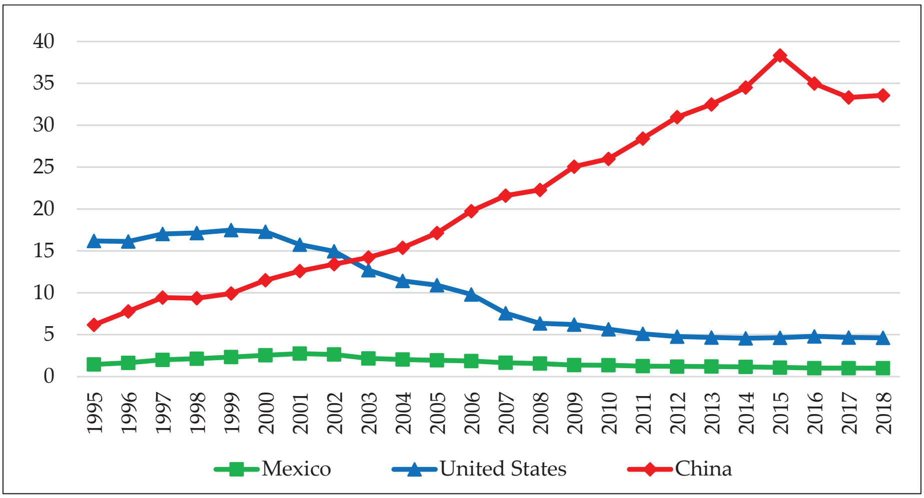

Figure 6. Participation of China, Mexico, and the United States as origin of the global final demand value added in the textile and apparel industry (%). 1995–2018 Source: Authors based on TiVA (OECD 2021).

Moreover, the comparison between the share in final global demand and the share in world exports shows opposite trends. The Mexican and the U.S. value-added loses relative importance, while value-added originating in China is becoming increasingly relevant, both at the export level and from the final global demand for textile products.

These results are in line with the general hypothesis that there is a process of consolidation of China's hegemony in the global textile market to the detriment of the participation of the United States and Mexico. Furthermore, these findings show that this process is being mainly conducted through the flows of value added in GVCs.

3.4. Econometric Results

This section presents the econometric estimations, first, for the U.S., Mexico, and China, and second, for the global case. These analyses aim to explore the influence of variables such as FDI, textile tariffs, and textile wages on GVC participation. Concerning participation, as indicated in the methodological section, a distinction is made between total and backward participation in GVCs.

As discussed in Section 2, two econometric models have been estimated to capture some of the determinants of total participation (TPART) and backward participation (BPART). The independent variables selected are the tariff rates applied to the textile sector, where TARIF1 refers to textile raw materials (model I) and TARIF2 refers to textile products (model II). The ratio of inward FDI stock to GDP (FDI) and labor costs to value added in the textile industry (LABC). The main descriptive statistics for these variables are shown in Table 1.

Table 1. Main statistics of the variables included in the model.

Variable	Mean	Median	S.D.	Min.	Max.
TPART	50.5	50.4	9.76	25.9	81.9
BPART	28.0	27.4	10.1	3.91	54.8
TARIF1	6.90	4.99	5.89	0.00	60.00
TARIF2	14.86	11.65	10.80	0.00	100.00
FDI	70.6	32.2	198.	0.613	1.99×10^3
LABC	54.1	55.6	15.2	19.8	133.5

Source: Author's elaboration based on information from OECD and UNCTAD.

The following subsections present the estimation results of the two models specified in Section 2. Section 3.4.1 presents the results for the group of countries formed by the United States, Mexico, and China, while Section 3.4.2 presents the results for the complete sample of 61 countries, adopting a global approach.

3.4.1. The Case of Mexico, the U.S., and China

Table 2 presents the estimation results for the specific case of Mexico, the U.S., and China. For each of the two models, there are two dependent variables: total (TTPART) and backward (BPART) GVC participation. The results for tariff variables are, apparently, counterintuitive, revealing a positive impact on GVC participation rates. However, this apparent contradiction may be justified by the heterogeneous nature and different profiles of international insertion of these three countries: Mexico, with higher backward participation in the textile and apparel GVC, and higher tariffs affecting textile manufacturing than the U.S. and China. Meanwhile, lower labor costs in Mexican textiles may be counterbalancing their higher tariffs. Regarding FDI, the results are in line with the positive effect expected for this variable, which reveals a positive and significant effect on backward participation.

Table 2. Results of the econometric estimation (time fixed effects) for total (TPART) and backward (BPART) GVC participation. Period: 1995–2018. Countries: 3 (China, Mexico, and United States) Number of observations: 72.

	Total GVC Participation (TPART)		Backward GVC Participation (BPART)	
	Model I	**Model II**	**Model I**	**Model II**
Constant	36.9778 (3.3072) ***	37.2433 (2.7351) ***	10.6792 (3.4515) ***	10.8526 (2.7861) ***
TARIF1	0.4301 (0.1175) ***	—	0.4814 (0.1226) ***	—
TARIF2	—	0.2689 (0.0494) ***	—	0.3058 (0.0503) ***
FDI	0.0445 (0.0467)	0.0569 (0.0422)	0.1209 (0.0488) **	0.1360 (0.0430) ***
LABC	0.0799 (0.0463) *	0.0518 (0.0417)	0.0636 (0.0483)	0.0320 (0.0424)
R^2	0.83	0.86	0.95	0.96
rho	0.67	0.67	0.83	0.82

Source: Author's elaboration based on information from OECD and UNCTAD Note: The standard error is indicated in brackets. "***", "**" and "*" indicate significance at 0.01%, 0.05% and 0.1%, respectively.

3.4.2. The Global Case

This subsection extends the analysis carried out in the previous point to a broader set of countries. In particular, a total of 61 countries are considered, grouped into developed and developing countries, thus, capturing different participation patterns, according to income level. Tables 3 and 4 present the main results of the estimation of the two models presented in expression (2) for three groups: full sample, developed economies, and developing economies (see Appendix B). It is of particular interest to consider both groups individually as the effects of FDI may vary on the performance of the textile industry, which is generally considered a central sector in the industrialization process of developing economies (Raei et al. 2019). As in Table 2, there are two cases for each model, considering total (Table 3) and backward (Table 4) GVC participation.

Table 3. Results of the econometric estimation (time fixed effects) for total GVC participation (TPART). Period: 1995–2018. Countries: 61. Number of observations: 1464.

	Full Sample (61 Countries)		**Developed Economies (39 Countries)**		**Developing Economies (22 Countries)**	
	Model I	**Model II**	**Model I**	**Model II**	**Model I**	**Model II**
Constant	51.0136 (0.8364) ***	50.5985 (0.8683) ***	50.5143 (1.2384) ***	48.7253 (1.295) ***	45.9248 (1.2792) ***	44.0946 (1.3576) ***
TARIF1	−0.1778 (0.0287) ***	—	0.1352 (0.1024)	—	−0.1565 (0.0313) ***	—
TARIF2	—	−0.0255 (0.0169)	—	0.1828 (0.0421) ***	—	−0.0217 (0.0197)
FDI	−0.0023 (0.0008) ***	−0.0021 (0.0008) ***	−0.0022 (0.0008)	−0.0016 (0.0008) **	0.0438 (0.0126) ***	0.0622 (0.0130) ***
LABC	0.0166 (0.0152)	0.0084 (0.0153)	0.0461 (0.0202) **	0.0511 (0.0198) ***	−0.0168 (0.0231)	−0.0182 (0.0236)
R^2	0.83	0.83	0.80	0.81	0.79	0.78
rho	0.77	0.78	0.77	0.77	0.74	0.75

Source: Author's elaboration based on information from OECD and UNCTAD Note: The standard error is indicated in brackets. "***", and "**" indicate significance at 0.01%, and 0.05%, respectively.

Table 4. Results of the econometric estimation (time fixed effects) for backward GVC participation (BPART). Period: 1995–2018. Countries: 61. Number of observations: 1464.

	Full Sample (61 Countries)		**Developed Economies (39 Countries)**		**Developing Economies (22 Countries)**	
	Model I	**Model II**	**Model I**	**Model II**	**Model I**	**Model II**
Constant	30.6691 (0.8099) ***	30.4297 (0.8515) ***	31.1911 (1.1231) ***	29.9897 (1.2062) ***	26.7746 (1.3834) ***	24.6514 (1.4766) ***
TARIF1	−0.2621 (0.0278) ***	—	−0.5200 (0.0928) ***	—	−0.1909 (0.0339) ***	—
TARIF2	—	−0.0653 (0.0166) ***	—	−0.0157 (0.0392)	—	−0.0299 (0.0214)
FDI	−0.0010 (0.0008)	−0.0009 (0.0008)	−0.0014 (0.0007) *	−0.0010 (0.0008)	0.0481 (0.0136) ***	0.0697 (0.0141) ***
LABC	−0.0151 (0.0147)	−0.0263 (0.0150) *	0.0146 (0.0184)	−0.0023 (0.0184)	−0.0262 (0.0249)	−0.0281 (0.0257)
R^2	0.85	0.85	0.83	0.82	0.87	0.86
rho	0.81	0.82	0.79	0.79	0.84	0.84

Source: Author's elaboration based on information from OECD and UNCTAD Note: The standard error is indicated in brackets. "***" and "*" indicate significance at 0.01% and 0.1%, respectively.

According to the results presented in Tables 3 and 4, the tariff protection level is revealed as a key variable in the participation degree in GVCs, regardless of whether total or backward participation is considered a dependent variable. In the full sample models, for developing and developed countries in the case of backward participation, a higher level of tariff protection would lead to lower GVC participation in textiles through a trade-reducing effect concerning the economic costs associated with export and import flows.

This result is particularly relevant considering that the textile industry, especially in developing countries, is sensitive to internal and external cost changes. However, in the case of the developed countries presented in Table 3, the results suggest that higher tariff protection on textile products (but not on raw materials) would increase total GVC participation. These results may make sense insofar as the participation of developed countries in the textile value chain is more linked to design and branding tasks and less to the manufacture of textile products. Cost sensitivity does not only stand out among developing countries. As Lawless and Morgenroth (2019) point out, Brexit is an excellent example of how a change in trade tariffs can affect cost-sensitive sectors, such as food and textiles.

Something similar happens with the other variables incorporated in the model. In the case of FDI, it is generally significant in all cases, while labor costs are significant in the case of developing countries. Regarding FDI, a greater inflow would imply higher participation in GVC in the case of developing countries; however, in developed economies, a greater inflow of FDI would have the opposite effect. In this regard, it should be noted that developed countries are, in many cases, the final markets for textile products. In this sense, an increase in FDI in these countries could well reduce the global textile insertion of developed economies and strengthen their domestic market. Finally, it should be noted that when analyzing the influence of FDI on the total GVC participation without discriminating by sector of activity, the literature supports a significant and positive relationship (Martínez-Galán and Fontoura 2019; Okah Efogo et al. 2022).

Labor costs are significant, especially in developing countries, because their insertion in the textile sector occurs mainly through a cost strategy. Thus, a higher labor cost would imply lower participation in GVCs (Fukase 2013; Javed and Atif 2021; McCaig 2011). However, regarding developed countries, labor costs are not a useful variable to explain the insertion of the textile sector in GVCs since, in this group of countries, the industry does not depend on its competitiveness in terms of costs, but rather on its innovative capacity, design, and brand prestige (Padilha and Gomes 2016; Vila and Kuster 2007).

4. Conclusions

This paper is based on the recognition of the important role played by GVC in production processes, because of international fragmentation of production, which particularly affects manufacturing sectors such as the textile and apparel industry. In this sense, different conclusions can be drawn.

From a comparative perspective, the analysis of the GVC participation shows interesting results. China, and especially, Mexico and the United States present high levels of total participation, with different trends. However, from this sectoral approach, it also can be observed an accentuation of general GVC insertion patterns. In this sense, the United States and China show clearer forward participation, while Mexico follows a different pattern, with a predominance of backward participation. This feature is common in countries specializing in manufacturing and assembly tasks.

Another relevant finding is China's growing importance in the origin of the value-added incorporated in Mexican and U.S. textile exports. This result shows a steady increase in the backward linkages of Factory North America with the Asian giant. The contrary occurs in Mexico and especially the United States, which currently have residual importance as the origin of the value-added of China's textile exports and have even seen this importance decrease in the period of analysis (1995–2018).

These results are consistent with the characterization of the decline of Mexican and U.S. textile manufacturing, due to the presence of fragmented global trade flows not positively exploited. However, this contrasts with the innovative and competitive Asian textile and apparel industry, which adopts the full package strategy and FDI in a positive sense to increase its value-added with empowerment in productive fragmentation.

The analysis of the participation in the final global demand for textile products provides additional findings. The most remarkable result is that the Chinese share shows a clear and steady increase, approaching 40% of global demand. However, Mexico and the

United States show the opposite trend, becoming nearly residual. This contrasts with their respective shares in the textile GVC, where there are no significant differences in the levels of participation.

From an extended (61 countries) and explanatory perspective, the econometric analysis reveals tariff rates as a key factor for textile GVC participation, particularly in developing economies. However, the effect of other variables, such as FDI stock and labor costs, depends, to a large extent, on the group of countries considered. While among developed economies, FDI does not help to provide a simple understanding of their insertion in textile GVCs which, in developing economies, is a significant factor. This may be due to the type of tasks performed by each country. While the first group focuses on tasks such as branding and design, the developing countries mainly perform manufacturing tasks.

Although labor costs only reveal a significant and negative effect on backward participation for the full sample, this variable shows a negative impact mainly for developing countries, which is consistent with the expected influence of this factor. The counterintuitive result obtained for this variable for developed economies can be explained because the total GVC participation of these countries is driven by high labor costs (e.g., offshoring of manufacturing tasks to other countries), which causes a positive (and significant) effect on their total GVC participation. In contrast, the lower level of labor costs positively influences the GVC participation of developing economies.

Finally, this research highlights the profound changes the textile sector has undergone in a highly competitive and globalized context. In particular, it allows us to understand the changing role of economies in the international fragmentation of production, where countries follow different insertion patterns, with uneven results.

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

Funding: This research has been supported by the ICEDE research group, to which the authors belong, Galician Competitive Research Group financed by Xunta de Galicia (Ref. ED431C 2022/15).

Data Availability Statement: All data have been obtained from publicly available databases, such as TiVA (OCDE), UNCTAD, and WTO.

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

Appendix A

GVC Participation and Corresponding OECD TiVA Indicators (Extracted from Martins Guilhoto et al. 2022)

Backward participation in GVCs, percentage (DEXFVApSH): Foreign VA embodied in exports, as % of total gross exports of the exporting country. This indicator is calculated for the total value of source and exporting industries; it is estimated as the ratio between the VA contents of imports from the source country p and the gross exports of the exporting country c. This indicator is estimated as: DEXFVApSHc,p = EXGR_BSCIc,p/EXGRc $\times$ 100, where EXGR_BSCIc,p is the total VA from country p embodied in the total exports of exporting country c, and EXGRc is the total gross exports of exporting country c.

Forward participation in GVCs, percentage (FEXDVApSH): Domestic VA embodied in foreign exports as a share (%) of total gross exports of the value-added source country. This indicator is calculated for the total value of source and exporting industries; it is estimated as the VA contents of exports originating in the source country, and embodied in the exports of the exporting country, divided by the gross exports of the source country. This indicator is estimated as follows: FEXDVApSHc,p = EXGR_BSCIc,p/EXGRc $\times$ 100, where EXGR_BSCIc,p is the total VA from country c embodied in the exports of country p, and EXGRc is the total gross exports of the value added source country c.

Appendix B

Table A1. List of the 61 countries/territories and considered in the econometric analysis.

Developed Economies (39)		Developing Economies (22)	
Australia	Japan	Argentina	Thailand
Austria	Korea	Brazil	Tunisia
Belgium	Latvia	Bulgaria	Turkey
Canada	Lithuania	Cambodia	
Taiwan	Malta	Chile	
Croatia	The Netherlands	China	
Cyprus	New Zealand	Colombia	
Czech Republic	Norway	Costa Rica	
Denmark	Poland	India	
Estonia	Portugal	Kazakhstan	
Finland	Saudi Arabia	Laos	
France	Singapore	Malaysia	
Germany	Slovak Republic	Mexico	
Greece	Slovenia	Morocco	
Hong Kong	Spain	Peru	
Hungary	Sweden	Philippines	
Iceland	Switzerland	Romania	
Ireland	United Kingdom	Russian Federation	
Israel	United States	South Africa	
Italy			

References

Amador, João, and Sónia Cabral. 2014. *Global Value Chains: Surveying Drivers and Measures*. ECB Working Paper 1739. Frankfurt: European Central Bank (ECB). ISBN 978-92-899-1147-4.

Amendolagine, Vito, Andrea Presbitero, Roberta Rabellotti, Marco Sanfilippo, and Adrian Seric. 2017. *FDI, Global Value Chains, and Local Sourcing in Developing Countries*. IMF Working Paper 17/284. Washington, DC: International Monetary Fund.

Antrás, Paul. 2020. Conceptual aspects of global value chains. *The World Bank Economic Review* 34: 551–74. [CrossRef]

Arrighetti, Alessandro, Fabio Landini, and Andrea Lasagni. 2014. Intangible assets and firm heterogeneity: Evidence from Italy. *Research Policy* 43: 202–13. [CrossRef]

Arroyo, Pilar, and Lourdes Cárcamo. 2010. La evolución histórica e importancia económica del sector textil y del vestido en México. In *Economía y Sociedad*. Morelia: Universidad Michoacana de San Nicolás de Hidalgo, pp. 51–68.

Baldwin, Richard. 2016. *The Great Convergence: Information Technology and the New Globalization*. Cambridge: Belknap Press.

Banga, Rashmi. 2014. Linking into Global Value Chains Is Not Sufficient: Do You Export Domestic Value Added Contents? *Journal of Economic Integration* 29: 267–97. [CrossRef]

Castro-Gonzáles, Segundo, and Dale Mathews. 2013. Cómo influye la Tasa de Cambio Real en la competitividad de las exportaciones de confecciones textiles entre Asia y Centroamérica. *Estrategias* 11: 103–19. [CrossRef]

Chen, Minglu, and David Goodman. 2018. Reporting Mexico in China: 'China's economy is developing robustly and rapidly'. *México y la Cuenca del Pacífico* 8: 143–63.

Choksy, Umair, Muhammad Ayaz, Omar Al-Tabbaa, and Mahour Parast. 2022. Supplier resilience under the COVID-19 crisis in apparel global value chain (GVC): The role of GVC governance and supplier's upgrading. *Journal of Business Research* 150: 249–67. [CrossRef]

Cohen, Morris, and Hau Lee. 2020. Designing the right global supply chain network. *Manufacturing and Service Operations Management* 22: 15–24. [CrossRef]

Del Prete, Davide, Giorgia Giovannetti, and Enrico Marvasi. 2017. Global value chains participation and productivity gains for North African firms. *Review of World Economics* 153: 675–701. [CrossRef]

Durand, Cédric, and William Milberg. 2020. Intellectual monopoly in global value chains. *Review of International Political Economy* 27: 404–29. [CrossRef]

Fernández, Karina, and Gary Gereffi. 2019. Global value chain analysis: A primer. In *Handbook on Global Value Chains*. Cheltenham: Edward Elgar Publishing.

Frederick, Stacey, and Gary Gereffi. 2011. Upgrading and restructuring in the global apparel value chain: Why China and Asia are outperforming Mexico and Central America. *International Journal of Technological Learning, Innovation and Development* 4: 67–95. [CrossRef]

Fuji, Gerardo, Eduardo Candaudap, and Claudia Gaona. 2005. Exportaciones, industria maquiladora y crecimiento económico en México a partir de la década de los noventa. *Investigación Económica* 64: 125–56.

Fukase, Emilio. 2013. *Export Liberalization, Job Creation and the Skill Premium: Evidence from the U.S.-Vietnam Bilateral Trade Agreement*. Washington, DC: World Bank. [CrossRef]

Gereffi, Gary. 2019. Global value chains, development, and emerging economies. In *Business and Development Studies*. London: Routledge, pp. 125–58.

Gereffi, Gary, and Karina Fernandez-Stark. 2016. Global Value Chain Analysis: A Primer. Duke CGGC (Center on Globalization, Governance & Competitiveness). Available online: https://bit.ly/30GKIXM (accessed on 9 May 2022).

Gereffi, Gary, and Joonkoo Lee. 2016. Economic and Social Upgrading in Global Value Chains and Industrial Clusters: Why Governance Matters. *Journal of Business Ethics* 133: 25–38. [CrossRef]

Gereffi, Gary, and Xubei Luo. 2015. Risks and Opportunities of Participation in Global Value Chains. *Journal of Banking and Financial Economics* 2: 51–63. [CrossRef]

Gereffi, Gary, and Olga Memedovic. 2003. *The Global Apparel Value Chain: What Prospects for Upgrading by Developing Countries*. Vienna: United Nations Industrial Development Organization, pp. 1–40.

Gómez Chiñas, Carlos, and Juan González García. 2017. Competencia y competitividad de las exportaciones de México y China en el mercado estadounidense: Nueva evidencia. *México y la cuenca del pacífico* 6: 79–105. [CrossRef]

Gómez-Galvarriato, Aurora. 1999. Industrial development under institutional frailty: The development of the Mexican textile industry in the nineteenth century. *Revista de Historia Economica-Journal of Iberian and Latin American Economic History* 17: 191–223. [CrossRef]

Gracia-Hernández, Maximiliano. 2011. Competencia entre México y China en el interior del mercado de Estados Unidos. *Economía, Sociedad y Territorio* 13: 37–78. [CrossRef]

Hansen, Lawrence. 2020. The Origins of the Maquila Industry in Mexico. *Migrationsforschung–Interdisziplinär and Diskursiv: Internationale Forschungserträge zu Migration in Wirtschaft, Geschichte und Gesellschaft* 1: 123.

Inomata, Satoshi, and Daria Taglioni. 2019. *Technological Progress, Diffusion, and Opportunities for Developing Countries: Lessons from China*. Global Value Chain Development Report 2019. Geneva: World Trade Organization, pp. 83–101.

Javed, Ayesha, and Rao Atif. 2021. Global Value Chain: An Analysis of Pakistan's Textile Sector. *Global Business Review* 22: 879–92. [CrossRef]

Jona-Lasinio, Cecilia, Stefano Manzocchi, and Valentina Meliciani. 2016. Intangible Assets and Participation in Global Value Chains: An Analysis on a Sample of European Countries. *Rivista di Politica Economica* 7–9: 65–96.

Koopman, Robert, Zhi Wang, and Shang-Jin Wei. 2014. Tracing Value-Added and Double Counting in Gross Exports. *The American Economic Review* 104: 459–94. [CrossRef]

Kowalski, Przemyslaw, Javier Lopez Gonzalez, Alexandros Ragoussis, and Cristian Ugarte. 2015. *Participation of Developing Countries in Global Value Chains. Implications for Trade and Trade-Related Policies*. OECD Trade Policy Papers, No. 179. Paris: OECD Publishing.

Lawless, Martina, and Edgar Morgenroth. 2019. The product and sector level impact of a hard Brexit across the EU. *Contemporary Social Science* 14: 189–207. [CrossRef]

Li, Xin, Bo Meng, and Zhi Wang. 2019. Recent patterns of global production and GVC participation. In *Technological Innovation, Supply Chain Trade, and Workers in a Globalized World*. Geneva: World Trade Organization, pp. 9–44.

Lu, Sheng. 2013. *Moving to "Made in China" to "Design in China"*. FASH455 Global Apparel and Textile Trade and Sourcing. Newark: University of Delaware. Available online: https://shenglufashion.com/2013/08/10/moving-from-made-in-china-to-design-in-china/ (accessed on 1 April 2022).

Lu, Sheng. 2018. *Changing Role in the World Textile and Apparel Supply Chain*. FASH455 Global Apparel and Textile Trade and Sourcing. Newark: University of Delaware. Available online: https://shenglufashion.com/2018/11/02/chinas-changing-role-in-the-world-textile-and-apparel-supply-chain/ (accessed on 25 March 2022).

Lund, Susan, James Manyika, Jonathan Woetzel, Jacques Bughin, Mekala Krishnan, Jeongmin Seong, and Mac Muir. 2019. Globalization in transition: The future of trade and value chains. In *Globalization in Transition: The Future of Trade and Value Chains*. San Francisco: McKinsey Global Institute, pp. 35–51.

Ma, Huimin, Xiang Wu, Yan Li, Huang Han, Wu Han, Xiong Jie, and Zhang Jinlong. 2018. Strategic plan of "Made in China 2025" and its implementation. In *Analyzing the Impacts of Industry 4.0 in Modern Business Environments*. Hershey: IGI Global, pp. 1–23. ISBN 9781522534686.

Martínez-Galán, Enrique, and María Fontoura. 2019. Global value chains and inward foreign direct investment in the 2000s. *The World Economy* 42: 175–96. [CrossRef]

Martins Guilhoto, Joaquim, Colin Webb, and Norihiko Yamano. 2022. *Guide to OECD TiVA Indicators*, 2021 ed. OECD Science, Technology and Industry Working Papers No. 2022/02. Paris: OECD Publishing. [CrossRef]

McCaig, Brian. 2011. Exporting out of poverty: Provincial poverty in Vietnam and U.S. market access. *Journal of International Economics* 85: 102–13. [CrossRef]

Montón, Jorge. 2015. Offshoring in Textile Industry. Doctoral dissertation, Universidad Politécnica de Valencia, Valencia, Spain. Available online: http://hdl.handle.net/10251/47923 (accessed on 1 May 2022).

Morrison, Andrea, Carlo Rabellotti, and Roberta Pietrobelli. 2008. *Global Value Chains and Technological Capabilities: A Framework to Study Industrial Innovation in Developing Countries*. Aalborg: GLOBELICS, pp. 1–27.

OECD. 2013. *Interconnected Economies: Benefiting from Global Value Chains*. Paris: OECD Publishing. Available online: https://www.oecd.org/sti/ind/interconnected-economies-GVCs-synthesis.pdf (accessed on 25 March 2022).

OECD. 2021. TiVA Indicators: 2021 Edition. Available online: https://www.oecd.org/sti/ind/measuring-trade-in-value-added.htm (accessed on 25 March 2022).

Okah Efogo, Françoise, Kwami Wonyra, and Evans Osabuohien. 2022. Foreign direct investment and participation of developing countries in global value chains: Lessons from the last decade. *International Review of Applied Economics* 36: 264–84. [CrossRef]

Padilha, Carolina, and Giancarlo Gomes. 2016. Innovation culture and performance in innovation of products and processes: A study in companies of textile industry. *RAI Revista de Administração e Inovação* 13: 285–94. [CrossRef]

Pahl, Stefan, and Marcel P. Timmer. 2020. Do Global Value Chains Enhance Economic Upgrading? A Long View. *The Journal of Development Studies* 56: 1683–705. [CrossRef]

Pathikonda, Vilas, and Thomas Farole. 2017. The Capabilities Driving Participation in Global Value Chains. *Journal of International Commerce Economics and Policy* 8: 1750006. [CrossRef]

Pino, Juan Carlos. 2020. *China, un Actor no Invitado al TLCAN-TMEC. University Repository*. México: Universidad Nacional Autónoma de México (UNAM). Available online: http://ru.iiec.unam.mx/5056/1/4-199-Pino.pdf (accessed on 15 March 2022).

Pipkin, Seth, and Alberto Fuentes. 2017. Spurred to upgrade: A review of triggers and consequences of industrial upgrading in the global value chain literature. *World Development* 98: 536–54. [CrossRef]

Ponte, Stefano, Gary Gereffi, and Gale Raj-Reichert. 2019. *Introduction to the Handbook on Global Value Chains*. Cheltenham: Edward Elgar Publising Limited, pp. 1–27. [CrossRef]

Raei, Faezeh, Anna Ignatenko, and Borislava Mircheva. 2019. *Global Value Chains: What Are the Benefits and Why Do Countries Participate?* Washington, DC: International Monetary Fund.

Rahman, Jesmin, and Tianli Zhao. 2013. *Export Performance in Europe: What Do We Know from Supply Links?* IMF Working Papers 13/62. Washington, DC: International Monetary Fund.

Robertson, Raymond, Timothy Halliday, and Sindhu Vasireddy. 2020. Labour market adjustment to third-party competition: Evidence from Mexico. *The World Economy* 43: 1977–2006. [CrossRef]

Rodil, Óscar. 2017. La inserción conjunta de América Latina y China en las cadenas globales de valor. *Revista Comercio Exterior BANCOMEXT* 13: 8–14.

Rodil, Óscar, and Ana Gómez. 2021. La inserción del sector textil de México, Estados Unidos y China en las cadenas globales de valor. *Revista Economía y Desafíos del Desarrollo* 1: 77–94.

Rodríguez, Carlos, and Lizbeth Fernández. 2006. Manufactura textil en México: Un enfoque sistémico. *Revista Venezolana de Gerencia* 11: 335–351. [CrossRef]

Stehrer, Rober, and Roman Stöllinger. 2015. *The Central European Manufacturing Core: What Is Driving Regional Production Sharing?* FIW Research Reports 2014/15, N°02. Vienna: Research Centre International Economics.

Taglioni, Daria, and Deborah Winkler. 2014. *Making Global Value Chains Work for Development*. Poverty Reduction and Economic Management. The World Bank 143. Washington, DC: World Bank.

Vázquez, Gabriela, Osvaldo Torres, and Sara Cuenca. 2015. La competitividad de los productos textiles mexicanos en el marco de la Alianza del Pacifico, 1980–2014. *RECAI Revista de Estudios en Contaduría, Administración e Informática* 4: 43–67.

Vila, Natalia, and Ines Kuster. 2007. The importance of innovation in international textile firms. *European Journal of Marketing* 41: 17–36. [CrossRef]

Vrh, Nataša. 2018. What drives the differences in domestic value added in exports between old and new E.U. member states? *Economic Research-Ekonomska Istraživanja* 31: 645–663. [CrossRef]

Whitfield, Lindsay, Kristoffer Marslev, and Cornelia Staritz. 2021. *Can Apparel Export Industries Catalyse Industrialisation? Combining GVC Participation and Localisation*. SARChI Industrial Development Working Paper Series. Johannesburg: The South African Research Chair in Industrial Development (SARChI).

WTO. 2014. *The Rise of Global Value Chains. World Trade Report 2014. Trade and Development: Recent Trends and the Role of the WTO*. Geneva: World Trade Organization, pp. 78–127.

Yi, Kei-Mu. 2003. Can Vertical Specialization Explain the Growth of World Trade? *Journal of Political Economy* 111: 52–102. [CrossRef]

economies

MDPI

Article

A High-Resolution Lead-Lag Analysis of US GDP, Employment, and Unemployment 1977–2021: Okun's Law and the Puzzle of Jobless Recovery

Knut Lehre Seip [1,*] and Dan Zhang [2]

[1] Faculty of Technology Arts and Design, Oslo Metropolitan University, P.O. Box 4 St. Olavs Plass, Pilestredet Park 33, 0176 Oslo, Norway
[2] Faculty of Social Sciences, Oslo Business School, Finance and Accounting, Oslo Metropolitan University, P.O. Box 4 St. Olavs Plass, Pilestredet Park 33, 0176 Oslo, Norway
* Correspondence: knut.lehre.seip@oslomet.no; Tel.: +47-67238816

Abstract: Okun's law is formulated as the ratio between GDP and unemployment (UE): $\beta = f(\text{GDP/UE})$. It is used to investigate the relations between output and labor input across regions or across business cycles. Based on results by James D. Hamilton we replaced the United States UE with employment (EM) for the years 1977 to 2021 and examined how employment changed relative to output during recessions and recoveries. We found that (i) EM was leading GDP before and lagging GDP after all recessions, except the 2020 recession. (ii) The $\beta_E(9)$ = GDP/EM for rolling ordinary linear regression over 9 months decreases just after a recession and then recovers over 2- to 4-year periods. (iii) The two series showing that EM $\rightarrow$ GDP and $\beta_E(9) < 0.5$ coincided in the 34 months that partly preceded and partly followed five of six NBER recession dates, providing a probability of $\approx$0.0002 to coincide with the recessions by chance. Thus, the two series may be used to support forecasts of coming recessions. Since EM precedes GDP and labor productivity declines before recessions, a policy recommendation for avoiding "jobless recovery" is that employment should not increase more rapidly than the real economy.

Keywords: Okun's law; employment; unemployment; US recessions; jobless recovery; job search; business cycles

Citation: Seip, Knut Lehre, and Dan Zhang. 2022. A High-Resolution Lead-Lag Analysis of US GDP, Employment, and Unemployment 1977–2021: Okun's Law and the Puzzle of Jobless Recovery. *Economies* 10: 260. https://doi.org/10.3390/economies10100260

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 7 September 2022
Accepted: 12 October 2022
Published: 20 October 2022

1. Introduction

Economic recessions are events that seriously affect employment and output. Since the early 1990s, recoveries from recessions in the United States were followed by persistently weak employment growth (Galí et al. 2012). Policymakers and economists refer to this phenomenon as jobless recovery and it remains a puzzle to experts. In this article, we examine if there are behavioral traits among employers or employees that have the potential to be managed so that some of the detrimental effects of recessions can be abated. We do this by examining the employment and output in the context of Okun's law.

Okun's law depicts an empirically observed relation between the gross domestic product (GDP) and unemployment (UE), which is specified as:

$$\frac{\left(\overline{\text{GDP}} - \text{GDP}\right)}{\overline{\text{GDP}}} = \beta\left(\text{UE} - \overline{\text{UE}}\right) \quad (1)$$

where $\overline{\text{GDP}} - \text{GDP}$ is the output, GDP, gap and $\text{UE} - \overline{\text{UE}}$ is the unemployment (UE) gap. In short, β = GDP gap/UE gap. The β is Okun's coefficient or Okun's elasticity coefficient.

Recently, Hamilton (2018, p. 838) showed that the cyclical component of employment (EM) started to decline significantly before the NBER business cycle peak for essentially every recession. Thus, lead-lag (LL) relations and LL (GDP, EM) show that EM is leading

GDP before a recession. Inspired by Hamilton (2018)'s results, we study GDP versus EM and examine the relation β_E = GDP/EM and its relation to jobless recovery following a recession. Both EM and UE are often included in forecasting algorithms for GDP (Camacho and Perez-Quiros 2010; Hamilton 2018) and thus their detailed behavior is an important issue in applied economics. Our findings provide insights into the relationship between GDP and employment as well as jobless recovery.

For four recessions, we found that employment is a leading variable of GDP, and labor productivity decreases or levels off relative to its potential growth. The leading role of employment ceases after the recessions, and the change in LL relations causes spikes in β_E = GDP/EM. There are 34 months out of 547 where EM leads GDP and β_E shows peaks at the same time, and these 34 months partly precede and partly follow the beginning of the NBER recessions, except the last COVID-19 recession in 2020.

The present study differs from other studies in that we examine (i) the ordinary linear regression (OLR) β coefficients for GDP/UE and GDP/EM over running time windows (9 months) and we study (ii) the LL relations between GDP, EM, and UE over very short time horizons (9 months). Most other studies restrict their study to GDP and UE, and they use longer time windows, e.g., decades, (Cazes et al. 2013; Donayre and Panovska 2021). We also calculate labor productivity with a running time window of 9 months, and lastly, we embed the LL results in a principal component analysis (PCA) "map" of the US economy to place the LL results into a wider context.

A Literature and Hypotheses

Various attributes of Okun's law have been studied extensively in the literature and summaries have been made, for example in Donayre and Panovska (2021) and Obst (2020). Here, we simply summarize attributes assigned to Okun's law with respect to regional differences and differences across business cycle regimes. The latter studies relate to regime shifts, recession and shocks that are identified in the cycles.

Maza (2022) found regional differences across Europe in Okun's law, defined as $\beta = (\beta_{UE} - \Delta UE - \alpha - \mu)/\Delta GDP$, α constant, μ error term. The β-values were clustered geographically and with high values in, e.g., Germany. However, Elhorst and Emili (2022) show for the Netherlands that there are spillover effects among regions, thus a finer spatial solution than nations may be helpful to understand the relation between output and unemployment.

Across business cycles for the US economy from 1949 to 2020, Donayre and Panovska (2021) identify three regimes and thus two structural breaks, and they show that there is an increase (steepening) of Okun's β across expansions, mild recessions, and deep recessions. That is, GDP changes relatively more than the UE during a deep recession than during an expansion. In contrast, Søgner (2001) found no structural breaks in the Austrian economy from 1977 to 1995 and concludes that there is a stable Okun's β coefficient across time in Austria. Ziegenbein (2021) studies the effects of six types of macroeconomic shocks on Okun's β and found that shock type, e.g., financial shocks and government spending shocks, affect both the reaction time and the duration (0 to 20 quarters) of changes in the Okun's β, away from its "neutral" value (zero in Ziegenbein (2021)). A major conclusion is that while GDP and employment declined with similar speed during a recession, output recovered faster than employment, thus giving rise to the term "jobless recovery".

