Risk Measures with Applications in Finance and Economics

Edited by
Michael McAleer and Wing-Keung Wong
Printed Edition of the Special Issue Published in *Sustainability*

MDPI

Risk Measures with Applications in Finance and Economics

Risk Measures with Applications in Finance and Economics

Special Issue Editors

Michael McAleer
Wing-Keung Wong

MDPI

Special Issue Editors

Michael McAleer

Wing-Keung Wong

Asia University

Asia University

Taiwan

Taiwan

This is a reprint of articles from the Special Issue published online in the open access journal *Sustainability* (ISSN 2071-1050) from 2017 to 2018 (available at: https://www.mdpi.com/journal/sustainability/special_issues/Risk_finance_sustainability)

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**, *Article Number*, Page Range.

Contents

About the Special Issue Editors

Michael McAleer has an extensive resume beginning with a Ph.D. in Economics (1981), from Queen's University, Canada. He has filled the following positions at the following institutions: Research Chair and Professor in the Department of Finance, Asia University, Taiwan; Erasmus Visiting Professor of Quantitative Finance at the Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Netherlands; Adjunct Professor in the Department of Economic Analysis and ICAE, Complutense University of Madrid, Spain; Adjunct Professor in the Department of Mathematics and Statistics, University of Canterbury, New Zealand; and IAS Adjunct Professor at the Institute of Advanced Sciences, Yokohama National University, Japan. On numerous occasions, he has been a Distinguished Visiting Professor at: the University of Tokyo, Kyoto University, and Osaka University in Japan; the University of Padova in Italy; Complutense University of Madrid in Spain; Ca' Foscari University of Venice in Italy; the University of Zurich in Switzerland; and the University of Hong Kong, the Chinese University of Hong Kong, and Hong Kong University of Science and Technology in Hong Kong. He was elected a Distinguished Fellow of the International Engineering and Technology Institute (DFIETI), as well as a Fellow of: the Academy of the Social Sciences in Australia (FASSA), the International Environmental Modelling and Software Society (FIEMSS), the Modelling and Simulation Society of Australia and New Zealand (FMSSANZ), the Tinbergen Institute in the Netherlands, the *Journal of Econometrics, and Econometric Reviews*. He is the Editor-in-Chief of six international journals, is on the editorial boards of a further 40+ international journals, and has guest co-edited numerous special issues of the following journals: *Journal of Econometrics (Elsevier), Econometric Reviews (Taylor and Francis), Environmental Modelling and Software (Elsevier), Mathematics and Computers in Simulation (Elsevier), North American Journal of Economics and Finance (Elsevier), International Review of Economics and Finance (Elsevier), Annals of Financial Economics (World Scientific), Journal of Risk and Financial Management (MDPI), Sustainability (MDPI), Energies (MDPI), Risks (MDPI), Journal of Economic Surveys (Wiley), Economic Record (Wiley), and China Finance Review International (Emerald)*. In terms of academic publications, he has published 750+ journal articles and books in the fields of economics, theoretical and applied financial econometrics, quantitative finance, risk and financial management, theoretical and applied econometrics, theoretical and applied statistics, time series analysis, energy economics and finance, sustainability, carbon emissions, climate change econometrics, forecasting, informatics, data mining and bibliometrics.

Wing-Keung Wong obtained his Ph.D. from the University of Wisconsin in Madison, USA. He is a Chair Professor in the Department of Finance at Asia University and made an appearance in "Who's Who in the World." He is ranked in the top 1% of economists by the Social Science Research Network and is on the lists of the top Taiwanese, Asian, and world economists according to RePEc. His research areas include financial economics, econometrics, investment theory, risk management, and operational research. He has published 200+ papers in numerous international journals. Wing-Keung's time serving international academies, governments, societies, and universities has garnered him invitations to provide consults to government departments and corporations and to give lectures and seminars as well as present papers at several universities and institutions. He has been editor, advisor, and associate editor of international journals, has solely or jointly supervised several overseas graduate students and has been appointed external reviewer and external examiner by numerous universities.

Preface to "Risk Measures with Applications in Finance and Economics"

Risk Measures play a vital role in many fields in Economics and Finance. Using different risk measures could compare the performances of different variables through the analysis of empirical real-world data. For example, risk measures could help to form effective monetary and fiscal policies, and to develop pricing models for financial assets, such as equities, bonds, currencies, and derivative securities.

A Special Issue of "Risk Measures with Applications in Finance and Economics" will be devoted to advancements in the mathematical and statistical development of risk measures with applications in Finance and Economics. This Special Issue will bring together theory, practice and applications of risk measures.

Michael McAleer, Wing-Keung Wong
Special Issue Editors

Article

The Effects of Health Status on Life Insurance Holdings in 16 European Countries

Saruultuya Tsendsuren [1], Chu-Shiu Li [2,*], Sheng-Chang Peng [3] and Wing-Keung Wong [4,5,6,7]

[1] Business Development Division, Golomt Bank of Mongolia, Ulaanbaatar 15160, Mongolia;
    saka_can@yahoo.com
[2] Department of Risk Management and Insurance, College of Finance and Banking, National Kaohsiung
    University of Science and Technology, Kaohsiung 824, Taiwan
[3] Department of Risk Management and Insurance, School of Management, Ming Chuan University,
    Taipei 111, Taiwan; scpeng@mail.mcu.edu.tw
[4] Department of Finance, College of Management, Fintech Center, and Big Data Research Center,
    Asia University, Taichung 41354, Taiwan; wong@asia.edu.tw
[5] Department of Medical Research, China Medical University, Taichung 404, Taiwan
[6] Department of Economics and Finance, Hang Seng Management College, Hong Kong 999077, China
[7] Department of Economics, Lingnan University, Tuen Mun, Hong Kong 999077, China
* Correspondence: chushiu.li@gmail.com; Tel.: +886-7-601-1000 (ext. 33017)

Abstract: This study examines the relationships among three health status indicators (self-perceived health status, objective health status, and future health risk) and life insurance holdings in 16 European countries. Our results show that households with poor self-perceived health status and high future health risk are less likely to purchase life insurance in the entire sample as well as in the subsample for countries with a national health system (NHS). In non-NHS countries, those households that have high future health risk are less inclined to purchase life insurance. In terms of preferences for types of life insurance policies (term life, whole life, both, or none) in the whole sample, poor self-perceived health status and high future health risk are less inclined to hold only term life insurance policy. In addition, poor self-perceived health status and high future health risk have a negative impact on holdings of both types of life insurance. Our findings reveal that there is no adverse selection problem in the life insurance market, especially in European countries with NHS.

1. Introduction

Life insurance has a special standing among households, used to hedge against the loss of income resulting from an unexpected death [1] Life insurance often helps to carry out family responsibilities such as educating children, paying off mortgage or other debt, and providing revenue for survivors [2].

From prior studies on the relationship between health status indicators and medical insurance purchases, poor health status is negatively associated with the purchase of medical insurance in the US [3] and Europe [4]. Buchmueller et al. [5] observe that those with private health insurance have lower hospital utilization than those without private health insurance in Australia. In China, rural residents enrolled in The New Cooperative Medical Scheme have higher probability of shifting from working for others to being self-employed and from being temporarily employed to being self-employed [6].

The main purpose of this study is to explore the effects of three health status indicators, self-perceived health status (SPH), objective health status (OHS), and future health risk (FHR) on life insurance holdings in 16 European countries (The detailed definitions of SPH, OHS, and FHR are included in the Section 3). We also investigate the impact of these three health status indicators on the decision to purchase different types of life insurance (term, whole, or both types of life insurance policies). The data used in this study is from the Survey of Health, Aging, and Retirement in Europe (SHARE). Prior literature reveals that different national health systems (NHS) offer differing degrees of risk protection [7,8]. Therefore, we examine whether NHS impacts on the relationship between health status and life insurance holdings.

The important contributions of this paper are as follows: First, to the best of our knowledge, this is the first paper to examine the effects of three different health status indicators on the demand for life insurance in European countries. Second, we use SHARE household data from 16 European countries to compare the results of other determinants on life insurance ownership and the types of life insurance, as well as previous studies based on data from only one country. The use of SHARE data represents significant improvements over previous studies based on data from individual countries. Third, we examine and compare the responses of households in NHS and non-NHS countries to explore the effect of NHS on life insurance holdings. Finally, our empirical results may provide policy implications for insurers in European countries in that the marketing strategies for life insurance should consider not only demographic factors, but also household health status and national health insurance coverage.

Our findings clearly support our hypotheses that SPH and high FHR are negatively associated with the decision to hold life insurance in the pooled data and in the subsample of NHS countries (In our regression models, when we consider these three health status indicators one by one, each has negative correlation with life insurance holdings. However, when we consider the three health status indicators together, the coefficient of OHS becomes insignificantly different from zero). However, among households in non-NHS countries, only FHR has a negative effect on life insurance purchase. Moreover, elderly households with high FHR have high probability to hold life insurance in the whole sample, as well as in the subsample of non-NHS countries.

There are some interesting results in terms of the demand for different types of life insurance (term life only, whole life only, or both types) in the whole sample. The estimated marginal effects reveal that all three health status indicators are negatively related to holding only term life policies (There are similar regression results for life insurance holdings. When we consider the three health status indicators together, there is no effect of OHS on the holding of term life insurance only). Households with poor SPH or high FHR are less likely to own both types. However, no health status indicator is related to households with whole life only. Our empirical evidence may provide policy implications for insurers in European countries. For example, marketing strategies should consider not only demographic factors but also household health status indicators and NHS. Finally, our empirical evidence reveals that there is no existing adverse selection problem in life insurance markets especially among NHS countries in Europe.

The rest of this paper is organized as follows: Section 2 provides a review of existing literature and hypothesis development. Section 3 includes a discussion of the research methods. In Section 4, we present the empirical results. Finally, Section 5 is the conclusion.

2. Literature Review and Hypothesis Development

This section begins with a brief review of the literature followed by the hypotheses tested in this study. One stream of the literature on life insurance demand focuses on aggregated country analysis and concludes that income per capita, young dependency ratio, social security system, interest rate, and inflation are the main factors that affect the demand for life insurance in different countries [9–12]. Another stream of the literature uses household or individual data for one specific country to determine the demographic factors (such as age, education, marital status, numbers of children) and economic factors (such as income and net wealth) that are associated with the decision

to hold life insurance [13,14]. However, very few papers examine the association between health status and the holding of life insurance. Fang and Kung [15] use eight health conditions to define individual health status, including high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychological disorder, and arthritis. They demonstrate that healthy individuals are more likely to purchase life insurance than unhealthy individuals in the US.

2.1. Health Status Indicators and Life Insurance Holding Behaviour

The concept of health encompasses more than the absence of disease. It includes social, psychological, and economic well-being [16]. Good health indicates satisfaction with life and general acceptance, while poor health refers to a low quality of life or dissatisfaction with life. Furthermore, economic or social factors are the main determinants of good health [16]. Being married and effective health care have the strongest impact on people's positive perceptions of health [17].

The subjective measure of health status is SPH, which refers to a single-item health measure in which individuals rate the current status of their own health on a five-point scale from excellent (or very good) to very poor. Some indicators provide direct evidence of the health status of individuals, including previous and current diseases (diagnosed by physicians), collectively termed OHS.

It is well known that elderly perceiving their health in positive terms tend to overestimate their health, while others tend to report poorer health than those with similar OHS [18]. Thus, the relationship between SPH and OHS is complex. Individuals with poor SPH and high FHR should anticipate higher out-of-pocket health expenditures than similar individuals with low FHR. Individuals are generally unable to dynamically insure against FHR and medical expenditure risk [19].

Some empirical studies identify health risks as an important factor in precautionary participation in the financial market [7,8,20–24]. With respect to health status, most of the previous literature, except Atella et al. [7], considers the effects of current health status on portfolio decision, without investigating the roles of FHR and OHS. The elderly are less likely to increase income risks when they face a much higher health risk. In other words, when health risks cannot be easily avoided, investors may tend to underestimate their exposure to avoidable risks and financial risks.

2.2. Hypothesis Development

2.2.1. Health Status and Life Insurance Purchase

In real life, insured people may overstate their health condition and hide some information related to poor health. Therefore, in the underwriting process, life insurance premiums are normally based on two risk factors, gender and age, which may not reflect actuarial life insurance premiums.

Compared with SPH, OHS is a more realistic method of expressing an individual's health status, and can serve as a global measure [18]. It is common for insured to be required to have a health examination or to submit medical reports to the insurer during the process of underwriting under certain conditions, such as above a certain age or with higher coverage. This implies that households with higher health risks (OHS or FHR) pay higher life insurance premiums based on their real health condition. Although the purchase date of life insurance is not included in SHARE data, our study sample consists of households with members who are at least 50 years old. Thus, we expect that most face the uncertainty of adjusted premium through the process of underwriting when they purchase life insurance. We expect a lower probability of purchasing a life insurance policy when an individual has a higher OHS or FHR and, thus, we set the following hypothesis:

Hypothesis 1. *Among three health status indicators (SPH, OHS, and FHR), OHS or FHR is negatively associated with life insurance holdings.*

The perception of health risk is not only a function of current and expected health status, but also of the extent of national health insurance coverage. Atella et al. [7] demonstrate that households in

countries with a less protective healthcare system, based on background risk and poor SPH, have less incentive to invest in risky financial assets. In such cases, the decision to hold risky assets is driven by SPH rather than OHS, which is consistent with the theoretical underpinnings of background risk. In addition to current perceived health, Atella et al. [7] find that households consider FHR in their financial portfolios, especially in non-NHS countries. This suggests an important role for NHS in shaping household portfolio decisions.

Thus, the aims of this paper are to further examine the role of NHS and to investigate the differences between NHS and non-NHS countries. We expect that households with poor health status are less likely to buy life insurance in countries with NHS, and, thus, we set the following hypothesis:

Hypothesis 2. *By examining the impact of NHS, all three health status indicators (SPH, OHS, and FHR) are negatively associated with life insurance holdings, especially in NHS countries.*

2.2.2. Other Factors and Life Insurance Purchases

Education

Most previous studies show a positive relationship between educational level and life insurance demand [10]. Li et al. [11] demonstrate that educational level is positively related to life insurance demand in OECD countries (including 30 European countries). However, Çelik and Kayali [25] find a negative relationship between educational level and life insurance purchases from 2000 to 2006 in European countries. In this study, we expect a positive association between educational level and life insurance holdings in European countries.

Bequest Motive

The main function of life insurance is to provide funds for carrying out family responsibilities in the event of the premature death of a wage earner. The proxies of the bequest motive contain three variables: being married, having children, and a subjective preference for leaving bequests. Life insurance policies (especially term life insurance) are mainly bought for bequest purposes. According to a review by Zietz [26], two papers reveal a negative connection between marital status and life insurance. In contrast, two studies find a positive association between the bequest motive and personal life insurance demand. Inkmann and Michaelides [27] reveal a positive correlation between the demand for life insurance and bequest motive. A more recent study highlights the positive correlation between family members and life insurance demand [28]. Based on this empirical evidence, we expect positive effects of marital status and with child on the demand for life insurance.

Income and Net Wealth

Income is probably the most influential determinant for purchasing life insurance in terms of the ability to pay premiums. Thus, much of the literature shows positive correlation between income level and life insurance demand [26,29]. Çelik and Kayali [25] also find that income is the central variable which affects life insurance purchases in European countries. However, from a review of 12 studies by Zietz [26] regarding the association between wealth and consumption of life insurance, there is no consistent result or correlation. Heo et al. [30] indicate that the amount of insurance purchase increases with net wealth. Shi et al. [28] indicate that both household current income and wealth have positive correlations with life insurance holdings.

Pension

Few studies analyse the relationship between public pension system and life insurance consumption. Among households with low public pension, purchasing life insurance can serve to increase bequest. Thus, there is a higher tendency for self-employed individuals in Germany who

are not covered by the public pension system to buy life insurance and accumulate their wealth to reach higher wealth levels [31]. Andersson and Eriksson [32] also show that compulsory pension reduces the demand for life insurance. Sauter et al. [13] indicate that the impact of public pension as an income source on life insurance demand depends on the relative levels of savings and bequest motive.

Life Expectancy

Li et al. [11] indicate that longer life expectancy is associated with a lower demand for life insurance in OECD countries. In contrast, Inkmann and Michaelides [27] find that term life insurance purchases decrease with higher survival probabilities among elderly households in England. Beck and Webb [9] observe that life expectancy has no connection with life insurance consumption across countries. Thus, we expect the effect of life expectancy on life insurance purchase to be uncertain.

Religion

Based on the literature, the effect of religion on the demand for life insurance varies. Burnett and Palmer [33] indicate that households without religious beliefs have a more positive attitude toward purchasing higher levels of life insurance coverage than those with religious beliefs in the US. In addition, life insurance consumption is significantly lower in Islamic nations [34] and Muslim populations [9,29]. However, Loke and Goh [35] (2011) consider ethnicity as the proxy for religion and demonstrate that both Indians and Chinese are inclined to hold life insurance policies compared to Malays. Thus, we expect that the effect of religion on the demand for life insurance varies due to the differences in religious beliefs (In the SHARE questionnaire, there is generalization of questions pertaining to religious participation. Therefore, religions are not separated into specific categories).

3. Materials and Methods

3.1. Materials

This study uses data from Wave 4 (2010–2011) of SHARE, a survey of households from 16 European countries. It also contains previous information from Wave 1 and Wave 2 (Data from Wave 1 (2004) of SHARE is from 11 countries: Austria, Denmark, France, Germany, Greece, Italy, the Netherlands, Spain, Sweden, Switzerland, and Belgium. Three new European Union members, the Czech Republic, Poland, and Ireland, are included in Wave 2 of SHARE (2006-2007). Wave 3 (2008-2009), SHARELIFE, collects detailed retrospective life histories in 13 countries. All questions are standardized across countries, allowing for consistent international comparisons.). The initial data on life insurance holdings is from households in 11 countries in Wave 1 (2004–2005). Any changes in life insurance holding statuses between Wave 2 (2006–2007) and Wave 4 (2010–2011) are noted. In particular, if a household initially has life insurance holdings in Wave 1, but no life insurance holdings in Wave 2, we consider this household as without life insurance in Wave 4. As changes in life insurance holdings are likely to be related to marital status, we use the marital status specified in Wave 4. In addition, our inference is based on health status measured at the time of the interview, while life insurance purchase is a decision made beforehand.

We analyze the purchasing of life insurance based on the information provided by households in the following 16 countries: Denmark, Sweden, Austria, Belgium, France, Germany, the Netherlands, Switzerland, Poland, the Czech Republic, Italy, Spain, Hungary, Portugal, Slovenia, and Estonia, in Wave 4 (Certain numbers of observations are removed from the panel respondents participating in both waves, particularly for the primary countries Greece and Ireland in which respondents participate in the initial waves but not in Wave 4). SHARE is conducted among households with at least one member aged 32 or more. We focus on the overall financial situation of households and those with respondents who are aged 50 to 90, eliminating observations with missing values for any of the variables relevant to our analysis. Our overall sample consists of 34,341 households.

SHARE is an international, multidisciplinary, and balanced longitudinal survey of various countries in Europe, developed to address research issues on aging. As the main structure of the SHARE survey is generic, the instrument is fixed, and all questions are standardized across countries, our findings allow for consistent international comparisons. SHARE provides comprehensive information on standard demographic variables, health, cognition, intensity of social interaction, and a variety of economic and financial variables, including net wealth, gross income, and household total consumption (For all waves, SHARE interviewers conduct computer-assisted personal interviews to collect most of the data. The structure of the computer-assisted personal interviewing instrument is generic, the instrument is fixed, and only the language used varies among the countries. A detailed description of SHARE data and methodology is published in Börsch-Supan, et al. [36]. Data is available to registered users on the SHARE website (http://www.share-project.org)).

In this paper, health risk is evaluated based on medical expenditures, which affect a household's decision to buy life insurance. Health risk is a function of current and expected health statuses and medical expenditures. These depend not only on health risk, but also on health insurance coverage.

To examine how health risk affects life insurance holdings, we classify countries into two groups: (1) with publicly supported NHS, which offers full coverage; and (2) with NHS that does not provide full coverage (non-NHS). Rather, several forms of private health insurance cover medical expenditures. This raises the overall degree of background risk and hence life insurance holdings may decrease. We split the sample using a method similar to that described by Atella et al. [7] and Bressan et al. [8], distinguishing between countries with NHS with full coverage (Sweden, Spain, Italy, Denmark, Czech Republic, Poland, Hungary, Portugal, Slovenia, Estonia) and countries with NHS with partial coverage (Austria, Germany, the Netherlands, France, Switzerland, Belgium).

Consequently, we expect that an important effect of NHS is on the household decision to hold/buy life insurance. This enables us to investigate whether households are willing to buy life insurance when the financial consequences of health risk are diminished by a highly protective NHS.

In this study, household propensity to purchase life insurance is the dependent variable. We then focus on the health status variables: SPH (the overall assessment by respondents of their health in general), OHS (current overall health status based on the number of chronic diseases), and FHR (as measured by average number of risky behaviours and chronic diseases).

Statistical analysis is applied at the household level, based on responses by household financial respondents. Particularly, financial transfer and asset questions are answered by financial respondents on behalf of the household. Life insurance holdings and types of life insurance variables are also based on financial respondents' responses [11].

3.2. Variables

This section describes the variables based on the characteristics of the households in the whole sample which includes NHS and non-NHS countries. We define three health status variables (SPH, OHS, and FHR) by following the study of Atella et al. [7] who examine the association between health status and portfolio choices in NHS and non-NHS countries separately.

In addition to examining the effects of health status variables on life insurance holdings, we investigate holdings of three categories of life insurance. Basically, life insurance can be classified into term life and whole life. Term life is insurance with a fixed period without cash value after the policy is terminated, but the policyholder can receive claim payment for certain risks during the policy's effective period. Whole life insurance accumulates cash value during the policy period and pays death benefits if the insured dies.

The variables used in this paper are defined as follows: (The detailed information of all variables in this study is shown in the Appendix A, Table A1.)

Life insurance holding: a dummy variable that equals 1 if household holds life insurance and 0 otherwise.

Types of life insurance: a category variable from 1 to 3 (1 = term, 2 = whole, 3 = both).

SPH dummy: Self-perceived health status, categorical: from 1 "very good" to 5 "very bad". We define SPH = 1 if poor self-perceived health (indicating level 3, 4 or 5), SPH = 0 if good health (indicating level 1 or 2).

OHS: OHS is a determinant of current overall health status that considers not only SPH status, but also the numbers of chronic diseases. This study looks at eight types of chronic diseases, including high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychological disorder, and arthritis. Following the procedure used in Zhang et al. [37], the predicated health indicator is obtained from the following formula. $\hat{H}_i^*$ is re-scaled to value in [0, 1]:

$$\widetilde{H}_i^* = \frac{\hat{H}_i^* - \hat{H}^{min}}{\hat{H}^{max} - \hat{H}^{min}},$$

where $\hat{H}^{max}$ and $\hat{H}^{min}$ are, respectively, the largest and the smallest predicted values. The association between life insurance and health can be analysed using the adjusted health indicator $\widetilde{H}_i^*$ as the well-being measurement. Thus, households with poor health are more likely to have higher OHS value (Attela et al. [7] use a more complicated term "weighted number of chronic diseases", where the weights are derived according to the degree of severity of disease and the implied disability [38].).

FHR: FHR is evaluated by increasing function of the average number of risky behaviours (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis) and decreasing function of average household grip strength. The higher the score, the greater the number and severity of perceived problems (Readers may refer to Attela et al. [7] for more information on the definition of FHR.).

Age $\geq$ 65: a dummy variable that equals 1 if the age of household respondent is 65 or older and 0 if the age of household respondent is less than 65.

Higher education: a dummy variable that equals 1 if the number of years of education completed is more than or equal to 10 and 0 otherwise.

Marital status: a scale from 1 to 4 (1 = married, 2 = divorced, 3 = widowed and 4 = never married).

With children: a dummy variable that equals 1 if household includes child(ren) and 0 otherwise.

Household size: the total number of household members.

Household income per capita: the monthly household income divided by household size.

Net wealth per capita: the monthly household net wealth divided by household size.

Pension: household pension.

Probability of receiving inheritance: the probability of receiving any positive amount of inheritance.

Higher life expectancy: a dummy variable that equals 1 if household expects to live at least 10 years and 0 otherwise.

Social activity: a dummy variable that equals 1 if household has social interaction and 0 otherwise.

Religious participation: a dummy variable that equals 1 if household participates in religious activities and 0 otherwise.

Non-NHS: a dummy variable that equals 1 if household is in non-NHS country and 0 if household is in NHS country.

Health spending from coverage: the country-level data of health spending from government or compulsory schemes.

Out-of-pocket health spending: the country-level data of health spending from voluntary schemes or household out-of-pocket payments.

Future retirement age: the country-level data of the future retirement age for a person who entered the labour force at age 20 (average age if future retirement ages are different from men and women).

Gross pension replacement rate: the country-level data of the gross pension entitlement divided by gross pre-retirement earnings in term of country level (average value if gross pension replacement rates are different from men and women).

3.3. Methods

The following probit model [39] is used to examine whether health status variables are related to life insurance holdings:

$$
\begin{aligned}
\text{Pr (Life insurance holding)} &= X\beta + \varepsilon, \\
&= \beta_0 + \beta_1 \, SPH \, dummy + \beta_2 \, OHS + \beta_3 \, FHR + \beta_4 \, SPH \, dummy \times Higher \, education + \beta_5 \, OHS \\
&\quad \times Higher \, education + \beta_6 \, FHR \times Higher \, education + \beta_7 \, SPH \, dummy \times Age \geq 65 + \beta_8 \, OHS \\
&\quad \times Age \geq 65 + \beta_9 \, FHR \times Age \geq 65 + \beta_{10} \, Non\text{-}NHS + \varepsilon,
\end{aligned}
\tag{1}
$$

where Y is the binary response variable, life insurance holdings, and β_i (i = 1, 2, 3) are the parameters of the three health status variables (*SPH dummy*, *OHS*, and *FHR*). Interactions of different variables and country variable are included in (1). ε is an error term.

Atella et al. [7] test the effects of health status variables with age by splitting the data into distinctive groups to analyse whether FHR varies by educational level in terms of portfolio choice. It is interesting to evaluate how the three health status variables (*SPH dummy*, *OHS*, and *FHR*) interact with some demographic characteristics (*Education*, $Age \geq 65$) to impact on the decision to hold life insurance. To examine the holding of life insurance, health status variables are multiplied by educational level and by age for the whole sample and NHS and non-NHS subsamples.

In addition, we follow Barasinska et al. [40] (This study focuses on individual risk attitudes and the composition of financial portfolios in Europe) to examine the marginal effects of *J* outcomes (*J* = 4) using the following multinomial logit (hump-shaped pattern) regression model to analyse the probability of observing a specific type of life insurance holding, $Prob(Y_j)$, in the pooled data from all 16 countries:

$$
Prob(Y_j) = \frac{\exp\left(X\beta_j\right)}{\sum_{n=1}^{J} \exp(X\beta_n)} \quad n = 0, \, 1, \, 2, \, \ldots, J; \, j = 1, \, 2, \, 3, \, 4; \, j \neq n,
\tag{2}
$$

in which *X* is the vector of explanatory variables that include health status variables (*SPH dummy*, *OHS*, and *FHR*) and other controls.

4. Results

4.1. Descriptive Statistics

Table 1 provides the summary statistics and means for all variables. Of the 34,341 households, 21% hold life insurance in the whole sample, 18.5% in NHS countries, and 24.3% in non-NHS countries. In the whole sample, households with life insurance have lower incidences of poor SPH, OHS, and FHR (29%, 0.07, and 0.25, respectively) compared with those without life insurance (43%, 0.11 and 0.33, respectively). To sum up, these three health status variables are significantly lower among those with life insurance than among those without life insurance in both NHS and non-NHS countries.

In general, younger age (50–64 compared to 65 and above), higher educational level, married status, with children, larger household, higher income and net wealth, lower pension, lower health spending from coverage, lower out-of-pocket health spending, lower future retirement age, lower gross pension replacement rate, higher probability of receiving inheritance, higher life expectancy (50% or higher), socially active, and more religious are more strongly associated with holding life insurance. However, households with certain religious beliefs tend to have life insurance holdings only in countries with NHS (Table 1).

Table 1. Demographic and economic characteristics of holding life insurance in NHS countries, Age 50+.

		Whole (N = 34,341)			NHS (N = 19,383)			Non-NHS (N = 14,958)		
		with LI	without LI	p-Value	with LI	without LI	p-Value	with LI	without LI	p-Value
		7216 (21%)	27,125 (79%)		3587 (18.5%)	15,796 (81.5%)		3629 (24.3%)	11,329 (75.7%)	
Health variables										
SPH dummy (%) (1 if poor health)	Poor	29.2	43.5	<0.001 &	35.1	52.2	<0.001 &	23.4	31.4	<0.001 &
	Good	70.8	56.5		64.9	47.8		76.6	68.6	
SPH discrete, mean (SD) (from 1 "very good" to 5 "very bad")		2.9 (1.1)	3.3 (1.1)	<0.001 @	3.0 (1.1)	3.5 (1.0)	<0.001 @	2.8 (1.0)	3.0 (1.0)	<0.001 @
Objective health status (OHS), mean (SD)		0.07 (0.1)	0.11 (0.1)	<0.001 @	0.08 (0.1)	0.12 (0.2)	<0.001 @	0.07 (0.1)	0.09 (0.1)	<0.001 @
Future health risk (FHR), mean (SD)		0.25 (0.5)	0.33 (0.7)	<0.001 @	0.25 (0.5)	0.32 (0.6)	<0.001 @	0.24 (0.6)	0.33 (0.8)	<0.001 @
Demographic variables										
Age, mean (SD)		64.1 (8.0)	69.9 (9.5)	<0.001 @	64.5 (8.0)	69.9 (9.4)	<0.001 @	63.8 (8.0)	69.8 (9.7)	<0.001 @
Age (%)	50 ≤ age < 65	59.2	32.6	<0.001 &	57.3	32.3	<0.001 &	61.1	33	<0.001 &
	≥65	40.8	67.4		42.7	67.7		38.9	67	
Education (%)	Lower	31	43.7	<0.001 &	28	44.4	<0.001 &	33.9	42.9	<0.001 &
	Higher	69	56.3		72	55.6		66.1	57.1	
Marital status (%)	Married	73.4	61.5	<0.001 &	76.6	62.3	<0.001 &	70.2	60.4	<0.001 &
	Divorced	11.8	12.1		9.6	10.8		13.9	13.8	
	Widowed	8.9	18.8		9	20.1		8.8	17	
	Never married	5.9	7.6		4.9	6.7		7	8.7	
With children (%)		91.3	88.8	<0.001 &	93.5	90.7	<0.001 &	89.1	86.3	<0.001 &
Household size, mean (SD)		2.3 (1.0)	2.0 (1.0)	<0.001 @	2.4 (1.1)	2.1 (1.1)	<0.001 @	2.2 (1.0)	1.9 (0.9)	<0.001 @
Household income per capita, mean (SD)		0.210 (0.35)	0.185 (0.337)	<0.001 @	0.143 (0.170)	0.132 (0.21)	0.004 @	0.276 (0.450)	0.260 (0.447)	0.060 @
Net wealth per capita, mean (SD)		1.454 (3.36)	1.098 (3.608)	<0.001 @	1.202 (4.15)	0.821 (4.13)	<0.001 @	1.703 (2.303)	1.486 (2.661)	<0.001 @
Pension, mean (SD)		0.054 (0.133)	0.074 (0.159)	<0.001 @	0.040 (0.092)	0.049 (0.086)	<0.001 @	0.067 (0.162)	0.108 (0.219)	<0.001 @
Future retirement age, mean (SD)		65.3 (3.04)	65.7 (2.88)	<0.001 @	65.3 (3.82)	65.9 (3.35)	<0.001 @	65.4 (1.98)	65.3 (0.01)	<0.001 @
Gross pension replacement rate, mean (SD)		58.2 (18.0)	59.4 (17.3)	<0.001 @	54.6 (16.7)	58.3 (3.02)	<0.001 @	61.8 (18.5)	61.0 (18.5)	<0.001 @
Health spending from coverage, mean (SD)		0.484 (0.620)	0.529 (0.647)	<0.001 @	0.740 (0.763)	0.801 (0.785)	<0.001 @	0.226 (0.270)	0.333 (0.431)	<0.001 @
Out-of-pocket health spending, mean (SD)		0.141 (0.170)	0.158 (0.179)	<0.001 @	0.215 (0.198)	0.238 (0.205)	<0.001 @	0.067 (0.086)	0.101 (0.131)	<0.001 @
Prob. of receiving inheritance, mean (SD)		22.0 (34.4)	11.5 (26.7)	<0.001 @	18.4 (32.7)	8.8 (23.5)	<0.001 @	25.6 (35.6)	15.3 (30.2)	<0.001 @
Higher life expectancy (%)		81.2	69.7	<0.001 &	76.5	64.7	<0.001 &	85.8	76.5	<0.001 &
Socially active (%)		39.4	27.1	<0.001 &	35	20.8	<0.001 &	43.6	35.9	<0.001 &
Religious participation (%)		14.5	13.1	0.003 &	15.2	12	<0.001 &	13.8	14.6	0.20 &

Note: The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good or 2 for good). The discrete SPH variable is measured on a scale of 1 to 5 (1. Very Good, 2. Good, 3. Fair, 4. Bad, 5. Very Bad). OHS variable is a determinant of current overall health as this variable not only considers SPH status but also the numbers of chronic diseases. FHR is evaluated by the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. & Chi-Square Test; @ t-test. Monetary amounts are PPP-adjusted and in thousand Euros. Household income per capita, Net wealth per capita, Pension, Health spending from coverage, and Out-of-pocket spending are transferred via natural logarithm. Source: the SHARE data and OECD. Stat (https://stats.oecd.org/Index.aspx).

To examine the strength of association for life insurance purchase and explanatory variables, we calculate Pearson's correlation. The results are shown in Table A2. There is significantly negative association between life insurance holdings and the three health status variables.

4.2. The Whole Sample

Table 2 reports the marginal effects of the variables of interest on the decision to hold life insurance for the whole sample. We first examine the marginal effects of each health status indicator (SPH, OHS, or FHR) separately on the decision to hold life insurance and then examine all three health status indicators simultaneously. For simplification, we only display the estimation results of model including three health status indicators simultaneously in Table 2. The estimation results for each health status indicator (SPH, OHS, or FHR) are provided in Appendix A Table A3.

In Table 2, when we consider all three health status variables simultaneously, the marginal effects of SPH and FHR decrease the probability of purchasing life insurance by 2.1% and 1.2%, respectively. It is important to note that SPH is the most influential factor among the three variables in the decision to hold life insurance.

Younger households ($50 \leq$ age < 65) are more likely to hold life insurance. The marginal effects of age over 65 decrease the probabilities of purchasing life insurance by 12.7%. Our results confirm that the marginal effects of higher educational level increase the probabilities of owning life insurance by 2.2%.

In addition, compared with households in which respondents are divorced or widowed, married households are more likely to hold life insurance, consistent among all three health status variables. We also find that households with children have a higher propensity to hold life insurance the probabilities of purchasing life insurance increase by 2.8%, respectively. Moreover, the probability of holding life insurance increases with both income and net wealth. In addition, the probabilities of receiving inheritance and higher life expectancy (50% or higher) are associated with higher probabilities of owning life insurance.

Furthermore, household pension income is significantly and negatively associated with the decision to buy life insurance and households that are more socially active and possess religious beliefs are more inclined to purchase life insurance. Health spending from coverage (out-of-pocket health spending) is significantly and positively (negatively) associated with the decision to purchasing life insurance, indicating that households are more likely to hold life insurance when there is more (less) health spending from government or compulsory scheme (voluntary schemes or household out-of-pocket payments). In addition, for the impact of national pension system, our results show that future retirement age has a negative and significant association, whereas gross pension replacement rate has positive and significant association, with life insurance holding. Finally, people in advanced countries are more inclined to own life insurance, when compared with those in emerging countries.

4.3. Analysis for Interaction Effects

The second column in Table 2 presents the results for the interaction terms among the three health status indicators and non-NHS countries. Unhealthy households (in terms of bad perceived health or poor objective health status) are less likely to own life insurance, except those with higher SPH located in non-NHS countries. This indicates that those households with poor self-perceived health status are more likely purchase life insurance in non-NHS countries than in NHS countries.

Table 2. Probit regression of purchasing life insurance by whole sample, Age 50+.

Dep. var.: Life Insurance Holding	The Whole Sample (N = 34,341)	
	Model 1	Model 2
SPH dummy	−0.0205 ***	−0.0346 ***
	(0.005)	(0.006)
OHS	0.00152	−0.0162
	(0.018)	(0.023)
FHR	−0.0121 ***	−0.0100 *
	(0.004)	(0.006)
Demographic variables		
Age ≥ 65	−0.127 ***	−0.124 ***
	(0.007)	(0.007)
Higher education	0.0216 ***	0.0190 ***
	(0.005)	(0.005)
Marital status (ref: Never married)		
Married	0.0566 ***	0.0564 ***
	(0.009)	(0.009)
Divorced	0.0142	0.0133
	(0.010)	(0.010)
Widowed	−0.00574	−0.00729
	(0.010)	(0.010)
With children	0.0281 ***	0.0286 ***
	(0.008)	(0.008)
Household income per capita	0.0225 ***	0.0230 ***
	(0.003)	(0.003)
Household income per capita squared	0.000	0.000
	(0.000)	(0.000)
Net wealth per capita	0.0148 ***	0.0151 ***
	(0.001)	(0.001)
Net wealth per capita squared	0.000	0.000
	(0.000)	(0.000)
Prob. of receiving inheritance	0.000630 ***	0.000649 ***
	(0.000)	(0.000)
Pension	−0.00423 ***	−0.00449 ***
	(0.001)	(0.001)
Future retirement age	−0.0166 ***	−0.0154 ***
	(0.001)	(0.001)
Gross pension replacement rate	0.000484 **	0.000579 ***
	(0.000)	(0.000)
Health spending from coverage	0.0907 ***	0.0859 ***
	(0.007)	(0.007)
Out-of-pocket health spending	−0.0938 ***	−0.0870 ***
	(0.008)	(0.008)
Higher Life expectancy	0.0370 ***	0.0378 ***
	(0.005)	(0.005)
Social activity	0.0401 ***	0.0400 ***
	(0.005)	(0.005)
Religious participation	0.0276 ***	0.0276 ***
	(0.006)	(0.006)

Table 2. *Cont.*

Dep. var.: Life Insurance Holding	The Whole Sample (N = 34,341)	
	Model 1	Model 2
Advanced (ref: Emerging)	0.0275 ***	
	(0.008)	
Non-NHS (ref: NHS)		0.00235
		(0.006)
SPH dummy × Non-NHS		0.0327 ***
		(0.008)
OHS × Non-NHS		0.044
		(0.035)
FHR × Non-NHS		0.004
		(0.007)
Pseudo R^2	0.101	0.102
Log likelihood	−15,842.956	−15,852.956

Note: This table reports mean marginal effects evaluated at each observation. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4 or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 for good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

4.4. NHS versus Non-NHS Countries

We next investigate whether NHS impacts on the relation between health status and life insurance holdings, by analyzing the data of the two subgroups: NHS versus non-NHS countries. We follow the same MNL model to observe whether there is significant impact by any explanatory variable on the dependent variable in the whole sample.

Table 3 shows that the estimated marginal effects are different from the results of the whole sample. When considering all three health status indicators together, we find that SPH and FHR health status variables are significantly and negatively associated with life insurance only in NHS countries (2.7% and 0.9%, respectively). The estimation results for each health status indicator (SPH, OHS, or FHR) are provided in Table A4 displayed in the Appendix A. However, only FHR has a significant influence on the decision to purchase life insurance (1.7% decrease) for non-NHS countries.

Table 3. Probit regression of purchasing life insurance by different NHS, Age 50+.

Dep. var.: Binary for Holding Life Insurance	Non-NHS (N = 14,958)	NHS (N = 19,383)
SPH dummy	−0.0038	−0.0267 ***
	(0.008)	(0.006)
OHS	0.0406	−0.0179
	(0.031)	(0.021)
FHR	−0.0166 ***	−0.00868 *
	(0.005)	(0.005)
Demographic variables		
Age ≥ 65	−0.152 ***	−0.0911 ***
	(0.011)	(0.008)
Higher education	0.0126 *	0.0238 ***
	(0.007)	(0.006)

Table 3. *Cont.*

Dep. var.: Binary for Holding Life Insurance	Non-NHS (N = 14,958)	NHS (N = 19,383)
Marital status (ref: Never married)		
Married	0.0532 ***	0.0529 ***
	(0.014)	(0.012)
Divorced	0.0212	−0.0016
	(0.016)	(0.014)
Widowed	−0.00777	−0.0144
	(0.016)	(0.013)
With children	0.0299 ***	0.0304 ***
	(0.011)	(0.011)
Income per capita	0.0174 ***	0.0289 ***
	(0.004)	(0.003)
Income per capita squared	0.000	0.000
	(0.000)	(0.000)
Net wealth per capita	0.0184 ***	0.0161 ***
	(0.001)	(0.001)
Net wealth per capita squared	0.000	0.000
	(0.000)	(0.000)
Prob. of receiving inheritance	0.000543 ***	0.000891 ***
	(0.000)	(0.000)
Pension	−0.00552 ***	−0.00456 ***
	(0.001)	(0.001)
Future retirement age	−0.0202 ***	0.00498 ***
	(0.003)	(0.002)
Gross pension replacement rate	0.00143 ***	−0.00500 ***
	(0.000)	(0.000)
Health spending from coverage	0.116 ***	0.00702
	(0.014)	(0.011)
Out-of-pocket health spending	−0.154 ***	0.0152
	(0.015)	(0.011)
Higher Life expectancy	0.0449 ***	0.0348 ***
	(0.009)	(0.006)
Social activity	0.0269 ***	0.0644 ***
	(0.007)	(0.007)
Religious participation	−0.0156 *	0.0328 ***
	(0.009)	(0.008)
Pseudo R^2	0.093	0.123
Log likelihood	−7510.732	−8133.964

Note: Mean marginal effects evaluated at each observation. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 for good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, * $p < 0.1$ Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

The results of the other explanatory variables in NHS countries are comparable with the findings of the whole sample, except for pension. In contrast, our empirical findings related to the relationship between religion and life insurance demand are unclear in non-NHS countries.

We investigate the interaction effects, focusing on those among the three health status variables and two demographic variables of older age (Age ≥ 65) and higher level of education, following Atella et al. [7]. This study highlights the impact of age on household decisions to hold risky assets. The primary intention is to examine those interactions and whether there are differences in health

factors depending on if the person in the household deciding on life insurance purchase is elderly or highly educated.

Table 4 illustrates that there is no significant result for the interaction terms among the three health status variables and the higher level of education. We further test the interaction effects of health status with old age (Age $\geq$ 65) on the probability of owning life insurance in the whole sample and in the two subsamples, NHS and non-NHS countries, separately, shown in Table 5. Interestingly, the marginal effects of the interaction between FHR and Age $\geq$ 65 on life insurance are strongly (1%) significant for both whole sample and in the subsample of non-NHS countries when considering the health status variables together. We present the estimation results by considering each health status individually, provided in Appendix A Table A6. This implies that elderly households with high probability of future risk tend to purchase life insurance. However, it is very difficult for insurers to distinguish among the types of health situations that insured may present with in the future.

Table 4. Interaction terms with three health status indicators and higher education, whole sample, NHS countries, and Non-NHS countries, Age 50+.

Dep. var.: Binary for Holding Life Insurance	The Whole Sample (N = 34,341)	Non-NHS (N = 14,958)	NHS (N = 19,383)
SPH dummy	−0.0155 **	0.000457	−0.0224 **
	(0.008)	(0.013)	(0.010)
OHS	0.0106	0.0476	−0.00458
	(0.025)	(0.043)	(0.030)
FHR	−0.0101 *	−0.0155 **	−0.00664
	(0.005)	(0.008)	(0.008)
Demographic variables			
Age≥65	−0.127 ***	−0.152 ***	−0.0912 ***
	(0.007)	(0.011)	(0.008)
Higher education	0.0273 ***	0.0165 *	0.0304 ***
	(0.006)	(0.009)	(0.008)
SPH dummy × Higher education	−0.00795	−0.00723	−0.00634
	(0.010)	(0.017)	(0.012)
OHS × Higher education	−0.0185	−0.0141	−0.0263
	(0.035)	(0.060)	(0.041)
FHR × Higher education	−0.00366	−0.0023	−0.00369
	(0.007)	(0.011)	(0.010)
Other controls	Marital status, with child, Household income per capita, Household income per capita squared, Net wealth per capita, Net wealth per capita squared, Probability of receiving inheritance, Pension, Future retirement age, Gross pension replacement rate, Health spending from coverage, Out-of-pocket health spending, Higher life expectancy, Social activity, Religious participation, and Country dummy		
Pseudo R^2	0.101	0.093	0.124
Log likelihood	−15,858.79	−7510.494	−8133.286

Note: Means of marginal effects evaluated at each observation. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: the SHARE data and OECD. Stat (https://stats.oecd.org/Index.aspx).

Table 5. Interaction terms with three health status indicators and age ≥ 65, whole sample, NHS countries, and Non-NHS countries, Age 50+.

Dep. var.: Binary for Holding Life Insurance	The Whole Sample (N = 34,341)	Non-NHS (N = 14958)	NHS (N = 19383)
SPH dummy	−0.0211 ***	−0.00878	−0.0190 **
	(0.007)	(0.013)	(0.009)
OHS	0.0204	0.0285	0.0118
	(0.032)	(0.056)	(0.037)
FHR	−0.0243 ***	−0.0367 ***	−0.0121
	(0.007)	(0.010)	(0.009)
Demographic variables			
Age ≥ 65	−0.130 ***	−0.165 ***	−0.0822 ***
	(0.008)	(0.012)	(0.010)
Higher education	0.0217 ***	0.0127 *	0.0237 ***
	(0.005)	(0.007)	(0.006)
SPH poor $\times$ Age ≥ 65	0.00113	0.00949	−0.0146
	(0.010)	(0.017)	(0.012)
OHS $\times$ Age ≥ 65	−0.0277	0.0111	−0.0408
	(0.038)	(0.067)	(0.045)
FHR $\times$ Age ≥ 65	0.0178 **	0.0297 ***	0.00487
	(0.008)	(0.012)	(0.011)
Other controls	Marital status, with child, Household income per capita, Household income per capita squared, Net wealth per capita, Net wealth per capita squared, Probability of receiving inheritance, Pension, Future retirement age, Gross pension replacement rate, Health spending from coverage, Out-of-pocket health spending, Higher life expectancy, Social activity, Religious participation, and Country dummy		
Pseudo R^2	0.101	0.094	0.124
Log likelihood	−15,857.09	−7506.71	−8132.09

Note: Means of marginal effects evaluated at each observation. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

4.5. Marginal Effect of the Preference for Type of Life Insurance

The marginal effects of each of the three health status variables on the probability of holding different types of insurance are estimated for the whole sample, including term life insurance, whole life insurance, both (term life and whole life insurance) policies, and no life insurance. When we consider only OHS indicator in our regression model, our results show that OHS is only significantly and negatively related to both types of life insurance, the estimation results are shown in Table A7 exhibited in the Appendix A. When considering all three health status variables, shown in Table 6, the results reveal that households with poor SPH or high FHR risk are less likely to hold only term life or both types of life insurance. It seems that poor health statuses (for SPH and FHR indicators) are not supportive of holding only term life insurance. A possible explanation is that term life policy is associated with relatively lower premiums compared to whole life policy and insurers have stricter underwriting process in terms of evaluating the health condition of the insured. Thus, households with poor health status are less likely to buy a term life policy. In contrast, those with no life insurance have lower SPH and higher FHR. However, we do not find any significant evidence for the relationship between the decision to own only whole life insurance and health status.

Table 6. Marginal effects of purchasing different types of life insurance, whole sample, N = 34,341, Age 50+.

Dep. var.: Binary for Holding Life Insurance	Term	Whole	Both	No LI *
SPH dummy	−0.00700 ***	0.00228	−0.0169 ***	0.0217 ***
	(0.003)	(0.004)	(0.004)	(0.005)
OHS	0.0133	−0.00313	−0.00454	−0.00561
	(0.009)	(0.013)	(0.013)	(0.018)
FHR	−0.00565 ***	0.000574	−0.00887 ***	0.0139 ***
	(0.002)	(0.003)	(0.003)	(0.004)
Demographic variables				
Age ≥ 65	−0.0328 ***	−0.0343 ***	−0.0579 ***	0.125 ***
	(0.003)	(0.005)	(0.005)	(0.007)
Higher education	0.00514 **	0.0114 ***	0.00663 **	−0.0232 ***
	(0.002)	(0.003)	(0.003)	(0.005)
Marital status (ref: Never married)				
Married	0.00913 **	0.0324 ***	0.0161 **	−0.0576 ***
	(0.004)	(0.006)	(0.007)	(0.009)
Divorced	0.0107 **	0.0148 **	−0.0115	−0.0139
	(0.005)	(0.007)	(0.007)	(0.010)
Widowed	0.0019	0.0115 *	−0.0187 ***	0.00532
	(0.005)	(0.007)	(0.007)	(0.010)
With children	0.00685 **	−0.00343	0.0238 ***	−0.0272 ***
	(0.003)	(0.006)	(0.005)	(0.008)
Income per capita	0.00570 ***	0.00521 ***	0.0101 ***	−0.0210 ***
	(0.001)	(0.002)	(0.002)	(0.003)
Income per capita squared	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Net wealth per capita	0.00467 ***	0.00868 ***	0.00234 ***	−0.0157 ***
	(0.001)	(0.001)	(0.001)	(0.001)
Net wealth per capita squared	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Prob. of receiving inheritance	0.000136 ***	0.000285 ***	0.000235 ***	−0.000657 ***
	(0.000)	(0.000)	(0.000)	(0.000)
Pension	−0.000789 ***	−0.00146 ***	−0.00222 ***	0.00447 ***
	(0.000)	(0.000)	(0.000)	(0.001)
Future retirement age	−0.00628 ***	−0.00318 ***	−0.00630 ***	0.0158 ***
	(0.001)	(0.001)	(0.001)	(0.001)
Gross pension replacement rate	0.00149 ***	−0.000885 ***	−0.0000854	−0.000518 ***
	(0.000)	(0.000)	(0.000)	(0.000)
Health spending from coverage	−0.0423 ***	0.0641 ***	0.0574 ***	−0.0792 ***
	(0.004)	(0.005)	(0.005)	(0.008)
Out-of-pocket health spending	0.0267 ***	−0.0335 ***	−0.0781 ***	0.0850 ***
	(0.004)	(0.005)	(0.006)	(0.008)
Higher Life expectancy	0.00634 **	0.0183 ***	0.0119 ***	−0.0365 ***
	(0.003)	(0.003)	(0.003)	(0.005)
Social activity	0.0191 ***	0.00756 **	0.0127 ***	−0.0394 ***
	(0.002)	(0.003)	(0.003)	(0.005)
Religious participation	0.00278	0.0185 ***	0.00329	−0.0245 ***
	(0.003)	(0.005)	(0.004)	(0.006)
Advanced markets (ref: Emerging markets)	0.0452 ***	−0.0268 ***	0.0145 ***	−0.0330 ***
	(0.003)	(0.006)	(0.005)	(0.008)
Log likelihood			−22,411.1	

Note: Mean marginal effects evaluated at each observation. No LI * means households without life insurance. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 for good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** *p* < 0.01, ** *p* < 0.05, * *p* < 0.1. Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

Our results show no clear confirmation of the association between any of the three health status variables and whole life insurance holdings. Households without life insurance holdings are more likely to have serious health conditions or risks. Moreover, households with term life insurance only are more likely to be young, highly educated, married, socially active and more religious, with children, higher income and net wealth, higher probability of receiving inheritance, and higher life expectancy. In addition, households with whole life insurance are more likely to have married, divorced, or widowed respondents and young, highly educated, socially active and more religious, higher income and net wealth, higher probability of receiving inheritance, and higher life expectancy. Furthermore, owning both types of life insurance is significantly associated with younger age, married status, with children, higher income and net wealth, higher probability of receiving inheritance, higher life expectancy, and being socially active.

The signs of demographic variables for the households having no life insurance are, in general, the opposite of those for the households purchasing different types of insurance. The characteristics of the households with no life insurance holdings include older age, lower educational level, unmarried status, with fewer children, having less financial possibility but higher pension income, lower probability of receiving inheritance, lower life expectancy, socially inactive and less religious.

5. Discussion

To our knowledge, this is the first study that examines the effects of three different health status indicators on the demand for life insurance in European countries. Our findings show that two health status indicators (SPH and FHR) are negatively associated with the decision to hold life insurance in the whole sample, but only in NHS countries on subgroup analysis. Our results are consistent with the findings of Fang and Kung [15], in which healthy individuals are more likely to have life insurance in the US. However, results of subgroup analysis show that only FHR has a decreasing effect on the purchase of life insurance in non-NHS countries.

Consistent with Inkmann and Michaelides [27], we find a negative relationship between age and households owning life insurance. An earlier study by Zeitz [26] did not reveal evidence of an association between educational level and the holding of life insurance. Later on, Ward and Zurbruegg [29] and Li et al. [11] demonstrated that life insurance demand is lower (higher) among households with lower (higher) educational level. Thus, this study provides additional confirmation that educational level is positively associated with life insurance demand. Moreover, consistent with the findings in the literature on the bequest motive [27,28], our results reveal that households with married respondents or with children have a higher propensity to hold life insurance.