Several studies also offer explanations for the "jobless recoveries", and the mechanisms that could explain jobless recoveries can be divided into six categories: (i) Demographic traits. For example, Maza (2022) found that, across all regions in Europe, the β coefficient significantly depended on the participation of women and youth in the workforce and the industry's contribution to GDP. (ii) Legal factors were found by Maza (2022) to contribute only when the UK was exempted from the sample. However, Vaubel (2008) shows that legislative acts related to social provisions increased in the EU from 1970 to 2003. Cazes et al. (2013) suggests that whereas Okun's β increased sharply in the US, Canada, Spain, and other economies that were severely affected by the Great Recession in 2008, in countries such as Germany and the Netherlands, with strong (legislative) employment protection, it did not. However, Mukoyama and Sahin (2009) quote studies that show the ratio of benefit claims to total UE in the US has declined over the post-war era, and thus employment protection in terms of unemployment compensation may not play a prominent role for explaining jobless recovery in the US. (iii) Worker behavior may play a role in several respects. With references to Yellen (1991) Mukoyama and Sahin (2009) propose that unemployment may be caused by the unemployed searching for good, rent-paying jobs rather than working at the poor jobs following a recession. Capsada-Munsech and Valiente (2020) examine variations in employee willingness to participate in vocational education and training (VET) and thus to obtain skills that meet demand. Elhorst and Emili (2022) suggest that employed persons may work shorter hours. (iv) Employer behavior will interact with employee behavior. For example, firms in Germany and Austria support more vocational training than southern European countries (Capsada-Munsech and Valiente 2020, p. 171). (v) The mismatch between human resources available and human resources required was studied by Lazear and Spletzer (2012). However, they show that unemployment was a cyclic phenomenon during the 2007–2009 recession rather than due to a mismatch between labor requirements and labor supply. Gimbel and Sinclair (2020), studying the period 2014 to 2019, suggest that a mismatch that may have been an issue around 2014 declined after the Great Recession. However, since Okun's β changes with the development of the GDP, e.g., expansions and recessions, the average β_E coefficient will also depend on the economic policy, management, and probably luck, in the region studied. (iv) For the present study, we do not believe that the underground economy will have a great influence on employment and its recovery after a recession, but it may be important in countries with a large and established underground economy.

We develop three hypotheses that all relate to US recessions. The first hypothesis, H1, is that employment leads GDP (EM $\rightarrow$ GDP), and if employment is decreasing faster than GDP before a recession, β(GDP, EM) will tend to increase $\uparrow$ (Abel et al. (1998, p. 658 on x/y relations). The rationale for the hypothesis is the results shown by Hamilton (2018) that employment was leading GDP before recession. Our second hypothesis, H2, is that GDP leads employment (GDP $\rightarrow$ EM), and β_E(GDP, EM) decreases$\downarrow$ after a recession and during an expansion. The rationale is that Cazes et al. (2013, p. 6) show that for many countries, unemployment is likely to rise (and employment to decrease (our interpretation)) during a recession. However, if employment decreases less rapidly than GDP, β_E (GDP/EM) will tend to increase. Our third hypothesis, H3, is that by embedding the lead-lag results in a "map" of US macroeconomy, we will obtain clues as to which macroeconomic variables determine the LL relations between GDP and EM. The rationale is that Seip et al. (2019), by using the "map" method, found macroeconomic conditions for why leading indexes failed to predict industrial production in Germany.

The rest of the article is organized as follows. In Section 2, we present the data we used, and in Section 3, we outline the methods used with emphasis on the relatively novel LL method. In Section 4, we present the results from the application of the high-resolution LL method and, as far as we know, a novel application of Okun's law to GDP, EM, and UE. The results are discussed in Section 5, and Section 6 summarizes and concludes.

2. Data

All the following data are retrieved from the St. Louis Federal Reserve database between 14 June 2022 and 10 July 2022. We use two sets of data. The first set is US employment rate, EP, US unemployment rate (UE) and real gross domestic product (GDP). The other set is used to draw a "map" of the US economy. We emphasize variables that may have implications for the interpretation of Okun's law and we include time series that could allow us to examine causal mechanisms for "jobless recovery".

Okun's law data. UE represents the number of unemployed as a percentage of the nonfarm labor force. EM is the number of employed in thousand persons. Real GDP is measured in billions of chained 2012 dollars.

US economy data. The strings of letters following the acronyms we use are the identification code used by St. Louis Fed. We have chosen twelve macroeconomic variables. In addition to real GDP, we chose industrial production (IP); working hours (WH)—HOANBS; inflation (INF) represented by the consumer price index (CPI)-CPI; US government expenditures (EXP)—W068RCQ27SBEA; federal government tax receipts (TRE)—W006RC1Q027SBEA; federal government: current expenditures (CE)—FGEXPND; federal debt as total public debt (PD)—GFDEBTN; federal funds rate (FF); and monetary supply (M2). Data for labor productivity (LP)—OPHNFB is an index for output per hour and were used to compare productivity during recession phases. Data for union affiliation as percentage of employed were only available from 2011 to 2021 and ranged from 10.3 to 11.8% (https://data.bls.gov/cgi-bin/surveymost) accessed on 15 August 2022, and were therefore not used. For wage spread, we used the Gini index, which is only sporadically available for the US before 1991. A high index value suggests a high degree of inequality. Mukoyama and Sahin (2009, p. 203) show a curve for the 90–10% residual wage inequality from 1970 to 2002, and we extended the Gini index to the time window 1977 to 2022 based on the author's figure. We use the recession dating from NBER, and characteristics for the recessions are shown in Table 1.

Table 1. Recessions in USA, 1977 to 2022. Beginning, end and duration are NBER data. GDP decline, labor productivity, jobless depth and jobless duration is measured as the anomaly from the linear detrended series.

Beginning	End	Duration	GDP Decline (Peak to Trough)	EM Leads GDP before Recession (1)	b_E (GDP,EM) Coefficient Peaks	b_{LP} (LP,t) before Rec. (2)	Jobless Depth (3)	Jobless Duration (4)
		Months	Relative Values	0–12 Months	<0.5	0–12 Months	Relative Values	Months
Jan 1980	July 1980	6	−2.2	3.00	−0.30	−0.041	−0.18	14
July 1981	Nov 1982	16	−2.7	12.00	0.18	−0.022	−0.53	18
July 1990	March 1991	8	−1.4	5.00	−0.16	0.000	−0.83	12
March 2001	Nov 2001	8	−0.3	11.00	−0.64	−0.010	−1.27	68 (5)
Dec 2007	June 2009	18	−5.1	9.00	0.48	0.019	−2.54	144 (5)
Feb 2020	April 2020	2	−19.2	1.00	−0.1	−0.015	−4.62	12
Average		9.67 ± 5.40	−5.15 ± 1.79	6.83 ± 3.87	−0.25 ± 0.31	−0.012 ± 0.020	−1.66 ± 0.91	44.67 ± 58.83

Number of months EM leads GDP during one years before a NBER recession. Peak value relative to a detrended β_E. Deviation from linear trend. Months to pre-recession values. Employment does not recover to pre-recession values before next recession. Duration is the time between the two recessions.

3. Methodology

We use a relatively novel technique for calculating high resolution LL relations between cyclic time series. The method relates to a dual representation of the time series, x(t) and y(t), first as a series depicted as a function of time and second as depicted in a phase plot with one series on the x-axis and the other series on the y-axis. Time in the phase plot is then shown as the trajectories between points. For a quick intuitive illustration, see: https://en.wikipedia.org/wiki/Lissajous_curve#/media/File:Lissajous_phase.svg (accessed on 15 February 2022).

Seip et al. (2018) describe the method in detail, but recently Krüger (2021) has described an LL method that is based on wavelet techniques and on the same dual representation of paired time series. Figure 1a shows two sine functions with identical cycle lengths, but the dashed curve is shifted a few time steps forward, and we have added a small amount of stochasticity to make the example a little more realistic. The bold curve that is to the left of the target sine function (dashed curve) is a leading series (Figure 1a). With the leading series (bold) on the x-axis and the lagging series (dashed) on the y-axis, the trajectories in the phase plot rotate counterclockwise (positive per definition, Figure 1b). Thus, we can identify LL relations by the way that trajectories rotate in the phase space. For time series normalized to unit standard deviation, the trajectories will form an ellipse-like curve with the major axis either in the 1:1 direction or the −1:1 direction. For shifted perfect sine functions with common cycle periods, the minor axis will show the phase shifts between the sine functions.

The rotational direction represented by the angle θ between two successive vectors, $\mathbf{v}_1$ and $\mathbf{v}_2$, through three consecutive observations in the trajectory is calculated with Equation (2):

$$\theta = sign(\mathbf{v}_1 \times \mathbf{v}_2) \cdot \text{Arccos}\left(\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1||\mathbf{v}_2|}\right) \tag{2}$$

The vectors are calculated as (yi − yi−1)/(xi − xi−1) with i = 2, 3, ...

We define a measure of LL strength as

$$\text{LL} = (\text{N}_+ - \text{N}_-)/(\text{N}_+ + \text{N}_-) \tag{3}$$

where N_+ and N_- is the number of positive and negative angles, θ, in a set of n consecutive observations in the two series. Using n = 9 and with N_+ = 9 and N_- = 0, we obtain LL = (9 − 0)/(9 + 0) = 1. The number 9 is a trade-off between the goal of identifying LL relations for short time windows and the goal of identifying a confidence band. In the time series mode, it means that one series leads the other for nine consecutive observations. In the phase representation, it means that when the two series are plotted as trajectories in the phase plot, the trajectories will rotate persistently in one direction.

In Figure 1c is the angles, and θ(3) in Figure 1b is shown as a function of time (light blue bars) with the LL(9) relations as dark blue bars. The angles are measured in radians (range $-\pi$, π) and the LL relations are in the interval [−1 to 1]. Note, for example, that the first angle is negative, showing a clockwise rotation. The dark blue bars are all negative, showing that the added noise is not sufficient to impair the overall clockwise rotation.

Pro-cyclic and counter-cyclic relations: If the OLR β coefficient is positive, the two series are pro-cyclic. If the β coefficient is negative, the two series are counter-cyclic. Figure 1d shows that an LL relation can be positive both for pro-cyclic and counter-cyclic series. Since a phase plot with GDP on the x-axis and EM on the y-axis will show an ellipsoid with its major axis in the 1:1 direction, the average β_E coefficient is 1 per definition. However, β_E coefficients based on short time windows of the series may deviate from 1 and thus give rise to interpretations of how coefficients change with the economy, e.g., expansions or recessions.

Figure 1e shows two sine functions where one function has a constant argument, a = sin (0.25t), whereas the other function, b = sin (0.25t + ϕ), has a variable argument, ϕ, that varies from a positive to a negative value. Figure 1b shows the ratio of the slopes a(t)/b(t) as a function of time (time window: nine time steps) compared to the average of the two sine functions. The growth rate of a/b is Δa/a − Δb/b (Abel et al. (1998, p. 658)). It is seen that at the downturn side of the sine function the ratio a/b first shows a positive peak and then a negative peak. At the upturn side it first shows a negative peak and then a positive peak.

Confidence interval. The 95% confidence interval (CI) is based on the probability that two uniformly stochastic series will show a persistent rotation in one direction. It is calculated with Monte Carlo simulations, applying Equations (2) and (3) to two uniformly stochastic series, and the confidence limits are the asymptotic values for 1000 replicates. Values of LL < −0.32 and LL > 0.32 suggest that for time series longer than nine time steps, the LL values are significant at the 95% level. When time series are smoothed with a smoothing algorithm, the probability that consecutive angles will have the same sign increases, so the CI does not strictly apply to smoothed series. However, by comparing LL relations for smoothed series with LL relations for unsmoothed series, the confidence in the LL relations for the smoothed series may be enhanced.

Detrending. Since we want to study interannual to decennial time windows corresponding to the typical duration of recessions (Burns and Mitchell 1946) we detrended the 12 US economy macroeconomic data series with a linear or a quadratic function, depending upon the large-scale form of the time series. Several detrending methods are available, but linear detrending is simple and will not introduce anomalies that have no economic relevance.

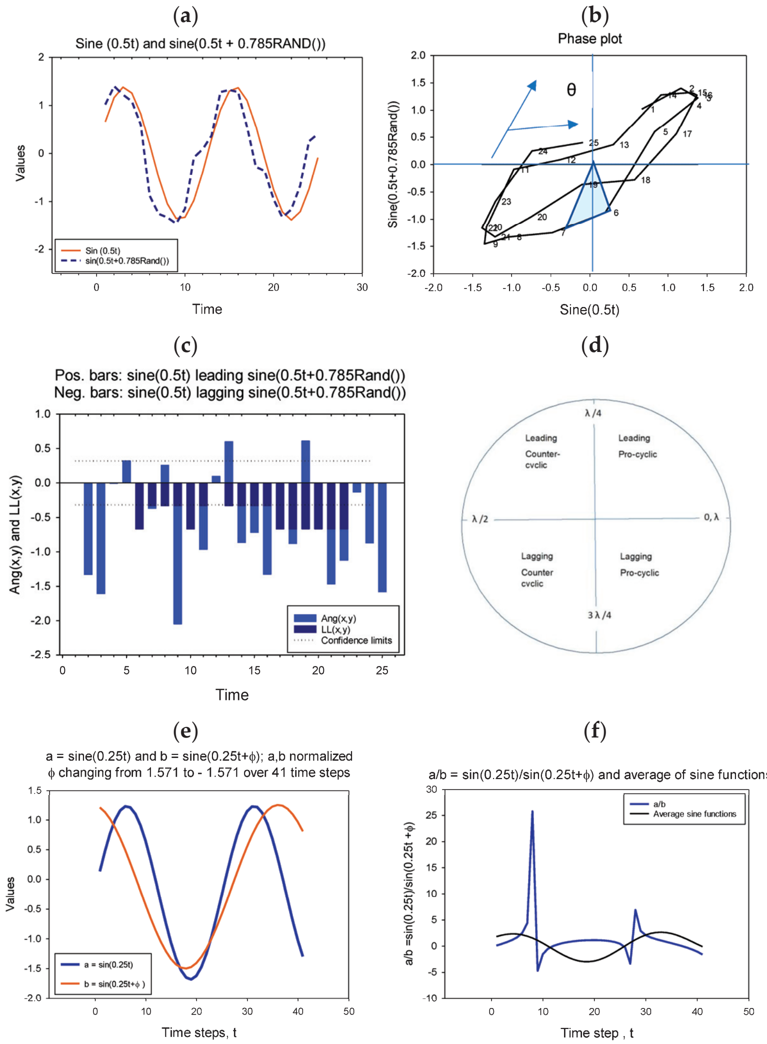

Figure 1. (**a**) Example: Calculating leading-lagging (LL) relations and LL-strength. Two sine functions: the smooth curve is a simple sine function, sin (0.5t), the dashed curve has the form sin (0.5t + ϕ × RAND()) where ϕ = +0.785 (**b**) In a phase plot with sin (0.5t) on the x-axis and the sin (0.5t + ϕ RAND()) on the y-axis, the time series rotates counterclockwise; θ is the angle between two consecutive trajectories. The wedge suggests the angle between the origin and lines to observations 6 and 7. (**c**) Angles between successive trajectories (light blue bars) and LL strength (dark blue bars). Dashed lines suggest confidence limits for persistent rotation in the phase plot and persistent leading or lagging relations in the time series plot. (**d**) LL relations and pro cyclic/counter cyclic relations between two cyclic series as a function of the phase shift between them. (**e**) Two sine functions; sin (0.25t + ϕ), the blue function has ϕ = 0 and the red function has ϕ shifted gradually from a positive to a negative value, thus the two functions shift in being a leading function. (**f**) The ratio a/b = sin (0.25t)/sin (0.25t + ϕ) as a function of time and the average of the two sine functions. (**a**–**d**) are redrawn after Seip and Grøn (2019) and Seip and Wang (2022).

Smoothing. We use the LOESS-smoothing algorithm. The algorithm has two parameters: the parameter (f), which shows how large fraction of the series is that is used as a moving window; and the parameter (p), which shows the polynomial degree used for interpolation. We always use p = 2. With 540 months ≈ 45 years of observations and f = 0.1, the moving time window is ≈50 time steps. We use the LOESS-smoothing algorithm as implemented in SigmaPlot, but the algorithm is implemented in many statistical packages. Since we always use the parameter p = 2, we use the nomenclature LOESS(f) for LOESS smoothing.

Principal component analysis (PCA). Both the data series for the US economy and the data that were extracted for time windows around the recession periods were analyzed with PCA and presented in PCA loading and score plots. However, we only have six recession periods, so we present the main results also as scattergrams to see if outliers affect the regression results. All data for the US economy were LOESS(0.1)-smoothed to avoid sharp peaks that could cause failures in the PCA algorithm. The PCA calculates new variables that are orthogonal.

4. Results

We first show the initial preparation of the observed data, and thereafter the LL results for GDP, EM, and UE. Third, we show the results of embedding the portion of the time windows where GDP leads EM, and the portion of the time windows where β_E (GDP/EM) is less than 0.5 in a "map" of the US economy. Recall that with the time series centered and normalized to unit standard deviation, graphs for x = GDP and y = EM will form an elliptic form with the major axis in the 1:1 direction. Therefore, we evaluate the anomalies from the $\beta_E = 1.0$ line. Last, we examine how labor productivity relates to recession characteristics.

4.1. Data Preparation

To extract information from GDP and UE, GDP will most often have to be detrended. Mukoyama and Sahin (2009) use the Hodrick–Prescott (HP) filter to extract a trend. Ziegenbein (2021) uses a quadratic deterministic trend for GDP and EM. For the LL and β_E-analysis, we did not detrend the data since the LL method we use can be applied to time series with similar trends. However, we center and normalize the data to unit standard deviation. With the high-resolution LL method, we can identify LL relations over three synoptic observations in the paired series (nine time steps to obtain the confidence interval). Other studies use either the whole series or time windows that are considerably longer than ours, e.g., decadal scales, (Cazes et al. 2013; Donayre and Panovska 2021).

4.2. LL Relations

In Figure 2a, the upper two series show GDP and EM time series centered, normalized and slightly LOESS(0.1)-smoothed. The lower line shows the $\beta_E(9)$ coefficients for an OLR over a rolling 9-month time window. The β_E coefficient shows characteristic peak anomalies for paired cyclic curves where one curve is sifted in time relative to the other. The droplines show when the NBER recessions starts. Figure 2b shows the LL relations for the GDP and EM series in Figure 2a. The dark grey bars show θ(3) and the light grey bars show LL(9). GDP leads EM pseudo-significantly 58% of the time and GDP leads EM pseudo-significantly 34% of the time. Table 1 shows the number of months during a 12-month period where θ(3) < 0. Figure 2c shows the residual time series for LOESS(0.8)-smoothed GDP. This series and the UE series are, in addition, slightly LOESS(0.1)-smoothed to avoid sharp peaks.

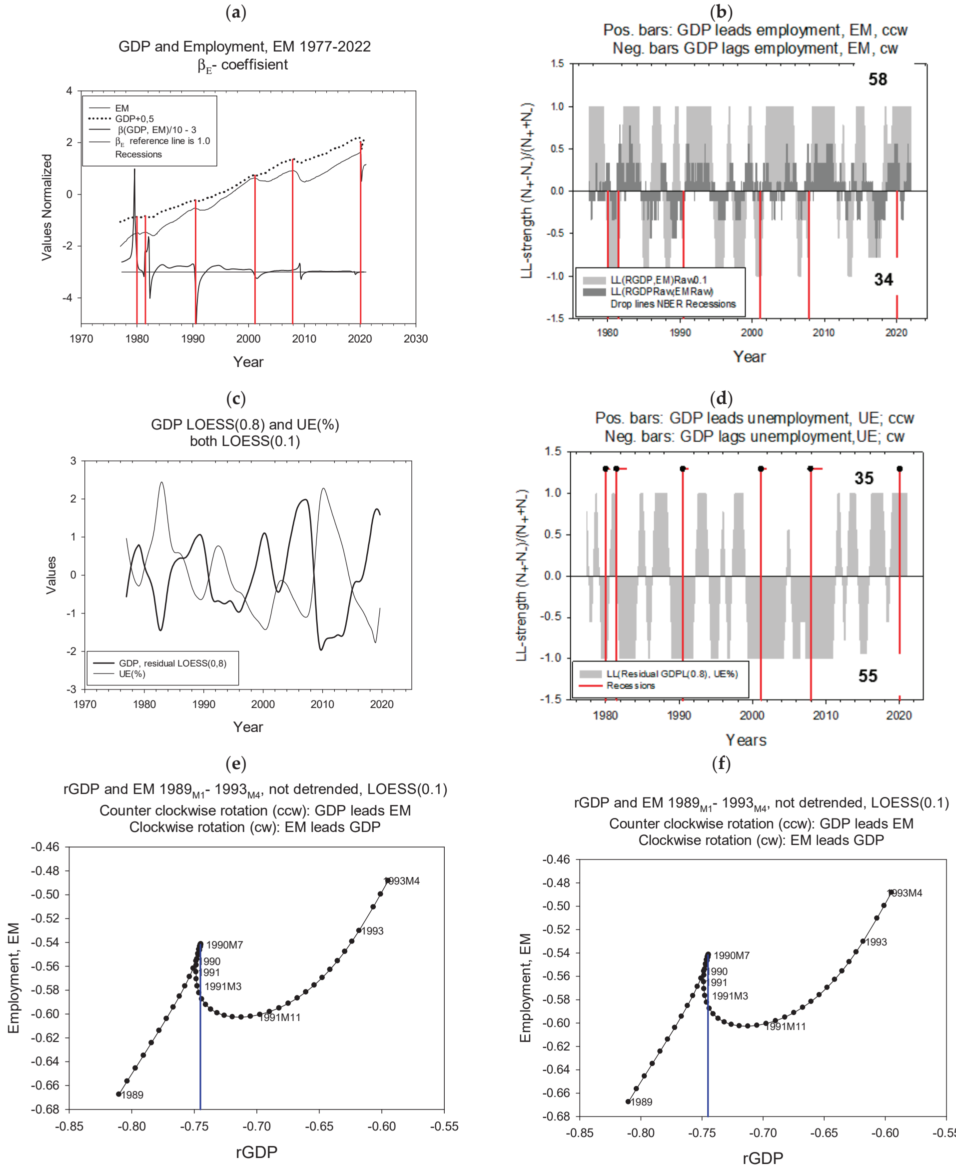

Figure 2. Lead-lag relations between GDP, Employment, EM and unemployment, UE. (**a**) GDP and EM, both LOESS(0.1)-smoothed. Drop lines show NBER recession. (**b**) LL relations between GDP and EM, both LOESS(0.1)-smoothed (grey) and raw, unsmoothed (dark grey). Numbers show percentage "pseudo significant" LL relations (see text). Drop-lines show the beginning of recessions. OLR between grey and black bars give R = 0.30, $p < 0.001$. (**c**) GDP-LOESS (0.8) residual and UE (%) both series LOESS(0.1)-smoothed. (**d**) LL(GDP LOESS (0.8) residual, UE), both series LOESS(0.1)-smoothed and normalized to unit standard deviation. GDP leads UE during the periods 1997–1991; 1993–1995, 1997–2008; 2011–. Red horizontal lines are recession periods. Droplines shows beginning of recessions. (**e**) Phase plot for GDP and EM, not detrended, 1989–1993M3, dropline show the beginning of the 1990 recession. (**f**) Same as (**e**), but with GDP and UE.

Figure 2d shows the LL relations between the two series. GDP leads UE pseudo-significantly 35% of the time and GDP lags UE pseudo-significantly 55% of the time. The LL(GDP,UE) relation shows that UE may lead GDP both before and after a recession.

The LL relations for GDP and EM could in principle be the inverse of the LL relations for GDP and UE. To see how the differences would show up in a phase plot, we plotted the GDP/EM series and the GDP/UE series for time windows around the 1990 recession in two phase plots for the years 1989M1 to 1993M4 (Figure 2e,f). The drop lines show the timing of the July 1990 recession. The graph shows both associations and rotational directions. The results for the GDP/EM pair are shown embedded in a "map" of the US economy in Figure 3a. The blue lines show that the recessions, with the 2020 recession as an exception, are associated with a leading relation for EM to GDP (EM $\rightarrow$ GDP), but EM also leads GDP during other periods. Figure 2f shows that higher GDP generally leads to lower UE. However, there is no discontinuity around the 1990 recession as it was with the GDP/EM pair.

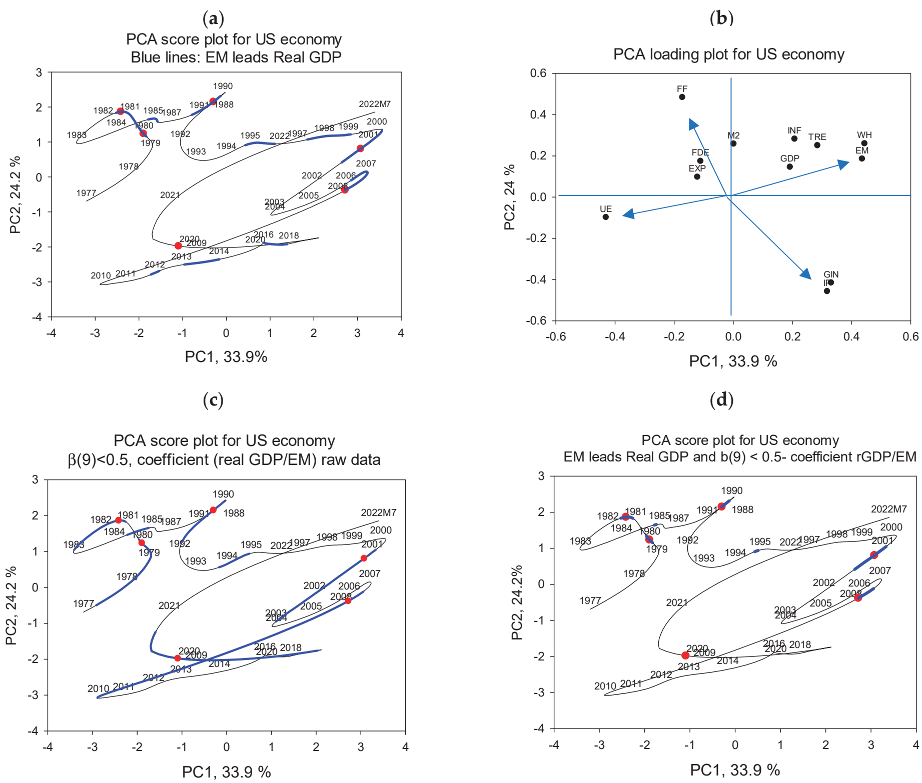

Figure 3. Principal component plots of US Economy 1977 to 2022. (**a**) Score plot for US economy. Blue lines show when employment leads real GDP: Red dots shows US NBER recessions. (**b**) Loading plot for US economy. (**c**) Same as a, but $\beta_E(9) < 0.5$. See text. (**d**) same as (**a**), but both EM leads GDP and $\beta_E(9) < 0.5$. UE = unemployment, FDE = federal debt, EXP = federal expenditures, FF = fed's interest rate, M2 = monetary supply, M2, INF = inflation (we use the consumer price index), GDP = gross domestic product, TRE = tax receipts, GIN = Gini's index, but reconstructed from Mukoyama and Sahin (2009), IP = industrial production, WH = working hours, EM = employment. $\beta_E(9)$ coefficient is for running GDP/EM.

4.3. The β_E Coefficient

The β_E coefficients are calculations over 9 months. Figure 2a shows that the β_E coefficient is generally greater than 1.0 before a recession and then becomes less than 1.0 for a period of 4 to 52 months (21 $\pm$ 17 months) after the recession. Table 1 shows the values of the β_E *coefficient* when it has its lowest value around the recession periods. In Figure 3c, the result is embedded in the "map" of the US economy. The distribution of the blue lines suggests that the period with the β_E coefficient less than 0.5 is mostly associated with recession periods. Figure 3d shows that time windows where both EM leads GDP and $\beta_E(9)$ is less than 0.5 are concentrated around the recessions. The 2020 recession is an exception.

4.4. US Economy Results

The "map" of the US economy seen in the loading plot on Figure 3b provides a reasonable stylized picture of an economy. There are some obvious relations. EM and UE point in opposite directions and align roughly with the x-axis. FF and IP point in opposite directions and align roughly with the y- axis. GDP, EM, WH and TRE all have high values at the same time. However, the close connection between IP and GIN may not be intuitive. A PCA loading plot of cyclic series will have some other characteristics relative to series with Gaussian distributions. If two identical and perfect sine functions are shifted a quarter of a cycle length ($\lambda/4$) relative to each other, their symbol representation will be connected to the origin with lines that are at 90^{o} to each other (Seip and Grøn 2019). Thus, FF and EM may have similar cycle periods, but will be shifted $\lambda/4$ relative to each other.

Recessions appear to occur in all parts of the "map", but they are appearing after sharp bends in the trajectories. For the 1980, 1981 and 1990 recessions, trajectories are counterclockwise, whereas the trajectories rotate clockwise for the 2000, 2007 and 2020 recessions. The two series showing that EM $\rightarrow$ bGDP and $\beta_E(9) < 0.5$ coincide in 34 months out of 547 months (Figure 3d). The 34 months partly preceded and partly followed five of six NBER recession dates, providing a probability of $\approx$0.0002 to coincide with the recessions by chance.

4.5. Labor Productivity

Labor productivity is generally regarded as a leading variable to the business cycle, (Abel et al. 1998, p. 321). Figure 4a and Table 1 show that the detrended LP is declining one year before most recessions, except the 2008 recession. On a decennial time scale, LP shows an overall decrease during the period 1977 to 1997 (the stagflation period under Volcker $\approx$ 1975–1985 and "the great moderation" period under Greenspan, $\approx$1985–1997, (McNown and Seip 2011) then increases until 2010 and decreases again until 2018. For the 12 months preceding a recession, the β_E coefficient is zero per definition for linearly detrended series, but it is on average (minus) 0.10 $\pm$ 0.22, and all β_E coefficients are either negative or close to zero. These values can be compared to the average 12-month increase of 0.28 in LP during the period 1996 to 2010.

We compared the five recession characteristics: recession depth and duration, LP depth and duration, and LP. However, we have only six sets of observation series, suggesting that outliers may play a dominant role. Figure 4b shows that outliers actually play a role for jobless recovery depth versus recession depth. Excluding the recessions in 2007 and 2020, it appears that jobless recovery depth is associated with recession depth, but the regression is not significant, and neither were the other five ((4 $\times$ 3/2) $-$ 1) regressions.

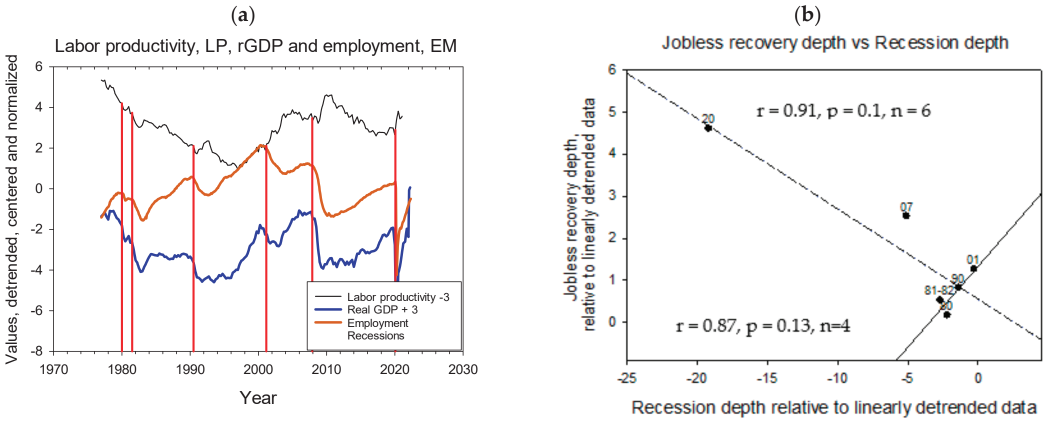

Figure 4. Jobless recovery and recession depth. (**a**) Time series for linearly detrended labor productivity, GDP, and employment. Red droplines show recessions. (**b**) Jobless recovery depth as a function of recession depth. Separate regressions for all recessions and a subset where the two last recessions, 2007 and 2020, are excluded.

5. Discussion

There is no firm conclusion about how the LL relations between GDP, UE or EM should be. However, changes in LL relations between GDP, UE and EM may be associated with the recession periods. Not all recessions are similar, and our study shows that the recent 2020 recession caused by the COVID-19 pandemic was different. The time window around a recession can been divided into four periods: first an expansion in GDP, then two periods during the downturn in GDP, and then a second expansion period where the economy recovers.

The first expansion is the one that triggers a monetary or financial response because the economy is overheating, e.g., a rise in FF (Taylor and Williams 2010). The first recession period is characterized by an uncertainty in the reality of the recession. During the third period, the recession is a certainty. The fourth period is the expansion period that terminates in “natural” growth and employment rates.

5.1. Lead-Lag Relations

A leading role for a causal effect is a prerequisite, but not a sufficient, criterion for causation. However, the leading role is often offered as a strong argument for a causal effect (Sugihara et al. 2012). We examined the LL relations between the two pairs: real GDP and EM, and GDP (detrended) and UE (percentages). Our results for EM support Hamilton’s (2018, p. 838) finding that the cyclical component of EM starts to decline before the NBER business cycle peak for essentially every recession (here, before 1981, 1990, 2001 and 2007 recessions). We also found that whereas EM leads GDP before a recession, it lags GDP after the recession.

Figure 2e,f showed two characteristics for paired time series during the period 1989 to 1993, that is, around the 1990 recession. The first shows how EM and UE vary with GDP (slopes in the graph), and it shows which of the variables are leading the others (rotational directions). The graphs confirm the LL—the results are in Figure 2b,d. In the following discussion, we use traditional “stylized facts” to explain the time series movements we observe. Following that, we will discuss possible causal factors examined in the literature.

Before July 1990M7, the first expansion, EM increases as GDP increases, and the rotational direction is clockwise (CW), but not easy to identify visually (Figure 2e). The graph confirms the LL results in Figure 2b. The interpretation could be that firms fill vacant positions as production or demand for services increases, but the overall US labor productivity anomaly decreases (from a linear trend) during the period one year before the 1990M7 recession. We did not find a first recession period where business was slow to cut workers. (Note that EM increases before 1990M1 (Figure 4a), whereas Figure 3a,b shows a movement towards higher UE at the same time). However, there are 12 variables in the PCA plot that show a movement towards a higher UE, so there are probably other variables that are responsible for the movement, e.g., working hours, as suggested by Obst (2020, p. 229) Following the trajectories in Figure 2e, it is seen that during the whole recession period from 1990M7, both GDP and EM decrease, but whereas GDP increases after March 1991M3, EM increases just slowly after November 1991M11 and is not up to pre-recession values before November 1992M11. Thus, for one year there is a "jobless recovery" during the second expansion period. During this expansion (the recovery), firms are slow to rehire as they use current workers more intensively. Rotations are counterclockwise, showing that GDP is leading EM from 1990M9 to 1993M4 (outside graph: to 1994M10). The labor production anomaly increases from 1991M1 to 1992M7. We have not discussed the effects of an increase in the underground economy during a recession, nor its potential impact on the recovery process. However, it may contribute a significant amount in countries with large underground economies, e.g., Seip and Orsi (2022).

For the GDP–UE pair (Figure 2f), GDP decreases and UE increases, rotations are clockwise, and UE leads GDP. However, it is the trough in unemployment that would imply a peak in GDP, thus, the interpretation should be that (minus) UE apparently leads GDP or GDP leads UE, i.e., it is business growth that creates hiring.

Our results are based on EM and UE data downloaded in September 2021, however Ahn and Hamilton (2022) conclude that current unemployment measures underestimate the number of people that are unemployed, and the magnitude of the bias is larger when the true unemployment rate is higher). Ahn and Hamilton (2022)' s revised UE series running from 2001 to 2020 showed similar LL(GDP, UE) relations as in Figure 2d, except that from 2006 to 2010 GDP was leading UE. Mimicking the verbal Ahn and Hamilton assessments for the whole period of 1977 to 2020 and replacing UE with UE1.1, the results for the period did not change the LL(GDP, UE) patterns appreciably.

In contrast to the LL relations for GDP and EM, Elhorst and Emili (2022) found that for the Netherlands, output growth leads UE. However, the results reported here apply to employment and not to unemployment, and Figure 2e,f and Appendix A show that EM and UE provide different results when put into similar contexts.

5.2. The β_E-Coefficient

We chose to calculate the β_E *coefficient* over 9 months. Examining the phase plot for the period 1989M1 to 1992M10 in Figure 2e and following the $\beta_E(9)$ from left to right during the same time window in Figure 2a, the β_E will first show a small positive peak, then after 1990 a sharp negative peak and last a horizontal "no change" trend.

5.3. LL-Relations and the β_E-Coefficient

The following results apply to the five recessions during the period 1977 to 2010, but the COVID-19 recession in 2020 is an exception for all results. Figure 3d shows months where both EM leads GDP and the β_E *coefficient* is less than 0.5, that is, less than the "neutral" slope of 1.0. This result support both our hypotheses. Hypothesis 1 (**H1**), that the EM leads GDP and EM decreases faster than GDP before and during a recession was supported. The NBER recessions come after sharp bends in the trajectories and away from optimal EM and GDP in Figure 3a (NBER recession dates are shown by red dots in Figure 3a–d). The leading relation for EM before a recession was also found by Hamilton (2018). Hypothesis 2, (**H2**), that GDP leads EM and β_E (GDP/EM) decreases before and during an expansion, was also supported (Figure 2a,b). The COVID-19 recession in 2020 was, in contrast to the "classical economy overheating" recessions, initiated by the COVID-19 pandemic and caused at least partly by a supply side shock (Hobbs 2020).

5.4. US Economy

Jobless abatement measures: Hypothesis 3, (**H3)**, that we could infer macroeconomic conditions where leading or lagging relations between EM and GDP were significant, was not supported. Changes in LL relations were closely associated with recessions but distributed over the whole economy "map". However, the finding that the NBER recession dates come after a sharp bend in the trajectories describing the US economy in Figure 3 may offer some clues. Several studies suggest various factors that may explain variations in Okun's β coefficient, such as changes in working hours and labor productivity (Cazes et al. 2013; Elhorst and Emili 2022), nominal wages, (Donayre and Panovska 2021 Okun's beta and wages covary positively, p. 9), labor legislation, (Cazes et al. 2013) and a mismatch in job search (Mukoyama and Sahin 2009; Gimbel and Sinclair 2020). The selection of variables used to build the PCA plots was chosen primarily to construct and validate the score plot for the US economy. We were not able to extract reasonable information from the loading plot in Figure 3b that could be used to associate any of the variables with the "jobless recovery" issue.

5.5. Policy Implications

Observations. First, the leading relations of EM $\rightarrow$ GDP before five of the six recessions and the leading relation of GDP $\rightarrow$ EM after recessions provided clear signatures to the Okun's $\beta_E(t)$ signature. Second, periods before recessions were associated with labor productivity that was less than optimal. Third, we found that months where both EM $\rightarrow$ GDP and Okun's β_E(GDP,EM) were less than 0.5 were closely associated with the five recessions before 2010. The COVID-19 recession in 2020 was an exception. Fourth, months that are common for the two events (34 months) precisely identify recessions, and fifth, jobless recoveries were pronounced for all five recessions.

Interpretations. Since EM leads GDP before a recession, it suggests that hiring is made before business increases. This is supported by a decrease, or slowing down, in labor productivity. During economic recovery, hiring is slower, providing a background for the term "jobless recovery".

Diagnostics predictions and abatements. The 34 months where EM $\rightarrow$ GDP and $\beta_E < 0.5$ matched the recessions, indicating that some of the months precede the recession. Thus, inspecting the two series: (i) LL relations between GDP and EM and (ii) Okun's law in the format β_E = GDP/EM, may help us to predict a recession or support that a recession is real. However, the COVID-19 recession in 2020 provided an important exception to this "rule", perhaps due to the different nature of this atypical recession that was caused by a pandemic.

We could suggest two abatement measures for jobless recoveries. The first is that management should not hire workers during expansion periods that show a sign of a coming recessions. An assessment of the "natural" rise in labor productivity (the trend) is probably not easy to identify but examining labor productivity anomalies may still offer clues as to when new hirings should be made with care. The second measure is to implement legal procedures that make hiring and firing employees less easy. However, employment policy is a multicriteria question. The utility of being temporarily employed before a recession may be higher than the disutility of not being rehired for several months during a recession recovery (see, e.g., (Ball 2015)).

The results were not realized when EM was replaced by UE and real GDP by GDP. For a rationale, see Appendix A.

5.6. Further Studies

The present study addresses unemployment issues in terms of LL relations and OLR regressions between GDP/EM and time. Our results refer to the US economy, but a comparison with other countries along the same lines would be interesting, for example, comparing Okun's $\beta_E(9)$ coefficient in the US with the corresponding $\beta_E(9)$ coefficients in

Germany and the Netherlands. Volatilities in employment are largest in the US (0.0389), lower in the Netherlands (0.0232) and the lowest in Germany (0.0115).

We found that UE both leads and lags GDP, which would complicate shifting one series relative to the others to identify LL relations for the whole series (Figure 2d). For example, Obst (2020) includes two quarter delays between GDP and UE in analyzing the relation between GDP and UE. Figure A1a in Appendix A shows that even within time windows where the LL relations are consistent, cycle periods and phase shifts vary.

We also compared the two graphs GDP vs. UE (left in Figure A1) and GDP and EM (right in Figure A1), and it is seen that the graph for GDP and UE shows the overall "traditional" inverse relation between GDP and UE, whereas the GDP and EM graph does not directly reflect that UE and EM should be inverse measures of the same phenomenon. However, a contributing factor may be that if unemployed persons drop out of the labor force, they are no longer counted as unemployed in the unemployment statistics (Elhorst and Emili 2022). Thus, replacing UE with EM in Okun's law appears to change the overall relation beyond the fact that the two variables EU and EM express inverse characteristics of labor participation. However, a discussion of these results is beyond the scope of the present study.

6. Conclusions

Loss of jobs is an important issue and optimum employment is among the three most important goals for the US federal reserve. Okun's law, as expressed by its β coefficient, is traditionally formulated by the ratio between GDP and the UE gap, and over the whole period of study or for decennial time scales. Inspired by the results of Hamilton (2018), we replaced the normal measure of unemployment (UE) with employment (EM) and calculated a $\beta_E(9)$ = real GDP/EM. We found that EM became a leading variable to GDP before a recession, whereas GDP became a leading variable to EM after a recession. The shifts in the leading relation between GDP and EM around a recession also create peaks in the $\beta_E(9)$ coefficient around most "classical" recessions (the 2020 recession caused by the COVID-19 pandemic was an exception). All the recessions were also characterized by a decrease in labor productivity relative to an "optimal" but realized 1997–2010 value. Our findings of the leading role of EM to GDP before a recession and the loss in labor productivity suggest that hiring employees during periods with a heated economy may cause subsequent jobless recovery. To alleviate jobless recoveries, one should make hiring decisions with caution when there is a sign of a coming decision or labor productivity anomalies.

Author Contributions: Conceptualization, K.L.S., Methodology, K.L.S.; software, K.L.S.; validation K.L.S. and D.Z., Formal analysis, K.L.S. and D.Z., Investigation, K.L.S. and D.Z.; Resources, K.L.S. and D.Z.; Data curation, K.L.S. and D.Z.; Writing original draft, K.L.S. Review and Editing D.Z.; Visualization, K.L.S., Supervision, K.L.S. and D.Z.; Project administration, K.L.S. and D.Z.; calculations, K.L.S. and D.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Oslo Metropolitan University.

Data Availability Statement: All data and all calculations are available from the author.

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

Appendix A

The appendix contains: Comparing rGDP to EM and UE.

Comparing rGDP to EM and UE

The EM were significantly shifted 1.6 time steps backward relative to rGDP in 16*% of the time, and EM was significantly shifted forward 1.7 time steps relative to rGDP 33 % of the time.

(**a**) (**b**)

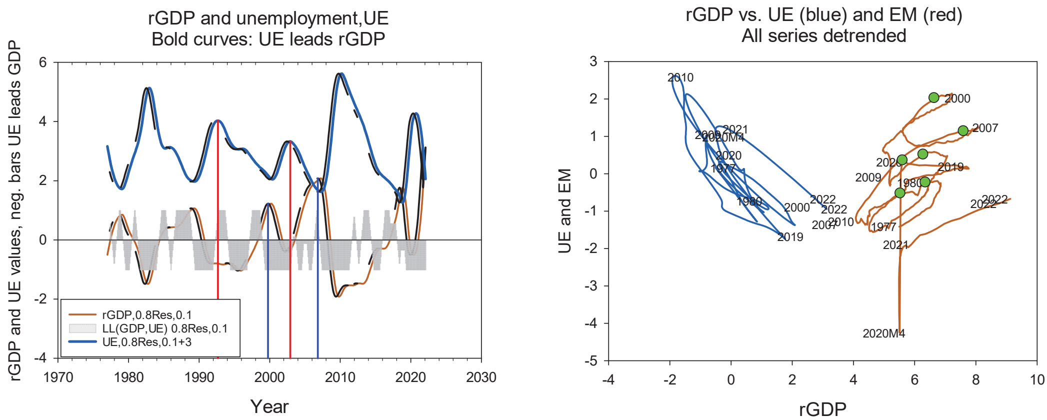

Figure A1. GDP and unemployment. (**a**) time series in bold shows part of the series where UE leads GDP. Distance between the two first peaks (red and blue dropdown lines) are 86 months and distance between the two next peaks are 47 months. (**b**) GDP versus UE and GDP versus EM. Green circles show NBER recessions.

References

Abel, Andrew, Ben Bernanke, and Robert McNabb. 1998. *Macroeconomics*. Harlow: Addison-Wesley.

Ahn, Hie Joo, and James D. Hamilton. 2022. Measuring labor-force participation and the incidence and duration of unemployment. *Review of Economic Dynamics* 44: 1–32. [CrossRef]

Ball, Laurence. 2015. *Monetary Policy for a High-Pressure Economy. Policy Futures*. Washington: Center on Budget and Policy Priorities, p. 13.

Burns, Arthur F., and Wesley C. Mitchell. 1946. *Measuring Business Cycles*. New York: National Bureau of Economic Research.

Camacho, Maximo, and Gabriel Perez-Quiros. 2010. Introducing the Euro-Sting: Short-Term Indi-cator of Euro Area Growth. *Journal of Applied Econometrics* 25: 663–94. [CrossRef]

Capsada-Munsech, Queralt, and Oscar Valiente. 2020. Sub-National Variation of Skill For-mation Regimes: A Comparative Analysis of Skill Mismatch Across 18 European Regions. *European Education* 52: 166–79. [CrossRef]

Cazes, Sandrine, Sher Verick, and Fares Al Hussami. 2013. Why did unemployment respond so differently to the global financial crisis across countries? Insights from Okun's Law. *IZA Journal of Labor Policy* 2: 1–18. [CrossRef]

Donayre, Luiggi, and Irina Panovska. 2021. Recession-specific recoveries: L's, U's and everything in between. *Economics Letters* 209: 110145. [CrossRef]

Elhorst, J. Paul, and Silvia Emili. 2022. A spatial econometric multivariate model of Okun's law. *Regional Science and Urban Economics* 93: 103756. [CrossRef]

Galí, Jordi, Frank Smets, and Rafael Wouters. 2012. Slow recoveries: A structural interpretation. *Journal of Money, Credit and Banking* 44: 9–30. [CrossRef]

Gimbel, Martha E., and Tara M. Sinclair. 2020. Mismatch in Online Job Search. Paper presented at the LERA 2020 Meetings, Virtual, June 13–16. Sa Diego: Labor and Employment Relations Assocciation (LERA).

Hamilton, James D. 2018. Why You Should Never Use the Hodrick-Prescott Filter. *Review of Economics and Statistics* 100: 831–43. [CrossRef]

Hobbs, Jill E. 2020. Food supply chains during the COVID-19 pandemic. *Canadian Journal of Agricultural Economics-Revue Canadienne D Agroeconomie* 68: 171–76. [CrossRef]

Krüger, J. Jens. 2021. A Wavelet Evaluation of Some Leading Business Cycle Indicators for the German Economy. *Journal of Business Cycle Research* 17: 203–319. [CrossRef]

Lazear, Edward P., and James R. Spletzer. 2012. *The United States Labor Market: Status quo or a New Normal? The Changing Policy Landscape*. Jacson Hole: The Federal Reserve Bank of Kansas City.

Maza, Adolfo. 2022. Regional Differences in Okun's Law and Explanatory Factors: Some Insights from Europe. *Interna-tional Regional Science Review* 45: 016001762210823. [CrossRef]

McNown, Robert, and Knut Lehre Seip. 2011. Periods and structural breaks in US economic history 1959–2007: A data driven identification. *J. Policy Modelling* 33: 169–82. [CrossRef]

Mukoyama, Toshihiko, and Ayşegül Sahin. 2009. Why did the average duration of unemployment become so much longer? *Journal of Monetary Economics* 56: 200–9. [CrossRef]

Obst, Thomas. 2020. Dynamic version of Okun's law in the EU15 countries—The role of delays in the unemployment-output nexus. *Scottish Journal of Political Economy* 69: 225–41. [CrossRef]

Seip, Knut L., and Hui Wang. 2022. The North Atlantic Oscillations: Lead–Lag Relations for the NAO, the AMO, and the AMOC—A High-Resolution Lead–Lag Analysis. *Climate* 10: 63. [CrossRef]

Seip, Knut L., Øyvind Grøn, and Hui Wang. 2018. Carbon dioxide precedes temperature change during short-term pauses in multi-millennial palaeoclimate records. *Palaeogeography Palaeoclimatology Palaeoecology* 506: 101–11. [CrossRef]

Seip, Knut L., Yunus Yilmaz, and Michael Schroder. 2019. Comparing Sentiment- and Behavioral-Based Leading Indexes for Industrial Production: When Does Each Fail? *Economies* 7: 104. [CrossRef]

Seip, Knut L., and Øyvind Grøn. 2019. Atmospheric and Ocean Dynamics May Explain Cycles in Oceanic Oscillations. *Climate* 7: 77. [CrossRef]

Seip, Knut Lehre, and Renzo Orsi. 2022. Effects of tax changes on GDP and underground economy: Italy 1982 to 2006. *submitted*.

Søgner, Leopold. 2001. Okun's law. Does the Austrian unemployment-GDP relationship exhibit structural breaks? *Empirical Economics* 26: 553–64. [CrossRef]

Sugihara, George, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. 2012. Detecting Causality in Complex Ecosystems. *Science* 338: 496–500. [CrossRef]

Taylor, John B., and John C. Williams. 2010. *Simple and Robust Rules for Monetary Policy*. Working Papers. San Francisco: Federal Reserve Bank of San Francisco.

Vaubel, Roland. 2008. The political economy of labor market regulation by the European Union. *Review of International Organizations* 3: 435–65. [CrossRef]

Yellen, Janet L. 1991. Comments and discussion. *Brookings Papers on Economic Activity* 2: 127–33.

Ziegenbein, Alexander. 2021. Macroeconomic shocks and Okun's Law. *Economic Letters* 202: 1–4. [CrossRef]

economies

Article

The Industrial Pattern of Italian Regions: A Disaggregated Sectoral Analysis Based on Input–Output Tables

Clio Ciaschini [1], Margherita Carlucci [2], Francesco Maria Chelli [1], Giuseppe Ricciardo Lamonica [1] and Luca Salvati [3,*]

[1] Department of Economic and Social Sciences, Polytechnic University of Marche, 60121 Ancona, Italy
[2] Department of Social and Economic Sciences, Sapienza University of Rome, 00185 Rome, Italy
[3] Department of Methods and Models for Economics, Territory and Finance, Sapienza University of Rome, 00185 Rome, Italy
* Correspondence: luca.salvati@uniroma.it

Abstract: Italy joined the so-called 'Industry 4.0' European framework in 2016, which designed and approved a national plan to regulate this key issue for regional development. To better support such a framework, the present study attempts to quantify the contribution of the Italian regions to the output formation process. More specifically, a multi-sectoral Input–Output (IO) model that supports national policies was proposed to cumulatively consider 29 industries that partition the Italian economy into representative branches at the level of administrative regions. Elementary input data were derived from the inter-sectoral table of the economy released by the Italian National Institute of Statistics (ISTAT). The economic outcomes of the Italian regions were estimated using a non-survey procedure, based on Flegg Location Quotients, to determine the upstream and downstream positions of each industry at country and regional levels. Indices grounded on the Hypothetical Extraction Method (HEM) further delineated the role each industry plays in the regional economy. The empirical findings of this study demonstrate how non-survey IO regionalization and the resulting industry-based indices provide appropriate knowledge for regional development policies.

Keywords: Input–Output; regionalization; Hypothetical Extraction Method; linkage analysis

Citation: Ciaschini, Clio, Margherita Carlucci, Francesco Maria Chelli, Giuseppe Ricciardo Lamonica, and Luca Salvati. 2022. The Industrial Pattern of Italian Regions: A Disaggregated Sectoral Analysis Based on Input–Output Tables. *Economies* 10: 300. https://doi.org/10.3390/economies10120300

Academic Editors: Ralf Fendel, Sajid Anwar and Robert Czudaj

Received: 3 October 2022
Accepted: 14 November 2022
Published: 28 November 2022

1. Introduction

The COVID-19 pandemic has produced devastating effects, whose consequences depend on the capability of the economic structures to react to the macroeconomic shocks affecting both supply and demand (Li et al. 2021; Notteboom et al. 2021; Pham et al. 2021). The measures imposed by governments to contain the spreading of the virus are examples of strict supply-side constraints (Chen et al. 2021). As soon as the lockdown would end and the containment measures gradually be lifted (Seetharaman 2020), the pre-crisis levels of activity could be expected to recover (Sigala 2020; Maples et al. 2021; Song et al. 2021). These dynamics apply to the economic system as a whole, although the restrictive effects have had different sectoral impacts (Bashir et al. 2020). Such impacts, spread over the whole economy at a magnitude that depends on the relevance of individual sectors, depend on the inter-sectoral relation characteristics of the production system (Hirschman 1958; Dean et al. 2021; Belitski et al. 2022). A comprehensive assessment of the inter-industry relationship characteristics of the Italian economy is the base for a more exhaustive analysis of COVID-19s effect on production sectors (Ascani et al. 2021; Bragatto et al. 2021; Cutrini and Salvati 2021).

At the global level, Italy ranks high within the industrialized countries (Bigerna 2013). The most significant industries that undoubtedly shape and lead inter-industry interactions are manufacturing and, among services, those related to wholesale and retail trade, as well as transport (Archibugi et al. 1991; Cainelli and Leoncini 1999; ISTAT 2016). Manufacturing relies on specialized and high-quality products that are realized by a network of small- and medium-size enterprises (Bertolini and Giovannetti 2006). In the last decades, globalization has further stimulated the expansion of manufacturing (Ghisellini and Ulgiati 2020). However, since 2018, this expansion wave decelerated because of marked instances influencing the stability of international markets (Cainelli et al. 2018; Basso 2020; D'Ingiullo and Evangelista 2020). The inward-looking American commercial policies, the economic outcomes of Brexit, the latent tensions between the USA and China, and the uncertain political outcomes of national elections in many European countries resulted in insecure landscapes for economic interactions. Within this uncertain climate, Italy suffered from political instability, economic stagnation, and a lack of structural reforms (Rubino and Vitolla 2018; Cainelli et al. 2019; Ciffolilli et al. 2019). In fact, even before the Great Recession (e.g., between 2001 and 2007), Italy was growing by less than 1.3% per year, on average. The 2008 crisis exacerbated the weakness of the Italian economy (Da Roit and Iannuzzi 2022). In 2009, its gross domestic product declined by 5.5%, and has been recovering slowly and only partly (Crespí-Cladera et al. 2021). Moreover, the crisis has deepened the divide between affluent and industrialized Northern regions and disadvantaged, agriculture-dependent, Southern regions, which has been particularly evident for gross domestic product, unemployment, and capital formation (Dei Ottati 2018).

Stagnation affects both public investments in infrastructures and private capital formation, notwithstanding the support provided by the incentives to the 4.0 digital conversion of manufacturing (Confindustria Report 2019). However, internal weaknesses and the increasing concentration of industrial development in small districts—mainly in Northern regions—did not prevent Italy from becoming the seventh world manufacturing power by 2018, and the ninth in the world for export capacity (Table 1).

Table 1. Value added and Exports (2018) in percent values of world's total (source: our elaboration on Confindustria Report 2019).

	Value Added		Exports	
1	China	28.5	China	12.8
2	USA	17.2	USA	8.6
3	Japan	8.1	Germany	8.0
4	Germany	6.1	Japan	3.8
5	South-Korea	3.1	The Netherlands	3.7
6	India	3.0	South-Korea	3.1
7	**Italy**	**2.3**	France	3.0
8	France	2.1	Hong Kong	2.9
9	Unted Kingdom	1.9	**Italy**	**2.8**
10	Indonesia	1.6	United Kingdom	2.5

In this context, the 4.0 transformation of manufacturing has played a role in improving firms' efficiency. Technology 4.0 was aimed at fastening decision processes and facilitating new forms of interaction between humans and machines to connect the entire value chain within the firms (Ghisellini and Ulgiati 2020). Italy joined the European framework of Industry 4.0 in 2016, enforcing a National Plan called 'Industria 4.0'. Hyper–amortization investments, estimated to account for 10 billion euros, have been the main measure stimulating firms' development (Da Roit and Iannuzzi 2022).

The regional framework in which economic activities are performed also has relevance, together with the dimensions of firms within the region. Italian regions consist of levels of territorial subdivision and public authorities, i.e., public bodies with legal status and wide autonomy regulated by the Constitution (Salvati and Zitti 2009). The 20 Italian administrative regions (see Figure 1) can be grouped into three macro-areas (Zambon et al. 2017). The North macro-area comprises nine regions: Liguria, Lombardy, Piedmont, Aosta Valley, Veneto, Friuli-Venezia Giulia, Emilia-Romagna, and the two autonomous provinces of Trento and Bolzano in the Trentino-Alto-Adige/Sudtirol region. The Central macro-area includes Tuscany, Latium, Marche and Umbria. The South/Islands macro-area includes Abruzzo, Basilicata, Calabria, Campania, Molise, Apulia, Sicily, and Sardinia (Salvati et al. 2008).

Figure 1. A map illustrating the geography of Italian administrative regions (Source: own elaboration on Eurispes (2022)).

According to a qualified source of regional analysis in Italy, "there is no doubt that the social and economic disunity of Italy still remains nowadays the most apparent and the most ignored structural limit" (Eurispes 2022). The most relevant inconsistency of Italy is that a part of it (corresponding to 41% of the territory) lives in disadvantaged socio-economic conditions. A measure of the relative economic performance of each region can be quantified using per-capita GDP as shown in Table 2, where the 2019 percent gap of the regional per-capita GDP with respect to the country figure is given. In 1951, per-capita GDP in Southern Italy amounted to 53 per cent of country GDP. In 1973, this economic aggregate reached the lowest peak (51 per cent), and that result was never observed again.

Table 2. Per-capita GDP departure (%) of Italian Regions from the national average in 2019 (ISTAT 2020).

Macro-Area	Region	Per-Capita GDP Gap	Macro-Area	Region	Per-Capita GDP Gap
North	Bolzano	+62.2	South	Abruzzo	−11.8
	Aosta Valley	+34.3		Basilicata	−24.6
	Lombardy	+33.9		Sardinia	−27.6
	Trento	+31.5		Molise	−28.8
	Emilia-Romagna	+25.1		Apulia	−35.7
	Veneto	+14.7		Campania	−35.9
	Liguria	+11.2		Sicily	−39.0
	Piedmont	+8.6		Calabria	−41.5
	Friuli-V. Giulia	+8.1			
Center	Latium	+15.8	Italy		29,000
	Tuscany	+8.8			
	Marche	−3.2			
	Umbria	−12.8			

Our contribution sheds light on the role of each of the 29 industries, described in Table 3, characterizing the technological structure of the Italian economy by considering each of the 20 administrative regions separately. Defining and quantifying the role and relevance of each activity allow for a convenient comparison with national dynamics, as well as an appropriate analysis of the role of each activity within and between macro-areas (Lamonica and Chelli 2018). For this aim, on the one hand, we regionalized the Italian Input-Output 2016 Table, with a disaggregation level corresponding to the 29 representative sectors estimated at the level of Italian administrative regions, by using Flegg Location Quotient methodology (Lamonica et al. 2020).

Table 3. A list of industry denominations used in this study.

Industries	Denominations
S1	Crop and animal production, hunting and related service activities
S2	Fishing and acquaculture
S3	Mining and quarring
S4	Manufacture of food products, beverages, and tobacco products
S5	Manufacture of textiles and wearing apparel
S6	Manufacture of wood and of products of wood, paper and paper products and printing
S7	Manufacture of coke and refined petroleum products
S8	Manufacture of rubber, plastic products, and other non-metallic mineral products
S9	Manufacture of fabricated metal products, except machinery and equipment
S10	Manufacture of computer, electronic and optical products, electrical equipment machinery and equipment n.e.c.
S11	Manufacture of transport equipment
S12	Manufacture of furniture, Other manufacturing Repair and installation of machinery and equipment
S13	Electricity, gas, steam, and air conditioning supply
S14	Water collection, treatment, and supply
S15	Construction
S16	Wholesale and retail trade and repair of motor vehicles and motorcycles
S17	Trasportation and storage
S18	Accommodation and food service activities
S19	Information and communication
S20	Financial and insurance activities
S21	Real estate activities
S22	Professional, scientific, and technical activities
S23	Administrative and support service activities
S24	Public Administration and defence; compulsory social security
S25	Education
S26	Human health and social work activities
S27	Arts, entertainment and recreation
S28	Other services activities
S29	Activities of households as employers; undifferentiated goods/services producing activities of households for own use

By performing a non-survey regionalization, this approach avoided time-consuming and costly data collection. On the other hand, with the aim of revealing and quantifying the regional/sectoral role of each activity, we performed a linkage analysis determining the sector potential in the national rank. The specific roles of key production branches in the Italian regions were also determined. Linkage analysis was performed by adopting the most suitable definitions and operational frames for the aims of our study. In particular, we use the outcomes of linkage definition based on the Hypothetical Extraction Method (HEM) approach.

This paper is structured as follows. The next section briefly presents the most relevant literature on linkage analysis. Section 3 illustrates the regionalization technique and the methodology adopted to derive Hypothetical Extraction Method (HEM) coefficients. Section 4 illustrates the empirical results of this analysis. Section 5 discusses the main research findings and concludes the paper.

2. Literature

2.1. Linkage Analysis and the Hypothetical Extraction Method

In the context of multiple industries interacting within the economic system, the relevance of production sectors, especially local ones and their contributions to economic growth, requires a measure that delineates both forward and backward linkages in order to empirically analyze the complete role of an industry and to evaluate the technological connections between economic sectors (Kay et al. 2007). Within this framework, intersectoral linkages are regarded as techno-economic connections between industries that are embodied in the exchange of tangibles and intangibles (Hauknes and Knell 2009). Intersectoral linkages, therefore, estimate interdependencies between sectors, which affect the paths of vertical specialization (Reis and Rua 2009). In the literature on regional economics, different indicators have been suggested and widely adopted for the identification of key sectors. These sectors, characterized by relevant intermediate purchases (backward linkages) and sales (forward linkages), are more likely than the other sectors to propagate growth impulses all over the economy (OECD 2021). Golan et al. (1994) integrated this definition to include three further conditions to appropriately define a key cluster, i.e., (i) a well-developed domestic market, (ii) a competitive local business climate, and (iii) efficient production factors.

This broader definition takes the size of the linkages as an approximation for the potential benefits of stimulating the sector. This should imply that the taxpayer cost of acquiring these benefits would be the same among sectors and among forward and backward linkages. In reality, stimulating wider sectors is more expensive than smaller ones; hence, key sector measures have to be corrected in consideration of their dimensions to more accurately achieve this objective (Temurshoev and Oosterhaven 2014). In addition, industries of the same size do not necessary require similar policy measures. This fact has been considered in the net backward linkages, which correct the standard (gross) backward linkages for the size of final demand, assuming that a relatively wide final demand is more easily stimulated than a small-sized one (Oosterhaven 2004, 2007). Lastly, the creation of benefits of sizeable backward linkages requires demand-stimulating measures, whereas the generation of benefits of wide forward linkages requires a further improvement in productivity, i.e., price reducing policies to strengthen output growth. Obviously, the cost of these quite different policy measures per unit of potential benefit, i.e., per linkage measures, will not be the same. Hence, selecting key sectors requires much more analysis than just establishing which sectors have the largest forward and backward linkages. To address this issue, many key sectors' measures have been proposed in the literature (Golan et al. 1994).

On the one side, the various measures result from methodological enhancements, such as the substitution of the direct backward linkages (Chenery and Watanabe 1958) with the total backward linkages originated by the column sums of the Leontief-inverse (Rasmussen 1956), or the substitution of the row sums of the Leontief-inverse (Rasmussen 1956) with the row sums of the Ghosh-inverse (Beyers 1976), in the case of total forward linkages (Jones 1976). On the other side, these measures originate from different labelling of the same measure in different works. Among these measures, it is worth mentioning the output-to-output multiplier (Miller and Blair 2009), analogous to the total flow multiplier (Szyrmer 1992), is comparable to the earlier Hypothetical Extraction Method (HEM) of whole sectors (Strassert 1968; Schultz 1977; Temurshoev 2010), which was later reformulated by Meller and Marfán (1981), Cella (1984), and Clements (1990).

In the present work, we adopted the latter, straightforward and flexible version of the HEM, since it allows for the extraction of any subset of transactions, instead of a mere removal of full rows and columns, (Miller and Lahr 2001; Gallego and Lenzen 2005). The HEM identifies the 'keyness' of a sector through the hypothetical output loss in the economic system due to the abrupt stop of the related activity, i.e., assuming all sales to (and purchases from) the other sectors are set to zero. The Hypothetical Extraction Method (HEM) was adopted in this study to assess the position occupied by various economic sectors within a given (country or regional) economy. This method is regarded as an improvement of the Classical Multiplier Method (Rasmussen 1956), which measures the 'keyness' of a sector only in terms of simple averages of direct and indirect technical coefficients. HEM, indeed, weights the 'keyness' of a sector by assuming its external linkages, i.e., sales and purchases from all other sectors, as null (Guerra and Sancho 2010).

The output loss deriving from this extreme condition quantifies the underlying network of economic linkages (Miller and Lahr 2001) and provides a measure of 'keyness' (Miller and Blair 2009). Therefore, the HEM evaluates the extent at which the total output of the economy would change (e.g., decrease) if a j-th sector is removed from the economic system.

The bulk of this approach lies in the inverse Leontief matrix, i.e., $L = (I - A^n)^{-1}$ in the first case and $L = (I - R)^{-1}$ in the second case, where A^n and R, are the matrices of national and regional direct input coefficients, respectively. The generic L_{ij} entry of the L matrix measures the total requirement (multiplier), both direct and indirect, of goods and services produced by the i-th industry needed to satisfy one unit of final use of the j-th sector. Consequently, the j-th column-sum (L_j) of L measures the total requirements of the j-th sector to produce one (final) production unit. In other words, it is the extent to which a unitary increase in the final demand of the j-th sector causes a production increase in all sectors. On the contrary, the row-sum of the L matrix ($L_i.$) measures the total production requirements of the i-th sector needed to off-set a unitary increase in the final uses of each product. In other words, output magnitude increases in the i-th sector if the final demand of all sectors increases by one unit. Initially, this was modelled in an input–output context by deleting row and column j from the A matrix of the technical coefficients (Ali et al. 2019).

To this regard, let $\overline{A}_{(j)}$ be the $(k-1) \times (k-1)$ matrix without the sectors j and $\overline{f}_{(j)}$ in the correspondingly reduced final demand vector (see Miller and Blair 2009 for details), then, the total output in the 'reduced' economy reads as $\overline{x}_{(j)} = \left[I - \overline{A}_{(j)}\right]^{-1} \overline{f}_{(j)}$. On the contrary, in the full k-sector model, the total output is $x = [I - A]^{-1} f$. Consequently, $\boldsymbol{i}'x - \boldsymbol{i}'\overline{x}_{(j)}$ (where $\boldsymbol{i}$ is a column vector of ones) is an aggregate measure of the economy loss (reflecting a decrease in gross output value) if sector j disappears, and is in turn an indirect, overall estimate of the multi-dimensional linkages of the j-th sector.

Normalization by total gross output ($\boldsymbol{i}'x$) and multiplication by 100 provides an estimate of the percent loss in total economic activity. The hypothetical extraction approach can also be used to measure backward and forward linkage components separately. We assume that the j-th sector buys no intermediate inputs from any production sector by replacing the j-th column in A with zeroes. Then, the following index is a candidate measure of (aggregate) backward linkage for the j-th sector:

$$\overline{BL}_j = \frac{\boldsymbol{i}'x - \boldsymbol{i}'\overline{x}_{(j)}}{\boldsymbol{i}'x} 100 \tag{1}$$

Similarly, replacing the i-th row of the output coefficients matrix ($B = \left[\frac{x_{ij}}{x_i}\right]$) with zeroes and denoting this matrix as $\overline{B}_{(I)}$ makes $\overline{x}'_{(i)} = v'\left[I - \overline{B}_{(i)}\right]^{-1}$ the row vector whose entries are the sectoral production when the i-th sector is removed from the economy, i.e., the total production of all other sectors if the i-th sector sells no intermediate input. An aggregate measure of a given sector (i) forward linkage is:

$$\overline{FL}_i = \frac{\boldsymbol{i}'x - \overline{x}'_{(i)}\boldsymbol{i}}{\boldsymbol{i}'x} 100 \tag{2}$$

For the sake of comparison, the previous indices were normalized as follows:

$$\overline{\overline{BL}}_j = \frac{\overline{BL}_j}{\frac{1}{k}\sum_{j=1}^{k} \overline{BL}_j} \tag{3}$$

$$\overline{\overline{FL}}_i = \frac{\overline{FL}_i}{\frac{1}{k}\sum_{i=1}^{k} \overline{FL}_i} \tag{4}$$

The index reported in Equation (1), known as the 'backward linkage' (or 'dispersion power'), measures the activation degree of a given economic sector. Values greater than 1 indicate the importance of a given sector in the regional economy, because it requires a production level from the other sectors above the average. By contrast, the more the index falls below 1, the less important the sector considered is.

The index reported in Equation (2), known as the 'forward linkage' (or 'dispersion sensitivity'), measures the level at which the output of one sector is used as input for the remaining production sectors, and thus measures the degree of reaction characteristic of a given economic sector. As in the previous case, the greater the index is than 1, the more important the corresponding sector is, because it supplies its production to the other sectors at a level which exceeds the general average. By contrast, the more the index falls below 1, the less important the sector considered is. The joint analysis of these two indices makes it possible to determine how an individual sector is woven into the economic structure of a region, as well as its relative importance. Based on these premises, we define:

- 'Key Sectors', with values of both forward and backward linkages higher than 1;
- 'Low Impact' sectors, with both forward and backward linkages lower than 1;
- 'Prime Vendor' sectors, with forward linkages higher than 1 and backward linkages lower than 1;
- 'Prime User' sectors with forward linkages lower than 1 and backward linkages higher than 1.

2.2. Regionalization of the Italian Input-Output Table

Regionalized Input–Output matrices make it possible to perform a linkage analysis that compares the importance of each sector in providing and buying goods and services for the remaining sectors. Unfortunately, the Italian Institute of Statistics (ISTAT) does not provide regionalized Input-Output Tables. As remarked by Hewings and Jensen (1988), non-survey methods regionalizing a national Input-Output Table (NIOT) have been developed, with the aim of avoiding the huge costs and considerable release delays associated with the construction of regional tables through direct surveys. Non-survey methodologies were based (i) on Location Quotients (LQs) or (ii) on constrained matrix-balancing approaches. The former methodologies included Simple and Cross-Industry LQs (SLQ and CILQ), along with refinements such as Round (RLQ) formula (Round 1983), Flegg FLQ formula (Flegg et al. 1995; Flegg and Webber 2000), as well as the augmented FLQ (AFLQ) approach (Flegg and Webber 2000). These methods hinged on the assumption that regions and nations employed the same production technology, with the implication that regional input coefficients only differed from their national counterparts for the fact that each region imports goods and services from other regions (Cuello et al. 1992). By contrast, constrained matrix-balancing procedures estimate unknown data from limited initial information and are subject to a set of linear constraints (e.g., Salvati and Zitti 2009). The most popular techniques include RAS and Cross-Entropy (CE) approaches (Schultz 1977) and those based on minimizing squared or absolute differences (Golan et al. 1994). However, such methods are more time-consuming than the LQ-based approach and normally require the solution of a constrained non-linear optimization problem, whereas the LQ-based methods are quick and simple to apply. The present study concentrates on the FLQ method briefly explained in Section 3, since it is one of the best-performing LQ-based approaches (Bonfiglio and Chelli 2008; Flegg and Tohmo 2016).