In addition, consistent with the literature [10,11,26], our findings also indicate that households with higher income level or higher net wealth have higher life insurance demand. Interestingly, we find that households with religious beliefs are more inclined to hold life insurance. However, this finding differs from those of previous studies on Islamic countries and Muslim communities [9,29,34]. Thus, from the literature and our results, we conclude that the decision to purchase life insurance might depend on particular religious beliefs.

Moreover, our results demonstrate that both SPH and FHR health status indicators are significantly and negatively associated with life insurance holdings in the entire sample as well as in the subsample in NHS countries. However, only FHR has a significantly negative impact on the decision to purchase life insurance in non-NHS countries. Our findings contradict those of Atella et al. [7], in which the relationship between health status variables and risky financial assets holdings are ambiguous in NHS countries, but the holdings of risky financial assets decrease in households, based on SPH and FHR in non-NHS countries. Our findings also differ from those of Bressan et al. [8] who reveal that SPH is significantly and negatively related to direct and indirect stockownership in both NHS and non-NHS countries. The implication of our study is that the decision to hold life insurance or risky assets among households differs from the assessment of financial tools from the risk management point of view.

6. Conclusions

Compared with previous works of other determinants on life insurance ownership based on data from only one country, our study examines household life insurance holdings in 16 European countries and focuses on three health status variables. We also investigate the impact of health status on the purchase of different types of life insurance (term life only, whole life only, or both types) and characteristics of the health care system in the countries of residence. In all cases, we include a set of standard socio-economic and demographic variables as control variables. Our whole sample is classified into NHS and non-NHS countries, depending on national health insurance coverage.

Our results demonstrate that both SPH and FHR health status variables are negatively related to life insurance holdings in the whole sample and in NHS subgroups, indicating that unhealthy households are less likely to hold life insurance. In non-NHS countries, only FHR has negative effects. This implies that households treat national health insurance as a self-insurance mechanism, substituting market insurance for NHS. We also find that older households with higher FHR are more inclined to hold life insurance in the pooled data and non-NHS sample.

There are some interesting findings on the demand for different types of life insurance in the whole sample. The results demonstrate that both SPH and FHR health status indicators have negative impacts on holding term life only and on holding both term life and whole life insurance policies. In contrast, there is no relationship between health status and holding whole life only. It seems that there is no adverse selection problem in the life insurance market, especially in European countries with NHS. Our empirical results may provide policy implications for insurers in European countries in that the marketing strategies for life insurance should consider not only demographic factors, but also household health status and national health insurance coverage.

Appendix A

Table A1. List of variables.

	Variables	Description
Dependent variables	Life insurance holding	We consider that an individual has life insurance if that individual positively answered the question, "Do you [or/or/or/or] [your/your/your/your] [husband/wife/partner/partner] currently own any life insurance policies?"
	SPH dummy	Self-perceived health status, categorical: from 1 "very good" to 5 "very bad". The dummy variable measures poor self-perceived health and takes the value of 1 whenever average household Perceived Health Status is "fair" (3), "bad" (4), or "very bad" (5), and takes the value 0 otherwise.
	SPH discrete	Self-perceived health status, categorical: from 1 "very good" to 5 "very bad".
Main independent variables	Objective health status (OHS)	OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. The higher the score the greater the number and severity of perceived problems.
	Future Health Risk (FHR)	The variable FHR is evaluated by the increasing function of the average number of risky behaviours (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. The higher the score the greater the number and severity of perceived problems. We construct this health variable according to the method described by Atella et al. (2012).
	Age	In order to construct age variable, we use respondent's birth year. Age ranges from 50 to 90 years, with two subgroups: younger (50 < age < 65) and elderly (aged ≥ 65).
	Education	Education systems in SHARE countries are highly heterogeneous. We follow Atella et al. (2012) to rely on the completed years of education rather than educational degrees. How many years have you been in full time education (including receiving tuition, engaging in practical work or supervised study or taking examinations). Possible range is 0–25 years. High level of education is defined as the number of completed years of education higher than or equal to 10 and low level of education is equal to compulsory full education, which is 9 years.
	Marital status	What is your marital status? 1. Married and living together with spouse, 2. Registered partnership, 3. Married, and living separated from spouse, 4. Never married, 5. Divorced, 6. Widowed. We define married individuals as those providing one of the first three answers.
	With children	How many children do you have that are still alive? Please count all biological, foster, adopted and stepchildren.
Demographic variables	Household size	Total number of household members.
	Household income ('000s euro)	We focus on wave 4, thus Version B (thinc2) is more suitable to our study. Version B is obtained from the one shot question on total household monthly income.
	Net wealth	The sum of the financial and real assets of the household (including the value of owned property) net of financial liabilities.
	Pension ('000s euro)	Annual old age & early retirement pensions (ypen1).

Table A1. *Cont.*

Variables	Description
Future retirement age	The future retirement age for a person who entered the labor force at age 20. Source: OECD.Stat (https://stats.oecd.org/Index.aspx).
Gross pension replacement rate	The gross replacement rate is defined as gross pension entitlement divided by gross pre-retirement earnings. Source: OECD.Stat (https://stats.oecd.org/Index.aspx).
Health spending from coverage (billion euro)	The country-level data of health spending from government or compulsory schemes. Source: OECD.Stat (https://stats.oecd.org/Index.aspx).
Out-of-pocket health spending (billion euro)	The country-level data of health spending from voluntary schemes or household out-of-pocket payments. Source: OECD. Stat (https://stats.oecd.org/Index.aspx).
Prob. of inheritance	Thinking about the next ten years, what are the chances that you will receive any inheritance, including property and other valuables (0–100 percent)?
Higher life expectancy	To define life expectancy, we use the question "What are the chances that you will live to be age [75/80/85/90/95/100/105/110/120] or more?" Possible choices are between 0-100. High life expectancy defined as higher than or equal to 50 percent of the individuals expecting to live at least 10 years. Less than 50 percent defines low life expectancy.
Lifestyle variables	
Socially active	Social participation is defined by the question of "Gone to a sport, social or other kind of club and political or community in the past twelve months?". We construct this variable by following Christelis et al. (2010), and Atella et al. (2012).
Religious participation	The general question for religion "Taken part in activities of a religious organization" (church, synagogue, mosque etc.)" following Christelis et al. (2010), and Atella et al. (2012).

Table A2. Correlation coefficients of variables.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
A Life insurance holdings	1																							
B OBS	-0.093*	1																						
C SPH dummy	-0.119*	0.400*	1																					
D FHR	-0.051*	0.154*	0.153*	1																				
E Age≥65	-0.245*	0.226*	0.257*	0.138*	1																			
F Higher education	0.105*	-0.128*	-0.110*	0.202*	0.090*	1																		
G Marital status	-0.102*	0.099*	0.0795	-0.127*	-0.066*	0.159*	1																	
H With children	0.032*	-0.000	0.007	0.094*	0.043*	0.006	-0.059*	1																
I Household size	0.101*	-0.051*	-0.083*	-0.000	-0.000	-0.293*	-0.010	-0.380*	1															
J Household income per capita	0.096*	-0.229*	-0.104*	-0.046*	0.002	-0.027*	0.018*	-0.408*	0.202*	1														
K Household income per capita squared	0.096*	-0.229*	-0.104*	-0.209*	-0.041*	-0.027*	0.059*	0.063*	-0.089*	-0.297*	1.000*													
L Net wealth per capita	0.129*	-0.200*	-0.120*	-0.183*	-0.081*	-0.010	0.059*	0.063*	-0.089*	-0.047*	0.193*	1												
M Net wealth per capita squared	0.129*	-0.200*	-0.120*	-0.183*	-0.081*	-0.010	0.123*	-0.123*	-0.000	-0.047*	0.193*	0.193*	1.000*											
N Pension	-0.170*	0.097*	0.140*	0.079*	0.026*	0.630*	0.123*	-0.123*	-0.000	-0.215*	0.014*	0.014*	0.046*	1										
O Probability of receiving inheritance	0.148*	-0.190*	-0.127*	-0.167*	-0.054*	-0.289*	-0.075*	0.056*	0.009	0.090*	0.117*	0.117*	0.046*	0.046*	1									
P Health spending from coverage	0.032	-0.157*	-0.081*	-0.175*	0.017*	0.022*	0.124*	-0.085*	-0.000	0.024*	0.241*	0.241*	0.113*	0.113*	-0.210*	1								
Q Out-of-pocket health spending	0.024*	-0.151*	-0.079*	-0.170*	0.017*	0.013*	-0.110*	-0.027*	0.024*	0.042*	0.232*	0.233*	0.093*	0.093*	-0.000	0.100*	1							
R Higher life expectancy	-0.048*	-0.105*	-0.026*	-0.064*	-0.010	0.013*	-0.147*	-0.061*	-0.036*	-0.011*	0.056*	0.056*	0.092*	0.09*	-0.010*	0.070*	0.975*	1						
S Future retirement age	-0.028*	-0.109*	-0.041*	-0.075*	0.017*	0.09	-0.104*	-0.048*	-0.010	-0.020*	0.100*	0.100*	0.034*	0.034*	-0.020*	0.030*	0.249*	0.169*	1					
T Gross pension replacement rate	0.104*	-0.286*	-0.211*	-0.258*	-0.078*	-0.231*	-0.190*	-0.019*	-0.019*	0.063*	0.110*	0.110*	0.120*	0.120*	-0.10*	0.151*	0.338*	0.299*	0.746*	1				
U Social activity	0.109*	-0.238*	-0.114*	-0.204*	-0.065*	-0.088*	0.066*	-0.097*	0.010	-0.000	0.146*	0.146*	0.156*	0.156*	-0.010	0.130*	0.096*	0.104*	0.071*	0.089*	1			
V Religious participation	0.016*	-0.031*	-0.031*	-0.032*	-0.000	0.050*	0.129*	-0.064*	0.020*	0.043*	-0.00	-0.00	0.009	0.009	0.020*	0.002	0.043*	0.085*	0.064*	0.0887	0.128*	1		
W Non-NHS	0.042*	-0.249*	-0.116*	-0.222*	2E-04	0.013*	-0.184*	-0.046*	-0.063*	-0.042*	0.347*	0.347*	0.152*	0.152*	-0.030*	0.162*	0.732*	0.742*	0.396*	0.528*	0.181*	0.130*	1	
X																					0.014*	0.057*	0.020*	1

Note: * is significance level at 0.05. Some of the control variables are highly correlated. For example, the correlation coefficient between age and SPH good is 0.23. This is likely to cause multi-collinearity problem during multivariate regression analyses. However, we checked the variance inflation factors (VIF). There are no VIF values greater than 5, indicating no multi-collinearity among the variables.

Table A3. Probit regression of purchasing life insurance by whole sample, Age 50+.

Dep. var.: Binary for Holding Life Insurance	The Whole Sample (N = 34,341)					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
SPH dummy	−0.0220 *** (0.005)			−0.0376 *** (0.006)		
OHS		−0.0292 * (0.017)			−0.0635 *** (0.021)	
FHR			−0.0136 *** (0.004)			−0.0144 *** (0.006)
Demographic variables						
Age ≥ 65	−0.128 *** (0.007)	−0.129 *** (0.007)	−0.129 *** (0.007)	−0.125 *** (0.007)	−0.126 *** (0.007)	−0.126 *** (0.007)
Higher education	0.0220 *** (0.005)	0.0232 *** (0.005)	0.0232 *** (0.005)	0.0193 *** (0.005)	0.0204 *** (0.005)	0.0204 *** (0.005)
Marital status (ref: Never married)						
Married	0.0565 *** (0.009)	0.0570 *** (0.009)	0.0571 *** (0.009)	0.0563 *** (0.009)	0.0568 *** (0.009)	0.0568 *** (0.009)
Divorced	0.014 (0.010)	0.0142 (0.010)	0.0143 (0.010)	0.013 (0.010)	0.0133 (0.010)	0.0133 (0.010)
Widowed	−0.00627 (0.010)	−0.00636 (0.010)	−0.00622 (0.010)	−0.00783 (0.010)	−0.00794 (0.010)	−0.00789 (0.010)
With children	0.0280 *** (0.008)	0.0285 *** (0.008)	0.0285 *** (0.008)	0.0285 *** (0.008)	0.0288 *** (0.008)	0.0287 *** (0.008)
Household income per capita	0.0226 *** (0.003)	0.0236 *** (0.003)	0.0235 *** (0.003)	0.0232 *** (0.003)	0.0242 *** (0.003)	0.0242 *** (0.003)
Household income per capita squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Net wealth per capita	0.0149 *** (0.001)	0.0152 *** (0.001)	0.0151 *** (0.001)	0.0152 *** (0.001)	0.0154 *** (0.001)	0.0154 *** (0.001)
Net wealth per capita squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Prob. of receiving inheritance	0.000633 *** (0.000)	0.000646 *** (0.000)	0.000644 *** (0.000)	0.000651 *** (0.000)	0.000665 *** (0.000)	0.000662 *** (0.000)
Pension	−0.00419 *** (0.001)	−0.00412 *** (0.001)	−0.00419 *** (0.001)	−0.00442 *** (0.001)	−0.00438 *** (0.001)	−0.00440 *** (0.001)
Future retirement age	−0.0165 *** (0.001)	−0.0165 *** (0.001)	−0.0167 *** (0.001)	−0.0153 *** (0.001)	−0.0153 *** (0.001)	−0.0155 *** (0.001)

Table A3. *Cont.*

Dep. var.: Binary for Holding Life Insurance	The Whole Sample (N = 34,341)					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Gross pension replacement rate	0.000472 ** (0.000)	0.000466 ** (0.000)	0.000484 ** (0.000)	0.000564 *** (0.000)	0.000578 *** (0.000)	0.000602 *** (0.000)
Health spending from coverage	0.0904 *** (0.007)	0.0906 *** (0.007)	0.0912 *** (0.007)	0.0858 *** (0.007)	0.0886 *** (0.007)	0.0903 *** (0.007)
Out-of-pocket health spending	−0.0936 *** (0.008)	−0.0937 *** (0.008)	−0.0941 *** (0.008)	−0.0870 *** (0.008)	−0.0888 *** (0.008)	−0.0901 *** (0.008)
Higher Life expectancy	0.0374 *** (0.005)	0.0402 *** (0.005)	0.0405 *** (0.005)	0.0382 *** (0.005)	0.0410 *** (0.005)	0.0415 *** (0.005)
Social activity	0.0405 *** (0.005)	0.0426 *** (0.005)	0.0424 *** (0.005)	0.0404 *** (0.005)	0.0426 *** (0.005)	0.0424 *** (0.005)
Religious participation	0.0277 *** (0.006)	0.0285 *** (0.006)	0.0283 *** (0.006)	0.0277 *** (0.006)	0.0282 *** (0.006)	0.0279 *** (0.006)
Advanced (ref: Emerging)	0.0272 *** (0.008)	0.0292 *** (0.008)	0.0298 *** (0.008)			
Non-NHS (ref: NHS)				0.00357 (0.006)	0.00906 (0.006)	0.0157 *** (0.006)
SPH dummy × Non-NHS				0.0362 *** (0.009)		
OHS × Non-NHS					0.0827 ** (0.033)	
FHR × Non-NHS						0.00145 (0.007)
Pseudo R²	0.101	0.101	0.101	0.102	0.102	0.101
Log likelihood	−15,865.41	−15,868.79	−15,869.03	−15,859.64	−15,873.96	−15,871.59

Note: This table reports mean marginal effects evaluated at each observation. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 for good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

Table A4. Probit regression of purchasing life insurance by different NHS, Age 50+.

Dep. var.: Life Insurance Holding	Non-NHS (N = 14,958)			NHS (N = 19,383)		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
SPH dummy	−0.00278 (0.008)			−0.0295 *** (0.006)		
OHS		0.0245 (0.029)			−0.0521 *** (0.020)	
FHR			−0.0160 *** (0.005)			−0.0117 ** (0.005)
Demographic variables						
Age ≥ 65	−0.153 *** (0.011)	−0.154 *** (0.011)	−0.151 *** (0.011)	−0.0923 *** (0.008)	−0.0938 *** (0.008)	−0.0956 *** (0.008)
Higher education	0.0128 * (0.007)	0.0134 * (0.007)	0.0124 * (0.007)	0.0242 *** (0.006)	0.0264 *** (0.006)	0.0271 *** (0.006)
Marital status (ref: Never married)						
Married	0.0531 *** (0.014)	0.0532 *** (0.014)	0.0532 *** (0.014)	0.0527 *** (0.012)	0.0536 *** (0.012)	0.0536 *** (0.012)
Divorced	0.0209 (0.016)	0.0207 (0.016)	0.0213 (0.016)	−0.00188 (0.014)	−0.00138 (0.014)	−0.00175 (0.014)
Widowed	−0.00816 (0.016)	−0.00869 (0.016)	−0.00726 (0.016)	−0.015 (0.013)	−0.0146 (0.013)	−0.0152 (0.013)
With children	0.0300 *** (0.011)	0.0302 *** (0.011)	0.0298 *** (0.011)	0.0302 *** (0.011)	0.0307 *** (0.011)	0.0308 *** (0.011)
Household income per capita	0.0174 *** (0.004)	0.0176 *** (0.004)	0.0174 *** (0.004)	0.0291 *** (0.003)	0.0299 *** (0.003)	0.0299 *** (0.003)
Household income per capita squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Net wealth per capita	0.0186 *** (0.001)	0.0187 *** (0.001)	0.0184 *** (0.001)	0.0162 *** (0.001)	0.0165 *** (0.001)	0.0166 *** (0.001)
Net wealth per capita squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Prob. of receiving inheritance	0.00054 *** (0.000)	0.00054 *** (0.000)	0.00054 *** (0.000)	0.000896 *** (0.000)	0.00092 *** (0.000)	0.00092 *** (0.000)
Pension	−0.0054 *** (0.001)	−0.0054 *** (0.001)	−0.0054 *** (0.001)	−0.00455 *** (0.001)	−0.0044 *** (0.001)	−0.00446 *** (0.001)
Future retirement age	−0.0200 *** (0.003)	−0.0199 *** (0.003)	−0.0204 *** (0.003)	0.00499 *** (0.002)	0.00474 ** (0.002)	0.00452 ** (0.002)
Gross pension replacement rate	0.00140 *** (0.000)	0.00141 *** (0.000)	0.00142 *** (0.000)	−0.00501 *** (0.000)	−0.00494 *** (0.000)	−0.00492 *** (0.000)

Table A4. *Cont.*

Dep. var.: Life Insurance Holding	Non-NHS (N = 14,958)			NHS (N = 19,383)		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Health spending from coverage	0.116 ***	0.115 ***	0.117 ***	0.007	0.012	0.0139
	(0.014)	(0.014)	(0.013)	(0.011)	(0.011)	(0.011)
Out-of-pocket health spending	−0.154 ***	−0.153 ***	−0.155 ***	0.0151	0.0112	0.00968
	(0.015)	(0.015)	(0.015)	(0.011)	(0.011)	(0.011)
Higher Life expectancy	0.0447 ***	0.0463 ***	0.0440 ***	0.0356 ***	0.0384 ***	0.0400 ***
	(0.009)	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)
Social activity	0.0273 ***	0.0279 ***	0.0268 ***	0.0647 ***	0.0680 ***	0.0683 ***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Religious participation	−0.0155	−0.0154	−0.0154	0.0328 ***	0.0340 ***	0.0336 ***
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Pseudo R^2	0.093	0.093	0.093	0.123	0.123	0.123
Log likelihood	−7516.404	−7516.1	7511.61	−8135.99	−8145.81	−8146.52

Note: Mean marginal effects evaluated at each observation. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 for good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

Table A5. Interaction terms with three health status indicators and higher education, whole sample, NHS countries, and Non-NHS countries, Age 50+.

Dep. var: Life Insurance Holding	The Whole Sample (N = 34,341)			Non-NHS (N = 14,958)			NHS (N = 19,383)		
SPH dummy	-0.0158 ** (0.007)			0.00203 (0.012)			-0.0235 ** (0.009)		
OHS		-0.0103 (0.024)			0.0377 (0.040)			-0.0281 (0.028)	
FHR			-0.0109 ** (0.005)			-0.0142 * (0.008)			-0.00833 (0.008)
Demographic variables									
Age ≥ 65	-0.128 *** (0.007)	-0.129 *** (0.007)	-0.129 *** (0.007)	-0.153 *** (0.011)	-0.154 *** (0.011)	-0.151 *** (0.011)	-0.0924 *** (0.008)	-0.0937 *** (0.008)	-0.0957 *** (0.008)
Higher education	0.0257 *** (0.006)	0.0265 *** (0.005)	0.0246 *** (0.005)	0.0152 * (0.009)	0.0154 * (0.009)	0.0134 * (0.008)	0.0284 *** (0.008)	0.0309 *** (0.007)	0.0287 *** (0.007)
SPH dummy × Higher education	-0.0101 (0.009)			-0.00832 (0.015)			-0.00926 (0.011)		
OHS × Higher education		-0.0358 (0.032)			-0.0253 (0.055)			-0.0448 (0.038)	
FHR × Higher education			-0.00505 (0.007)			-0.00339 (0.010)			-0.00591 (0.010)
Other controls	Marital status, with child, Household income per capita, Household income per capita squared, Net wealth per capita, Net wealth per capita squared, Probability of receiving inheritance, Pension, Future retirement age, Gross pension replacement rate, Health spending from coverage, Out-of-pocket health spending, Higher life expectancy, Social activity, Religious participation, and Country dummy								
Pseudo R^2	0.101	0.101	0.101	0.101	0.093	0.093	0.123	0.123	0.123
Log likelihood	-15,864.79	-15,874.13	-15,868.79	-7511.56	-7515.99	-7516.2	-8135.65	-8145.12	-8146.35

Note: Means of marginal effects evaluated at each observation. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

Table A6. Interaction terms with three health status indicators and Age ≥ 65, whole sample, NHS countries, and Non-NHS countries, Age 50+.

Dep. var: Life Insurance Holding	The Whole Sample (N = 34,341)			Non-NHS (N = 14,958)			NHS (N = 19,383)		
SPH dummy	-0.0223 *** (0.007)			-0.0121 (0.012)			-0.0192 ** (0.008)		
OHS		-0.0233 (0.030)			-0.00657 (0.053)			-0.0181 (0.035)	
FHR			-0.0258 *** (0.007)			-0.0371 *** (0.010)			-0.0137 (0.009)
Demographic variables									
Age ≥ 65	-0.128 *** (0.007)	-0.134 *** (0.007)	-0.134 *** (0.007)	-0.157 *** (0.012)	-0.157 *** (0.012)	-0.161 *** (0.011)	-0.0838 *** (0.009)	-0.0891 *** (0.009)	-0.0964 *** (0.009)
Higher education	0.0220 *** (0.005)	0.0232 *** (0.005)	0.0232 *** (0.005)	0.0131 * (0.007)	0.0133 * (0.007)	0.0124 * (0.007)	0.0240 *** (0.006)	0.0264 *** (0.006)	0.0271 *** (0.006)
SPH poor ×Age ≥ 65	0.000605 (0.009)			0.016 (0.015)			-0.0187 * (0.011)		
OHS × Age ≥ 65		-0.00845 (0.035)			0.0435 (0.062)			-0.0498 (0.042)	
FHR × Age ≥ 65			0.0177 ** (0.008)			0.0312 *** (0.011)			0.0028 (0.011)
Other controls	Marital status, with child, Household income per capita, Household income per capita squared, Net wealth per capita, Net wealth per capita squared, Probability of receiving inheritance, Pension, Future retirement age, Gross pension replacement rate Health spending from coverage, Out-of-pocket health spending, Higher life expectancy, Social activity, Religious participation, and Country dummy								
Pseudo R^2	0.101	0.101	0.101	0.093	0.093	0.093	0.123	0.123	0.123
Log likelihood	-15,865.4	-15,866.48	-15,874.73	-7515.85	-7515.85	-7507.71	-8134.48	-8145.09	-8146.49

Note: Means of marginal effects evaluated at each observation. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

Table A7. Marginal effects of purchasing different types of life insurance, whole sample, N = 34,341, Age 50+.

Dep. var.: Life Insurance Holding	Term	Whole	Both	No LI*	Term	Whole	Both	No LI*	Term	Whole	Both	No LI*	Term	Whole	Both	No LI*
SPH dummy	−0.00654 *** (0.002)	0.00203 (0.003)	−0.0184 *** (0.003)	0.0229 *** (0.005)										−0.0289 ** (0.012)		0.027 (0.017)
OHS					0.00182 (0.009)	0.000103 (0.012)										
FHR									−0.006 *** (0.002)	0.000728 (0.002)		0.0157 *** (0.004)		−0.0105 *** (0.003)	0.00788 ** (0.003)	−0.0248 *** (0.005)
Demographic variables																
Age ≥ 65	−0.0328 *** (0.003)	−0.0344 *** (0.005)	−0.0585 *** (0.005)	0.126 *** (0.007)	−0.0335 *** (0.003)	−0.0342 *** (0.005)	−0.0593 *** (0.005)	0.127 *** (0.009)	−0.0330 *** (0.003)	−0.0343 *** (0.005)	−0.0597 *** (0.005)	0.127 *** (0.007)	−0.0330 *** (0.003)	−0.0343 *** (0.005)	−0.0597 *** (0.005)	0.127 *** (0.007)
Higher education	0.00513 ** (0.002)	0.0115 *** (0.003)	0.00691 ** (0.003)	−0.0235 *** (0.005)	0.00569 *** (0.002)	0.0113 *** (0.003)	0.00782 ** (0.003)	−0.0249 *** (0.005)	0.00554 ** (0.002)	0.0114 *** (0.003)	0.00788 ** (0.003)	−0.0248 *** (0.005)	0.00554 ** (0.002)	0.0114 *** (0.003)	0.00788 ** (0.003)	−0.0248 *** (0.005)
Marital status (ref: Never married)																
Married	0.00909 ** (0.004)	0.0324 *** (0.006)	0.0160 ** (0.007)	−0.0575 *** (0.009)	0.00924 ** (0.004)	0.0325 *** (0.006)	0.0165 ** (0.007)	−0.0582 *** (0.009)	0.00927 ** (0.004)	0.0324 *** (0.006)	0.0166 ** (0.007)	−0.0583 *** (0.009)	0.00927 ** (0.004)	0.0324 *** (0.006)	0.0166 ** (0.007)	−0.0583 *** (0.009)
Divorced	0.0107 ** (0.005)	0.0148 ** (0.007)	−0.0117 * (0.007)	−0.0138 (0.010)	0.0108 ** (0.005)	0.0148 ** (0.007)	−0.0115 (0.007)	−0.0141 (0.010)	0.0108 ** (0.005)	0.0147 ** (0.007)	−0.0115 (0.007)	−0.014 (0.010)	0.0108 ** (0.005)	0.0147 ** (0.007)	−0.0115 (0.007)	−0.014 (0.010)
Widowed	0.00183 (0.005)	0.0114 * (0.007)	−0.0191 *** (0.007)	0.00581 (0.010)	0.00164 (0.005)	0.0115 * (0.007)	−0.0189 *** (0.007)	0.00578 (0.010)	0.00196 (0.005)	0.0115 * (0.007)	−0.0190 *** (0.007)	0.00556 (0.010)	0.00196 (0.005)	0.0115 * (0.007)	−0.0190 *** (0.007)	0.00556 (0.010)
With children	0.00680 ** (0.003)	−0.00344 (0.006)	0.0238 *** (0.005)	−0.0271 *** (0.008)	0.00698 ** (0.003)	−0.00356 (0.006)	0.0241 *** (0.005)	−0.0275 *** (0.008)	0.00696 ** (0.003)	−0.00357 (0.006)	0.0242 *** (0.005)	−0.0276 *** (0.008)	0.00696 ** (0.003)	−0.00357 (0.006)	0.0242 *** (0.005)	−0.0276 *** (0.008)
Household income per capita	0.00575 *** (0.001)	0.00520 *** (0.002)	0.0102 *** (0.002)	−0.0212 *** (0.003)	0.00604 *** (0.001)	0.00511 *** (0.002)	0.0111 *** (0.002)	−0.0222 *** (0.003)	0.00596 *** (0.001)	0.00513 *** (0.002)	0.0110 *** (0.002)	−0.0221 *** (0.003)	0.00596 *** (0.001)	0.00513 *** (0.002)	0.0110 *** (0.002)	−0.0221 *** (0.003)
Household income per capita squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Net wealth per capita	0.00468 *** (0.001)	0.00870 *** (0.001)	0.00242 *** (0.001)	−0.0158 *** (0.001)	0.00480 *** (0.001)	0.00865 *** (0.001)	0.00261 *** (0.001)	−0.0161 *** (0.001)	0.00475 *** (0.001)	0.00865 *** (0.001)	0.00258 *** (0.001)	−0.0160 *** (0.001)	0.00475 *** (0.001)	0.00865 *** (0.001)	0.00258 *** (0.001)	−0.0160 *** (0.001)
Net wealth per capita squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Prob. of receiving inheritance	0.00013 *** (0.000)	0.00029 *** (0.000)	0.00024 *** (0.000)	−0.0006 *** (0.000)	0.00014 *** (0.000)	0.00028 *** (0.000)	0.00024 *** (0.000)	−0.00067 *** (0.000)	0.00013 *** (0.000)	0.00028 *** (0.000)	0.00024 *** (0.000)	−0.0007 *** (0.000)	0.00013 *** (0.000)	0.00028 *** (0.000)	0.00024 *** (0.000)	−0.0007 *** (0.000)
Pension	−0.000757 ** (0.001)	−0.00147 *** (0.000)	−0.00220 *** (0.000)	0.00443 *** (0.001)	−0.000745 ** (0.000)	−0.00147 *** (0.000)	−0.00215 *** (0.000)	0.00436 *** (0.001)	−0.000757 ** (0.000)	−0.00146 *** (0.000)	−0.00221 *** (0.000)	0.0044 *** (0.001)	−0.000757 ** (0.000)	−0.00146 *** (0.000)	−0.00221 *** (0.000)	0.0044 *** (0.001)
Future retirement age	−0.00622 *** (0.001)	−0.00318 *** (0.001)	−0.00624 *** (0.001)	0.0156 *** (0.001)	−0.00624 *** (0.001)	−0.00317 *** (0.001)	−0.00624 *** (0.001)	0.0157 *** (0.001)	−0.00628 *** (0.001)	−0.00317 *** (0.001)	−0.00640 *** (0.001)	0.0159 *** (0.001)	−0.00628 *** (0.001)	−0.00317 *** (0.001)	−0.00640 *** (0.001)	0.0159 *** (0.001)
Gross pension replacement rate	0.00148 *** (0.000)	−0.00088 *** (0.000)	−0.000092 (0.000)	−0.00050 *** (0.000)	0.00148 *** (0.000)	−0.00088 *** (0.000)	−0.000098 (0.000)	−0.000495 ** (0.000)	0.00149 *** (0.000)	−0.000885 *** (0.000)	−0.0000812 (0.000)	−0.000520 *** (0.000)	0.00149 *** (0.000)	−0.000885 *** (0.000)	−0.0000812 (0.000)	−0.000520 *** (0.000)
Health spending from coverage	−0.0423 *** (0.004)	0.0640 *** (0.005)	0.0572 *** (0.005)	−0.0789 *** (0.008)	−0.0425 *** (0.004)	0.0640 *** (0.005)	0.0579 *** (0.005)	−0.0794 *** (0.008)	−0.0425 *** (0.004)	0.0641 *** (0.005)	0.0585 *** (0.005)	−0.0800 *** (0.008)	−0.0425 *** (0.004)	0.0641 *** (0.005)	0.0585 *** (0.005)	−0.0800 *** (0.008)
Out-of-pocket health spending	0.0267 *** (0.004)	−0.0335 *** (0.005)	−0.0780 *** (0.006)	0.0848 *** (0.008)	0.0268 *** (0.004)	−0.0334 *** (0.005)	−0.0784 *** (0.006)	0.0850 *** (0.008)	0.0268 *** (0.004)	−0.0335 *** (0.005)	−0.0788 *** (0.006)	0.0855 *** (0.008)	0.0268 *** (0.004)	−0.0335 *** (0.005)	−0.0788 *** (0.006)	0.0855 *** (0.008)
Higher Life expectancy	0.00629 ** (0.003)	0.0183 *** (0.003)	0.0123 *** (0.003)	−0.0369 *** (0.005)	0.00761 *** (0.003)	0.0180 *** (0.003)	0.0145 *** (0.003)	−0.0400 *** (0.005)	0.00716 *** (0.003)	0.0180 *** (0.003)	0.0150 *** (0.003)	−0.0402 *** (0.005)	0.00716 *** (0.003)	0.0180 *** (0.003)	0.0150 *** (0.003)	−0.0402 *** (0.005)
Social activity	0.0193 *** (0.002)	0.00757 *** (0.003)	0.0130 *** (0.003)	−0.0398 *** (0.005)	0.0201 *** (0.002)	0.00734 *** (0.003)	0.0148 *** (0.003)	−0.0422 *** (0.005)	0.0197 *** (0.002)	0.00735 ** (0.003)	0.0148 *** (0.003)	−0.0418 *** (0.005)	0.0197 *** (0.002)	0.00735 ** (0.003)	0.0148 *** (0.003)	−0.0418 *** (0.005)

Table A7. *Cont.*

Dep. var.: Life Insurance Holding	Term	Whole	Both	No LI *	Term	Whole	Both	No LI *	Term	Whole	Both	No LI *
Religious participation	0.00294	0.0184 ***	0.00332	−0.0247 ***	0.00308	0.0184 ***	0.00399	−0.0255 ***	0.00306	0.0184 ***	0.00383	−0.0253 ***
	(0.003)	(0.005)	(0.004)	(0.006)	(0.003)	(0.005)	(0.004)	(0.006)	(0.003)	(0.005)	(0.004)	(0.006)
Advanced markets (ref: Emerging markets)	0.0449 ***	−0.0267 ***	0.0144 ***	−0.0325 ***	0.0459 ***	−0.0269 ***	0.0156 ***	−0.0346 ***	0.0459 ***	−0.0269 ***	0.0161 ***	−0.0351 ***
	(0.003)	(0.006)	(0.005)	(0.008)	(0.003)	(0.006)	(0.005)	(0.008)	(0.003)	(0.006)	(0.005)	(0.008)
Log likelihood		−22,422.112				−22,439.99				−22,430.01		

Note: Mean marginal effects evaluated at each observation. ✴ No LI means households without life insurance. Monetary amounts are PPP-adjusted and in thousand Euros. The dummy SPH variable refers to SPH = 1 if poor health (indicating level 3, 4, or 5) and SPH = 0 if good health (indicating level 1 for very good, 2 for good). OHS variable is the determinant of current overall health as this variable not only considers the SPH status but also the numbers of chronic diseases. FHR evaluates the increasing function of the average number of risky behaviors (smoking, drinking, and a sedentary lifestyle), the household's Asymptomatic Objective Health Status (measured as the average number of diseases, blood pressure, blood cholesterol, and osteoporosis), and decreasing function of average household grip strength. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: the SHARE data and OECD.Stat (https://stats.oecd.org/Index.aspx).

References

1. Yaari, M.E. Uncertain lifetime, life insurance, and the theory of the consumer. *Rev. Econ. Stud.* **1965**, *32*, 137–150. [CrossRef]

2. Parrish, S. Permanent life insurance in financial planning: Time for a new model. *J. Financ. Serv. Prof.* **2014**, *68*, 30–32.

3. Fang, H.; Keane, M.P.; Silverman, D. Sources of Advantageous Selection: Evidence from the Medigap Insurance Market. *J. Political Econ.* **2008**, *116*, 303–350. [CrossRef]

4. Paccagnella, O.; Rebba, V.; Weber, G. Voluntary private health insurance among the over 50s in Europe. *Health Econ.* **2013**, *22*, 289–315. [CrossRef] [PubMed]

5. Buchmueller, T.C.; Fiebig, D.G.; Jones, G.; Savage, E. Preference heterogeneity and selection in private health insurance: The case of Australia. *J. Health Econ.* **2013**, *32*, 757–767. [CrossRef] [PubMed]

6. Su, B.; Thierry, G.H.; Chen, Q.; Zhao, Q. The new cooperative medical scheme and self-employment in rural China. *Sustainability* **2017**, *9*, 304. [CrossRef]

7. Atella, V.; Brunetti, M.; Maestas, N. Household portfolio choices, health status and health care systems: A cross-country analysis based on SHARE. *J. Bank. Financ.* **2012**, *36*, 1320–1335. [CrossRef] [PubMed]

8. Bressan, S.; Pace, N.; Pelizzon, L. Health status and portfolio choice: Is their relationship economically relevant? *Int. Rev. Financ. Anal.* **2014**, *32*, 109–122. [CrossRef]

9. Beck, T.; Webb, I. Economic, demographic, and institutional determinants of life insurance consumption across countries. *World Bank Econ. Rev.* **2003**, *17*, 51–88. [CrossRef]

10. Hwang, T.; Greenford, B. A cross-section analysis of the determinants of life insurance consumption in Mainland China, Hong Kong, and Taiwan. *Risk Manag. Insur. Rev.* **2005**, *8*, 103–125. [CrossRef]

11. Li, D.; Moshirian, F.; Nguyen, P.; Wee, T. The demand for life insurance in OECD countries. *J. Risk Insur.* **2007**, *74*, 637–652. [CrossRef]

12. Kjosevski, J. The determinants of life insurance demand in central and southeastern Europe. *Int. J. Econ. Financ.* **2012**, *4*, 237–247. [CrossRef]

13. Sauter, N.; Walliser, J.; Winter, J. Tax incentives, bequest motives, and the demand for life insurance: Evidence from a natural experiment in Germany. *J. Pension Econ. Financ.* **2015**, *14*, 525–553. [CrossRef]

14. Jappelli, T.; Pistaferri, L. Tax incentives and the demand for life insurance: Evidence from Italy. *J. Public Econ.* **2003**, *87*, 1779–1799. [CrossRef]

15. Fang, H.; Kung, E. *Why Do Life Insurance Policyholders Lapse? The Roles of Income, Health and Bequest Motive Shocks*; NBER Working Paper; National Bureau of Economic Research: Cambridge, MA, USA, 2012.

16. Lindström, M. Marital status, social capital, material conditions and self-rated health: A population-based study. *Health Policy* **2009**, *93*, 172–179. [CrossRef] [PubMed]

17. Vuong, Q.H.; Vuong, T.T.; Ho, T.M.; Nguyen, H.V. Psychological and socio-economic factors affecting social sustainability through impacts on perceived health care quality and public health: The case of Vietnam. *Sustainability* **2017**, *9*, 1456. [CrossRef]

18. Wu, S.; Wang, R.; Zhao, Y.; Ma, X.; Wu, M.; Yan, X.; He, J. The relationship between self-rated health and objective health status: A population-based study. *BMC Public Health* **2013**, *13*, 320. [CrossRef] [PubMed]

19. Goldman, D.; Maestas, N. Medical expenditure risk and household portfolio choice. *J. Appl. Econ.* **2013**, *28*, 527–550. [CrossRef] [PubMed]

20. Rosen, H.S.; Wu, S. Portfolio choice and health status. *J. Financ. Econ.* **2004**, *72*, 457–484. [CrossRef]

21. Edwards, R.D. Health risk and portfolio choice. *J. Bus. Econ. Stat.* **2008**, *26*, 472–485. [CrossRef]

22. Cardak, B.A.; Wilkins, R. The determinants of household risky asset holdings: Australian evidence on background risk and other factors. *J. Bank. Financ.* **2009**, *33*, 850–860. [CrossRef]

23. Fan, E.; Zhao, R. Health status and portfolio choice: Causality or heterogeneity? *J. Bank. Financ.* **2009**, *33*, 1079–1088. [CrossRef]

24. Christelis, D.; Jappelli, T.; Padula, M. Cognitive abilities and portfolio choice. *Eur. Econ. Rev.* **2010**, *54*, 18–38. [CrossRef]

25. Çelik, S.; Kayali, M.M. Determinants of demand for life insurance in European countries. *Probl. Perspect. Manag.* **2009**, *7*, 32–37.

26. Zietz, E.N. An examination of the demand for life insurance. *Risk Manag. Insur. Rev.* **2003**, *6*, 159–191. [CrossRef]

27. Inkmann, J.; Michaelides, A. Can the life insurance market provide evidence for a bequest motive? *J. Risk Insur.* **2012**, *79*, 671–695. [CrossRef]
28. Shi, X.; Wang, H.J.; Xing, C. The role of life insurance in an emerging economy: Human capital protection, assets allocation and social interaction. *J. Bank. Financ.* **2015**, *50*, 19–33. [CrossRef]
29. Ward, D.; Zurbruegg, R. Law, politics and life insurance consumption in Asia. *Geneva Pap. Risk Insur. Issues Pract.* **2002**, *27*, 395–412. [CrossRef]
30. Heo, W.; Grable, J.E.; Swarn, C. Life insurance consumption as a function of wealth change. *Financ. Serv. Rev.* **2013**, *22*, 389–404.
31. Sommer, M. *Savings Motives and the Effectiveness of Tax Incentives—An Analysis Based on the Demand for Life Insurance in Germany*; MEA Discussion Papers; Fakultät für Rechtswissenschaft und Volkswirtschaftslehre: Mannheim, Germany, 2007.
32. Andersson, L.F.; Eriksson, L. The compulsory public pension and the demand for life insurance: The case of Sweden, 1884–1914. *Econ. Hist. Rev.* **2015**, *68*, 244–263. [CrossRef]
33. Burnett, J.J.; Palmer, B.A. Examining life insurance ownership through demographic and psychographic characteristics. *J. Risk Insur.* **1984**, *51*, 453–467. [CrossRef]
34. Browne, M.J.; Kim, K. An international analysis of life insurance demand. *J. Risk Insur.* **1993**, *60*, 616–634. [CrossRef]
35. Loke, Y.J.; Goh, Y.Y. Purchase decision of life insurance policies among Malaysians. *Int. J. Soc. Sci. Hum.* **2012**, *2*, 415–420. [CrossRef]
36. Börsch-Supan, A.; Hank, K.; Jürges, H. A new comprehensive and international view on ageing: Introducing the 'Survey of Health, Ageing and Retirement in Europe'. *Eur. J. Ageing* **2005**, *2*, 245–253. [CrossRef] [PubMed]
37. Zhang, Y.H.; Li, C.S.; Liu, C.C.; Peng, S.C. The impact of health on work in China: A study using pilot survey data. *Geneva Pap. Risk Insur. Issues Pract.* **2013**, *38*, 857–870. [CrossRef]
38. Sprangers, M.A.; de Regt, E.B.; Andries, F.; van Agt, H.M.; Bijl, R.V.; de Boer, J.B.; Miedema, H.S. Which chronic conditions are associated with better or poorer quality of life? *J. Clin. Epidemiol.* **2000**, *53*, 895–907. [CrossRef]
39. Bliss, C.I. The method of probits. *Science* **1934**, *79*, 38–39. [CrossRef] [PubMed]
40. Barasinska, N.; Schäfer, D.; Stephan, A. Individual risk attitudes and the composition of financial portfolios: Evidence from German household portfolios. *Q. Rev. Econ. Financ.* **2012**, *52*, 1–14. [CrossRef]

Article

Does Sustainability Engagement Affect Stock Return Volatility? Evidence from the Chinese Financial Market

Junru Zhang, Hadrian Geri Djajadikerta and Zhaoyong Zhang *

School of Business and Law, Edith Cowan University, Joondalup WA 6027, Australia; j.zhang@ecu.edu.au (J.Z.); h.djajadikerta@ecu.edu.au (H.G.D.)

* Correspondence: zhaoyong.zhang@ecu.edu.au

Abstract: This paper examines the impact of firms' sustainability engagement on their stock returns and volatility by employing the EGARCH and FIGARCH models using data from the major financial firms listed in the Chinese stock market. We find evidence of a positive association between sustainability engagement and stock returns, suggesting firms' sustainability news release in favour of the market. Although volatility persistence can largely be explained by news flows, the results show that sustainability news release has the significant and largest drop in volatility persistence, followed by popularity in Google search engine and the general news. Sustainability news release is found to affect positively stock return volatility. We also find evidence that market expectation can be driven by the dominant social paradigm when sustainability is included. These findings have important implications for market efficiency and effective portfolio management decisions.

Keywords: corporate sustainability; news release; stakeholder theory; stock return volatility; EGARCH-m

1. Introduction

The growing international focus on corporate sustainability and social responsibility has triggered a trend toward requiring firms to engage corporate sustainable practices. However, the increase in attention from both industry and academia has led to an increasing number of studies on the association between sustainability engagement and firm risk (for instance, Feldman et al. [1]). According to Godfrey [2], sustainability engagement can provide insurance-like protection to preserve financial performance by generating moral capital and goodwill in the long term. Sustainability engagement can help reduce corporate adverse cash flow [3] and the cost of capital [4], as well as increase efficiency in waste reduction [5], control long-term risk, and refine long-term risk management [6]. Additionally, the existing literature shows that sustainability engagement can improve market confidence, thereby reducing stock market return risk through volatility. For instance, Harjoto and Jo [7] found that legalized sustainability exposure due to governmental requirements is in favour of the market because the information is more likely to be genuine and less costly to access. Theodoulidis et al. [8] indicated that sustainability engagement information increases market confidence as long-run-oriented business strategy eliminates stock market speculators. However, when a company discloses information about sustainability engagement, there are two general arguments among the existing studies about how the market interprets the information. First, the optimistic view highlights the importance of sustainability engagement in reducing stock market return volatility: even if sustainability information does not increase the present value of a company, it can potentially maximise future value by stretching the existing wealth creation of a business [9]. The commitment to sustainability can be captured by market participants over time, in which sustainability strategies and engagement can increase company

performance in the long run, thereby increasing firm value and market confidence and reducing stock return volatility [4]. However, sustainability information can be influenced by information asymmetry in a semi-strong efficient market, causing investors to have dispersed opinion of the released sustainability information [7]. Market interpretation of sustainability information might lead to higher volatility in stock price return, because information released can be viewed as a strategy for management intrinsic values [10], thereby causing a bubble in stock prices [11].

In recent years the Chinese government has undertaken a series of initiatives and procedures encouraging firms to release sustainability reports. The two Chinese stock exchanges have also taken a leading role in requiring all listed companies to engage in environmental, social and governance (ESG) reporting since 2006. This study intends to examine the impact of such sustainability engagement through news releases on firms' stock return volatility, using data from the Chinese stock market. To form the basis of the theory behind this study, we consider two main views of sustainability and their association with stock return volatility; one is the information asymmetry view of sustainability associated with Crane [12] and Orlitzky [11], who suggest that the heterogeneous definitions of corporate social responsibility (CSR) can only exacerbate the problem of information asymmetry, in which case sustainability news is not different from other general news such as noise. The other view we consider is from Godfrey [2] and Jo and Na [10], where sustainability can reduce stock return volatility by providing shareholders with insurance-like protection for relationship-based intangible assets. By considering the background in China, we adopted the stakeholder theory from Freeman [13] with a view that powerful stakeholders are able to influence the dominant social paradigm, and therefore can alter the expectations of the market and drive corporate activities towards their expectations. We relate this study to the above arguments and believe they are not mutually exclusive, given a market that is stable in political turnover, where investors are influenced by the dominant social paradigm. In this context, social forces, sustainability dynamics and information can improve market confidence where there is information asymmetry. We consider the following theoretical grounds in this study. First, information asymmetry in a semi-strong market leads to information dispersal among investors, and in general bad news causes more volatility than good news. We argue that such an effect from information asymmetry can be reduced if a society is influenced substantively by its dominant social paradigm. Second, we relate the risk-reduction approach to the stakeholder theory to posit a negative association between sustainability engagement and stock return volatility.

Recently there have been studies on the relationship between news sentiment and changes in asset dynamics [14–17]. In particular, Riordan et al. [16] argue that, compared with positive messages, negative newswire messages are particularly informative and have a more significant impact on high-frequency asset price discovery and liquidity. Ho et al. [14] examined the dynamic relationship between firm-level return volatility and public news sentiment. Ho et al. [15] examined the impact of public information flows on the volatility of the bilateral Chinese RMB-U.S. dollar exchange rates in the spot, non-deliverable forward (NDF) and futures markets. The purpose of this study is to examine the impact from sustainability engagement information on firm stock return volatility, and to contrast the risk reduction effect against other types of news. In particular, to assess the impact of sustainability information on stock return volatility this study adopts the EGARCH variance-in-mean model to examine the association between sustainability engagement information and return volatility using data from the Chinese stock market. We find evidence of a positive association between sustainability engagement and stock returns. The results show that firms with higher frequencies of sustainability news release are associated with higher stock return, suggesting that the market takes firms' sustainability news release positively. This contrasts with the findings on market response to other types of general news releases. Sustainability news releases can reduce the volatility persistence and positively affect the return volatility. This study makes three major contributions. First, it provides new evidence in support of the stakeholder theory under the theoretical framework of market efficiency and information asymmetry. Second, it contributes to the existing

debate on whether the market responds asymmetrically to sustainability news releases. Using data from the Chinese stock market, we find evidence that investors take firms' sustainability engagement information as a positive indicator of strong and healthy performance in the future. Finally, this study is among the first to collect and use daily news releases fitting into the concept of corporate sustainability and to document the dynamic effect of firms' sustainability engagement information on stock return volatility with robust results. These findings have important implications for the efficiency of the Chinese stock market and investors' effective portfolio investment decisions.

The remainder of the paper is structured as follows. Section 2 provides a brief literature review and proposes several hypotheses, followed by discussions on methodology and data in Section 3. In Section 4, we discuss the empirical results. The last section concludes the paper.

2. Literature Review and Empirical Predictions

2.1. Economic Consideration of Corporate Sustainability

The concept of sustainability is derived from a normative concept, which occupies a continuum where it is understood quantitatively through an economic dimension, and qualitatively through a development dimension [18]. Economic-oriented considerations focus on the monetary terms and looking at economic growth related to corporate activities (i.e., whether sustainability adds financial benefits to shareholders); more precisely, sustainability adjudicates to stretch this wealth creation in the long run. In the modern economic model, sustainability is initially mentioned by DesJardins [19], but it is considered to constrain firms in profit maximization. This concept is then further extended to development, and the extension considers generating financial values without creating excessive environmental and social damage and aims to ensure that the firm is using natural resources without producing waste that exceeds the capacity at the expense of the ecological system [19]. In other words, there needs to be a balance between the growth of the economy and the development towards sustainability. Due to the unique institutional background in China, the perception of the social actors on corporate sustainability can be largely influenced by the dominant social paradigm, thereby incorporating sustainability into long-run business strategy and performance [18]. Corporate sustainability activities are assessed by whether they are translated into the long term through market-worthiness, creating reputation or maintaining legitimacy within the local community [20]. Companies' product chains and the associated natural resources inputs and outputs are severely influenced because they are framed by the dominant social paradigm in this institutional background [21].

In addition, the dominant social paradigm can shape and lead social action towards sustainability even if the movements do not fit perfectly into the theoretical concept of sustainability and generate short-term economic benefits [22]. An example is the valuation of social and human capital [23]. Although companies are required to report the social and human (labour) resources occupied in business operations, the legal requirements are largely linked to monetary terms that can hardly reflect the welfare situation of the reporting companies. For product responsibility, assessment of lifecycle is also an issue for many emerging economies, especially in China, which intensively uses resources obtained for low prices and produces exchange goods at higher prices. Due to the existence of these problems, the social actors (in this case, the reporting companies) are likely to focus solely on the dominant social paradigm and alter their understanding through disclosure about sustainability. The information exposure and corporate actions from the reporting companies are likely to be constrained by releasing information solely on the norms described in the paradigm, which may not often be related to corporate sustainability but what is required and expected by the government and political legitimacy. Similarly, the stakeholders may link information materiality with the extent to which companies have met and fulfilled the social paradigm in this context. In this study, we propose that the Chinese governing bodies are at a dominant level in the social paradigm, and market participants and corporate sustainability actions are framed by a certain higher-level social group

in the economy. Given that the reporting requirements are derived from a political perspective (or considerations) on corporate sustainability, we posit that the market's expectation for material sustainability information is led by the dominant social paradigm in China.

2.2. Institutional Background

To address public concerns about environmental and social issues that have arisen both nationally and internationally, the Chinese government has undertaken a series of controls and initiatives to strengthen the concept of corporate sustainable development. In China, the notion of corporate sustainability reporting was first proposed in 2006 with the amendment of the Company Law of the People's Republic of China, Article 5 of the General Law. Later, in 2006, in the Chinese Communist Party's Sixth Plenary Session, the creation of a harmonious society was proposed with the focus of being socially responsible, particularly for business enterprises [24]. As a response to the national plan, both the Shanghai Stock Exchange (SHSE) and the Shenzhen Stock Exchange (SZSE) issued social reporting guidelines in 2006 and 2008, respectively, to create an appropriate system for corporate sustainability reporting. Although policies were introduced after the two stock exchanges announced their reporting guidelines, the meaning and definition of corporate sustainability were not clearly specified; in particular, there was no clear indication of how and what to report in a corporate sustainability disclosure [25]. To address further public concerns on the transparency of corporate sustainability information, in 2008, both stock exchanges began setting mandatory corporate sustainability disclosures for a subset of listed companies, including the top 100 companies in the SZSE 100 index, companies in the SHSE Corporate Governance index firms listed in overseas stock markets, and all financial firms [26]. These requirements are important milestones in promoting and adopting sustainability reporting standards in China, and they are consistent with the goals and objectives of the government for building a harmonious socialist society by 2020. In the recent 19th CPC National Congress, President Xi specifically emphasized in his report the issues with the monotonous economic structure and the high carbon dependency in China. The Chinese government has drawn up a new blueprint in its national plan, vowing that "through the efforts of the concept of harmonious society, we will firmly establish a road of sustainable development with Chinese characteristics of conservation, recycling, low carbon, ecology, and environmental friendliness" [27]. A series of books and learning guides were published after the Congress, aiming to provide interpretation of the new measures and policies to be implemented by the government. Based on the series of policies and initiatives bonded with the national approach to sustainability, in this study we posit that the Chinese government includes corporate sustainability in the dominant social paradigm that influences social actors' decision-making.