2.3. Regionalization Methodologies Using Location Quotiens

In this section, we review the most used location quotient (LQ) methods to estimate a Regional Input-Output Table (RIOT) representative of the 20 Italian regions. We used the most recent NIOT released by the Italian Institute of Statistics (ISTAT) for the year 2016. In Table 2, we show the national and regional IOT for an economic system of k sectors in block matrix notation, where:

National Input–Output Table

$$\begin{pmatrix} \mathbf{X}^n & \mathbf{f}^n & \mathbf{x}^n \\ (\mathbf{v}^n)' & 0 & 0 \\ (\mathbf{x}^n)' & 0 & 0 \end{pmatrix}$$

Regional Input–Output Table

$$\begin{pmatrix} \mathbf{X}^r & \mathbf{f}^r & \mathbf{x}^r \\ \mathbf{M}^r & 0 & 0 \\ (\mathbf{v}^r)' & 0 & 0 \\ (\mathbf{x}^r)' & 0 & 0 \end{pmatrix}$$

- $X^n = \left[x^n_{ij}\right]$ is the matrix whose entries are the total internal flows for intermediate use from the i-th sector to the j-th sector at the national level;
- $X^r = \left[x^r_{ij}\right]$ is the matrix whose entries are the flows of intermediate use from the i-th sector to the j-th sector at the regional level, with both i and j sectors located in region r;
- f^n and f^r are the national and regional final demand vectors;
- $M^r = \left[imp^r_{ij}\right]$ is the matrix of imported (intermediate) inputs produced by the i-th sector of the other regions and acquired by the regional j-th sector;
- $(v^n)'$ and $(v^r)'$ are row vectors whose entries are the primary input by sector at the national and regional level.

Moreover, we define $A^n = \left[a^n_{ij} = \frac{x^n_{ij}}{x^n_j}\right]$, $R = \left[r_{ij} = \frac{x^r_{ij}}{x^r_j}\right]$, and $M^r = \left[m^r_{ij} = \frac{imp^r_{ij}}{x^r_j}\right]$ as the matrices whose entries are the national technical coefficients, the regional input coefficients, and the regional import coefficients, respectively. Assuming that only NIOT (A^n) and the vector of the regional total sectorial output (x^r_j, j = 1, ... , k) are known, the LQ methods estimate the matrix of the regional input coefficients R by adjusting the national technical coefficient in the following way:

$$\hat{r}_{ij} = a^n_{ij} q_{ij} \tag{5}$$

where q_{ij} represents the degree of modification of the national coefficient. Interregional import coefficients (the entries of M^r) are estimated as the difference between the national and the estimated regional input coefficient:

$$m^r_{ij} = a^n_{ij} - \hat{r}_{ij} \tag{6}$$

The first LQ method introduced in the literature (Flegg and Tohmo 2016) was the Simple Location Quotient (SLQ), where the regional input coefficients are estimated as:

$$\hat{r}_{ij} = \begin{cases} a^n_{ij} \cdot SLQ_i & \text{if } SLQ_i < 1 \\ a^n_{ij} & \text{if } SLQ_i \geq 1 \end{cases} \tag{7}$$

and where SLQ_i is defined as:

$$SLQ_i = \frac{\frac{x^r_i}{x^r}}{\frac{x^n_i}{x^n}} \tag{8}$$

and x^r_i and x^n_i are the total outputs of the i-th regional and national sector, respectively, where $x^n = \sum_{i=1}^{k} x^n_i$ and $x^r = \sum_{i=1}^{k} x^r_i$.

Several other LQ methods have been proposed in the literature (Miller and Blair 2009). However, earlier studies (Bonfiglio and Chelli 2008; Hermannsson 2016; Morrissey 2016; Jahn 2017) have demonstrated how FLQ provides more accurate results than the other LQ methods and, based on such evidence, this method was chosen to estimate the 20 Italian RIOTs.

The basic idea underlying FLQs is that a region's propensity to import from other domestic regions is inversely and non-linearly related to its relative size (Ciommi et al. 2019). By incorporating explicit adjustments for interregional trade, the method provides more accurate estimates of regional input coefficients. As with other non-survey techniques, the main aim of the FLQ approach is to delineate an optimal frame to estimate input–output tables that are representative of the regional economic structure (Lamonica et al. 2020). FLQ coefficients can be expressed as follows:

$$FLQ_{ij} = \begin{cases} CILQ_{ij}\lambda & \text{for } i \neq j \\ SLQ_{ij}\lambda & \text{for } i = j \end{cases} \quad (9)$$

where λ stands for the relative size of the region and takes the following form:

$$\lambda = \left[\log_2\left(1+\frac{x^r}{x^n}\right)\right]^{\delta} \quad (10)$$

and

$$CILQ_{ij} = \frac{x_i^r/x_i^n}{x_j^r/x_j^n} = \frac{SLQ_i}{SLQ_j}$$

based on Flegg et al. (1995). Here, δ ($0 \leq \delta < 1$) is a sensitivity parameter that controls the degree of convexity in Equation (5). Referring to Flegg et al. (1995) for details, the larger the value of δ, the lower the value of λ, so that greater adjustments of regional imports are made. The implementation of the FLQ formula is carried out in line with other LQ methods:

$$\hat{r}_{ij} = \begin{cases} a_{ij}^n FLQ_{ij} & \text{if } FLQ_{ij} < 1 \\ a_{ij}^n & \text{if } FLQ_{ij} \geq 1 \end{cases} \quad (11)$$

To apply the FLQ, a value for the unknown parameter (δ) has to be chosen. A number of empirical studies (Flegg and Webber 2000; Flegg et al. 2016; Flegg and Tohmo 2016; Jahn et al. 2020) were devoted to find appropriate values of δ. In consideration of their results, a value of δ = 0.3 was considered appropriate in our case.

3. Data and Indicators

This study relied on the 2016 Italian Input-Output Table (IOT) using a disaggregation nomenclature of 63 sectors, and was retrieved from an official database (ISTAT 2020). The 63 sectors' classification was based on the NACE Rev.2 Statistical classification of economic activities in the European Community (EUROSTAT 2008). Unfortunately, the only available data at the regional level are related with the employment number and the added value of the 29 sectors' disaggregation (NACE Rev. 2). As a consequence, the National Input-Output Table was then reaggregated in 29 sectors that were used as the starting point for regionalization by means of the Flegg Location Quotient (FLQ), as pointed out in Section 2.2, which allowed for a comparative analysis at regional scale.

The regional sectorial employment number was used as a generalization of the national coefficient based on the fact that "in cases where regional output data are not consistently available, or where analysts feel it is appropriate, other measures of regional and national economic activity are often used—including employment (probably the most popular), personal income earned, value added, and so on, by sector" (Miller and Blair 2009, p. 349). Moreover, the practical calculation of Equations (1) and (2) requires the sectorial final demand as input data at the regional level. Unfortunately, these data are unavailable for Italy. To overcome this drawback, we assumed the regional share of the i-th sector as coinciding with the share allocated for the whole country economy. This assumption was derived from the fundamental input–output relationship $\mathbf{v}'i = i'f$, and following the argumentations of Round (1983). Thus, the regional sectoral final demand was estimated as follows:

$$f_i^r = v^r \frac{f_i^n}{\sum_{i=1}^k f_i^n} \quad (12)$$

where $v^r = (v^r)'i$ is the regional (total) added value.

4. Results

The pandemic-driven economic recession affected all areas of Italy. However, the drop in GDP has been partly attenuated at the regional level through the measures adopted by the national government and European authorities. The campaign of vaccination, the progressive easing of restrictions aimed at the containement of contagion, and the perseverance in the measures benefitting households and firms helped in sustaining the economic recovery. According to the quarterly indicator of the regional economy (ITER) elaborated by the Bank of Italy, recovery was particularly evident in Northern Italy. Exports have grown in all areas and investments appear higher than those planned. Positive signals were observed on incomes and consuption expenditures. Savings have continued to be addressed, due in large part to liquid financial instruments as deposits (Banca d'Italia 2020). National results in terms of linkage, that we obtained from the empirical analysis, are shown in Table 4. The forward and backward linkages were computed for the 29 industries of the Italian economy. In Table 5, following the taxonomy defined in the methodological Section 2.1, industries were classified in the relevant panel according to the following denominations: Key Sectors, Low Impact, Prime Vendors and Prime Users.

Table 4. Forward and backward linkage results for the 29 industries constituting the national economy.

INDUSTRIES		ITALY_FL	ITALY_BL
S1	Crop and animal production, hunting and related service activities	0.908	0.452
S2	Fishing and acquaculture	0.452	0.015
S3	Mining and quarring	0.846	0.080
S4	Manufacture of food products, beverages, and tobacco products	0.998	1.847
S5	Manufacture of textiles and wearing apparel	0.734	0.913
S6	Manufacture of wood and of products of wood, paper and paper products and printing	0.907	0.588
S7	Manufacture of coke and refined petroleum products	1.325	1.431
S8	Manufacture of rubber, plastic products, and other non-metallic mineral products	1.032	0.929
S9	Manufacture of fabricated metal products, except machinery and equipment	1.475	1.544
S10	Manufacture of computer, electronic and optical products, electrical equipment machinery and equipment n.e.c.	1.265	2.292
S11	Manufacture of transport equipment	0.819	1.342
S12	Manufacture of furniture, Other manufacturing Repair and installation of machinery and equipment	0.712	0.783
S13	Electricity, gas, steam, and air conditioning supply	1.054	0.937
S14	Water collection, treatment, and supply	0.745	0.432
S15	Construction	1.037	2.101
S16	Wholesale and retail trade and repair of motor vehicles and motorcycles	1.772	2.907
S17	Trasportation and storage	1.721	1.656
S18	Accommodation and food service activities	0.732	1.091
S19	Information and communication	1.196	1.078
S20	Financial and insurance activities	1.493	0.842
S21	Real estate activities	1.085	0.561
S22	Professional, scientific, and technical activities	1.864	1.211
S23	Administrative and support service activities	1.386	1.026
S24	Public Administration and defence; compulsory social security	0.651	0.809
S25	Education	0.512	0.229

Table 4. *Cont.*

INDUSTRIES		ITALY_FL	ITALY_BL
S26	Human health and social work activities	0.620	1.135
S27	Arts, entertainment and recreation	0.640	0.424
S28	Other services activities	0.559	0.329
S29	Activities of households as employers; undifferentiated good and services producing activities of households for own use	0.443	0.000

Table 5. Classification of Italian industries according to their role in the economic interactions.

	FL < 1	FL > 1
BL < 1	Low impact (I panel) (S1) Crop and animal production, hunting and related service activities (S2) Fishing and acquaculture (S3) Mining and quarring (S5) Manufacture of textiles and wearing apparel (S6) Manufacture of wood and of products of wood, paper and paper products and printing (S12) Manufacture of furniture, Other manufacturing Repair and installation of machinery (S14) Water collection, treatment and supply (S24) Public Administration and defence; compulsory social security (S25) Education (S27) Arts, Entertainment and recreation (S28) Other services activities (S29) Activities of households as employers; undifferentiated good and services producing activities of households for own use	Prime Vendors (II panel) (S8) Manufacture of rubber, plastic products and other non-metallic mineral products (S13) Electricity, gas, steam and air conditioning supply (S20) Financial and insurance activitites (S21) Real estate activities
BL > 1	Prime Users (III panel) (S4) Manufacture of food products, beverages and tobacco products (S11) Manufacture of transport equipment (S18) Accomodation and food service activities (S26) Human health and social work activities	Key Sectors (IV panel) (S7) Manufacture of coke and refined petroleum products (S9) Manufacture of fabricated metal products, except machinery and equipment (S10) Manufacture of computer, electronic and optical products, electrical equipment (S15) Construction (S16) Wholesale and retail trade and repair of motor vehicles and motorcycles (S17) Trasportation and storage (S19) Information and communication (S22) Professional, scientific, and technical activities (S23) Administrative and support service activities

Each industry was assumed to be part of a network that developed through inter-industry interactions. A set of interactions was given by the inflow of commodities, from raw materials to finished products, realized by the industry's intermediate purchases, and used for producing the industry's total output. These interactions, which define the role of each industry in the inter-sectoral interactions in the upstream supply chain, were then given by the backward linkage coefficient quantified in the last column of Table 4. The second set of interactions was the downstream network that involved processing the materials collected during the upstream stage into a finished product and the actual sale of the industry's total output to other industries. The last column of Table 4 displays the capability of each industry in activating the other industries downstream. In order to synthetize the features of each industry in the interaction, we rearranged the results in Table 4 according the linkage value, as shown in Section 2.1. Table 5 shows the resulting industry classifications at the national economy.

From this table, nine industries emerged as Key Sectors, and provided a relevant impulse to the production process in terms of both upstream and downstream interaction. Four industries emerged as Prime Users and Prime Vendors, respectively. Twelve industries actually proved to be Low Impact activities. More information could be attained by adopting the regional viewpoint. Considering a regional perspective, results similar to those obtained for the national economy, shown in Tables 4 and 5, were observed. In this way, a further development got results for each sector according to the role in the economic interaction and according to the regional allocation of the economic activity, since the regional economic context easily influenced the efficiency of the economic interactions among industries. Here, we regrouped the five macro-regions into three groups: North, Center, and South. The islands were considered within the Southern region.

When compared with national outcomes, some specificities emerge. Table 6 shows the Low Impact sectors for all the Italian regions. As expected, regional outcomes reflected, in most cases, the national ones. Nevertheless, some regional aspects that did not reflect the national trend should be taken into consideration. The industry S1—crop and animal production, hunting and related service activities, and agricultural and hunting—result was classified as a 'Prime Vendor', (Table 7) since it provided its output to other industries, and its activities were allocated prevalently in the eight regions of the south. The textile industry (S5) emerged as a 'Prime User' (Table 8) in Veneto, Tuscany, Umbria, Abruzzo, Campania, and Apulia, and as a 'Prime Vendor' in Marche. The manufacture of rubber, plastic products, and other non-metallic mineral products (S8) was classified as a 'Low Impact' sector in Aosta Valley, Liguria, Sicily, Sardinia, Tuscany, Latium, Calabria, Campania, and Apulia, and as a 'Prime User' in Marche.

The accommodation and food service activities (S18) were classified as 'Prime Vendors' in Marche. In all the other regions of Italy, they were 'Prime Users', with the exception of Latium ('Low Impact'). In both Calabria and Latium, public administration (S24) was classified as a 'Prime User,' while in all other regions it was a low impact one. These findings are coherent with the geography of Latium, whose economic system gravitates to Rome, the Italian capital city, where most of central administrations are located. Public administration is regarded as the main customer for the providers of intermediate inputs; while in Calabria, a similar outcome may have depended on the limited development of the industrial system.

Additional specificities can be found with reference to S4—food, beverage and tobacco industry—emerging as a 'Key Sector' in all Italian regions, with the exception of Liguria, Lombardy, Tuscany, and Latium, where it was regarded as a 'Prime User' (Table 7) and in Marche, where it was classified as a 'Prime Vendor' (Table 8). With reference to the coke and petroleum industry (S7), exceptions emerged for Trentino Alto Adige, Umbria, Campania, Apulia, and Calabria, where this sector was classified as a 'Prime User', as well as for Aosta Valley, Friuli-Venezia-Giulia, and Basilicata, where it was classified as a 'Low Impact' sector.

The manufacture of metals (S9), in the well-known 'Industrial Triangle' encompassing Lombardy, Piedmont and Liguria, emerged as a 'Key Sector' (Table 9) because of the intense links between naval industries, machineries, aerospace, and automobiles concentrated in the area. Exceptions, in reference to the manufacture of computer and electronic devices (S10) with respect to the national classification (Key sector), emerged for Latium, Molise, Campania, Apulia, Basilicata, Calabria, Sicily, and Sardinia, where the sector emerged as a 'Prime User'. Electricity and gas (S11) in seven regions (Veneto, Friuli Venezia Giulia, Emilia Romagna, Tuscany, Marche, Abruzzo, and Campania), emerged as a 'Low Impact' sector. This was different from the national level in the remaining regions, where the sector was classified as a 'Key Sector.' Information and communication (S19) was revealed as a 'Prime Vendor' sector, while in Umbria, Marche, Abruzzo, Molise, Puglia, and Basilicata, it was classified as a 'Low Impact' Sector and as a 'Prime User' in Campania. The only region which included financial and insurance activities within 'Key Sectors' was Latium, while in Marche this sector was included within 'Prime Users'. Peculiarities for administrative and support service activities (S3) were observed in Liguria, Trentino Alto Adige, Veneto, Umbria, Molise, Calabria, and Sicily, where this activity was a 'Prime Vendor' and in Marche, where administrative and support services were a 'Prime User'.

Table 6. Low Impact Sectors in the Italian regions.

SECTORS	S1		S2		S3		S5		S6		S7		S8		S9		S12		S13		S14		S18		S19		S24		S25		S27		S28		S29	
	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL
NORTHERN REGIONS																																				
EMILIA R.	0.92	0.47	0.01	0.01	0.39	0.05	0.46	0.86	0.84	0.57	-	-	-	-	-	-	0.52	0.80	0.67	0.57	0.57	0.41	-	-	-	-	0.24	0.82	0.10	0.23	0.32	0.39	0.21	0.30	0.00	0.00
FRIULI V. G.	0.91	0.43	0.01	0.01	0.28	0.04	0.14	0.59	0.95	0.57	0.76	0.95	-	-	-	-	0.56	0.81	0.65	0.54	0.54	0.42	-	-	-	-	0.42	0.85	0.13	0.25	0.40	0.45	0.22	0.32	0.00	0.00
LIGURIA	0.61	0.38	0.02	0.01	0.27	0.04	0.04	0.57	0.36	0.32	-	-	0.76	0.78	-	-	0.53	0.76	-	-	0.62	0.45	-	-	-	-	0.44	0.95	0.13	0.28	0.44	0.47	0.25	0.36	0.00	0.00
LOMBARDIA	0.48	0.31	0.00	0.01	0.75	0.08	0.57	0.96	0.88	0.59	-	-	-	-	-	-	0.50	0.81	-	-	0.43	0.36	-	-	-	-	0.20	0.81	0.09	0.24	0.33	0.39	0.19	0.34	0.00	0.00
PIEMONTE	0.75	0.41	0.00	0.01	0.46	0.05	0.44	0.84	0.87	0.58	-	-	-	-	-	-	0.49	0.80	-	-	0.57	0.42	-	-	-	-	0.27	0.81	0.11	0.23	0.36	0.42	0.21	0.33	0.00	0.00
TRENTINO A. A.	-	-	0.00	0.01	0.55	0.06	0.14	0.63	0.96	0.56	-	-	-	-	-	-	0.44	0.81	-	-	0.36	0.34	-	-	-	-	0.42	0.81	0.14	0.25	0.30	0.41	0.21	0.36	0.00	0.00
V. D'AOSTA	0.91	0.40	0.00	0.01	0.45	0.06	0.04	0.59	0.85	0.49	0.08	0.85	0.45	0.52	-	-	0.28	0.73	-	-	0.54	0.42	-	-	-	-	0.49	0.91	0.14	0.30	0.48	0.48	0.26	0.35	0.00	0.00
VENETO	0.95	0.47	0.01	0.01	0.25	0.03	-	-	0.90	0.59	-	-	-	-	-	-	0.53	0.81	0.68	0.57	0.55	0.40	-	-	-	-	0.27	0.87	0.11	0.24	0.31	0.39	0.21	0.32	0.00	0.00
CENTRAL REGIONS																																				
LAZIO	0.66	0.38	0.01	0.01	0.79	0.08	0.10	0.61	0.62	0.44	-	-	0.54	0.69	0.82	0.88	0.31	0.71	-	-	0.59	0.45	0.60	0.99	-	-	-	-	0.14	0.29	0.45	0.50	0.26	0.38	0.00	0.00
MARCHE	0.45	0.85	0.01	0.01	0.05	0.41	-	-	0.59	0.89	-	-	-	-	-	-	0.83	0.54	0.66	0.80	0.42	0.62	-	-	0.87	0.87	0.86	0.38	0.23	0.13	0.42	0.35	0.31	0.22	0.00	0.00
TOSCANA	0.84	0.43	0.01	0.01	0.74	0.07	-	-	0.91	0.62	-	-	0.95	0.88	-	-	0.53	0.80	0.91	0.77	0.57	0.42	-	-	-	-	0.35	0.89	0.12	0.25	0.40	0.45	0.22	0.33	0.00	0.00
UMBRIA	-	-	0.00	0.01	0.44	0.06	-	-	0.92	0.60	-	-	-	-	-	-	0.53	0.80	-	-	0.63	0.44	-	-	0.82	0.84	0.42	0.86	0.13	0.25	0.32	0.41	0.23	0.34	0.00	0.00
SOUTHERN REGIONS																																				
ABRUZZO	-	-	0.01	0.01	0.77	0.07	-	-	0.90	0.59	-	-	-	-	-	-	0.41	0.80	0.99	0.82	0.61	0.43	-	-	0.67	0.76	0.42	0.82	0.13	0.23	0.37	0.43	0.23	0.32	0.00	0.00
BASILICATA	-	-	0.00	0.01	0.62	0.05	0.17	0.69	0.63	0.42	0.24	0.90	-	-	-	-	0.47	0.79	-	-	0.69	0.45	-	-	0.86	0.91	0.47	0.93	0.15	0.26	0.33	0.43	0.24	0.36	0.00	0.00
CALABRIA	-	-	0.02	0.01	0.73	0.06	0.08	0.67	0.57	0.40	-	-	0.74	0.69	-	-	0.30	0.73	-	-	0.72	0.46	-	-	-	-	-	-	0.17	0.29	0.35	0.44	0.28	0.36	0.00	0.00
CAMPANIA	-	-	0.02	0.01	0.32	0.05	-	-	0.73	0.49	-	-	0.79	0.72	-	-	0.43	0.78	0.99	0.78	0.63	0.44	-	-	-	-	0.44	0.93	0.15	0.26	0.43	0.48	0.25	0.33	0.00	0.00
MOLISE	-	-	0.02	0.01	0.71	0.07	0.16	0.68	0.79	0.53	-	-	-	-	-	-	0.31	0.75	-	-	0.45	0.38	-	-	0.91	0.89	0.43	0.83	0.15	0.25	0.32	0.42	0.25	0.34	0.00	0.00
PUGLIA	-	-	0.02	0.01	0.59	0.06	-	-	0.67	0.47	-	-	0.77	0.74	-	-	0.50	0.77	-	-	0.64	0.45	-	-	0.95	0.91	0.44	0.93	0.14	0.27	0.36	0.44	0.25	0.35	0.00	0.00
SARDEGNA	-	-	0.02	0.01	0.88	0.08	0.06	0.60	0.54	0.40	-	-	0.69	0.76	-	-	0.37	0.69	-	-	0.66	0.44	-	-	-	-	0.47	0.95	0.15	0.28	0.43	0.46	0.26	0.36	0.00	0.00
SICILIA	-	-	0.02	0.01	0.78	0.08	0.08	0.63	0.49	0.38	-	-	0.75	0.78	-	-	0.37	0.69	-	-	0.66	0.43	-	-	-	-	0.47	0.93	0.15	0.27	0.46	0.47	0.26	0.36	0.00	0.00

Table 7. Prime Vendors in the Italian regions.

SECTORS	S1		S4		S5		S8		S9		S11		S13		S18		S19		S20		S21		S23		S26	
	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL
NORTHERN REGIONS																										
EMILIA R.	-	-	-	-	-	-	1.14	0.92	-	-	-	-	-	-	-	-	1.04	0.92	1.94	0.82	1.19	0.57	-	-	-	-
FRIULI V. G.	-	-	-	-	-	-	1.21	0.81	-	-	-	-	-	-	-	-	-	-	2.20	0.93	1.25	0.63	-	-	-	-
LIGURIA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.17	0.91	1.40	0.62	1.66	0.95	-	-
LOMBARDIA	-	-	-	-	-	-	1.12	0.95	-	-	-	-	1.02	0.86	-	-	-	-	2.04	0.90	1.16	0.59	-	-	-	-
PIEMONTE	-	-	-	-	-	-	1.12	0.92	-	-	-	-	1.00	0.83	-	-	-	-	2.01	0.88	1.18	0.58	-	-	-	-
TRENTINO A. A.	1.11	0.47	-	-	-	-	1.03	0.81	-	-	-	-	1.30	0.99	-	-	1.02	0.97	2.34	0.96	1.27	0.64	1.54	0.95	-	-
V. D'AOSTA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.38	0.94	1.53	0.62	-	-	-	-
VENETO	-	-	-	-	-	-	1.18	0.89	-	-	-	-	-	-	-	-	1.05	0.95	1.99	0.84	1.23	0.60	1.61	0.98	-	-
CENTRAL REGIONS																										
LAZIO	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.43	0.66	-	-	-	-
MARCHE	-	-	1.89	0.94	1.02	0.63	-	-	-	-	1.37	0.51	-	-	1.07	0.54	-	-	-	-	-	-	-	-	1.20	0.37
TOSCANA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.06	0.96	2.03	0.87	1.26	0.59	-	-	-	-
UMBRIA	1.04	0.48	-	-	-	-	1.19	0.84	-	-	-	-	1.26	0.99	-	-	-	-	1.85	0.76	1.29	0.59	1.52	0.93	-	-
SOUTHERN REGIONS																										
ABRUZZO	1.02	0.49	-	-	-	-	1.20	0.93	-	-	-	-	-	-	-	-	-	-	1.58	0.64	1.20	0.56	-	-	-	-
BASILICATA	1.27	0.49	-	-	-	-	1.13	0.78	-	-	-	-	1.29	0.99	-	-	-	-	1.62	0.66	1.29	0.66	-	-	-	-
CALABRIA	1.17	0.46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.03	0.96	1.73	0.71	1.58	0.58	1.98	0.98	-	-
CAMPANIA	1.15	0.49	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.69	0.72	1.37	0.59	-	-	-	-
MOLISE	1.18	0.54	-	-	-	-	1.09	0.93	-	-	-	-	-	-	-	-	-	-	1.67	0.68	1.34	0.55	1.70	0.99	-	-
PUGLIA	1.15	0.50	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.83	0.76	1.37	0.58	-	-	-	-
SARDEGNA	1.13	0.52	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.12	0.96	1.72	0.71	1.45	0.57	-	-	-	-
SICILIA	1.10	0.50	-	-	-	-	-	-	1.03	0.99	-	-	-	-	-	-	1.02	0.94	1.86	0.76	1.45	0.57	1.76	0.95	-	-

Table 8. Prime Users in the Italian regions.

SECTORS	S4		S5		S7		S8		S10		S11		S18		S20		S21		S23		S24		S26	
	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL
NOTHERN REGIONS																								
EMILIA R.	-	-	-	-	-	-	-	-	-	-	0.73	1.43	0.52	1.13	-	-	-	-	-	-	-	-	0.34	1.18
FRIULI V. G.	-	-	-	-	-	-	-	-	-	-	0.75	1.48	0.56	1.12	-	-	-	-	-	-	-	-	0.40	1.11
LIGURIA	0.78	1.77	-	-	-	-	-	-	-	-	0.85	1.28	0.62	1.04	-	-	-	-	-	-	-	-	0.45	1.35
LOMBARDIA	0.96	1.66	-	-	-	-	-	-	-	-	0.50	1.28	0.43	1.07	-	-	-	-	-	-	-	-	0.30	1.18
PIEMONTE	-	-	-	-	-	-	-	-	-	-	0.71	1.40	0.43	1.09	-	-	-	-	-	-	-	-	0.35	1.16
TRENTINO A. A.	-	-	-	-	0.72	1.04	-	-	-	-	0.70	1.41	0.59	1.25	-	-	-	-	-	-	-	-	0.41	1.11
V. D'AOSTA	-	-	-	-	-	-	-	-	0.29	2.00	0.68	1.29	0.67	1.14	-	-	-	-	-	-	-	-	0.46	1.08
VENETO	-	-	0.61	1.00	-	-	-	-	-	-	0.36	1.24	0.55	1.15	-	-	-	-	-	-	-	-	0.34	1.20
CENTRAL REGIONS																								
LAZIO	0.51	1.61	-	-	-	-	-	-	0.62	1.91	0.44	1.11	-	-	-	-	-	-	-	-	0.45	1.01	0.44	1.36
MARCHE	-	-	-	-	-	-	0.91	1.20	-	-	-	-	-	-	0.84	2.03	0.59	1.26	0.87	1.37	-	-	-	-
TOSCANA	0.77	1.74	0.62	1.04	-	-	-	-	-	-	0.63	1.31	0.56	1.03	-	-	-	-	-	-	-	-	0.38	1.26
UMBRIA	-	-	0.64	1.03	0.75	1.01	-	-	-	-	0.35	1.26	0.57	1.26	-	-	-	-	-	-	-	-	0.39	1.14
SOUTHERN REGIONS																								
ABRUZZO	-	-	0.63	1.02	-	-	-	-	-	-	0.75	1.32	0.56	1.17	-	-	-	-	-	-	-	-	0.38	1.19
BASILICATA	-	-	-	-	-	-	-	-	0.50	2.07	0.87	1.18	0.61	1.42	-	-	-	-	-	-	-	-	0.49	1.13
CALABRIA	-	-	-	-	0.54	1.07	-	-	0.29	2.07	0.21	1.30	0.70	1.27	-	-	-	-	-	-	0.50	1.01	0.53	1.21
CAMPANIA	-	-	0.57	1.01	0.71	1.02	-	-	0.64	2.03	0.81	1.18	0.61	1.33	-	-	-	-	-	-	-	-	0.45	1.16
MOLISE	-	-	-	-	-	-	-	-	0.39	1.90	0.80	1.09	0.60	1.32	-	-	-	-	-	-	-	-	0.44	1.28
PUGLIA			0.59	1.03	0.77	1.09			0.53	1.96	0.80	1.17	0.61	1.33									0.44	1.19
SARDEGNA	-	-	-	-	-	-	-	-	0.28	1.84	0.07	1.21	0.64	1.29	-	-	-	-	-	-	-	-	0.47	1.34
SICILIA									0.63	2.07	0.20	1.30	0.64	1.26									0.47	1.31

Table 9. Key Sectors in the Italian regions.

SECTORS	S4		S7		S9		S10		S13		S15		S16		S17		S19		S20		S22		S23	
	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL	FL	BL
NORTHERN REGIONS																								
EMILIA R.	1.10	1.96	1.71	1.37	2.03	1.59	1.57	2.42	-	-	1.08	2.17	2.50	2.99	2.39	1.67	-	-	-	-	2.59	1.18	1.73	1.02
FRIULI V. G.	1.04	1.95	-	-	2.22	1.65	1.71	2.54	-	-	1.24	2.34	2.27	3.13	2.41	1.63	1.14	1.02	-	-	2.71	1.23	1.79	1.05
LIGURIA	-	-	1.68	1.39	1.06	1.01	1.34	2.20	1.31	1.05	1.44	2.29	2.63	3.42	2.85	1.91	1.14	1.04	-	-	3.01	1.24	-	-
LOMBARDIA	-	-	1.65	1.48	1.98	1.58	1.55	2.40	-	-	1.08	2.16	2.52	2.92	2.20	1.57	1.44	1.13	-	-	2.71	1.24	1.78	1.02
PIEMONTE	1.03	1.81	1.58	1.35	1.99	1.59	1.56	2.38	-	-	1.14	2.19	2.35	2.94	2.14	1.53	1.41	1.11	-	-	2.63	1.22	1.72	1.01
TRENTINO A. A.	1.31	2.15	-	-	1.91	1.47	1.10	2.29	-	-	1.38	2.23	2.62	3.14	2.34	1.56	-	-	-	-	2.65	1.17	-	-
V. D'AOSTA	1.09	2.00	-	-	2.32	1.59	-	-	1.47	1.09	1.63	2.24	2.34	3.64	3.12	1.80	1.38	1.12	-	-	2.73	1.19	1.81	1.02
VENETO	1.17	2.03	1.46	1.23	2.08	1.60	1.59	2.43	-	-	1.23	2.27	2.63	2.94	2.44	1.65	-	-	-	-	2.55	1.16	-	-
CENTRAL REGIONS																								
LAZIO	-	-	1.78	1.57	-	-	-	-	1.31	1.10	1.28	2.08	2.52	3.40	2.82	1.94	1.71	1.26	2.43	1.08	3.16	1.32	2.08	1.06
MARCHE	-	-	1.33	1.51	1.66	2.16	2.50	1.64	-	-	2.29	1.21	2.92	2.62	1.57	2.26	-	-	-	-	1.14	2.54	-	-
TOSCANA	-	-	1.72	1.51	1.39	1.20	1.18	2.17	-	-	1.16	2.18	2.56	3.07	2.42	1.72	-	-	-	-	2.78	1.23	1.81	1.06
UMBRIA	1.22	2.12	-	-	2.01	1.60	1.03	2.15	-	-	1.29	2.28	2.64	3.02	2.63	1.68	-	-	-	-	2.75	1.18	-	-
SOUTHERN REGIONS																								
ABRUZZO	1.21	2.07	1.64	1.50	1.85	1.51	1.07	2.19	-	-	1.25	2.24	2.47	3.02	2.34	1.64	-	-	-	-	2.45	1.10	1.73	1.01
BASILICATA	1.50	2.33	-	-	1.53	1.26	-	-	-	-	1.57	2.33	2.56	3.47	2.76	1.71	-	-	-	-	2.90	1.19	2.08	1.01
CALABRIA	1.06	2.17	-	-	1.19	1.03	-	-	1.52	1.19	1.70	2.25	3.01	3.37	3.32	1.96	-	-	-	-	2.85	1.13	-	-
CAMPANIA	1.38	2.31	-	-	1.28	1.12	-	-	-	-	1.43	2.23	2.79	3.29	2.86	1.83	1.17	1.03	-	-	2.88	1.22	1.90	1.04
MOLISE	1.42	2.33	1.89	1.56	1.11	1.03	-	-	1.36	1.05	1.44	2.19	2.50	3.30	2.86	1.85	-	-	-	-	2.60	1.12	-	-
PUGLIA	1.37	2.28	-	-	1.38	1.19	-	-	1.31	1.04	1.42	2.22	2.76	3.21	2.63	1.72	-	-	-	-	2.73	1.17	1.86	1.02
SARDEGNA	1.21	2.27	1.67	1.56	1.15	1.02	-	-	1.41	1.12	1.52	2.15	2.83	3.23	2.93	1.88	-	-	-	-	2.83	1.16	1.90	0.98
SICILIA	1.13	2.20	1.56	1.50	-	-	-	-	1.39	1.10	1.49	2.13	2.87	3.22	3.07	1.92	-	-	-	-	2.69	1.12	-	-

Our results confirmed the performance of the industries linked to the so-called "Made in Italy" designation, which was intended as high-value, and mostly consisted of artisan and non-routinely products realized exclusively in Italy (Salvati and Zitti 2011). According to recent official statistics, the agri-food system as a whole (including agroindustry, wholesale and retail trade, and catering), produced 522 billion euros, and accounted for 15% of the country's gross domestic product, which thus ensured a prominent position in Europe. Significant growth was also observed in the last decade for the food industry (S4), +12% value added and +8% production index, which doubled the production of manufacturing.

The contribution of agriculture (S1) and the food industry (S4) was also particularly evident in Italy, which displayed an absolute growth in sales (1.3%) by 324 billion euros (CREA 2020). The manufacture of food products, beverages, and tobacco products (S4) emerged as a 'Key Sector' in most Italian regions, except for Liguria, Lombardy, Tuscany, and Latium, which were North-Central regions where it emerged as a Prime User. The role of the Prime User in the economic interactions also applied to this activity at the national level (see Table 4). Although the economic systems proved to be able to regain their average performance at pre-pandemic levels, the agri-food system in Italy seemed to call for supporting policy actions. At the same time, the manufacturing of textiles and wearing apparel (S5) is another industry with a longstanding tradition. Sales of this sector accounted for 9% of total manufacture, with wool and linen as the dominant yarn productions. Artisan products and the export of footwear were recognized to have a prominent role in the sector. While this activity was classified as 'Low Impact' in most regions, the linkage indices at the national level amounted to FL = 0.734 and BL = 0.913. At the regional level, however, six regions emerged as 'Prime Users': Veneto, Tuscany, Umbria, Abruzzo, Campania, and Apulia. In Marche, a well-known shoe production region, the industry arose as a 'Prime Vendor'.

Delocalization processes affected the most recent dynamics of this sector. Usually, delocalization operates by displacing Italian production towards low-cost countries, which possibly reduces (or subtracts) the technological assets developed by the creativity of Italian workers. The liberalization of international commerce, which involved more than half of the textile firms, was an additional factor influencing the recent development of the fiber and yarn industries. The prominent role of Italy in the global rank should be preserved with targeted duties and other supportive measures (Chiaradia 2019). The metal product industry (S9) produced the most investment goods, through which technical innovation can be transmitted to all branches of the economy. In this way, this activity supports the intrinsic competitiveness of the entire manufacturing sector, whose growth depends on the latent capacity of the industry to grow and renew.

5. Discussion

Since the 1950s, as a consequence of industrialization in emerging economies, the need for a universally recognized method to measure inter-industry linkages began to emerge (Rubino and Vitolla 2018). This method was aimed at assessing the relationship between and within industries by promoting the balanced development of the economic system and by optimizing the industrial structure of the national economy (Lamonica and Chelli 2018). After a short description of the literature related to linkage analysis, we focused on a specific methodology based on the Hypothetical Extraction Method approach, which measures the relative importance of a given sector by taking into account its net importance to the external connections with the other sectors (Lamonica et al. 2020). By means of a 'non-survey' regionalization method, i.e., the Flegg Location Quotient, we regionalized the Input–Output matrix of 2016 and applied linkage analysis at both the national and regional levels to highlight the relevance of the weights in the location of the sectors, based on *a-priori* classes, namely Low Impact sectors, Prime Vendors, Prime Buyers, and Key Sectors (OECD 2021).

This methodology provided an economically robust tool for building the 20 Italian regional input–output tables, in a regional framework where innovation is basically created by larger firms, with limited innovation results of small- and medium-enterprises (SMEs) representing the dominant part of the industrial system (Cainelli et al. 2019). Given the burden of bureaucratic procedures and the relevant delays in the completion of the third industrial revolution (Salvati et al. 2017), the ICT revolution was consolidated in terms of infrastructure (both technical and administrative), even in a context where the Italian industrial sectors revealed a markedly fragmented structure (Ghisellini and Ulgiati 2020). Under such conditions, the most suitable candidate for Industry 4.0 provisions might be medium/large firms, including multi-nationals (ISTAT 2016). Through internally organized competences, this type of firm is prepared to deal with managerial and financial issues, and is equipped to deal with international trade procedures (D'Ingiullo and Evangelista 2020), while small firms need to hire external abilities and, possibly, find further credit sources to cover the commonly long delays in the operation of public administration (Da Roit and Iannuzzi 2022).

This could mean that, as policy beneficiaries, large businesses will eventually crowd out the small firms that constitute the backbone of the Italian economy (Ciffolilli et al. 2019), and have relevant limitations in terms of innovation diffusion (Ciaschini 2022). Such weak performance of SMEs has also been observed in applied works, such as Muscettola (2015) and Bartoloni et al. (2020), in which the patterns of growth of a representative panel of Italian manufacturing firms were investigated. We observed that, although the estimates suggest that small firms grow faster than larger ones, the applied results did not show a significant change in the average size of businesses at the end of the period under investigation (Bugamelli et al. 2018).

The slow pace in the realization of the ICT infrastructures was in turn influenced by austerity policies on public expenditures in the area of technology (Ciffolilli et al. 2019). Since the 1980s, the funds for private and public institutions and universities have not been considered as policy priorities by governments (e.g., Corona 2019). A recent diminution of 19% of public funding for research that took place in the period between 2008 and 2016 may confirm this assumption. However, the decline in public research and university activities has progressively stimulated an improvement in research for R&D (Bigerna 2013). However, due to the weakening of the system of public research, the scientific goals recently attained could be only temporary (Cutrini and Salvati 2021). The observed emigration of younger researchers, due to easier and more rewarding job opportunities (e.g., Recanatesi et al. 2015), as well as higher research funding, is a context where competences are recognized is a key issue at stake in this development dimension (Bertolini and Giovannetti 2006). The persistent weakness of Italian firms in the technological innovation of human capital and corporate governance—and especially the insufficient improvement of the context in which business activity develops—has certainly influenced the low increase in total factor productivity (Cainelli et al. 2019) that has characterized the unfavorable trend in hourly productivity, which definitely determines lazy growth—or even stagnation—in present times.

6. Conclusions

Based on the empirical data shown in this paper and the related discussion, some perplexities emerge on the topics of innovation, which are connected with the features of the Italian economy and social framework. The innovation policy is expected to strengthen regional innovation capabilities to increase regional competitiveness and nurture innovative and dynamic enterprises. Since its inception, the policy design has supported collaborative research and development (R&D), including through innovation clusters, and the promotion of partnerships in important areas such as the smart factory, Industry 4.0, life sciences, and the bio-economy. Yet, strong concentration in manufacturing and sophisticated/specific innovation activities within local core industries is at risk of decline, due to ongoing industrial transitions. Some specific features of the present Italian economy, characterized by the slowness of bureaucratic procedures with respect to other competing economies, as well as the swiftness of the political economic cycle (that weakens the monitoring process of innovative changes and influences the fragmented structure of sectors that are predominantly small- and medium- enterprises (SMEs)) tend to burden the innovation process and leave the entire load of innovation to the private sector. Policy should consider the results of regionalized IO exercises when designing general or disaggregated strategic instruments and measures aimed at fueling economic development through the leverage of industrial interlinkage improvements.

Author Contributions: Conceptualization, M.C. and F.M.C.; methodology, G.R.L. and C.C.; validation, C.C., L.S. and G.R.L. formal analysis, C.C.; resources, G.R.L., C.C. and L.S.; data curation, G.R.L. and C.C.; writing—original draft preparation, C.C.; writing—review and editing, C.C. and L.S. supervision, M.C., L.S. and F.M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in Istat: https://www.istat.it/it/archivio/253253, accessed on 13 November 2022.

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

References

Ali, Yousaf, Muhammad Sabir, and Noor Muhammad. 2019. A comparative input-output analysis of the construction sector in three developing economies of South Asia. *Construction Management and Economics* 37: 643–58. [CrossRef]

Archibugi, Daniele, Sergio Cesaratto, and Giorgio Sirilli. 1991. Sources of innovative activities and industrial organization in Italy. *Research Policy* 20: 299–313. [CrossRef]

Ascani, Andrea, Alessandra Faggian, and Sandro Montresor. 2021. The geography of COVID-19 and the structure of local economies: The case of Italy. *Journal of Regional Science* 61: 407–41. [CrossRef] [PubMed]

Banca d'Italia. 2020. Economie regionali Dinamiche recenti ed aspetti strutturali Banca d'Italia (2020) Economie regionali L'economia delle regioni Italiane 22. Rome: Bank of Italy. Available online: https://www.bancaditalia.it/pubblicazioni/economie-regionali/2020/2020-0022/index.html (accessed on 16 November 2022).

Bartoloni, Eleonora, Maurizio Baussola, and Luca Bagnato. 2020. Waiting for Godot? Success or failure of firms' growth in a panel of Italian manufacturing firms. *Structural Change and Economic Dynamics* 55: 259–75. [CrossRef]

Bashir, Muhammad Farhan, Benjiang Ma, and Luqman Shahzad. 2020. A brief review of socio-economic and environmental impact of COVID-19. *Air Quality, Atmosphere & Health* 13: 1403–9.

Basso, Gaetano. 2020. The Evolution of the Occupational Structure in Italy, 2007–17. *Social Indicators Research* 152: 673–704. [CrossRef]

Belitski, Maksim, Christina Guenther, Alexander S. Kritikos, and Roy Thurik. 2022. Economic effects of the COVID-19 pandemic on entrepreneurship and small businesses. *Small Business Economics* 58: 593–609. [CrossRef]

Bertolini, Paola, and Enrico Giovannetti. 2006. Industrial districts and internationalization: The case of the agri-food industry in Modena, Italy. *Entrepreneurship and Regional Development* 18: 279–304. [CrossRef]

Beyers, William B. 1976. Empirical identification of key sectors: Some further evidence. *Environment and Planning A* 8: 231–36. [CrossRef]

Bigerna, Simona. 2013. The regional growth-instability frontier in Italy. *Atlantic Economic Journal* 41: 463–64. [CrossRef]

Bonfiglio, Andrea, and Francesco Chelli. 2008. Assessing the behaviour of non-survey methods for constructing regional input–output tables through a Monte Carlo simulation. *Economic Systems Research* 20: 243–58. [CrossRef]

Bragatto, Paolo, Tomaso Vairo, Maria Francesca Milazzo, and Bruno Fabiano. 2021. The impact of the COVID-19 pandemic on the safety management in Italian Seveso industries. *Journal of Loss Prevention in the Process Industries* 70: 104393. [CrossRef]

Bugamelli, Matteo, Monica Amici, Fabrizio Colonna, Francesco D. Amuri, Francesco Manaresi, Giuliana Palumbo, Filippo Scoccianti, and Enrico Sette. 2018. *Productivity Growth in Italy: A Tale of a Slow-Motion Change*. Bank of Italy Occasional Paper 422. Rome: Bank of Italy. Available online: https://www.bancaditalia.it/pubblicazioni/qef/2018-0422/QEF_422_18.pdf?language_id=1 (accessed on 16 November 2022).

Cainelli, Giulio, and Riccardo Leoncini. 1999. Externalities and long-term local industrial development. Some empirical evidence from Italy. *Revue d'Economie Industrielle* 90: 25–39. [CrossRef]

Cainelli, Giulio, Roberto Ganau, and Anna Giunta. 2018. Spatial agglomeration, global value chains, and productivity. Micro-evidence from Italy and Spain. *Economics Letters* 169: 43–46. [CrossRef]

Cainelli, Giulio, Roberto Ganau, and Marco Modica. 2019. Does related variety affect regional resilience? New evidence from Italy. *The Annals of Regional Science* 62: 657–80. [CrossRef]

Cella, Guido. 1984. The input-output measurement of interindustry linkages. *Oxford Bulletin of Economics and Statistics* 46: 73–84. [CrossRef]

Chen, Ji, Jiayan Huang, Weihua Su, Dalia Štreimikienė, and Tomas Baležentis. 2021. The challenges of COVID-19 control policies for sustainable development of business: Evidence from service industries. *Technology in Society* 66: 101643. [CrossRef] [PubMed]

Chenery, Hollis B., and Tsunehiko Watanabe. 1958. International comparisons of the structure of production. *Econometrica* 26: 487–521. [CrossRef]

Chiaradia, Giuseppe. 2019. La delocalizzazione dei prodotti del tessile deve essere disincentivata con pesanti dazi sulle merci di qualità. Available online: https://www.lineaitaliapiemonte.it/2019/07/30/leggi-notizia/argomenti/editoriali/articolo/la-delocalizzazione-dei-prodotti-del-tessile-deve-essere-disincentivata-con-pesanti-dazi-sulle-merci.html (accessed on 11 December 2021).

Ciaschini, Clio. 2022. A quantitative evaluation of Industra 4.0 through simulation of QUEST III Italia. *Rivista Internazionale di Scienze Sociali, Vita e Pensiero* 3: 251–92. [CrossRef]

Ciffolilli, Andrea, Eleonora Cutrini, and Marco Pompili. 2019. Do European Funds support the formation of firms? New evidence from Italy. *Regional Science Policy & Practice* 11: 549–69.

Ciommi, Mariateresa, Francesco M. Chelli, and Luca Salvati. 2019. Integrating parametric and non-parametric multivariate analysis of urban growth and commuting patterns in a European metropolitan area. *Quality & Quantity* 53: 957–79.

Clements, Benedict J. 1990. On the decomposition and normalization of interindustry linkages. *Economics Letters* 33: 337–40. [CrossRef]

Confindustria Report. 2019. Dove va l'industria italiana. 2019. Rapporto sull'industria Italiana 2019. Available online: https://www.confindustria.it/home/centro-studi/temi-di-ricerca/tendenze-delle-imprese-e-dei-sistemi-industriali/tutti/dettaglio/rapporto-industria+-italiana+-2019 (accessed on 11 December 2021).

Corona, Piermaria. 2019. Global change and silvicultural research. *Annals of Silvicultural Research* 43: 1–3.

CREA. 2020. *Annuario della Agricoltura Italiana*. Rome: Consiglio per la Ricerca in Agricoltura e l'analisi dell'Economia Agraria.

Crespí-Cladera, Rafel, Alfredo Martín-Oliver, and Bartolomé Pascual-Fuster. 2021. Financial distress in the hospitality industry during the COVID-19 disaster. *Tourism Management* 85: 104301. [CrossRef]

Cuello, Federico A., Fayçal Mansouri, and Geoffrey JD Hewings. 1992. The identification of structure at the sectoral level: A reformulation of the Hirschman–Rasmussen key sector indices. *Economic Systems Research* 4: 285–96. [CrossRef]

Cutrini, Eleonora, and Luca Salvati. 2021. Unraveling spatial patterns of COVID-19 in Italy: Global forces and local economic drivers. *Regional Science Policy & Practice* 13: 73–108.

D'Ingiullo, Dario, and Valentina Evangelista. 2020. Institutional quality and innovation performance: Evidence from Italy. *Regional Studies* 54: 1724–36. [CrossRef]

Da Roit, Barbara, and Francesco E. Iannuzzi. 2022. One of many roads to industry 4.0? Technology, policy, organisational adaptation and worker experience in 'Third Italy' SMEs. *New Technology, Work and Employment*. Available online: https://onlinelibrary.wiley.com/doi/epdf/10.1111/ntwe.12241 (accessed on 16 November 2022).

Dean, Mark, Al Rainnie, Jim Stanford, and Dan Nahum. 2021. Industrial policy-making after COVID-19: Manufacturing, innovation and sustainability. *The Economic and Labour Relations Review* 32: 283–303. [CrossRef]

Dei Ottati, Gabi. 2018. Marshallian industrial districts in Italy: The end of a model or adaptation to the global economy? *Cambridge Journal of Economics* 42: 259–84. [CrossRef]

Eurispes. 2022. Rapporto Italia 2021. Rome. Available online: http:www.eurispes.eu (accessed on 13 September 2022).

EUROSTAT. 2008. NACE Rev.2 Statistical classification of economic activities in the European Community. Available online: https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF (accessed on 11 December 2021).

Flegg, Anthony T., and Chris D. Webber. 2000. Regional size, regional specialization and the FLQ formula. *Regional Studies* 34: 563–69. [CrossRef]

Flegg, Anthony T., and Timo Tohmo. 2016. Estimating Regional Input Coefficients and Multipliers: The Use of FLQ is Not a Gamble. *Regional Studies* 50: 310–25. [CrossRef]

Flegg, Anthony T., Chris D. Webber, and Matthew V. Elliott. 1995. On the appropriate use of location quotients in generating regional input–output tables. *Regional Studies* 29: 547–61. [CrossRef]

Flegg, Anthony T., Leonardo J. Mastronardi, and Carlos A. Romero. 2016. Evaluating the FLQ and AFLQ formulae for estimating regional input coefficients: Empirical evidence for the province of Córdoba, Argentina. *Economic Systems Research* 28: 21–37. [CrossRef]

Gallego, Blanca, and Manfred Lenzen. 2005. A consistent input–output formulation of shared producer and consumer responsibility. *Economic Systems Research* 17: 365–91. [CrossRef]

Ghisellini, Patrizia, and Sergio Ulgiati. 2020. Circular economy transition in Italy. Achievements, perspectives and constraints. *Journal of Cleaner Production* 243: 118360. [CrossRef]

Golan, Amos, George Judge, and Sherman Robinson. 1994. Recovering information from incomplete or partial multisectoral economic data. *The Review of Economics and Statistics* 76: 541–49. [CrossRef]

Guerra, Ana-Isabel, and Ferran Sancho. 2010. Measuring energy linkages with the hypothetical extraction method: An application to Spain. *Energy Economics* 32: 831–37. [CrossRef]

Hauknes, Johan, and Mark Knell. 2009. Embodied knowledge and sectoral linkages: An input–output approach to the interaction of high-and low-tech industries. *Research Policy* 38: 459–69. [CrossRef]

Hermannsson, Kristinn. 2016. Beyond intermediates: The role of consumption and commuting in the construction of local input–output tables. *Spatial Economic Analysis* 11: 315–39. [CrossRef]

Hewings, Geoffrey J., and Rodney C. Jensen. 1988. Emerging challenges in regional input-output analysis. *The Annals of Regional Science* 22: 43–53.

Hirschman, Albert O. 1958. *The Strategy of Economic Development*. New Haven: Yale University Press.

ISTAT. 2016. *Il Sistema delle Tavole Input-Output*. Rome: ISTAT.

ISTAT. 2020. *Regional Accounts | Years 2017–2019 (dataset)*. Rome: ISTAT. Available online: https://www.istat.it/it/archivio/253253 (accessed on 16 November 2022).

Jahn, Malte. 2017. Extending the FLQ formula: A location quotient-based interregional input–output framework. *Regional Studies* 51: 1518–29. [CrossRef]

Jahn, Malte, Anthony T. Flegg, and Timo Tohmo. 2020. Testing and implementing a new approach to estimating interregional output multipliers using input–output data for South Korean regions. *Spatial Economic Analysis* 15: 165–85. [CrossRef]

Jones, Leroy P. 1976. The measurement of Hirschmanian linkages. *The Quarterly Journal of Economics* 90: 323–33. [CrossRef]

Kay, David L., James E. Pratt, and Mildred E. Warner. 2007. Role of services in regional economy growth. *Growth and Change* 38: 419–42. [CrossRef]

Lamonica, Giuseppe Ricciardo, and Francesco Maria Chelli. 2018. The performance of non-survey techniques for constructing sub-territorial input-output tables. *Papers in Regional Science* 97: 1169–202. [CrossRef]

Lamonica, Giuseppe R., Maria C. Recchioni, Francesco M. Chelli, and Luca Salvati. 2020. The efficiency of the cross-entropy method when estimating the technical coefficients of input–output tables. *Spatial Economic Analysis* 15: 62–91. [CrossRef]

Li, Zhong-Fei, Qi Zhou, Ming Chen, and Qian Liu. 2021. The impact of COVID-19 on industry-related characteristics and risk contagion. *Finance Research Letters* 39: 101931. [CrossRef]

Maples, Joshua G., Jada M. Thompson, John D. Anderson, and David P. Anderson. 2021. Estimating covid-19 impacts on the broiler industry. *Applied Economic Perspectives and Policy* 43: 315–28.

Meller, Patricio, and Manuel Marfán. 1981. Small and large industry: Employment generation, linkages, and key sectors. *Economic Development and Cultural Change* 29: 263–74.

Miller, Ronald E., and Michael L. Lahr. 2001. A taxonomy of extractions. *Contributions to Economic Analysis* 249: 407–41.

Miller, Ronald E., and Peter D. Blair. 2009. *Input-Output Analysis: Foundations and Extensions*. Cambridge: Cambridge University Press.

Morrissey, Karyn. 2016. A location quotient approach to producing regional production multipliers for the Irish economy. *Papers in Regional Science* 95: 491–506. [CrossRef]

Muscettola, Marco. 2015. Difficulties for small firms to invest in research prerogatives. An empirical analysis of a sample of Italian firms. *Applied Economics* 47: 1495–510. [CrossRef]

Notteboom, Theo, Thanos Pallis, and Jean-Paul Rodrigue. 2021. Disruptions and resilience in global container shipping and ports: The COVID-19 pandemic versus the 2008–9 financial crisis. *Maritime Economics & Logistics* 23: 179–210.

OECD. 2021. *Regional Innovation in Piedmont, Italy: From Innovation Environment to Innovation Ecosystem OECD*. Paris: Regional Development Studies OECD Publishing.

Oosterhaven, Jan. 2004. *On the Definition of Key Sectors and the Stability of Net Versus Gross Multipliers*. Groningen: SOM Research School, University of Groningen.

Oosterhaven, Jan. 2007. The net multiplier is a new key sector indicator: Reply to De Mesnard's comment. *The Annals of Regional Science* 41: 273–83. [CrossRef]

Pham, Tien Duc, Larry Dwyer, Jen-Je Su, and Tramy Ngo. 2021. COVID-19 impacts of inbound tourism on Australian economy. *Annals of Tourism Research* 88: 103179. [CrossRef]

Rasmussen, Poul Nørregaard. 1956. *Studies in Inter-Sectoral Relations*. København: E. Harck Publisher, vol. 15.

Recanatesi, Fabio, Matteo Clemente, Efstathios Grigoriadis, Flavia Ranalli, Marco Zitti, and Luca Salvati. 2015. A fifty-year sustainability assessment of Italian agro-forest districts. *Sustainability* 8: 32. [CrossRef]

Reis, Hugo, and António Rua. 2009. An input–output analysis: Linkages versus leakages. *International Economic Journal* 23: 527–44. [CrossRef]

Round, Jeffery I. 1983. Nonsurvey techniques: A critical review of the theory and the evidence. *International Regional Science Review* 8: 189–212. [CrossRef]

Rubino, M., and F. Vitolla. 2018. Implications of network structure on small firms' performance: Evidence from Italy. *International Journal of Business and Management* 13: 46–56. [CrossRef]

Salvati, Luca, and Marco Zitti. 2009. Substitutability and weighting of ecological and economic indicators: Exploring the importance of various components of a synthetic index. *Ecological Economics* 68: 1093–99. [CrossRef]

Salvati, Luca, and Marco Zitti. 2011. Economic growth vs. land quality: A multidimensional approach in Italy. *Journal of Environmental Planning and Management* 54: 733–48. [CrossRef]

Salvati, Luca, Marco Petitta, Tomaso Ceccarelli, Luigi Perini, Federica Di Battista, and Maria Elisa Venezian Scarascia. 2008. Italy's renewable water resources as estimated on the basis of the monthly water balance. *Irrigation and Drainage: The Journal of the International Commission on Irrigation and Drainage* 57: 507–15. [CrossRef]

Salvati, Luca, Marco Zitti, and Margherita Carlucci. 2017. In-between regional disparities and spatial heterogeneity: A multivariate analysis of territorial divides in Italy. *Journal of Environmental Planning and Management* 60: 997–1015. [CrossRef]

Schultz, Siegfried. 1977. Approaches to identifying key sectors empirically by means of input-output analysis. *The Journal of Development Studies* 14: 77–96. [CrossRef]

Seetharaman, Priya. 2020. Business models shifts: Impact of COVID-19. *International Journal of Information Management* 54: 102173. [CrossRef]

Sigala, Marianna. 2020. Tourism and COVID-19: Impacts and implications for advancing and resetting industry and research. *Journal of Business Research* 117: 312–21. [CrossRef]

Song, Hyoung J., Jihwan Yeon, and Seoki Lee. 2021. Impact of the COVID-19 pandemic: Evidence from the US restaurant industry. *International Journal of Hospitality Management* 92: 102702. [CrossRef] [PubMed]

Strassert, Günter. 1968. Zur bestimmung strategischer sektoren mit hilfe von input-output-modellen. *Jahrbücher für Nationalökonomie und Statistik* 182: 211–15. [CrossRef]

Szyrmer, Janusz M. 1992. Input—Output coefficients and multipliers from a total-flow perspective. *Environment and Planning A* 24: 921–37. [CrossRef]

Temurshoev, Umed. 2010. Identifying optimal sector groupings with the hypothetical extraction method. *Journal of Regional Science* 50: 872–90. [CrossRef]

Temurshoev, Umed, and Jan Oosterhaven. 2014. Analytical and empirical comparison of policy-relevant key sector measures. *Spatial Economic Analysis* 9: 284–308. [CrossRef]

Zambon, Ilaria, Andrea Colantoni, Margherita Carlucci, Nathan Morrow, Adele Sateriano, and Luca Salvati. 2017. Land quality, sustainable development and environmental degradation in agricultural districts: A computational approach based on entropy indexes. *Environmental Impact Assessment Review* 64: 37–46. [CrossRef]

economies

MDPI

Article

Financial Development and Language Structures

Gregory W. Caskey [1] and Nabamita Dutta [2,*]

[1] Tommy & Victoria Baker School of Business, The Citadel, The Military College of South Carolina, 89 Hagood Avenue, Charleston, SC 29409, USA
[2] Department of Economics, College of Business Administration, University of Wisconsin-La Crosse, 1725 State St., La Crosse, WI 54601, USA
* Correspondence: ndutta@uwlax.edu

Abstract: Using cross country data, we explore the role of linguistic structures for the financial development of countries. Specially, we investigate if future time reference (FTR), the requirement of an obligatory future tense marking in languages, matters for financial development or not. Our results show that countries speaking weak FTR language or a language not needing a dedicated future tense marking have enhanced financial development relative to countries speaking strong FTR language. Discounting the future less or having a connection between the present and the future—characteristics of weak FTR languages—has implications for caring about saving and investment, having efficient property rights, protection of shareholders and cost of acquiring information. Our results are robust to multiple measures of financial development and inclusion of determinants of the same. Finally, results show that weak FTR language speaking countries benefit more when their financial development is low.

Keywords: linguistic structures; future time reference; financial development; discounting future

JEL Classification: O11; Z10; D53

Citation: Caskey, Gregory W., and Nabamita Dutta. 2022. Financial Development and Language Structures. *Economies* 10: 313. https://doi.org/10.3390/economies10120313

Academic Editor: Franklin G. Mixon

Received: 18 March 2022
Accepted: 4 November 2022
Published: 8 December 2022

1. Introduction

The findings related to the role of financial development for economic growth is ambiguous in the literature.[1] While an extensive set of studies have stressed the impact of financial development on economic growth to be positive (Bekaert et al. 2005; Christopoulos and Tsionas 2004; Arestis et al. 2001; Xu 2000; Levine et al. 2000; Levine and Zervos 1998), other studies find financial development's effect on growth to be insignificant (Ho and Saadaoui 2022; Demetriades and Hussein 1996). More recent investigations utilizing time series data exploring the impact of financial development on growth find the effect to be weak or even negative (Cevik and Rahmati 2020; Nwani and Bassey Orie 2016; Adeniyi et al. 2015; Samargandi et al. 2014; Quixina and Almeida 2014). Such findings are of interest, particularly because in the context of other development outcomes, financial development has been shown to be beneficial for poverty alleviation (Rewilak 2017), enhancing the effectiveness for aid-recipient countries (Nkusu and Sayek 2004), promoting foreign direct investment (Desbordes and Wei 2017) and generating comparative advantage for manufacturing economies (Beck 2002).

Given these findings, many studies have explored the determinants of financial development, including institutional determinants (Roe and Siegel 2011; Herger et al. 2008; Beck and Levine 2004; Galindo and Micco 2004; Rajan and Zingales 2003; Johnson et al. 2000), policy determinants (Chinn and Ito 2006; Boyd et al. 2001; Huybens and Smith 1999) and economic development-related determinants (Levine 1997, 2003, 2005; Jaffee and Levonian 2001). We add to this strand of literature by exploring the role of linguistic structures for financial development. Specifically, we explore a specific linguistic trait—future time reference—and investigate if it explains differences in financial development across countries along with other factors.

In recent years, a number of studies have investigated the effects of linguistic traits on economic outcomes (Chen 2013; Galor et al. 2016; Mavisakalyan et al. 2018). One in particular, future time reference (FTR), is a linguistic trait indicating whether languages require an obligatory future tense marking. Languages such as English and French are *strong* FTR languages, as these languages require a dedicated form when referring to future events. In English, we state *she will go to New York tomorrow*. The same in French is stated as *Elle ira à New York demain*. Words like "will go" and "ira" represent a marked future tense. In contrast, languages such as German and Finnish are *weak* FTR languages, as the same grammatical form can be used for the future and the present. In German, for example, the same sentence can be stated as *Morgen geht sie nach New York.* The word "geht" indicating "goes" can be used for present and future tense.

We hypothesize that financial development is benefitted in countries where agents speak "weak" FTR languages relative to countries speaking "strong" FTR languages. In the context of FTR, the notion of "temporal displacement" (Mavisakalyan et al. 2018) suggests that dedicated used of grammar to indicate future events can make the future very distant for the individual. In the case of these strong FTR languages, the future can potentially appear discontinuous to the present relative to speakers of weak FTR languages.

As emphasized in the literature, the five key functions of the financial system include ex-ante production of information and allocating capital, generating effective corporate governance, mobilization of savings into investment, efficient risk management and minimizing cost in financial transactions (Dutta and Meierrieks 2021; World Bank 2012). We argue that FTR has effects on these outcomes, particularly through the channel of individual differences in future discounting. For example, individuals speaking strong FTR languages may be likely to care less about savings and financial intermediaries under such language structures, are less likely to be bothered less about easing the cost of acquiring capital as that can be put away for the future.[2] Likewise, well-functioning property rights are considered to be an essential part of financial development. The incentive for structuring such operational property rights might be lacking for agents speaking strong FTR languages. Since the future does not seem connected to the present, essential functions of the financial system—allocating resources effectively, protecting investors, sharing of information—can appear less rewarding. Thus, financial development is likely to suffer in countries speaking strong FTR languages. In contrast, agents speaking weak FTR languages are likely to discount the future less, as they cognitively associate the connection between the present and the future relative to strong FTR language speakers, and thus, are incentivized to generate an effective financial system. Under weak FTR language structures, agents are likely to place effort in effectively allocating capital towards saving and investment, create systems that protect investors and shareholders, and work towards minimizing the cost of acquiring information about financial decision making.

Our contribution in this paper is adding to the extensive strand of studies that have looked into the determinants of financial development and highlighting the role of language structures. Our results show that countries speaking weak FTR languages are likely to have between 10 and 25% percentage more financial development (depending on the specific measure of financial development used) relative to countries speaking strong FTR languages. Our results are robust to the inclusion of an array of controls, including political institutions. We check our results to alternate measures of financial development assessing its different characteristics—financial depth, size of the financial system, efficiency and extent of equity market activities.

Section 2 provides a brief background and literature review. Section 3 explains data and the sources. In Section 4, we describe the empirical methodology and benchmark results. Robustness analysis is described in Section 5, and Section 6 concludes.

2. Literature Review

The idea that linguistic traits can influence thought, and thus have effects on human behavior, has long been investigated. Based on the works of de Saussure (1916) and Wittgenstein (1922), the Sapir–Whorf hypothesis (SWH) stresses the idea that language can influence thought. Subsequent lines of research have explored this idea in various context. Though the hypothesis has been supported by many studies, the seminal works of Chomsky (1957) and Pinker (1994) have contested these findings, stressing that languages do not shape human cognition or ways of thinking. In subsequent decades, the Linguistic Relativity Hypothesis (LRH) was developed, which advances the Sapir–Whorf hypothesis by stating that both human cognition and behavior can be shaped by languages. While LRH was regarded as misguided by some linguists and cognitive scientists (Mavisakalyan et al. 2018), a substantial and ever-growing body of literature emerged starting in the 1990s testifying to the validity of the theory (Levinson and Wilkins 2006; Kay and Regier 2006; Boroditsky et al. 2003; Slobin 2003; Levinson 1996).

The idea that there's a direct influence of language on cognition and behavior is at the heart of the Linguistic Relativity Hypothesis (LRH). In a nutshell, the LRH states that the structure of one's language has a systematic influence on cognition and behavior, and as such, different languages represent the world in different ways by emphasizing different aspects of reality. As a result, speakers of a certain language may be more sensitive to various features of the world relative to speakers of another language. For example, studies in psychology like Harner (1981) show that for children speaking English, the use of future tense begins as early as age 3. English is considered to be a strong FTR language or a language requiring a dedicated future marker. Szagun (1978) also investigate differences in FTR for English (strong FTR language) and German (weak FTR language). He found no differences in future verb usage among children but did find such differences being reflected among adults.

As Mavisakalyan and Weber (2018) point out, the studies in economics on the effects of language on social outcomes differ from those studies in linguistics and psychology in a few significant ways. Whereas the former utilize much larger sample sizes and focus on the connection between language and broader economic and social outcomes, those studies in linguistics and psychology tend to use smaller sample sizes and focus on smaller, more specific cognitive effects. While studies like Licht et al. (2007) and Tabellini (2008) have considered linguistic structures as a source of exogenous variation in culture (Mavisakalyan and Weber 2018), more recent studies have considered linguistic traits as proxies for culture (Bhalotra et al. 2015; Santacreu-Vasut et al. 2014). Some studies have focused on investigating the effect of linguistic traits on various outcomes, explaining the association through the channels of both culture and cognition, but without distinguishing well between the two (Hicks et al. 2015; Santacreu-Vasut et al. 2013). Studies like Chen (2013) and Mavisakalyan (2015) claim that linguistic structures affect behavior and, thus, outcomes by directly altering individual's cognition. Studies such as these form the base of the new LRH literature.

As an example of the future marker, we can give the example of English where specific words like "will" or "is going to" has to be used to indicate " it will snow tomorrow". On the other hand, a language like German can imply the same thing by stating *Morgen schneien es* and not using grammar to indicate marked future events. Chen (2013) emphasizes that languages requiring grammar to indicate marked future events or strong future time reference (FTR) languages have speakers that are less future-oriented behavior which, in turn, lead inefficient outcomes.[3] As examples of inefficient outcomes, Chen finds that individuals speaking strong FTR languages save less, have less wealth after retirement, smoke more, tend to be obese and engage in unsafe sex.[4]

Galor et al. (2016) mention that speakers of languages that do not have marked grammar use indicating future tense are likely to have long-term orientation. Long term orientation or the lack of it affects individual's discounting of future. A reduction in an agent's discount rates can be because of long term orientation. Speakers of strong

FTR languages are likely to discount much more relative to speakers of weak FTR languages (Mavisakalyan et al. 2018). As such, FTR is relevant to the long tradition of studies on the human tendency to discount future costs and rewards (Frederick et al. 2002; Kirby and Herrnstein 1995; Solnick et al. 1980; Ramsey 1928).[5]

3. Data

3.1. The Sample

Our sample consists of an unbalanced panel of 100 developed and developing countries over the period 2001 to 2018. The panel is unbalanced because our variables of interest are not available for each country for every year. Overall, we have 844 observations for 100 countries, or approximately 8.44 observations for each country. We compile our data from the World Bank's Global Financial Development Database (GFDD) and World Development Indicators (WDI), as well as from Chen (2013). Below we describe our variables of interest in detail.

3.2. The Main Variables

As our benchmark measure of financial development, we consider the most used measure in the literature—private credit to GDP (Ibrahim and Alagidede 2017a; Jauch and Watzka 2015; Adeniyi et al. 2015; Nikoloski 2012; Clarke et al. 2006; Levine 2005). Specifically, this variable is defined as "domestic private credit to the real sector by deposit money banks" as a percentage of GDP (World Bank 2012). Private credit does not include credits issued to governments and public enterprises, nor does it include credits issued by central banks. This variable is a common measure of financial dept, which captures the financial sector relative to the economy, and has been documented in the literature as having a strong association with long-term economic growth (Beck et al. 2009). A measure of efficient credit allocation, private credit to GDP, signals the credit worthiness of private institutions, as well as accessibility of the credit market to private individuals (Jauch and Watzka 2015). The mean of the variable for our sample is about 73.5 percent with the range of the variable being from 6 percent to about 300 percent,[6] and the median of the variable is about 65%. For example, countries like Nigeria, Algeria, Pakistan, Mexico, Argentina and Ghana all have less than 22.5% private credit, which is the 10th percentile value of our sample.

The other alternate measure we considered is *private money by deposit money bank and other financial institutions and other financial institutions to GDP*. This is a standard alternate indicator of financial depth that has been used in the finance and growth literature (Beck et al. 2000, 2009). The mean for the variable for our sample is 81.34%, and the median of the variable is about 71%. We consider additional financial development measures as part of robustness analysis, which we discuss in subsequent sections.

3.3. Independent Variable

Based on Chen (2013), we classify languages that need a dedicated future marking (such as English and French) as a strong FTR language. On the other hand, languages like German and Finnish that do not require dedicated grammar use to mark future events are categorized as weak FTR languages.[7] We construct a dummy, taking the value of 1 for weak FTR languages and 0 for strong FTR languages. We chose this as our independent variable because, as Chen (2013) indicates, agents' intertemporal preferences and decision making are represented via strong and weak FTR languages. Likewise, Mavisakalyan et al. (2018) argue that future tense (for strong FTR languages) can be used to indicate cultural factors, and that usage of such can effect speakers' cognition and behavior (or both).

For our sample of countries, 28.2 percent of observations are assigned a dummy of 1 (indicating weak FTR languages), while the remaining 71.8 are assigned 0 (representing strong FTR languages). Our sample has sufficient regional and continental variation. For example, a number of European countries including Denmark, Belgium, Estonia, Germany, Finland, Iceland and Luxembourg have weak FTR languages. Yet, other European countries like France, Czech Republic, Latvia, Greece, Italy, Lithuania, Poland and United Kingdom have strong FTR languages.

Following Chen (2013) and Mavisakalyan et al. (2018), the language considered for each country is the major spoken language. Chen (2013) mentions that, for the majority of countries in our sample, there is no intra-country variation in terms of FTR strength. This implies that, in most countries, either one language dominates or a common FTR structure is shared among the languages for multi-lingual countries. As an example of the latter, Chen points to the example of Canada. While the country has significant English and French speaking populations, both are strong FTR languages. Likewise, Mavisakalyan et al. (2018) mentions that since available information on multi-lingual countries is not easily available, checking results with an alternate measure—share of total population speaking a strong FTR language—reduces the sample. As part of robustness analysis, we consider this measure and check our results.

3.4. Controls

We follow the literature in our choice of benchmark controls, which consist of GDP per capita growth, urban population as a percentage of total population, labor force participation rates, trade openness and polity as a measure of political institutions. Studies like Huang and Temple (2005) and Svaleryd and Vlachos (2002), for instance, show that trade openness is a significant determinant of financial development. Based on both demand and supply side arguments, studies like Jung (1986), Goldsmith (1969), Gurley and Shaw (1967) and Patrick (1966) have stressed a causal relationship from economic growth to financial development. For this reason, we include GDP per capita growth as a measure of economic development within our benchmark controls. As a further measure of economic development (based on the demand side argument), we control for urban population as a percentage of total population. Lastly, since multiple studies have shown that democratic institutions are an important determinant for financial development, we consider the Polity 2 variable, which runs from −10 to +10 with higher values implying more democratic institutions (Begović et al. 2017; Bhattacharyya 2013; Yang 2011; Huang 2010; Clague et al. 1996).