2.3. Information Asymmetry and Stakeholder Theory

Even though the aim of sustainability is to generate a long-term business direction with less speculative behaviour, information asymmetry can cause sustainability information to become noise in an equity market. Corporate sustainability information is largely voluntary in most cases, so regulators may not have a strict approach to hard disclosure [28]. The flexibility in the voluntary form of sustainability information can be used as a manipulation of legitimacy by opportunistic management due to the misrepresentation of enforceable public accountability [29,30]. Investors can experience high information asymmetry in a market that is not fully efficient [7]. Studies have found that information misinterpretation can cause a reduction in stock price return and firm value. Under uncertain circumstances, investors can have a divergence of opinion about share prices, which leads to higher risk through return volatility and lower return [31]. Such divergence was later investigated by Grossman and Stiglitz [32], who found that the extent of disagreement is associated with the costs of information, e.g., information quality, information case, investment noise in risky assets and the number of investors involved. The study indicates that the scale of information can have a negative impact on the degree of divergence. Participants in the market are either unable or unwilling to screen out noise caused

by information asymmetry. Another counter-sustainability argument indicates that investors are speculative so that they deliberately choose not to filter out noise from sustainability information, in which CSR actions are manipulation-prone for certain short-run company economic benefits [11].

However, the problem caused by information asymmetry can be reduced if the dominant social paradigm is able to provide market efficiency. Based on stakeholder theory, corporate sustainability not only relates to its shareholders/debt holders, but also to any relevant stakeholders [33]. The normative stakeholder theory asserts that "regardless of whether stakeholder management leads to improved financial performance, managers should manage the business for the benefit of all stakeholders" [34] (p. 32); however, from the positive branch, the more salient the stakeholder, the more efforts will be exerted in terms of satisfying their needs [35]. One major facet of stakeholder theory involves recognising and identifying the association between the behaviour of a company and its impact on company stakeholders [36]. Freeman [13] indicated that there are two types of stakeholders who are influential to companies. The primary stakeholders have the control of scarce resources that a company is dependent on to survive. The other type is defined as secondary stakeholders, who have less control over companies' decision-making. Under stakeholder theory, management will choose to meet the expectations of the primary stakeholders as a priority because the secondary stakeholders have limited resources that companies are dependent on. In the context of China, the government has great influence on companies' business activities and strategies, which has a significant impact on investment performance and financial returns [37,38]. The government policies are often used as directions for corporate investment [39]. Hence, we argue that since China has adopted corporate sustainability into the dominant social paradigm, companies are likely to engage in corporate sustainability in order to meet the expectations of the primary stakeholders. Being sustainable means that firms take a long-term view of their business actions rather than being speculative about short-term performance [26]; hence, the market is expected to respond favourably to sustainability news releases as they signal to investors that corporate decisions from the reporting firms are likely to link sound sustainable strategy with expected performance outcomes. Firms that integrate sustainability initiatives with their business strategies are more likely to gain a good reputation for social and environmental issues, which helps to mitigate risk. Vast existing research shows that sustainability-engaged firms are more likely to perform better financially in the long run [40–44]. Provided the market is efficient and the investors are rational, sustainability-engaged firms are more likely to perform well in the long run, which subsequently increases market confidence. As a consequence, it also reduces the potential speculative behaviour because, in the short term, sustainability engagement is costly and is at the expense of the shareholders [26]. Thus, it is believed that such "socially responsible" firms can help build shareholders' investment confidence and attract more funding at a lower cost of capital than "less responsible" firms [4]. Based on these arguments, we state the following hypothesis:

Hypothesis 1 (H1). *Sustainability engagement increases market confidence and reduces speculation, thereby having a positive effect on stock return and decreasing return volatility.*

There are some very important assumptions for our first hypothesis. First, whether sustainability is considered in the social paradigm is crucial in this study. In Section 2.2, we have listed several significant initiatives by the Chinese government towards sustainability. These include the 2006 National People's Congress, the 2006 Chinese Company Law, the 2006 SSE Corporate Environmental Responsibility Reporting Guidelines, the 2008 SZSE Corporate Social Responsibility Reporting Guidelines, and the 19th CPC National Congress. Also, another assumption is the existence of a dominant social paradigm in China; it has been shown by numerous existing studies that the government has a substantial influence on corporate business activities and corporate decision-making [26,37,45,46]. When we contrast the impact of other information on market participants due to information asymmetry, we expect that the market will respond differently to sustainability news and general news such that the impact on return volatility will also be different.

Hypothesis 2 (H2). *According to information asymmetry, market participants react more volatilely to negative than positive general news due to their potentially speculative behaviour in comparison with the response to sustainability news.*

To test these hypotheses, in particular the dynamic relationship between sustainability engagement news and stock return volatility, we consider the generalized autoregressive conditional heteroskedasticity (GARCH) family model in this study. We first adopt the popular exponential GARCH (or EGARCH) specification for its ability to capture the most important stylized characteristics of volatility series, including asymmetry and leverage effects. We then consider the fractionally integrated generalized autoregressive conditional heteroscedastic (FIGARCH) models due to their long memory for examining persistence in stock return volatility. There are several existing studies that adopt volatility to determine the association with sustainability (see, for instance, Harjoto and Jo [7]; Jo and Na [10]; Becchetti et al. [47]), but few on the dynamics of conditional variance with daily stock return. This study is among the first to use daily sustainability news frequency as a proxy for corporate sustainability engagement to assess the impact on stock return volatility. For comparison purposes, we also use companies' general news frequency and Google search frequency to assess the impact.

3. Sample and Methodology

3.1. Sample and Return Series

This study examines the relationship between stock return volatility and sustainability news release of the listed financial firms in China. We obtained the daily stock price samples of the listed financial firms and their financial report information from the Bloomberg database. The sample period was 24 December 2007 to 21 March 2018. For sample selection, the initial sample included all the financial companies listed in both the Shanghai and the Shenzhen Stock Exchanges. We then excluded companies listed after 2008 (to ensure correct data on sustainability news releases) and those that were de-listed during the sample period. In addition, we also excluded listed companies that were relatively small in terms of total assets and market capitalization. This is because studies have found that sustainability engagements are positively associated with large firms both in terms of total assets [48] and market capitalisation [49]. As a result, our dataset consists of 30 listed financial companies with 80,190 daily observations in total, ranging from 1 December 2007 to 31 March 2018. The lists of our sample companies are shown in Appendixs A and B (Tables A1 and A2).

To mitigate the effect of size bias, we adopted a weighted approach to determine the stock return of the financial firms listed in the Chinese stock markets. The weight was assigned to each firm based on its market capitalization. This approach is supported by the stakeholder theory, which states that large entities are more likely to be influenced by stakeholder media [50], where instruments in communication are used to increase resources and corporate influences, leading to higher sensitivity to media exposure [51]. Let S_{it} denote firm i's stock price at time t; WI_{it} be the weight at time t, and Market cap_{it} be firm i's market capitalisation at time t. The stock return, r_{it}, is calculated as follows:

$$WI_{it} = \frac{Market\ cap_{it}}{\sum Market\ cap_{it}} + 1$$
$$r_{it} = WI_{it} \times 100 \times \log\left(\frac{S_{i,t}}{S_{i,t-1}}\right). \tag{1}$$

3.2. Sustainability News Information Arrival

The concept of sustainability development or corporate sustainability has been defined in various ways. The most commonly cited formulation for sustainability is provided by the World Commission on Environment and Development, that sustainable development shall "meet the needs of the present generation without compromising the ability of future generations to meet their own needs" [52] (p. 43). The interpretation of this definition is often related to human welfare and well-being, which should be

sustained in the long run without raising concerns about impacts on future generations [53]. In this study, we apply these definitions of corporate sustainability to the case of China, defining the major sustainability themes promoted by the Chinese government as the indicative measures of corporate sustainability. Based on the long-term aims of sustainability and the concept of corporate sustainable development in the Chinese context, we define corporate sustainable news as information from news releases about a company's development that fits into the scale of the sustainability themes promoted by the government (environmental protection and social responsibility). To ensure that sustainability information is genuine, the news release must specify that the sustainability activities are associated with and have been integrated into the company's business strategies and operations.

To collect sustainability news information, we have used Bloomberg News, Baidu News search engine and Google News search engine in the data collection process. In particular, we first collected news from Baidu News and Google News search engines by using a series of keywords that fit into the concept of corporate sustainability. The key words were developed based on a series of themes under 'sustainable development' that were recently promoted by the 19th National Congress in China, which covers the main issues in environmental protections and social responsibility. Because news about our sample companies is often released in Chinese and English, we used both languages in our keyword search for sustainability news. These keywords used in Baidu News and Google news include social responsibility (社会责任), labour (劳工), human rights (人权), product responsibility (产品责任), economy (经济), environment (环境), ecology (生态), nature (自然), green (绿色), emissions reduction (减排), energy saving (节能), and environmental protection (环保). Then, we collected and manually processed all the keyword-related news based on firm, time and relevance to determine the frequency of each sample firm's daily news associated with its sustainability engagement. In order to differentiate sustainability news and other types of company news, we applied the following two criteria: First, the news must be related to a company's sustainability practices that are directly associated with the government's sustainable development policies. Second, the news must mention that such activities are a part of the company's sustainability business strategy, so that the sustainability practices reported in the news are reflections of sustainability through business operations (e.g., banks may apply screening when introducing products to promote customers' business ideas towards 'new energy'). If the news fits into the criteria, we record its frequency. It is also worth noting that our news frequency dataset does not include news reproduced or reprinted by other media to avoid double counting of the same news from the original media reporter.

Apart from Baidu News and Google News, we also used the Bloomberg News database. We first searched for Company News, then collected the sustainability news under the ESG (environment, social and governance) category based on our previous criteria, and finally recorded the number of daily news items associated with corporate sustainability. Similarly, we excluded news reproduced or reprinted by other media to avoid counting it twice. Consequently, the final sustainability news proxy includes the sum of the amount of news from all three databases. Again, we used the weighted sustainability news frequency (*WSNF*) in this study to accommodate the size effect. It is defined as:

$$WSNF_{it} = WI_{it} \times SNF_{it}, \tag{2}$$

where SNF_{it} is the sustainability news frequency for firm i, and WI_{it} is the weight for firm i at time t.

3.3. Other News Information Arrival

To compare the news' influence on stock return volatility, we divided all news into sustainability news and other news information, which includes a company's general news and general public news of interest to the company. Companies' general daily news frequencies (*GNF*) were collected from the Bloomberg News database, which contains more than 1000 different news sources globally and over 90,000 web sources and social media. The major media include Dow Jones Newswires and the Wall Street Journal, as well as the central news agencies in China, e.g., The Central News

Agency, Xinhua News Agency, China News Agency, and ENET Communication Agency. Due to the comprehensiveness of the Bloomberg database, the frequencies of companies' general news were directly collected from Bloomberg News Trend.

Recently, Internet usage has emerged as an important source of information for public opinion about a company's performance. Traditionally, people's interest in a topic is collected through surveys to measure the extent of awareness or support from the public for a decision made by a company [54]. However, as search engine technology develops, Google has become a reliable and valuable resource for people to obtain information, and it has become the most popular search engine in the world [55]. In this paper, to compare the sustainability news impact, we constructed a Google Trends Frequency index (*GTF*) as a proxy of public interest in a company's general business activities. It measures the popularity of a company in the Internet search, and ranges from zero to 100 based on the frequency of a search item relative to its previous-day 'popularity'. *GTF* is adopted as another proxy for news information arrival given its representation of people's interest about a topic or theme [56]. In particular, we entered a sample company's name in the search engine in English, Simplified Chinese and Traditional Chinese to collect all the search results, then processed the news and finally calculated the final *GTF* index for each sample company accordingly.

A weighted index was also adopted in the calculation of GNF_{it} and GTF_{it}. At time t, the weighted *GNF* and *GTF* are:

$$WGNF_{it} = WI_{it} \times \log(GNF_{it})$$
$$WGTF_{it} = WI_{it} \times \log(GTF_{it}). \tag{3}$$

3.4. Methodology and Model Specification

To analyse the relationship between sustainability engagement news and stock return volatility over time, we considered the generalized autoregressive conditional heteroskedasticity (GARCH) family model in this study. Since the autoregressive conditional heteroskedasticity approach (ARCH) was first proposed by Engle [57], many significant theoretical and empirical developments have emerged in the literature [58–63]. GARCH family models have enjoyed popularity among academics because of their ability to capture some of the typical stylized facts of financial time series, such as volatility clustering [64], and also to take into account the feature of volatility over a long period of time and provide good in-sample estimates [65,66]. The symmetric univariate GARCH model originally proposed by Bollerslev [67] has been extended to incorporate various kinds of features, such as asymmetries, long memory persistence, and regime switches [63,68–70]. McAleer [71] reviews a wide range of models of financial volatility, univariate and multivariate, conditional and stochastic, and McAleer and Medeiros [70] discuss recent developments in modelling univariate asymmetric volatility.

The exponential GARCH (EGARCH) model proposed by Nelson [72] has become one of the two most widely estimated univariate asymmetric conditional volatility models for its ability to capture asymmetry and (possible) leverage [73,74]. Given that EGARCH is a discrete-time approximation to a continuous-time stochastic volatility process in logarithms, conditional volatility is guaranteed to be positive, but the model requires parametric restrictions to ensure that it can capture the (possible) leverage [73]. McAleer and Hafner [74] showed that EGARCH could be derived from a random coefficient complex nonlinear moving average (RCCNMA) process. Chang and McAleer [75] further derive the regularity conditions for asymmetry in EGARCH to show that, in practice, EGARCH always displays asymmetry, though not leverage. In order to testify the impact of condition volatility on stock return, a conditional variance term is added to the mean equation in the constructed EGARCH variance-in-mean model. In this study, we modify the EGARCH variance-in-mean model by adding the

sustainability news variable into the mean equation in order to examine the hypothetical associations between stock return, volatility and sustainability news release [15], specified as follows:

$$r_t = b_0 + b_\lambda h_t + b_1 News_t + \varepsilon_t$$
$$\varepsilon_t = \eta_t \sqrt{h_t}, \; \eta_t \overset{iid}{\sim} t(0, 1, v)$$
$$\log(h_t) = \omega + \alpha \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \beta \log(h_{t-1}) + \gamma \left\{ \left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| - E\left(\left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| \right) \right\},$$

(4)

where r_t is the daily return of stock price. $News_t$ stands for the weighted daily sustainability news frequency ($WSNF_t$), the weighted daily general news frequency ($WGNF_t$), and the weighted Google Trends frequency ($WGTF_t$), respectively, and ε_t is the standard error at time. h_t is the conditional volatility of ε_t at time t. η_t is the standardised residual of ε_t with zero mean, one-unit standardisation in student-t distribution, where the degree of freedom is v. Lastly, in the variance equation, the coefficient β captures the degree of volatility persistence that measures how quickly the present shock dissipates. EGARCH (1,1) is covariance stationary if $\beta < 1$; however, a relatively greater value in β implicates the present shock will influence volatility in the long run [72]. Asymmetry exists for EGARCH if $\alpha \neq 0$, while the leverage effect exists if $\alpha < 0$ and $\alpha < \gamma < -\alpha$ [74]. In the benchmark case that no news impact is examined, we should remove the news variable from the mean equation.

To address our research concerns and also serve as a robustness check of the EGARCH results, we then continued the study by employing the fractionally integrated GARCH (FIGARCH) to test the memory of the volatility of stock return. FIGARCH is based on the application of the fractional differencing operator to the autoregressive structure of the conditional variance by assuming that it follows a hyperbolic rather than exponential decay [76]. Extended from the family of GARCH models, Baillie et al. [76] proposed the FIGARCH model, which provides additional features for volatility clustering with good in-sample estimates [65,66]. Chang et al. [77] suggest that the FIGARCH(1,d,1) model outperforms its GARCH(1,1) counterpart (see also Ho et al. [14]). Since the introduction of the model, many significant empirical studies on long memory have emerged in the existing literature [14,78–81]. In this study, we adopt FIGARCH to investigate the long-term memory in the conditional volatility of the stock return and how volatility persistence is affected by a firm's sustainability news releases and other firm-specific general news releases. We modify the FIGARCH(1,d,1) model by including the news variables as follows:

$$r_t = \mu + b_1 News_t + \varepsilon_t, \text{ where } \varepsilon_t = \eta_t \sqrt{h_t},$$
$$b(L) h_t = \omega + \left[b(L) - \varnothing(L)(1 - L)^d \right] \varepsilon_t^2,$$
$$b(L) = 1 - b_1 L, \text{ and } \varnothing(L) = 1 - \varnothing_1 L,$$

(5)

where ε_t is the error at time t, h_t is the conditional volatility of ε_t at time t, η_t is an identical and independent sequence following a specific distribution, L is the lag operator, $(1 - L)^d$ is the fractional differencing operator, and d is the long-memory parameter. $News_t$ stands for the weighted daily sustainability news frequency ($WSNF_t$), the weighted daily general news frequency ($WGNF_t$), and the weighted Google Trends frequency ($WGTF_t$), respectively. The stationary long memory process for volatility is assessed through the parameter, d, which lies between 0 and 1. The FIGARCH model offers higher flexibility when modelling conditional variables due to the nests of covariance stationary GARCH when $d = 0$, where in the integrated GARCH (IGARCH), $d = 1$. The IGARCH process seems to be too restrictive as it implies infinite persistence of a volatility shock and in most of the empirical situations the volatility process is found to be mean-reverting [76]. Under the FIGARCH model, the persistence of shocks to the conditional variance, which is also referred to as a long memory process of persistence, is captured by a fractional differencing parameter d with a range from 0 to 1. When $d = 1$, a unit root is subjected, and it shows a permanent shock effect similarly to the IGARCH model; whereas when $d = 0$, an ordinary GARCH process ensures that no long-memory persistence is involved [76]. The FIGARCH model implies a slower hyperbolic rate of decay for a lagged shock in

the conditional variance equation so that the fractional differencing parameter provides important information about the pattern and speed with which shocks to volatility are propagated, which implies that the effect of a volatility shock is mean-reverting and is also quite persistent.

4. Data Analysis and Implications

4.1. Descriptive Statistics

Table 1 presents the descriptive statistics of the stock returns and news variables. Panel A reports the descriptive statistics of the daily stock returns of the financial companies. The mean return among the 30 sample firms is less than 0, and the median of 0 confirms the negative average return. The standard deviations of the returns are 0.8487 and 0.6449, indicating that they are considerably volatile. Both stock price returns are leptokurtic, indicating potential higher peak clustering, where the skewness of stock price return is slightly negative. The prices and daily returns of these stocks are presented in Figure 1. It can be seen that both stock prices and returns are more volatile during the Global Financial Crisis (GFC) period in 2008–2009 and also during China's stock market crash in the period 2015–2016. The stock prices are relatively flat during 2011–2015 (with the exception of 2013), as are the stock returns.

Table 1. Descriptive statistics of daily data.

Var.	Mean	Std. Dev	Median	Min.	Max.	Skew.	Kuro.	Obs.
			Panel A: Descriptive statistics of stock price return					
r_t	−0.0018	0.8487	0.0000	−4.7958	4.2516	−0.1240	7.4307	2673
			Panel B: Descriptive statistics of news frequency					
SNF_t	0.1349	0.1646	0.1000	0.0000	2.0000	3.0131	20.3135	2673
GNF_t	5.9828	4.6825	4.6500	0.0000	32.7500	1.4506	5.5448	2673
GTF_t	47.7899	15.3058	49.3681	9.0526	88.9474	−0.2663	2.3256	2673
			Panel C: Descriptive statistics of news variables					
$WSNF_t$	0.1468	0.1796	0.1034	0.0000	2.1997	3.0795	21.2409	2673
$WGNF_t$	1.4495	0.9289	1.5369	−2.9957	3.4889	−1.0312	5.2408	2673
$WGTF_t$	3.8585	0.3744	3.9530	2.2450	4.5300	−1.0619	3.7932	2673

Note: Var. stands for variables. Std. Dev stands for standard deviation. Min. stands for minimum. Max. stands for maximum. Skew. stands for skewness. Kuro. stands for Kurtosis and Obs. stands for the number of observations. r_t is the daily return of the sample financial firms. GNF_t is companies' general news frequency, and $WGNF_t$ is weighted general news frequency. SNF_t is sustainability news frequency, including the frequency on news particularly regarding corporate sustainability. $WSNF_t$ is weighted sustainability news frequency. GTF_t is Google Trends frequency, indicating a popularity search index in Google about a company. $WGTF_t$ is weighted Google Trends Frequency.

The descriptive statistics of all the news variables are shown in Panels B and C of Table 1. The mean of GNF_t is 5.9828, with a range between 0 and 32.75. As for the weighted variable, the standard deviation of $WGNF_t$ drops substantially to 0.9289, and the skewness changes from positive to slightly negative. For SNF_t, the mean is relatively low at 0.1349 compared to other general news, and the variable is significantly positively skewed, even using the weighted approach. GTF_t has a mean of 47.7899, which indicates that sample companies are generally popular in a Google search. When we consider the weighted approach, the standard deviation of the variable is reduced to 0.3744, in which a negative skewness is observed.
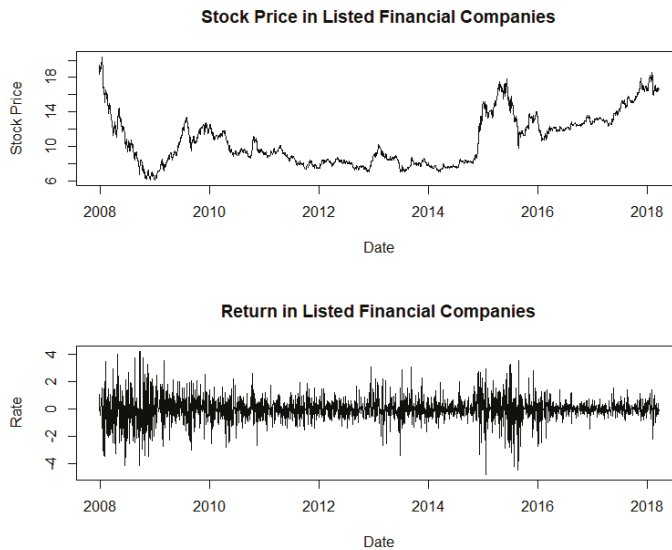
Figure 1. Daily stock price in financial companies and returns.

The frequencies of news releases of the three news variables are plotted in Figure 2. For sustainability news releases, the frequency was generally low during the GFC in 2008–2009, which may suggest that companies are less engaged in sustainability during the crisis periods. This was further shown in mid-2015 and 2016 during the Chinese stock market crisis, when the trends became lower and showed a flat pattern. Furthermore, the sustainability news frequency fluctuated extensively from 2010 to 2015. It is also noted that the peaks in sustainability news releases show some correspondence with the peaks in the stock prices particularly around 2013, 2014, 2016, and 2017. This preliminary visual observation indicates that there might be a positive association between sustainability news releases and firm stock prices. For GNF_t, the number of general news releases was generally low during the GFC period in 2008–2009, and a number of peaks were formed after the second half of 2010. Interestingly, the amount of company general news decreased extensively after the Chinese stock market crisis in 2016 when the stock price increased, while the amount of general news increased and fluctuated substantially between 2011 and 2015 when the stock price decreased. This observation may indicate that the financial firms are more inclined to release news when the stock price is low, whereas less news is released when the stock price is relatively high. Regarding Google searches, the overall trend is considerably stable during the sample period, with the exception of several notable drops. The frequency changes of Google searches show a similar pattern to the stock prices, suggesting that investors and business stakeholders or even the general public are more likely to use Google searches for company news when their stock prices are low.
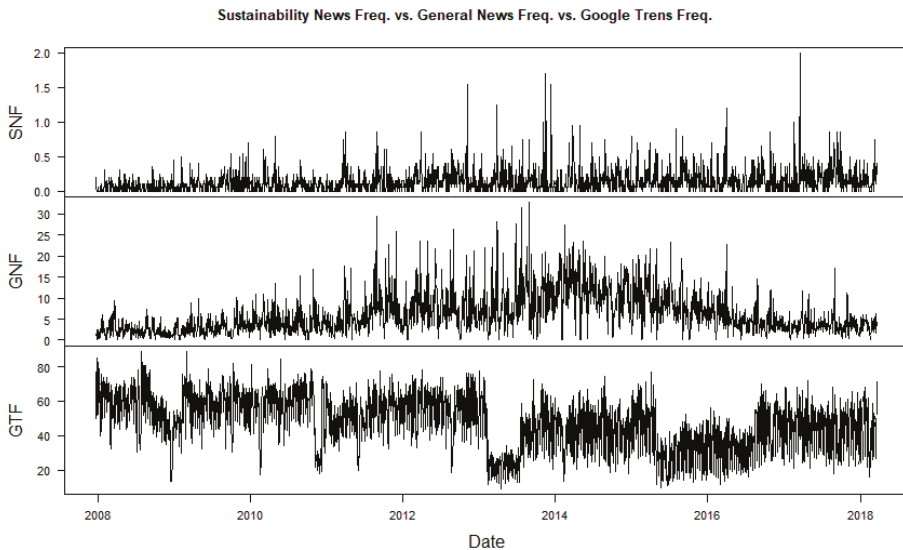
Sustainability News Freq. vs. General News Freq. vs. Google Trens Freq.

Figure 2. The number of sustainability news vs. general news vs. Google searches.

Given that the stock price returns are more volatile, especially during the GFC period in 2008–2009 and post-GFC in 2013, and the news variables do not show a distinctive pattern during the sample period, we follow Ho et al. [15] to assess if any distinctive patterns in the news series can be observed during the identified calm or turbulent state in stock returns by adopting a moving average window approach. It is arguable that volatility in a calm state is relatively smaller due to policies and regulation controls, whereas return in a turbulent state is more volatile. To investigate the relationship between the news variables and the stock return, we set the length of the moving window to 100 and calculate the mean of the first 100 observations. We then continue the process to calculate the mean of the 2nd to 101st observations and so on, until the mean of the last 100 observations is calculated. We plot the moving average of the news series and stock return in Figure 3, which will allow us to determine the different states based on the turning points. More specifically, we call the periods in which the moving average continues to increase the turbulent state, and the rest the calm state. It can be seen in Figure 3 that there is no steady pattern among the variables, with the exceptions of *GNF/SNF* and the return between 2008 and 2010. In contrast, the relationship between *GTF* and return shows a substantial fall during the same period. Hence, this study is less likely to be affected by policy switching and/or structural breaks. In addition, notably in Figure 3, the moving window shows a negative association between returns and *SNF* during the 2008–2009 GFC and the 2015–2016 financial market crisis in China. A possible explanation is that, during the crisis period, the market is more interested in information relating to firms' financial performance than to their sustainable practices; hence, the market is less reactive to sustainability disclosure [82]. This may also indicate that the market is more likely to have a diverse interpretation of information that is costly to access and to determine its genuineness [7]. As such, the market could not indicate whether sustainability engagement is truthful or bluffing. It may subsequently lead to different responses among the market participants, thereby increasing the volatility of the stock returns. It is also interesting to note that the extent of the negative association between return and *SNF* is relatively higher in the 2008–2009 GFC period than during the 2015–2016 crisis. This finding is consistent with our casual observation that sustainability in China has become a mega-trend, where investors and business stakeholders are more aware of sustainability issues and firms have become more critically observed by society, which has put more pressure on them to act in

sustainable ways and prioritize sustainability engagement in their business operations. This pattern can also be observed when we compare the trend for the moving window between *GNF* and return, and that for *GTF* and return during the two crisis periods.
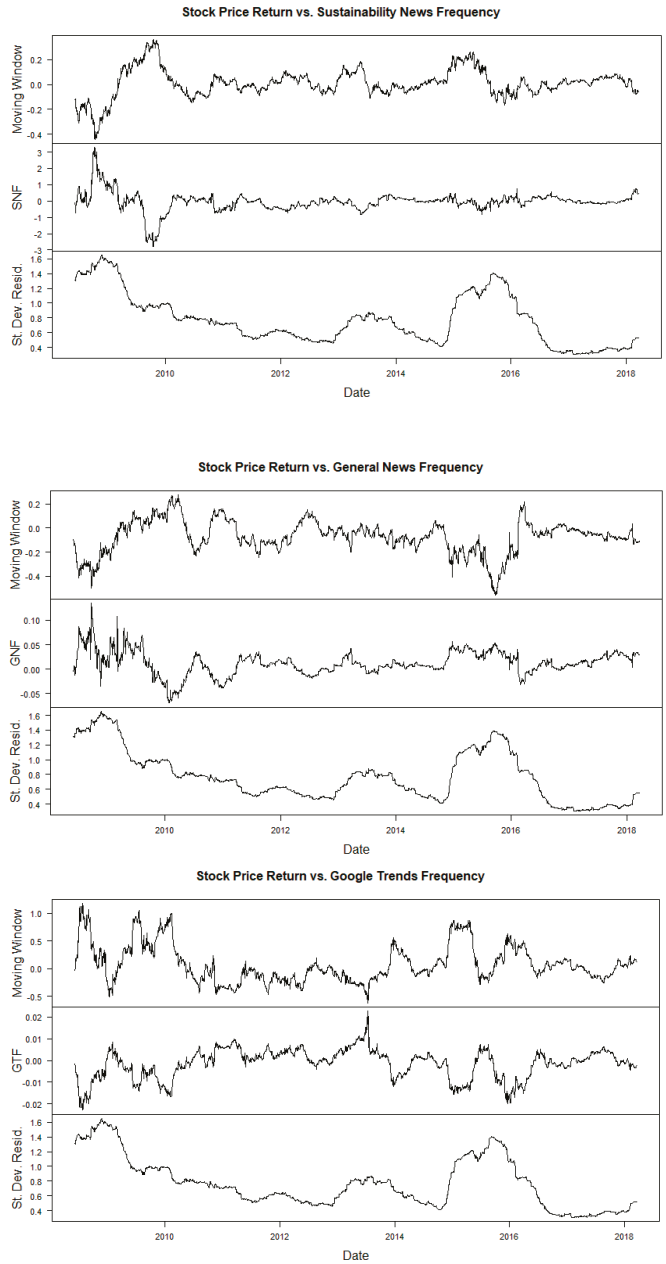

Figure 3. Moving average window of stock price return vs. news variables.

It is also interesting to note that the relationship between stock returns and sustainability news release varies across the different states. Both are relatively less correlated in the turbulent state and highly correlated during the calm state. This pattern may be associated with investors' short-term opportunistic behaviour and the government's direct market intervention, especially during the turbulent state. On the other hand, the observed distinctive pattern between stock returns and sustainability news flows since 2011 becomes more visible, which may be associated with the dominant social paradigm since then. In contrast, the correlations between firm-specific general news and popularity in Google searches and the stock return are less noticeable.

4.2. The EGARCH-M Framework

We begin with an estimation of the benchmark model and then the full model to assess the news impact on stock return and volatility. The optimal ARMA orders are determined based on the Akaike Information Criterion (AIC). The estimation of the parameters of the model was undertaken through the conditional maximum likelihood method. The stock price returns are firstly fitted into the benchmark model (without news variable) and then the full model with news variable. The news variable in the full model is proxied by the sustainability news variable, the general news variable, and the Google search-based variable, respectively. Table 2 reports the results.

Table 2. Estimation results of the EGARCH-m model.

Parameter	Sustainability News (WSNF)	General News (WGNF)	Google Search (WGTF)	Benchmark Model
b_0 (constant)	0.0558 ($p = 0.0000$)	0.5667 ($p = 0.0000$)	4.0442 ($p = 0.0000$)	0.0209 ($p = 0.1098$)
b_λ	0.1743 ($p = 0.0000$)	0.7709 ($p = 0.0000$)	0.5595 ($p = 0.0000$)	−0.0156 ($p = 0.4209$)
b_1	2.9819 ($p = 0.0000$)	1.7560 ($p = 0.0000$)	2.2062 ($p = 0.0000$)	−
ω(constant)	−0.1231 ($p = 0.0000$)	−0.2691 ($p = 0.0000$)	−0.5206 ($p = 0.0000$)	0.0070 ($p = 0.0005$)
α	0.1601 ($p = 0.0000$)	−0.1747 ($p = 0.0000$)	−0.3292 ($p = 0.0000$)	−0.0044 ($p = 0.0000$)
β	0.9679 ($p = 0.0000$)	0.6388 ($p = 0.0000$)	0.8177 ($p = 0.0000$)	0.9934 ($p = 0.0000$)
γ	−0.0530 ($p = 0.0000$)	0.0273 ($p = 0.0982$)	−0.1063 ($p = 0.0000$)	0.1284 ($p = 0.0000$)
Log lik.	1081.085	5567.7940	116.5570	−2881.932
AIC	−0.8094	2.0984	0.0488	2.1616

Note: The table shows estimation results of the following EGARCH Variance-in-mean (1,1) models. Log lik. is log likelihood. AIC is Akaike Information Criterion.

As shown in Table 2, the estimates of all the news variables are significantly positive. In comparison with the benchmark models, the log likelihood values are remarkably improved in the cases for the full models, and the values of the AIC also suggest that models with news variables, especially sustainability news, are preferred. More specifically, the results indicate that a sustainability news release can significantly increase stock returns. With the inclusion of the sustainability news variable, the persistence of the stock return volatility has been reduced to 0.9679 from 0.9934, as in the case with the benchmark model. The results seem to suggest that the market responds favourably to firms' sustainability news releases, as the news signals to the investors that firms are engaged in sustainable strategies. The finding of a significant positive relationship between sustainability news and stock returns suggests that public exposure of a company's sustainability engagement through news releases can increase shareholders' investment confidence and help mitigate stock return volatility. This is consistent with our Hypothesis 1, which predicts a positive relationship between sustainability engagement, market confidence, and stock returns. The results lend support to the findings of Mehran [83] and Jenter and Kanaan [84], and confirm that sustainability engagement-related information will increase market confidence and have impacts on value creation as sustainability-integrated firms are viewed as being more likely to care about creating long-lasting financial success by implementing sustainability in their strategy. The results also support the argument that information asymmetry can be overcome by the dominant social paradigm if sustainability has been included. The results are consistent with existing studies where information about sustainability engagement is negatively associated with firm risk through volatility [3,10]. This is especially the

case in China given its unique institutional features. As business activities are often influenced and to some extent driven by government policies [85], it is often observed that market performance is very sensitive to the direction changes of government policies. As such, with the recent promotion of "sustainable development" by the Chinese government, there would be a transition from the dominant social paradigm to the embodiment of the "sustaincentric paradigm" [86], a result of which would be a strong link between firms' sustainability engagement and stock returns as the market believes that sustainability-engaged firms will have high future returns. Furthermore, as sustainability engagement aims to boost wealth creation in the long term, decisions on sustainability engagement will strengthen asset performance in the future [18], which will lead to less speculative investment in the market and also help reduce return volatility. Therefore, we should accept our Hypothesis 1 given the evidence that news releases on corporate sustainability engagement increase market confidence, positively affect stock returns, and help reduce stock return volatility.

Similar results are found for the other two types of news variables. These findings confirm that news releases add explanatory power to the variance of stock returns, and are also consistent with the existing literature. The asymmetric effect is captured by α. As can be seen in Table 2, all the estimates for α are significantly different from zero, confirming the existence of an asymmetric effect. In the sustainability news model, α is significantly positive at 0.1601, and negative for the rest of the models, suggesting the presence of leverage effects in the models with general news and Google search news. As a special case of asymmetry, leverage captures the negative correlation between return shocks and subsequent shocks to volatility [74,87]. Our results provide further evidence that the stock return volatility is affected differently by positive and negative general news and Google search news, but is less likely to be affected by the sustainability news. This finding is consistent with existing studies on information asymmetry, where, in a semi-efficient market, investors are more responsive to negative news due to the management's speculative behaviour [11]. It also has important implications for firms' news release decisions as the market responds to different types of news differently. Under the notion of sustainable development, firms that engage sustainability outperform those less sustainable firms financially in the long run [4]. As such, the market participants would link sustainability engagement with corporate financial stability in the long run, which helps reduce speculative behaviour in the market, but it is not observed for the general news release. We find evidence that, due to the presence of asymmetric and leverage effects, the release of firm-specific general news and being 'popular' in a Google search may not always be favourable because negative news will cause increases in stock return volatility. In contrast, such leverage effect does not exist for sustainability news releases. In addition, sustainability news releases affect stock returns favourably, and the log likelihood values for the model with sustainability news variables are also remarkably improved compared to the benchmark model. This finding is consistent with stakeholder theory, and also with our Hypothesis 2 that market reaction is more volatile to negative than to positive general news in comparison with the response to sustainability news.

4.3. The FIGARCH Framework

To address our research concerns, and also as a robustness check for the EGARCH results, we continue this study by employing the FIGARCH framework to test the memory of the volatility of stock return. We fit our data into the FIGARCH specifications and reported the estimation results in Table 3. As can be seen in Table 3, the model without news variable has the largest estimate for the long memory parameter d. When news variables are included, the volatility persistence is reduced to below 0.5. The results further suggest that volatility persistence is largely explained by news flows. However, the impact of different types of news on the persistence is different. It is found that sustainability news releases cause the largest drop in volatility persistence, followed by Google search engine and then general news. This finding is consistent with our early conclusion and supports Hypothesis 2.

Table 3. Estimation results of the FIGARCH (1, d, 1) model.

Parameter	Sustainability News (WSNF)	General News (WGNF)	Google Search (WGTF)	Benchmark Model
M (constant)	0.0066 ($p = 0.0000$)	0.0055 ($p = 0.0000$)	0.0092 ($p = 0.0000$)	0.0224 ($p = 0.1447$)
b1	0.0233 ($p = 0.0000$)	−0.0014 ($p = 0.0000$)	−0.0107 ($p = 0.0000$)	—
ω (constant)	0.0014 ($p = 0.0000$)	0.0009 ($p = 0.0000$)	−0.0001 ($p = 0.1679$)	0.0022 ($p = 0.0082$)
α	0.0423 ($p = 0.0000$)	0.0353 ($p = 0.0000$)	0.0449 ($p = 0.0000$)	0.0000 ($p = 1.0000$)
β	0.9008 ($p = 0.0000$)	0.9039 ($p = 0.0000$)	0.9002 ($p = 0.0000$)	0.9395 ($p = 0.0000$)
d	0.4014 ($p = 0.0000$)	0.4398 ($p = 0.0982$)	0.4016 ($p = 0.0000$)	0.7007 ($p = 0.0000$)
Log lik.	−4019.846	−4036.078	−4235.052	−2884.184
AIC	3.0145	3.0266	3.1755	2.1640

This table presents the summary output for FIGARCH (1, d, 1) model fitted with normal distribution. *d* is a factional differencing parameter. For the explanation of the variables, please see Table 1.

Thus, it has been demonstrated that news releases relating to corporate sustainability engagement will affect stock return positively and reduce the volatility persistence. The results are consistent with and support the stakeholder theory.

5. Conclusions

In this study we have empirically examined the impact of sustainability engagement information on financial firms' stock returns and volatility by employing the EGARCH-M and FIGARCH models using data from the Chinese stock market. We posit that market participants are less likely to face information asymmetry when they deal with sustainability engagement news, as sustainability is adopted in the dominant social paradigm in the context of China. Information asymmetry exists due to the cost, accessibility and reliability of the information. We hypothesize that the promotion of sustainable development by the Chinese government increases information credibility towards corporate sustainability engagement, and relevant news can effectively increase market confidence, thereby reducing market dispersion through stock return volatility. Using a sample of the Chinese listed financial firms during 2007–2018, we find evidence of a positive association between sustainability engagement and stock returns, which is consistent with our hypotheses. The results show that firms with higher frequencies of sustainability news releases are associated with higher stock returns, suggesting that the market looks favourably on firms' sustainability news releases and sustainability engagement increases market confident and reduces stock return volatility. This contrasts with the market response to other types of news flows such as firm-specific general news releases and popularity in Google search engine. It is also found that stock return volatility is influenced by all types of news, and the volatility persistence is substantially reduced after incorporating news variables, suggesting that volatility persistence is mostly explained by news flows. In particular, sustainability news releases have a greater impact on volatility persistence than the other two types of news releases, and there is no evidence indicating the presence of a leverage effect for sustainability news releases. This finding is consistent with our hypothesis of an asymmetric market reaction to positive and negative news associated with firm-specific general news releases and popularity news in Google search engine in contrast to the response to sustainability news.

The results also support the argument that information asymmetry can be overcome by the dominant social paradigm if sustainability has been included. This finding is consistent with our casual observation that, due to its unique institutional features, the Chinese government's recent promotion of "sustainable development" has caused a transition from the dominant social paradigm towards the embodiment of the "sustaincentric paradigm" [86], which would lead to a strong link between firms' sustainability engagement and stock returns. This will lead to less speculative investment in the market and hence to low return volatility. These findings imply that the market expectation can be driven by a social paradigm, given that sustainability is considered as a priority task by the government, and hence it has important implications for market efficiency and effective portfolio investment decisions.

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

Appendix A

Table A1. Firm size of sample companies (total assets in millions of RMB).

Date	Stock Code									
	000001	**002142**	**600000**	**600015**	**600016**	**600036**	**601009**	**601169**	**601328**	**601601**
2007	352.54	75.511	914.98	592.34	918.83	1310.96	76.06	354.22	2110.44	322.34
2008	474.44	103.26	1309.43	731.63	1054.35	1571.79	93.70	417.02	2556.41	317.89
2009	587.81	163.35	1622.72	845.45	1426.39	2067.94	149.56	533.46	3309.13	397.18
2010	727.21	263.27	2191.41	1040.23	1823.73	2402.50	221.49	733.21	3951.59	475.71
2011	1258.18	260.50	2684.69	1244.14	2229.06	2794.97	281.79	956.49	4611.17	570.61
2012	1606.54	372.70	3145.71	1488.86	3212.00	3408.09	343.79	1119.96	5273.37	681.50
2013	1891.74	462.19	3680.13	1672.44	3226.21	4016.39	434.05	1336.76	5960.93	723.53
2014	2186.46	554.11	4195.92	1851.62	4015.13	4731.82	573.15	1524.43	6268.29	825.10
2015	2507.145	716.46	5044.35	2020.60	4520.68	5474.97	805.02	1844.90	7155.36	923.84
2016	2953.43	885.02	5857.26	2356.23	5895.87	5942.31	1063.90	2116.33	8403.16	1020.69
	601166	**601939**	**601988**	**601998**	**601398**	**601628**	**600837**	**601318**	**600030**	**600369**
2007	851.34	6598.18	5991.22	1011.18	8683.71	933.70	95.34	692.22	189.65	-
2008	1020.90	7555.45	6951.68	1319.57	9757.14	987.49	74.68	704.56	206.80	8.02
2009	1332.16	9623.36	8751.94	1775.03	11,785.05	1226.25	120.73	935.71	153.17	14.97
2010	1849.67	10,810.32	10,459.87	2081.31	13,458.62	1410.57	115.41	1171.62	148.28	22.77
2011	2408.80	12,281.83	11,829.79	2765.88	15,476.86	1583.90	98.97	2285.42	168.50	17.76
2012	3250.98	13,972.83	12,680.62	2959.93	17,542.21	1898.91	126.48	2844.26	271.35	17.25
2013	3678.30	15,363.21	13,874.30	3641.19	18,917.75	1972.94	169.12	3360.31	479.62	29.99
2014	4406.40	16,744.09	15,251.38	4138.81	20,609.95	2246.56	352.62	4005.91	616.10	58.20
2015	5298.88	18,349.49	16,815.60	5122.29	22,209.78	2448.31	576.44	4765.15	597.43	71.74
2016	6085.90	20,963.71	18,148.89	5931.05	24,137.26	2696.95	560.86	5576.90	189.65	70.99

Appendix B

Table A2. Firm size of sample companies (market capitalisation in millions of RMB).

Date	Stock Code									
	000001	**002142**	**600000**	**600015**	**600016**	**600036**	**601009**	**601169**	**601328**	**601601**
2007	80.76	-	229.94	80.47	214.58	556.59	35.081	126.79	639.48	380.77
2008	2.94	17.00	75.01	36.28	76.61	180.14	15.410	55.49	236.46	85.62
2009	75.70	43.73	191.52	61.98	175.30	344.65	35.54	12.04	425.43	220.80
2010	55.03	35.76	177.78	54.40	136.66	291.35	29.51	71.24	339.02	207.31
2011	79.87	26.42	158.37	76.92	155.57	259.50	27.55	57.79	275.04	162.24
2012	82.08	30.74	185.04	70.89	219.184	296.70	27.31	81.84	358.38	205.41
2013	100.42	26.62	175.90	76.31	213.213	283.85	24.02	66.10	300.21	182.35
2014	180.97	51.12	292.67	119.86	351.531	413.75	43.49	115.42	469.83	290.60
2015	171.56	60.49	340.80	129.72	329.453	441.44	59.59	133.44	412.80	255.54
2016	156.26	64.89	350.43	115.94	317.773	437.77	65.68	148.42	402.27	241.77
2017	228.37	90.29	369.54	115.40	293.156	717.57	65.65	151.17	368.74	375.35
	601166	**601939**	**601988**	**601998**	**601398**	**601628**	**600837**	**601318**	**600030**	**600369**
2007	259.30	1468.28	1443.64	328.35	2471.79	1485.46	225.94	706.49	295.95	2.22
2008	73.00	875.47	670.58	131.74	1186.85	542.69	66.727	211.77	119.15	1.73
2009	201.55	1375.34	1051.10	291.58	1836.22	911.17	157.89	416.88	210.65	36.15
2010	144.12	1464.68	922.21	192.81	1537.88	643.84	79.32	479.34	125.22	27.01
2011	135.05	1100.50	764.92	181.65	1436.39	483.21	60.97	294.82	107.75	20.04
2012	180.03	1247.06	803.68	193.08	1474.73	597.07	98.91	380.17	150.02	20.74
2013	193.19	1137.55	745.87	172.38	1302.25	455.91	107.33	369.40	144.90	23.06
2014	314.36	1290.58	1143.75	333.81	1682.45	892.52	218.02	593.40	361.07	62.91
2015	325.22	1124.09	1087.37	292.89	1574.85	745.55	167.12	658.05	224.93	55.89
2016	307.50	1336.77	982.49	283.94	1550.20	636.17	168.04	642.42	190.14	40.25
2017	352.95	14,499.25	1104.18	199.71	2126.29	577.94	136.35	1238.572	219.32	26.14

References

1. Feldman, S.J.; Soyka, P.A.; Ameer, P.G. Does improving a firm's environmental management system and environmental performance result in a higher stock price? *J. Investig.* **1997**, *6*, 87–97. [CrossRef]
2. Godfrey, P.C. The relationship between corporate philanthropy and shareholder wealth: A risk management perspective. *Acad. Manag. Rev.* **2005**, *30*, 777–798. [CrossRef]
3. Sharfman, M.P.; Fernando, C.S. Environmental risk management and the cost of capital. *Strateg. Manag. J.* **2008**, *29*, 569–592. [CrossRef]
4. Dhaliwal, D.S.; Li, O.Z.; Tsang, A.; Yang, Y.G. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *Account. Rev.* **2011**, *86*, 59–100. [CrossRef]
5. Ratnatunga, J.; Jones, S. An inconvenient truth about accounting: The paradigm shift required in carbon emissions reporting and assurance. In *Contemporary Issues in Sustainability Accounting, Assurance and Reporting*; Emerald Group Publishing: Bradford, UK, 2012; pp. 71–114.
6. Dhaliwal, D.S.; Radhakrishnan, S.; Tsang, A.; Yang, Y.G. Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. *Account. Rev.* **2012**, *87*, 723–759. [CrossRef]
7. Harjoto, M.A.; Jo, H. Legal vs. Normative CSR: Differential Impact on Analyst Dispersion, Stock Return Volatility, Cost of Capital, and Firm Value. *J. Bus. Ethics* **2015**, *128*, 1–20. [CrossRef]
8. Theodoulidis, B.; Diaz, D.; Crotto, F.; Rancati, E. Exploring corporate social responsibility and financial performance through stakeholder theory in the tourism industries. *Tour. Manag.* **2017**, *62*, 173–188. [CrossRef]
9. Mackey, A.; Mackey, T.B.; Barney, J.B. Corporate social responsibility and firm performance: Investor preferences and corporate strategies. *Acad. Manag. Rev.* **2007**, *32*, 817–835. [CrossRef]
10. Jo, H.; Na, H. Does CSR Reduce Firm Risk? Evidence from Controversial Industry Sectors. *J. Bus. Ethics* **2012**, *110*, 441–456. [CrossRef]
11. Orlitzky, M. Corporate social responsibility, noise, and stock market volatility. *Acad. Manag. Perspect.* **2013**, *27*, 238–254. [CrossRef]
12. Crane, A. *The Oxford Handbook of Corporate Social Responsibility*; Oxford Handbooks: Oxford, UK, 2008.
13. Freeman, R.E. The politics of stakeholder theory: Some future directions. *Bus. Ethics Q.* **1994**, *4*, 409–421. [CrossRef]
14. Ho, K.-Y.; Shi, Y.; Zhang, Z. How does news sentiment impact asset volatility? Evidence from long memory and regime-switching approaches. *N. Am. J. Econ. Financ.* **2013**, *26*, 436–456. [CrossRef]
15. Ho, K.-Y.; Shi, Y.; Zhang, Z. Does news matter in China's foreign exchange market? Chinese RMB volatility and public information arrivals. *Int. Rev. Econ. Financ.* **2017**, *52*, 302–321. [CrossRef]
16. Riordan, R.; Storkenmaier, A.; Wagener, M.; Zhang, S.S. Public information arrival: Price discovery and liquidity in electronic limit order markets. *J. Bank. Financ.* **2013**, *37*, 1148–1159. [CrossRef]
17. Tetlock, P.C. Does public financial news resolve asymmetric information? *Rev. Financ. Stud.* **2010**, *23*, 3520–3557. [CrossRef]
18. Lu, Y.; Abeysekera, I. What do stakeholders care about? Investigating corporate social and environmental disclosure in China. *J. Bus. Ethics* **2017**, *144*, 169–184. [CrossRef]
19. DesJardins, J. Corporate environmental responsibility. *J. Bus. Ethics* **1998**, *17*, 825–838. [CrossRef]
20. Chow, W.S.; Chen, Y. Corporate sustainable development: Testing a new scale based on the mainland Chinese context. *J. Bus. Ethics* **2012**, *105*, 519–533. [CrossRef]
21. Korhonen, J. The dominant economics paradigm and corporate social responsibility. *Corp. Soc. Responsib. Environ. Manag.* **2002**, *9*, 66–79. [CrossRef]
22. Korhonen, J. Should we measure corporate social responsibility? *Corp. Soc. Responsib. Environ. Manag.* **2003**, *10*, 25–39. [CrossRef]
23. Jha, A.; Cox, J. Corporate social responsibility and social capital. *J. Bank. Financ.* **2015**, *60*, 252–270. [CrossRef]
24. Gao, Y. Corporate social performance in China: Evidence from large companies. *J. Bus. Ethics* **2009**, *89*, 23–35. [CrossRef]
25. Bai, C.; Sarkis, J.; Dou, Y. Corporate sustainability development in China: Review and analysis. *Ind. Manag. Data Syst.* **2015**, *115*, 5–40. [CrossRef]
26. Chen, Y.-C.; Hung, M.; Wang, Y. The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China. *J. Account. Econ.* **2017**, *65*, 169–190. [CrossRef]

27. The 19th National Congress of the Communist Party of China. Available online: http://www.gov.cn/zhuanti/19thcpc/baogao.htm (accessed on 16 May 2018).

28. Clarkson, P.M.; Li, Y.; Richardson, G.D.; Vasvari, F.P. Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Account. Organ. Soc.* **2008**, *33*, 303–327. [CrossRef]

29. Orlitzky, M.; Siegel, D.S.; Waldman, D.A. Strategic corporate social responsibility and environmental sustainability. *Bus. Soc.* **2011**, *50*, 6–27. [CrossRef]

30. Owen, D.L.; O'Dwyer, B. Corporate social responsibility: The reporting and assurance dimension. In *The Oxford Handbook of Corporate Social Responsibility*; Oxford University Press: Oxford, UK, 2008; pp. 384–409.

31. Miller, E.M. Risk, uncertainty, and divergence of opinion. *J. Financ.* **1977**, *32*, 1151–1168. [CrossRef]

32. Grossman, S.J.; Stiglitz, J.E. On the impossibility of informationally efficient markets. *Am. Econ. Rev.* **1980**, *70*, 393–408. [CrossRef]

33. Cordeiro, J.J.; Tewari, M. Firm characteristics, industry context, and investor reactions to environmental CSR: A stakeholder theory approach. *J. Bus. Ethics* **2015**, *130*, 833–849. [CrossRef]

34. Hasnas, J. The normative theories of business ethics: A guide for the perplexed. *Bus. Ethics Q.* **1998**, *8*, 19–42. [CrossRef]

35. Mitchell, R.K.; Agle, B.R.; Wood, D.J. Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Acad. Manag. Rev.* **1997**, *22*, 853–886. [CrossRef]

36. Rodrigue, M. Contrasting realities: Corporate environmental disclosure and stakeholder-released information. *Account. Audit. Account. J.* **2014**, *27*, 119–149. [CrossRef]

37. Chen, S.; Sun, Z.; Tang, S.; Wu, D. Government intervention and investment efficiency: Evidence from China. *J. Corp. Financ.* **2011**, *17*, 259–271. [CrossRef]

38. Li, W.; Zhang, R. Corporate social responsibility, ownership structure, and political interference: Evidence from China. *J. Bus. Ethics* **2010**, *96*, 631–645. [CrossRef]

39. Liu, X.; Wang, C.; Wei, Y. Causal links between foreign direct investment and trade in China. *China Econ. Rev.* **2001**, *12*, 190–202. [CrossRef]

40. Kim, K.-H.; Kim, M.; Qian, C. Effects of corporate social responsibility on corporate financial performance: A competitive-action perspective. *J. Manag.* **2018**, *44*, 1097–1118. [CrossRef]

41. Lins, K.V.; Servaes, H.; Tamayo, A. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *J. Financ.* **2017**, *72*, 1785–1824. [CrossRef]

42. McGuire, J.B.; Sundgren, A.; Schneeweis, T. Corporate social responsibility and firm financial performance. *Acad. Manag. J.* **1988**, *31*, 854–872.

43. Oh, S.; Hong, A.; Hwang, J. An Analysis of CSR on Firm Financial Performance in Stakeholder Perspectives. *Sustainability* **2017**, *9*, 1023. [CrossRef]

44. Saeidi, S.P.; Sofian, S.; Saeidi, P.; Saeidi, S.P.; Saaeidi, S.A. How does corporate social responsibility contribute to firm financial performance? The mediating role of competitive advantage, reputation, and customer satisfaction. *J. Bus. Res.* **2015**, *68*, 341–350. [CrossRef]

45. Luo, Y.; Xue, Q.; Han, B. How emerging market governments promote outward FDI: Experience from China. *J. World Bus.* **2010**, *45*, 68–79. [CrossRef]

46. Unger, J. Bridges: Private business, the Chinese government and the rise of new associations. *China Q.* **1996**, *147*, 795–819. [CrossRef]

47. Becchetti, L.; Ciciretti, R.; Hasan, I. Corporate social responsibility, stakeholder risk, and idiosyncratic volatility. *J. Corp. Financ.* **2015**, *35*, 297–309. [CrossRef]

48. Tehulu, T.A. Determinants of financial sustainability of microfinance institutions in East Africa. *Eur. J. Bus. Manag.* **2013**, *5*, 152–158. [CrossRef]

49. Lourenço, I.C.; Branco, M.C.; Curto, J.D.; Eugénio, T. How does the market value corporate sustainability performance? *J. Bus. Ethics* **2012**, *108*, 417–428. [CrossRef]

50. Hunter, M.L.; Van Wassenhove, L.N.; Besiou, M.; Van Halderen, M. The agenda-setting power of stakeholder media. *Calif. Manag. Rev.* **2013**, *56*, 24–49. [CrossRef]

51. Besiou, M.; Hunter, M.L.; Van Wassenhove, L.N. A web of watchdogs: Stakeholder media networks and agenda-setting in response to corporate initiatives. *J. Bus. Ethics* **2013**, *118*, 709–729. [CrossRef]

52. Bruntland, G. *World Commission on Environment and Development (WCED, 1987): Our Common Future*; Oxford University Press: Oxford, UK, 1987.

53. Atkinson, G. Measuring corporate sustainability. *J. Environ. Plan. Manag.* **2000**, *43*, 235–252. [CrossRef]

54. Inglehart, R. Public support for environmental protection: Objective problems and subjective values in 43 societies. *PS Political Sci. Politics* **1995**, *28*, 57–72. [CrossRef]

55. Nghiem, L.T.; Papworth, S.K.; Lim, F.K.; Carrasco, L.R. Analysis of the capacity of Google Trends to measure interest in conservation topics and the role of online news. *PLoS ONE* **2016**, *11*, e0152802. [CrossRef] [PubMed]

56. Choi, M.J.; Torralba, A.; Willsky, A.S. Context models and out-of-context objects. *Pattern Recognit. Lett.* **2012**, *33*, 853–862. [CrossRef]

57. Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econ. J. Econ. Soc.* **1982**, *50*, 987–1007. [CrossRef]

58. Franses, P.H.; McAleer, M. Financial volatility: An introduction. *J. Appl. Econ.* **2002**, *17*, 419–424. [CrossRef]

59. McAleer, M.; Oxley, L. The econometrics of financial time series. *J. Econ. Surv.* **2002**, *16*, 237–243. [CrossRef]

60. Li, W.; Ling, S.; McAleer, M. Recent theoretical results for time series models with GARCH errors. *J. Econ. Surv.* **2002**, *16*, 245–269. [CrossRef]

61. Ling, S.; McAleer, M. Necessary and sufficient moment conditions for the GARCH (r, s) and asymmetric power GARCH (r, s) models. *Econ. Theory* **2002**, *18*, 722–729. [CrossRef]

62. Ling, S.; McAleer, M. Asymptotic theory for a vector ARMA-GARCH model. *Econ. Theory* **2003**, *19*, 280–310. [CrossRef]

63. Asai, M.; McAleer, M.; Medeiros, M.C. Asymmetry and long memory in volatility modeling. *J. Financ. Econ.* **2012**, *10*, 495–512. [CrossRef]

64. Marcucci, J. Forecasting stock market volatility with regime-switching GARCH models. *Stud. Nonlinear Dyn. Econ.* **2005**, *9*. [CrossRef]

65. Franses, P.H.; Van Dijk, D. Forecasting stock market volatility using (nonlinear) GARCH models. *J. Forecast.* **1996**, *15*, 229–235. [CrossRef]

66. French, K.R.; Schwert, G.W.; Stambaugh, R.F. Expected stock returns and volatility. *J. Financ. Econ.* **1987**, *19*, 3–29. [CrossRef]

67. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econ.* **1986**, *31*, 307–327. [CrossRef]

68. Bollerslev, T.; Mikkelsen, H.O. Modeling and pricing long memory in stock market volatility. *J. Econ.* **1996**, *73*, 151–184. [CrossRef]

69. Chan, F.; McAleer, M. Estimating smooth transition autoregressive models with GARCH errors in the presence of extreme observations and outliers. *Appl. Financ. Econ.* **2003**, *13*, 581–592. [CrossRef]

70. McAleer, M.; Medeiros, M.C. A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *J. Econ.* **2008**, *147*, 104–119. [CrossRef]

71. McAleer, M. Automated inference and learning in modeling financial volatility. *Econ. Theory* **2005**, *21*, 232–261. [CrossRef]

72. Nelson, D.B. Conditional heteroskedasticity in asset returns: A new approach. *Econ. J. Econ. Soc.* **1991**, *59*, 347–370. [CrossRef]

73. Martinet, G.G.; McAleer, M. On The Invertibility of EGARCH (p, q). *Econ. Rev.* **2018**, *37*, 824–849. [CrossRef]

74. McAleer, M.; Hafner, C.M. A one line derivation of EGARCH. *Econometrics* **2014**, *2*, 92–97. [CrossRef]

75. Chang, C.-L.; McAleer, M. The correct regularity condition and interpretation of asymmetry in EGARCH. *Econ. Lett.* **2017**, *161*, 52–55. [CrossRef]

76. Baillie, R.T.; Bollerslev, T.; Mikkelsen, H.O. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econ.* **1996**, *74*, 3–30. [CrossRef]

77. Chang, C.-L.; McAleer, M.; Tansuchat, R. Modelling long memory volatility in agricultural commodity futures returns. *Ann. Financ. Econ.* **2012**, *7*, 1250010. [CrossRef]

78. Shi, Y.; Ho, K.-Y. Modeling high-frequency volatility with three-state FIGARCH models. *Econ. Model.* **2015**, *51*, 473–483. [CrossRef]

79. Bentes, S.R. Measuring persistence in stock market volatility using the FIGARCH approach. *Phys. A Stat. Mech. Appl.* **2014**, *408*, 190–197. [CrossRef]

80. Cochran, S.J.; Mansur, I.; Odusami, B. Volatility persistence in metal returns: A FIGARCH approach. *J. Econ. Bus.* **2012**, *64*, 287–305. [CrossRef]

81. Belkhouja, M.; Boutahary, M. Modeling volatility with time-varying FIGARCH models. *Econ. Model.* **2011**, *28*, 1106–1116. [CrossRef]

82. Taylor, D.; Shan, Y.G. What drives the fledgling practice of social and environmental reporting by Chinese companies listed in Hong Kong? *Account. Account. Perform.* **2007**, *13*, 55.

83. Mehran, H. Executive compensation structure, ownership, and firm performance. *J. Financ. Econ.* **1995**, *38*, 163–184. [CrossRef]

84. Jenter, D.; Kanaan, F. CEO turnover and relative performance evaluation. *J. Financ.* **2015**, *70*, 2155–2184. [CrossRef]

85. Quitzow, R. Dynamics of a policy-driven market: The co-evolution of technological innovation systems for solar photovoltaics in China and Germany. *Environ. Innov. Soc. Trans.* **2015**, *17*, 126–148. [CrossRef]

86. Foley, H. Understanding Sustainability from a Global Perspective: Exploring the Role of Education for Sustainable Development within Contemporary Education in Ireland. Ph.D. Thesis, University College Cork, Cork, Ireland, 2017.

87. McAleer, M. Asymmetry and leverage in conditional volatility models. *Econometrics* **2014**, *2*, 145–150. [CrossRef]

Article

The Social Efficiency for Sustainability: European Cooperative Banking Analysis

Leire San-Jose [1,2,*], Jose Luis Retolaza [3] and Eric Lamarque [4]

[1] Faculty of Economics and Business (Sarriko), ECRI, University of the Basque Country, UPV/EHU, 48015 Bilbao, Spain

[2] Department of Accountancy, Finance and Economics, University of Huddersfield, Huddersfield HD1 3DH, UK

[3] Faculty of Economics and Business Administration, Economics Department, University of Deusto, 48014 Bilbo, Spain; joseluis.retolaza@deusto.es

[4] Management and the Governance of Financial Cooperatives (MGCF), Université Paris 1 Pantheon Sorbonne, 75013 Paris, France; eric.lamarque@univ-paris1.fr

[*] Correspondence: leire.sanjose@ehu.eus

Abstract: This paper seeks to establish the relationship between economic efficiency and social efficiency to analyze the sustainability of banking in Europe. The type-effect has been analyzed, as stakeholder value banks—cooperatives and saving banks—should not be less socially and economically efficient than commercial banks. This European analysis was made using the Bankscope database, as it provides a unique insight into the stakeholder view that clarifies, by an analysis of two-stage boundaries, that there is no single model of social and economic efficiency according to the type of financial entity in Europe. These findings contribute to the social cost paradox and shared value perspective, and more broadly to stakeholder theory. It is established that a tradeoff between economic and social efficiency is not needed. There are different behaviors in different European countries. Moreover, our results could lead to the development of social indicators of the sustainability aspects of organizations without resorting to traditional accounting.

Keywords: stakeholder theory; sustainability; risk; social efficiency; banking; cooperative banks; Data Envelopment Analysis (DEA)

1. Introduction

The situation of financial institutions is changing: regulation, governance, digitalization, and supervision are aspects that make banks change. These changes are still ongoing, but there is a need to assess whether business models are sustainable, even in adverse scenarios. The banking crisis is now at an end, enabling us to analyze the social and economic situation of financial institutions, and lay the foundations for a new story of banking in Europe. The efficiency of banks is one of the measures used to organize this sector, and if this measure is developed to achieve sustainability it will be marked a management line towards the purpose of sustainability of financial institutions, not only doing well but also doing good for all stakeholders. Traditionally it has applied bank efficiency from a general economic perspective, but for the purpose of this analysis, new, more social and sustainable aspects have been considered. Financial institutions that adopt a more social approach based on stakeholder value, namely cooperative and savings banks, tend to be secondary in nature [1,2], although in Europe they account for more than sixty percent of the market [3]. This paper will pay attention to them and use them to show a different view of the sustainability approach in banking.

Then, although most of the research regarding bank efficiency focuses on the economical view [4–6], a number of papers focus on the social efficiency of financial institutions [2,7,8]. In general,

a tradeoff between economic and social efficiency (probably because of the strong influence of financial theory), when one increase other decrease. But, to overcome this problem that we will address here, there are some previous theories; the paradox of social costs [9] and the shared-value perspective [10] state that sustainability should be obtained integrating both economic and social efficiencies. In line stakeholder theory [11], we establish that organizations, including financial entities, should create value for all stakeholders, and that the Triple Bottom Line [12] include the environment for future generations, as well.

To the best of our knowledge, no research has explicitly addressed the question of the social efficiency of European banks (an explanation for the concept of social efficiency is given in the next section), considering not only commercial, but also cooperative banks. Continuing with the work of previous authors [2,13–15], we aim to contribute, not only to empirical research on financial institutions, but also to demonstrate that achieving social efficiency does not necessarily imply a direct decline in an entity's economic goals; this is the primary interest of our research. The purpose of credit cooperatives, framed in the Social Economy, is to create value for workers and society as a whole This is in juxtaposition to commercial banks, where the shareholder value is the main objective [13]; hence, banking specialization (type) can influence in the level of social efficiency of each entity; then, the legal form could thus determine a specific behavior in this line. The country-effect will be also considered with the aim to develop a unique social efficiency model for European banking.

From certain perspectives may exist a direct link between sustainability and environment; however, from the Triple Bottom Line [16] there are three aspects of organizations fundamental for sustainability; economic, environmental, and social. The economic aspects are already incorporated in the classical financial theory; the environmental ones are referred to externalities in general, and the social ones; the focus of this paper, raised to the society inclusions as an important element in organizations; because the interests of stakeholders have been taken into account [11] for a broad sustainability purpose. It can be debatable the primacy of the different areas in relation to sustainability, but the relationship between sustainability in a broad sense and the social value of banking activities are inseparable. The inclusion of the interests of stakeholders is fundamental nowadays because organizations, in this case, banks, should answer their needs and return to them what they are asking for, at least because they are using the societal system for a banking purpose. Then, banks should make an to be socially responsible for sustainability in line that banks should return value to the stakeholders of banks, and the society in general, what they need: employment, less risks, society supports and wellbeing; among others [2].

We used the Bankscope database to obtain the variables to analyze the social efficiency and economic efficiency (profitability) of banking. The research period is 2014 because it is a year considered "out of the financial crisis effect" (see the literature [3,17] for a comparison analysis between pre- and postcrisis). An initial postcrisis picture (2008–2013) will lay the foundations for a future longitudinal study that undoubtedly will be of great importance for banking. However, prior to this, a year-base analysis will highlight the lack of connections between European countries in social terms. Our results will consequently strengthen future banking literature, particularly from a European social perspective. We have used Data Envelopment Analysis (DEA) and a Factorial Analysis of Variance to measure the efficiency of financial institutions.

This paper makes two contributions. Firstly, whilst previous studies have focused on bank efficiency to analyze economic efficiency [4], this paper considers another important aspect, namely social efficiency. Secondly, the European case provides unique information for analyzing the banking sector as a whole; as we used the population of financial institutions, the results have no sample bias, therefore shedding light on the real banking situation in which there is a country effect within a theoretically harmonized Europe in this highly regulated sector. It is to be expected that the European unification will entail a similar behavior of the entities in the group of EU countries in terms of achieving social efficiency. The results obtained have potentially major implications in order to encourage governance based on multiple stakeholder participation in financial institutions: stakeholder value

banks. The type and country effect should be analyzed in order to come up with a unique European banking efficiency model: European banking is not yet harmonized. This might contribute to the development of a sustainable European banking system in order to establish typologies, values, or regulations depending on the type of each financial institution (see a past paper [1] for an analysis of banking models in Europe).

The article is organized as follows. Section 2 reviews previous literature on the relationship between social efficiency and stakeholder theory, taking into consideration the inclusion of credit cooperatives. Section 3 explains the research hypothesis to establish the basis of the reasoning on the assumption made, the sample and methodology with the explanation of input/output data. In the next section, the empirical analysis results concerning country and type effects are shown, not only for economic efficiency/profitability, but also using social efficiency as a measure for analyzing the performance of financial institutions. After those results, the analysis of cooperative banks' social efficiency is also addressed in Section 4. Finally, in Section 5 a discussion is shown, and Section 6 ends with the conclusions, limitations, and recommendations for further research.

2. Literature Review and Framework

The efficiency of banks is a major issue that still remains unresolved, at least the social perspective of efficiency is a gap in the bank efficiency literature; our approach is based on carry on contributions in this line [18,19]. In the last decade the focus of this social view has been linked, for example, to specific types of financial institutions, such as microfinance institutions [16–18]. This is due to the fact that social purpose is inherent to them, and is intended to reduce poverty. In microfinance, studies that share this aim analyzed, not only the social efficiency of this type of institution, but also the relationship with economic efficiency and profitability. They concluded that those performances are correlated (economic and social efficiency), and that socially efficient microfinance institutions are not financially less efficient. This shows that in this type of institution at least, social responsibility does not penalize financial efficiency.

In this sense, and based on the Pareto social optimum [20], the notion of social efficiency is understood as the balance between resources for the purpose of the organization and generation of value for the society with those resources. Such resources could include equity and external funding, whilst those that generate value for stakeholders could be the amount of the loans, number of clients, or economic sustainability. In this regard, the organization is understood as a set of stakeholders with an aligned purpose, therefore the higher the profits generated for an entity's stakeholders—excepting negative externalities (Freeman, 1984)—the greater the social efficiency of an organization [2].

Then, our paper is related to the literature on bank social efficiency. In the model application of the DEA method to evaluate banks' efficiency presented below, the social value added approach has been chosen based on value for society. In this approach the equity (or more commonly, shareholders' equity) refers to the amount of capital contributed by the owners and accumulated reserves [2]. Bank deposits consist of total money placed into banking institutions for safekeeping providing liquidity and act as delegated monitors [21]. Both equity and deposits are defined as inputs [2] including the resources needed for bank activity [22]. The selection of outputs is based on social generation of value in banking based on stakeholders approach [19] then, customer loans, labor, social contribution, but considering the assumed risk level in the bank are defined outputs. Customer loans show the total lending of money by the bank to other entities, individuals, and/or organizations; labor refers to the number of jobs held by the entity throughout the year, and social contribution is the sum of the funds contributed by the bank to the public administration. Finally, the risk level is included as the difference between the risk admitted by each banking entity, and the provision of funds destined eventually to cover detected risk in each period. Hence, the optimal bank social efficiency trades off social value creation for stakeholders and the resources used for the activities of the bank. In the case of banking, although the values for stakeholders are based not only on market values, but also on non-market (i.e., free of charge use of financial entities dependencies, training for individuals and companies, or

newsletter), and emotional ones; due to the lack of generalized and normalized non-market social values, the social efficiency inputs and outputs of this research are limited to bank accounting-base data. In this regard, a second limitation should be taken into account, because the result may not be an optimal reflection of the social value generated by financial intermediation. This may be due to the atypical interest rates in which the financial market is immersed, with results that may not correctly reflect the value induced to third parties through the financing processes. This mismatch may have occurred in 2014, when interest essentially fails to reflect loan value due to the intervention of central banks. This argument is the reason why the volume of borrowed funds has been chosen as a proxy of the output, instead of the result obtained with the loans.

Once reviewed, the concept to analyze social efficiency, we proceed to review the studies analyzed in this paper: the geographic scope, European country, and the typology of financial institution (specialization).

2.1. Country-Effect Studies in Banking

There are a few European-based studies that focus on making a contribution to bank efficiency. For example, Chortareas et al. [23] has addressed the influence of financial freedom on European bank efficiency as a country-effect. They conclude that the freedom of a European country enjoys influences efficiency: free countries will have relatively higher levels of economic efficiency (cost reduction view). Another study conducted by Lozano-Vivas et al. [24] analyzed bank efficiency in ten European countries, concluding that it is lower than expected. Moreover, their findings indicated that environmental variables play an important role in explaining differences in efficiency. More recent analyses [14] have used the estimated profit and cost efficiencies of banks within a region as a proxy for financial quality, and have concluded that regions with more efficient banks are more resilient to Europe's financial and debt crisis. In addition, bank sector efficiency is related to economic growth. Galema and Koetter [25] used a stochastic production boundary model for European bank efficiency to indicate that the type of banking supervision (Single Supervisory Mechanism-European Central Bank (SSM-ECB) vs. National Competent Authorities (NCA) influences bank efficiency (cost and profit), SSM-ECB supervision means lower efficiencies. None of these papers have based their analyses on social aspects, a welfare state in which employment is generated, a level of infrastructure is maintained or social risks are limited. In this same line, a recent study, carried out by Fijałkowska et al. [26], established the performance between Corporate Financial Performance (CFP) and corporate social-environmental performance (CSP) for Central and Eastern European banks using the DEA as a methodology. The results suggest that a high economic efficiency entails a high socio-environmental efficiency, without necessarily creating an inverse relationship. The previous studies of Lozano-Vivas et al. [24] and Belke et al. [14] include the country-effect as a determinant for bank efficiency. We will continue their conclusions and focus on country-effect, which will contribute to the European banking harmonization level analysis. In this sense, bank taxes and risks are incorporated as outputs in the efficiency analysis.

2.2. Type-Effect Studies in Banking

In terms of type of banking institutions, a number of bank efficiency studies have been conducted. These include studies based on an analysis of Islamic banks [27,28]. In Islamic banking, the conclusions are not clear; there is no consensus over the comparison between commercial and Islamic banks in terms of financial efficiency (see for instance see a past paper [29]). There are some studies addressing savings bank efficiency [30,31]. They contend that when comparing saving banks and commercial banks it is important to control the geographical operational level, otherwise, we will establish or compare culturally, strategically, and tactically different financial institutions. Geography is therefore a relevant issue, particularly if we wish to provide an effective estimate for savings bank efficiency levels. Indeed, they display various financial characteristics with far-reaching implications for bank efficiency. Cuesta & Orea [31] have based their analysis on merger vs. no-merger, using savings

banks as the sample. They conclude that although merger firms are less efficient at first, they increase their technical efficiency and exceed in the end. The efficiency of cooperative banks has already been studied by Lang and Welzel [32] using panel data of German cooperative banks. They base their study on cost efficiency, rather than social efficiency, continuing the bank efficiency view of that decade to establish the economic efficiency of cooperative banks. Bos and Kool [33] have subsequently analyzed 401 cooperative banks in the Netherlands, and conducted both profit and cost efficiency analyses. Their control factors explain less than 10% of profit efficiency, even in a relatively small, homogeneous geographical area with banks close to each other. They conclude that a number of environmental factors have an impact on estimated efficiencies; the uncontrolled 90% of profit efficiency suggests that it is based on managerial inefficiencies. Other more recent studies by Manetti & Bagnoli [7] have analyzed Italian cooperative banks (a specific type of European credit cooperatives: Italian 'Banche di Credito Cooperativo'), and conclude that they are less efficient than traditional banks, probably because of their statutory commitments. After analyzing the distribution of value added for stakeholders (system strengthening, member, community, staff, and cooperative system), they have established that the efficiency of cooperative banks should be developed from a social point of view [34].

2.3. The Purpose of Our Study in Banking

In this line, our study considers a European bank analysis and two aspects in efficiency calculation: social efficiency and economic efficiency. Financial institution typology, studied in previous literature (Lang and Welzel [32] for cooperative banks; Tabak et al. [30] and Cuesta & Orea [31] for saving banks; and Chortareas et al. [23] and Lozano-Vivas et al. [24] for commercial banks; and Bal & Gölcükcü [35] for industrial banks), probably influences the social and economic efficiency relationship. Specifically, we have considered three types of financial institutions: commercial banks, savings banks, and cooperative banks (the former referred to as shareholder value bank and the latter two stakeholder value banks [3]), to find their differences, placing a particular focus on cooperative banks. Moreover, we have included, not only economic efficiency based on profitability (it is computed as the ratio of net income to tangible total assets), but also, and more exhaustively, social efficiency based on generation of value for bank stakeholders [36–38].

To sum up, the present investigation is relevant for the Stakeholder Theory [11]. On the one hand, in case there is evidence that there is no tradeoff between social and economic efficiency, the supposed social cost for shareholders would not remain a critical element to stakeholder theory and would be another contribution to the paradox of social costs [9]. On the other hand, a positive correlation between social and economic efficiency would be a "critical case" that would support stakeholder theory, in the absence of a longitudinal analysis. In turn, the fact that the economic and social efficiency of an entity may not necessarily be correlated, can contribute to justify the need for social and environmental accounting [SEA] [39] in the line of the triple bottom line [12]. This in itself will be a relevant contribution.

3. Hypothesis, Sample and Methodology

3.1. Hypothesis

The issue we face is to analyze whether type and country could influence, not only efficiency from a social perspective, but also profitability: resolving this question will contribute to stakeholder theory, and paradox of social cost and shared-value perspective. This is because some financial institutions, such as cooperative banks, are governed by stakeholders and not only shareholders (namely stakeholder value banks), where commercial banks are based more on hierarchy and linearly structured governance in which there is less or no stakeholder participation (namely shareholder value banks). To make contributions in this sense, we have resorted to statistical hypothesis testing using the hypothetical-deductive method. Prior to this, we employed the synthetic analytical method to identify the components of the problem and move them to a system of inputs and outputs.

In fact, several papers have debated the efficiency of cooperatives banks in relation to commercial banks [7,40]. There is some agreement that this type of institution is less efficient from an economic perspective, although greater social efficiency is expected [41].

Hypothesis 1 (H1). *Commercial banks have greater economic efficiency than cooperative banks.*

This hypothesis is consistent both with the theory of property rights [42] and with agency theory [43,44]. As Jensen [45] points out, a shareholder-oriented and controlled entity, such as a banking institution, can orient itself in a one-dimensional way toward a single objective, which will allow for more efficient management, at least compared with multiple objective management. Multiple objectives might consider for example, the priority of workers' interests [46] to the detriment of economic performance, as might occur when decision-making bodies do not correspond to capital, but mainly to workers and customers. Previous studies [5,13,14,47–49], have focused on the economic efficiency of banks that will prevail because of the sine-qua-non need for existence, the main objective of any bank. Cooperative banks focus on social aspects and are legitimated by society to create value for stakeholders [2,15,34]. Consequently, whilst it remains relevant, economic efficiency for cooperative banks is more instrumental than central.

Hypothesis 2 (H2). *Cooperatives banks have greater social efficiency than commercial banks in Europe.*

Consistent with classical social theory [50] such as the stakeholder theory [11,51], cooperatives tend to devote a significant part of their efforts to social improvement, or to stakeholders rather than capital gain [34]: namely workers, asset clients, liability customers, the social environment, and the public administration, among others. "Cooperative banks have a strong connection between risk taking and the moral narrative behind their organizational purpose, as values are commonly perceived to be an integral part of the business model" [34]: p. 22. It would therefore appear that the social outputs generated should be higher than those of commercial banks [2,7,32,33]. If this hypothesis were rejected, it would seriously question the social utility of cooperative banks, reducing their remit to the mere satisfaction of workers. It will be an evidence of the existence of social costs. Alternatively, it would show that market-based social determinants are not enough to establish the real social values of organizations.

There are not cooperative banks in all European countries, in those European countries that there are enough cooperatives banks (France, Spain, Germany, Italy, Austria, Portugal, Belgium, Denmark, and Finland) the third hypothesis is:

Hypothesis 3 (H3). *Credit cooperatives have superior social efficiency than commercial banks in each European country with credit cooperatives (France, Spain, Germany, Italy, Austria, Portugal, Belgium, Denmark, and Finland—subhypothesis).*

We have analyzed the differences in social efficiency for each country (France, Spain, Germany, Italy, Austria, Portugal, Belgium, Denmark, and Finland), considering that the social efficiency of cooperative banks is higher than efficiency of banks (see the literature [3] for reviewing descriptively the European banking system). Harmonization and new banking regulations provide European countries with the option of establishing a number of similarities across businesses in various European countries. However, there are cultural and environmental aspects, or technical aspects (regulation, taxes or policies) that could affect the social values of each type of financial institution. Previous studies by Lozano-Vivas et al. [24] and Belke et al. [14] include country-effect as a determinant for bank efficiency, and we will continue their research with more evidence to show the harmonization level of European banking.

Hypothesis 4 (H4). *There is a significant difference between European countries regarding the social efficiency of cooperative banks.*

With the aim of establishing possible differences across countries according to type, isolating the type and analyzing the country-effect only could enable us to predict, that in general terms, cooperative banks should be more efficient socially in those countries with a greater tradition in social affairs, such as France, Spain, Italy, or Germany [52,53].

The fundamental interest of this paper lies in determining the social efficiency of financial institutions, comparing bank typologies. However, there is no doubt that a further line of research should consist of identifying the various factors in each country that may influence the country effect, and we therefore believe it is necessary to conduct a qualitative analysis with a sample of significant banks from each country. In this sense, the determinant variables should include economic growth, welfare improvement, and, albeit in a negative sense, corruption. We have therefore selected a number of variables based on the results of Manetti & Bagnoli [7], whereby mutual and territorial aspects are relevant: a corruption index [54] and welfare data [55].

Hypothesis 5.1 (H5.1). *Greater corruption level in the country is negatively correlate with the efficiency of cooperative banks.*

Corruption has a negative impact on a country's economy [56] and naturally, also on its banking system. Taking a general index of country corruption level—the Corruption Perception Index, developed by Transparency International Association (for more information see https://www.transparency.org/research/cpi/overview), we analyzed the relationship between corruption level (measure by the inverse of Corruption Perceptions Index) and cooperative bank efficiency. A negative correlation is expected, whereby higher levels of corruption in countries will probably imply less efficient banking cooperatives.

Hypothesis 5.2 (H5.2). *There is a positive and significant correlation between the social efficiency of cooperative banks and Wealth-to-Well-Being Coefficient across EU-15 countries.*

Given the social approach of our research, rather than choosing an economic development measure such as GDP or its growth directly, we opted to use a social impact GDP indicator, namely the Wealth-to-Well-Being Coefficient, developed by the Boston Consulting Group, within the framework of Sustainable Economic Development Assessment (SEDA). This coefficient is obtained by comparing the SEDA score for a country's current welfare level with the expected score given the per capita GDP and the average for all countries. This provides a relative indicator of a country's effectiveness in converting wealth into welfare that benefits the population. Countries scoring higher than 1.0 offer greater levels of welfare than expected in accordance with their GDP, whilst those scoring below 1.0 provide lower levels of welfare than expected.

A positive relation between this indicator and the social efficiency of cooperative banks is expected, as their purpose is precisely to generate social welfare, essentially by satisfying their stakeholders' interests [53].

Hypothesis 5.3 (H5.3). *There is a positive and significant correlation between the social efficiency of cooperative banks and the Growth-to-Well-Being Coefficient across EU-15 countries.*

The Growth-to-Well-Being Coefficient was used, which was also developed by the Boston Consulting Group within the framework of SEDA. This coefficient compares each country's increase in its SEDA score with the expected score based on its GDP growth rate, given the average between the scores for recent progress and the GDP growth rates for all countries during the same period. The coefficient reveals whether a country has been able to convert growth into increased

welfare. As with the previous coefficient, countries scoring higher than 1.0 are experiencing greater improvements to welfare than forecast given their GDP growth rate between 2006 and 2015. A positive link between this indicator and the social efficiency of cooperative banks is predicted, as the higher the social efficiency of cooperative banks, the higher a country's forecasted increase in welfare, either due to traction or orientation towards the common good.

3.2. Sample

Our analysis focused on a single model of European bank efficiency, using data from Bankscope (Bureau van Dijk), and on financial entities (commercial banks, cooperative banks, and saving banks) in 2014 (2752 financial institutions with 38,528 observations) in EU-15 countries (Austria; Belgium, Denmark, France, Finland, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom). We have used all the data in the Bankscope using the criterion of activate entity, positive assets, and deposits of accounts in the analysis year; then, we used the population based on this known database (take into account the regular limitations of this type of databases). As the used data is the population we do not need to prove the independency of the data, because it is by definition independent. We have shown in the following table (see Table 1) the means of used variables, but we could provide a complete table. We have used Frontier Analyst® program for DEA analysis and SPSS for statistical analysis (univariate and multivariate analysis).

Table 1. Descriptive data of variables by country and type.

COUNTRY	TYPE	N	Equity	Deposits	Asset	Loan	Labor	Taxes	Risk	Profit
France	Banks	79	11,826	74,306	272,855	97,700	4956	362	1450	773
	Savings	18	5639	43,552	71,487	39,844	1717	156	318	327
	Cooperatives	65	20,790	114,126	378,549	141,292	7435	493	1235	1133
Spain	Banks	18	55,545	365,233	761,873	427,455	22,327	1306	10,418	3527
	Savings	14	9520	83,080	155,008	78,666	3196	203	5467	354
	Cooperatives	51	516	4887	6795	3563	209	4	71	25
Germany	Banks	99	7438	60,928	149,962	53,575	2532	127	598	254
	Savings	503	3114	24,988	42,113	23,894	1550	85	97	100
	Cooperatives	909	394	3167	5847	2837	187	13	25	22
Italy	Banks	66	9488	62,522	148,561	81,743	4879	266	5205	380
	Savings	31	2153	15,631	50,069	40,795	1115	71	1499	52
	Cooperatives	382	785	4247	9715	5762	309	6	479	14
Austria	Banks	36	4210	30,123	56,079	32,972	3021	111	717	282
	Savings	14	910	8114	13,098	8878	403	15	130	49
	Cooperatives	20	2631	14,004	37,343	18,372	914	26	541	29
Portugal	Banks	15	5873	48,710	82,291	51,937	3235	87	18	25
	Savings	79	820	11,514	13,060	5568	115	1	1	2
	Cooperatives	3	1918	16,340	23,189	11,100	1451	60	14	44
Belgium	Banks	16	10,705	97919	184,819	92,915	3196	331	2574	979
	Savings	3	2761	46,869	55,820	37,353	269	67	145	288
	Cooperatives	2	1004	6282	7499	4283	39	0	0	87
Denmark	Banks	26	7709	36,334	159,356	92,389	1616	131	1449	362
	Savings	29	287	1769	2310	1291	116	2	127	14
	Cooperatives	7	215	1107	1535	935	72	0	112	7
Finland	Banks	13	6085	33,763	157,186	43,221	1133	144	331	581
	Savings	14	158	1313	1807	1306	49	3	1	11
	Cooperatives	2	26,512	148,147	414,478	228,474	8930	1036	0	2649

3.3. Methodology

3.3.1. Two-Stage DEA Analysis

The empirical analysis has carried out using a two-step analysis; firstly a DEA analysis was applied in order to establish the relative efficiency of bank institutions. This was followed by a Factorial

Analysis of Variance to analyze the country and type-effect efficiency. A DEA analysis displays both strengths and weaknesses. On one hand, it is more flexible and there is no a pre-established relation between input and outputs that permits a quasi-real show of the relationship between variables; it is a welcome tool for extracting information from the empirical world [57] but is an untidy method for parametric regression. On the other hand, it is an extreme form method that assumes that if a DMU levels output with input, other DMUs should reach the same level (deterministic method); it does not directly imply homogeneity across DMUs, but is necessary to prevent inefficiencies stemming from nonuniform factors. Furthermore, variable selection is of fundamental importance as there are no suitable tests to estimate if the results of the analysis are stable or would vary significantly with other variables. Continuing with Stolp [57]: p. 115, "whether a given research tool is better or worse than another tool is really not the relevant point: what counts is the attitude that is brought to bear on the research". In this sense, we have not only provided statistical results, but also results that are coherent with the real situation.

Then, in the first stage we have carried a DEA analysis with the aim to develop the bank social efficiency and bank economic efficiency. It has been used some input/output variables that are shown in the Table 2 and argue and define after it.

Table 2. Inputs and outputs of Social and Economic (Profitability) Efficiencies.

	INPUTS	**OUTPUTS**
Social Efficiency for sustainability (SE) Definition: it is the balance between resources (input) and generation of value (outputs) for the society with those resources (inputs), being sustainable socially.	Equity (E) Deposits (D)	Customer Loans (CC) Labor (L) * Social Contribution/Taxes (SCT) Risk (R) **
Economic Efficiency (EE) Profitability Definition: it is the balance between the resources (assets) used to obtain the net profit.	Total Assets (TA)	Net Profit (P) ***

Notes: * the labor (L) is measure by headcounts. ** risk (R) is measured by obtained as the inverse of the sum of the contingent risks and commitments recognized by the different institutions. *** It is the benefit after taxes. The other variables are from the book accounts of banks without transformations.

Social Efficiency for Sustainability should include two major inputs, namely Equity and Deposits [58], thereby allowing for the control of funds related to corporate performance. Although other variables could be included as an input based on bank production theory, such as nondepositors borrowed funds, liquid assets, or financial services, the aim of this paper is to establish social welfare goals, and therefore we opted to start with the basic and minimum sustainable bank performance based on first-level needs: from our point of view, and based on McGuire et al. [58], these are the principal funds institutions need to generate value added.

Although it is not easy due to the absence of standardized indicators measuring the social value of organizations that show the added value generated by stakeholders; social efficiency is explained using the following outputs (as per the literature [2] as an ad-hoc selection based on the interests of the most important stakeholder groups—customers, employees, and the community—customers loans [59], labor, social contribution, profit, and risk. The customer loans (CC) input is relevant for social efficiency because it is the main financial resource of households and corporations [2,34], and is therefore necessary for social value. It is desirable, but not possible, to analyze the type of credit because depending on the aim as the level of social assistance will vary [60]. Nevertheless, as they could be necessary for social value, and because there is a lack of information regarding the exact purpose of the loans, they will all be considered in the same category. The second output, labor (L) is a clear indicator of social value [34]. It is important, insofar as society is based on work, as this guarantees people (bank workers in this case) the wherewithal to live. Social contribution (using taxes) (SC) is important for social efficiency because this output represents the funds that financial

institutions pay back to society with the aim of catering to citizens or society's needs [2]. The final output is risk. The risk is a negative output and outsourced by the entities, so if at a certain moment the risk is updated, the entity itself or a stakeholder of the financial entity will deal with the risk effect; then, it is the gap between provisioned and declared risk. Technically it is calculated as the difference between the risk perceived or declared and the amount of funds provisioned to cover the risks [5,48]. The risk assumed by financial institutions has involved huge cash bailouts, so although it is based on expectations, the citations are necessary because if the assumed risk is not taken into account and not reflected, it could have a negative and direct effect on the purchasing power of citizens and countries. Risk—obtained as the inverse of the sum of the contingent risks and commitments recognized by the different institutions—has previously been incorporated by Fiordelisi et al. [5] and San-Jose et al. [2]. See Table 3 for a mathematical representation of social efficiency using DEA.

Table 3. Data Envelopment Analysis (DEA) mathematical model for social efficiency.

Variables/Data	Equation
j = number of DMUs θ = efficiency rating yrj = amount of output r used by j unit [Customer Loans CC, Labor L, Social Contribution-Taxes SCT and Risk R) Xij = amount of input i used by j unit (Equity E and Deposits D) r = number of outputs from 1 to s i = number of inputs from 1 to m ur = coefficient or weight assigned by DEA to output r vi = coefficient or weight assigned by DEA to input i	For each DMU from 1 to n the Social Efficiency (maximizing the outputs) is shown as: $$\text{Max } \theta\left(j = 1 \rightarrow n\right) = \frac{u_1 \times CC_{10} + u_2 \times L_{20} + u_3 \times SCT_{30} - u_4 \times R_{40}}{v_1 \times E_{10} + v_2 \times D_{20}}$$

The economic view of bank efficiency has been measured using profitability, which can be conducted in a variety of ways, although, following Gutierrez-Nieto et al. [61], we measured it with a standard ratio: return on assets (ROA).

3.3.2. Second Stage: Factorial Analysis of Variance

In the second stage it has been used a Factorial Analysis of Variance applied. It is appropriate because one of the aims of this paper refers to study of the effect of two factors; country and type, and this analysis compares the means of two or more factors. Specifically, F tests are used to determine statistical significance of the factors and their interactions. Then, it also gives us information about their dependence or independence in the same experiment, which is crucial for the sustainability of European banking system, if it is based on harmony and homogenization. The tests are nondirectional in that the null hypothesis specifies that all means are equal and the alternative hypothesis simply states that at least one mean is different.

4. Results

This paper contains two levels of analysis: the first considers the overall social and economic efficiency of a European country, whilst the second is applied to specific aspects that could shed some light on the differences detected.

4.1. European Social Efficiency Analysis: A Path for Sustainability in Financial Area

Economic efficiency has been selected using a ratio of economic performance; the intention is to provide a general analysis and consensus for economic aspects of financial institutions. The following figure (see Figure 1) shows the means of the ROA of financial entities in each of the EU-15 countries. It is clear that there are some differences in this exploratory analysis in which the banks in some countries show higher ROAs: examples include Italy and Sweden, which contrast with other countries such as Greece or Portugal. The main countries are in a third group, with neither high nor low returns.
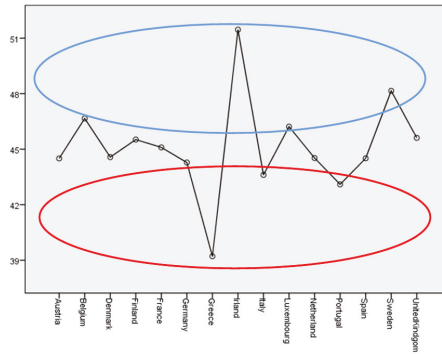
Figure 1. Profitability (ROA) analysis in European Countries: profile graph.

The analysis has compared the types of financial institutions with the aim of developing a unique bank efficiency model regardless of legal status. This means analysis reveals that there are no apparent differences across types in terms of economic efficiency measured by the profitability of financial entities (see Figure 2 to see the means in a visual form).
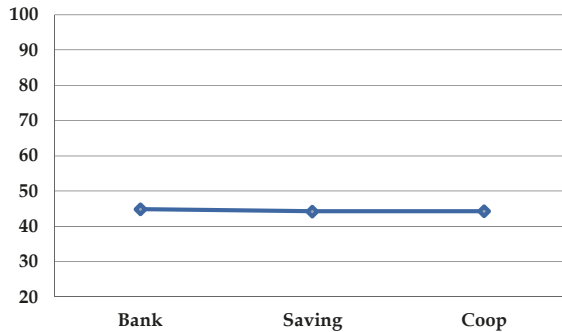

Figure 2. Economic Efficiency/Profitability: analysis by type of financial institution.

A multivariate analysis was conducted to determine both effects together, namely country (EU-15) and type (commercial banks, credit cooperatives, and savings banks). A Factorial Analysis of Variance was used whereby not only each factor is analyzed, but also their intersection (see Table 4).

Table 4. Multivariate analysis of economic efficiency: Factorial Analysis of Variance.

Origin	Type III of Sum of Squares	df	Quadratic Means	F	Sig.	Partial to Squared Eta
Corrected Model	3443.086	38	90.608	8.088	0.000	0.102
Intersection	319,453.149	1	319,453.149	28,517.022	0.000	0.913
Type	35.510	2	17.755	1.585	0.205	0.001 ***
Country	1876.861	14	134.061	11.967	0.000	0.058
Type * Country	861.991	22	39.181	3.498	0.000	0.028 **
Error	30,357.940	2710	11.202			
Total	5,453,484.370	2749				
Total corrected	33,801.026	2748				

Note: * Significant at $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

We used Factorial Analysis of Variance, revealing whether the iteration between type and country influences the dependent variable (economic efficiency/profitability). When an interaction is significant, as in this case, attention should be paid to the iteration rather than the main effects. In this case, the iteration is significant, and therefore, country and type together have some influence on the profitability of the financial institution. See the Appendix A for the robustness independent test of each dependent variable. The model explains a significant variance regarding the dependent variable, ROA, albeit not individually, as type is not a good control variable, and does not make any distinction.

The same method was applied to social efficiency. Following the DEA analysis, in which the selected inputs and outputs are used (as explained in a previous section), we compared social efficiency means in accordance with the institutions' country and type. The data used consisted of the population, thereby preventing sampling bias and data collection problems. Furthermore, the considerable size of the sample avoids normality problems. The DMUS are also higher than three times the number of inputs plus the number of outputs [57], thereby guaranteeing no sensitivity to specification. The efficiency surface is not sensitive to the sample, which means that in this case the frontier surface is robust. There is no autocorrelation (Durbin Watson is higher than 1.4: standing at more than 1.827), thereby avoiding the need to adjust or include more variables. The collinearity diagnosis establishes that the highest level of explanatory variable is 7.401, indicating no multicollinearity problem.

The European country effect using social efficiency as a dependent variable is shown in Figure 3, with no similarities in terms of social efficiency across those countries. They have shown visually the means of social efficiency for sustainability by country. Some countries display a high level of social efficiency, namely Austria, Belgium, Finland, France, Germany, Luxemburg, Portugal, Spain, and the UK. Other countries are not particularly efficient in social terms; examples include Denmark, Greece, and Italy.
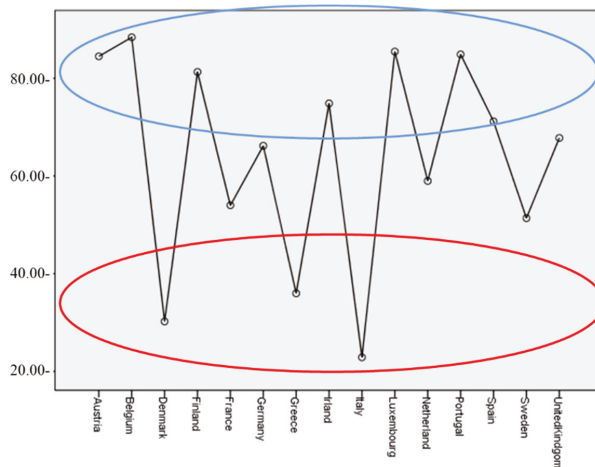

Figure 3. Social Efficiency for sustainability analysis in European Countries: profile graph.

The following figure (Figure 4) shows means by type. In general terms, commercial banks are more socially efficient than savings banks and cooperatives. We did not predict this in our hypothesis, but in this analysis at least, the initial results show the potentially social aspects of banks compared to theoretically more social typologies, such as savings banks and credit cooperatives. It should be remembered that the inputs and outputs used for developing social efficiency are based on market social value because there is no standardization of other aspects of social value in banking.
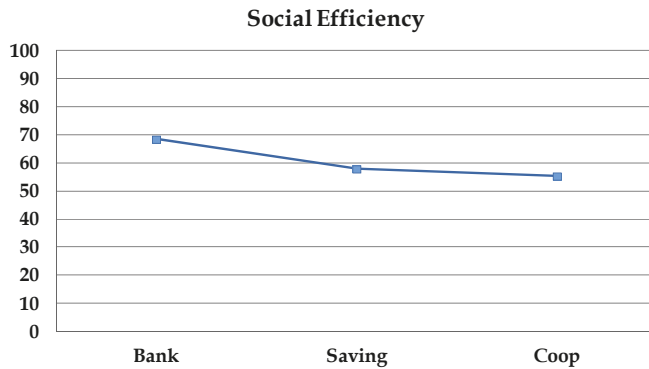
Figure 4. Social Efficiency for sustainability: analysis by type of financial institution.

Concerning social efficiency, we carried out a multivariate analysis to establish both effects together: country and type. This involved a Factorial Analysis of Variance in which not only each factor is analyzed, but also the intersection between them (see Table 5).

Table 5. Multivariate analysis of social efficiency for sustainability: Factorial Analysis of Variance.

Origin	Type III of Sum of Squares	df	Quadratic Means	F	Sig.	Partial to Squared Eta
Corrected Model	1132,242.643	38	29,795.859	24.510	0.000	0.256
Intersection	718,533.603	1	718,533.603	591.069	0.000	0.179
Type	3223.202	2	1611.601	1.326	0.266	0.001 ***
Country	377,493.668	14	26,963.833	22.181	0.000	0.103
Type * Country	81,397.755	22	3699.898	3.044	0.000	0.024 **
Error	3,294,415.588	2710	1215.652			
Total	13,873,241.818	2749				
Total corrected	4,426,658.231	2748				

Note: * Significant at $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

The iteration is significant, so this model explains a significant variance around social efficiency. However, country and type should not be considered separately: neither aspect is a good control variable due to the absence of differentiation. Therefore, country and type should be analyzed jointly. Figure 5 shows both variables together. It is shown visually the means of social efficiency for sustainability by type of financial entity and by country; both together.
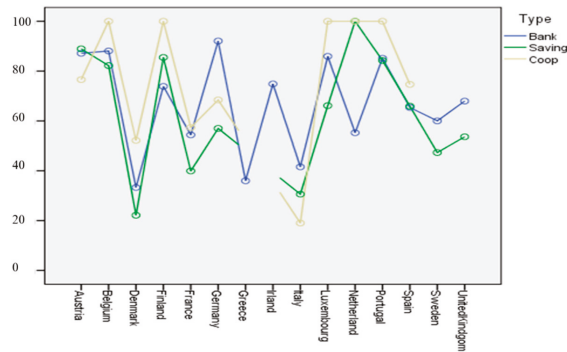
Figure 5. Country and type effect for social efficiency for sustainability: profile graph.

4.2. Country Effect in Social Efficiency for Sustainability between Banks and Cooperatives

With the aim of analyzing the banking situation in each country, we conducted a comparative means analysis of social efficiency. We used three types, as shown in Table 6, which indicate those applied to banks and credit cooperatives. There are only two countries where those financial institutions differ in terms of social efficiency, namely, Germany and Italy, and the relationship is opposite to that expected (banks outperform credit cooperatives). In the other countries (France, Spain, Austria, Portugal, Belgium, Denmark, and Finland), no significant differences were observed, although in some the social efficiency of cooperatives is higher than in banks (France, Spain, Portugal, Belgium, Denmark, and Finland). Some EU-15 countries do not have enough financial entities in each type, namely Greece, Ireland, Luxembourg, Sweden, the Netherlands, and the United Kingdom.

Table 6. Means comparison (Tamhane) between financial institution categories by country.

COUNTRY	TYPE	N	Mean Social Efficiency (SD)	Levene	F (Sign) (Inter-Groups)	Tamhane (Three Types, But Banks vs. Coop Shown Only)
France	Banks	79	54.406 (35.542)	9.695 ***	1.968 (0.143)	No Sign.
	Cooperatives	65	57.374 (32.239)			
Spain	Banks	18	65.397 (31.103)	2.117 *	0.566 (0.570)	No Sign.
	Cooperatives	51	74.667 (39.751)			
Germany	Banks	99	92.064 (20.980)	306.977 ***	38.708 (0.000)	Sign. ***
	Cooperatives	909	68.425 (38.932)			
Italy	Banks	66	41.555 (36.515)	32.982 ***	24.645 (0.000)	Sign. ***
	Cooperatives	382	19.015 (22.112)			
Austria	Banks	36	87.144 (28.237)	3.400 **	0.925 (0.402)	No Sign.
	Cooperatives	20	76.589 (37.001)			
Portugal	Banks	15	85.163 (30.927)	2.857 *	0.405 (0.668)	No Sign.
	Cooperatives	3	100.000 (0.000)			
Belgium	Banks	16	88.083 (22.179)	1.997	0.374 (0.693)	No Sign.
	Cooperatives	2	100.000 (0.000)			
Denmark	Banks	26	33.277 (30.964)	10.621 *	3.526 (0.036)	No Sign.
	Cooperatives	7	52.184 (45.718)			
Finland	Banks	13	73.880(36.061)	4.034 *	0.797 (0.461)	No Sign
	Cooperatives	2	100.000 (0.000)			

Note: * Significant at $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

To test the robustness of the previous results, we repeated the empirical research using a post-hoc test, such as Games-Howell [62], it was concluded that this post-hoc test is appropriate when there are doubts regarding the normality and homogeneity of variables. Similar results were achieved, as shown in Appendix B.

4.3. Analysis of Credit Cooperatives' Social Efficiency Across European Countries

Some countries may share similarities in terms of the social efficiency of credit cooperatives. A multiple means comparison was conducted of nine European countries with sufficient data on the social efficiency of credit cooperatives. Table 7 shows those countries displaying significant differences.

Table 7. Credit Cooperatives' social efficiency by country: a matrix of significance by pair comparisons among countries.

Country/Country	N	Mean	SD	I	D	Fr	G	S	A	B	Fi	P
Italy	382	19.015	22.112	1	0.980	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Denmark	7	52.184	45.718		1	1.00	1.00	1.00	1.00	0.696	0.696	0.696
France	65	57.374	32.239			1	0.311	0.382	0.814	0.000	0.000	0.000
Germany	909	68.425	38.932				1	1.00	1.00	0.000	0.000	0.000
Spain	51	74.667	39.751					1	1.00	0.001	0.001	0.001
Austria	20	76.589	37.001						1	0.321	0.321	0.321
Belgium	2	100	0							1	1.00	1.00
Finland	2	100	0								1	1.00
Portugal	3	100	0									1
Total	1441	55.237	41.237									

4.4. Country Effect of Cooperative Bank Efficiency

We sought to identify whether there are any country variables that can act as independent variables related to the social efficiency of credit cooperatives in order to gain a deeper insight into the impact of the country effect on social efficiency (see Table 8).

Table 8. Correlations among country indexes and country social efficiency means.

Indexes of Nine Countries: Italy, Denmark, France, Germany, Spain, Austria, Belgium, Finland, Portugal	Mean (Deviation)	Correlations Statistics (Pearson) with Social Efficiency 72.02 (26.946)	Significance
Corruption level (measure by the inverse of Corruption Perceptions Index- taken from Transparency International Corruption index) +	21.11 (19.601)	−6.22	***
Wealth-to-Well-Being Coefficient	1.083 (0.070)	0.799	***
Growth-to-Well-Being Coefficient	0.96 (0.221)	0.721	***

Note: Significant at *** $p < 0.001$; + other proxies could be used to test, such as indexes from World Bank (http://info.worldbank.org/governance/wgi/#home).

Interestingly, there is a significant correlation between the social efficiency of credit cooperatives and welfare, both in global terms and in terms of the growth of European countries. However, in recent years growth has been much lower in Italy than in Germany, a trend that looks set to continue in the years to come (see current issues of the IMF World Outlook). There may be some intermediate variables, such as the corruption index, which affect both ratios: social efficiency and welfare—both static and dynamic.

4.5. Hypothesis Testing

Each hypothesis is constructed in accordance with the previous hypothesis test with the aim to understand the unforeseen results obtained, so-called constructing hypothesis by testing. But in this paper we have shown them together in chronological order to be coherent with uses and custom in publications.