4. Empirical Methodology and Benchmark Results

4.1. Empirical Specification and Methodology

Based on our hypothesis, we test the following regression specification:

$$\mathrm{FD}_{it} = \beta_0 + \beta_1 FTR_i + controls + \varnothing_t + \epsilon_{it}$$

FD_{it} represents the specific financial development measure considered for country i in time t. Our main independent variable of interest is FTR, which represents the future time reference dummy. We remind our readers that FTR dummy takes 1 for weak FTR countries, and 0 for strong FTR countries. According to our hypothesis, we expect β_1 to be positive and significant. This would imply that financial development is higher in weak FTR countries (i.e., those with languages that do not require a dedicated form when referring to future events) relative to strong FTR countries. β_1 being negative and significant would mean the opposite.

Our benchmark measure of financial development is private credit to the real sector by deposit money banks as a percentage of GDP. As an alternate benchmark measure of financial development, we consider private credit to the real sector by deposit money banks and other financial institutions as a percentage of GDP. Our robustness analysis considers alternate measures of financial development.

Since linguistic features and financial development are likely the product of deeper, unobserved factors, we follow the literature in constructing our battery of controls, including observables that can help explain differences in financial development across countries. These, in turn, are unlikely to make the effect of FTR on financial development exogenous and uncorrelated with the error term. Controlling for observables that can help explain differences in financial development across countries is the first step to mitigate bias arising out of omitted variable bias.[8] As explained, our benchmark controls are GDP per capita growth, urban population as a percentage of total population, labor force participation rates, trade openness and polity as a measure of political institutions. These variables have been shown to matter for financial development. In the robustness section, we discuss further controlling for additional variables as well as mitigating the effect of unobserved heterogeneity with respect to the effects of FTR on financial development.

Our benchmark analysis consists of ordinary least squares (OLS) regressions. $\varnothing_t$, our time fixed effects, help us take into account time shocks. For example, global shocks, like the 2009 recession, that likely impact financial development should be captured in time fixed effects. In addition to our OLS specifications, we consider quantile regressions to make sure our results are not driven by the presence of outliers. We talk about endogeneity and how our findings should be interpreted in subsequent sections.

4.2. Benchmark Results

In Table 1, we present our first set of benchmark results with OLS regressions. The dependent variable considered is private credit to GDP. In column (1), we run a bivariate regression without any controls to assess the variation in financial development that is attributed to weak and strong FTR languages. Based on column (1), we find that, relative to strong FTR countries, weak FTR countries have 41 percent more private credit (as a percentage of GDP). Providing an example, this suggests that when compared against strong FTR countries (such as India), weak FTR countries (such as Indonesia) should have much more private credit. However, the coefficient of the FTR dummy in column (1) is capturing effects of other variables that also affect financial development. We add controls in subsequent columns. In column (2), we add labor force participation rate, and in column (3), we control for urban population as a percentage of total population. In column (4), GDP per capita growth is included, and in column (5), we control for trade as a percentage of GDP. Finally in column (6), we add the polity score, which is a measure of how relatively democratic a nation is in terms of its governance.

Table 1. Private Credit and FTR. OLS regressions with time fixed effects: The dependent variable is private credit by deposit money banks as a percentage of GDP. FTR is future time reference dummy with 1 indicating weak FTR countries and strong FTR countries. The controls for labor force participation rate (LFPR), urban population as a percentage of total population, GDP per capita growth, trade as percentage of GDP and polity. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$.

	(1)	(2)	(3)	(4)	(5)	(6)
Weak FTR	41.22 *** (3.355)	33.35 *** (3.306)	29.12 *** (3.484)	29.45 *** (3.491)	24.51 *** (3.597)	23.25 *** (3.620)
LFPR	—	0.827 *** (0.131)	0.778 *** (0.127)	0.785 *** (0.126)	0.911 *** (0.129)	0.801 *** (0.122)
Urban Population	—	—	0.402 *** (0.087)	0.329 *** (0.092)	0.299 *** (0.088)	0.232 *** (0.087)
GDP per cap. growth	—	—	—	−1.520 *** (0.508)	−1.812 *** (0.521)	−1.890 *** (0.515)
Trade	—	—	—	—	0.165 *** (0.024)	0.146 *** (0.025)
Polity	—	—	—	—	—	0.307 (0.298)
Constant	60.40 *** (7.559)	19.64 ** (9.414)	−6.692 (10.91)	−0.061 (11.23)	−13.49 (11.33)	−3.707 (11.19)

Table 1. *Cont.*

	(1)	(2)	(3)	(4)	(5)	(6)
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	778	778	778	777	769	748
R-squared	0.192	0.241	0.261	0.270	0.304	0.262

Once we include all of the aforementioned controls, the effect of FTR on private credit drops to 23 percentage points. This implies that weak FTR countries have 23 percentage points more private credit than strong FTR countries. Our controls are predominantly significant and are of expected sign and significance.[9]

To account for potential outliers driving our results, we replicate the specifications from Table 1 in Table 2 by using quantile regressions. The results are very similar. The magnitude of FTR dummy is marginally higher compared to Table 1.

Table 2. Private Credit and FTR. Quantile regressions with time fixed effects: The dependent variable is private credit by deposit money banks as a percentage of GDP. FTR is future time reference dummy with 1 indicating weak FTR countries and strong FTR countries. The controls for labor force participation rate (LFPR), urban population as a percentage of total population, GDP per capita growth, trade as percentage of GDP and polity. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$.

	(1)	(2)	(3)	(4)	(5)	(6)
Weak FTR	46.60 ***	38.34 ***	34.06 ***	33.66 ***	27.87 ***	26.28 ***
	(3.228)	(4.089)	(4.230)	(4.237)	(4.238)	(4.044)
LFPR	—	0.552 ***	0.389 ***	0.364 **	0.353 **	0.378 ***
		(0.146)	(0.146)	(0.147)	(0.145)	(0.140)
Urban Population	—	—	0.465 ***	0.446 ***	0.432 ***	0.424 ***
			(0.109)	(0.113)	(0.109)	(0.109)
GDP per cap. growth	—	—	—	−0.465	−1.129 **	−1.394 **
				(0.590)	(0.570)	(0.547)
Trade	—	—	—	—	0.164 ***	0.139 ***
					(0.032)	(0.033)
Polity	—	—	—	—	—	0.182
						(0.354)
Constant	60.12 ***	36.99 ***	14.70	16.85	5.088	5.892
	(8.152)	(12.10)	(14.01)	(14.25)	(14.10)	(13.81)
Observations	778	778	778	777	769	748

For example, in column (1), when we do not include any control variables, the magnitude of difference between weak FTR and strong FTR countries in terms of private credit is 46 percentage points. Once we control for all the variables, it drops to 26 percentage points. The sign and significance of the control variables remain similar to our previous table.

In Table 3, we consider an alternate measure of financial development—private credit to the real sector by deposit money banks and other financial institutions as a percentage of GDP. As mentioned earlier, this is a broader measure of financial depth. We consider both OLS and quantile regressions including all controls. Column (1) in Table 3 presents OLS regression, while column (2) presents quantile regressions. Here, we find that the impact of the FTR dummy for the OLS regression is stronger than in the case of quantile regression.

While weak FTR countries have 21 percentage points more private credit (including financial institutions) relative to strong FTR countries in the case of OLS regression, the effect drops to 15 percentage points in the case of quantile regression. Labor force participation rate (LFPR) and urban population are positive and significant in both regressions, while GDP per capita growth is negative and significant in both cases. Trade is not significant in Table 3 specifications.

Table 3. Private Credit (including Financial Institutions) and FTR. OLS and Quantile regressions with time fixed effects: The dependent variable is private credit by deposit money banks and other financial institutions as a percentage of GDP. FTR is future time reference dummy with 1 indicating weak FTR countries and strong FTR countries. The controls for labor force participation rate (LFPR), urban population as a percentage of total population, GDP per capita growth, trade as percentage of GDP and polity. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)	(2)
	OLS	**Quantile**
Weak FTR	21.76 ***	15.36 ***
	(4.326)	(5.561)
LFPR	0.869 ***	0.571 ***
	(0.135)	(0.193)
Urban Population	0.411 ***	0.614 ***
	(0.100)	(0.150)
GDP per cap. growth	−1.962 ***	−1.667 **
	(0.562)	(0.752)
Trade	0.043	0.049
	(0.028)	(0.046)
Polity	0.606 *	0.733
	(0.326)	(0.486)
Constant	−5.785	−10.72
	(13.33)	(19.00)
Observations	748	748
R-squared	0.224	NA

5. Robustness Analysis

For our robustness analysis, we conduct an array of tests to make sure our results are not sensitive additional controls, alternate fixed effects, or other measures of our linguistic variable (FTR). We start our robustness analysis by controlling for continent fixed effects within our benchmark specifications, the importance of which is emphasized by Ang (2019). Additionally, we consider regional dummies based on the country income classification by the World Bank, as is commonly used in the literature. As Mavisakalyan et al. (2018) point out, because linguistic features can be spatially correlated, this implies that linguistic features can be concentrated in certain areas. Thus, the effect of FTR can be biased due to geographic and climatic factors that are correlated. We present the results in Table 4 with both measures of financial development considered in our benchmark analysis.

Our main conclusions remain unchanged. The impact for private credit as evident from specification (1) is around 12 percentage points, which again implies that weak FTR countries have more private credit compared to strong FTR countries. As anticipated, controlling for continent fixed effects does reduce the magnitude of FTR dummy relative to previous specifications.

We continue our robustness analysis by controlling for additional variables to further mitigate omitted variable bias. Following the extensive literature on the determinants of financial development, we include different measures of human capital in Table 4 to bolster our benchmark set of controls. As Ibrahim and Sare (2018) found, human capital has a robust influence on financial development, thus creating greater demand for financial intermediation and services that constitute the process of financial development.[10] In Table 4 column (1), we consider a measure of human capital—net primary enrollment. As an alternate measure of human capital, secondary (net) enrollment in considered in column (2).

Table 4. Private Credit and FTR—Including additional controls. OLS regressions with time fixed effects: The dependent variable is private credit by deposit money banks as a percentage of GDP. FTR is future time reference dummy with 1 indicating weak FTR countries and strong FTR countries. The benchmark controls for labor force participation rate (LFPR), urban population as a percentage of total population, GDP per capita growth, trade as percentage of GDP and polity. The additional controls are school enrollment (primary), school enrollment (secondary), foreign direct investment inflows, constraints on the chief executive, durability (democracy) and inflation. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$.

	(1)	(2)	(3)	(4)	(5)
Weak FTR	12.46 ***	11.79 ***	12.94 ***	15.14 ***	13.17 ***
	(3.719)	(3.892)	(3.703)	(3.283)	(3.492)
LFPR	0.905 ***	1.361 ***	0.898 ***	0.514 ***	0.698 ***
	(0.126)	(0.139)	(0.127)	(0.122)	(0.119)
Urban Population	−0.095	−0.325 ***	−0.118	−0.238 **	−0.050
	(0.112)	(0.099)	(0.111)	(0.105)	(0.107)
GDP per cap. growth	−4.297 ***	−3.438 ***	−4.270 ***	−3.580 ***	−4.153 ***
	(0.642)	(0.598)	(0.635)	(0.567)	(0.517)
Trade	0.188 ***	0.123 ***	0.179 ***	0.163 ***	—
	(0.027)	(0.022)	(0.026)	(0.027)	
Polity	1.554 ***	−0.347	—	—	—
	(0.327)	(0.376)			
School enrollment (primary)	2.186 ***	—	2.185 ***	1.946 ***	2.094 ***
	(0.280)		(0.282)	(0.285)	(0.264)
School enrollment (secondary)	—	1.515 ***	—	—	—
		(0.111)			
Constraints (chief exec)	—	—	5.424 ***	—	3.029 ***
			(1.023)		(0.994)
Durable (dem)	—	—	—	0.369 ***	—
				(0.054)	
FDI	—	—	—	—	0.142
					(0.141)
Inflation	—	—	—	—	15.34 ***
					(1.354)
Constant	−212.6 ***	−110.5 ***	−232.5 ***	−161.1 ***	−332.4 ***
	(26.53)	(15.54)	(27.76)	(25.72)	(27.49)
Observations	587	524	585	587	587
R-squared	0.392	0.476	0.397	0.462	0.468

Likewise, political institutions have been shown to be an important determinant of financial development in the literature. As Pagano and Volpin (2001) point out, self-interested policy makers can intervene in financial markets for promotion of group interests. Rajan and Zingales (2003) also emphasize the role that interest groups can play in financial development. As Huang (2010) argues, the presence of a stronger elite group favors the interests of elites and restricts democratic participation. Greater shift of power towards elite groups potentially makes the system more autocratic and results in greater obstacles for financial development. In this context, Girma and Shortland (2008) have shown that both democracy and regime change matters for financial development.

With these in mind, we check the sensitivity of our findings with alternate measures of political institutions (other than polity, which we used in our benchmark regressions). The first variable we use is constraints on the chief executive. Based on the data and definition provided by Marshall et al. (2019), the variable, "refers to the extent of institutionalized constraints on the decision-making powers of chief executives, whether individuals or collectivities". It ranges from 1 to 7, with higher numbers denoting greater constraints as measured by the ability in which "accountability groups" may impose limitations. For example, in Western democracies, these typically take the form of legislatures. Column (3) of Table 4 considers the constraint measure instead of polity.

As an alternate measure, we consider the durability of the political system. Based on the definition set forth by Marshall et al. (2019), it is measured as the number of years "since the last substantive change in authority characteristics (defined as a 3-point change in the POLITY score)". We consider this measure in column (4). Finally, in column (5) of the table, we consider inflation as an additional control variable.

For all the specifications in Table 4, we consider private credit by deposit money banks (excluding other financial institutions) as the dependent variable. As we can see from the table, the coefficient of FTR is positive and significant for all specifications. In terms of magnitude, for weak FTR countries, weak FTR countries have between 11 and 15 percentage points more private credit. These findings, given our additional controls, suggest that greater constraints and a more durable political system enhance financial development. Likewise, the effects of both education measures are positive and significant as well.

Next, we next consider alternate measures of financial development as dependent variables. We present these results in Table 5. The first alternate measure we consider is liquid liabilities as a percentage of GDP. The measure, used by King and Levine (1993), is the broadest indicator of financial intermediation as it encompasses currency as well as interest bearing liabilities of banks and other financial intermediaries. One other alternate measure considered is bank credit as a percentage of bank deposits. As Beck et al. (2009) state, this measure indicates the ratio of claims on the private sector to deposits in money banks. It is a measure of the efficiency of the financial system as it assesses the extent to which "banks intermediate society's savings into private sector credits". The other two measures considered are equity market indicators, including stock market capitalization and stock market total value, both as a percentage of GDP. The size of the equity markets relative to the size of the economy are captured by these two indicators. The first measure, stock market capitalization to GDP, assessing activity of the stock market equals total shares traded on the stock market as a percentage of GDP. The second measure, stock market total value to GDP, also measures activity of the stock market but in terms of trading volume as a share of national output.

In Table 5, we find that the FTR dummy is positive and significant across all measures. The effect of the FTR dummy is the strongest in the case of the liquid liabilities measure, while the magnitude is the least in the case of stock market capitalization measure. In the case of liquid liabilities, weak FTR countries have 34 percentage points more liquid liabilities relative to strong FTR countries. In the case of bank deposits, the magnitude is about 23 percentage points.

Table 5. Alternate Measures of Financial Development and FTR. OLS regressions with time fixed effects: The dependent variables are liquid liabilities (% of GDP), stock market capitalization (% of GDP), stock market total value traded (% of GDP) and bank deposits as percentage of bank credits in columns (1), (2), (3) and (4), respectively. FTR is future time reference dummy with 1 indicating weak FTR countries and strong FTR countries. The benchmark controls for labor force participation rate (LFPR), urban population as a percentage of total population, GDP per capita growth, trade as percentage of GDP and polity. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)	(2)	(3)	(4)
	Liquid Liabilities	**Stock Market Cap.**	**Stock Market Total Val.**	**Bank Deposits**
Weak FTR	35.84 ***	9.224 **	18.77 ***	23.44 ***
	(5.558)	(3.896)	(5.146)	(6.098)
LFPR	−0.307 *	1.179 ***	0.980 ***	1.237 ***
	(0.168)	(0.134)	(0.141)	(0.302)
Urban Population	0.581 ***	0.450 ***	0.137	−0.273
	(0.126)	(0.113)	(0.105)	(0.176)
GDP per cap. growth	−0.470	0.282	−0.185	0.268
	(0.659)	(0.482)	(0.552)	(0.660)
Trade	0.601 ***	0.053 **	−0.207 ***	0.092
	(0.133)	(0.025)	(0.032)	(0.080)
Polity	−1.223 ***	0.355	0.130	−1.913 **
	(0.414)	(0.270)	(0.398)	(0.882)
Constant	1.542	−42.24 ***	5.338	70.73 ***
	(18.04)	(11.62)	(16.61)	(16.98)
Observations	744	601	600	729
R-squared	0.362	0.280	0.210	0.127

For the final part of our robustness analysis, we investigate if strong or weak FTR matters differently for high or low levels of financial development. In other words, do countries speaking weak FTR languages benefit more if they have *lower* levels of financial development relative to countries speaking weak FTR countries but have *higher* levels of financial development? We run quantile regressions for the 25th, 50th and 75th percentiles of financial development based on our sample. In Table 6, we consider our benchmark measure of financial development—private credit. We also consider the other broader measure of private credit—private credit to the real sector by deposit money banks and other financial institutions as a percentage of GDP. In column (1), we consider the private credit measure, and in column (2), we consider the private credit (plus financial institutions) measure. We report the results for our main variable of interest, FTR dummy, for the 25th, 50th and 75th percentiles of financial development.[11]

We find that the FTR dummy has the strongest effect for both measures of financial development for the 25th percentile. For both measures, weak FTR language speaking countries with financial development in the 25th percentile have about 26–27% more private credit or private credit (plus financial institutions) relative to strong FTR language countries in the same financial development percentile. For countries in the highest percentile (75th percentile) of financial development, weak FTR language speaking countries also benefit more than strong FTR language speaking countries but by a lesser magnitude. Thus, across all specifications, we observe that weak FTR languages are associated with enhanced financial development relative to strong FTR countries.

Table 6. Financial Development Percentiles and Private Credit. OLS Regressions with time fixed effects: The dependent variables are liquid liabilities (% of GDP), stock market capitalization (% of GDP), stock market total value traded (% of GDP) and bank deposits as percentage of bank credits in columns (1), (2), (3) and (4), respectively. FTR is future time reference dummy with 1 indicating weak FTR countries and strong FTR countries. The benchmark controls for labor force participation rate (LFPR), urban population as a percentage of total population, GDP per capita growth, trade as percentage of GDP and polity. Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$.

	(1)	(2)
	Private Credit	**Private Credit (including fin. Inst.)**
25th Percentile		
Weak FTR	26.391 ***	27.667 ***
	(6.562)	(6.084)
O.25 Pseudo R	0.155	0.134
50th Percentile		
Weak FTR	26.277 ***	15.355 **
	(6.040)	(7.199)
O.50 Pseudo R	0.171	0.154
75th Percentile		
Weak FTR	16.236 ***	17.922 **
	(4.301)	(7.203)
O.75 Pseudo R	0.179	0.186
Observations	748	748

6. Conclusions

Given the implications of financial development for economic growth and varied development outcomes (Ibrahim and Alagidede 2017b; Mishra and Narayan 2015; Masten et al. 2008; Rioja and Valev 2004; Calderon and Liu 2003), the factors that shape a country's economic development remains an important research question to consider. Our results add to the literature on the determinants of financial development by finding that linguistic structures of countries play an important role in affecting financial development. Specifically, our results show that countries speaking weak future time reference (FTR) language experience enhanced financial development relative to countries speaking strong future time reference (FTR) languages. In light of this, weak FTR languages discount the future relatively less and maintain the connection between present and the future. Due to this, individual speakers of these languages are more likely to work towards creating and bettering property rights institutions, investor protection, efficient corporate governance and information symmetry for all participants in financial markets. In terms of policy implications, it does not seem reasonable to build policies to change language structures. Linguistic structures are exogenous factors and prevalent in countries over the very long term. Yet, being aware of how such language structures can affect financial development can help policy makers to create an environment that can mitigate the adverse effect of strong FTR languages.

We want to point out that this is a preliminary analysis exploring the relationship between linguistic traits and financial development. We want to remind our readers again that our results and economic interpretations should be read as significant correlation between the variables and not as causation. Future studies can establish identification considering external instruments or matching techniques. We have stuck to the benchmark measure of FTR used in the literature, which is the dummy indicating strong or weak FTR countries based on major spoken language. Additional nuanced measures considering language family or verb ration is beyond the scope of our analysis. We hope this study leads to further research on the important topic of language structures and financial markets and institutions.

Author Contributions: N.D. obtained the data and conducted the empirical analysis. N.D. and G.W.C. discussed the results and wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: N. Dutta gratefully acknowledges the financial support from Menard Family Initiative (MFI), University of Wisconsin-La Crosse.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data supporting reported results is available upon request.

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

Notes

1 "Financial development" generally refers to a financial system's ability to engage in information and capital allocation, effective corporate governance, the mobilization of savings into investment, efficient risk management, and minimizing the costs of financial transactions (Dutta and Meierrieks 2021; World Bank 2012).

2 For instance, see Dar and Sahu (2022), who highlight three possible channels through which the effects of FTR are mediated upon financial inclusion, including patience, educational attainment, and tax morale.

3 Importantly, Chen (2013) points out that differences between weak and strong FTR language speakers, "… do not reflect innate cognitive nor early cultural differences between speakers of different languages …".

4 In response to questions regarding spurious correlations in Chen (2013), a study by Roberts et al. (2015) introduced additional controls and robustness analyses to further account for the idea that cultures and languages co-evolve, as opposed languages being strictly exogenous. Controlling for the geographic and historical interrelatedness of languages. Their results yield weaker results than Chen (2013), but continued to find a considerably robust correlation between linguistic features and future-oriented decision making.

5 While a number of studies make the connection between FTR and future decision-making, several recent studies provide contrary evidence. See for instance Chen et al. (2019), Angerer et al. (2021), and Jäggi et al. (2022), who examined the linguistic-savings hypothesis and did not find evidence in support of the hypothesis.

6 Similar to variables like Foreign Direct Investment as a percentage of GDP or Trade as a percentage of GDP, for small countries with small GDP, the private credit share can be much larger.

7 Chen's classification of languages that need an obligatory future time reference is adopted from the European Science Foundation's Typology of Languages in Europe (EUROTYP) project (Dahl 2000). In the case of non-European languages, Chen uses established cross-linguistic analyses. Of the latter, see Cyffer et al. (2009); Nurse (2008); Bybee et al. (1994); Dahl (1985); Dahl and Dienes (1984).

8 That is, while our results are not causal in nature, they are strongly correlative across numerous specifications.

9 For example, the positive and significant association between trade openness and financial development has been document by several studies (Baltagi et al. 2009; Huang and Temple 2005).

10 These findings have been supported with country specific data. For Turkey, for example, see Eryiğit et al. (2015); for India, see Arora and Ratnasiri (2011).

11 Keeping space constraint in mind, additional results are available on request.

References

Adeniyi, Oluwatosin, Abimbola Oyinlola, Olusegun Omisakin, and Festus O. Egwaikhide. 2015. Financial development and economic growth in Nigeria: Evidence from threshold modelling. *Economic Analysis and Policy* 47: 11–21. [CrossRef]

Ang, Desmond. 2019. Do 40-Year-Old Facts Still Matter? Long-Run Effects of Federal Oversight under the Voting Rights Act. *American Economic Journal: Applied Economics* 11: 1–53. [CrossRef]

Angerer, Silvia, Daniela Glätzle-Rützler, Philipp Lergetporer, and Matthias Sutter. 2021. The effects of language on patience: An experimental replication study of the linguistic-savings hypothesis in Austria. *Journal of the Economic Science Association* 7: 88–97. [CrossRef]

Arestis, Philip, Panicos O. Demetriades, and Kul B. Luintel. 2001. Financial development and economic growth: The role of stock markets. *Journal of Money, Credit and Banking* 33: 16–41. [CrossRef]

Arora, Rashmi Umesh, and Shyama Ratnasiri. 2011. Financial Development, Financial Inclusion and Human Capital: How Close is the Link? A Study of India. Paper Presented in the Indian Economy; Australian National University. Available online: https://www.crawford.anu.edu.au/acde/asarc/pdf/papers/conference/INDIA2011_05.pdf (accessed on 1 September 2022).

Baltagi, Badi H., Panicos O. Demetriades, and Siong Hook Law. 2009. Financial development and openness: Evidence from panel data. *Journal of Development Economics* 89: 285–96. [CrossRef]

Beck, Thorsten. 2002. Financial development and international trade: Is there a link? *Journal of International Economics* 57: 107–31. [CrossRef]

Beck, Thorsten, and Ross Levine. 2004. Stock markets, banks, and growth: Panel evidence. *Journal of Banking and Finance* 28: 423–42. [CrossRef]

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. 2009. *Financial Institutions and Markets across Countries and over Time: Data and Analysis*. Policy Research Working Paper: No. 4943. Washington, DC: World Bank.

Beck, Thorsten, Ross Levine, and Norman Loayza. 2000. Finance and the sources of growth. *Journal of Financial Economics* 58: 261–300. [CrossRef]

Begović, Boris, Zorica Mladenović, and Danica Popović. 2017. *Democracy, Distance from the Technological Frontier, and Economic Growth: Some Empirical Results*. CLDS Working Paper #0217. Belgrade: Center for Liberal-Democratic Studies.

Bekaert, Geert, Campbell R. Harvey, and Christian Lundblad. 2005. Does financial liberalization spur growth? *Journal of Financial Economics* 77: 3–55. [CrossRef]

Bhalotra, Sonia, Damian Clarke, Joseph F. Gomes, and Atheendar Venkataramani. 2015. Maternal mortality and female life expectancy: The importance of gender inequality. *Unpublished manuscript*.

Bhattacharyya, Sambit. 2013. Political origins of financial structure. *Journal of Comparative Economics* 41: 979–94. [CrossRef]

Boroditsky, Lera, Lauren A. Schmidt, and Webb Phillips. 2003. Sex, Syntax, and Semantics. In *Language in Mind: Advances in the Study of Language and Cognition*. Edited by Dedre Gentner and Susan Goldin-Meadow. Cambridge, MA: MIT Press.

Boyd, John H., Ross Levine, and Bruce D. Smith. 2001. The impact of inflation on financial sector performance. *Journal of Monetary Economics* 47: 221–48. [CrossRef]

Bybee, Joan L., Revere Dale Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Calderon, César, and Lin Liu. 2003. The direction of causality between financial development and economic growth. *Journal of Development Economics* 72: 321–34. [CrossRef]

Cevik, Serhan, and Mohammad H. Rahmati. 2020. Searching for the finance–growth nexus in Libya. *Empirical Economics* 58: 567–81. [CrossRef]

Chen, Josie I., Tai-Sen He, and Yohanes E. Riyanto. 2019. The effect of language on economic behavior: Examining the causal link between future tense and time preference in the lab. *European Economic Review* 120: 103307. [CrossRef]

Chen, M. Keith. 2013. The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets. *American Economic Review* 103: 690–731. [CrossRef]

Chinn, Menzie D., and Hiro Ito. 2006. What matters for financial development? Capital controls, institutions, and interactions. *Journal of Development Economics* 81: 163–92. [CrossRef]

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague and Paris: Mouton.

Christopoulos, Dimitris K., and Efthymios G. Tsionas. 2004. Financial development and economic growth: Evidence from panel unit root and cointegration tests. *Journal of Development Economics* 73: 55–74. [CrossRef]

Clague, Christopher, Philip Keefer, Stephen Knack, and Mancur Olson. 1996. Property and Contract Rights under Democracy and Dictatorship. *The Journal of Economic Growth* 1: 243–76. [CrossRef]

Clarke, George R. G., Lixin Colin Xu, and Heng-Fu Zou. 2006. Finance and Income Inequality: What Do the Data Tell Us? *Southern Economic Journal* 72: 578–96. [CrossRef]

Cyffer, Norbert, Georg Ziegelmeyer, and Erwin Ebermann. 2009. *Negation Patterns in West African Languages and Beyond*. Amsterdam: John Benjamins Publishing.

Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.

Dahl, Östen. 2000. The Grammar of Future Time Reference in European Languages. In *Tense and Aspect in the Languages of Europe*. Edited by Östen Dahl. Berlin: Mouton de Gruyter.

Dahl, Östen, and Dora Kós Dienes. 1984. *Selected Working Papers from the Tense-Mood-Aspect Project*. Stockholm: Institute of Linguistics, University of Stockholm.

Dar, Shafkat Shafi, and Sohini Sahu. 2022. The effect of language on financial inclusion. *Economic Modelling* 106: 105693. [CrossRef]

de Saussure, Ferdinand. 1916. *Course in General Linguistics*. Edited by Charles Bally and Albert Sechehaye Harris. Translated by Roy Harris. 1983. La Salle: Open Court.

Demetriades, Panicos O., and Khaled A. Hussein. 1996. Does financial development cause economic growth? Time-series evidence from 16 countries. *Journal of Development Economics* 51: 387–411. [CrossRef]

Desbordes, Rodolphe, and Shang-Jin Wei. 2017. The effects of financial development on foreign direct investment. *Journal of Development Economics* 127: 153–68. [CrossRef]

Dutta, Nabamita, and Daniel Meierrieks. 2021. Financial development and entrepreneurship. *International Review of Economics, and Finance* 73: 114–26. [CrossRef]

Eryiğit, Sibel Balı, Yasin Eryiğit Kadir, and Dülgeroğlu Ercan. 2015. Local Financial Development and Capital Accumulations: Evidence from Turkey. *Panoeconomicus* 62: 339–60. [CrossRef]

Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. Time discounting and time preference: A critical review. *Journal of Economic Literature* 40: 351–401. [CrossRef]

Galindo, Arturo, and Alejandro Micco. 2004. Creditor Protection and Financial Markets: Empirical Evidence and Implications for Latin America. *Federal Reserve Bank of Atlanta, Economic Review* Q2: 29–37.

Galor, Oded, Ömer Özak, and Assaf Sarid. 2016. *Geographical Origins and Economic Consequences of Language Structures*. CESifo Working Paper Series No. 6149. Munich: Center for Economic Studies and the Ifo Institute.
Girma, Sourafel, and Anja Shortland. 2008. The political economy of financial development. *Oxford Economic Papers* 60: 567–96. [CrossRef]
Goldsmith, Raymond W. 1969. *Financial Structure and Development*. New Haven: Yale University Press.
Gurley, John G., and Edward S. Shaw. 1967. Financial structure and economic development. *Economic Development and Cultural Change* 34: 333–46. [CrossRef]
Harner, Lorraine. 1981. Children Talk about the Time and Aspect of Actions. 1981. *Child Development* 52: 498–506. [CrossRef]
Herger, Nils, Roland Hodler, and Michael Lobsiger. 2008. What Determines Financial Development? Culture, Institutions or Trade. *Review of World Economics* 144: 558–87.
Hicks, Daniel L., Estefania Santacreu-Vasut, and Amir Shoham. 2015. Does mother tongue make for women's work? Linguistics, household labor, and gender identity. *Journal of Economic Behavior and Organization* 110: 19–44. [CrossRef]
Ho, Sy-Hoa, and Jamel Saadaoui. 2022. Bank credit and economic growth: A dynamic threshold panel model for ASEAN countries. *International Economics* 170: 115–28. [CrossRef]
Huang, Yongfu. 2010. Political institutions and financial development: An empirical study. *World Development* 38: 1667–77. [CrossRef]
Huang, Yongfu, and Jonathan R. W. Temple. 2005. Does External Trade Promote Financial Development? CEPR Discussion Paper No. 5150. Available online: https://ssrn.com/abstract=785244 (accessed on 15 September 2022).
Huybens, Elisabeth, and Bruce D. Smith. 1999. Inflation, Financial Markets and Long-Run Real Activity. *Journal of Monetary Economics* 43: 283–315. [CrossRef]
Ibrahim, Muazu, and Paul Alagidede. 2017a. Financial Development, Growth Volatility and Information Asymmetry in Sub-Saharan Africa: Does Law Matter? *South African Journal of Economics* 85: 570–88. [CrossRef]
Ibrahim, Muazu, and Paul Alagidede. 2017b. Financial sector development, economic volatility and shocks in sub-Saharan Africa. *Physica A* 484: 66–81. [CrossRef]
Ibrahim, Muazu, and Yakubu Awudu Sare. 2018. Determinants of financial development in Africa: How robust is the interactive effect of trade openness and human capital? *Economic Analysis and Policy* 60: 18–26. [CrossRef]
Jaffee, Dwight, and Mark Levonian. 2001. Structure of banking systems in developed and transition economies. *European Financial Management* 7: 161–81. [CrossRef]
Jäggi, Tiziana, Sayaka Sato, Christelle Gillioz, and Pascal Mark Gygax. 2022. Is the future near or far depending on the verb tense markers used? An experimental investigation into the effects of the grammaticalization of the future. *PLoS ONE* 17: e0262778.
Jauch, Sebastian, and Sebastian Watzka. 2015. Financial development and income inequality: A panel data approach. *Empirical Economics* 51: 291–314. [CrossRef]
Johnson, Simon, Peter Boone, Alasdair Breach, and Eric Friedman. 2000. Corporate Governance in the Asian Financial Crisis. *Journal of Financial Economics* 58: 141–86. [CrossRef]
Jung, Woo S. 1986. Financial development and economic growth: International evidence. *Economic Development and Cultural Change* 34: 336–46. [CrossRef]
Kay, Paul, and Terry Regier. 2006. Language, thought and colour: Recent developments. *Trends in Cognitive Sciences* 10: 51–54. [CrossRef] [PubMed]
King, Robert G., and Ross Levine. 1993. Finance and growth: Schumpeter might be right. *Quarterly Journal of Economics* 108: 717–38. [CrossRef]
Kirby, Kris N., and Richard J. Herrnstein. 1995. Preference reversals due to myopic discounting of delayed reward. *Psychological Science* 6: 83–89. [CrossRef]
Levine, Ross. 1997. Financial development and economic growth: Views and agenda. *Journal of Economic Literature* 35: 688–726.
Levine, Ross. 2003. *More on Finance and Growth: More Finance, More Growth?* St. Louis: The Federal Reserve Bank of St. Louis.
Levine, Ross. 2005. Finance and growth: Theory and evidence. In *Handbook of Economic Growth*. Edited by Philippe Aghion and Steven N. Durlauf. Amstedam: Elsevier, pp. 865–934.
Levine, Ross, and Sara Zervos. 1998. Stock markets, banks, and economic growth. *American Economic Review* 88: 537–58.
Levine, Ross, Norman Loayza, and Thorsten Beck. 2000. Financial intermediation and growth: Causality and causes. *Journal of Monetary Economics* 46: 31–77. [CrossRef]
Levinson, Stephen C. 1996. Frames of reference and Molyneux's question: Crosslinguistic evidence. In *Language and Space*. Cambridge, MA and London: MIT Press, pp. 109–69.
Levinson, Stephen C., and David P. Wilkins. 2006. *Grammars of Space: Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press.
Licht, Amir N., Chanan Goldschmidt, and Shalom H. Schwartz. 2007. Culture rules: The foundations of the rule of law and other norms of governance. *Journal of Comparative Economics* 35: 659–88.
Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2019. *Polity IV Project: Political Regime Characteristics and Transitions, 1800–2018*. Center for Systemic Peace: Available online: https://www.systemicpeace.org/inscr/p4manualv2018.pdf (accessed on 1 September 2022).
Masten, Arjana Brezigar, Fabrizio Coricelli, and Igor Masten. 2008. Non-linear growth effects of financial development: Does financial integration matter? *Journal of International Money and Finance* 27: 295–313. [CrossRef]

Mavisakalyan, Astghik. 2015. Gender in Language and Gender in Employment. *Oxford Development Studies* 43: 403–24. [CrossRef]
Mavisakalyan, Astghik, and Clas Weber. 2018. Linguistic Structures and Economic Outcomes. *Journal of Economic Survey* 32: 916–39. [CrossRef]
Mavisakalyan, Astghik, Yashar Tarverdi, and Clas Weber. 2018. Talking in the present, caring for the future: Language and environment. *Journal of Comparative Economics* 46: 1370–87. [CrossRef]
Mishra, Sagarika, and Paresh Kumar Narayan. 2015. A nonparametric model of financial system and economic growth. *International Review of Economics and Finance* 39: 175–91. [CrossRef]
Nikoloski, Zlatko. 2012. Financial Sector Development and Inequality: Is there a Financial Kuznets curve? *Journal of International Development* 25: 897–911. [CrossRef]
Nkusu, Mwanza, and Selin Sayek. 2004. Local Financial Development and the Aid-Growth Relationship. IMF Working Paper No. 04/238. Available online: https://ssrn.com/abstract=879063 (accessed on 1 September 2022).
Nurse, Derek. 2008. *Tense and Aspect in Bantu*. Oxford: Oxford University Press.
Nwani, Chinazaekpere, and Jacob Bassey Orie. 2016. Economic growth in oil-exporting countries: Do stock market and banking sector development matter? Evidence from Nigeria. *Cogent Economics and Finance* 4: 1153872. [CrossRef]
Pagano, Marco, and Paolo Volpin. 2001. The Political Economy of Finance. *Oxford Review of Economic Policy* 17: 502–19. [CrossRef]
Patrick, Hugh T. 1966. Financial development and economic growth in underdeveloped countries. *Economic Development and Cultural Change* 14: 174–89. [CrossRef]
Pinker, Steven. 1994. *The Language Instinct: How the Mind Creates Language*. New York: William Morrow and Company.
Quixina, Yuri, and A. Almeida. 2014. *Financial Development and Economic Growth in a Natural Resource Based Economy: Evidence from Angola*. FEP Working Papers No: 542. Porto: School of Economics and Management, University of Porto.
Rajan, Raghuram G., and Luigi Zingales. 2003. The great reversals: The politics of financial development in the twentieth century. *Journal of Financial Economies* 69: 5–50. [CrossRef]
Ramsey, Frank Plumpton. 1928. A mathematical theory of saving. *Economic Journal* 38: 543–49. [CrossRef]
Rewilak, Johan. 2017. The Impact of Financial Crises on the Poor. *Journal of International Development* 7: 169–76. [CrossRef]
Rioja, Felix, and Neven Valev. 2004. Does one size fit all? A re-examination of the finance and growth relationship. *Journal of Development Economics* 74: 429–47. [CrossRef]
Roberts, Seán G., James Winters, and Keith Chen. 2015. Future tense and economic decisions: Controlling for cultural evolution. *PLoS ONE* 10: e0132145. [PubMed]
Roe, Mark J., and Jordan I. Siegel. 2011. Political Instability: Its Effects on Financial Development, Its Roots in the Severity of Economic Inequality. *Journal of Comparative Economics* 39: 279–309. [CrossRef]
Svaleryd, Helena, and Jonas Vlachos. 2002. Markets for risk and openness to trade: How are they related? *Journal of International Economics* 57: 369–95. [CrossRef]
Samargandi, Nahla, Jan Fidrmuc, and Sugata Ghosh. 2014. Financial development and economic growth in an oil-rich economy: The case of Saudi Arabia. *Economic Modelling* 43: 267–78. [CrossRef]
Santacreu-Vasut, Estefania, Amir Shoham, and Victor Gay. 2013. Do female/male distinctions in language matter? Evidence from gender political quotas. *Applied Economics Letters* 20: 495–98. [CrossRef]
Santacreu-Vasut, Estefania, Oded Shenkar, and Amir Shoham. 2014. Linguistic gender marking and its IB ramifications. *Journal of International Business Studies* 45: 1170–78. [CrossRef]
Slobin, Dan I. 2003. Language and thought online: Cognitive consequences of linguistic relativity. In *Language in Mind: Advances in the Investigation of Language and Thought*. Edited by Dedre Gentner and Susan Goldin-Meadow. Cambridge, MA: MIT Press, pp. 157–91.
Solnick, Jay V., Catherine H. Kannenberg, David A. Eckerman, and Marcus B. Waller. 1980. An experimental analysis of impulsivity and impulse control in humans. *Learning and Motivation* 11: 61–77. [CrossRef]
Szagun, Gisela. 1978. On the frequency of use of tenses in English and German children's spontaneous speech. *Child Development* 49: 898–901. [CrossRef]
Tabellini, Guido. 2008. Presidential address institutions and culture. *Journal of the European Economic Association* 6: 255–94. [CrossRef]
Wittgenstein, Ludwig. 1922. *Tractatus Logico-Philosophicus*. London: Routledge and Kegan Paul.
World Bank. 2012. *Global Financial Development Report 2013: Rethinking the Role of the State in Finance*. Available online: https://www.worldbank.org/en/publication/gfdr (accessed on 15 September 2022).
Xu, Zhenhui. 2000. Financial development, investment, and economic growth. *Economic Inquiry* 38: 331–44. [CrossRef]
Yang, Benhua. 2011. Does democracy foster financial development?: An empirical analysis. *Economic Letters* 112: 262–65. [CrossRef]

economies

Article

Income-Related Inequality in Health Care Utilization and Out-of-Pocket Payments in China: Evidence from a Longitudinal Household Survey from 2000 to 2015

Miaoqing Yang [1,*] and Guido Erreygers [2,3]

[1] Centre for Guidelines, National Institute for Health and Care Excellence, 2 Redman Place, London E20 1JQ, UK
[2] Department of Economics, Faculty of Applied Economics, University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium
[3] Centre for Health Policy, University of Melbourne, 207 Bouverie Street, Carlton, VIC 3010, Australia
* Correspondence: miaoqing.yang@nice.org.uk

Abstract: In recent decades, China has experienced rapid economic growth and rising health inequality. The government has introduced a nationwide health care reform aimed at achieving affordable and equitable basic health care for all. This paper investigates income-related inequality in health care utilization and out-of-pocket (OOP) payments and explores the underlying factors that drive the inequalities. Using data running from 2000 to 2015 and covering nine of thirty-one provinces in China, we calculate indices to measure income-related inequality and adopt a regression-based decomposition approach to explore the sources of inequality. We find pro-rich inequality in the use of preventive care and pro-poor inequality in the use of folk doctors. In addition, the better-off have preferential access to higher level hospitals, while the use of primary care facilities is more concentrated among the poor. The poor are also found to face a heavier financial burden since they tend to spend a larger share of their income on OOP payments. Education, employment and geographic regions all appear to contribute to the total inequality. Our results indicate that affordability remains a common barrier for the poor to access health care, and that the inequality is largely driven by socio-economic factors.

Keywords: income-related inequality of health; health care utilization; out-of-pocket payments; decomposition analysis; China

Citation: Yang, Miaoqing, and Guido Erreygers. 2022. Income-Related Inequality in Health Care Utilization and Out-of-Pocket Payments in China: Evidence from a Longitudinal Household Survey from 2000 to 2015. *Economies* 10: 321. https://doi.org/10.3390/economies10120321

Academic Editors: Ralf Fendel, Robert Czudaj and Sajid Anwar

Received: 25 October 2022
Accepted: 6 December 2022
Published: 13 December 2022

1. Introduction

Many low- and middle-income countries are seeking ways to pursue the goal of equity in health care utilization. The Chinese experience is important to understand how health care inequality changed during the course of a transition from a command to a market economy. China's market-oriented reforms in the late 1970s brought higher efficiency in the economy and dramatically increased household income; however, they also led to health inequalities due to market failures in health sectors and the removal of social health insurance as a safety net. Health care financing relied heavily on out-of-pocket (OOP) payments, and the poor usually had limited access to necessary health services (Gong and Brixi 2005). Both demand-side subsidies and supply-side infrastructure investments disproportionately served the better-off (Wagstaff et al. 2009), leading to a widening gap in health status and utilization across income groups (Liu et al. 1999; Gao et al. 2002; Zhang and Kanbur 2005; Tang et al. 2008; Wagstaff 2009a). Along with aging and disease transitions from infectious to chronic conditions, the poor were far more vulnerable to the financial and physical consequences of illness.

In response to increasing pressure for equitable access to quality care, from the end of the 1990s onwards the Chinese government introduced a series of health care reforms that incorporated a number of pro-poor measures. Since the effects of these health care reforms differ across income groups, it is important to examine the pre- and post-reform changes in the distribution of health care utilization and medical expenditure and to estimate the contributions of various factors to the observed inequality. An increasing body of studies in recent years attempted to compare the income-related inequality in terms of the use of outpatient and inpatient care (Wang et al. 2012; Xie 2011; Zhang et al. 2015; Zhou et al. 2011; Chen et al. 2015), preventive care (Yang 2013), maternal health services (Li et al. 2015; Shen et al. 2014) and treatment of major chronic conditions (Elwell-Sutton et al. 2013; Xie et al. 2014). Previous evidence showed that the health care reforms were characterised by an overall improvement in insurance coverage, but the rich still seemed to have better access to health care compared to the poor. The impact of health insurance on financial protection appeared to be limited, and in some cases there was even a widening of social disparities in health care access among the insured.

While previous studies have provided important insights into the socio-economic differentials in health care access, most of them focus on either earlier periods of the health care reforms or a limited number of geographic regions, so that we still know little about the longer term trends for the total population following the reforms. In this paper, we use a recent dataset over a 15-year period from 2000 to 2015 to capture the long-term impact of policy changes. The data also cover nine provinces that spread across the eastern, middle and western areas of China, and therefore provide a much broader picture of health care inequality. We assess the evolution of health care inequality and its determinants during the period of rapid economic development and the implementation of the nationwide health care reforms. By examining how types of health services and choice of facilities differ among people with different income levels, we explore both financial and non-financial access barriers related to insurance coverage and quality of care. We calculate both rank- and level-dependent indices to measure income-related inequality and to obtain robust results (Erreygers and Kessels 2017). We also adopt a direct regression-based decomposition of inequality indices to explore the sources of inequality, taking into account the correlation of health and income (Kessels and Erreygers 2019). Our empirical findings could feed back into the policy making process in China and other developing countries to move towards an efficient and equitable health care system.

2. Background

The Chinese government launched three main social health schemes to achieve universal coverage: the Urban Employee Basic Medical Insurance (UEBMI) for the urban employees and retirees, the New Cooperative Medical Scheme (NCMS) for the rural residents and the Urban Resident Basic Medical Insurance (URBMI) for the unemployed urban population (including students and children). Initiated in 1998, UEBMI is a compulsory scheme based on employment, but it only provides coverage to formal-sector workers in urban areas and leaves the majority of the population uninsured. During the 2000s, two new voluntary insurance schemes were introduced, both of which were heavily subsidised by the central and local governments. The NCMS was launched in 2003 and expanded rapidly from 13 to 97.5% of the rural population of about 800 million between 2003 and 2008 (Yang and Wu 2017). The URBMI targeted 420 million urban residents who were not covered by the UEBMI (e.g., the elderly, students, children and the unemployed) and was first implemented in 79 pilot cities in 2007 (Dong 2009). It was then extended to other cities and covered about 49% of urban residents in 2015 (Si 2021). Universal coverage was nearly achieved under these three schemes by the end of 2015, but the reimbursement and types of services covered remained limited. The benefit packages varied geographically, but a typical package covered inpatient services and catastrophic outpatient services. Beneficiaries needed to bear most of their outpatient expenses and about half of their inpatient costs. The average reimbursement rates for inpatient care ranged from 65–68% for the UEBMI, 44–48% for the URBMI and 38–44% for the NCMS, considering the deductibles, co-payments and reimbursement cap (Yip et al. 2012). Over time, the government aimed to gradually extend insurance coverage to more types of care and to reduce co-payments. The central government also assisted local governments in relatively poor regions of western and middle provinces, while the funding of insurance premiums fell solely on local governments in eastern provinces.

With effective risk-sharing at the community/city level, the insurance schemes have the potential to offer better protection for individuals from low- and middle-income backgrounds. In addition, the insurance reimbursement varies with the level and type of health facilities, with much more generous reimbursement rates for low-level facilities. The various reimbursement rate schedules provide incentives for people to seek care from primary care facilities and purchase generic drugs in order to contain overall medical costs. However, due to inadequate resources and insufficient medical training, primary care facilities tend to provide low quality care and are more likely to misdiagnose or inappropriately treat their patients (Sylvia et al. 2014; Li et al. 2017a; Sylvia et al. 2017; Wong et al. 2017). During the past decades, they mainly catered for economically disadvantaged patients, who are highly price-elastic and tend to compromise on quality. In absence of a strict referral or gatekeeping health care system, wealthy patients would rather bypass primary care and seek care at high-level hospitals even for minor conditions (Babiarz et al. 2010; Sylvia et al. 2017). To strengthen the primary care system, recent reforms increased government funding for community health centres in cities and township health centres and village clinics in rural areas. A number of measures, such as imposing strict licensing requirements and promoting regular in-service medical trainings, were implemented to update the clinicians' professional knowledge and skills that are essential to the provision of appropriate patient care (Yi et al. 2020). These measures encourage people to switch from hospitals to primary care facilities and aim to benefit more low-income households. Therefore, both the demand- and supply-side measures have the potential to improve the general population's access to health care and to reduce socio-economic disparities in health and health care.

3. Methodology

3.1. Data and Variables

We use data from the China Health and Nutrition Survey (CHNS), a large-scale panel dataset that employed a multi-stage stratified sampling method to select households from 12 provinces and municipal cities in China, spread across the eastern, central and western regions. The selected households were followed for 10 waves (from 1989 to 2015), and surveyed on a wide range of topics including demographics, socio-economic characteristics and health outcomes (Popkin et al. 2010).

In this study, we use six waves of data from 2000 to 2015 and exclude three newly added municipal cities from 2011. Data before 2000 are not used since the structure of the early questionnaires was different from that of the following waves. We measure income-related inequality for the use of different types of health services (formal medical care, preventive care, folk doctors and inpatient care) and different levels of health facilities (from low-level to high-level facilities: village clinics/community health centres, township health centres, county hospitals and city hospitals). We also consider the burden of OOP payments, defined as the proportion of the OOP payments in the last month to the total monthly per capita household income following previous literature (Wagstaff and Lindelow 2008). We censored the maximum value of this variable at 100% in order to eliminate extremely high OOP values for individuals belonging to households with a very low income (3.25% of the sample). To measure living standards, we adjust household income by applying the Organisation for Economic Co-operation and Development (OECD)-modified equivalence scale to household income, assigning a value of 1 to the household head, 0.5 to each additional adult and 0.3 to each child (OECD n.d.). The CHNS income measure is net monetary income received by the household members and includes income from farming, fishing, gardening, livestock and small commercial household business. Notice that we exclude households with negative income, so the final sample size consists of 24,762 individuals, 6789 households and 67,856 person-wave observations. About 34.9% of respondents were only interviewed once in the survey, 20.2% twice, 13.8% three times, 9.8% four times, 10.2% five times, and 11.3% in all waves. The attrition rate is quite high, and individuals reporting more use of formal medical care and higher level of OOP payments were more likely to drop out of the sample.

In the decomposition analysis we explore how the major determinants of individuals' care-seeking behaviour are associated with income-related health inequality. We look at health-related (number of major diseases, number of symptoms and illness status during the past four weeks preceding the survey), demographic (age, gender, number of children, ethnicity, urban/rural residential status, marital status) and socio-economic characteristics (house ownership, education levels, employment, occupation, social health insurance coverage and geographics). Specifically, health status is thought to be the most important factor that drives utilization of health care. Since self-rated health status is not available in the CHNS 2009 and 2011 surveys, we use the number of major diseases and illness status during the past four weeks preceding the survey as proxy variables for patients' health status (O'Donnell and Propper 1991; Van Doorslaer et al. 1992). We create 14 variables by interacting age categories with gender and control for other demographics such as ethnicity, urban/rural residential status, marital status. The socio-economic characteristics include house ownership, education levels and job status. Household ownership can be a good indicator of the households' wealth level in addition to income, especially for rural households who live from subsistence farming and have little or no income. Better education can either lead to an increase in health care use or better health status that results in lower needs for health care. Job status is particularly relevant in the Chinese context since most state welfare benefits (including UEBMI) are associated with the types of industry. For example, we would expect that state government officials are more likely than self-employed businessmen to have better access to health care facilities because they are granted more generous state welfare benefits. We also investigate the impact of geographic factors by dividing the nine provinces into three groups: eastern, middle and western regions. The eastern coastal area is generally more affluent and supplied with better quality medical infrastructures and services than the middle and western areas. The coverage of the three social health insurance schemes is also included: UEBMI and URBMI for urban residents and NCMS for rural residents. Table 1 shows the definitions of all variables and their summary statistics.

Table 1. Number of observations, mean, standard deviation of all variables, pooling all years from 2000 to 2015.

Variables	Definition	N [1]	Mean	SD [1]
Outcome variables				
Last 4 weeks: formal medical care use	Dummy, 1 if used and 0 otherwise	67,517	0.118	0.323
Last 4 weeks: preventive care use	Dummy, 1 if used and 0 otherwise	67,856	0.038	0.190
Last year: folk doctor use	Dummy, 1 if used and 0 otherwise	67,856	0.030	0.170
Last 4 weeks: inpatient care use	Dummy, 1 if used and 0 otherwise	67,856	0.011	0.103
Last 4 weeks: village clinics/community health centres use	Dummy, 1 if used and 0 otherwise	67,856	0.030	0.170

Table 1. *Cont.*

Variables	Definition	N [1]	Mean	SD [1]
Last 4 weeks: town hospitals use	Dummy, 1 if used and 0 otherwise	67,856	0.020	0.139
Last 4 weeks: county hospitals use	Dummy, 1 if used and 0 otherwise	67,856	0.015	0.121
Last 4 weeks: city hospitals use	Dummy, 1 if used and 0 otherwise	67,856	0.025	0.157
Last 4 weeks: OOP [2] burden	Proportion (%), Monthly OOP of the household/annual household income	67,856	0.019	0.100
Independent variables				
Equivalised household income (thousands)	Annual household income adjusted by an equivalence factor	67,856	29.306	96.984
Female × age 18 below	Interaction term between age and gender	67,851	0.084	0.278
Female × age 18–24	Interaction term between age and gender	67,851	0.028	0.166
Female × age 25–34	Interaction term between age and gender	67,851	0.060	0.237
Female × age 35–44	Interaction term between age and gender	67,851	0.088	0.283
Female × age 45–54	Interaction term between age and gender	67,851	0.095	0.294
Female × age 55–64	Interaction term between age and gender	67,851	0.079	0.269
Female × age 65 above	Interaction term between age and gender	67,851	0.075	0.263
Male × age 18 below (reference group)	Interaction term between age and gender	67,851	0.098	0.297
Male × age 18–24	Interaction term between age and gender	67,851	0.028	0.166
Male × age 25–34	Interaction term between age and gender	67,851	0.056	0.229
Male × age 35–44	Interaction term between age and gender	67,851	0.079	0.269
Male × age 45–54	Interaction term between age and gender	67,851	0.088	0.283
Male × age 55–64	Interaction term between age and gender	67,851	0.075	0.264
Male × age 65 above	Interaction term between age and gender	67,851	0.065	0.247
Number of major diseases	Number of diseases: hypertension, diabetes, heart disease, stroke, bone fracture	67,856	0.167	0.453
Sickness in the last month	Dummy, 1 if sick and 0 otherwise	67,414	0.189	0.391
Number of symptoms in the last month	Number of symptoms: fever, cough, diarrhoea, asthma and headache, etc.	67,856	0.252	0.649
Private dwelling	Dummy, 1 if owning a house/flat and 0 otherwise	67,686	0.918	0.275
Social health insurance	Dummy, 1 if insured and 0 otherwise	67,832	0.513	0.500

Table 1. *Cont.*

Variables	Definition	N [1]	Mean	SD [1]
Ethnicity	Dummy, 1 if Han ethnicity and 0 if ethnic minority	67,581	0.861	0.346
Marital status	Dummy, 1 if married and 0 otherwise	57,856	0.781	0.414
Education: illiterate (reference group)	Dummy, 1 if illiterate and 0 otherwise	67,856	0.182	0.386
Education: primary school	Dummy, 1 if finished primary school and 0 otherwise	67,856	0.197	0.398
Education: junior high school	Dummy, 1 if finished junior high school and 0 otherwise	67,856	0.295	0.456
Education: senior high school	Dummy, 1 if finished senior high school and 0 otherwise	67,856	0.173	0.379
Education: university degree	Dummy, 1 if had a university degree and 0 otherwise	67,856	0.058	0.233
In employment	Dummy, 1 if currently working and 0 otherwise	56,493	0.595	0.491
Occupation: white collar worker	Dummy, 1 if white collar worker and 0 otherwise	56,525	0.105	0.307
Occupation: farmer	Dummy, 1 if farmer and 0 otherwise	56,525	0.263	0.440
Rural resident	Dummy, 1 if rural residents and 0 otherwise	67,372	0.598	0.490
Geographics: east region	Dummy, 1 if living in the eastern region and 0 otherwise	67,856	0.227	0.419
Geographics: middle region	Dummy, 1 if living in the middle region and 0 otherwise	67,856	0.458	0.498
Geographics: west region (reference group)	Dummy, 1 if living in the western region and 0 otherwise	67,856	0.274	0.446
Number of children aged 0–4	Number of children aged 0–4 living in the household	67,856	0.780	1.169
Number of children aged 5–14	Number of children aged 5–14 living in the household	67,856	1.853	2.069

[1] N refers to the total number of observations that pool the repeated observations of the same individuals over the years, and SD to the standard deviation. [2] OOP stands for out-of-pocket expenditures.

3.2. Measurement of Inequality

Our first goal in this paper is to measure to what extent the health outcomes we have selected are related to incomes, and to examine whether and how these relationships have changed over the study period. Put differently: is there any evidence that wealthy people tend to have better access to health services than poor people, and have the health reforms changed anything? It is customary to use indices to measure the degree of socioeconomic inequality. Since we are looking at the joint distribution of health and income, the indices must be of the bivariate type. Positive index values indicate that health and income are positively correlated, and negative values that they are negatively correlated. Due to a lack of reliability in self-reported health measures in the setting of low-income countries (Van Doorslaer and O'Donnell 2011), we focus on measuring inequality in the allocation of health care resources without standardizing for the differences in health needs.

Broadly speaking, two types of bivariate indices of socioeconomic inequality of health can be distinguished: rank-dependent indices, such as the well-known concentration index (CI) (Wagstaff et al. 1991), and level-dependent indices. Rank-dependent indices measure the degree of correlation between health levels and income ranks, and can be expressed as weighted sums of health levels, where the weights are defined by a function of the income ranks (Coveney et al. 2016). The standard (or relative) CI is usually applied to non-negative ratio-scale health variables (Erreygers and Van Ourti 2011). Given that our health care and OOP burden variables are bounded variables, we use a modified version of the CI developed for bounded health variables (Erreygers 2009; Erreygers and Van Ourti 2011). However, since rank-dependent indices only rely on income ranks, they ignore relevant information about the levels of income (Erreygers and Kessels 2017). Level-dependent indices are similar to rank-dependent indices, but take into account the income levels rather than the income ranks. They too can be expressed as weighted sums of health levels, but the weights are now determined by a function of the income levels. These indices exploit more information about the income distribution and measure both income and health consistently by their levels. Additionally, in this case we have to use a modified version appropriate for bounded variables. The precise definitions of the indices calculated in this paper can be found in Appendix A.

3.3. Decomposition of Inequality

Our second goal is to increase our understanding of the determinants of income-related inequalities. To this end, we decompose the inequality indices by means of demographic, socio-economic and health-related variables at the individual level. The conventional regression-based decomposition approach rests on a regression of the health variable only (Wagstaff et al. 2003), and for this reason has been subjected to criticism (Erreygers and Kessels 2013). In recent years, two alternative methods have been developed. The first is based on the recentred influence function approach (Heckley et al. 2016) and has already been applied to Chinese data (Cai et al. 2017). In this study we employ a new approach, based on a regression of a composite variable that incorporates both health and income (Kessels and Erreygers 2019). The idea is that this variable can be interpreted as an indicator of an individual's deviation from a reference position in the income-health space, with the reference position determined by average health and average income. The exact definitions of the dependent variables of our decomposition regressions can be found in Appendix A.

We apply ordinary least square (OLS) regressions to estimate the marginal effects of each individual variable on the inequality index. Previous studies found there is little difference between OLS and non-linear models for decomposition analysis, while the approximation techniques required by non-linear models might introduce additional errors (Van Doorslaer et al. 2004; Van Doorslaer and Masseria 2004; Van Doorslaer et al. 2000). A positive (negative) regression coefficient means that the associated explanatory variable is positively (negatively) correlated with both income and health. In contrast to what is often done in applications of the conventional regression-based decomposition technique, we do not estimate the contribution of each factor to the inequality indices. Instead, we calculate the logworth values based on the p value of the F tests to evaluate the relative importance of the (groups of) variables in influencing the correlation between the income and health dimensions.

4. Results

4.1. Income-Related Inequality in Health Care Utilization and OOP Burden

Tables 2 and 3 present the rank- and level-dependent indices measuring income-related inequality for health care utilization and OOP burden. Broadly speaking, both rank- and level-dependent indices give similar results in terms of the direction of inequality and its significance. There is substantial pro-rich inequality in the use of preventive care and pro-poor inequality in the use of folk doctors. The results suggest that low-income people have limited access to preventive services and are more likely to use folk doctors, who are traditional Chinese medical practitioners in rural areas. They are usually less qualified providers who received minimal basic medical and paramedical training and had no more than middle or high school education (Li et al. 2017b). They offer cheaper services compared to formal health providers, but also tend to conduct unnecessary or even dangerous practices. However, some folk doctor care also includes traditional Chinese medicine, which is considered appropriate in some clinical settings (Cui et al. 2004; Chen et al. 2008; Xiang et al. 2019; Harmsworth and Lewith 2001).

Table 2. Rank-dependent indices for income-related inequality of health care utilization and medical expenditure in China.

Health Variable [1]	2000	2004	2006	2009	2011	2015
Formal care	—	−0.008	−0.042 ***	−0.028 ***	0.007	−0.025 **
	—	0.014	0.012	0.010	0.020	0.012
Preventive care	0.017 ***	0.030 ***	0.015 *	0.024 ***	0.047 ***	0.002
	0.004	0.006	0.009	0.006	0.010	0.006
Folk doctors	−0.002	−0.017 ***	−0.032 ***	−0.026 ***	−0.027 ***	−0.030 **
	0.001	0.005	0.010	0.007	0.009	0.012
Inpatient care	0.001	0.000	0.003	−0.002	−0.003	−0.006 **
	0.002	0.002	0.003	0.003	0.003	0.003
Village clinics/community health centres	−0.007 **	−0.020 **	−0.034 ***	−0.026 ***	−0.025 ***	−0.016 ***
	0.003	0.007	0.006	0.007	0.006	0.006
Town hospitals	−0.006 ***	−0.007 *	−0.014 ***	−0.010 ***	0.014	−0.007
	0.002	0.004	0.004	0.004	0.009	0.006
County hospitals	0.003	0.005	−0.003	0.004	0.003	0.001
	0.002	0.005	0.004	0.003	0.005	0.004
City hospitals	0.008 **	0.021 **	0.010 *	0.017 ***	0.030 ***	0.010 **
	0.004	0.011	0.005	0.006	0.011	0.005
OOP burden	−0.012 ***	−0.045 ***	−0.054 ***	−0.060 ***	−0.049 ***	−0.056 ***
	0.005	0.008	0.008	0.006	0.007	0.007
N	13,457	10,807	10,311	10,505	11,544	11,232

[1] For each outcome, the first row shows coefficients and the second standard errors. * indicates statistically significant at the 10% level, ** at the 5% level and *** at the 1% level.

Table 3. Level-dependent indices for income-related inequality of health care utilization and medical expenditure in China.

Health Variable [1]	2000	2004	2006	2009	2011	2015
Formal care	—	0.000	−0.016 ***	−0.010 **	0.005	−0.010
	—	0.006	0.005	0.005	0.009	0.008
Preventive care	0.009 ***	0.014 ***	0.006	0.006 **	0.022 ***	0.002
	0.002	0.003	0.004	0.002	0.004	0.004
Folk doctors	0.000	−0.008 ***	−0.011 ***	−0.012 ***	−0.011 ***	−0.016 ***
	0.001	0.002	0.004	0.003	0.003	0.005
Inpatient care	0.002	0.001	0.000	0.000	0.000	0.002
	0.001	0.001	0.001	0.001	0.001	0.006
Village clinics/ community health centres	−0.003 ***	−0.008 ***	−0.013 ***	−0.008 ***	−0.011 ***	−0.006 *
	0.001	0.002	0.002	0.003	0.002	0.003
Town hospitals	−0.002 **	−0.003 **	−0.005 ***	−0.004 ***	0.006	−0.005 **
	0.001	0.001	0.002	0.001	0.004	0.002
County hospitals	0.003 *	0.002	−0.002	0.000	0.001	−0.001
	0.002	0.003	0.001	0.002	0.002	0.002
City hospitals	0.006 *	0.009 *	0.005	0.008 ***	0.013 **	0.007
	0.003	0.005	0.003	0.003	0.005	0.005
OOP burden	−0.003	−0.016 ***	−0.021 ***	−0.022 ***	−0.016 ***	−0.024 ***
	0.002	0.003	0.002	0.002	0.002	0.003
N	13,457	10,807	10,311	10,505	11,544	11,232

[1] For each outcome, the first row shows coefficients and the second standard errors. * indicates statistically significant at the 10% level, ** at the 5% level and *** at the 1% level.

Significant values of the inequality indices are found for the health categories preventive care, folk doctors, village clinics/community health centres, and city hospitals. For clarity we represent the values in Figures 1 and 2, respectively for rank-dependent indices and level-dependent indices. In terms of health facility use, the direction of inequality varies by provider levels where pro-poor inequality is observed for the use of village clinics/community health centres and pro-rich inequality for the use of city hospitals. Wealthier people seem to have better access to high-level hospitals that offer more sophisticated care and require higher OOP expenditures.

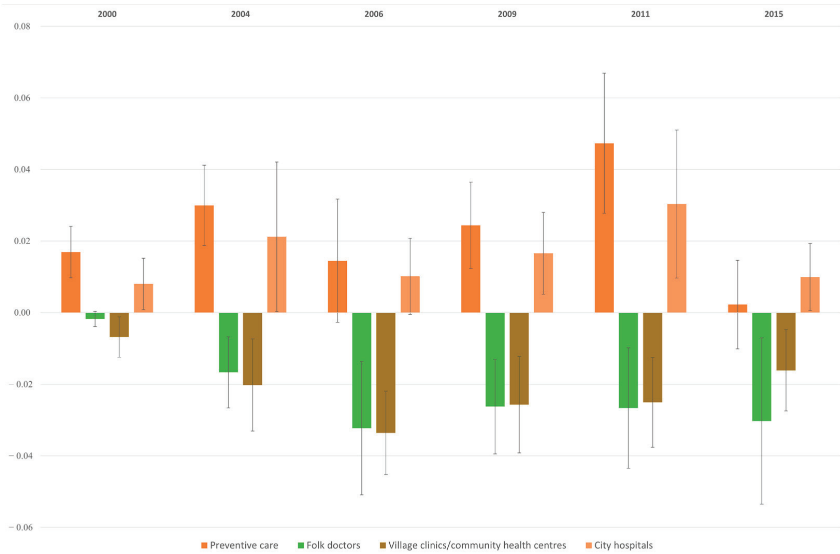

Figure 1. Rank-dependent indices for income-related inequality of health care utilization in China.

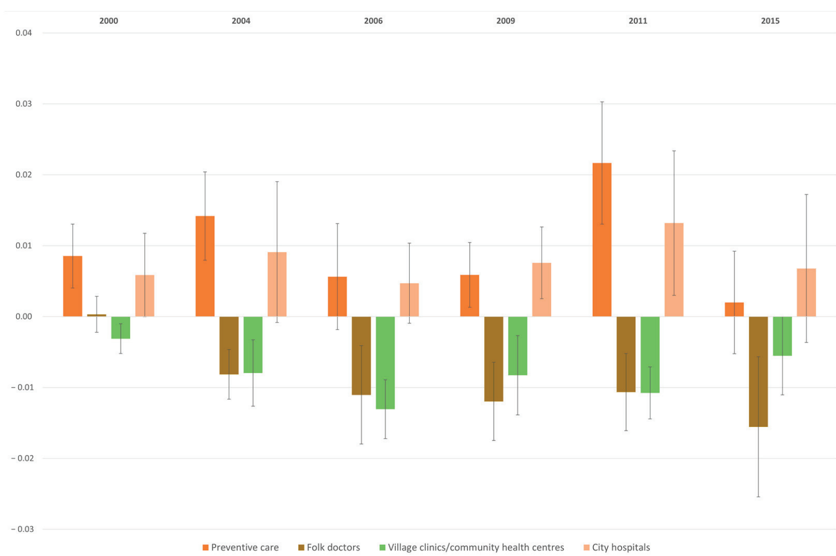

Figure 2. Level-dependent indices for income-related inequality of health care utilization in China.

Even though the rich tend to use more expensive health care facilities than the poor, OOP expenditures seem to impose a higher weight on the poor than on the rich. As we defined the OOP burden as the proportion of the absolute amount of OOP payments relative to the per capita household income, a pro-poor distribution of the OOP burden indicates that in relative terms OOP expenditures tend to fall more heavily on the poor than on the rich (see Figure 3). In spite of the rapid expansion of social health insurance and other health care reform efforts, the disparities in health care utilization across incomes remain similar over time.

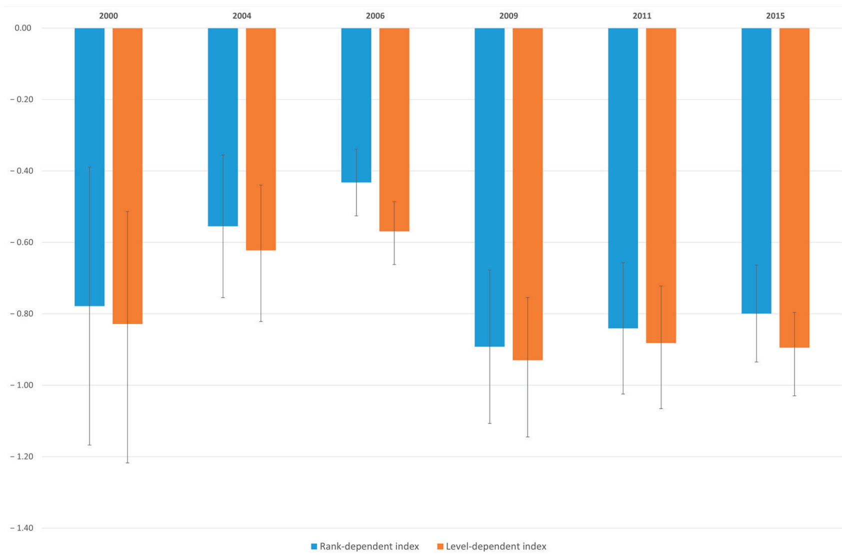

Figure 3. Rank- and level-dependent indices for income-related inequality of OOP burden in China.

4.2. Decomposition of Income-Related Inequality in OOP Burden

The regression-based decomposition analysis we introduced above can be applied to income-related inequality of both health care utilization and OOP burden, allowing us to identify which covariates have a significant effect on each form of inequality. Here, we apply it to the burden of OOP only. Note that since OOP burden is a bounded ill-health variable, we have to define the dependent variable of our regression as explained in Appendix A. Income is excluded as an explanatory variable since it would distort the explanation of the correlation between income and health (Erreygers and Kessels 2013). For each wave of the survey we estimate two regressions, one for the rank-dependent index and the other for the level-dependent index, based on the same set of independent variables. We present both the marginal effects and the logworth values for key covariates to assess their importance as explanatory variables in Table 4 (rank-dependent indices) and Table 5 (level-dependent indices). The full results of all covariates can be found in Appendix B. A logworth value larger than 1.3 indicates that a variable (or group of variables) is significant at the 5% level and a logworth value above 2 indicates significance at the 1% level. For variables with more than two categories, the logworth values are combined over the categories.

Table 4. Effect of selected demographic and socioeconomic variables in the decomposition of the rank-dependent indices for OOP burden.