Hypothesis 1 (H1). *Commercial banks have greater economic efficiency than credit cooperatives.*

This is maintained the null hypothesis, because there are no significant differences between credit cooperatives and banks regarding economic efficiency.

Hypothesis 2 (H2). *Credit cooperatives have greater social efficiency than commercial banks in Europe.*

The null hypothesis is not rejected; there are no significant differences among types of financial entities in terms of social efficiency.

Hypothesis 3 (H3). *Credit cooperatives have superior social efficiency than commercial banks in each European countries with credit cooperatives (France, Spain, Germany, Italy, Austria, Portugal, Belgium, Denmark, and Finland—subhypothesis).*

There is some doubt regarding this relationship because there are no significant differences between credit cooperatives and banks regarding social efficiency in most European countries, and in those with significant differences; Germany and Italy, the banks are more socially efficient than credit cooperatives. So, the hypothesis is rejected.

Hypothesis 4 (H4). *There is a significant difference between European countries regarding the social efficiency of credit cooperatives.*

The null hypothesis is rejected because there are differences between European countries in terms of the social efficiency of credit cooperatives (see Table 6).

Hypothesis 5.1 (H5.1). *Greater corruption level in the country is negatively correlate with the efficiency of cooperative banks.*

Hypothesis 5.2 (H5.2). *There is a positive and significant correlation between the social efficiency of credit cooperatives and Wealth-to-Well-Being Coefficient across EU-15 countries.*

Hypothesis 5.3 (H5.3). *There is a positive and significant correlation between the social efficiency of credit cooperatives and the Growth-to-Well-Being Coefficient across EU-15 countries.*

In exploratory terms, we might contend that there is a relationship between these indexes (corruption index, wealth-to-well-being, and growth-to-well-being) and social efficiency in European countries. It is to be expected that social efficiency in those countries with less corruption will be higher, whilst those countries with high welfare coefficients will generate greater social efficiency. This indicates the possibility of linking social efficiency with specific country aspects; in other words, the connection between banking systems and country welfare (negatively or positively measured; corruption and well-being coefficients). Future research should be conducted to identify the mechanisms of the possible influence of these variables on efficiency, and to investigate whether, conversely, banking efficiency could also be a cause of greater welfare or lower levels of corruption. To confirm those aspect a regression analysis including the control variables will be necessary.

5. Discussion

This paper assesses the efficiency of banking in Europe in 2014 by using the boundary method, under European harmonization. The research focuses on social efficiency for sustainability. Specifically, we have paid attention to credit cooperatives because their strategy is based on social values for being sustainable. We have also endeavored to analyze the country-effect in Europe. We conclude that European banking is not yet harmonized. In the line of Lozano-Vivas et al. [24], we obtained

evidence that allows us to encourage the development of policies towards the harmonization of the banking system in Europe, at least if we advocate a more social economy. The geographical effect and stakeholder participation based on the interest in responding to stakeholders have important implications for policymakers because one policy does not necessarily fit all. Instead, it is important to establish the determining factors that make possible a new vision of sustainability-oriented banking. We suggest certain country indexes as potential moderating variables that could establish the social efficiency in banking of some European countries; corruption and well-being index.

The paper's contribution is relevant both to scholars and practitioners. Related to scholars, we first contribute to the literature through a preliminary exploration of how social efficiency could be developed, and how this concept establishes certain differences depending on the type of financial institution and the country (European). We introduce social efficiency for sustainability concept measured according to accounting-based data (market based social values). However, and such as second contribution the actual accounting-based data is not of a sufficiently high quality in order to show the whole social story of banking; more exhaustive data are needed to show how stakeholder interests are accomplished in banking. Cooperative banks have segregated more exhaustive internal information, a fact that could represent the most important social value of these entities; nevertheless, if they are not public and harmonized, they are not fully used. The benchmarking and improvement options and welfare for society is not possible. It is therefore necessary to develop proper social value measures for hybrid organizations that complement their economic and social results, such as credit cooperatives. Those new social indicators could be for example: first, all those costs with social function that do not have an economic interest (transactional); second, the social value generated for the stakeholders outside the market (nontransactional); and the third, the emotional value contributed to the people who interact in organizations (relational).

6. Conclusions, Limitations, and Future Research

Thus, the findings evidence that there is no tradeoff between social efficiency and economic efficiency, although we have not been able to confirm that there is a positive relationship between them. We contribute to the applicability of the stakeholder theory since one of its main obstacles is eliminated: the possible conflict between economic and social efficiency. We falsify the condition of necessity; which means that there does not always have to be opposition between social and economic efficiency. This evidences that social costs are a paradox; another new contribution from the banking sector in this case. However, we cannot confirm that there is correlation between economic efficiency and social efficiency as proposed by the shared-value perspective.

Finally, our research has a series of limitations. The social value metric based only on accounting-based data, the selected inputs and outputs that inform social efficiency are not accepted by all researchers due to the lack of literature on this topic. Our study is also limited to punctual data analysis, the situation in 2014. This is because our aim was to lay the foundations for a deeper longitudinal analysis in postcrisis Europe. Furthermore, the reporting bias is one of the most important limitations because of difficulties to obtaining population data about social for sustainability elements.

In addition to the aforementioned need for future social measures to demonstrate the social value of banks and a longitudinal analysis, a further area of research could consist of analyzing the country effect in relation to institution effect. This research could be conducted in line with the work of Belke et al. [14]. A comparative analysis of transnational financial institutions would enable us to determine the degree of stability of this efficiency in the various countries they operate in, or in contrast, whether they are highly differential. A further aspect for consideration could be the extent to which they correlate with the development indexes of these countries. Such analyses would enable us to determine the degree to which attempts to harmonize the European Banking Union are proving successful, as well as the role the possible country-based differentiation in efficiencies could play as a risk absorption mechanism. Despite being of major interest for the topic addressed here, the work should focus on non-cooperative banks, as transnationality is not a defining feature of cooperative credit institutions.

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

Appendix A

Table A1. Robustness Statistics: profitability and Social Efficiency.

Test/Variable	Profitability		Social Efficiency	
	Type-Effect	**Country-Effect**	**Type-Effect**	**Country-Effect**
Welch	2.961	15.593 ***	24.103 ***	80.930 ***
Brown-Forsythe	4.191 *	5.752 ***	22.994 ***	68.427 ***

Note: Significant at * $p < 0.05$; *** $p < 0.001$.

Appendix B

Table A2. Games–Howell test.

					p	95% Confidence Level
Social Efficiency: Type-Effect						
Games-Howell	Bank	Saving	10.585	2.122	0.000	5.61
		Coop	13.133	1.911	0.000	8.65
	Saving	Bank	−10.585	2.122	0.000	−15.56
		Coop	2.549	1.790	0.329	−1.65
	Coop	Bank	−13.133	1.911	0.000	−17.62
		Saving	−2.549	1.790	0.329	−6.75
Profitability: type-effect					p	95% confidence level
Games-Howell	Bank	Saving	0.658	0.278	0.048	0.01
		Coop	0.549	0.274	0.112	−0.10
	Saving	Bank	−0.658	0.278	0.048	−1.31
		Coop	−0.109	0.102	0.536	−0.35
	Coop	Bank	−0.549	0.274	0.112	−1.19
		Saving	0.109	0.102	0.536	−0.13

Social Efficiency: country-effect			Mean Differences	Standard Error	p
Games-Howell Test.	**F**	**p**			
Austria	0.925	0.402	Bank-Saving	−1.792	8.943	0.978
			Coop-Bank	10.556	9.519	0.516
			Coop-Saving	12.347	11.238	0.522
Belgium	0.374	0.693	Bank-Saving	5.884	18.644	0.948
			Coop-Bank	−11.916	5.545	0.113
			Coop-Saving	17.800	17.800	0.645
Denmark	3.526	0.036	Bank-Saving	11.128	6.977	0.260
			Coop-Bank	−18.907	18.316	0.580
			Coop-Saving	−30.035	17.618	0.274
Finland	0.797	0.461	Bank-Saving	−11.607	12.772	0.640
			Coop-Bank	−26.120	10.002	0.055
			Coop-Saving	14.513	7.942	0.199
France	1.968	0.143	Bank-Saving	14.484	6.852	0.101
			Coop-Bank	−2.968	5.655	0.859
			Coop-Saving	17.452 *	6.852	0.039
Germany	38.708	0.000	Bank-Saving	35.071 *	2.751	0.000
			Coop-Bank	23.639 *	2.473	0.000
			Coop-Saving	11.433 *	2.189	0.000
Italy	24.645	0.000	Bank-Saving	10.946	6.649	0.233
			Coop-Bank	22.540 *	4.635	0.000
			Coop-Saving	−11.594	5.028	0.069

Table A2. *Cont.*

Social Efficiency: country-effect				Mean Differences	Standard Error	*p*
Games-Howell Test.	F	*p*				
Luxemburg	0.601	0.552				
Netherland	1.197	0.322				
Portugal	0.405	0.668	Bank-Saving	0.903	8.666	0.994
			Coop-Bank	−14.837	7.985	0.187
			Coop-Saving	−15.739 *	3.367	0.000
Spain	0.566	0.570	Bank-Saving	−0.514	12.426	0.999
			Coop-Bank	−9.270	9.205	0.577
			Coop-Saving	−8.756	11.474	0.729
Sweden	1.873	0.176				
United Kingdom	0.148	0.702				
When there are less than two entities is not possible to apply this test.						

Note: Significant at * $p < 0.05$.

References

1. Ayadi, R.; Groen, W.P.D. Stress Testing, Transparency, and Uncertainty in European Banking. *Oxf. Handb. Econ. Inst. Transpar.* **2014**. [CrossRef]

2. San-Jose, L.; Retolaza, J.L.; Torres Pruñonosa, J. Efficiency in Spanish banking: A multistakeholder approach analysis. *J. Int. Financ. Mark. Inst. Money* **2014**, *32*, 240–255. [CrossRef]

3. Ayadi, R.; Llewellyn, D.T.; Schmidt, R.H.; Arbak, E.; De Groen, W.P. *Investigating Diversity in the Banking Sector in Europe: Key Developments, Performance and Role of Cooperative Banks*; CEPS: Brussels, Belgium, 2010; ISBN 978-94-6138-042-5.

4. Berger, A.N.; Humphrey, D.B. Efficiency of financial institutions: International survey and directions for future research. *Eur. J. Oper. Res.* **1997**, *98*, 175–212. [CrossRef]

5. Fiordelisi, F.; Marques-Ibanez, D.; Molyneux, P. Efficiency and risk in European banking. *J. Bank. Financ.* **2011**, *35*, 1315–1326. [CrossRef]

6. Ramly, Z.; Chan, S.-G.; Mustapha, M.Z.; Sapiei, N.S. Women on boards and bank efficiency in ASEAN-5: The moderating role of the independent directors. *Rev. Manag. Sci.* **2017**, *11*, 225–250. [CrossRef]

7. Manetti, G.; Magnoli, L. Mutual and social efficiency of italian co-operative banks: An empirical analysis. *Ann. Public Coop. Econ.* **2013**, *84*, 289–308. [CrossRef]

8. Lebovics, M.; Hermes, N.; Hudon, M. Are Financial and Social Efficiency Mutually Exclusive? A Case Study of Vietnamese Microfinance Institutions. *Ann. Public Coop. Econ.* **2016**, *87*, 55–77. [CrossRef]

9. Pava, M.L.; Krausz, J. The association between corporate social-responsibility and financial performance: The paradox of social cost. *J. Bus. Ethics* **1996**, *15*, 321–357. [CrossRef]

10. Porter, M.E.; Kramer, M.R. Creating Shared Value. How to reinvent capitalism—And unleash a wave of innovation and growth. *Harv. Bus. Rev.* **2011**, *89*, 62–77.

11. Freeman, R.E. *Strategic Management: A Stakeholder Approach*; Pitman: Boston, MA, USA, 1984; ISBN 978-0-273-01913-8.

12. Elkington, J. Partnerships from cannibals with forks: The triple bottom line of 21st-century business. *Environ. Qual. Manag.* **1998**, *8*, 31–51. [CrossRef]

13. Fiordelisi, F. Shareholder value efficiency in European banking. *J. Bank. Financ.* **2007**, *31*, 2151–2171. [CrossRef]

14. Belke, A.; Haskamp, U.; Setzer, R. Regional bank efficiency and its effect on regional growth in "normal" and "bad" times. *Econ. Model.* **2016**, *58*, 413–426. [CrossRef]

15. Wijesiri, M.; Martínez-Campillo, A.; Wanke, P. Is there a trade-off between social and financial performance of public commercial banks in India? A multi-activity DEA model with shared inputs and undesirable outputs. *Rev. Manag. Sci.* **2017**, 1–26. [CrossRef]

16. Elkington, J. Towards the sustainable corporation: Win-win-win business strategies for sustainable development. *Calif. Manag. Rev. Berkeley* **1994**, *36*, 90. [CrossRef]

17. Tsionas, E.G.; Mamatzakis, E.C. Adjustment costs in the technical efficiency: An application to global banking. *Eur. J. Oper. Res.* **2017**, *256*, 640–649. [CrossRef]

18. Gutierrez-Goiria, J.; San-Jose, L.; Retolaza, J.L. Social Efficiency in Microfinance Institutions: Identifying How to Improve It. *J. Int. Dev.* **2017**, *29*, 259–280. [CrossRef]

19. Bibi, U.; Balli, H.O.; Matthews, C.D.; Tripe, D.W.L. Impact of gender and governance on microfinance efficiency. *J. Int. Financ. Mark. Inst. Money* **2018**, *53*, 307–319. [CrossRef]

20. Chang, H.F. A Liberal Theory of Social Welfare: Fairness, Utility, and the Pareto Principle. *Yale Law J.* **2000**, *110*, 173–235. [CrossRef]

21. Casu, B.; Molyneux, P. A comparative study of efficiency in European banking. *Appl. Econ.* **2003**, *35*, 1865–1876. [CrossRef]

22. Gorton, G.; Pennacchi, G. Financial Intermediaries and Liquidity Creation. *J. Financ.* **1990**, *45*, 49–71. [CrossRef]

23. Chortareas, G.E.; Girardone, C.; Ventouri, A. Financial freedom and bank efficiency: Evidence from the European Union. *J. Bank. Financ.* **2013**, *37*, 1223–1231. [CrossRef]

24. Lozano-Vivas, A.; Pastor, J.T.; Pastor, J.M. An Efficiency Comparison of European Banking Systems Operating under Different Environmental Conditions. *J. Product. Anal.* **2002**, *18*, 59–77. [CrossRef]

25. Galema, R.; Koetter, M. European Bank Efficiency and Performance: The Effects of Supranational Versus National Bank Supervision. In *The Palgrave Handbook of European Banking*; Palgrave Macmillan: London, UK, 2016; pp. 257–292. ISBN 978-1-137-52143-9.

26. Fijałkowska, J.; Zyznarska-Dworczak, B.; Garsztka, P.; Fijałkowska, J.; Zyznarska-Dworczak, B.; Garsztka, P. Corporate Social-Environmental Performance versus Financial Performance of Banks in Central and Eastern European Countries. *Sustainability* **2018**, *10*, 772. [CrossRef]

27. Alam, N.; Rizvi, S.A.R. Empirical Research in Islamic Banking: Past, Present, and Future. In *Islamic Banking*; Palgrave CIBFR Studies in Islamic Finance; Palgrave Macmillan: Cham, Switzerland, 2017; pp. 1–13, ISBN 978-3-319-45909-7.

28. Widiarto, I.; Emrouznejad, A. Social and financial efficiency of Islamic microfinance institutions: A Data Envelopment Analysis application. *Socioecon. Plan. Sci.* **2015**, *50*, 1–17. [CrossRef]

29. Johnes, J.; Izzeldin, M.; Pappas, V.; Tsionas, M. Measuring Efficiency Convergence in Islamic and Conventional Banks: Cross-Country Evidence. 2017. Available online: https://ssrn.com/abstract=2960018 (accessed on 30 August 2018).

30. Tabak, B.M.; Miranda, R.B.; Fazio, D.M. A geographically weighted approach to measuring efficiency in panel data: The case of US saving banks. *J. Bank. Financ.* **2013**, *37*, 3747–3756. [CrossRef]

31. Cuesta, R.A.; Orea, L. Mergers and technical efficiency in Spanish savings banks: A stochastic distance function approach. *J. Bank. Financ.* **2002**, *26*, 2231–2247. [CrossRef]

32. Lang, G.; Welzel, P. Mergers Among German Cooperative Banks: A Panel-based Stochastic Frontier Analysis. *Small Bus. Econ.* **1999**, *13*, 273–286. [CrossRef]

33. Bos, J.W.B.; Kool, C.J.M. Bank efficiency: The role of bank strategy and local market conditions. *J. Bank. Finance* **2006**, *30*, 1953–1974. [CrossRef]

34. Minto, A. The spirit of the law over its letter: The role of culture and social norms in shielding cooperative banks from systemic shocks. *Law Financ. Mark. Rev.* **2016**, *10*, 16–26. [CrossRef]

35. Bal, H.; Gölcükcü, A.; Bal, H.; Gölcükcü, A. Data Envelopment Analysis: An Application to Turkish Banking Industry. *Math. Comput. Appl.* **2002**, *7*, 65–72. [CrossRef]

36. Freeman, R.E. The New Story of Business: Towards a More Responsible Capitalism. *Bus. Soc. Rev.* **2017**, *122*, 449–465. [CrossRef]

37. Freeman, R.E.; Ginena, K. Rethinking the Purpose of the Corporation: Challenges from Stakeholder Theory. *Notizie di Politeia* **2015**, *31*, 9–18.

38. San-Jose, L.; Retolaza, J.L.; Freeman, R.E. Stakeholder Engagement at Extanobe: A Case Study of the New Story of Business. In *Stakeholder Engagement: Clinical Research Cases*; Issues in Business Ethics; Springer: Cham, Switzerland, 2017; pp. 285–310, ISBN 978-3-319-62784-7.

39. Gray, R. Rob Social, environmental and sustainability reporting and organisational value creation? Whose value? Whose creation? *Account. Audit. Account. J.* **2006**, *19*, 793–819. [CrossRef]

40. Amonarriz, C.A.; Landart, C.I.; Cantín, L.N. Cooperatives proactive social responsibility in crisis time: How to behave? *REVESCO Rev. Estud. Coop.* **2017**. [CrossRef]

41. Richez-Battesti, N.; Leseul, G. Cooperative Banks in France: Emergence, Mutations and Issues. In *Credit Cooperative Institutions in European Countries*; Contributions to Economics; Springer: Cham, Switzerland, 2016; pp. 55–81, ISBN 978-3-319-28783-6.

42. Demsetz, H. Toward a Theory of Property Rights. In *Classic Papers in Natural Resource Economics*; Palgrave Macmillan: London, UK, 1974; pp. 163–177. ISBN 978-1-349-41750-6.

43. Jensen, M.C.; Meckling, W.H. Theory of the firm: Managerial behavior, agency costs and ownership structure. *J. Financ. Econ.* **1976**, *3*, 305–360. [CrossRef]

44. Galbraith, J.K. *The New Industrial State*; Princeton University Press: Princeton, NJ, USA, 2015; ISBN 978-1-4008-7318-0.

45. Jensen, M.C. Value Maximization, Stakeholder Theory, and the Corporate Objective Function. *Bus. Ethics Q.* **2002**, *12*, 235–256. [CrossRef]

46. Parmar, B.L.; Keevil, A.; Wicks, A.C. People and Profits: The Impact of Corporate Objectives on Employees' Need Satisfaction at Work. *J. Bus. Ethics* **2017**. [CrossRef]

47. Asmild, M.; Zhu, M. Controlling for the use of extreme weights in bank efficiency assessments during the financial crisis. *Eur. J. Oper. Res.* **2016**, *251*, 999–1015. [CrossRef]

48. Hughes, J.P.; Lang, W.W.; Mester, L.J.; Moon, C.-G.; Pagano, M.S. Do bankers sacrifice value to build empires? Managerial incentives, industry consolidation, and financial performance. *J. Bank. Financ.* **2003**, *27*, 417–447. [CrossRef]

49. Luo, Y.; Tanna, S.; De Vita, G. Financial openness, risk and bank efficiency: Cross-country evidence. *J. Financ. Stable* **2016**, *24*, 132–148. [CrossRef]

50. Reckwitz, A. Toward a Theory of Social Practices: A Development in Culturalist Theorizing. *Eur. J. Soc. Theory* **2002**, *5*, 243–263. [CrossRef]

51. Freeman, R.E.; Harrison, J.S.; Wicks, A.C.; Parmar, B.L.; Colle, S.D. *Stakeholder Theory: The State of the Art*; Cambridge University Press: Cambridge, MA, USA, 2010; ISBN 978-1-139-48411-4.

52. Kalmi, P. The Role of Stakeholder Banks in the European Banking Sector. In *Institutional Diversity in Banking*; Palgrave Macmillan Studies in Banking and Financial Institutions; Palgrave Macmillan: Cham, Switzerland, 2017; pp. 33–50. ISBN 978-3-319-42072-1.

53. Monzon, J.L.; Chaves, R. The European social Economy: Concept and dimensions of the third sector. *Ann. Public Coop. Econ.* **2008**, *79*, 549–577. [CrossRef]

54. Transparency International. Corruption Perceptions Index-Transparency 278 International's Corruption. Available online: https://www.transparency.org (accessed on 5 September 2018).

55. Boston Consulting Group. Welfare data (Wealth-279 to-Well-Being Coefficient and Growth-to-Well-Being Coefficient, as Static and Improvement 280 Coefficient of Welfare). Available online: https://www.bcg.com/ (accessed on 1 July 2018).

56. Mauro, P. Corruption and Growth. *Q. J. Econ.* **1995**, *110*, 681–712. [CrossRef]

57. Stolp, C. Strengths and weaknesses of data envelopment analysis: An urban and regional perspective. *Comput. Environ. Urban Syst.* **1990**, *14*, 103–116. [CrossRef]

58. McGuire, J.B.; Sundgren, A.; Schneeweis, T. Corporate Social Responsibility and Firm Financial Performance. *Acad. Manag. J.* **1988**, *31*, 854–872. [CrossRef]

59. Preston, L.E.; O'Bannon, D.P. The Corporate Social-Financial Performance Relationship: A Typology and Analysis. *Bus. Soc.* **1997**, *36*, 419–429. [CrossRef]

60. Cowton, C.J.; San-Jose, L. On the Ethics of Trade Credit: Understanding Good Payment Practice in the Supply Chain. *J. Bus. Ethics* **2017**, *140*, 673–685. [CrossRef]

61. Gutiérrez-Nieto, B.; Serrano-Cinca, C.; Molinero, C.M. Social efficiency in microfinance institutions. *J. Oper. Res. Soc.* **2009**, *60*, 104–119. [CrossRef]

62. Everitt, P.D.; Landau, S.; Everitt, B.S.; Landau, S.L.; Everitt, D. *A Handbook of Statistical Analyses Using SPSS*; Taylor & Francis: Abingdon, UK, 2004; ISBN 978-1-58488-369-2.

Article

Specification Testing of Production in a Stochastic Frontier Model

Xu Guo [1], Gao-Rong Li [2], Michael McAleer [3,4,5,6,7] and Wing-Keung Wong [8,9,10,11,*]

[1]   School of Statistics, Beijing Normal University, Beijing 100875, China; xguo12@bnu.edu.cn
[2]   Beijing Institute for Scientific and Engineering Computing, Beijing University of Technology, Beijing 100124, China; ligaorong@bjut.edu.cn
[3]   Department of Finance, College of Management, Asia University, Taichung 41354, Taiwan; michael.mcaleer@gmail.com
[4]   Discipline of Business Analytics, University of Sydney Business School, Sydney NSW 2006, Australia
[5]   Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands
[6]   Department of Economic Analysis and ICAE, Complutense University of Madrid, 28040 Madrid, Spain
[7]   Institute of Advanced Sciences, Yokohama National University, Yokohama 240-8501, Japan
[8]   Department of Finance and Big Data Research Center, Asia University, Taichung 41354, Taiwan
[9]   Department of Medical Research, China Medical University Hospital, Taichung 41354, Taiwan
[10] Department of Economics and Finance, Hang Seng Management College, Sha Tin 999077, Hong Kong, China
[11] Department of Economics, Lingnan University, Tuen Mun 999077, Hong Kong, China
*   Correspondence: wong@asia.edu.tw

Abstract: Parametric production frontier functions are frequently used in stochastic frontier models, but there do not seem to be any empirical test statistics for the plausibility of this application. In this paper, we develop procedures to test whether or not the parametric production frontier functions are suitable. Toward this aim, we developed two test statistics based on local smoothing and an empirical process, respectively. Residual-based wild bootstrap versions of these two test statistics are also suggested. The distributions of technical inefficiency and the noise term are not specified, which allows specification testing of the production frontier function even under heteroscedasticity. Simulation studies and a real data example are presented to examine the finite sample sizes and powers of the test statistics. The theory developed in this paper is useful for production managers in their decisions on production.

Keywords: production frontier function; stochastic frontier model; specification testing; wild bootstrap; smoothing process; empirical process; simulations

JEL Classification: C0; C13; C14; D81

1. Introduction

Since the seminal works of [1,2], stochastic frontier analysis (SFA) has been a very appealing and popular approach for studying productivity and efficiency analysis. Greene [3] extended the stochastic frontier model by allowing the one-sided component of the disturbance to have a two-parameter gamma distribution rather than the less-flexible half-normal distribution. Greene [4] extended the model further by using a nonlinear specification. For an up-to-date introduction and literature review, see [5,6].

Consider the following SFA model:

$$Y = m(X) - U + V, \tag{1}$$

where Y is the log of output, X is the log of inputs of dimension p, $m(\cdot)$ is an unknown smooth production frontier function, U is the inefficiency term, and V represents random noise. Assume that the positive random variable, U, and the symmetric noise term, V, are conditionally independent, given the inputs X, and $E(V|X) = 0$.

Parametric SFA models specify the functional form of the production frontier function, $m(\cdot)$, as well as the distributions of the inefficiency term, U, and the independent noise, V. A fully parametric SFA framework sacrifices flexibility, and has been criticized as a major deficiency of SFA models (see details in [7]).

Some authors have discussed how to test the distributional assumptions on U and/or V. For instance, Wang et al. [8] developed the Pearson χ^2 and Kolmogorov–Smirnov tests for the distribution of U. Chen et al. [9] proposed a centered residuals-based method of moments to test the distributional assumptions on both U and V (see also [10–13]). However, it should be noted that all these procedures are based on the assumed parametric form of the production frontier function. If the parametric assumption on $m(\cdot)$ is not valid, the conclusions can be inaccurate and misleading.

On the other hand, there have been attempts to reduce the parametric restrictions on the production frontier function. Fan et al. [14] introduced the quasi-likelihood method, where the production frontier is not specified, but distributional assumptions are imposed on the stochastic components. Kumbhakar et al. [15] proposed a local maximum likelihood method but without parametric assumptions on the production frontier function, while using semi-parametric assumptions about U and V.

Recently, Simar et al. [16] developed a nonparametric least squares method to avoid the high computational complexity involved in the local maximum likelihood method in [15]. Another merit of the method of [16] is that only local distributional assumptions on U are needed, although symmetry is still necessary for V. Nonetheless, it should be realized that the methods discussed above would not be necessary if the hypothetical parametric model was satisfied. Studying the "wrong skewness phenomenon" in stochastic frontiers (SF), Bonanno et al. [17] proposed a more general and flexible specification of the SF model by introducing dependences between the two error components and asymmetry of the random error.

The studies discussed above call for the specification testing of the production frontier function. Parametric specifications for the frontier are appealing because they offer easy economic interpretation of the production process. Furthermore, due to well-established theories, easy computation, and interpretation, parametric SFA models have been dominant in the area of productivity and efficiency analysis. Specification testing can also be used to validate the accuracy of some production theory, such as Cobb–Douglas, CES, Translog, and related functions. There is literature on specification testing for conventional regression models (see [18] for a useful review). However, it would seem that there is as of yet no analysis that discusses this problem for SFA models.

In this paper, we aim to develop procedures to test whether the production frontier function can be described by some known parametric functions. To be precise, the null hypothesis is given as:

$$H_0 : m(X) = g(X, \beta_0), \tag{2}$$

for some β_0 against the alternative hypothesis:

$$H_1 : m(X) \neq g(X, \beta), \tag{3}$$

for any β, where $g(X, \beta)$ is a known smooth function with unknown d-dimensional parameter β.

Two test statistics are proposed, based on local smoothing and global smoothing, respectively. To apply these two test statistics in practice, we suggest the residual-based wild bootstrap. A merit of our procedure is that, even under heteroscedasticity, the test statistics can still detect the alternative hypothesis efficiently. To the best of our knowledge, this is a novel contribution to the

literature. The theory developed in this paper is useful for production managers in their decisions on production [19].

The remainder of the paper is organized as follows. In Section 2, we construct the test statistics and describe the residual-based wild bootstrap. In Section 3, simulation results are reported to examine the finite sample performance of the test statistics. An empirical application is given in Section 4, and Section 5 concludes the paper.

2. Test Statistics

To focus on specification testing of the production frontier function, we first discuss the estimation procedures for the parametric SFA model without specific distributional assumptions on U and V.

2.1. Estimation

Let $\mu_U(X) = E(U|X)$, $\epsilon = V - U + \mu_U(X)$, and $r_1(X) = Y - \epsilon$. Note that $E(\epsilon|X) = 0$ always holds. We can then rewrite model (1) under the null hypothesis as follows:

$$Y^1 = Y + \mu_U(X) = g(X, \beta) + \epsilon.$$

For the data set (Y^1, X), the model is the traditional parametric regression model. If we can obtain the value of $\mu_U(X)$, then we can estimate the parameter β by using nonlinear least squares based on (Y^1, X). Thus, the most important and difficult part is how to estimate $\mu_U(X)$. To achieve this goal, we adopt the approach that was recently proposed by [16].

Under the null hypothesis, model (1) can also be rewritten as:

$$Y = r_1(X) + \epsilon,$$

where $E(\epsilon|X) = 0$ still holds, which is the standard nonparametric regression model. We can obtain the estimator of $r_1(X)$, $\hat{r}_1(X)$, by using nonparametric methods such as kernels, local polynomials, and/or splines. Although there are several nonparametric methods for regression models, in the following we focus on kernel-type estimators given by $\hat{r}_1(x) = \sum_{i=1}^{n} W_{ni}(x)Y_i$, with:

$$W_{ni}(x) = \frac{K_h(x - X_i)}{\sum_{j=1}^{n} K_h(x - X_j)},$$

and $K_h(\cdot) = K(\cdot/h)/h^p$, with $K(\cdot)$ the kernel function, and h being the bandwidth.

Under the symmetry assumption on V, and the conditional independence of U and V given X, we have the following:

$$
\begin{aligned}
E(\epsilon^2|X) &= var_U(X) + var_V(X), \\
E(\epsilon^3|X) &= -E[(U - \mu_U(X))^3|X],
\end{aligned}
$$

where $var_U(X)$ and $var_V(X)$ denote the conditional variances of U and V given X, respectively.

Denote $r_j(X) = E(\epsilon^j|X)$ for $j = 2$ and 3. After estimation of $r_1(X)$, we can obtain the residuals, $\hat{\epsilon} = Y - \hat{r}_1(X)$. By adopting appropriate nonparametric techniques, we can estimate the functions $r_j(X)$ for $j = 2$ and 3 consistently. Define:

$$\hat{r}_j(x) = \sum_{i=1}^{n} W_{ni}(x)(Y - \hat{r}_1(X_i))^j,$$

for $j = 2$ and 3. Note that if $\mu_U(X)$ is a function of $E[(U - \mu_U(X))^3|X]$, then we can easily estimate $\hat{r}_3(X)$. To achieve this goal, local parametric assumptions on the types of distributions of $U|x$ are necessary.

Assume that $U|x \sim |N(0, \sigma_U^2(x))|$ and that, conditionally on X, U and V are independent, which is the same paradigm as in [15]. As a result, we have:

$$\mu_U(X) = E(U|X) = \sqrt{\frac{2}{\pi}}\sigma_U(X),$$

$$E(\epsilon^2|X) = \frac{\pi - 2}{\pi}\sigma_U^2(X) + var_V(X),$$

$$E(\epsilon^3|X) = \sqrt{\frac{2}{\pi}}\left(1 - \frac{4}{\pi}\right)\sigma_U^3(X) \le 0.$$

From the above equations, we can obtain the following:

$$\hat{\sigma}_U(X) = \max\left\{0, \left[\sqrt{\frac{\pi}{2}}(\frac{\pi}{\pi - 4})\hat{E}(\epsilon^3|X)\right]^{1/3}\right\},$$

$$\hat{\mu}_U(X) = \sqrt{\frac{2}{\pi}}\hat{\sigma}_U(X)$$

(for further details, see [16]).

After estimating $\hat{\mu}_U(X)$, we can estimate β by using nonlinear least squares based on the data points, $\{(\hat{Y}_i^1, X_i)|i = 1, \cdots, n\}$. Defining $\hat{Y}_i^1 = Y_i + \hat{\mu}_U(X_i)$, let $\epsilon_0 = Y^1 - g(X, \beta)$ to obtain the residuals under the null hypothesis, $\hat{\epsilon}_{0i} = \hat{Y}_i^1 - g(X_i, \hat{\beta})$.

2.2. Construction

Under the null hypothesis, we can easily obtain:

$$E(\epsilon_0|X) = E(Y + \mu_U(X) - g(X, \beta)|X) = E(g(X, \beta) + V - U + \mu_U(X) - g(X, \beta)|X) = 0,$$

while under the alternative hypothesis, we obtain:

$$
\begin{aligned}
E(\epsilon_0|X) &= E(Y + \mu_U(X) - g(X, \beta)|X) = E(m(X) + V - U + \mu_U(X) - g(X, \beta)|X) \\
&= m(X) - g(X, \beta) \ne 0.
\end{aligned}
$$

The above observations form the basis of the construction of the new test statistics. We introduce the local smoothing-based test statistic. Note that under the null hypothesis, we have:

$$E(\epsilon_0 E(\epsilon_0|X)f(X)) = E[E^2(\epsilon_0|X)f(X)] = 0,$$

where $f(X)$ is the density function of X. Under the alternative hypothesis, the first term in the above equation must be positive. Thus, the empirical counterpart of this term can be used as the test statistic. By using the leave-one-out kernel estimator of $f(X)$ and $E(\epsilon_0|X)$, the following test statistic is constructed:

$$T_{n1} = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j \ne i}^{n} K_h(X_i - X_j)\hat{\epsilon}_{0i}\hat{\epsilon}_{0j}.$$

The type of test statistic given above is introduced in [20], and was proposed independently by [21]. In classical regression models, it can be shown that the distribution of T_{n1} converges to a centered normal as $n \to \infty$. However, we should note that in the context of the SFA model, the asymptotic properties of T_{n1} can be complex due to the existence of the term $\mu_U(X)$. To formally study the asymptotic properties of T_{n1}, we need to investigate the impact of the nonparametric estimation of $\hat{\mu}_U(X)$ on the estimation of β explicitly. In this paper, we focus on investigating the numerical performance of T_{n1}, and leave the theoretical project for future research.

We can construct an empirically-based test statistic. Note that under the null hypothesis, the following equation holds:

$$E(\epsilon_0 I(X \le x)) = 0, \quad \forall x \in \mathbb{R}^p.$$

This motivates the construction of the residual-based empirical process, as follows:

$$R_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\epsilon}_{0i} I(X_i \le x).$$

Then, the Cramér–von Mises-type test statistic can be defined by:

$$T_{n2} = \int (R_n(x))^2 dF_n(x), \tag{4}$$

where $F_n(x)$ is the empirical distribution based on $\{X_1, X_2, \cdots, X_n\}$.

Similarly, in classical regression models, it can be shown that the defined empirical process $R_n(x)$ converges to a centered continuous Gaussian process, and the test statistic converges to the functional of this Gaussian process (see details in [22]), but the covariance function of the Gaussian process would be changed. We leave the formal theoretical analysis for future research.

We follow the residual-based wild bootstrap method (see details in [23]) to determine whether to reject the null hypothesis using the following steps:

Step 1. Obtain $\hat{\mu}_U(X)$, $\hat{\beta}$, and $\hat{\epsilon}_0$ by using the approach proposed in Section 2.1, and then construct $T_{ni}, i = 1, 2$, as in Section 2.2.

Step 2. Generate bootstrap observations, $Y_i^* = g(X_i, \hat{\beta}) - \hat{\mu}_U(X_i) + \hat{\epsilon}_{0i} \times e_i$. Here $\{e_i\}_{i=1}^{n}$ is a sequence of i.i.d. random variables with zero mean, unit variance, and independent of the sequence $\{Y_i, X_i\}_{i=1}^{n}$. Usually, $\{e_i\}_{i=1}^{n}$ can be chosen to be i.i.d. Bernoulli variates with:

$$P(e_i = \frac{1 - \sqrt{5}}{2}) = \frac{1 + \sqrt{5}}{2\sqrt{5}}, \qquad P(e_i = \frac{1 + \sqrt{5}}{2}) = 1 - \frac{1 + \sqrt{5}}{2\sqrt{5}}.$$

Step 3. Let $T_{ni}^*, i = 1, 2$ be defined similarly as $T_{ni}, i = 1, 2$, based on the bootstrap sample, $\{Y_i^*, X_i\}_{i=1}^{n}$.

Step 4. Repeat Steps 2 and 3, B times, and calculate the p-value as $p_i^B = \#\{T_{ni}^* > T_{ni}\}/B$.

3. Simulations

We now perform simulations to examine the finite sample performance of the proposed test statistics.

Study 1

$$H_{11} : Y = 5 + 5X + a \exp\{X^2\} - U + V,$$
$$H_{12} : Y = 5 + 5X + a \sin\{4\pi X\} - U + V.$$

The value $a = 0$ corresponds to the null hypothesis, and $a \ne 0$ to the alternative. In the above models, we take $X \sim U(0, 1)$, $U \sim |N(0, 1)|$, and $V \sim N(0, \sigma_V^2)$, where $\sigma_V = 0.75 \times \sqrt{(\pi - 2)/\pi}$. For the models, under the null hypothesis, $a = 0$, this is Example 1 in [15]. For H_{11}, the sample size is taken to be 100, and $a = \{0.0, 0.3, \cdots, 1.5\}$ to examine the size and power performance of the proposed test statistics, T_{n1} and T_{n2}. For H_{12}, we consider $n = 50$ and 100, and the sequence of a is taken to be $a = \{0.0, 0.2, \cdots, 1.0\}$.

In the simulation study, the number of replications was 2000. For each replication, $B = 500$ bootstrapped samples were generated. In the nonparametric regression estimation, the kernel function was taken to be $K(u) = 15/16(1 - u^2)^2$, if $|u| \le 1$; and 0, otherwise. The bandwidth was taken to be

$h = \hat{\sigma}(X) \times n^{-1/5}$ for simplicity, where $\hat{\sigma}(X)$ is the empirical estimator of the standard deviation of X. The nominal level of α was set at 0.05.

The simulation results are presented in Table 1.

Table 1. Simulated sizes and powers of the proposed test statistics T_{n1} and T_{n2} for Study 1.

H_{11}	$n = 100$	
a	T_{n1}	T_{n2}
0.0	0.0490	0.0530
0.3	0.0730	0.0950
0.6	0.1370	0.2370
0.9	0.2685	0.4170
1.2	0.4255	0.6430
1.5	0.6445	0.8400

H_{12}	$n = 50$		$n = 100$	
a	T_{n1}	T_{n2}	T_{n1}	T_{n2}
0.0	0.0510	0.0480	0.0540	0.0450
0.2	0.1240	0.0770	0.1920	0.1390
0.4	0.3590	0.2100	0.7010	0.4280
0.6	0.7190	0.4060	0.9640	0.8410
0.8	0.9170	0.6880	0.9990	0.9840
1.0	0.9790	0.8550	1.0000	0.9980

From the table, we have the following observations. First, for all situations considered, the empirical sizes of the two test statistics were all close to the nominal level. This implies that the proposed test statistics had accurate size. Second, when we consider empirical power, we can see clearly that the proposed tests were very sensitive to the alternative, such that when the value of a increased, power increased quickly. For model H_{11}, the second test statistic, T_{n2}, had higher power than the first test statistic, T_{n1}. However, for H_{12}, T_{n1} was more powerful. For model H_{12}, when the sample size was $n = 100$, the power performance of both tests improved compared with sample size $n = 50$.

Study 2

Consider the same models as in Study 1, but now introduce heteroscedasticity in the distribution of the technical inefficiency. Here, we have $U|X = x \sim |N(0, (1 + x)^2)|$. We should note that under the null hypothesis, $a = 0$, is Example 2 in [15]. This study investigates the impact of heteroscedasticity on the performance of the two proposed test statistics. Other settings are the same as in Study 1.

The simulation results are shown in Figure 1. For comparison, we also plot the simulation results of these two test statistics in Study 1.
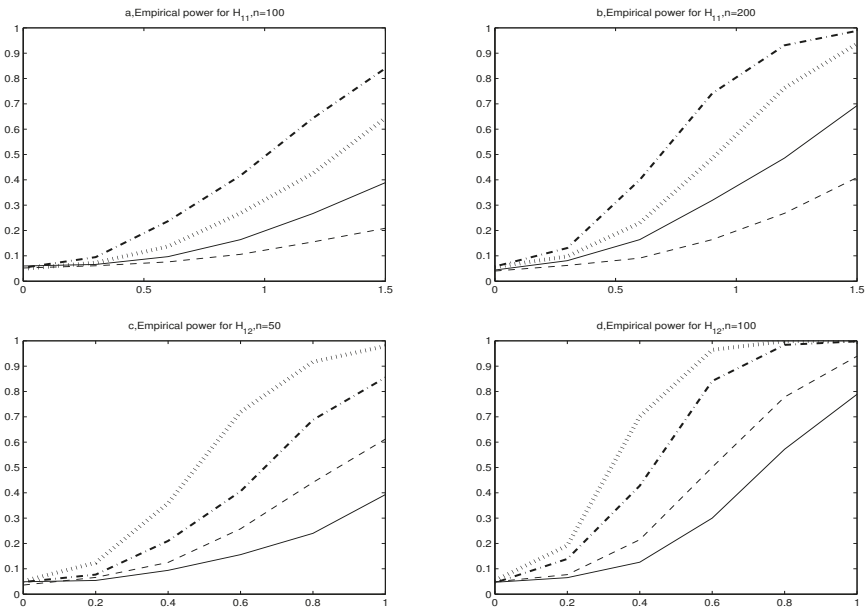
Figure 1. Powers of test statistics with H_{11} and $n = 100$ (**top-left corner**), H_{11} and $n = 200$ (**top-right corner**), H_{12} and $n = 50$ (**lower-left corner**), and H_{12} and $n = 100$ (**lower-right corner**), respectively. The dashed, dotted, solid, and dot–dashed lines represent the results of T_{n1} for Study 2 and Study 1, and T_{n2} for Study 2 and Study 1, respectively.

From this figure, we conclude that the powers of the two test statistics decreased significantly compared with the results in Study 1. This suggests that heteroscedasticity in the distribution of the technical inefficiency can have a negative impact on power performance. We can also see that for H_{11}, T_{n2} performed better than T_{n1}, while for H_{12}, T_{n1} was more powerful. These observations suggest that the two new test statistics should be viewed as complementary to each other.

4. Empirical Application

A rice production data set is available online, as described in the Preface of [24] (p. xvi, further details on the data can be found in Appendix 2 of [24]). The data set was recently analyzed in [8] to calculate goodness-of-fit tests for the distribution of technical inefficiency. Here we use this data set to check whether the Cobb–Douglas model is plausible.

Following [8,24], three inputs (area, labor, and fertilizer) and one output (tons of freshly threshed rice) were used, denoted by $X = (X_1, X_2, X_3) = (AREA, LABOR, NPK)$, and $Y = PROD$, respectively. The Cobb–Douglas model is given as follows:

$$\ln Y = \beta_0 + \sum_{i=1}^{3} \beta_i \ln X_i - U + V.$$

In our context, the null hypothesis is:

$$H_0 : m(X) = \beta_0 + \sum_{i=1}^{3} \beta_i \ln X_i.$$

For sample size $n = 344$, the values of $T_{ni}, i = 1, 2$, were 1.8062 and 616.5035, and the corresponding p-values were 0.160 and 0.774, respectively. Since both p-values were larger than 0.05, a Cobb–Douglas model is plausible. This implies that for the data set we used in our illustration, the relationship between the log output and log inputs can be considered as linear.

5. Concluding Remarks

Though SFA models have been used widely in many disciplines (e.g., economics, finance, and statistics), a formal specification testing procedure for the production frontier function has not been available. This paper develops two new test statistics by adopting local smoothing and global smoothing methods, respectively.

The asymptotic properties of the two test statistics under the null hypothesis, fixed alternative hypothesis, and local alternative hypothesis have not been investigated. The existence of the inefficiency term, U, makes the analysis complicated. We leave these interesting and important theoretical studies to future research.

Without explicit asymptotic distributions under the null hypothesis, we must rely on resampling approaches to calibrate the critical values. To this end, the residual-based wild bootstrap is suggested. The new proposed test statistics allow specification testing of the production frontier function, even under heteroscedasticity. The simulation studies showed that the sizes of the two test statistics are quite close to the nominal level, and that the powers are also satisfactory—even when the sample size is relatively small ($n = 50$). The theory developed in this paper is useful for production managers (see details in [25–27]) in their decisions on production [19] and for investors [28] in their decision making in their investment.

Model building is always a key concern for theoretical and practical studies. In this paper, we investigate whether a parametric production frontier function is suitable in the analysis. Lai et al. [29] considered the model selection criterion for the stochastic frontier models. Later on, Lai et al. [30] suggested using the model-averaged estimator based on the multimodel inference to estimate stochastic frontier models. Parmeter et al. [31] also suggested the use of this approach.

References

1. Aigner, D.J.; Lovell, C.A.K.; Schmidt, P. Formulation and estimation of stochastic frontier models. *J. Econom.* **1977**, *6*, 21–37.
2. Meeusen, W.; Broeck, J.V.D. Efficiency estimation from Cobb-Douglas production functions with composed error. *Int. Econ. Rev.* **1977**, *18*, 435–444.
3. Greene, W.H. A Gamma-distributed stochastic frontier model. *J. Econom.* **1990**, *46*, 141–163.

4. Greene, W.H. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *J. Econom.* **2005**, *126*, 269–303.

5. Kumbhakar, S.; Lovell, C.A.K. *Stochastic Frontier Analysis*; Cambridge University Press: Cambridge, UK, 2000.

6. Fried, H.; Lovell, C.A.K.; Schmidt, S. *The Measurement of Productive Efficiency and Productivity Change*; Oxford University Press: New York, NY, USA, 2008.

7. Simar, L.; Wilson, P. Statistical approaches for nonparametric frontier models: A guided tour. *Int. Stat. Rev.* **2015**, *83*, 77–110.

8. Wang, W.S.; Amsler, C.; Schmidt, P. Goodness of fit tests in stochastic frontier models. *J. Prod. Anal.* **2011**, *35*, 95–118.

9. Chen, Y.T.; Wang, H.J. Centered-residuals-based moment tests for stochastic frontier models. *Econom. Rev.* **2012**, *31*, 625–653.

10. Schmidt, P.; Lin, T.F. Simple tests of alternative specifications in stochastic frontier models. *J. Econom.* **1984**, *24*, 349–361.

11. Coelli, T.J. Estimators and hypothesis tests for a stochastic frontier function: A Monte Carlo analysis. *J. Prod. Anal.* **1995**, *6*, 247–268.

12. Lee, L.F. A test for distributional assumptions for the stochastic frontier functions. *J. Econom.* **1983**, *22*, 245–267.

13. Kopp, R.J.; Mullahy, J. Moment-based estimation and testing of stochastic frontier models. *J. Econom.* **1990**, *46*, 165–183.

14. Fan, Y.; Li, Q.; Weersink, A. Semiparametric estimation of stochastic production frontier. *J. Bus. Econ. Stat.* **1996**, *14*, 460–468.

15. Kumbhakar, S.C.; Park, B.U.; Simar, L.; Tsionas, E.G. Nonparametric stochastic frontiers: A local likelihood approach. *J. Econom.* **2007**, *137*, 1–27.

16. Simar, L.; Van Keilegom, I.; Zelenyuk, V. Nonparametric least squares methods for stochastic frontier models. *J. Prod. Anal.* **2017**, *47*, 189–204.

17. Bonanno, G.; De Giovanni, D.; Domma, F. The 'wrong skewness' problem: A re-specification of stochastic frontiers. *J. Prod. Anal.* **2017**, *47*, 9–64.

18. González-Manteiga, W.; Crujeiras, R.M. An updated review of goodness-of-fit tests for regression models. *Test* **2013**, *22*, 361–411.

19. Tsekouras, K.; Chatzistamoulou, N.; Kounetas, K. Productive performance, technology heterogeneity and hierarchies: Who to compare with whom. *Int. J. Prod. Econ.* **2017**, *193*, 465–478.

20. Zheng, J.X. A consistent test of functional form via nonparametric estimation techniques. *J. Econom.* **1996**, *75*, 263–289.

21. Fan, Y.; Li, Q. Consistent model specification tests: Omitted variables and semiparametric functional forms. *Econometrica* **1996**, *64*, 865–890.

22. Stute, W. Nonparametric model checks for regression. *Ann. Stat.* **1997**, *25*, 613–641.

23. Stute, W.; Gonzáles-Manteiga, W.; Presedo-Quindimil, M. Bootstrap approximation in model checks for regression. *J. Am. Stat. Assoc.* **1998**, *93*, 141–149.

24. Coelli, T.J.; Prasada Rao, D.S.; O'Donnell, C.J.; Battese, G.E. *An Introduction to Efficiency and Productivity Analysis*, 2nd ed.; Springer: New York, NY, USA, 2005.

25. Guo, X.; Wong, W.K.; Xu, Q.F.; Zhu, L.X. Production and hedging decisions under regret aversion. *Econ. Model.* **2015**, *51*, 153–158.

26. Moslehpour, M.; Pham, V.K.; Wong, W.K.; Bilgicli, I. E-purchase intention of Taiwanese consumers: Sustainable mediation of perceived usefulness and perceived ease of use. *Sustainability* **2018**, *10*, 234, doi:10.3390/su10010234.

27. Li, C.S.; Wong, W.K.; Peng, S.C.; Tsendsuren, S. The effects of health status on life insurance holding in 16 European countries. *Sustainability* **2018**, in press.

28. Li, Z.; Li, X.; Hui, Y.C.; Wong, W.K. Maslow portfolio selection for individuals with low financial sustainability. *Sustainability* **2018**, *10*, 1128, doi:10.3390/su10041128.

29. Lai, H.P.; Huang, C.J. Likelihood ratio tests for model selection of stochastic frontier models. *J. Prod. Anal.* **2010**, *34*, 3–13.

30. Lai, H.P.; Huang, C.J. Estimation of stochastic frontier models based on multimodel inference. *J. Prod. Anal.* **2012**, *38*, 273–284.

31. Parmeter, C.F.; Wan, A.T.; Zhang, X.Y. Model Averaging Estimators for the Stochastic Frontier Model, Working Paper. 2016. Available online: https://www.bus.miami.edu/_assets/files/repec/WP2016-09.pdf (accessed on 27 July 2018).

Does Business Group Affiliation Matter for Superior Performance? Evidence from Pakistan

Ishtiaq Ahmad [1], Judit Oláh [2,*], József Popp [3] and Domicián Máté [4]

1. Institute of Accounting and Finance, Faculty of Economics and Business, Ihrig Károly PhD School, University of Debrecen, 4028 Debrecen, Hungary; ishtiaq.ahmad@econ.unideb.hu
2. Institute of Applied Informatics and Logistics, Faculty of Economics and Business, University of Debrecen, 4032 Debrecen, Hungary
3. Institute of Sectoral Economics and Methodology, Faculty of Economics and Business, University of Debrecen, 4032 Debrecen, Hungary; popp.jozsef@econ.unideb.hu
4. Institute of Accounting and Finance, Controlling Department, Faculty of Economics and Business, University of Debrecen, 4028 Debrecen, Hungary; mate.domician@econ.unideb.hu
* Correspondence: olah.judit@econ.unideb.hu; Tel.: +36-20-286-9085

Abstract: Business groups have been described as improving the value of the affiliated firms they control, which is often beyond the capability of standalone firms. The purpose of the current study is to analyze the financial performance of affiliates of diversified Pakistani business groups relative to standalone firms. The current study employs data from 284 Pakistani listed non-financial firms from 2008–2015. In order to test the hypotheses, two dependent variables are used, namely, accounting (Return on Assets (ROA)) and stock market (Tobin's Q) measures of performance. Specifically, this study probes and compares the performance measures of group member and standalone firms. The findings of the study suggest that business group memberships have statistically significant effects on accounting and stock market measures of firm performance. In addition, size and sales growth have an increasing effect on the performance of firms. We believe that business groups in Pakistan are efficient economic actors and can be considered responses to high transaction costs and market failures.

Keywords: business groups; financial performance; group-affiliated; institutional voids

1. Introduction

Due to economic liberalization and globalization, corporate firms understand the intense competition they face: they need to diversify risk in order to achieve economies of scope and scale. Companies have to search for new markets, leverage resources to gain a competitive edge, and intensify the connections between firms by mergers, investments, and cross-shareholdings. One appropriate way of achieving these goals is to form a business group. By forming a business group, the affiliated firms use collaborative efforts between member firms to acquire favorable financial and intangible resources and capabilities. In fact, business groups create economies of scale and scope in order to minimize their Transaction Costs and increase the efficiencies of asset allocation. These collaborative efforts result in the maximization of firms' value and financial performance [1].