Variable [1]	2000 Coefficient	2000 Logworth	2004 Coefficient	2004 Logworth	2006 Coefficient	2006 Logworth	2009 Coefficient	2009 Logworth	2011 Coefficient	2011 Logworth	2015 Coefficient	2015 Logworth
Private dwelling	−0.004	**1.45**	−0.005	**2.32**	−0.003	1.05	−0.013	**9.87**	−0.007	**4.72**	−0.007	**4.90**
Social health insurance	−0.013	**6.11**	−0.007	**3.58**	−0.008	**15.94**	−0.002	1.26	−0.004	**2.18**	0.002	0.70
Ethnicity	−0.013	**14.85**	−0.005	**3.18**	−0.002	0.86	−0.001	0.47	−0.009	**8.94**	−0.008	**7.95**
Marital status	−0.001	0.34	−0.005	**3.21**	−0.007	**7.15**	−0.009	**10.97**	−0.013	**22.46**	−0.011	**14.57**
Primary school	−0.009	**35.44**	−0.006	**51.53**	−0.005	**64.32**	−0.006	**51.27**	−0.007	**55.83**	−0.006	**40.49**
Junior high school	−0.013		−0.011		−0.011		−0.011		−0.012		−0.010	
Senior high school	−0.021		−0.022		−0.023		−0.023		−0.020		−0.017	
University degree	−0.039		−0.039		−0.037		−0.035		−0.035		−0.030	
In employment	−0.015	**20.89**	−0.016	**33.38**	−0.013	**21.96**	−0.016	**32.83**	−0.014	**25.10**	−0.013	**24.27**
White collar worker	−0.010	**5.77**	−0.012	**9.67**	−0.012	**9.95**	−0.010	**6.15**	−0.009	**6.26**	−0.011	**8.65**
Farmer	0.025	**61.17**	0.016	**28.33**	0.012	**16.32**	0.014	**21.46**	0.014	**22.18**	0.007	**4.29**
Rural resident	0.011	**13.89**	0.016	**41.57**	0.012	**25.13**	0.011	**18.02**	0.018	**51.25**	0.014	**34.62**
East region	−0.002	**60.27**	−0.015	**43.46**	−0.009	**15.38**	−0.013	**28.68**	−0.005	**56.34**	0.004	**21.26**
Middle region	0.018		0.001		0.000		−0.001		0.013		0.010	
Number of observations	9886		8561		8444		8717		9595		9377	
Adjusted R-squared	0.2225		0.2466		0.2415		0.2070		0.2563		0.2388	
F statistic	90.96		90.04		89.27		73.15		106.32		94.58	

[1] For each wave, the first column shows the estimated marginal effects and the second the logworth values. Logworth values in bold indicate significance at or below 5% level.

Table 5. Effect of selected demographic and socioeconomic variables in the decomposition of the level-dependent indices for OOP burden.

	2000		2004		2006		2009		2011		2015	
Variable [1]	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth
Private dwelling	−0.033	**1.31**	−0.040	**1.69**	−0.014	0.25	−0.202	**4.71**	−0.124	**2.12**	−0.573	**1.59**
Social health insurance	−0.039	0.85	−0.016	0.41	−0.077	**7.48**	−0.014	0.18	−0.130	**2.69**	0.482	1.12
Ethnicity	−0.087	**7.50**	−0.046	**2.74**	−0.038	1.22	−0.073	**1.53**	−0.238	**7.60**	−0.681	**2.78**
Marital status	0.003	0.06	−0.026	1.13	−0.051	**1.87**	−0.180	**7.23**	−0.249	**9.43**	−0.596	**2.21**
Primary school	−0.039	**51.90**	−0.027	**39.00**	−0.038	**35.36**	−0.075	**17.14**	−0.071	**35.13**	0.172	**3.05**
Junior high school	−0.068		−0.066		−0.090		−0.162		−0.163		0.221	
Senior high school	−0.141		−0.146		−0.198		−0.331		−0.320		−0.093	
University degree	−0.498		−0.383		−0.448		−0.473		−0.822		−0.978	
In employment	−0.102	**9.85**	−0.131	**20.66**	−0.104	**6.80**	−0.269	**15.73**	−0.237	**8.66**	−0.349	1.15
White collar worker	−0.091	**4.83**	−0.182	**19.18**	−0.206	**12.69**	−0.304	**10.13**	−0.362	**10.69**	−0.581	**1.44**
Farmer	0.150	**22.52**	0.128	**18.00**	0.133	**10.08**	0.290	**16.26**	0.296	**11.38**	0.103	0.16
Rural resident	0.080	**7.74**	0.158	**35.70**	0.101	**8.73**	0.111	**3.88**	0.316	**19.02**	0.993	**8.49**
East region	−0.025	**18.55**	−0.157	**36.95**	−0.112	**8.09**	−0.151	**6.29**	0.031	**26.02**	0.939	**7.99**
Middle region	0.088		−0.006		−0.026		−0.021		0.318		0.869	
Number of observations	9886		8561		8444		8717		9595		9377	
Adjusted R-squared	0.1400		0.1987		0.1404		0.0964		0.1528		0.0401	
F statistic	51.76		69.49		45.79		29.88		55.63		12.60	

[1] For each wave, the first column shows the estimated marginal effects and the second the logworth values. Logworth values in bold indicate significance at or below 5% level.

Demographic and socio-economic factors such as education levels, employment status, occupation, residential regions, are among the most important determinants of income-related inequality in OOP burden across all years. Positive coefficients indicate that variables have a positive marginal effect on the observed inequality, i.e., tend to make income-related inequality of OOP burden less pro-poor, and the opposite holds for negative coefficients. For example, suffering from more major diseases tends to make income-related inequality of OOP burden more pro-poor, although the effect is often insignificant. The magnitudes of the coefficients in the decomposition of the level-dependent index in 2015 appear to be larger than the ones in other years. This might be due to the relatively high values of the OOP payments and household incomes in 2015. For both indices, factors that contribute to more pronounced pro-poor inequality in OOP burden seem to be higher education levels, being employed, having white-collar jobs and living in cities, given that the marginal effects of the associated variables are all negative. Social health insurance coverage appears to reduce the pro-poor inequality in OOP burden in some of the years, but in others the effects remain rather limited.

5. Discussion

The measurement and explanation of income-related inequality in health care attracts much policy interest, especially if reducing health inequality is high on the agenda. In this study, we track changes in the distribution of health care resources across income levels over a 15-year period, during which the government adopted a series of measures that gradually led to a profound reform of the health care system in China. We also explore the influence of various factors (e.g., demographic, socio-economic and health-related characteristics) on inequality.

Our study provides fresh evidence on the uneven access to health care and facilities. The results for the two types of indices we have used to estimate the extent of income-related inequality of health point in the same direction: high-income people tend to obtain more preventive care and use more hospital services, while low-income patients mostly seek care from village clinics/community health centres and folk doctors. The gap of preventive care between the rich and poor might be due to the fact that preventive care is only partially reimbursed by insurance and requires more cost-sharing. Although folk doctor care is not covered by insurance either, it is usually less costly and more accessible compared with hospital services, especially in remote rural areas. This utilization pattern is consistent with evidence from developed countries, with a pro-poor distribution in primary care use and pro-rich distribution in the use of specialised care (Van Doorslaer et al. 2000; Van Doorslaer et al. 2004; Van Doorslaer and Masseria 2004). The pro-poor distribution of the OOP burden is in line with the findings of utilization patterns, indicating that affordability remains a common barrier for the poor to access health care. Further decomposition analysis of the pro-poor inequality in OOP burden suggests that the inequality is largely driven by demographic and socio-economic factors. Higher levels of education, employment and occupation tend to be associated with a larger extent of pro-poor inequality of the OOP burden. In line with the findings of previous research, our results suggest that the recent expansion of social health insurance has a limited impact on the reduction of this inequality (Coté et al. 2013; Cai et al. 2017).

The study findings need to be interpreted in the light of the following limitations. First, we use equivalised household income to measure living standard, but in low- and middle-income countries income is not always a dependable indicator of a household's socioeconomic status, especially when day labour with volatile incomes and subsistence farming and fishing are common (Wagstaff 2009b; Wagstaff et al. 2003). However, in the context of China, income measures were regarded as more reliable than household expenditure since expenditure data might be distorted by the high saving rates of Chinese households (Sun et al. 2010; Yang 2013). Second, the nine provinces included in our analyses vary considerably in terms of demographics and economic development levels, so that comparing households' incomes in fairly rich and prosperous areas in the eastern region with those in worse-off and more rural areas without accounting for the differences in purchasing power might be problematic. Third, although the data have a longitudinal (panel) structure, this feature is not exploited in the empirical analysis since health care utilization was only reported by people who fell sick during the study period. Therefore, our results indicate the association between income-related inequality in the burden of OOP payments and various demographic and socio-economic characteristics and do not intend to obtain causal inference. Fourth, in the survey OOP payments are reported for the last four weeks only, and therefore there is a high risk of random high expenditure and random zero expenditure. Previous literature also pointed out that CHNS has a much lower OOP level on average compared to other household surveys in China because it might ignore the expenditure of people who were still in hospital at the time of the interview (Wagstaff and Lindelow 2008). However, with a longer reporting period (e.g., one year), OOP might also suffer from recall bias. As far as OOP payments are concerned, data on the previous month is the only source we could obtain. We need to bear in mind that we might underestimate or overestimate the OOP burden given the limitations stated above. Fifth, health care utilization behaviour is usually shaped by both financial (e.g., price elasticity, income levels, insurance coverage) and non-financial factors (e.g., health care need, quality of care, availability of transportation, health care personnel and infrastructure). In this paper, we focus on financial access to care, but evidence is lacking with respect to whether or not the health care reforms led to any change in non-financial access barriers and how this varied across different socioeconomic groups. Finally, primary care facilities and high-level specialised hospitals tend to serve different types of patients so that the observed inequality in facility use might also be related with differential levels of health care needs across income. It would be valuable to obtain more objective and reliable quality measurements for a rigorous assessment of the scale of inequality in health care use. The above issues could be the subject of future research through well-designed surveys and field studies conducted in more recent years.

6. Conclusions

Inequality in health care is a common challenge worldwide, especially in low- and middle-income countries that are looking for means of ensuring access to basic health care and protecting poor patients from health payment-induced impoverishment. Our findings have high relevance in the debate over the use of publicly sponsored health insurance programmes in tackling income dependence of health care use in China and other developing countries. An important policy lesson drawn from the study is that broad insurance coverage at population level does not necessarily lead to equal access to good quality health care. Our findings show there are still inequalities in the use of preventive care and hospital services across people from different income groups, indicating that the poor are faced with a heavy financial burden due to high insurance co-payments and insufficient coverage. Insufficient coverage of preventive care among the poor could lead to a disease-poverty trap as minor conditions would develop into severe illnesses that require specialist care from high-level hospitals and long-term use of medication (Xu et al. 2007). Early detection through screening or diagnostic tests could be a more cost-effective strategy compared with expensive acute care to tackle the challenges of the recent epidemiological transitions from infectious diseases to non-communicable diseases. Therefore, preventive care should be an integral part of a comprehensive insurance coverage to adjust for the socio-economic gradient in disease burdens (Yang 2013). To reduce the socio-economic gap in the access to health services, it is important to extend benefit packages to preventive care and hospital services. The expansion can be achieved incrementally as government subsidies and insurance premiums increase over the years, so that the means to extend the types of services covered by the insurance are compatible with the means to achieve equity. In addition, a well-functioning primary care system would provide more affordable and good-quality health care for patients from vulnerable socio-economic groups. Compared to investing most of the public resources in specialised hospitals, strengthening the delivery of basic needs-oriented primary care is a more viable way to benefit the majority of the patients.

In recent years, the Chinese government has attached greater importance to achieving a more balanced allocation of resources to primary care clinics and high-level hospitals by increasing funding for strengthening community health centres in cities and village clinics and township hospitals in rural areas. However, there is still a lack of well-trained personnel in many primary care facilities so that they cannot sufficiently meet the needs of the wide population (Mossialos et al. 2016). A comprehensive insurance coverage for health services combined with a strong primary care delivery system could help reduce disparities in health and health care across incomes. Even though Chinese policymakers have already started to address some of the issues identified above, stronger and more positive policy responses still need to be developed to close the socioeconomic gap in the access to health resources.

Author Contributions: Conceptualization, M.Y. and G.E.; methodology, M.Y. and G.E.; formal analysis, M.Y. and G.E.; data curation, M.Y.; writing—original draft preparation, M.Y.; writing—review and editing, G.E.; funding acquisition, M.Y. and G.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by PhD scholarship from the Lancaster University Management School.

Informed Consent Statement: The data collectors of the China Health and Nutrition Survey (CHNS) obtained informed consent from all subjects involved in the study.

Data Availability Statement: The data used by this study are publicly available at: http://www.cpc.unc.edu/projects/china. Data were accessed and downloaded in January 2020.

Acknowledgments: Financial support by PhD scholarship from Lancaster University Management School is gratefully acknowledged. This research uses data from China Health and Nutrition Survey (CHNS). We thank the National Institute for Nutrition and Health, China Centre for Disease Control and Prevention, Carolina Population Centre (P2C HD050924, T32 HD007168), the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, R24 HD050924, and R01-HD38700) and the NIH Fogarty International Centre (D43 TW009077, D43 TW007709) for financial support for the CHNS data collection and analysis files from 1989 to 2015 and future surveys, and the China-Japan Friendship Hospital, Ministry of Health for support for CHNS 2009, Chinese National Human Genome Centre at Shanghai since 2009, and Beijing Municipal Centre for Disease Prevention and Control since 2011.

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

In this appendix we provide the formulas of the indices we have calculated and of the dependent variables we have used in the regression-based decompositions. More details can be found in Kessels and Erreygers (2019).

We consider a population of n individuals, labelled by the subscript $i = 1, 2, \ldots, n$. Let h stand for health, y for income, and r for income rank (with the poorest person having rank 1, the second poorest rank 2, etc.). Since all health outcome variables of the paper are bounded between 0 and 1, we have used the bounded-variable versions of the indices. The rank-dependent index is equal to:

$$R = \frac{4}{n}\sum_{i=1}^{n}\left[\frac{(2r_i - 1)}{n} - 1\right]h_i \tag{A1}$$

while the level-dependent index is equal to:

$$L = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{y_i}{\mu(y)} - 1\right]h_i \tag{A2}$$

where $\mu(y)$ represents mean income.

The decomposition analysis in the paper is applied to OOP burden, which is an ill-health variable. Let z be this variable. The dependent variable d^R of the decomposition regression for the rank-dependent index is then defined as:

$$d_i^R = 4\left[1 - \mu(z) - \frac{(2r_i - 1)}{n}(1 - z_i)\right] \tag{A3}$$

where $\mu(z)$ is the mean of the ill-health variable. Likewise, the dependent variable d^L of the decomposition regression for the level-dependent index is equal to:

$$d_i^L = 1 - \mu(z) - \frac{y_i}{\mu(y)}(1 - z_i) \tag{A4}$$

Appendix B

Table A1. Effect of selected demographic and socioeconomic variables in the decomposition of the rank-dependent indices for OOP burden (full results).

	2000		**2004**		**2006**		**2009**		**2011**		**2015**	
Variable [1]	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth
Female × age 18 below	−0.006		−0.028		0.000		−0.006		−0.016		0.055	
Female × age 18–24	−0.001		0.003		0.007		0.000		0.005		−0.006	
Female × age 25–34	0.004		0.002		0.010		0.000		0.004		0.000	
Female × age 35–44	0.003		0.003		0.006		−0.003		0.005		0.006	
Female × age 45–54	−0.005		−0.004		0.003		−0.005		0.001		0.005	
Female × age 55–64	0.000		−0.008		0.003		−0.010		−0.005		0.006	
Female × age 65 above	0.012	**22.03**	−0.001	**10.84**	0.003	**3.93**	−0.009	**3.81**	0.000	**5.91**	0.008	**8.74**
Male × age 18–24	0.000		0.000		0.000		0.000		0.000		0.000	
Male × age 25–34	0.007		0.008		0.009		0.000		−0.001		0.006	
Male × age 35–44	0.015		0.008		0.010		0.000		0.006		0.007	
Male × age 45–54	0.003		0.004		0.009		0.000		0.007		0.009	
Male × age 55–64	0.006		0.000		0.008		−0.001		0.003		0.012	
Male × age 65 above	0.014		0.003		0.007		−0.004		0.004		0.012	
Number of major diseases	−0.002	0.52	0.000	0.15	−0.001	0.60	−0.001	0.83	0.000	0.22	−0.001	0.49
Sickness in the last month	0.013	**3.72**	0.010	**7.74**	0.013	**12.72**	0.013	**13.26**	0.015	**14.12**	0.012	**11.88**
Number of symptoms in the last month	0.011	**5.47**	0.006	**9.91**	0.003	**2.38**	0.002	**1.59**	0.002	1.10	0.003	**2.69**
Private dwelling	−0.004	**1.45**	−0.005	**2.32**	−0.003	1.05	−0.013	**9.87**	−0.007	**4.72**	−0.007	**4.90**
Social health insurance	−0.013	**6.11**	−0.007	**3.58**	−0.008	**15.94**	−0.002	1.26	−0.004	**2.18**	0.002	0.70
Ethnicity	−0.013	**14.85**	−0.005	**3.18**	−0.002	0.86	−0.001	0.47	−0.009	**8.94**	−0.008	**7.95**
Marital status	−0.001	0.34	−0.005	**3.21**	−0.007	**7.15**	−0.009	**10.97**	−0.013	**22.46**	−0.011	**14.57**
Primary school	−0.009		−0.006		−0.005		−0.006		−0.007		−0.006	
Junior high school	−0.013	**35.44**	−0.011	**51.53**	−0.011	**64.32**	−0.011	**51.27**	−0.012	**55.83**	−0.010	**40.49**
Senior high school or above	−0.021		−0.022		−0.023		−0.023		−0.020		−0.017	
University degree	−0.039		−0.039		−0.037		−0.035		−0.035		−0.030	
In employment	−0.015	**20.89**	−0.016	**33.38**	−0.013	**21.96**	−0.016	**32.83**	−0.014	**25.10**	−0.013	**24.27**

Table A1. *Cont.*

Variable [1]	2000 Coefficient	2000 Logworth	2004 Coefficient	2004 Logworth	2006 Coefficient	2006 Logworth	2009 Coefficient	2009 Logworth	2011 Coefficient	2011 Logworth	2015 Coefficient	2015 Logworth
White collar worker	−0.010	**5.77**	−0.012	**9.67**	−0.012	**9.95**	−0.010	**6.15**	−0.009	**6.26**	−0.011	**8.65**
Farmer	0.025	**61.17**	0.016	**28.33**	0.012	**16.32**	0.014	**21.46**	0.014	**22.18**	0.007	**4.29**
Rural resident	0.011	**13.89**	0.016	**41.57**	0.012	**25.13**	0.011	**18.02**	0.018	**51.25**	0.014	**34.62**
East region Middle region	−0.002 0.018	**60.27**	−0.015 0.001	**43.46**	−0.009 0.000	**15.38**	−0.013 −0.001	**28.68**	−0.005 0.013	**56.34**	0.004 0.010	**21.26**
Number of children aged 0–4 in the household	0.001	0.82	−0.001	**2.19**	0.001	**1.82**	0.000	0.05	0.001	**1.73**	0.002	**5.56**
Number of children aged 5–14 in the household	0.002	**8.47**	0.001	**1.25**	0.000	0.86	0.000	0.59	0.001	**2.68**	0.003	**32.08**
Number of observations Adjusted R-squared F statistic	9886 0.2225 90.96		8561 0.2466 90.04		8444 0.2415 89.27		8717 0.2070 73.15		9595 0.2563 106.32		9377 0.2388 94.58	

[1] For each wave, the first column shows the estimated marginal effects and the second the logworth values. Logworth values in bold indicate significance at or below 5% level.

Table A2. Effect of selected demographic and socioeconomic variables in the decomposition of the level-dependent indices for OOP burden (full results).

	2000		2004		2006		2009		2011		2015	
Variable [1]	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth
Female × age 18 below	−0.049		−0.096		0.000		0.011		0.126		0.910	
Female × age 18–24	0.015		0.002		0.029		0.096		0.352		−0.977	
Female × age 25–34	−0.012		−0.034		0.030		0.104		0.193		−0.782	
Female × age 35–44	0.054		0.003		−0.031		−0.014		0.166		−0.452	
Female × age 45–54	−0.014		−0.061		−0.056		−0.052		0.140		0.120	
Female × age 55–64	0.002		−0.085		−0.040		−0.127		0.035		0.166	
Female × age 65 above	0.064	**9.44**	−0.030	**8.20**	−0.023	**1.68**	−0.128	**2.31**	0.178	**2.92**	0.574	**2.84**
Male × age 18–24	0.000		0.000		0.000		0.000		0.000		0.000	
Male × age 25–34	−0.002		0.016		−0.004		0.075		0.094		−0.630	
Male × age 35–44	0.130		0.066		0.024		0.018		0.202		−0.578	
Male × age 45–54	0.043		0.021		0.027		0.046		0.246		0.038	
Male × age 55–64	0.037		−0.017		0.010		0.010		0.210		0.540	
Male × age 65 above	0.053		0.014		0.038		−0.089		0.244		0.520	
Number of major diseases	0.010	0.29	0.011	0.45	−0.017	0.57	−0.019	0.39	−0.031	0.68	0.007	0.02
Sickness in the last month	0.009	0.10	0.031	1.06	0.070	**2.17**	0.107	**1.87**	0.149	**2.18**	0.309	0.63
Number of symptoms in the last month	0.064	**2.18**	0.037	**3.90**	0.021	0.82	0.026	0.52	0.045	0.73	0.050	0.13
Private dwelling	−0.033	**1.31**	−0.040	**1.69**	−0.014	0.25	−0.202	**4.71**	−0.124	**2.12**	−0.573	**1.59**
Social health insurance	−0.039	0.85	−0.016	0.41	−0.077	**7.48**	−0.014	0.18	−0.130	**2.69**	0.482	1.12
Ethnicity	−0.087	**7.50**	−0.046	**2.74**	−0.038	1.22	−0.073	**1.53**	−0.238	**7.60**	−0.681	**2.78**
Marital status	0.003	0.06	−0.026	1.13	−0.051	**1.87**	−0.180	**7.23**	−0.249	**9.43**	−0.596	**2.21**
Primary school	−0.039		−0.027		−0.038		−0.075		−0.071		0.172	
Junior high school	−0.068	**51.90**	−0.066	**39.00**	−0.090	**35.36**	−0.162	**17.14**	−0.163	**35.13**	0.221	**3.05**
Senior high school	−0.141		−0.146		−0.198		−0.331		−0.320		−0.093	
University degree	−0.498		−0.383		−0.448		−0.473		−0.822		−0.978	
In employment	−0.102	**9.85**	−0.131	**20.66**	−0.104	**6.80**	−0.269	**15.73**	−0.237	**8.66**	−0.349	1.15
White collar worker	−0.091	**4.83**	−0.182	**19.18**	−0.206	**12.69**	−0.304	**10.13**	−0.362	**10.69**	−0.581	**1.44**
Farmer	0.150	**22.52**	0.128	**18.00**	0.133	**10.08**	0.290	**16.26**	0.296	**11.38**	0.103	0.16

Table A2. *Cont.*

	2000		2004		2006		2009		2011		2015	
Variable [1]	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth	Coefficient	Logworth
Rural resident	0.080	**7.74**	0.158	**35.70**	0.101	**8.73**	0.111	**3.88**	0.316	**19.02**	0.993	**8.49**
East region	−0.025	**18.55**	−0.157	**36.95**	−0.112	**8.09**	−0.151	**6.29**	0.031	**26.02**	0.939	**7.99**
Middle region	0.088		−0.006		−0.026		−0.021		0.318		0.869	
Number of children aged 0–4 in the household	−0.007	0.66	−0.015	**2.71**	0.003	0.15	−0.034	**2.60**	0.019	0.75	0.233	**2.69**
Number of children aged 5–14 in the household	0.028	**16.80**	0.005	**1.19**	0.008	1.41	0.009	0.80	0.015	**1.12**	0.171	**3.99**
Number of observations	9886		8561		8444		8717		9595		9377	
Adjusted R-squared	0.1400		0.1987		0.1404		0.0964		0.1528		0.0401	
F statistic	51.76		69.49		45.79		29.88		55.63		12.60	

[1] For each wave, the first column shows the estimated marginal effects and the second the logworth values. Logworth values in bold indicate significance at or below 5% level.

References

Babiarz, Kimberly Singer, Grant Miller, Hongmei Yi, Linxiu Zhang, and Scott Rozelle. 2010. New evidence on the impact of China's New Rural Cooperative Medical Scheme and its implications for rural primary healthcare: Multivariate difference-in-difference analysis. *BMJ* 341: c5617. [CrossRef] [PubMed]

Cai, Jiaoli, Peter C. Coyte, and Hongzhong Zhao. 2017. Decomposing the causes of socioeconomic-related health inequality among urban and rural populations in China: A new decomposition approach. *International Journal for Equity in Health* 16: 128. [CrossRef]

Chen, Mingsheng, Guixia Fang, Lidan Wang, Zhonghua Wang, Yuxin Zhao, and Lei Si. 2015. Who Benefits from Government Healthcare Subsidies? An Assessment of the Equity of Healthcare Benefits Distribution in China. *PLoS ONE* 10: e0119840. [CrossRef] [PubMed]

Chen, Zhi, Kai Gu, Ying Zheng, Wei Zheng, Wei Lu, and Xiao Ou Shu. 2008. The use of complementary and alternative medicine among Chinese women with breast cancer. *The Journal of Alternative and Complementary Medicine* 14: 1049–55. [CrossRef]

Coté, Pierre, Xiaoqing Yang, Vicki Kristman, Sheilah Hogg-Johnson, Dwayne Van Eerd, Mana Rezai, and Marjan Vidmar. 2013. The association between workers' compensation claims involving neck pain and future health care utilization: A population-based cohort study. *Journal of Occupational Rehabilitation* 23: 547–56. [CrossRef] [PubMed]

Coveney, Max, Pilar García-Gómez, Eddy Van Doorslaer, and Tom Van Ourti. 2016. Health Disparities by Income in Spain Before and After the Economic Crisis. *Health Economics* 25: 141–58. [CrossRef]

Cui, Yong, Xiao-Ou Shu, Yutang Gao, Wanqing Wen, Zhi-Xian Ruan, Fan Jin, and Wei Zheng. 2004. Use of complementary and alternative medicine by chinese women with breast cancer. *Breast Cancer Research and Treatment* 85: 263–70. [CrossRef]

Dong, Keyong. 2009. Medical insurance system evolution in China. *China Economic Review* 20: 591–97. [CrossRef]

Elwell-Sutton, Timothy M., Chao Qiang Jiang, Wei Sen Zhang, Kar Keung Cheng, Tai H. Lam, Gabriel M. Leung, and Catherine Mary Schooling. 2013. Inequality and inequity in access to health care and treatment for chronic conditions in China: The Guangzhou Biobank Cohort Study. *Health Policy and Planning* 28: 467–79. [CrossRef]

Erreygers, Guido. 2009. Correcting the concentration index. *Journal of Health Economics* 28: 504–15. [CrossRef]

Erreygers, Guido, and Roselinde Kessels. 2013. Regression-based decompositions of rank-dependent indicators of socioeconomic inequality of health. *Health and Inequality* 21: 227–59.

Erreygers, Guido, and Roselinde Kessels. 2017. Socioeconomic Status and Health: A New Approach to the Measurement of Bivariate Inequality. *International Journal of Environmental Research and Public Health* 14: 673. [CrossRef] [PubMed]

Erreygers, Guido, and Tom Van Ourti. 2011. Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: A recipe for good practice. *Journal of Health Economics* 30: 685–94. [CrossRef] [PubMed]

Gao, Jun, Juncheng Qian, Shenglan Tang, Bo Eriksson, and Erik Blas. 2002. Health equity in transition from planned to market economy in China. *Health Policy and Planning* 17: 20–29. [CrossRef]

Gong, Sen, and Hana Brixi. 2005. *China: Health, Poverty and Economic Development*. Beijing: Office of the World Health Organization Representative in China and Social Development Department of China State Council Development Research Centre.

Harmsworth, K., and George T. Lewith. 2001. Attitudes to traditional Chinese medicine amongst Western trained doctors in the People's Republic of China. *Social Science & Medicine* 52: 149–53. [CrossRef]

Heckley, Gawain, Ulf-G. Gerdtham, and Gustav Kjellsson. 2016. A general method for decomposing the causes of socioeconomic inequality in health. *Journal of Health Economics* 48: 89–106. [CrossRef]

Kessels, Roselinde, and Guido Erreygers. 2019. A direct regression approach to decomposing socioeconomic inequality of health. *Health Economics* 28: 884–905. [CrossRef]

Li, Chao, Lingxia Zeng, Michael John Dibley, Duolao Wang, Leilei Pei, and Hong Yan. 2015. Evaluation of socio-economic inequalities in the use of maternal health services in rural western China. *Public Health* 129: 1251–57. [CrossRef]

Li, Xi, Jiapeng Lu, Shuang Hu, K. K. Cheng, Jan De Maeseneer, Qingyue Meng, Elias Mossialos, Dong Roman Xu, Winnie Yip, Hongzhao Zhang, and et al. 2017a. The primary health-care system in China. *The Lancet* 390: 2584–94. [CrossRef]

Li, Yanping, Vasanti Malik, and Frank B. Hu. 2017b. Health Insurance in China: After Declining In The 1990s, Coverage Rates Rebounded To Near-Universal Levels By 2011. *Health Affairs* 36: 1452–60. [CrossRef]

Liu, Yuanli, William C. Hsiao, and Karen Eggleston. 1999. Equity in health and health care: The Chinese experience. *Social Science & Medicine* 49: 1349–56.

Mossialos, Elias, Yanfeng Ge, Jia Hu, and Liejun Wang. 2016. *Pharmaceutical Policy in China Challenges and Opportunities for Reform*. Geneva: World Health Organization.

O'Donnell, Owen, and Carol Propper. 1991. Equity and the Distribution of United Kingdom National Health Service Resources. *Journal of Health Economics* 10: 1–19. [CrossRef] [PubMed]

OECD. n.d. *What Are Equivalence Scales?* Edited by OECD Project on Income Distribution and Poverty. Paris: OECD.

Popkin, Barry M., Shufa Du, Fengying Zhai, and Bing Zhang. 2010. Cohort Profile: The China Health and Nutrition Survey—Monitoring and understanding socio-economic and health change in China, 1989–2011. *International Journal of Epidemiology* 39: 1435–40. [CrossRef] [PubMed]

Shen, Yuan, Hong Yan, Klemetti Reija, Qiang Li, Shengbin Xiao, Jianmin Gao, and Zhongliang Zhou. 2014. Equity in use of maternal health services in Western Rural China: A survey from Shaanxi province. *BMC Health Services Research* 14: 155. [CrossRef]

Si, Wei. 2021. Public health insurance and the labor market: Evidence from China's Urban Resident Basic Medical Insurance. *Health Economics* 30: 403–31. [CrossRef]

Sun, Xiaoyun, Adrian C. Sleigh, Gordon A. Carmichael, and Sukhan Jackson. 2010. Health payment-induced poverty under China's New Cooperative Medical Scheme in rural Shandong. *Health Policy and Planning* 25: 419–26. [CrossRef]

Sylvia, Sean, Hao Xue, Chengchao Zhou, Yaojiang Shi, Hongmei Yi, Huan Zhou, Scott Rozelle, Madhukar Pai, and Jishnu Das. 2017. Tuberculosis detection and the challenges of integrated care in rural China: A cross-sectional standardized patient study. *PLoS Medicine* 14: e1002405. [CrossRef]

Sylvia, Sean, Yaojiang Shi, Hao Xue, Xin Tian, Huan Wang, Qingmei Liu, Alexis Medina, and Scott Rozelle. 2014. Survey using incognito standardized patients shows poor quality care in China's rural clinics. *Health Policy and Planning* 30: 322–33. [CrossRef]

Tang, Shenglan, Qingyue Meng, Lincoln Chen, Henk Bekedam, Tim Evans, and Margaret Whitehead. 2008. Tackling the challenges to health equity in China. *The Lancet* 372: 1493–501. [CrossRef]

Van Doorslaer, Eddy, Adam Wagstaff, and Frans Rutten. 1992. *Equity in the Finance and Delivery of Health Care: An International Perspective*. Oxford: Oxford University Press.

Van Doorslaer, Eddy, Adam Wagstaff, Hattem Van der Burg, Terkel Christiansen, Diana De Graeve, Inge Duchesne, Ulf-G. Gerdtham, Michael Gerfin, José Geurts, and Lorna Gross. 2000. Equity in the delivery of health care in Europe and the US. *Journal of Health Economics* 19: 553–83. [CrossRef]

Van Doorslaer, Eddy, and Cristina Masseria. 2004. *Income-Related Inequality in the Use of Medical Care in 21 OECD Countries*. Paris: OECD.

Van Doorslaer, Eddy, and Owen O'Donnell. 2011. Measurement and Explanation of Inequality in Health and Health Care in Low-Income Settings. In *Health Inequality and Development*. Edited by Marc McGillivray, Indranil Dutta and David Lawson. London: Palgrave Macmillan.

Van Doorslaer, Eddy, Xander Koolman, and Andrew M. Jones. 2004. Explaining income-related inequalities in doctor utilisation in Europe. *Health Economics* 13: 629–47. [CrossRef] [PubMed]

Wagstaff, Adam. 2009a. *Reforming China's Rural Health System, Directions in Development—Human Development*. Washington, DC: The World Bank.

Wagstaff, Adam. 2009b. Reranking and Pro-Poor Growth: Decompositions for China and Vietnam. *The Journal of Development Studies* 45: 1403–25. [CrossRef]

Wagstaff, Adam, and Magnus Lindelow. 2008. Can insurance increase financial risk?: The curious case of health insurance in China. *Journal of Health Economics* 27: 990–1005. [CrossRef]

Wagstaff, Adam, Eddy Van Doorslaer, and Naoko Watanabe. 2003. On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. *Journal of Econometrics* 112: 207–23. [CrossRef]

Wagstaff, Adam, Pierella Paci, and Eddy Van Doorslaer. 1991. On the measurement of inequalities in health. *Social Science & Medicine* 33: 545–57. [CrossRef]

Wagstaff, Adam, Winnie Yip, Magnus Lindelow, and William C. Hsiao. 2009. China's health system and its reform: A review of recent studies. *Health Economics* 18: S7–S23. [CrossRef]

Wang, Yang, Jian Wang, Elizabeth Maitland, Yaohui Zhao, Stephen Nicholas, and Mingshan Lu. 2012. Growing old before growing rich: Inequality in health service utilization among the mid-aged and elderly in Gansu and Zhejiang Provinces, China. *BMC Health Services Research* 12: 302. [CrossRef]

Wong, William C. W., Shanzhu Zhu, Jason J. Ong, Minghui Peng, Cindy L. K. Lam, Michael R. Kidd, Martin Roland, and Sunfang Jiang. 2017. Primary care workforce and continuous medical education in China: Lessons to learn from a nationwide cross-sectional survey. *BMJ Open* 7: e015145. [CrossRef]

Xiang, Yuening, Zimu Guo, Pengfei Zhu, Jia Chen, and Yongye Huang. 2019. Traditional Chinese medicine as a cancer treatment: Modern perspectives of ancient but advanced science. *Cancer Medicine* 8: 1958–75. [CrossRef]

Xie, E. 2011. Income-related inequalities of health and health care utilization. *Frontiers of Economics in China* 6: 131–56. [CrossRef]

Xie, Xin, Qunhong Wu, Yanhua Hao, Hui Yin, Wenqi Fu, Ning Ning, Ling Xu, Chaojie Liu, Ye Li, Zheng Kang, and et al. 2014. Identifying Determinants of Socioeconomic Inequality in Health Service Utilization among Patients with Chronic Non-Communicable Diseases in China. *PLoS ONE* 9: e100231. [CrossRef] [PubMed]

Xu, Ke, David B. Evans, Guido Carrin, Ana Mylena Aguilar-Rivera, Philip Musgrove, and Timothy Evans. 2007. Protecting Households From Catastrophic Health Spending. *Health Affairs* 26: 972–83. [CrossRef] [PubMed]

Yang, Wei. 2013. China's new cooperative medical scheme and equity in access to health care: Evidence from a longitudinal household survey. *International Journal for Equity in Health* 12: 20. [CrossRef] [PubMed]

Yang, Wei, and Xun Wu. 2017. Providing comprehensive health insurance coverage in rural China: A critical appraisal of the New Cooperative Medical Scheme and ways forward. *Global Policy* 8: 110–16. [CrossRef]

Yi, Hongmei, Paiou Wu, Xiaoyuan Zhang, Dirk E. Teuwen, and Sean Sylvia. 2020. Market competition and demand for skills in a credence goods market: Evidence from face-to-face and web-based non-physician clinician training in rural China. *PLoS ONE* 15: e0233955. [CrossRef]

Yip, Winnie Chi-Man, William C. Hsiao, Wen Chen, Shanlian Hu, Jin Ma, and Alan Maynard. 2012. Early appraisal of China's huge and complex health-care reforms. *The Lancet* 379: 833–42. [CrossRef] [PubMed]

Zhang, Xiaobo, and Ravi Kanbur. 2005. Spatial inequality in education and health care in China. *China Economic Review* 16: 189–204. [CrossRef]

Zhang, Xin, Qunhong Wu, Yongxiang Shao, Wenqi Fu, Guoxiang Liu, and Peter C. Coyte. 2015. Socioeconomic Inequities in Health Care Utilization in China. *Asia-Pacific Journal of Public Health* 27: 429–38. [CrossRef]

Zhou, Zhongliang, Jianmin Gao, Ashley Fox, Keqin Rao, Ke Xu, Ling Xu, and Yaoguang Zhang. 2011. Measuring the equity of inpatient utilization in Chinese rural areas. *BMC Health Services Research* 11: 201. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Economies Editorial Office
E-mail: economies@mdpi.com
www.mdpi.com/journal/economies

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34

www.mdpi.com

ISBN 978-3-0365-6893-5