In the literature of business groups, a well-defined and widely accepted definition of the business group is 'a set of legally independent firms bound together by some formal and informal ties' [2]. A business group is an organizational form, that is, a collection of officially declared independent firms, and these firms work under the common financial and administrative control of certain families [3]. This study follows the definition provided by Khanna and Yafeh [4] in relation to emerging markets, which considers emerging markets as a 'transactional battlefield', where buyers and sellers do not come

together comfortably due to a lack of the specialized intermediaries in the market that generally assist with and advise on transactions between counterparties. Rehman [5] provided an initial identification of group membership. Previous studies have also referenced the same source to identify membership of business groups in the context of Pakistan [6].

Furthermore, the business group is an accepted phenomenon in different countries of the world. It is recognized under different names in many countries, for example, chaebol in Korea, keiretsu in Japan, business houses in Indian, and the 'twenty-two families' of Pakistan [7]. White [8] also proposed that the economic influence of Pakistan is concentrated in 'the 22 families' when considering domestic economic issues.

The business group is an important business form that prevails in both developing and developed countries. In a normative assumption view, group affiliation should increase the value of affiliated firms in the context of developing countries [9]. On the other hand, based on the literature of Transaction Cost economics, Williamson [10] and Coase [11] proposed the opposite view of group membership's influence on firm performance. Hence, in the case of developed countries, group affiliation outcomes resulted in high Transaction Costs and negative corporate performance. Thus, an empirical question arises that motivates scholars to analyze whether or not group affiliation positively affects the financial performance of firms in emerging economies.

Accordingly, performance comparison outcomes are different in relation to standalone firms, for example, in India, Chile, Korea, and Turkey group affiliation improves the performance of member firms. Orbay and Yurtoglu [12] reported that, in Turkey, group affiliation can be seen to have improved the investment performance and market value of firms. Other studies focused on Korea (Chang and Choi [13], Chang and Hong [3]), and others on India, namely, Khanna and Palepu [14,15]. All of them argued that business groups can be a source of value in emerging markets because they effectively fill in the institutional voids resulting from market inefficiencies. However, the performance of Japanese Keiretsu member firms is lower than standalone firms. Moreover, in China, business group membership has no effect on accounting performance [16]. Most of the available literature refers to Khanna and Rivkin [17]. Thus, in emerging economies research studies are based on the notion that groups are widely available in countries with weak institutional control and imperfect market conditions [18].

However, the existing literature is equivocal at best in presenting the impact of business groups on firm performance. Thus, how business group affiliation affects financial performance in emerging markets remains an open question. In order to fill this gap in the literature, the current study examines the performance outcomes of business group affiliation in Pakistan. We contend that Pakistan offers an excellent setting to test these phenomena for several reasons. Firstly, Pakistan represents an ideal case of the co-existence of standalone firms and large business groups, both contributing significantly to the country's economic activities. Saeed et al. [19] documented that business groups account for a major part of the private sector of the economy and hold a leading edge in terms of overall economic development and political favors. In addition, since Pakistan became independent, the owners of several business groups have migrated from India and run their businesses in Pakistan (1947). Therefore, business groups have a long history and strong roots in the Pakistani economy.

Secondly, diversified business groups are common in most developing economies. However, their role is poorly understood in India and Pakistan. The only exception is White [8], who revealed a statistically insignificant difference between the profitability of group and non-group firms. Hence, there is a pressing need to fill this research gap. The current study is amongst the first to explore the effect of group affiliation on performance using the most recent data set of Pakistani firms. The main contribution of this study is to show that business group membership can be an inevitable organizational response to institutional voids which enables group-affiliated firms to grow and prosper in an uncertain economic environment. Pakistan's economy faces different challenges, including energy crises, terrorism incidents, and political interference and governance issues. This situation has impeded Pakistan's economic and trading activities, which has not only resulted in higher transaction

costs for the corporate sector but also caused problems in production cycles, which results in significant delays in fulfilling export orders around the globe. Consequently, economic growth has slowed, and demand for imports reduced, accompanied by declining tax collection and foreign direct investment. Considering these economic conditions, it is meaningful to compare and evaluate the performance of group-affiliated and standalone firms and find out whether business group affiliation is a panacea for firms operating in such an economic environment.

The rest of the paper is organized as follows: Section 2 discusses a review of the literature, together with theoretical perspectives and empirical studies, conducted in different countries with the objective of exploring the relationship between group affiliation and the performance of firms. Following this, Section 3 discusses the sources of data and the criteria applied in the selection of the sample. An appropriate methodology to investigate the relationship between variables is also explained. Section 4 discusses the results of the study in order to answer the question of whether group-affiliated firms are more profitable than standalone firms. Finally, Section 5 ends with a conclusion and suggestions for future studies.

2. Literature Review

2.1. Theoretical Framework

Khanna and Palepu [20] coined the term 'institutional voids', and described them as the lack of intermediaries that connect buyers and sellers for efficient economic transactions. Institutional voids may create hurdles or certain opportunities for specific elements of the market. This provides an alternative justification for the presence of business groups in emerging economies.

Importantly, empirical studies have emphasized that business groups offset institutional voids by internalizing product, capital, and labor markets [21–23]. Nevertheless, when product, labor, and capital markets suffer due to the failing of these institutions, such practices lead to high Transaction Costs and the business group is one approach adopted in order to fill these institutional voids [14].

Considering the significance of Institutional Voids, a growing number of studies exist in the literature, which emphasize the association between business group affiliation and the performance outcomes of firms. Institutional voids theory suggests that prevailing voids in labor, product, and capital markets will not affect all firms equally. Rather, such voids have a strong negative influence on the performance of standalone firms relative to group-affiliated firms, since group-affiliated firms receive various benefits from each other, such as loans, debt guarantees, equity investments, and internal business trade [24]. Accordingly, it would be a rational approach for business groups to trade internally, to respond to market failures by protecting group-member firms from unusual external shocks to minimize risk and to increase performance.

Lee, Peng, and Lee [25] argued that during an institutional transition phase, the formal rules and regulations change, and increasing costs and uncertainty are expected. In a meta-analysis based on 141 studies, Carney et al. [26] related business group relationship with performance in 28 countries. They reported that the cost of group membership marginally balances its benefits in the form of improved financial performance, and that there were performance deviations to a certain degree at the firm and group levels.

The Transaction Cost perspective is based on the idea that firms strive to minimize the cost of exchanging resources within the economic environment [11]. Business groups are justified on the basis of Transaction Cost Theory by focusing on the differences at the overall level of Transaction Costs across countries affected by institutional voids [15]. In accordance with this approach, the business group is the right structure to deal with certain market failures that increase the overall Transaction Costs of an economy in different areas (labor, capital, and product markets) [27]. In addition, Yiu et al. [28] argued that the Transaction Cost approach has become a familiar viewpoint when rationalizing business groups in developing economies.

Consistent with the theory, if the level of Transaction Costs is high in an economy, then more economic activities are assumed to be carried out through an internally created market, as compared to the external market in the case of lower Transaction Costs [29]. In line with this assumption, previous studies have shown that internal capital markets have played a key role in business groups. Examples of this are the study by Shin and Park [30] on Korean chaebols and the study by Hoshi et al. [31] on Japanese Keiretsu. Therefore, business groups provide an efficient framework to capitalize investment opportunities at low transaction costs by investing in new ventures and ensuring the efficient allocation of funds generated through the internal capital market, as well as the external capital market.

Internal capital markets not only lower financial constraints for group-member firms, but also keep providing capital at low interest rates with soft protective covenants. Hence, the creation of an internal capital market lowers dependence on external market capital, which in turn strengthens their position compared to standalone firms. Zattoni, Pedersen, and Kumar [32] took a sample of Indian firms and observed that, in the presence of market and formal institutional imperfections, business groups perform better financially than standalone firms. However, business groups disappoint when it comes to confirming their superior performance when markets become more efficient.

2.2. Hypotheses Development

Business groups can be witnessed in many forms and sizes, with their diversity featuring challenges over time. Meanwhile, proportional returns in terms of profit are recognized to a greater degree in developing countries, where labor and financial markets are imperfect. In the comprehensive study by Khanna and Rivkin [17] related to business group affiliation and corporate performance, based on a sample of 14 countries, the effects of business groups were seen to differ from 4.2% (Mexico) to 31.1% (Indonesia). Moreover, Chang and Hong [3] found that business group effects account for between 5.7% and 9.7% of Korean firms' performance; importantly, this effect disappeared over a long period. In addition, the intensity of the business group effect is greater in small-sized business groups.

Comparing country-specific findings conducted in India and Korea, different strengths of the effect of business group membership on the performance of firms were witnessed [17]. Previous study findings, which are commonly seen as being in favor of the positive outcomes of group membership, supported their conclusions regarding the capability of business groups to overcome institutional voids in emerging economies. In China, it has also been concluded that the effect of group membership is positive on firm value [33]. In addition, He et al. [16] have reported that in China, group membership has a low and significant effect on firm accounting value.

However, Khanna and Yafeh [2] observed a negative association between group membership and firm performance in half of the ten emerging economies in their sample. Jia et al. [34] showed that business groups may be parasites that expropriate minority shareholders in the group, or may be paragons that support transactions and operations in and outside of the group when facing difficult economic and institutional environments. The equivocal impact of group membership was observed in earlier studies, for example, Careny et al. [26], which proposed that the association between group membership and firm performance may be more complex than has previously been empirically and theoretically modelled. They found that the effect of group membership varies substantial among countries: it is positive in Chile, Colombia, Hong Kong, Indonesia, Sweden, and Turkey; while it turns negative in Nigeria, France, Japan, and South Korea; and insignificant in Belgium, China, India, Taiwan, Thailand, and the Philippines. In addition, Mursitama [35] found a negative effect of business group membership on the performance of Indonesian business groups. In addition, Ma et al. [36] provided evidence from 1119 publicly listed Chinese firms that group membership has a statistically significant and negative influence on firm performance. Table A1 provides a summary of these studies.

Essentially, this brief review of the literature has reported mixed findings, offering evidence for both positive and negative associations between group membership and performance. Thus, many opportunities exist to increase understanding of the relationship between group membership and

financial performance through greater scrutiny when institutional voids are more severe. Consistent with the theory, empirical evidence supports the hypothesis that firms affiliated with a group located in an emerging economy have a higher financial performance than standalone firms. The Institutional and Transaction Cost theories emphasize that business groups may add value to member firms by filling the voids left by the missing institutions that support the efficient working of markets [37]. Therefore, it is expected that group membership positively affects the performance of group-affiliated firms in Pakistan.

Hypothesis 1. *Firms affiliated with business groups are more profitable than standalone firms.*

Firm size is taken to represent the capacity of economies of scale and scope accumulating to large firms. If large group-affiliated firms capitalize these two measures, the size of the firm will positively affect the performance of firms. The size of a business group affects firm performance [15]. On the positive side, Baumol [38] has documented that firm size positively affects the performance of firms by arguing that the benefits of large firms derive from their market power and greater access to external capital markets. Chu [39], in the Taiwanese context, concluded that group membership in the case of large-sized business groups leads to better stock market performance. On the contrary, Samuels and Smyth [40] suggested a negative relationship between firm size and profitability.

In their study, Claessens et al. [41] also used a sample of 2000 firms from nine East Asian economies, empirically analyzing the interaction effect of group affiliation and size on the value of firms. The results of interaction terms between group affiliation and size are statistically insignificant. Recently, scholars have also applied the interaction effect between group affiliation and size on firm value and reported that the interaction term has a statistically significant and positive influence on firm value [42], since large firms receive more advantages from group membership, such as easy access to external capital markets and greater economies of scale and scope. Therefore, we anticipated that the large size of a firm moderates the relationship between group affiliation and financial performance.

Hypothesis 2. *The relationship between group affiliation and affiliate performance is positively moderated by the size of firms.*

Hadlock and James [43] proposed that firms choose debt financing compared to equity financing, predominantly because the owners of firms prefer the dilution of earnings to the dilution of ownership. Therefore, this study applied indicators of leverage in order to measure the level of debt carried by a firm to reflect the availability of capital raised [44,45]. A greater ratio of debt-to-assets increases the chances of financial distress and bankruptcy and thus limits a firm's capacity to financially support its investment opportunities by borrowing [46]. Therefore, a negative sign is predicted for leverage measures in connection with performance measures.

Hypothesis 3. *The relationship between group affiliation and affiliate performance is negatively moderated by the leverage of firms.*

This study makes an initial effort to address the issue by investigating the influence of sales growth on accounting and stock market measures of performance. We ask two questions: (i) Does sales growth positively affect performance? (ii) Is the positive impact of sales growth on the performance of group-affiliated firms higher or lower in case of group affiliation? Using a sample of Keiretsu member firms, Aoki [47] reported by that group affiliation does not facilitate higher sales growth rates. A review of the literature posits different findings, offering both positive [48–50] and negative associations [51] between growth and profitability. Pakistani business groups focused on the sales growth of firms, particularly when searching for new markets and moving into new business ventures. Thus, it is expected that firms affiliated with a business group gain more from sales growth relative to standalone firms.

Hypothesis 4. *The relationship between group affiliation and affiliate performance is positively moderated by the sales growth of firms.*

3. Data Sources and Methodology

3.1. Sources of Data

This study analyses a large sample of group-member firms and standalone firms listed on the Pakistan Stock Exchange. Previously, the Pakistan Stock Exchange was known as the Karachi Stock Exchange. Then, three stock exchanges, namely, the Karachi Stock Exchange, the Lahore Stock Exchange, and the Islamabad Stock exchange, merged to become the Pakistan Stock Exchange (PSX), on 11 January 2016. The sample data is collected from the State Bank of Pakistan-Financial Statements Analysis of Companies (Non-Financial). This data is administered and published by the State Bank of Pakistan (SBP), as the Central Bank of Pakistan. The document contains data from the financial statements of non-financial firms and this data is comparable to the annual reports submitted to the Securities and Exchange Commission of Pakistan (SECP). More importantly, firms in Pakistan have to report their data to the SECP annually, thus transparency and accuracy of data is also required.

3.2. Data Collection and Sample Specification

Private limited firms have been excluded from the sample due to a lack of available data. The study sample also excludes financial, real estate, and utility firms, and firms that are subsidiaries of foreign firms. Financial services firms are not part of the sample since their accounting scheme is not compatible with that of firms in other industries. The returns of financial firms are not similar and cannot be compared with other sectors of the economy [52]. This study sample includes only public limited firms from the private sector in Pakistan. Thus, following various studies, firms operating in the financial services sector, firms affiliated with multinational patents, and firms that are owned partially or fully by the government are not part of the study sample [53].

Based on these facts, the study covers 284 public limited firms listed on the Pakistan Stock Exchange (PSX) for the period 2008–2015. The study sample consists of 284 firms, 143 (50.35%) of which are affiliated with a business group and 141 (49.65%) of which are standalone firms. The total numbers of observations in this study is 2272. In food and tobacco industries, out of 35 firms, 16 are group-affiliated and 19 are standalone firms. More importantly, in the sample, 74 firms are active in basic industries, including petroleum, of which 38 are group-affiliated firms and 36 are standalone firms. The textile industry comprises the major share, with 1032 observations from 129 firms of which 56 are group-affiliated, and 73 are standalone firms.

3.3. Methodologies

This study is based on unbalanced panel data analyzed primarily by the pooled ordinary squares (OLS) regression method to estimate the relationship between dependent and independent variables. The pooled OLS regression is appropriate for examining the effect of group affiliation on the performance of group-member firms, and there are no unique attributes of individuals within the measurement set. In this case, group affiliation is a dummy variable.

Firstly, to compare the performance of group-member firms and standalone firms, an independent sample *t*-test is applied to mean differences. Then, pooled regression is estimated to empirically analyze the effect of group affiliation on the performance of member firms. Earlier studies related to the performance of business groups have applied the pooled regression estimation technique at a firm level [54–56]. The performance comparison of group firms and standalone firms is applied by using a dummy variable; thus, a value of 1 indicates that a firm is a member of a group, while zero indicates it is a standalone firm. Therefore, group membership is a dummy variable distinguishing between affiliated firms and standalone firms.

Based on the review of the literature and business group theories the main hypothesis of the study is to investigate whether group-member firms perform better financially than standalone firms do. It is assumed that in emerging economies business group membership positively affects the performance of group members. The study estimates model 1 and 2 using regression analysis to explore the effect of group membership on the financial performance of firms.

$$\text{ROA}_{i,t} = \beta_o + \beta_1 DGroup_{i,t} + \beta_2 SIZE_i + \beta_3 SGRW_{i,t} + \beta_4 LEV_{i,t} + \beta_5 DInd + \varepsilon_i \tag{1}$$

$$\text{Tobin's Q}_{i,t} = \beta_o + \beta_1 DGroup_{i,t} + \beta_2 SIZE_i + \beta_3 SGRW_{i,t} + \beta_4 LEV_{i,t} + \beta_5 DInd + \varepsilon_i \tag{2}$$

where the dependent variables are Return on Assets (ROA) and Tobin's Q. ROA refers to the accounting based performance of a firm, and measures earnings before interest and taxes divided by total assets. Tobin's Q represents the stock market measure of firm performance, which is estimated as the market value of equity and the book value of debt divided by the book value of assets. We used the natural logarithm transformation of Tobin's Q, since, the log-transformed Tobin's Q ratio has shown better statistical distribution properties than raw Tobin's Q ratio [57,58]. The Group Affiliation (DGroup) dummy is the variable of interest and is a time-invariant dummy variable, showing the membership of firms. SIZE is the natural logarithm of total assets. It indicates the size of the firm. Sales Growth (SGRW) is represented by the sales growth of the firm and represents current year sales minus last year sales divided by last year sales. Leverage (LEV) is the capital structure of a firm, that is, the total debt divided by total assets. DInd shows each of the listed branches at a two-digit level of SIC (see Table A2). Lastly, ε is the error term. This study introduces interactive (cross-effect) variables within baseline models 1 and 2.

In particular, all firm-level control variables used, such as size, growth, and leverage, are interacted with group-affiliated dummy variables to catch the group affiliation relationship. Therefore, models 5–10 analyze the interaction between group affiliation and control variables to determine their effect on the profitability of firms. Table 1 shows the definitions of each dependent and independent variable, with its source.

Table 1. The definitions and sources of the variables.

Variables (Acronyms)	Definitions	Sources
Return on Assets (ROA)	Earnings before interest and taxes divided by total assets	Financial Statement Analysis (SBP)
Tobin's Q	Market value of equity plus book value of debt divided by total assets	Pakistan Stock Exchange (PSX)
Group Affiliation (DGroup)	Dummy variable that takes value 1 if the firm is affiliated with a Pakistani business group, 0 otherwise	Rehman (2016) [5]
Leverage (LEV)	Total debt divided by total assets	Financial Statement Analysis (SBP)
Firm Size (SIZE)	Natural Logarithm of total assets	Financial Statement Analysis (SBP)
Sales Growth (SGRW)	(Current year sales + Last year sales) divided by Last year sales	Financial Statement Analysis (SBP)

4. Analysis and Results

4.1. Descriptive Statistics

The *t*-test statistics are used to analyze the differences in the means of group member and standalone firms' performance and control variables. It can be seen that group-affiliated firms have a significantly higher Return on Assets with a mean value of 5.008 than standalone firms with 1.663. The second performance factor is measured by Tobin's Q, which is used to estimate the market performance of firms. Group-member firms appear to have higher Tobin's Q ratios, with a mean value of 4.132, than standalone firms with 3.467. The comparison of performance measures between group-member firms and standalone firms is shown in Table 2.

Consequently, it is hypothesized that member firms are more profitable than standalone firms. In particular, the results of the *t*-test indicate that group firms are significantly more profitable in terms of accounting performance (ROA) and stock market performance (Tobin's Q) than standalone firms. Thus, it is suggested that group affiliation improves member firms' profitability. The performance difference is statistically significant at the 1% level. It can also be seen that group-affiliated firms are larger than standalone firms; as measured by total assets, the difference is statistically significant at the 1% level. In addition, in terms of growth—measured by current year sales minus last year sales divided by last year's sales—the difference between affiliated and unaffiliated firms is statistically significant at 5%. This difference explains the advantages of economies of scale and scope for group-member firms. Moreover, the difference in the total debt between group-affiliated and unaffiliated firms is also analyzed, with the debt level in relation to total assets higher in unaffiliated firms than in group-affiliated firms. The overall results reveal that higher profitability, a larger size, and a better solvency position are important determinants of business group affiliation.

Table 2. Comparison of the statistics of the key variables.

Variables	Entire Sample (n = 284)		Affiliated Firms (n = 143)		Standalone Firms (n = 141)		*t*-Statistics
	Mean	SD	Mean	SD	Mean	SD	
ROA	3.347	9.707	5.008	9.488	1.663	9.642	−8.335 ***
Tobin's Q	3.802	3.605	4.132	3.777	3.467	3.391	−4.411 ***
SIZE	14.339	2.541	14.947	2.700	13.723	2.204	−11.824 ***
SGRW	0.094	0.285	0.109	0.270	0.078	0.299	−2.598 **
LEV	0.724	0.848	0.612	0.576	0.838	1.043	6.397 ***

Source: authors' own estimations. *** significance at the 1% Level, ** significance at the 5% Level.

4.2. Correlation Analysis

Previous studies have provided empirical evidence that group affiliation improves member firms' performance [9]. Moreover, several studies have provided a positive correlation between group affiliation and accounting performance and the stock market performance of group-member firms [59]. In this study, the correlation between group affiliation and accounting performance and stock market performance is statistically significant at 5%.

Therefore, positive correlations with both performance measures support the first hypothesis that group affiliation positively affects member firms' performance when compared to standalone firms. Moreover, the correlation coefficient between group affiliation and the size of firms is 0.24, suggesting a moderate correlation between them (See Table 3). However, a negative correlation is observed between total debt and accounting performance and stock market performance measures. This suggests that an increasing level of debt decreases the financial performance and value of firms.

Table 3. Results of the pairwise correlation matrix of Return on Assets (ROA).

	DGroup	ROA	Tobin's Q	Size	Sales Growth	Leverage
DGroup	1					
ROA	0.1723 *	1				
Tobin's Q	0.0922 *	0.2999 *	1			
Size	0.2409 *	0.1962 *	0.1097 *	1		
Sales Growth	0.0545 *	0.3202 *	0.0314 *	0.1044 *	1	
Leverage	−0.1844 *	−0.3744 *	−0.1084 *	−0.0538 *	−0.0838 *	1

Source: authors' own estimations. * Significance at the 5% Level.

4.3. Regression Analysis

This section presents the results of regression analysis calculated by using pooled OLS regression, and the importance of group affiliation in terms of financial performance. Taking group affiliation as the main variable, the regression is performed between group affiliation and performance measures with and without considering control variables. Table 4 reports the results of baseline models 1 and 2 taking ROA and Tobin's Q as dependent variables. The results of the first hypothesis, regarding whether firms affiliated with business groups have higher accounting and stock market performance than standalone firms, are reported in columns (1) and (3) of Table 4.

Table 4. Regression results of Equations (1) and (2).

Variable	ROA		Tobin's Q	
	(1)	(2)	(3)	(4)
DGroup	3.345 ***	2.062 ***	0.193 ***	0.139 ***
	(8.34)	(5.30)	(5.11)	(3.82)
SIZE		0.386 ***		0.047 ***
		(5.02)		(6.59)
SGRW		9.558 ***		0.009
		(14.07)		(0.35)
LEV		−1.748 ***		−0.201 ***
		(−7.83)		(−9.52)
DInd		Yes		Yes
DYear		Yes		Yes
Intercept	1.663 ***	−2.233	0.900 ***	0.138
	(5.84)	(−0.91)	(33.58)	(1.25)
Companies	284	284	284	284
Observations	2272	2272	2272	2272
Adj. R²	0.0293	0.1861	0.011	0.1552
F-Value	69.48 ***	31.54 ***	26.12 ***	30.81 ***
Breusch-Pagan (BP) test	0.29	0.04	0.894	0.139
Variance Inflation Factor (VIF)	1.00	1.92	1.00	1.42
Shapiro-Wilk (SW) test	0.195	0.132	0.110	0.212

Source: authors' own estimations. *t*-values are reported in parentheses. *** significance at the 1% Level.

As is shown in columns (2) and (4) of Table 4 the results are reported with control variables. The results support the first hypothesis (H1) regarding the fact that group affiliation improves the firm performance of group-member firms. As shown in columns (1) and (3) of Table 3, for accounting and stock market performance measures, the effect of group affiliation is statistically significant ($p < 0.01$) and positive. The results indicate that group affiliation has a statistically significant positive influence

on firm profitability (3.345, *t*-value 8.34) and market performance (0.193, *t*-value 5.11). In addition, the results of group affiliation with control variables are also statistically significant. As shown in columns (2) and (4) of Table 4, the regression results with control variables support the first hypothesis (H1), that the coefficient of group affiliation has a positive effect on the accounting performance (2.062, *t*-value 5.30) and market performance of firms (0.139, *t*-value 3.82).

The results of control variables are also significant. Size has a statistically significant positive effect on the profitability (0.386, *t*-value 5.02) and market performance of the firm (0.047, *t*-value 6.59). Therefore, we can conclude that the size of a firm matters for its financial performance. Earlier, Lang, Ofek, and Stulz [60] reported the positive effect of growth on firm value. Therefore, it was expected that sales growth and size are positively associated with the value of the firm. The sales growth coefficient is statistically significant in the case of accounting based performance (9.558, *t*-value 14.07), but insignificant in the case of market based performance. Thus, it is implied that sales growth contributes positively to the profitability of firms, as is evidenced by the positive coefficient of the sales growth variable. Amongst other control variables, the coefficient of leverage has a statistically significant negative effect on the financial performance of firms. The results suggest that as debt ratio increases, the performance of the firm decreases.

We considered that the positive affect of group affiliation on member firms' financial performance is derived from different channels, such as internal capital markets, parent office globalization, marketing channels, and professional human resources. We might attribute our findings to the fact that external markets have been relatively less sophisticated in Pakistan. In order to avoid the constraints on arm-length lending, business groups are responsible for providing access to capital and obtaining funded through internal capital markets for investment in high-yielding opportunities. Moreover, the bond market is not mature in Pakistan due to high administration costs, and a lack of technological development, transparency, and liquidity; and the expectations of inflation and the regular devaluation of PKR currency have hindered foreign investment. Thus, the only source for debt financing is bank loans. It is important to mention that almost every large business group in Pakistan has its own bank, that is, it is able to arrange loans and bank guarantees. Hence, affiliated financial institutions, besides providing internal capital markets, create an advantage over external capital markets for group firms in the form of loan guarantees, low interest rates, and almost non-existent protective covenants. This mechanism of cross-subsidies improves the overall financial performance of group-member firms.

Another interesting factor of Pakistani business groups is that they have parent offices outside Pakistan. These parent offices facilitate in increasing export sales and investment and coordinate activities relating to the adoption of modern technology, as there is no government support for technology upgrading and research and development. Therefore, in collaboration with multinational firms they are able to use modern technology to increase the productivity of their group-affiliated firms. Considering the linkages with the international market, business groups provide a baseline for international exposure for member firms, including the access to international markets so that they can learn about and capitalize on market opportunities. Standalone firms cannot easily access these knowledge-based advantages. Therefore, group membership supports member firms in their transactions with international clients in foreign markets and attracts clients from a wider range of foreign markets than is the case with standalone firms.

In Pakistan, business groups promote group-wide advertising, which focuses on the overall image of a business group rather than highlighting an individual member firm. As a result, group-wide advertising also creates economies of scale and scope. An example of this is the Sitara Group's advertising. After the advertisement of each affiliate, there is a message from the Sitara group of companies, first emphasizing an individual member firm and then promoting the overall image of the business group. This message promotes the idea that the quality of their products is excellent. In addition, Pakistan is amongst the top exporters of textiles around the globe, which assists in the market positioning of the Sitara brand name in different industries such as textiles, chemical products,

and energy. Similarly, the Hundai Group's advertising highlights the idea that the manufacturer operates from 'from chip to ship'. Chang and Hong [24] found that group investment in advertising and R&D activities contribute to the economic performance of group-member firms.

Interestingly, the owners of business groups send their children abroad for higher education, preferably to English speaking countries such as UK, USA, Australia, and Canada. After completion of their education, they join the business group as a manager. Then, after five-years they are appointed as directors of different group-affiliated firms and work as interlocking directors. Eventually, they emerge as a professional human resource for business groups. Thus, from a human perspective, framing a sound internal management and control system is also critical for business groups in cases where the number of professional managers in the market is limited. As well as assisting in control, interlocking directors encourage member firms to share resources and the flow of information, which ultimately influences their performance in the group.

Therefore, interlocking directorates work as a tool to align objectives between the parent firm and group-member firms. Moreover, business groups have the capacity to appoint government officials as directors on their boards to support member firms in dealing comfortably with legal, monitoring, and enforcement issues. This indicates that business groups are capable of dealing with the voids prevalent in product, capital, and labor markets.

Note: the table above shows the results of baseline (Equations (1) and (2)) models using pooled regression. The sample period is from 2008 to 2015. There are two dependent variables, the first of which is the accounting based performance measure return on assets (ROA). This variable measures earnings before interest and taxes divided by total assets. The second dependent variable is a stock market based performance measure, Tobin's Q, measuring the market value of equity plus the book value of debt divided by the book value of total assets. The independent variables are Group dummies (DGroup), size of firms (SIZE), sales growth (SGRW), leverage (LEV), industry, and time dummies. DGroup is a dummy variable, where 1 denotes that a firm is affiliated with a business group and zero that it is not. Size is measured by the natural logarithm of total assets. Sales growth is measured by current year sales minus last year sales divided by last year sales. Leverage is measured by total liabilities divided by total assets. DInd shows the industry dummies at a two-digit level of SIC. DYear substitutes the year dummies between 2008 and 2015.

(BP-test): the Breusch and Pagan test is used to check heteroscedasticity in the linear regression models [61]. The VIF-test checks the multicollinearity in the independent variables, expressed as the variance inflation factor (VIF). Five considers a two-digit critical value of VIF. The *p*-values of the Shapiro-Wilk normality test offer standard tools for detecting any violation of standard regression assumptions. Each of the residuals has a normal distribution.

Note: (BP-test): the Breusch and Pagan test is used to check heteroscedasticity in the linear regression models [61]. The VIF-test checks the multicollinearity in the independent variables, expressed as the variance inflation factor (VIF). Five considers a two-digit critical value of VIF. The *p*-values of the Shapiro-Wilk normality test offer standard tools for detecting the violation of standard regression assumptions. Each of the residuals has a normal distribution at the 5% level.

In order to explore the possible interaction effects, the size, leverage, and sales growth variables are interacted with the main variable under examination, that is, group affiliation. In Table 5, the interaction DGroup × SIZE is investigated, to analyze its influence on the financial performance of firms. As shown in columns (5) and (8) of Table 5 the coefficient of the interaction term between group dummy and size is positive and statistically significant (ROA $\beta = 0.167$, *t*-value 7.06, Tobin's Q $\beta = 0.223$, *t*-value 10.77). The results support the second hypothesis, namely, that the relationship between group membership and affiliates is positively moderated by the size of firms. Thus, the results indicate that large firms receive more advantages from group membership, such as easy access to external capital markets and greater economies of scale and scope.

Claessens et al. [41] used a sample of 2000 firms from nine East Asian economies to empirically analyze the interaction effect of group affiliation and size on the value of firms. However, the results

of the interaction terms between group affiliation and size were statistically insignificant. Recently, other researchers have also applied the interaction effect between group affiliation and size to firm value [42] and reported that the interaction term has a statistically significant and positive influence on firm value.

In Table 5, to test hypothesis 3, the interaction between group affiliation and leverage (DGroup $\times$ LEV) is introduced. In line with our expectations, in columns (6) and (9) the coefficient of the interaction term between group dummy and leverage is negative and statistically significant for accounting (ROA $\beta = -2.512$, *t*-value -4.46.) and stock market measures of performance (Tobin's Q $\beta = -0.212$, *t*-value -4.54). It is implied that a high debt ratio negatively affects and lowers the performance of affiliated firms. In other words, a one unit increase in firms' leverage tends to decrease firms' profitability performance, and if there are two examined firms the affiliated firm has a better performance than the non-affiliated one. Hypothesis 3 is supported.

Table 5. Regression Results of Equations (1) and (2) with Using Interactive Variables.

Variable	ROA			Tobin's Q		
	(5)	(6)	(7)	(8)	(9)	(10)
DGroup $\times$ SIZE	0.167 *** (7.06)			0.223 *** (10.77)		
DGroup $\times$ LEV		−2.512 ** (−4.46)			−0.212 *** (−4.54)	
DGroup $\times$ SGRW			8.705 ** (9.36)			0.186 ** (2.08)
SIZE		1.397 *** (11.62)	0.952 *** (8.40)		0.044 *** (6.36) ***	0.039 *** (3.88)
SGRW	9.097 *** (14.06)	9.612 *** (14.1)		0.151 ** (2.58)	0.132 ** (2.16)	
LEV	−9.831 *** (−17.36)		−9.890 *** (−17.19)	−0.112 ** (−2.18)		−0.176 *** (−9.49)
DInd	Yes	Yes	Yes	Yes	Yes	Yes
DYear	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	6.221 ** (2.85)	−14.379 *** (−5.17)	−4.986 * (−1.83)	2.344 *** (10.70)	0.785 *** (6.56)	0.140 (0.94)
Companies	284	284	284	284	284	284
Observations	2272	2272	2272	2272	2272	2272
Adj. R^2	0.259	0.179	0.236	0.125	0.237	0.206
F-Value	50.62 ***	32.02 ***	44.97 ***	22.72 ***	45.23 ***	37.90 ***
Breusch-Pagan (BP) test	0.13	1.48	1.82	0.065	0.20	0.15
Variance Inflation Factor (VIF)	1.82	2.43	1.91	1.54	1.88	1.43
Shapiro-Wilk (SW) test	0.096	0.153	0.163	0.112	0.193	0.221

Source: authors' own estimations. *t*-values are reported in parentheses. *** significance at the 1% Level, ** significance at the 5% Level, * significant at the 10% Level.

In order to test hypothesis 4, the interaction between group affiliation and sales growth (DGroup $\times$ SGRW) is also shown. Columns 7 and 10 present the results of DGroup $\times$ SGRW. The coefficient of the interactive term is positive and statistically significant (ROA $\beta = 8.705$, *t*-value 9.36, Tobin's Q

$β$ = 0.186, t-value 2.08). Hypothesis 4 is also supported. The interaction between group affiliation and firm characteristics, such as the size of the firm, sales growth, and capital structure, is statistically significant for performance measures. Sales growth and the size of the group-affiliated firms have a greater influence on the financial performance of firms than they do with non-affiliated firms.

The multi-collinearity amongst the independent (interaction and other financial) variables are tested by the variance inflation factor (VIF) in each case and their maximum individual values are reported in Table 5. The VIF values for each regression coefficient ranged from a low of 1.00 to a high of 2.43, and indicated that the collinearity problem is controlled, in which one predictor variable in a multiple regression model can be linearly predicted from the others with a substantial degree of accuracy. Hence, there is no particularly collinearity amongst the independent interaction and other control variables. This suggests that the VIF values are at acceptable levels [62], and in this case there is no need for centering such interaction models [63]. All of them are included in the final model. The Breusch and Pagan tests are also applied in order to test the existence of heteroscedasticity. In the present study, the $χ^2$ values suggest that the statistics are at an acceptable level ($p > 0.05$) and there is no heteroscedasticity.

5. Discussion

Researchers have offered different views of business groups, portraying them as parasites, villains, and anachronisms, or as paragons, heroes, and avatars. Our study findings provide compelling evidence of a direct positive link between group membership and firms' financial performance. In addition, this relationship is found to be strongly moderated by firm specific factors such as size, leverage, and sales growth. In line with our expectations as stated in H1, we find that firms affiliated with business groups are more profitable than standalone firms. The findings are consistent with the earlier studies. For example, Chang and Choi [13] reported a positive effect of group affiliation on the performance of chaebol firms. Moreover, Chittoor, Kale, and Puranam [64] and Manikandan and Ramachandran [65] also found that group-member firms have a better accounting and stock market performance. In the context of an emerging economy, such as Pakistan, group-affiliated firms perform better financially than standalone firms do.

Specifically, our study contributes to the business group literature in three ways. Firstly, the financial performance of group-affiliated and standalone firms are investigated to show the influence of group membership on accounting and stock market measures. The results of our study show that the performance outcome of business group affiliation is contingent upon various firm characteristics, such as size, leverage, and sale growth. Thus we stress the need to study the impact of group affiliation in the presence of other factors that may shape the outcomes of business affiliation.

Secondly, business groups support member firms in avoiding the severe institutional voids prevalent in the emerging economy of Pakistan. The free flow of capital, sharing intangible resources (such as R&D and advertising), and interlocking directors within group member firms generate considerable economies of scale and scope. An interesting observation is that most of the group member firms are more mature and large, as these member firms have been operating since the independence of Pakistan. Importantly, they have their own financial and technical resources.

Lastly, we found that standalone firms are more locally oriented and less diversified than their group affiliated counterparts, which explains their survival and the performance discount they incur. Business groups collaborate in the form of international joint ventures which benefit member firms by offering access to new markets and advanced technology, increased capacity, and the sharing of risks. In the case of Pakistan, standalone firms do not perform on a par with their group affiliated peers.

6. Conclusions

This research paper seeks to provide empirical evidence on the effect of group membership on the performance of firms in Pakistan. By using a sample of 284 Pakistani firms as the research sample, this study suggests that group membership is beneficial for member firms. Moreover, the benefits of group membership are linked to the size of business-group firms. In the case of large group member firms, the effect of business group membership is more influential than it is with small group member firms.

This study compares the profitability of group-member firms with standalone firms using an independent sample *t*-test for mean differences. The results support the hypothesis that group-affiliated firms are more profitable compared to standalone firms. Moreover, the results of interaction terms are also statistically significant, which implies that the size and sales growth of group firms positively contribute to the financial performance of firms.

Thus, the findings of the study suggest two important explanations. First, like most developing economies, business groups are able to overcome the inefficiencies related to emerging markets, such as imperfections in the markets regarding product, capital, and labor [63]. Second, in emerging economies, poor judicial systems lead to low trust, making personal ties more important and trustworthy than trust institutions [66].

The results of this study have vitally important implications for practitioners—managers, macroeconomic policymakers, academicians and theorists. Specifically, weak governance tends to discourage private sector investment and reduce economic efficiency. Importantly, governance issues are significantly related to institutional voids. These institutional voids provide opportunities to groups to benefit and create advantages over standalone firms. These advantages are created through sharing financial resources and intangible resources (R&D, advertising), appointing interlocking directors, and collaborating with multinational firms. Thus, these measures enable group-affiliated firms to respond positively to institutional voids by making themselves a part of trustworthy networks in order to reduce financial risks.

Nevertheless, Chari and David [53] claimed that a negative relationship between pro-market reforms and the sustainability of superior profits exists in an emerging economy. The decline in the sustainability of superior profits also shows that pro-market reforms bring significant threats in addition to offering various opportunities, such as a greater availability of production factors and greater freedom to enter and operate businesses. The empirical results also supported a significant difference in the superior and sustainable economic performance among firms in developed and developing countries [67]. Increasing evidence of climate change is forcing businesses to play an active role in reducing sustainability burdens and preserving their resources for future generations [68]. The greater investment in research and development (R&D) [69] and in marketing & advertising are firm-level resources [70] that can provide a measure of protection against the destruction in the sustainability of superior profits which is associated with pro-market reforms in such emerging economies.

Like other research studies, this study has its limitations. It is an empirical study, which is based on a single country framework of Pakistan. Thus, it would be valuable to extend this study by employing data from both financial and non-financial firms and, accordingly, comparing Pakistan with other emerging economies, such as India and Bangladesh, particularly because the Pakistani and Indian economies have very similar features. Therefore, a replication of this study in other emerging economies may allow these results to be generalized. Besides, the differences between manufacturing and non-manufacturing business group firms, or the characteristics of high tech companies, could also be explored [71]. From this perspective, there is a need for improved methods to determine additional—that is, environmental—risk effects on their financial performance [72]. Furthermore, it would be important to consider that competition does not only occur among companies, but also among Small and Medium Enterprises (SMEs) [73].

Appendix A

Table A1. Summary Table of the Literature Review.

Author	Objective	Method	Key Findings
Khanna and Palepu (2000) [15]	To analyze the performance of group-affiliated firms relative to standalone firms.	Multiple regression analysis.	Finds that accounting and stock market measures of firm performance initially decline and subsequently increase once group diversification exceeds a certain level.
Khanna and Rivkin (2001) [17]	To examine the effects of group affiliation on profitability.	Ordinary least squares (OLS).	Finds that business group affiliation affects the economic performance in 12 of the markets.
Gunduz and Tatoglu (2003) [54]	To compare the performance of affiliates of diversified Turkish business groups with that of unaffiliated firms.	ANOVA Multiple regression analysis.	Reports that firms affiliated with diversified business groups do not significantly differ from unaffiliated firms in terms of accounting and stock market measures of performance.
Chu (2004) [39]	To investigate the influence of group affiliation on performance of firms.	Multiple regression analysis.	Finds a U-shape relationship between group affiliation and profitability in emerging economies.
Khanna and Yafeh (2005) [2]	To examine whether business groups facilitate mutual insurance among group-affiliated firms.	Weighted least square (WLS) regression is used for the analysis.	Finds substantial evidence of risk sharing by Japanese, Korean, and Thai groups, but little evidence of it elsewhere.
Claessens et al. (2006) [41]	To investigate the benefits and costs of group affiliation.	Multiple regression analysis.	Finds that mature and slow-growing firms with ownership structures gain more from group affiliation, while young and high-growth firms lose more.
Zattoni et al. (2009) [32]	To analyze how business group affiliation affects performance in India in the post-reform era i.e., from 1990 to 2006.	Applied SAS procedure (Time-series and cross-section regression) with variance component model.	Finds (1) benefits of business group affiliation are evident in the early phase of institutional transition (2) older group-affiliated firms are better able to cope with institutional transition than younger group-affiliated firms.

Table A1. *Cont.*

Author	Objective	Method	Key Findings
Carney et al. (2011) [26]	To study business group affiliation and performance.	Weighted least squares (WLS).	Finds that affiliates perform better in contexts with underdeveloped financial and labor markets.
He et al. (2013) [16]	To discover whether group-affiliated firms tend to outperform standalone firms.	Fixed effect OLS regression	Finds that business group membership has no effect on accounting performance
Elango et al. (2016) [9]	To study the impact of specific business group characteristics on the performance of group-affiliated firms.	Hierarchical linear models (HLM).	Finds (1) membership in a group contributes 6% of the performance variation of affiliated firms; (2) the importance of the business group to performance varies with the extent of group diversification, age and size.

Table A2. Sample Distribution across Industries.

Industry	Two-Digit SIC Code	Number of Firms	Percentage of Entire Sample
Food & Tobacco	1, 2, 9, 20, 21, 54	35	12
Basic Industries including Petroleum	10, 12, 13, 14, 24, 26, 28, 29, 33	74	26
Construction	15, 16, 17, 32, 52	20	7
Textile & Trade	22, 23, 31, 51, 53, 56, 59	129	45
Consumer Durables	25, 30, 36, 37, 39, 50, 55, 57, 34, 35, 38	7	3
Transportation	40, 41, 42, 44, 45, 47	17	6
Services	72, 73 75, 76, 80, 82, 87, 89	2	1
Others	No specific SIC code	0	0
Entire Sample		284	100

References

1. Cheung, Y.-L.; Rau, P.R.; Stouraitis, A. Tunneling, propping, and expropriation: evidence from connected party transactions in Hong Kong. *J. Financ. Econ.* **2006**, *82*, 343–386. [CrossRef]
2. Khanna, T.; Yafeh, Y. Business Groups and Risk Sharing around the World. *J. Bus.* **2005**, *78*, 301–340. [CrossRef]
3. Chang, S.J.; Hong, J. How much does the business group matter in Korea? *Strateg. Manag. J.* **2002**, *23*, 265–274. [CrossRef]
4. Khanna, T.; Yafeh, Y. Business Groups in Emerging Markets: Paragons or Parasites? *J. Econ. Lit.* **2007**, *45*, 331–372. [CrossRef]
5. Rehman, S. *Who Owns Pakistan?: Fluctuating Fortunes of Business Mughals*; Sani H. Panhwar: Islamabad, Pakistan, 1998; ISBN B007HF9GRO.
6. Masulis, R.W.; Pham, P.K.; Zein, J. Family Business Groups around the World: Financing Advantages, Control Motivations, and Organizational Choices. *Rev. Financ. Stud.* **2011**, *24*, 3556–3600. [CrossRef]
7. Granovetter, M. Coase revisited: Business groups in the modern economy. *Ind. Corp. Chang.* **1995**, *4*, 93–130. [CrossRef]
8. White, L.J. *Industrial Concentration and Economic Power in Pakistan*; Princeton University Press: Princeton, NJ, USA, 2015; ISBN 9781400871797.
9. Elango, B.; Pattnaik, C.; Wieland, J.R. Do business group characteristics matter? An exploration on the drivers of performance variation. *J. Bus. Res.* **2016**, *69*, 3205–3212. [CrossRef]

10. Williamson, O.E. The Economics of Organization: The Transaction Cost Approach. *Am. J. Sociol.* **1981**, *87*, 548–577. [CrossRef]

11. Coase, R.H. The Nature of the Firm. *Economica* **1937**, *4*, 386–405. [CrossRef]

12. Orbay, H.; Yurtoglu, B.B. The Impact of Corporate Governance Structures on the Corporate Investment Performance in Turkey. *Corp. Gov. Int. Rev.* **2006**, *14*, 349–363. [CrossRef]

13. Chang, S.J.; Choi, U. Strategy, Structure and Performance of Korean Business Groups: A Transactions Cost Approach. *J. Ind. Econ.* **1988**, *37*, 141. [CrossRef]

14. Khanna, T. Why Focused Strategies May Be Wrong For Emerging Markets. *Harv. Bus. Rev.* **1997**, *75*, 41–48.

15. Khanna, T.; Palepu, K. Is Group Affiliation Profitable in Emerging Markets? An Analysis of Diversified Indian Business Groups. *J. Financ.* **2000**, *55*, 867–891. [CrossRef]

16. He, J.; Mao, X.; Rui, O.M.; Zha, X. Business groups in China. *J. Corp. Financ.* **2013**, *22*, 166–192. [CrossRef]

17. Khanna, T.; Rivkin, J.W. Estimating the performance effects of business groups in emerging markets. *Strateg. Manag. J.* **2001**, *22*, 45–74. [CrossRef]

18. Smelser, N.J.; Swedberg, R. *The Handbook of Economic Sociology*; Princeton University Press: Princeton, NJ, USA, 2005; ISBN 9780691121260.

19. Saeed, A.; Belghitar, Y.; Clark, E. Political Connections and Leverage: Firm-level Evidence from Pakistan. *Manag. Decis. Econ.* **2015**, *36*, 364–383. [CrossRef]

20. Khanna, T.; Palepu, K.G.; Bullock, R.J. *Winning in Emerging Markets: A Road Map for Strategy And Execution*; Harvard Business Press: Brighton, MA, USA, 2010; ISBN 1422166953.

21. Doh, J.P.; Lawton, T.C.; Rajwani, T. Advancing Nonmarket Strategy Research: Institutional Perspectives in a Changing World. *Acad. Manag. Perspect.* **2012**, *26*, 22–39. [CrossRef]

22. Fisman, R.; Khanna, T. Facilitating development: The role of business groups. *World Dev.* **2004**, *32*, 609–628. [CrossRef]

23. Elango, B.; Pattnaik, C. Building capabilities for international operations through networks: A study of Indian firms. *J. Int. Bus. Stud.* **2007**, *38*, 541–555. [CrossRef]

24. Chang, S.; Singh, H. Corporate and industry effects on business unit competitive position. *Strateg. Manag. J.* **2000**, *21*, 739–752. [CrossRef]

25. Lee, K.; Peng, M.W.; Lee, K. From diversification premium to diversification discount during institutional transitions. *J. World Bus.* **2008**, *43*, 47–65. [CrossRef]

26. Carney, M.; Gedajlovic, E.R.; Heugens, P.P.M.A.R.; Van Essen, M.; Van Oosterhout, J. Business Group Affiliation, Performance, Context, and Strategy: A Meta-analysis. *Acad. Manag. J.* **2011**, *54*, 437–460. [CrossRef]

27. Leff, N.H. Industrial Organization and Entrepreneurship in the Developing Countries: The Economic Groups. *Econ. Dev. Cult. Chang.* **1978**, *26*, 661–675. [CrossRef]

28. Yiu, D.W.; Lu, Y.; Bruton, G.D.; Hoskisson, R.E. Business Groups: An Integrated Model to Focus Future Research. *J. Manag. Stud.* **2007**, *44*, 1551–1579. [CrossRef]

29. Williamson, O.E. Transaction-Cost Economics: The Governance of Contractual Relations. *J. Law Econ.* **1979**, *22*, 233–261. [CrossRef]

30. Shin, H.-H.; Park, Y.S. Financing constraints and internal capital markets: Evidence from Korean 'chaebols'. *J. Corp. Financ.* **1999**, *5*, 169–191. [CrossRef]

31. Hoshi, T.; Kashyap, A.; Scharfstein, D. Corporate Structure, Liquidity, and Investment: Evidence from Japanese Industrial Groups. *Q. J. Econ.* **1991**, *106*, 33–60. [CrossRef]

32. Zattoni, A.; Pedersen, T.; Kumar, V. The Performance of Group-affiliated Firms during Institutional Transition: A Longitudinal Study of Indian Firms. *Corp. Gov. An Int. Rev.* **2009**, *17*, 510–523. [CrossRef]

33. Yu, H.; van Ees, H.; Lensink, R. Does group affiliation improve firm performance? The case of Chinese state-owned firms. *J. Dev. Stud.* **2009**, *45*, 1615–1632. [CrossRef]

34. Jia, N.; Shi, J.; Wang, Y. Coinsurance Within Business Groups: Evidence from Related Party Transactions in an Emerging Market. *Manag. Sci.* **2013**, *59*, 2295–2313. [CrossRef]

35. Mursitama, T.N. Creating relational rents: The effect of business groups on affiliated firms' performance in Indonesia. *Asia Pac. J. Manag.* **2006**, *23*, 537–557. [CrossRef]

36. Ma, X.; Yao, X.; Xi, Y. Business group affiliation and firm performance in a transition economy: A focus on ownership voids. *Asia Pac. J. Manag.* **2007**, *23*, 467–483. [CrossRef]

37. Kim, H.; Hoskisson, R.E.; Tihanyi, L.; Hong, J. The Evolution and Restructuring of Diversified Business Groups in Emerging Markets: The Lessons from Chaebols in Korea. *Asia Pac. J. Manag.* **2004**, *21*, 25–48. [CrossRef]

38. Baumol, W.J. *Business Behavior, Value and Growth*; Macmillan: London, UK, 1958; ISBN B0000CKDLI.

39. Chu, W. Are Group-Affiliated Firms Really More Profitable than Nonaffiliated? *Small Bus. Econ.* **2004**, *22*, 391–405. [CrossRef]

40. Samuels, J.M.; Smyth, D.J. Profits, Variability of Profits and Firm Size. *Economica* **1968**, *35*, 127. [CrossRef]

41. Claessens, S.; Fan, J.P.H.; Lang, L.H.P. The benefits and costs of group affiliation: Evidence from East Asia. *Emerg. Mark. Rev.* **2006**, *7*, 1–26. [CrossRef]

42. Ma, X.; Lu, J.W. Business group affiliation as institutional linkages in China's emerging economy: A focus on organizational traits and institutional conditions. *Asia Pac. J. Manag.* **2017**, *34*, 675–697. [CrossRef]

43. Hadlock, C.J.; James, C.M. Do Banks Provide Financial Slack? *J. Financ.* **2002**, *57*, 1383–1419. [CrossRef]

44. Myers, S.C. Determinants of corporate borrowing. *J. Financ. Econ.* **1977**, *5*, 147–175. [CrossRef]

45. Myers, S.C.; Majluf, N.S. Corporate financing and investment decisions when firms have information that investors do not have. *J. Financ. Econ.* **1984**, *13*, 187–221. [CrossRef]

46. Froot, K.A.; Scharfstein, D.S.; Stein, J.C. A Framework for Risk Management. *J. Appl. Corp. Financ.* **1994**, *7*, 22–33. [CrossRef]

47. Aoki, M. *The Economic Analysis of the Japanese Firm*; North-Holland: Amsterdam, The Netherlands, 1984; ISBN 0444868224.

48. Cowling, M. The Growth-Profit Nexus. *Small Bus. Econ.* **2004**, *22*, 1–9. [CrossRef]

49. Chandler, G.N.; Jansen, E. The founder's self-assessed competence and venture performance. *J. Bus. Ventur.* **1992**, *7*, 223–236. [CrossRef]

50. Mendelson, H. Organizational Architecture and Success in the Information Technology Industry. *Manage. Sci.* **2000**, *46*, 513–529. [CrossRef]

51. Markman, G.D.; Gartner, W.B. Is Extraordinary Growth Profitable? A Study of *Inc. 500* High-Growth Companies. *Entrep. Theory Pract.* **2002**, *27*, 65–75. [CrossRef]

52. Vissa, B.; Greve, H.R.; Chen, W.-R. Business Group Affiliation and Firm Search Behavior in India: Responsiveness and Focus of Attention. *Organ. Sci.* **2010**, *21*, 696–712. [CrossRef]

53. Chari, M.D.R.; David, P. Sustaining superior performance in an emerging economy: An empirical test in the Indian context. *Strateg. Manag. J.* **2012**, *33*, 217–229. [CrossRef]

54. Gunduz, L.; Tatoglu, E. A comparison of the financial characteristics of group affiliated and independent firms in Turkey. *Eur. Bus. Rev.* **2003**, *15*, 48–54. [CrossRef]

55. Carney, M.; Shapiro, D.; Tang, Y. Business Group Performance in China: Ownership and Temporal Considerations. *Manag. Organ. Rev.* **2009**, *5*, 167–193. [CrossRef]

56. Farías, P. Business group characteristics and firm operating performance: evidence from Chile. *Acad. Rev. Latinoam. Adm.* **2014**, *27*, 226–235. [CrossRef]

57. Hirsch, B.T.; Seaks, T.G. Functional Form in Regression Models of Tobin's Q. *Rev. Econ. Stat.* **1993**, *75*, 381–385. [CrossRef]

58. Anton, S.G. The Impact of Enterprise Risk Management on Firm Value: Empirical Evidence from Romanian Non-financial Firms. *Eng. Econ.* **2018**, *29*. [CrossRef]

59. Gonenc, H.; Kan, O.; Karadagli, E. Business Groups and Internal Capital Markets. *Emerg. Mark. Financ. Trade* **2007**, *43*, 63–81. [CrossRef]

60. Lang, L.; Ofek, E.; Stulz, R.M. Leverage, investment, and firm growth. *J. Financ. Econ.* **1996**, *40*, 3–29. [CrossRef]

61. Breusch, T.S.; Pagan, A.R. A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica* **1979**, *47*, 1287. [CrossRef]

62. Hair, J.F. *Multivariate Data Analysis*; Prentice Hall: Upper Saddle River, NJ, USA, 2010; ISBN 0138132631.

63. Tate, R.L. Limitations of Centering for Interactive Models. *Sociol. Methods Res.* **1984**, *13*, 251–271. [CrossRef]

64. Chittoor, R.; Kale, P.; Puranam, P. Business groups in developing capital markets: Towards a complementarity perspective. *Strateg. Manag. J.* **2015**, *36*, 1277–1296. [CrossRef]

65. Manikandan, K.S.; Ramachandran, J. Beyond institutional voids: Business groups, incomplete markets, and organizational form. *Strateg. Manag. J.* **2015**, *36*, 598–617. [CrossRef]

66. Fukuyama, F. *Trust: The Social Virtues and the Creation of Prosperity*; Free Press: New York, NY, USA, 1995; ISBN 9780029109762.

67. Ruiz, S.; Arvate, P.; Xavier, W. Superior economic performance in developed and developing countries. *Int. J. Emerg. Mark.* **2017**, *12*, 93–107. [CrossRef]

68. Jayanti, R.K.; Rajeev Gowda, M.V. Sustainability dilemmas in emerging economies. *IIMB Manag. Rev.* **2014**, *26*, 130–142. [CrossRef]

69. Korzh, N.; Mostenska, T.; Bilan, Y. Resource-Based View in Managing Financial Component of Corporate Capital. *Polish J. Manag. Stud.* **2017**, *16*, 133–146. [CrossRef]

70. Moravcikova, D.; Krizanova, A.; Kliestikova, J.; Rypakova, M. Green Marketing as the Source of the Competitive Advantage of the Business. *Sustainability* **2017**, *9*, 2218. [CrossRef]

71. Cygler, J.; Sroka, W. Coopetition Disadvantages: The Case of the High Tech Companies. *Eng. Econ.* **2017**, *28*, 494–504. [CrossRef]

72. Hoti, S.; McAleer, M.; Pauwels, L.L. Modelling environmental risk. *Environ. Model. Softw.* **2005**, *20*, 1289–1298. [CrossRef]

73. Kot, S. Sustainable supply chain management in small and medium enterprises. *Sustainability* **2018**, *10*, 1143. [CrossRef]

Article

IPO Underpricing After the 2008 Financial Crisis: A Study of the Chinese Stock Markets

Rui Li [1], Wei Liu [2,3,*], Yong Liu [4] and Sang-Bing Tsai [5,*]

[1] School of Economics, Peking University, Beijing 100871, China; liruipku@pku.edu.cn
[2] Discipline of International Business, The University of Sydney, Sydney 2006, Australia
[3] School of Business, Wuyi University, Nanping 354300, China
[4] Political Economics Institute, Wuhan University of Science and Technology, Wuhan 430065, China; michael_ly@wust.edu.cn
[5] Zhongshan Institute, University of Electronic Science and Technology of China, Zhongshan 528402, China
* Correspondence: wei.liu2@sydney.edu.au (W.L.); sangbing@hotmail.com (S.-B.T.)

Abstract: A firm's capability of raising funding is closely related to its sustainable development. With a more efficient allocation of funding among the whole society, social resources will be better utilized. Initial Public Offering (IPO) can indeed be an effective means of raising capital for corporate ventures. Using 1069 firms which completed IPOs on Chinese stock exchanges between 1st January 2004 and 1st January 2013, we investigate the difference in IPO underpricing before and after the 2008 financial crisis. Based on OLS regression models, we find that the IPOs are less underpriced in the post-crisis period. We examine the moderating effects of firm size on the difference in IPO underpricing between pre- and post-crisis periods, finding that small firms experienced less IPO underpricing than large firms after the financial crisis. After applying different model specifications such as Robust and OProbit regressions, the results remain consistent. Our study contributes to understanding the dynamics and influences of the financial crisis on firms' IPO cost from the perspective of information asymmetry.

Keywords: IPO underpricing; financial crisis; information asymmetry; financial risks

1. Introduction

Initial Public Offering (IPO) is an important channel for firms to obtain direct funding in capital markets [1,2]. With the rapid development of Chinese stock markets over the past 20 years, many firms have adopted IPO as an effective source of capital funding. During IPOs, information asymmetry is a serious issue which hinders the process of funding from potential investors [3]. Potential investors possess less information than the firm who undertakes the IPO [4]. Potential investors face a higher level of uncertainty regarding the firm's profitability and performance. As investors are more exposed to risk, they will only submit purchase orders at a discounted stock price. To encourage potential investors to participate in IPOs, the underwriter of the IPO firm has to set an offer price lower than the intrinsic value of the share price [3,5]. The difference between the intrinsic value and the offer price serves as a risk premium of information asymmetry for potential investors. IPO underpricing is regarded as an indirect cost for firms during the process of capital funding [6]. Ritter (1987) has shown that increasing information disclosure before IPOs could reduce the cost of capital funding in U.S. stock markets [6]. Ang and Brau (2002) found a negative relationship between the transparency of a firm's information and IPO costs, significantly influencing corporate financial performance and sustainable development [7]. A firm's capability of raising initial funding is closely related to its sustainable development. With a more effective allocation of funding among the whole society, social resources will be better utilized by firms with a higher efficiency, thus creating more values. Moreover, the IPO

market plays a critical part in sustainable economic growth. If the IPO market dries up, this can have long lasting negative effects on the evolution of innovative industries. Innovative firms might run into liquidity problems, and the speed of commercialization of technological innovations might slow down. Ultimately, a country's economic growth path can be negatively affected [8].

The 2008 financial crisis is typically regarded as the worst financial crisis since the Great Depression of the 1930s [9–11]. The global financial crisis of 2008 was caused by the expansion of subprime mortgages to high-risk borrowers under the situation of information asymmetry. In other words, there was asymmetry of information spun throughout the 2008 financial crisis. The crisis not only resulted in the collapse of famous and giant financial institutions, such as Lehman Brothers, but also impeded global credit markets and required intensive government interventions. After the crisis, governments around the world introduced a series of regulatory proposals and policies to require more information disclosure and increase the transparency of transactions [12]. Besides, the ensuing period after the 2008 financial crisis was driven by a revolution of information and communication technology (ICT). With the developments and advances in technology, information about firms became more transparent after the crisis [12].

While there is an extensive body of research on the determinants of IPO underpricing and the impacts of the financial crisis separately [3,7,13–18], the extant literature has paid little attention to the relationship between financial crisis and a firm's funding costs or IPO underpricing issues. Song and Lee (2012) studied the long-term effect of the 1998 Asian financial crisis on corporate cash holdings [19]. They divided the sample into well-established firms before the crisis and IPO firms during and after the crisis, finding that the crisis has dramatically changed firms' cash-holding policies. IPO companies engaging in aggressive income-increasing earnings management are proved to have a significantly worse market-based performance. For these companies, personal liquidity concerns are an important factor in IPO decisions during the economic crisis [20]. Blocker and Sandner (2009) found that the financial crisis is related to a 20% decrease in the average amount of funds raised per funding round [8]. So far, there is no study investigating the role of the financial crisis on IPO underpricing from the aspect of information asymmetry. The collapse of Lehman Brothers in September 2008 and other events leading to the financial crisis provide a good research opportunity to address this question in greater detail. To fill the literature gap, we investigate the influence of information on a firm's cost of capital funding. Our main study addresses the specific role played by the global financial crisis of 2008.

Utilizing 1069 firms going public on Chinese stock exchanges between January 2004 and January 2013, we study the difference in IPO underpricing before and after the financial crisis of 2008. The results suggest that IPOs are significantly less underpriced in the post-crisis period. Moreover, our empirical study goes beyond the original model of IPO underpricing by revealing the moderating effects of firm size on the relationship between the financial crisis and IPO underpricing. It is found that small firms experienced less IPO underpricing than large firms after the financial crisis.

Firstly, to the best of our knowledge and the literature in hand, this study is one of the few to study the nexus between the financial crisis and IPO underpricing from an empirical perspective [8,19,20]. Previous literature has studied the influences of financial crisis on cash holding and market-based performance. Blocker and Sandner (2009) studied the effect of crisis on the funding of US internet start-ups [8]. However, IPO firms cover a much wider range of industries. Different from the previous research, this study directly focuses on the funding cost of a firm and provides a new insight into the analysis of IPO costs. Secondly, publicly traded equity represents one of the most important sources of external capital to facilitate firm investment [21]. Although previous research has identified a significant relationship between information and a firm's IPO cost, we revisit this topic from a novel perspective of financial crisis and highlight the impact of information transparency in reducing a firm's cost of capital funding. Thirdly, we find that firms of different sizes are affected by the financial crisis to different degrees, which reveals that firm size plays an important role in the process of IPOs.

The remainder of this study is organized as follows: Section 2 presents the development of our hypotheses and provides the details of our theoretical arguments. Section 3 details the data and research methods used in our study. Section 4 reports the results of the empirical estimations and robustness checks. Section 5 presents the discussion and draws conclusions.

2. Theories and Hypotheses

2.1. Literature Review

Firms adopt an Initial Public Offering (IPO) as an efficient tool to raise direct funding in capital markets [1,2]. Normally, initially offered shares are underpriced compared with the market price. IPO underpricing refers to the difference between the price at which the shares are sold to investors during the offering procedure and the price at which the shares are traded in the secondary market. IPO underpricing has been empirically investigated in numerous countries and the results reveal that this phenomenon occurs all over the world [19–23]. A firm's information and the IPO costs significantly influence corporate financial performance and sustainable development [7]. It is commonly acknowledged that the root of underpricing is information asymmetry [3,5,16]. The theory of information asymmetry is used to illustrate the asymmetric distribution of relevant information on the market in the incomplete information market. The concept of information asymmetry originates from George A. Akerlof (1978) [24]. According to the theory of information economics, both borrowers and lenders face information asymmetry in the process of financing. Information asymmetry results in moral hazard and adverse selection [25]. The information asymmetry between different participants has produced different theoretical bases. First is the principal-agent theory which involves information asymmetry between the issuers and the underwriters, assuming that the underwriters have more information about the potential market demand and market conditions than the issuers. Issuers do not have this information due to their lack of market demand information and they have to negotiate the price to ensure the success of the issuance. Therefore, the underpricing is the remuneration for the underwriters. Secondly, the information asymmetry between the issuers and the investors points to the signaling theory [26]. High-quality company initiative signaling strategies further promote IPO underpricing. High-quality companies use IPO underpricing as a signal of its value to attract potential investors in the secondary market. These high-quality companies adopt subsequent issuances to compensate for the cost of underpricing. Thirdly, the information asymmetry among different investors results in Rock's "winner curse" hypothesis [5]. Investors who lack information are more likely to subscribe for new shares with higher prices than value, and are faced with "winner curses".

More importantly, information asymmetry has severe impacts on companies' funding cost and financing capability. Myers and Majluf (1984) have suggested that when the capital market is not perfect, there is information asymmetry between the company's external investors and insiders, which makes the cost of external financing higher than the cost of internal financing [27]. In order to ensure investors' lack of information, issuers have to underprice the shares during the IPO process. Because of the increase in financing costs, the company's net present value drops. As a result, the company's investment level will be reduced. Hubbard (1998) obtained a full picture of the relationship between imperfections in capital markets and corporate investment [28]. In summary, research on the capital market information asymmetry model and incentive problems shows that information cost determines the degree of financing constraints faced by firms. Therefore, reducing information asymmetry not only enhances firms' capability of raising funding, but also helps in resource allocations and contributes to sustainable economic growth.

The degree of IPO underpricing in developing countries is much greater than in developed countries. The average underpricing level in developed countries is around 15%, and the average underpricing level in some emerging countries is around 60%, while the IPO underpricing level in China's stock market is even higher [19]. The average IPO underpricing was closer to 100% from 1987 to 1995 [29,30]. The Chinese IPO process is the same as that in Western countries. The IPO process

includes several procedures such as selecting underwriters, setting the offer price, allocating the shares, and trading on the secondary market [31]. Nevertheless, IPOs in China require approval from the CSRC, which is different from the Securities and Exchange Commission's registration requirements for initial offerings in the U.S. [32,33]. The CSRC examines the quality of new securities, such as evaluating their issuers' profitability and potential risks. New issuers can only begin the IPO process after they obtain CSRC approval. Moreover, the institutional environment in China is very different from that in North America and Europe. For example, the Chinese Government can control the IPO offer price via the "Guidance window". Firms in China are more subject to government regulations than in Western countries [34].

2.2. The 2008 Financial Crisis and IPO Underpricing

The financial crisis became clearly visible in September 2008, when the bankruptcy of Lehman Brothers was announced. Immediately after that, the giant insurance company American International Group (AIG) suffered a liquidity crisis following a downgrade in its credit rating. Following this, many other financial institutions in the US and around the world were severely affected, losing large portions of their value, and could only be saved from bankruptcy by government funds. Stock prices declined, and a recession began [8]. The global financial crisis of 2008 was caused by the expansion of subprime mortgages to high-risk borrowers. Borrowers have an advantage of information over lenders because the former knows more about the investment projects that they want to undertake. Before the financial crisis, the market was typified by massive information asymmetry and built up innumerable layers of bad mortgages [7,35]. In other words, there was asymmetry of information spun throughout the 2008 financial crisis. After the crisis, governments around the world introduced a series of regulatory proposals and policies to require more information disclosure and increase the transparency of transactions. The degree of IPO underpricing in developing countries is much greater than in developed countries. The average underpricing level in developed countries is around 15%, and the average underpricing level in some emerging countries is around 60%, while the IPO underpricing level in China's stock market is even higher [19]. The average IPO underpricing was closer to 100% from 1987 to 1995 [29,30]. The Chinese IPO process is the same as that in Western countries. The IPO process includes several procedures such as selecting underwriters, setting the offer price, allocating the shares, and trading on the secondary market [31]. Nevertheless, IPOs in China require approval from the CSRC, which is different from the Securities and Exchange Commission's registration requirements for initial offerings in the U.S. [32,33]. The CSRC examines the quality of new securities, such as evaluating their issuers' profitability and potential risks. New issuers can only begin the IPO process after they obtain CSRC approval. Moreover, the institutional environment in China is very different from that in North America and Europe. For example, the Chinese Government can control the IPO offer price via the "Guidance window". Firms in China are more subject to government regulations than in Western countries [34].

Meanwhile, the world experienced an ICT (information and communication technology) surge in 2009 [9,36], which logically suggests that the period following the 2008 financial crisis has been driven by revolutions and developments in ICT. Due to the changes in regulation and advances in technology, information about firms became more transparent after the crisis. IPO underpricing is regarded as an indirect cost for firms during the process of capital funding [6,37]. Ritter (1987) has shown that increasing information disclosure before IPOs could reduce the cost of capital funding in the U.S. stock market [6]. Ang and Brau (2002) found a negative relationship between the transparency of a firm's information and its IPO costs [7]. Thus, we formulate Hypothesis 1:

Hypothesis 1. *IPO underpricing decreases after the 2008 financial crisis compared with the period before the crisis.*

2.3. Moderating Effect of Firm Size

Firm size is an important issue in the relationship between IPO underpricing and information asymmetry. Larger firms, as compared to smaller firms, represented less uncertainty and asymmetric information for potential investors before the financial crisis [3]. After the financial crisis, information and communication technologies (ICTs) have experienced a great surge of revolution. The development of ICTs has more significant impacts on small firms than on large firms [36,38]. Moreover, after the financial crisis, small firms have been subject to stricter regulations and have been required to disclose more information [39,40]. The asymmetric information of small firms is supposed to have decreased to a greater extent after the crisis. Therefore, we argue that firm size moderates the difference in IPO underpricing between pre- and post-crisis periods. Then, we derive Hypothesis 2:

Hypothesis 2. *IPO underpricing decreases more for smaller firms after the 2008 financial crisis.*

3. Data and Methods

3.1. Sample Selection

Our sample includes 1069 firms completing IPOs on Chinese stock exchanges between 1 January 2004 and 1 January 2013. We have chosen this time window because the policy restricting maximum return on the first day of IPO to 44% was issued in 2013 and implemented in 2014 by Shanghai and Shenzhen stock exchanges. We have collected firm information and IPO trading data from the China Stock Market Accounting Research (CSMAR) database (http://www.gtarsc.com/) [41].

3.2. Variable Measurement

To distinguish the difference in IPO underpricing before and after the crisis, we have chosen the year of 2008 as the relevant cutoff, which is excluded in the time window. The variable, post-crisis, is a dummy coded as 1 if the year is after 2008 and 0 otherwise [34,42].

The dependent variable in our analysis is IPO underpricing. There is an excess rate of return on the stock offerings in China, which means that the offer price is significantly lower than the closing price on the first day of IPO. Underpricing refers to the difference between the price initially offered and the closing price on the first day of trading [42–46]. IPO underpricing is commonly measured by the return on the first day of IPO, illustrating the degree of undervaluation of the offer price [6]. The market adjusted return excluding the factor of market price is a more accurate indicator of IPO underpricing [47]. Thus, we use the market adjusted return on the first day of IPO to represent IPO underpricing, calculated as follows:

$$IPO\ underpricing = [(closing\ price_t - offer\ pricing_t)/offer\ price_t - (market\ closing\ index_t - market\ closing\ index_{t-1})/maket\ closing\ index_{t-1}] \tag{1}$$

where the subscript t denotes the first day of IPO. Figure 1 depicts the Kernel density of IPO underpricing before and after the financial crisis. As is shown, the kernel density of IPO underpricing in the pre-crisis period is located to the right of that in the post-crisis period. Thus, we can observe from Figure 1 that the average IPO underpricing after the financial crisis was lower than that before the crisis.

Following previous studies, such as Arthurs et al. (2008) and Jia et al. (2014) [33,48], we control for the effect of firm size and value, including the firm's number of staff, capital at registration, and net value per share before IPO [49,50]. Underwrite cost indicates that the direct cost of IPO is controlled in our analysis [51]. We also introduce several variables to control for offering and trading conditions, including IPO volume, IPO price, lottery rate, and turnover on the first day of IPO [52,53]. Table 1 presents the description of the variables in our analysis. Table 2 shows the statistical summary of these variables. The mean of IPO underpricing before the crisis is 1.464 and after the crisis is 0.358.

As would be expected, IPO underpricing decreased by a noticeable degree after the crisis. To avoid the problem of collinearity of the variables, we estimate the Variance Inflating Factor (VIF) and report the statistics in Table 3. Table 3 shows that all values of VIF are less than 10, which confirms that there are no collinearity issues among the variables.
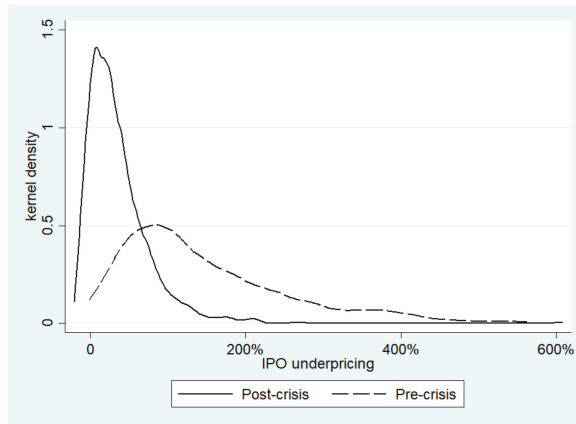
Figure 1. Kernel density estimation (KDE) estimates the probability density function of IPO underpricing. The solid curve demonstrates the density function of IPO before the financial crisis and the dashed curve demonstrates the density function of IPO before the financial crisis.

Table 1. Description of variables.

Variable Abbreviation	Variable Definition	Dimension
IPO underpricing	Market adjusted IPO underpricing	%
POST	Time indicator of Post-crisis	N/A
NOS	Number of staffs	Ten Thousand
NVPS	Net value per share	Yuan
CAR	Capital at registration	Million Yuan
IPOV	IPO volume	Billion
IPOP	IPO price	Yuan
UC	Underwrite cost	Million Yuan
TOFD	Turnover on the first day of IPO	%
LR	Lottery rate of IPO	%

Table 2. Statistical summary of variables.

Variables	Pre-Crisis		Post-Crisis	
	Mean	S.D.	Mean	S.D.
IPO underpricing	1.464	1.068	0.358	0.431
NOS	1.258	5.284	0.304	1.822
NPVS	2.472	0.894	3.089	1.171
CAR	6500	35481	741.0	10041
IPOV	0.396	1.385	0.120	1.006
IPOP	10.24	5.810	26.06	14.90
UC	66.84	153.6	57.00	57.81
TOFD	0.675	0.109	0.701	0.204
LR	0.389	0.592	1.456	3.059
Observations	188		881	

Table 3. Summary of the variance inflating factor.

Variable	VIF	VIF
Ln(CAR)	5.03	0.19899
Ln(UC)	5.01	0.199618
NOS	3.03	0.329525
IPOV	3.01	0.332324
IPOP	2.32	0.430528
POST	1.66	0.601727
NVPS	1.27	0.789147
TOFD	1.17	0.856115
LR	1.17	0.856265
Mean VIF	2.63	

3.3. Research Design

To obtain sophisticated results, this paper employs the Ordinary Least Square model (OLS) to perform an empirical analysis for Hypothesis 1. Model 1 (Equation (2)) examines the difference in IPO underpricing between pre- and post-crisis periods. We take the logarithms of capital at registration and the underwrite cost.

$$
\begin{aligned}
IPO\ underpricing_i = \alpha + \beta_{1R}POST_i + \beta_{2R}NOS_i + \beta_{3R}NVPS_i + \beta_{4R}Ln(CAR_i) \\
+ \beta_{5R}IPOV_i + \beta_{6R}IPOP_i + \beta_{7R}Ln(UC_i) + \beta_{8R}TOFD_i + \beta_{9R}LR_i + \varepsilon_R
\end{aligned}
\tag{2}
$$

To examine Hypothesis 2, we use two variables—the number of staff and capital at registration to measure firm size, which indicates that smaller firms are equal to the firms with a smaller number of staff and less capital at registration. Then, we introduce their interaction terms and post-crisis into Model 2 (Equation (3)) and Model 3 (Equation (4)) to test the argument, respectively.

$$
\begin{aligned}
IPO\ underpricing_i = \alpha + \beta_{1S}POST_i + \beta_{2S}NOS_i + \beta_{3S}NVPS_i + \beta_{4S}Ln(CAR_i) \\
+ \beta_{5S}IPOV_i + \beta_{6S}IPOP_i + \beta_{7S}Ln(UC_i) + \beta_{8S}TOFD_i + \beta_{9S}LR_i + \beta_{10S}NOS_i \times POST_i + \varepsilon_S
\end{aligned}
\tag{3}
$$

$$
\begin{aligned}
IPO\ underpricing_i = \alpha + \beta_{1T}POST_i + \beta_{2T}NOS_i + \beta_{3T}NVPS_i + \beta_{4T}Ln(CAR_i) \\
+ \beta_{5T}IPOV_i + \beta_{6T}IPOP_i + \beta_{7T}Ln(UC_i) + \beta_{8T}TOFD_i + \beta_{9T}LR_i + \beta_{10T}Ln(CAR_i) \times POST_i + \varepsilon_T
\end{aligned}
\tag{4}
$$

In addition, Robust and Ordered Probit (OProbit) regressions are used for robustness checks to test whether the results are stable. Robust regression has the same model specifications as the OLS model. To perform the OProbit regressions, the values of IPO underpricing are categorized into five quartiles with ordered scores from 1 to 5 [54]. If the results obtained from two different model specifications are the same, we can safely conclude that the results from the analysis are robust.

4. Results

4.1. Regression Analyses and Results

Table 4 provides the OLS regression results of the three models. Model 1 shows the difference in IPO underpricing between pre- and post-crisis periods. The estimated coefficient of POST is significantly negative and equal to -0.948, indicating that the average of IPO underpricing in the Chinese stock market decreased by 94.80% after the financial crisis. Model 2 adds the two-way interaction firm's number of staff and post-crisis. The joint effect of a firm's number of staff and post-crisis is significantly positive ($\beta = 0.003$, $p < 0.10$). Consistent with the result in Model 2, Model 3 shows that the estimated interaction coefficient of a firm's capital at registration and post-crisis is statistically significantly positive ($\beta = 0.152$, $p < 0.01$). The results of Models 2 and 3 validate that firm size moderates the difference in IPO underpricing between pre- and post-crisis periods.

Specifically, small firms experienced less IPO underpricing than large firms after the financial crisis, because small firms reduced information asymmetry to a greater degree in the post-crisis period.

Table 4. Estimation results of OLS regressions.

Variables	Model 1	Model 2	Model 3
POST	−0.948 ***	−0.971 ***	−3.893 ***
	(0.055)	(0.056)	(0.596)
NOS	0.004	0.003	0.020 *
	(0.010)	(0.010)	(0.011)
NVPS	−0.045 ***	−0.048 ***	−0.045 ***
	(0.016)	(0.016)	(0.016)
Ln(CAR)	0.048	0.044	−0.076 *
	(0.030)	(0.030)	(0.039)
IPOV	0.003	−0.025	−0.020
	(0.026)	(0.030)	(0.026)
IPOP	0.003 *	0.003 *	0.003 *
	(0.002)	(0.002)	(0.002)
Ln(UC)	−0.305 ***	−0.291 ***	−0.239 ***
	(0.055)	(0.056)	(0.056)
TOFD	1.035 ***	1.032 ***	1.049 ***
	(0.091)	(0.091)	(0.090)
LR	−0.004	−0.005	−0.007
	(0.006)	(0.006)	(0.006)
NOS × POST		0.003 *	
		(0.001)	
Ln(CAR) × POST			0.152 ***
			(0.031)
Observations	1069	1069	1069
R-squared	0.481	0.483	0.493
Adj R-squared	0.477	0.478	0.488

Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.2. Robustness Checks

To better prove Hypothesis 1, we calculate the IPO underpricing tendency of each year in our data and display the result in Table 5. From Table 5, it can be observed that from the year 2008, the underpricing tendency exhibits a downwards trend, which supports the hypothesis that IPO underpricing decreases after the 2008 financial crisis compared with the period before the crisis.

Table 5. IPO underpricing tendency.

2004	2005	2006	2007	2008 (Cutoff)	2009	2010	2011	2012
0.738	0.505	0.802	1.822	1.234	0.718	0.412	0.216	0.267

To satisfy the normality distribution of the residuals in OLS models, we use Robust regression to check the validity of our results [55]. Model 1 in Table 6 shows the difference in IPO underpricing between pre- and post-crisis periods. The estimated coefficient of POST is significantly negative. Model 2 adds the two-way interaction firm's number of staff and post-crisis to the Robust regression. The joint effect of a firm's number of staff and post-crisis is significantly positive (β = 0.014, $p < 0.05$). Model 3 shows that the estimated interaction coefficient of a firm's capital at registration and post-crisis

is statistically significantly positive ($\beta = 0.042$, $p < 0.01$). Overall, the empirical results of Robust regressions are consistent with our Hypotheses 1 and 2.

Table 6. Estimation results of Robust regressions.

Variables	Model 1	Model 2	Model 3
POST	−0.543 ***	−0.552 ***	−1.365 ***
	(0.0269)	(0.0274)	(0.297)
NOS	−0.00534	−0.00828 *	−0.00147
	(0.00497)	(0.00493)	(0.00521)
NVPS	−0.0187 **	−0.0179 **	−0.0158 **
	(0.00777)	(0.00776)	(0.00776)
Ln(CAR)	0.0218	0.0227	−0.0130
	(0.0150)	(0.0149)	(0.0194)
IPOV	0.0123	0.000772	0.00852
	(0.0127)	(0.0148)	(0.0129)
IPOP	0.00122	0.00125	0.00124
	(0.000807)	(0.000800)	(0.000806)
Ln(UC)	−0.109 ***	−0.102 ***	−0.0943 ***
	(0.0272)	(0.0273)	(0.0280)
TOFD	0.830 ***	0.811 ***	0.817 ***
	(0.0451)	(0.0447)	(0.0450)
LR	−0.00923 ***	−0.0169 ***	−0.0175 ***
	(0.00305)	(0.00303)	(0.00306)
NOS × POST		0.0140 **	
		(0.00710)	
Ln(CAR) × POST			0.0419 ***
			(0.0153)
Observations	1069	1069	1069
R-squared	0.565	0.577	0.583
Adj R-squared	0.562	0.573	0.579

Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

As the empirical analysis is tested by OLS models, we use Ordered Probit (OProbit) regressions to check whether the results of the main effect and interaction terms are robust. The results of the robustness checks are shown in Table 7. Model 1 tests the main effect between financial crisis and IPO underpricing; the results show that the coefficient of post-crisis is negative and significant ($\beta = -2.150$, $p < 0.01$), which is consistent with the result of the OLS analysis. Models 2 and 3 test the moderating effect of firm size, where the number of staff is positively significant ($\beta = 0.006$, $p < 0.10$). However, the moderating effect of capital at registration is positive but not significant in the confidence interval from 0 to 0.1. The result is less than 0.152 and the difference is not quite significant compared with the OLS result. Thus, the magnitude of this result can still be regarded as robust, though the significance reduces due to the different calculation methods involved in the analyses. The values of Pseudo R-square in the three models are not less than 0.350, which reflects enough explanatory power of the whole regression. Therefore, the checks confirm that the results of Hypotheses 1 and 2 are robust.

Table 7. Estimation results of OProbit regressions.

Variables	Model 1	Model 2	Model 3
POST	−2.150 ***	−2.191 ***	−3.913 ***
	(0.136)	(0.139)	(1.382)
NOS	0.018	0.013	0.026
	(0.022)	(0.021)	(0.022)
NVPS	−0.039	−0.045	−0.038
	(0.036)	(0.036)	(0.036)
Ln(CAR)	0.177 ***	0.164 **	0.101
	(0.068)	(0.069)	(0.091)
IPOV	0.037	0.022	0.042
	(0.073)	(0.067)	(0.072)
IPOP	0.008 **	0.008 **	0.008 **
	(0.004)	(0.004)	(0.004)
Ln(UC)	−0.409 ***	−0.394 ***	−0.374 ***
	(0.124)	(0.124)	(0.127)
TOFD	6.310 ***	6.295 ***	6.297 ***
	(0.309)	(0.309)	(0.308)
LR	−0.492 ***	−0.504 ***	−0.499 ***
	(0.045)	(0.045)	(0.045)
NOS × POST		0.006 *	
		(0.003)	
Ln(CAR) × POST			0.091
			(0.071)
Observations	1069	1069	1069
LR Chi-square	1203.27	1205.96	1204.91
Pseudo R-square	0.350	0.351	0.351

Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5. Discussion

5.1. Contributions and Implications

In the context of the Chinese IPO process, this paper makes unique contributions. Firstly, to the best of our knowledge and the literature in hand, this study pioneers in studying the relationship between the 2008 financial crisis and IPO underpricing from an empirical perspective. Thus, this study provides a new insight into the analysis of IPO cost.

Secondly, although previous research has identified a significant relationship between information asymmetry and a firm's IPO cost, we revisit this topic from the novel perspective of the 2008 financial crisis and highlight the impact of information transparency in reducing a firm's cost of capital funding. This paper contributes to understanding the dynamics and influences of the financial crisis on the stock market from the perspective of information asymmetry.

Thirdly, we find that firms of different sizes are differently affected by the financial crisis, which reveals that firm size plays an important role in the IPO process. In this sense, our study has practical implications for firms going through the IPO process in a transitional economy with developing information technology and improving capital market regulations.

5.2. Limitations and Future Study Directions

Although several interesting findings are presented in our study, the possible limitations should be noted. Firstly, although we clearly address the impact of the 2008 financial crisis on Chinese stock

markets, some scholars have argued that China withstood this great recession. They point out that the huge stimulus package put in place by the Chinese government in 2008 meant that China suffered relatively little from the financial crisis. The Chinese government also used state-owned enterprises as a fiscal instrument to implement an aggressive stimulus program in 2009. Further investigation in other developing countries is encouraged to examine the robustness of our findings.

Moreover, we have tested the moderating effects of firm size in the relationship between financial crisis and IPO underpricing. The finding shows that firm-level characteristics are associated with the degree of information asymmetry. However, environmental and institutional factors could also result in the varying degrees of information asymmetry and thus moderate the effect of financial crisis on the IPO process. An examination of these factors may provide deeper insights and is worthy of consideration in future analyses.

6. Conclusions

Utilizing 1069 firms completing IPOs on Chinese stock exchanges between January 2004 and January 2013, we have studied the difference in IPO underpricing before and after the financial crisis of 2008. The results suggest that IPOs are significantly less underpriced in the post-crisis period. Moreover, our empirical study goes beyond the original model of IPO underpricing by revealing the moderating effects of firm size on the relationship between the financial crisis and IPO underpricing. The findings demonstrate that small firms experienced less IPO underpricing than large firms after the 2008 financial crisis. In order to check the robustness, we have calculated the IPO underpricing tendencies each year and performed Robust and OProbit regressions. All results suggest that our empirical analyses are consistent.

IPO is still a relatively new but important activity in emerging markets such as China, but it has become an element vital to the economy. Many questions are left unanswered, thus offering good opportunities for future research. Does the influence of financial crises on IPO activity differ among regions and industries? How do IPO firms receiving funding during the financial crisis differ from firms that had received funding before the financial crisis? How do firms respond to the changes posed by the financial crisis and the difficulties encountered in the search for IPO funding? In addition to IPO firms, will the financial crisis have a similar impact on the funding process of newly emerged business such as start-ups?

Author Contributions: In this study, R.L. conceived and designed the study framework; Y.L. prepared the data for the analysis and provided the method; W.L. analyzed the data; S.-B.T. revised the paper.

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

References

1. Daily, C.M.; Certo, S.T.; Dalton, D.R.; Roengpitya, R. IPO underpricing: A meta-analysis and research synthesis. *Entrep. Theory Pract.* **2003**, *27*, 271–295. [CrossRef]
2. Reilly, F.K.; Hatfield, K. Investor experience with new stock issues. *J. Financ. Anal.* **1969**, *25*, 73–80. [CrossRef]
3. Ritter, J.R.; Welch, I. A review of IPO activity, pricing, and allocation. *J. Financ.* **2002**, *57*, 1795–1828. [CrossRef]
4. Easley, D.; O'hara, M. Information and the cost of capital. *J. Financ.* **2004**, *59*, 1553–1583. [CrossRef]
5. Rock, K. Why new issues are underpriced. *J. Financ. Econ.* **1986**, *15*, 187–212. [CrossRef]
6. Ritter, J.R. The costs of going public. *J. Financ. Econ.* **1987**, *19*, 269–281. [CrossRef]
7. Ang, J.S.; Brau, J.C. Firm transparency and the costs of going public. *J. Financ. Res.* **2002**, *25*, 1–17. [CrossRef]
8. Block, J.; Sandner, P. What is the effect of the financial crisis on venture capital financing? Empirical evidence from US Internet start-ups. *Venture Cap.* **2009**, *11*, 295–309. [CrossRef]

9.   Taylor, J.B. *The Financial Crisis and the Policy Responses: An Empirical Analysis of What Went Wrong*; National Bureau of Economic Research: Cambridge, MA, USA, 2009.

10.   Vyas, D. The timeliness of write-downs by U.S. financial institutions during the financial crisis of 2007–2008. *J. Account. Res.* **2011**, *49*, 823–860. [CrossRef]

11.   Kirkpatrick, G. The corporate governance lessons from the financial crisis. *OECD J. Financ. Mark. Trends* **2009**, *1*, 61–87. [CrossRef]

12.   Erkens, D.H.; Hung, M.; Matos, P. Corporate governance in the 2007–2008 financial crisis: Evidence from financial institutions worldwide. *J. Corp. Financ.* **2012**, *2*, 389–411. [CrossRef]

13.   Beatty, R.P.; Ritter, J.R. Investment banking, reputation, and the underpricing of initial public offerings. *J. Financ. Econ.* **1986**, *15*, 213–232. [CrossRef]

14.   Su, D. Adverse-selection versus signaling: Evidence from the pricing of Chinese IPOs. *J. Econ. Bus.* **2004**, *56*, 1–19. [CrossRef]

15.   Benveniste, L.M.; Wilhelm, W.J., Jr. A comparative analysis of IPO proceeds under alternative regulatory environments. *J. Financ. Econ.* **1990**, *28*, 173–207. [CrossRef]

16.   Loughran, T.; Ritter, J. Why has IPO underpricing changed over time? *Financ. Manag.* **2004**, *33*, 5–37. [CrossRef]

17.   Reinhart, C.M.; Rogoff, K.S. Is the 2007 US sub-prime financial crisis so different? An international historical comparison. *Am. Econ. Rev.* **2008**, *98*, 339–344. [CrossRef]

18.   Reinhart, C.M.; Rogoff, K.S. The aftermath of financial crises. *Am. Econ. Rev.* **2009**, *99*, 466–472. [CrossRef]

19.   Song, K.R.; Lee, Y. Long-term effects of a financial crisis: Evidence from cash holdings of East Asian firms. *J. Financ. Quant. Anal.* **2012**, *47*, 617–641. [CrossRef]

20.   Ahmad-Zaluki, N.A.; Campbell, K.; Goodacre, A. Earnings management in Malaysian IPOs: The East Asian crisis, ownership control, and post-IPO performance. *Inter. J. Account.* **2011**, *46*, 111–137. [CrossRef]

21.   He, J.J.; Tian, X. The dark side of analyst coverage: The case of innovation. *J. Financ. Econ.* **2013**, *109*, 856–878. [CrossRef]

22.   Engelen, P.J.; Van Essen, M. Underpricing of IPOs: Firm-, issue-and country-specific characteristics. *J. Bank Financ.* **2010**, *34*, 1958–1969. [CrossRef]

23.   Hahn, T.; Ligon, J.A.; Rhodes, H. Liquidity and initial public offering underpricing. *J. Bank Financ.* **2013**, *12*, 4973–4988. [CrossRef]

24.   Akerlof, G.A. The market for "lemons": Quality uncertainty and the market mechanism. *Uncertain. Econ.* **1978**, *235*, 237–251. [CrossRef]

25.   Baron, D.P. A model of the demand for investment banking advising and distribution services for new issues. *J. Financ.* **1982**, *37*, 955–976. [CrossRef]

26.   Allen, F.; Faulhaber, G.R. Signalling by underpricing in the IPO market. *J. Financ. Econ.* **1989**, *23*, 303–323. [CrossRef]

27.   Myers, S.C.; Majluf, N.S. Corporate financing and investment decisions when firms have information that investors do not have. *J. Financ. Econ.* **1984**, *13*, 187–221. [CrossRef]

28.   Hubbard, R.G. Capital Market Imperfections and Investment. *J. Econ. Lit.* **1998**, *36*, 193–225.

29.   Su, D.; Fleisher, B.M. An empirical investigation of underpricing in Chinese IPOs. *Pac.-Basin. Financ. J.* **1999**, *7*, 173–202. [CrossRef]

30.   Su, D.; Fleisher, B.M. Why does return volatility differ in Chinese stock markets? *Pac.-Basin. Financ. J.* **1999**, *7*, 557–586. [CrossRef]

31.   Certo, S.T.; Holcomb, T.R.; Holmes, R.M., Jr. IPO research in management and entrepreneurship: Moving the agenda forward. *J. Manag.* **2009**, *35*, 1340–1378. [CrossRef]

32.   Tian, L. Regulatory underpricing: Determinants of Chinese extreme IPO returns. *J. Empir. Financ.* **2011**, *18*, 78–90. [CrossRef]

33.   Jia, M.; Zhang, Z. Donating money to get money: The role of corporate philanthropy in stakeholder reactions to IPOs. *J. Manag. Stud.* **2014**, *51*, 1118–1152. [CrossRef]

34.   Liu, W.; Wei, Q.; Huang, S.-Q.; Tsai, S.-B. Doing Good Again? A Multilevel Institutional Perspective on Corporate Environmental Responsibility and Philanthropic Strategy. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1283. [CrossRef] [PubMed]

35.   Crotty, J. Structural causes of the global financial crisis: A critical assessment of the 'new financial architecture'. *Camb. J. Econ.* **2009**, *33*, 563–580. [CrossRef]

36. Southern, A.; Tilley, F. Small firms and information and communication technologies (ICTs): Toward a typology of ICTs usage. *New Tech. Work Employ.* **2000**, *15*, 138–154. [CrossRef]
37. Welch, I. Seasoned offerings, imitation costs, and the underpricing of initial public offerings. *J. Financ.* **1989**, *44*, 421–449. [CrossRef]
38. Cragg, P.; King, M.; Hussin, H. IT alignment and firm performance in small manufacturing firms. *J. Strat. Inf. Syst.* **2002**, *11*, 109–132. [CrossRef]
39. Berglöf, E.; Pajuste, A. What do firms disclose and why? Enforcing corporate governance and transparency in Central and Eastern Europe. *Oxf. Rev. Econ. Pol.* **2005**, *21*, 178–197. [CrossRef]
40. Hackston, D.; Milne, M.J. Some determinants of social and environmental disclosures in New Zealand companies. *Account. Audit. Account. J.* **1996**, *9*, 77–108. [CrossRef]
41. China Stock Market Accounting Research (CSMAR) Database. Available online: http://www.gtarsc.com/ (accessed on 27 September 2017).
42. Ivashina, V.; Scharfstein, D. Bank lending during the financial crisis of 2008. *J. Financ. Econ.* **2010**, *97*, 319–338. [CrossRef]
43. Brautzsch, H.U.; Günther, J.; Loose, B.; Ludwig, U.; Nulsch, N. Can R&D subsidies counteract the economic crisis?–Macroeconomic effects in Germany. *Res. Policy* **2015**, *44*, 623–633.
44. Ibbotson, R.G. Price performance of common stock new issues. *J. Financ. Econ.* **1975**, *2*, 235–272. [CrossRef]
45. Ibbotson, R.G.; Sindelar, J.L.; Ritter, J.R. Initial public offerings. *J. Corp. Appl. Financ.* **1988**, *1*, 37–45. [CrossRef]
46. Ritter, J.R. Initial public offerings. *Contemp. Financ. Dig.* **1998**, *2*, 5–30.
47. Aggarwal, R.K.; Krigman, L.; Womack, K.L. Strategic IPO underpricing, information momentum, and lockup expiration selling. *J. Financ. Econ.* **2002**, *66*, 105–137. [CrossRef]
48. Arthurs, J.D.; Hoskisson, R.E.; Busenitz, L.W.; Johnson, R.A. Managerial agents watching other agents: Multiple agency conflicts regarding underpricing in IPO firms. *Acad. Manag. J.* **2008**, *51*, 277–294. [CrossRef]
49. Booth, J.R.; Chua, L. Ownership dispersion, costly information, and IPO underpricing. *J. Financ. Econ.* **1996**, *41*, 291–310. [CrossRef]
50. Jegadeesh, N.; Weinstein, M.; Welch, I. An empirical investigation of IPO returns and subsequent equity offerings. *J. Financ. Econ.* **1993**, *34*, 153–175. [CrossRef]
51. Ting, Y.U.; Tse, Y.K. An empirical examination of IPO underpricing in the Chinese A-share market. *China Econ. Rev.* **2006**, *17*, 363–382.
52. Chang, E.; Chen, C.; Chi, J.; Young, M. IPO underpricing in China: New evidence from the primary and secondary markets. *Emerg. Mark. Rev.* **2008**, *9*, 1–16. [CrossRef]
53. Darmadi, S.; Gunawan, R. Underpricing, board structure, and ownership: An empirical examination of Indonesian IPO firms. *Manag. Financ.* **2013**, *39*, 181–200. [CrossRef]
54. Liu, W.; Shi, H.-B.; Zhang, Z.; Tsai, S.-B.; Zhai, Y.; Chen, Q.; Wang, J. The Development Evaluation of Economic Zones in China. *Int. J. Environ. Res. Public Health* **2018**, *15*, 56. [CrossRef] [PubMed]
55. Verardi, V.; Croux, C. Robust regression in Stata. *Stata J.* **2009**, *9*, 439–453. [CrossRef]

Article

Regime-Switching Determinants for Spreads of Emerging Markets Sovereign Credit Default Swaps

Jason Z. Ma [1], Xiang Deng [2], Kung-Cheng Ho [2,*] and Sang-Bing Tsai [3,4,*]

[1] School of Finance, Collaborative Innovation Center of Industrial Upgrading and Regional Finance (Hubei),
     Zhongnan University of Economics and Law, 182 South Nanhu Rd., Wuhan 430073, China;
     zhejma@hotmail.com
[2] School of Finance, Zhongnan University of Economics and Law, 182 South Nanhu Rd.,
     Wuhan 430073, China; xdeng@zuel.edu.cn
[3] Zhongshan Institute, University of Electronic Science and Technology of China, Zhongshan 528402, China
[4] Research Center for Environment and Sustainable Development of China Civil Aviation, Civil Aviation
     University of China, Tianjin 300300, China
*   Correspondence: kcho731101@163.com (K.-C.H.); sangbing@hotmail.com (S.-B.T.);
     Tel.: +86-760-8832-8103 (S.-B.T.)

Abstract: Using the Markov regime switching approach, we investigate the dependency of short term sovereign credit default swap (SCDS) spread changes on a nation's country-specific fundamental factors, local, regional and macroeconomic global factors. We find that the significance of the determinants of SCDS spread changes differ across the two states of our regime-switching model. Specifically, in the good state, the weekly SCDS spread changes are mainly determined by local, regional and fundamental factors; whereas global variables have a stronger influence in the bad regime. In particular, US market returns play a dominant role in influencing the SCDS spread change in the bad state suggesting loss aversion and flight–to–quality behavior of investors. We then examine the cross-sectional differences of the above regime switching effect based on country-specific characters and find that the regime switching effect is associated with a nation's country-specific characters such as openness, economic size and so forth.

---

1. Introduction

The sovereign default of Greece and the ongoing credit crisis in the Euro Zone have raised people's concern on sovereign credit risk. Sovereign credit risk is determined by the country's ability and willingness to re-pay its debt owing to creditors and is reflected in the spread paid for protection offered by the corresponding Sovereign Credit Default Swap (SCDS). Credit risk indicated by a nation's SCDS spread essentially reflects the same fundamental economic condition and market information as the yield of the underlying government bonds. SCDS spread is considered to be a timelier measure of sovereign credit risk than government bond yield spread. Adler and Song [1] compare the behavior of emerging market SCDS spreads and the corresponding bond yields and reject the widely accepted parity relationship between SCDS spreads and bond yields in the literature. Ammer and Cai [2] examine the relationship between SCDS spreads and bond yields for nine emerging market sovereigns and find that these two measures of credit risk deviate significantly in the short run with the former leading the later in price discovery [3]. They attribute such deviation to the higher liquidity in trading of SCDS.

In this paper, we examine the deterministic factors that affect the *variation* of a nation's credit risk as captured by its short-term SCDS spread changes using Markov regime switching model. We focus

on the sovereign credit market of emerging nations which are by far the most liquid. Evidence shows that the SCDS spread changes are affected by different factors in different economic and market conditions and cannot be fully explained by country specific economic fundamental variables [4–6]. Other than country specific economic fundamental variables global factors do play a significant role in influencing the sovereign risk of emerging countries. There are two possible channels through which global factors may exert their influence. First of all, the global effect could be the result of the fundamental economic relation between the emerging country and its global trading partners. Second, it may as well be through the actions in the international financial markets. We expect to witness an elevation of the sovereign risk of an emerging nation if foreign investors lose their appetite on the local financial assets. This type of global effect is expected to be time varying with the effect being more salient during the downturn and/or volatile market condition, in which the market is more prone to a flight-to-quality phenomenon. None of the above-mentioned studies explores the determinants of sovereign credit risk in a state-contingent framework.

Our study contributes to the literature by examining how the influence of different factors may vary in different states of the markets using Markov regime switching model. In particular, rather than classifying the state based on exogenous information that may not be directly relevant to the SCDS market, we let the data to speak for themselves by using the Markov regime switching model to identify the good versus the bad states of the market. This study is one of the most comprehensive empirical studies on SCDS covering a total of 11 emerging market countries across different geographical regions and at different stage of economic development. The time period under investigation is also one of the longest in the literature.

Markov regime switching model is used in a variety of economic and finance research. Goldfeld and Quandt [7] introduce the Markov model for switching regressions in the econometric analysis. Cosslett and Lee [8] use Markov switching in their discrete time models. Hamilton [9] applies Markov switching model to explain the dependence of real output growth on business cycle. In a subsequent paper, Hamilton [10] formally develops the statistical representation to use discrete-time and discrete-space Markov chain to model the transition of unobservable regime switching states in time series data. Since then, the Markov switching framework is widely exploited in a number of studies to model different financial time series that exhibit regime varying effect. For example, previous research uses it to examine stock market returns [11–14]. Clarida et al. [15] investigate the regime shifting effect in the term structure of interest rates. From our knowledge, regime switching model has not been used to study the time series behavior of SCDS. The use of the regime switching model allows us to capture the potential state-contingent behavior of SCDS spread allowing for the influence of different explanatory variables to vary under different economic and market conditions.

In analyzing the regime switching effect of the explanatory variables we also witness a significant difference among the countries in terms of the extent of which these explanatory variables are associated with a country's SCDS spread change. To identify the determinants of these cross-sectional differences we conduct a cross sectional analysis to investigate the variation of the significance of explanatory variables across several categories of sub-groups of our sample of emerging market countries. We expect that the more open an economy and the more it is integrated to the global economy, the stronger will be the influence of regional and global factors on its SCDS spread change. We also expect that there is likely to be a size effect, the smaller the economy, the more vulnerable it is to the regional and global shocks. The findings of this cross-sectional research will be useful for emerging market investors formulating sovereign credit risk management strategy that is specific to the characteristics of each emerging market country.

Both a country's SCDS spread and its government bond spread indicate its credit worthiness. A number of theoretical models are developed to price sovereign debt [16–21]. Duffie and Singleton [22] construct reduced-form models which apply term structure model of interest rates to value corporate and sovereign bonds. Duffie et al. [23] develop a framework to price sovereign bond that takes into

account several risk factors including default, restructuring and liquidity risk. Pan and Singleton [24] explore the nature of default arrival intensity and recovery value implicit in the term structures of SCDS spreads by applying the framework they develop earlier [23,25]. They examine several emerging market countries and show that a single-factor model captures most of the variation in the term structures of spreads of these countries and the risk premiums associated with the unpredictable variation in default arrival intensity are found to be economically significant and highly correlated with several economic measures of the global and local financial market. Delatte et al. [26] and Blommestein et al. [27] assess the influence of the SCDS market on the borrowing cost of SCDS issuing countries during the European sovereign crisis. They conclude that the more severe the distress the more dominant the SCDS market is in the information transmission between SCDS and bond markets. Theoretically the pricing of low-grade bonds issued by emerging economies ought to have no difference to that of developed economies due to the economy of integration. By setting up a series of panel error-correction models, González-Rozada and Eduardo [28] find that global factors, such as the international business cycle, are the determining factors of these spreads. The spread of high yield corporate bonds in developed markets are seen as a reflection of the market sentiment, also referred to as the risk appetite, in this paper. Another explanatory variable, global liquidity, are measured by the international interest rates. And the influence of contagion is taken into account as well, since there was a super excellent systemic event, the 1998 Russian default. The empirical results show that risk appetite and international liquidity explain around 30 percent of the long-run variability of emerging market spreads. And contagion from crisis with systemic effects has a negative influence on spreads. Godlewski [29] proposes a brand new perspective of investigating the connection between bank capital and credit risk. It is rather significant of the regulatory, institutional and legal mechanisms in driving bank capitalization and credit risk taking behavior.

Our study has practical implications that are important to global credit portfolio managers. First, by being able to pinpoint the determinants of the change in SCDS spread in a state-contingent framework, global credit portfolio managers can have a deeper understanding of the evolution of sovereign default risk that is crucial in affecting the risk-return tradeoff of their credit portfolios. Second, the understanding of state-dependent factors for SCDS spread changes can help global credit portfolio managers to formulate dynamic trading strategies that vary across different states of market conditions. Finally, portfolio managers who like to hedge their global credit portfolios using liquid SCDS may find our findings important as the results suggest the need to consider regime dependent hedge ratios to effectively manage credit risk exposure.

The remainder of the paper is organized as follows. Section 2 describes the methodology. Section 3 describes the data and research design. Section 4 provides estimates of OLS regression models for the determination of sovereign CDS spreads and testing Markov regime switching model in Section 5. Section 6 shows cross sectional analysis and Section 7 concludes the paper.

2. Methodology

We consider the following two-state. The number of states of Markov chain can be extended to be larger than 2. Markov regime switching regression model introduced by Hamilton [30] for the weekly change of the spread of SCDS (y_{s_t}) written on a particular emerging market sovereign.

$$y_{s_t} = \beta_{0,s_t} + \sum_{i=1}^{K} \beta_{i,s_t} * x_{i,t} + \varepsilon_{s_t} \tag{1}$$

where x_i's are the factors affecting the SCDS spread change of the country. Indicator variable $s_t = 1$ or 2 denotes the two possible regime switching states which are unobservable and ε_{s_t} is the normally distributed error term with zero mean and standard deviation σ_{s_t} for each $s_t = 1, 2$. All the coefficients and the error term ε_t are allowed to switch between the two states. The transition probability from state 1(2) to state 2(1) over the time period t to $t + 1$ is governed by the Markov transition probability p_{12} (p_{21}),

which is assumed to be constant over time. The distribution of y_{s_t} is fully described by σ_{s_t}, β_{0,s_t}, β_{i,s_t}, p_{11} and p_{22} and $0 < p_{11} < 1$, $0 < p_{22} < 1$. The transition matrix P is therefore represented by

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \tag{2}$$

where $p_{11} + p_{12} = 1$ and $p_{21} + p_{22} = 1$.

Since we can never be certain about what s_t is at any given time t, we can only infer what s_t might be based on what we observe at time t. The probability of having s_t at a given time t to be in regime j is given by

$$\xi_{jt} = Pr(s_t = j | \Omega_t; \theta) \tag{3}$$

where $j = 1, 2$ and Ω_t is the information observed from time 0 up to time t including both the dependent and independent variables and θ is the set of population parameters of the regime switching regression. That is,

$$\theta = (\beta_{i,1}, \ \beta_{i,2}, p_{11}, p_{22}, \ \sigma_1, \ \sigma_2)' \tag{4}$$

Since the regime of the state can either be 1 or 2, the two probabilities $\xi_{1,t}$ and $\xi_{2,t}$ always sum to 1.

The probabilities can be inferred iteratively from $t = 1, 2, \ldots, T$. Under Gaussian assumption of the error terms for the two regimes, the conditional densities needed to perform the iteration are given by:

$$\eta_{j,t} = f(y_t | s_t = j, \ \Omega_{t-1}; \theta) = \frac{1}{\sqrt{2\pi}\sigma_j} exp[-\frac{(y_t - x_t'\beta_j')^2}{2\sigma_j^2}] \tag{5}$$

Thus, the conditional density of the observation is the probability weighted sum of both states, which is:

$$f(y_t | \Omega_{t-1}; \theta) = \sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij}\xi_{i,t-1}\eta_{j,t} \tag{6}$$

The log likelihood function associated with the iteration is then:

$$Log f(\theta) = \sum_{t=0}^{T} log \ f(y_t | \Omega_{t-1}; \theta) = \sum_{t=0}^{T} log(\sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij}\xi_{i,t-1}\eta_{jt}) \tag{7}$$

The parameters θ can be estimated by maximizing the log likelihood function of Equation (7) [31].

3. Data

We use weekly data from the beginning of May 2001 to the end of December 2012 of 11 representative emerging countries in four different geographic regions, namely Asia (China, Korea and Malaysia), Europe (Poland, Russia and Turkey), Latin America (Brazil, Colombia and Venezuela) and Middle East/Africa (Israel and South Africa). The benefit of using weekly rather than daily data is that the former is less noisy than the latter. Monthly data on the other hand will not give us sufficient data points for the regime switching analysis.

The SCDS data are collected from Markit Financial Information Services. In the regressions, we use the weekly changes of SCDS spreads as the dependent variable. Table 1 summarizes the descriptive statistics of the levels of the weekly SCDS spreads of the 11 emerging market countries being studied spanning periods from May 2001 to end of 2012. As can be seen, the SCDS spreads vary considerably across countries with China the lowest (with mean value of 60.02 basis points) and Venezuela the highest (with mean value of 860.52 basis points). The variation in spreads of each country during the sample period is quite substantial as evidenced by the large difference between the maximum and minimum spread values. We observe the spikes in spreads during the period of 2002–2003 and the period of 2008–2009 as a result of the fall of Enron and Leman Brothers respectively. Table 2 summarizes

the descriptive statistics of the weekly changes in SCDS spreads of the 11 countries. The means of the weekly SCDS changes are small in general but the variations in the changes are substantial for all the 11 countries. The high measures of skewness and kurtosis suggest non-normal distributions of SCDS spread changes and reaffirm the regime switching behavior of the SCDS spread changes.

Table 1. Descriptive statistics of the level of sovereign credit default swap (SCDS) spread.

	Brazil	China	Colombia	Israel	Korea	Malaysia	Poland	Russia	South Africa	Turkey	Venezuela
Mean	465.49	60.02	311.86	99.17	90.45	86.52	85.69	244.74	143.24	375.30	860.52
Median	169.83	51.50	179.77	98.75	78.22	81.75	51.67	179.53	140.10	252.16	819.18
Max	3717.13	277.31	1373.22	272.86	708.64	505.40	415.00	1063.64	654.96	1348.33	3218.44
Min	61.14	9.35	67.61	16.92	14.39	11.96	8.17	37.95	24.87	116.78	119.89
St. Dev.	632.45	44.60	251.18	59.35	77.04	58.80	79.96	204.59	85.60	282.63	550.40
Skewness	2.86	1.76	1.61	0.36	2.87	1.53	1.30	1.56	1.68	1.48	1.05
Kurtosis	8.79	3.81	2.81	−0.86	12.16	5.30	1.11	1.76	5.27	1.18	1.85

Notes: This table summarizes the descriptive statistics of the level of weekly SCDS spreads (in basis points) of the 11 countries being studied spanning the time period from May 2001 to end of 2012.

Table 2. Descriptive statistics of SCDS spread change.

	Brazil	China	Colombia	Israel	Korea	Malaysia	Poland	Russia	South Africa	Turkey	Venezuela
Mean	−2.24	−0.09	−1.35	−0.19	−0.15	−0.22	0.06	−2.06	−0.38	−2.04	−1.76
Median	−1.95	−0.13	−1.02	−0.07	−0.43	−0.39	−0.13	−0.99	−0.50	−2.71	−3.30
Max	153.97	17.06	76.62	23.57	43.44	25.13	43.33	64.78	32.23	80.47	192.44
Min	−130.59	−18.15	−69.58	−22.63	−32.55	−23.31	−36.15	−72.15	−36.85	−81.60	−173.98
St. Dev.	41.99	5.67	23.73	7.39	11.45	8.44	11.25	21.96	11.27	28.99	64.24
Skewness	0.53	−0.08	0.34	0.13	0.69	0.29	0.63	−0.29	−0.19	0.16	0.28
Kurtosis	4.97	3.21	3.09	3.41	4.89	2.45	6.28	3.30	2.71	1.67	2.15

Notes: This table summarizes the descriptive statistics of the weekly SCDS spread change (in basis points) of the 11 countries being studied. The spread change data was winsorized at 2% and 98% window.

We consider a number of explanatory variables in the regressions including local financial variables, fundamental economic variables and global financial variables [32–38]. Among the many local financial variables we select, local stock market return and the change in exchange rate against US Dollar (USD) are selected as the explanatory variables representing local market. As widely acknowledged in the literature [4,5], the changes in SCDS spreads tend to be associated with the changes in local financial variables such as local stock index and exchange rates. Local stock indices are denominated in local currency and exchange rates are quoted as local currency per USD. A higher return of the local stock market indicates good market condition that results in a tightening of SCDS spread. We therefore expect the local stock market return and the change in SCDS spreads to move in opposite directions. On the contrary, increasing local currency exchange rate (as denoted by local currency value per USD), suggesting depreciating local currency value and deteriorating local economy, is expected to be related to an increase in SCDS spread.

Besides the above two financial market variables, we also consider the sovereign credit rating of the country as assigned by Standard & Poor's (S&P's) as another potential variable in explaining the variation of the country's SCDS. A country's sovereign rating is considered to be a measure of the fundamental economic and political outlook of the country. It therefore captures information regarding the long-term fundamental condition of the country that may not be captured by the above financial market variables. We expect an improvement in the credit rating (e.g., from A to AA) to be associated with a decrease in the country's SCDS spread.

A country's sovereign risk is also affected by regional and global factors through interactions in international trades, international financial market and geopolitical incidence [39–45]. The world has become more and more integrated. All countries (emerging markets with no exception) have all kinds

of economic and political relation with other countries. One of the contributions of this study is in the examination of how the local, regional, versus global factors are related to a country's sovereign risk under different states of the SCDS market. To achieve this objective, besides the local financial and fundamental variables mentioned above, we also consider the role played by several global financial market variables. Following Longstaff et al. [4] and Fender et al. [5], we use the US stock market return, change in US T-Bill yield and the change of VIX to proxy for the global financial market changes VIX is the CBOE volatility index defined as the forward-looking volatility of US stock market return. A higher return on the US stock market indicates good global market conditions so does an increase in the US T-Bill yield. We therefore expect increases of US stock market return and T-Bill yield lead to a tightening of the SCDS spreads. On the other hand, increasing VIX means a worsening outlook of the global market hence leading to a widening of SCDS spreads for all countries.

We include regional average SCDS as an explanatory variable to capture the regional effect. Economies in the same geographic vicinity (e.g., China, Korea and Malaysia within Asia) are expected to be more integrated with each other than with countries outside the region. For each country, we calculate the regional SCDS spread change as the average SCDS spread change of the other countries in the same region. To better capture the effect of regional influence, we consider both the raw average regional spread change and the residuals of the average regional spread changes after controlling for the global effects. The residual is obtained by running an ordinary least square (OLS) regression of the average regional spread changes against the above global variables (i.e., US stock market return, US T-Bill yield and VIX change). We expect a country's SCDS spread to move in the same direction as its regional SCDS spread.

Table 3 summarizes the explanatory variables providing their descriptions, expected sign of coefficients in the model and data sources.

Table 3. Explanatory variables.

Variable	Expected Sign	Description	Data Source
R_{local}	−	The country's local stock index return	Bloomberg
ΔFX	+	Weekly exchange rate percentage change (per USD)	Bloomberg
$\Delta CDS_{Regional}$	+	Average regional CDS spread excluding the subject country	Markit
$\Delta Rating$	−	Sovereign Rating change (positive change means credit improvement)	S&P
$Res_{R_{local}}$	−	Residual of R_{local} regressed on global variables	Bloomberg
$Res_{\Delta FX}$	+	Residual of ΔFX regressed on global variables	Bloomberg
$Res_{\Delta CDS_{Regional}}$	+	Residual of $\Delta CDS_{Regional}$ regressed on global variables	Bloomberg
$R_{S\&P}$	−	US Stock SP500 weekly return	Bloomberg
ΔVIX	+	Weekly VIX percentage change	Bloomberg
$\Delta TYield$	−	US T-Bill yield weekly difference	Federal Reserve

Notes: This table provides description of the explanatory variables, expected sign of each variable in the model and the data source of the variables.

4. Explaining CDS Return with a OLS Model

We provide the results of two OLS regressions here as benchmark for the regime switching models to be reported later in Section 5. In the first OLS regression (Equation (8)), we regress the weekly SCDS spread change on only the local (both financial and fundamental) and regional variables to see how much the change in sovereign risk can be explained by local and regional factors.

$$\Delta CDS_{i,t} = b0 + b1 * R_{local_{i,t}} + b2 * \Delta FX_{i,t} + b3 * \Delta Rating_{i,t} + b4 * \Delta CDS_{Regional_{i,t}} + \varepsilon_{i,t} \qquad (8)$$

Table 4 shows the result of regression Equation (8). The coefficients for the variable R_{local} are negative for all countries and are all significant at the 1% level (except for Israel at 2%). This is consistent with our expectation that an increase in the return of the local stock market indicates good market condition resulting in a tightening of SCDS spreads. For the variable ΔFX, seven out of the 11 countries have positive and statistically significant coefficients, which is consistent with our expectation that

currency depreciation and sovereign risk are positively related. Three of the remaining four countries have positive coefficients; albeit not statistically significant. The insignificant results for China, Russia and Venezuela could be due to the fact that their pegged exchange rate policies incur no significant variations of exchange rates during the sample period. For the variable $\Delta Rating$, nine countries have the expected negative coefficients but with only one country (Turkey) being statistically significant. This generally insignificant result could be due to the fact that credit rating for most countries tends to stay unchanged for a long period of time. The coefficients for $\Delta CDS_{Regional}$ are significantly positive for all countries, which suggests strong regional economic integration and is consistent with the expectation that a country's SCDS spread moves in the same direction as its regional SCDS spread. Finally, the high adjusted R-squared suggests a substantial amount of the variation of SCDS spreads is explained by these local and regional factors.

Table 4. OLS result of local and regional variables.

Country	Intercept		R_{local}		ΔFX		$\Delta Rating$		$\Delta CDS_{Regional}$		$\overline{R^2}$
	b0	p	b1	p	b2	p	b3	p	b4	p	
Brazil	−1.67	0.19	−1.91	0.00	5.50	0.00	−13.70	0.28	0.19	0.00	0.45
China	−0.03	0.86	−0.35	0.00	−0.59	0.56	−0.36	0.87	0.13	0.00	0.41
Colombia	−0.60	0.42	−0.94	0.00	2.37	0.00	−2.11	0.87	0.18	0.00	0.41
Israel	−0.17	0.49	−0.22	0.02	0.24	0.29	3.26	0.45	0.18	0.00	0.32
Korea	0.04	0.90	−0.38	0.00	2.45	0.00	−2.19	0.64	0.32	0.00	0.51
Malaysia	0.05	0.84	−0.88	0.00	2.50	0.00	−4.98	0.24	0.25	0.00	0.50
Poland	0.27	0.44	−0.52	0.00	1.43	0.00	−0.30	0.97	0.11	0.00	0.40
Russia	−1.54	0.01	−1.21	0.00	0.53	0.38	4.51	0.38	0.45	0.00	0.52
South Africa	−0.23	0.50	−0.94	0.00	1.29	0.00	−5.40	0.37	0.39	0.00	0.52
Turkey	−1.62	0.05	−1.25	0.00	4.24	0.00	−37.18	0.00	0.41	0.00	0.50
Venezuela	0.26	0.92	−2.76	0.00	0.75	0.16	−2.54	0.51	0.39	0.00	0.11

Notes: This table presents the OLS regression result of the weekly SCDS change of the 11 emerging market countries on local and regional variables in Equation (8). Left column of each variable reports the estimated coefficient and right column reports the *p*-value of *t*-test.

In the second OLS regression, we regress the weekly SCDS spread change not only on the local and regional variables but also on the global variables. To clearly separate the impact of the local and regional factors from the impact of the global factors on a country's SCDS spread, we use the residuals of the local and regional variables obtained from first regressing each of these variables against the three global factors as our explanatory variables representing the pure local and regional effects. For example, in order to strip out the global effects, we first regress the regional CDS spread change on the US Stock Return, changes in T-Bill yield and VIX (Equation (9)) and use the residuals (ε) of this regression as a new explanatory variable-regional CDS residual, in the OLS regression of each country's SCDS change.

$$\Delta CDS_{Regional_{i,t}} = a0 + a1 * R_{S\&P_{i,t}} + a2 * \Delta VIX_{i,t} + a3 * \Delta TYield_{i,t} + \varepsilon_{i,t} \qquad (9)$$

Note that the regional CDS residual (denoted as $Res_{\Delta CDS_{Regional}}$) is orthogonal to $R_{S\&P}$, ΔVIX and $\Delta TYield$. Thus, using this regional CDS residual allows us to eliminate the effect of global factors on regional CDS in explaining the change in a country's SCDS spread. The same applies to local stock return residual, $Res_{R_{local}}$ and the residual for exchange rate percentage change, $Res_{\Delta FX}$, which are obtained in a similar fashion.

The second OLS regression can be expressed as:

$$\begin{aligned} \Delta CDS_{i,t} = \ & b0 + b1 * Res_{R_{local_{i,t}}} + b2 * Res_{\Delta FX_{i,t}} + b3 * \Delta Rating_{i,t} + b4 \\ & * Res_{\Delta CDS_{Regional_{i,t}}} + b5 * R_{S\&P_{i,t}} + b6 * \Delta VIX_{i,t} + b7 * \Delta TYield_{i,t} \\ & + \varepsilon_{i,t} \end{aligned} \qquad (10)$$

Table 5 shows the result of regression Equation (10). The sign and significance of the coefficients of the local and regional variables are essentially consistent with those of the first OLS regression results as reported above when the global factors are ignored. Now we turn to the global variables in Equation (10). The coefficients for $R_{S\&P}$ are all negative and statistically significant. This is consistent with our expectation that a higher US stock market return indicates good global market conditions leading to lower sovereign risk. Except for Israel, all the countries have positive coefficients for ΔVIX and this is consistent with our expectation of the positive relation between VIX, as a global fear factor and sovereign risks. Nevertheless, only the coefficients for China, Russia and Turkey are statistically significant. Note that China, Russia and Turkey are relatively larger economies within our sample of countries and are well integrated into the global economy. We therefore expect these countries are likely to be more sensitive to global risk outlook as indicated by VIX being the forward-looking volatility of the US stock market return. The coefficients for $\Delta TYield$ are all positive and significant for most countries (except for Israel, Poland and South Africa), which is again consistent with our expectation.

Table 5. OLS result of all variables.

Country	Intercept		$Res_{R_{local}}$		ΔRes_{FX}		$\Delta Rating$		$Res_{\Delta CDS_{Regional}}$		$R_{S\&P}$		ΔVIX		$\Delta TYield$		$\overline{R^2}$
	b	p	b1	p	b2	p	b3	p	b4	p	b5	p	b6	p	b7	p1	
Brazil	−2.39	0.07	−3.11	0.00	6.03	0.00	−10.73	0.42	0.04	0.43	−4.50	0.00	0.21	0.15	−29.43	0.01	0.42
China	−0.12	0.50	−0.37	0.00	−0.30	0.78	0.48	0.84	0.05	0.00	−0.85	0.00	0.04	0.05	−4.15	0.01	0.36
Colombia	−1.40	0.08	−0.79	0.00	2.70	0.00	−4.65	0.74	0.12	0.00	−3.81	0.00	0.06	0.46	−18.10	0.02	0.31
Israel	−0.17	0.51	−0.04	0.68	0.62	0.01	4.00	0.39	0.09	0.00	−1.14	0.00	−0.02	0.48	−2.50	0.31	0.23
Korea	−0.19	0.59	−0.50	0.00	3.52	0.00	−1.65	0.74	0.06	0.00	−1.71	0.00	0.06	0.11	−8.24	0.01	0.45
Malaysia	−0.28	0.29	−0.98	0.00	3.72	0.00	−2.77	0.55	0.06	0.00	−1.14	0.00	0.04	0.13	−10.77	0.00	0.41
Poland	0.07	0.85	−0.63	0.00	1.75	0.00	−0.64	0.94	0.10	0.00	−1.77	0.00	0.04	0.34	−2.25	0.50	0.37
Russia	−2.25	0.00	−1.84	0.00	−0.08	0.90	5.91	0.28	0.25	0.00	−3.27	0.00	0.14	0.06	−18.64	0.00	0.45
South Africa	−0.39	0.28	−0.84	0.00	1.41	0.00	−3.26	0.60	0.12	0.00	−1.85	0.00	0.06	0.11	−3.25	0.32	0.41
Turkey	−1.94	0.02	−1.48	0.00	5.45	0.00	−36.30	0.00	0.30	0.00	−3.70	0.00	0.21	0.02	−13.18	0.09	0.48
Venezuela	−2.18	0.36	−2.76	0.00	0.76	0.15	−1.54	0.68	0.33	0.00	−6.22	0.00	0.44	0.10	−45.40	0.04	0.17

Notes: This table presents the OLS regression result of the weekly SCDS change of the 11 emerging market countries on all variables including local, regional and global variables in Equation (10). Left column of each variable reports the estimated coefficient and right column reports the *p*-value of *t*-test.

The OLS regression results show that local, regional and global factors are all important in determining the spread change of SCDS consistent with the findings in Longstaff et al. (2011) and Fender et al. (2012). We now turn to the Markov regime switching model to study how these factors evolve with the switching of market regimes.

5. Markov Regime Switching Analysis

We use a two-state Markov regime switching model to explain how the weekly change of a country's SCDS spread is related to the set of explanatory variables. To better capture and identify the effect of individual variables on the change of SCDS spreads, we categorize the variables into three groups (i.e., local, regional and global) and examine the effects of different subsets of these variables. Our goal is to find out if and how the explanatory power of these groups of variables differs across the two regimes. Specifically, we consider the regime-switching model using:

(a)   Only local and regional variables
(b)   All local, regional and global variables.

As confirmed by our preliminary OLS regression results reported in Section 4, SCDS spread change is affected by local, regional and global factors. The research question we are asking here is: Do these different groups of factors behave differently across different states of the market?

Equation (11) depicts model specification (a) where the local financial and fundamental variables, namely local stock return (financial), exchange rate change (financial), rating change (fundamental)

and the regional SCDS changes are used as explanatory variables, while leaving out the global financial market variables.

$$\Delta SCDS_{s_t} = \beta_{0,s_t} \quad + \beta_{1,s_t} * R_{local\,i,t}$$
$$+ \beta_{2,s_t} * \Delta FX_{i,t} \; + \; \beta_{3,s_t} * \Delta Rating_{i,t} + \beta_{4,s_t} * \Delta CDS_{Regional_{i,t}} \; + \varepsilon_{s_t} \tag{11}$$

We consider a two-state regime switching regression model where all coefficients and error terms are allowed to take on different values in the two states as denoted by s_t. The good state is defined as the market condition that is characterized by tightening SCDS spreads (negative changes) and low volatility, while the bad state is the market condition with widening SCDS spreads (positive changes) and high volatility. We calibrate this regime-switching model for the SCDS spread changes of each country and the results are reported in Table 6. Our findings regarding the regime switching effect of each explanatory variable.

Table 6. Regime switching regression summary—model specifications (a).

Country	*Intercept*	R_{local}	ΔFX	$\Delta Rating$	$\Delta CDS_{Regional}$
	fi_0	fi_1	fi_2	fi_3	fi_4
Brazil	−0.222	−1.023 ***	0.827 **	−11.170 *	0.114 ***
	3.741	−1.403 **	10.724 ***	−24.400	0.140 ***
China	0.124 *	−0.015	0.366	−8.602	0.713 ***
	6.911	−1.425	0.654	5.446	0.104 ***
Colombia	−0.433	−0.804 ***	1.097 ***	−2.832	0.191 ***
	2.067	−1.878 **	6.743 ***	2.492	0.206 ***
Israel	−0.184 ***	4.390 *	0.176 ***	15.496	0.151 ***
	0.823 *	−0.352 **	−0.288	1.451	0.353 ***
Korea	−8.611	4.276	0.333 ***	4.973 ***	1.104 ***
	0.144	0.231	−2.378	25.795	1.676 ***
Malaysia	0.241 ***	−0.117 ***	9.244	−25.140	0.943 ***
	−0.285	−6.037	2.074 **	−7.935 ***	0.589 ***
Poland	−5.518 ***	−4.139 ***	6.499 ***	−0.785 ***	1.557 ***
	0.108 ***	3.197 ***	1.159 ***	3.197 ***	0.583 ***
Russia	0.484	−0.597 ***	−0.863 *	−1.991	0.713 ***
	1.949	−2.241 ***	0.751	−17.560	0.632 ***
South Africa	0.221	−0.337 ***	0.285 **	−26.220 ***	0.350 ***
	−1.974	−1.173 ***	−0.263	25.655	1.832 ***
Turkey	−0.303	−0.881 ***	2.816 ***	−27.830 ***	0.699 ***
	2.871	2.757 *	12.517 ***	21.942	0.798 ***
Venezuela	2.702 ***	−0.895 ***	0.848	−4.209 ***	0.939 ***
	−5.347	0.233	0.283	4.407	0.285 ***

Notes: This table summarizes the Markov regime switching regression results of specification in Equation (11). First row of each country reports the estimated coefficients for the good Markov state and second row reports those of the bad state. ***, **, * denote statistical significance at 1%, 5% and 10%. Significance is based on t-statistics.

R_{local}—*Local stock index return:* We expect local stock return affect SCDS change in a negative way, that is, a positive local stock return indicating a good market condition, hence the SCDS spread should tighten (negative change). The coefficient is indeed negative in the good state for all countries except Israel and Korea. For most of the countries, the coefficient is also statistically significant in the good state. Taking as an example, for Brazil the estimated coefficient is −1.023 which means that each percentage point increase in the Brazil local stock market return is associated with a 1.023 basis point decrease in Brazil SCDS spread. The effect is found to be weaker in the bad state, the coefficient for quite a few countries (e.g., China, Korea, Malaysia and Venezuela) are insignificant. It seems that the local stock index return is more influential to a SCDS spread change when the economy is good, while in the bad time, other factors weigh in (refer to below discussion on global factors).

ΔFX—*Exchange rate percentage change (domestic/USD):* Venezuela and China adopt a pegging currency policy which renders no meaningful effect of ΔFX on their SCDS spread changes. Ignoring these two countries, we observe positive exchange rate change associated with positive

SCDS spread change; and similar to local stock market return as outlined above, we witness the same regime-contingent behavior for the effect of exchange rate change that it is in general significant in both good and bad states but is more influential in the good state than in the bad state of the SCDS market. For example, Israel, Korea, Poland, South Africa all report positive and strongly significant coefficients.

$\Delta Rating$—*Rating Change*: In the good state, the coefficient for rating change is negative and significant for five countries. In the bad state, the negative effect is only significant for Malaysia. It seems that the SCDS spread change of most of the countries is not significantly related to rating change but if it does, it mostly happens in the good state. Note that rating is a fundamental factor capturing a country's political, economic and other country-specific characters. These characters change infrequently, thus any foreseeable significant change may have already been captured in the SCDS spread before the actual rating changes.

$\Delta SCDS_{Regional}$—*Regional CDS*: This variable has significant effect for all countries in both the good state and the bad state. It affirms that countries in close geographic vicinity have strong relations with each other. No matter the economies is in a good time or bad time, these countries are strongly inter-coupled together.

In general, the above findings suggest that the local and fundamental variables have stronger influence on the SCDS change in the good state than in the bad state. This is consistent with our expectation that the governing role of local and fundamental variables may be weakened as global factors exert more influence during market downturn (i.e., in the bad state of our regime-switching process).

We hypothesize that global variables have stronger influence in a bad regime of SCDS spread change. In model specification (b), we test this hypothesis by using not only local, fundamental and regional variables but also including global factors in our regime-switching model (see Equation (12)). The estimation results are reported in Table 7.

$$
\begin{aligned}
\Delta SCDS_{s_t} = \beta_{0,s_t} \quad & + \beta_{1,s_t} * Res_{R_{local_{i,t}}} \\
& + \beta_{2,s_t} * Res_{\Delta FX_{i,t}} + \beta_{3,s_t} * \Delta Rating_{i,t} + \beta_{4,s_t} * Res_{\Delta CDS_{Regional_{i,t}}} \\
& + \beta_{5,s_t} R_{S\&P_{i,t}} + \beta_{6,s_t} * \Delta VIX_{i,t} + \beta_{7,s_t} * \Delta TYield_{i,t} + \varepsilon_{s_t}
\end{aligned}
\tag{12}
$$

$Res_{R_{local}}$—*Local stock index return residual*: Now the coefficient of this variable is negative and significant for 10 countries (with Israel being the exception) in the good state with Poland having the most negative coefficient of -5.265 with Israel being the exception. Only 5 countries have both negative and significant coefficient for Res_R_{local} in the bad state. This demonstrates that, after stripping out the global effect in the local stock index return, it has a stronger effect on a country's SCDS spread change in the good state while tends to be weaker in the bad state. This reinforces our expectation that the local stock market is more influential on SCDS spread change in the good state.

$Res_{\Delta FX}$—*Exchange rate percentage change residual*: After removing the global factor influence, the coefficient of exchange rate percentage change residual is positive and significant for 10 countries out of 11 (except for China) in the good state. But in the bad state, it has positive and significant effect for only five countries, namely Brazil, Israel, Korea, Malaysia and Poland. Consistent with the previous results, we conclude that exchange rate percentage change residual contributes to SCDS spread change strongly in the good state and but relatively weakly in the bad state. This is consistent with our expectation that the governing role of local factors is limited in the bad state with the contemporaneous influence of global factors.

$\Delta Rating$—*Rating change*: The expected negative effect of rating change is significant in six countries (Colombia, Israel, Korea, Poland, Russia and Turkey) in the good state. It is significant in only one country, Russia, in the bad state. This tells us that rating change is also a good state player which is consistent with our expectation.

$Res_{\Delta CDS_{Regional}}$—*regional CDS residual*: It can be seen that regional SCDS residual has significant and positive effect in ten countries in the good state but the positive effect is significant only for four countries in the bad state, namely Malaysia, Poland, South Africa and Israel. After removing

the effect of the global factors, regional SCDS residual is better able to capture the regional effect and influences the SCDS spread more heavily in the good state. This is consistent with our hypothesis that regional factor is expected to influence SCDS spread change more in the good state than in the bad state.

$R_{S\&P500}$—*US Stock S&P 500 Return*: The negative effect of S&P 500 stock index returns on emerging market countries' SCDS spread change is overwhelmingly significant in almost all countries in both good and bad states. But a closer examination at the coefficients reveals that S&P 500 stock index return contributes much stronger in the bad state than in the good state. The magnitude of the coefficient for the bad state is typically 2 to 5 times that for the good state. For a few countries, the difference between good and bad states is even larger. For example, for China, the coefficient is −0.281 in the good state however it is −2.016 in the bad state. The results show that the impact of US stock market return on China's SCDS spread change magnifies to ~seven folds in the bad state than in the good state. The findings are consistent with our expectation that global factors are more important in determining emerging market's SCDS spread change in the bad state.

ΔVIX—*VIX percentage change:* VIX is a measure of the implied volatility of S&P 500 index options. It represents the market's expectation of US stock return volatility over the next 30 days period and is often referred to as the fear index. We expect VIX to also play a strong role in affecting emerging market's SCDS spread change. But surprisingly, the effect is much weaker than that of S&P 500 return. From Table 7, we see that ΔVIX is only significant with the expected positive effect for three countries in the good state (Colombia, Korea and Malaysia). It is significant in the bad state for only two countries (Israel and Korea) in the expected direction. This suggests that SCDS spread change is only weakly sensitive to VIX movement contrary to its sensitivity to stock index return.

$\Delta TYield$—*US T Bill yield:* In general, the effect of US T-Bill yield is also weak. This variable is significant for five countries (Brazil, China, Korea, Russia and Venezuela) in the good state, while in the bad state it is significant for Israel and Russia but not in the expected direction.

Table 7. Regime switching regression summary—model specifications (b).

Country	*Intercept*	$Res_{R_{local}}$	$Res_{\Delta FX}$	$\Delta Rating$	$Res_{\Delta CDS_{Regional}}$	$R_{S\&P500}$	ΔVIX	$\Delta TYield$
	fi$_0$	fi$_1$	fi$_2$	fi$_3$	fi$_4$	fi$_5$	fi$_6$	fi$_7$
Brazil	−1.249 ***	−0.727 ***	1.831 ***	4.928	1.830	−3.659 ***	9.038	−10.121 ***
	−30.230 ***	−0.339	11.684 ***	80.565	−0.599	−7.194 ***	−1.394 **	−27.261
China	−0.425 ***	−4.452 *	−0.849	−3.004	1.851 ***	−0.281 ***	−3.719 ***	−3.607 ***
	0.146	−0.825 ***	−1.612	−5.086	−2.971*	−2.016 ***	−5.887	2.612
Colombia	−1.479	−0.809 ***	1.636 ***	−9.779 *	8.008 ***	−2.505 ***	0.138 ***	−4.788
	1.966	0.194	2.634	−24.57	0.128	−3.774 ***	−4.181	10.758
Israel	−0.201 **	0.119 ***	0.350 ***	−2.416 ***	4.047 ***	−0.315 ***	−0.008	−1.491
	0.550 ***	5.197	0.710 **	5.217	0.101 ***	−1.180 ***	8.821 **	5.632 ***
Korea	−0.103	−0.499 ***	1.264 ***	−13.801 ***	2.5884 **	−0.583 ***	4.928 **	−8.623 ***
	−0.122	1.096 *	6.489 ***	11.219	1.303	−2.071 ***	0.589 ***	44.304
Malaysia	−0.328 **	−0.186 **	1.358 **	−12.320	4.525 *	−0.242	3.545 *	−0.471
	0.765 *	−0.485 **	4.432 ***	5.237	4.559 ***	−1.961 ***	−0.114 **	−3.576
Poland	−0.124 **	−5.265 **	0.142 ***	−6.770 *	1.244 ***	−0.129 ***	9.006	2.443
	3.139	−0.612 *	1.974 ***	0.167	7.933 **	−3.787 ***	−0.265 ***	1.535
Russia	0.198 ***	−1.205 ***	1.958 ***	−3.312 ***	0.419 ***	−3.890 ***	−5.852 ***	−11.701 ***
	6.435 ***	1.279 ***	−4.397 ***	−8.370 ***	−0.757 ***	4.162 ***	−0.113 ***	34.018 ***
South Africa	−0.118	−0.629 ***	1.186 ***	−2.156	0.169 ***	−1.826 ***	−8.533	−0.806
	−3.322	−4.190	−0.618	−0.305	0.454*	−7.107*	−0.695	−2.407
Turkey	−0.609	−1.054 ***	5.317 ***	−54.330 ***	0.569 ***	−5.202 ***	6.530	−2.242
	−5.299	−7.917 ***	−3.203	−116.201	0.487	21.228 ***	9.172	19.485
Venezuela	0.188	−1.057 ***	1.595 **	−1.214	0.981 ***	−5.199 ***	3.802	−24.240*
	0.943	−7.450 **	−1.023	2.805	0.110	−16.270 ***	−1.288	17.403

Notes: This table summarizes the Markov regime switching regression results of specification in Equation (12). First row of each country reports the estimated coefficients for the good state and second row reports those of the bad state. ***, **, * denote statistical significance at 1%, 5% and 10%. Significance is based on t-statistics.

The above findings show that global influence magnifies itself mainly through the US stock index return. Especially, the effect is exacerbated in the bad state. This is consistent with loss aversion

and flight-to-quality-assets behavior of investors when the SCDS market becomes volatile. The findings in this section are consistent with our hypothesis. The SCDS spread change of emerging market countries is more subject to the changes of local, fundamental and regional variables when the market is in the good regime; while in the bad regime, the global effect as represented by the US stock index return, is dominant in determining the SCDS spread change. The other global variables such as the change in the VIX index and the change in the US T-Bill yield have limited influence on SCDS spread change regardless of the state of the market.

Figure 1 plots the smoothed probability of Regime 1, the good state, P[$St = 1$], fitted to the 11 countries' CDS spread changes for the Regime Switching Model specification (b) which is specified in Equation (12). The values of the smoothed probability series are typically very close to either zero (Regime 2, bad state) or one (Regime 1, good state) and the smoothed probability series do not frequently switch between the good state and the bad state. The smoothed probability is of interest in economically interpreting the regime switching behavior of the CDS spread changes and determines if and when regime switches occur. During the 2008–2009 financial crisis, all the 11 countries entered regime 2 (bad state) for a certain period of time and then exit the bad state during the 2010–2011 recovery period. During the 2003–2007 economic expansion all countries were in the good state.
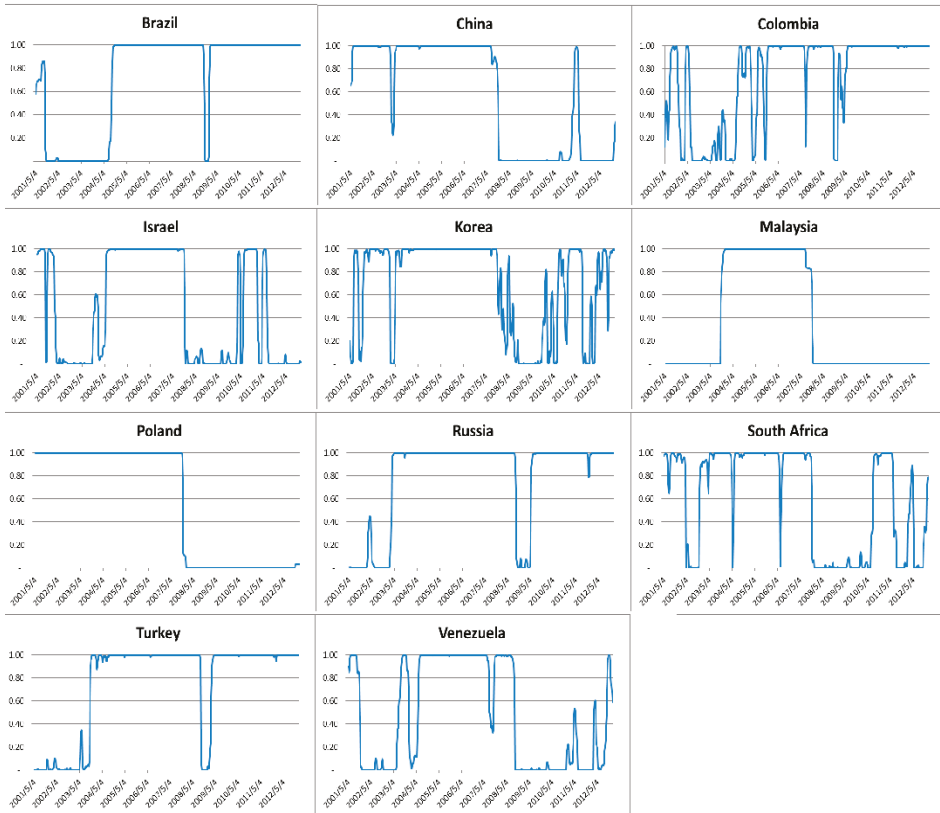

Figure 1. Smoothed Probability of Regime 1 (good state) for Markov-Switching Model.

6. Cross Sectional Analysis

From the empirical analysis in the previous section we see that all the explanatory variables have an impact on the SCDS spread changes where local and regional variables show more influence in the good state and global variables have more influence in the bad state. We also witness a significant difference among the countries in terms of the extent of which these explanatory variables are associated with a country's SCDS spread change. What are the determinants of these cross-sectional differences? The answer to this research question will be useful for emerging market investors in formulating sovereign credit risk management strategy that is specific to the characteristics of each emerging market country. First of all, we expect the more open an economy and the more it is integrated to the global economy, the stronger will be the influence of regional and global factors on its SCDS spread change. Second, there is likely to be a size effect, the smaller the economy, the more vulnerable it is to the regional and global shocks. Thus, both regional and global factors may play a more important role in affecting smaller country's SCDS spread changes. Finally, there may be a regional effect. For example, due to their geographical and/or cultural characteristics, Asian countries may behave differently from European countries in terms of the determinants of their sovereign credit risks.

In conducting our cross-sectional analysis, we classify our countries into different subgroups independently based on four country-specific indicators representing openness/global integration (Kaopen Index; trade-to-GDP ratio; foreign direct investment (FDI) to GDP ratio) and the size of the economy (as proxied by GDP). We then examine and compare the sensitivities of SCDS spread change to the representative explanatory variables across the subgroups. Table 8 summarizes the average values of four market and economic indicators of the 11 countries during period of 2001 to 2012. From Table 8, we observe significant cross-sectional variations of country characteristics as captured by these indicators. For example, the trade-to-GDP ratio of Malaysia is almost eight times that of Brazil; whereas the FDI ratio of Israel is again almost eight times that of Korea.

Table 8. Average values of indicators for each country.

Country	Kaopen Index	Import + Export (% of GDP)	FDI, Net Inflows (% of GDP)	GDP (MM)
Brazil	0.03	20.56	2.71	1,320,903.88
China	−1.17	67.82	3.67	2,753,506.67
Colombia	−0.29	29.20	3.56	200,148.08
Israel	2.13	55.35	3.83	175,681.59
Korea	0.13	74.00	0.49	862,100.87
Malaysia	−0.19	162.42	3.01	183,676.38
Poland	−0.05	64.96	3.46	364,760.02
Russia	−0.13	45.62	2.69	1,085,489.40
South Africa	−1.17	41.90	1.82	261,131.15
Turkey	−0.72	47.65	1.81	535,179.14
Venezuela	−0.62	25.75	1.05	202,169.33

Notes: This table summarizes average values of four market and economic indicators of the 11 countries being studied during the period of 2001 to 2012.

To examine how these country-specific factors is related to the influence of different variables on SCDS, we divide the countries into two subgroups. The first subgroup of each indicator consists of six countries with lower values of the indicator and the second subgroup consists of the remaining five countries with higher values of the indicator. Table 9 shows the sub-grouping of countries for each indicator.

As outlined in the previous sections, Equation (12) is our most comprehensive regime switching model that incorporates all local, regional and global factors. Based on regression results of Equation (12) as shown in Table 7, we select the four most significant explanatory variables to conduct the cross-sectional analysis. The four variables are $Res_{R_{local}}$, $\Delta Rating$, $Res_{\Delta CDS_{Regional}}$ and $R_{S\&P}$. Table 10 shows the average of the coefficients of these four variables (obtained from running our regime-switching model of Equation (12)) of the countries within each subgroup. The columns labelled by S1 (S2) consist of results for the good (bad) state. For example, the average coefficient of Res_R$_{local}$ for the closed group for Kaopen in the good state (S1) is denoted as −1.364 **. The closed group for Kaopen has six countries namely China, South

Africa, Turkey, Venezuela, Colombia and Malaysia and their coefficients of Res_R_{local} in the good state (S1) are respectively, -4.452 *, -0.629 ***, -1.054 ***, -1.057 ***, -0.809 *** and -0.186 * as shown in Table 7. The average in value therefore equals to -1.364 and the average in statistical significance is at 5% level (i.e., **). Below is a summary of the main findings from examining the average variations of the average coefficients of these four variables across subgroups of each indicator.

Table 9. Sub-grouping of Countries.

	Kaopen	Import + Export (% of GDP)	FDI, Net Inflows (% of GDP)	GDP Size
subgroup of lower indicator value	China South Africa Turkey Venezuela	Brazil Venezuela Colombia South Africa	Korea Venezuela Turkey South Africa	Israel Malaysia Colombia Venezuela
	Colombia Malaysia	Russia Turkey	Russia Brazil	South Africa Poland
subgroup of higher indicator value	*Russia* Poland Brazil Korea	*Israel* Poland China Korea	*Malaysia* Poland Colombia China	*Turkey* Korea Russia Brazil
	Israel	Malaysia	Israel	China

Notes: This table shows the sub-grouping of countries for each indicator. We divide the countries into two subgroups for the four indicators. The first subgroup consists of six countries with lower value of the indicator and the second subgroup consists of the remaining five countries with higher value of the respective indicator.

$Res_{\Delta CDS_{Regional}}$: This factor, in general, has larger impact on open countries than closed countries for the openness indicator subgroups. This is especially true in the bad state (S2), consistent with the expectation that the open countries are more integrated with the regional economies while the contagion effect being more salient in the bad state. Comparing the two size indicator subgroups, $Res_{\Delta CDS_{Regional}}$ has larger impact on small countries both in good and bad states. This is expected because, the smaller the economy, the easier it could be influenced by the surrounding economies. Especially in a global crisis, smaller countries with less diversified economies would be more affected because economic links are more important for such countries than larger countries. Larger countries are expected to be less affected by the surrounding economies than by its own local factors. Besides, larger countries tend to have more diverse economic composition and thus less susceptible to industry-specific shocks that propagate across borders.

$\Delta Rating$: Rating change has stronger impact in the good state than in the bad state for all indicator subgroups. This is consistent with our expectation that fundamental factors as captured by rating plays a stronger role in the good state, whereas its effect is weakened in the bad state as other financial factors dominate. The influence of rating is also more significant for closed than for open countries, suggesting fundamental factors are more influential in dictating the sovereign risk of closed countries.

Table 10. Cross sectional analysis of selected explanatory variables.

	Indicator	Subgroup	$Res_{R_{local}}$		$\Delta Rating$		$Res_{\Delta CDS_{Regional}}$		$R_{S\&P}$	
			S1	S2	S1	S2	S1	S2	S1	S2
Openness/	kaopen	Closed	-1.364 **	-3.445 *	-13.802	-23.029	2.683 **	0.461	-2.542 **	-1.65 **
Global Integration		Open	-1.515 **	1.324	-4.274 *	17.760	2.026 *	1.596 *	-1.715 **	-2.014 **
	Import + Export (% of GDP)	Closed	-0.913 **	-3.070 *	-10.977 *	-11.021	1.996 **	-0.030	-3.713 **	-1.492 **
		Open	-2.056 *	0.874 *	-7.663 *	3.351	2.851 **	2.184 *	-0.31 *	-2.203 **
	FDI, net inflows (% of GDP)	Closed	-0.861 **	-2.920 *	-11.648 *	-5.055	1.092 **	0.166	-3.393 **	-1.208 **
		Open	-2.118 *	0.693 *	-6.858 *	-3.809	3.935 **	1.949 *	-0.694 *	-2.543 **
Size	GDP size	Small	-1.304 **	-1.224 *	-5.776 *	-1.910	3.162 **	2.214 *	-1.702 **	-5.679 **
		Big	-1.587 **	-1.341 *	-13.904 *	-7.583	1.451 *	-0.507	-2.723 **	2.821 **

Notes: This table shows the average of the coefficients of the selected four explanatory variables (obtained from running our regime-switching model of Equation (12)) of the countries within each subgroup. The columns labelled by S1 (S2) consist of results for the good (bad) state.

$Res_{R_{local}}$ and $R_{S\&P}$: For open countries, in the bad state, their SCDS spreads are more affected by $R_{S\&P}$; whereas in the good state, they are more affected by $Res_{R_{local}}$. This asymmetry reflects the market sentimental effect of flight-to-quality that only manifests itself in the bad state. For closed countries, their SCDS spreads are more affected by $Res_{R_{local}}$ in the bad state, while $R_{S\&P}$ plays a stronger role in the good state. Finally, we find that $R_{S\&P}$ has the strongest influence for small countries in the bad state. This could be attributed to the fact that economic links are more important for smaller countries than larger countries with diverse economies. The strong effect of $R_{S\&P}$ in the good state suggests that the S&P 500 return captures fundamental global improvement that even benefits closed economies.

7. Conclusions

The weekly change of emerging market sovereign CDS spreads is affected by many market and economic variables. This paper examines the effect of a broad range of such variables including local financial, fundamental and global financial variables. The objective of the paper is to find the varying behavior of these variables on emerging market sovereign CDS in a two-state Markov regime switching environment.

We find that local, regional and fundamental variables such as local stock index return, exchange rate change, regional SCDS spread and credit rating change of the country influence the SCDS change more when the market is in a good state. Whereas global variables, such as US stock index return, have in general stronger influence in a bad state. Especially, when the regime is in a bad state, the single factor of US stock market return dominates other factors and its significance is much larger in the bad state than it is in the good state. This is consistent with the risk aversion and flight-to-quality assets behavior of investors when global market becomes volatile.

We also conduct cross sectional analysis to examine the behavior of the same explanatory variable on countries of different macroeconomic characters and reveal valuable findings. First, we find that more open countries are more integrated with the regional economies with the contagion effect being more salient in the bad state. Second, smaller countries with less diversified economies would be more affected a global crisis because economic links are more important for such countries than larger countries. Third, the influence of rating is more significant for closed than for open countries indicating that fundamental factors are more influential in dictating the sovereign risk of closed countries. Finally, we find that the market sentimental effect of flight-to-quality magnifies in the bad state and that $R_{S\&P}$ has the strongest influence for small countries in the bad state which could be attributed to the fact that economic links are more important for smaller countries than larger countries with diverse economies.

Our work opens a new page for studies on how CDS spreads vary with regimes and how various factors play their roles. Further research can certainly be done to improve the results in this paper as well as to expand the discussions in this paper.

Author Contributions: Writing: J.Z.M. and X.D.; Providing idea and data: J.Z.M. and X.D.; Revising and editing: K.-C.H. and S.-B.T.

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

References

1. Adler, M.; Song, J. The behavior of emerging market sovereigns' credit default swap premiums and bond yield spreads. *Int. J. Financ. Econ.* **2010**, *15*, 31–58. [CrossRef]
2. Ammer, J.; Cai, F. Sovereign CDS and bond pricing dynamics in emerging markets: Does the cheapest-to-deliver option matter? *J. Int. Financ. Mark. Inst. Money* **2011**, *21*, 369–387. [CrossRef]
3. Coudert, V.; Gex, M. Credit default swap and bond markets: Which leads the other? *Financ. Stab. Rev.* **2010**, *14*, 161–167.

4.   Longstaff, F.A.; Pan, J.; Peterson, L.H.; Singleton, K.J. How sovereign is sovereign credit risk. *Am. Econ. J. Macroecon.* **2011**, *3*, 75–103. [CrossRef]

5.   Fender, I.; Hayo, B.; Neuenkirch, M. Daily pricing of emerging market sovereign CDS before and during the global financial crisis. *J. Bank. Financ.* **2012**, *36*, 2786–2794. [CrossRef]

6.   Tobback, E.; Naudts, H.; Daelemans, W.; Fortuny, E.J.; Martens, D. Belgian economic policy uncertainty index: Improvement through text mining. *Int. J. Forecast.* **2018**, *34*, 355–365. [CrossRef]

7.   Goldfeld, S.M.; Quandt, R.E. A Markov model for switching regressions. *J. Econ.* **1973**, *1*, 3–15. [CrossRef]

8.   Cosslett, S.R.; Lee, L.F. Serial correlation in latent discrete variable models. *J. Econ.* **1985**, *27*, 79–97. [CrossRef]

9.   Hamilton, J.D. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **1989**, *57*, 357–384. [CrossRef]

10.   Hamilton, J.D. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 1994.

11.   Turner, C.M.; Startz, R.; Nelson, C.R. A markov model of heteroskedasticity, risk, and learning in the stock market. *J. Financ. Econ.* **1989**, *25*, 3–22. [CrossRef]

12.   Perez-Quiros, G.; Timmermann, A. Firm size and cyclical variations in stock returns. *J. Financ.* **2000**, *55*, 1229–1262. [CrossRef]

13.   Alexander, C.; Dimitriu, A. Indexing, cointegration and equity market regimes. *Int. J. Financ. Econ.* **2005**, *10*, 213–231. [CrossRef]

14.   Alexander, C.; Kaeck, A. Regime dependent determinants of credit default swap spreads. *J. Bank. Financ.* **2008**, *32*, 1008–1021. [CrossRef]

15.   Clarida, R.H.; Sarno, L.; Taylor, M.P.; Valente, G. The role of asymmetries and regime shifts in the term structure of interest rates. *J. Bus.* **2006**, *79*, 1193–1224. [CrossRef]

16.   Campbell, J.; Taskler, G. Equity volatility and corporate bond yields. *J. Financ.* **2003**, *58*, 2321–2350. [CrossRef]

17.   Fuertes, A.; Kalotchou, E. Optimal design of early warning systems for sovereign debt crises. *Int. J. Forecast.* **2007**, *23*, 85–100. [CrossRef]

18.   Cremers, M.; Driessen, J.; Maenhout, P.; Weinbaum, D. Individual stock-option prices and credit spreads. *J. Bank. Financ.* **2008**, *32*, 2706–2715. [CrossRef]

19.   Brutti, F. Sovereign defaults and liquidity crises. *J. Int. Econ.* **2011**, *84*, 65–72. [CrossRef]

20.   Mendoza, E.; Yue, V. A general equilibrium model of sovereign default and business cycles. *Q. J. Econ.* **2012**, *127*, 889–946. [CrossRef]

21.   Qian, Z.; Luo, Q. Regime-dependent determinants of China's sovereign credit default swap spread. *Emerg. Mark. Financ. Trade* **2016**, *52*, 10–21. [CrossRef]

22.   Duffie, D.; Singleton, K. Modeling the term structure of defaultable bonds. *Rev. Financ. Stud.* **1999**, *12*, 687–720. [CrossRef]

23.   Duffie, D.; Peterson, L.; Singleton, K. Modeling sovereign yield spreads: A case study of Russian debt. *J. Financ.* **2003**, *58*, 119–159. [CrossRef]

24.   Pan, J.; Singleton, K. Default and recovery implicit in the term structure of sovereign CDS spreads. *J. Financ.* **2008**, *63*, 2345–2384. [CrossRef]

25.   Edwards, S. LDC foreign borrowing and default risk: An Empirical investigation, 1976–1980. *Am. Econ. Rev.* **1984**, *74*, 726–734.

26.   Delatte, A.L.; Gex, M.; Lopez-Villavicencio, A. Has the CDS market influenced the borrowing cost of European countries during the sovereign crisis? *J. Int. Money Financ.* **2012**, *31*, 481–497. [CrossRef]

27.   Blommestein, H.; Eijffinger, S.; Qian, Z. Regime-dependent determinants of Euro area sovereign CDS spreads. *J. Financ. Stab.* **2016**, *22*, 10–21. [CrossRef]

28.   González-Rozada, M.; Yeyati, E.L. Global factors and emerging market spreads. *Econ. J.* **2008**, *118*, 1917–1936. [CrossRef]

29.   Godlewski, C.J. Bank capital and credit risk taking in emerging market economies. *J. Bank. Regul.* **2005**, *6*, 128–145. [CrossRef]

30.   Hamilton, J.D. *Regime Switching Models*; Palgrave Dictionary of Economics: San Diego, CA, USA, 2005.

31.   Perlin, M. MS Regress—The MATLAB Package for Markov Regime Switching Models. Available online: http://ssrn.com/abstract=1714016 (accessed on 26 December 2010).

32.   Tsai, S.-B. Using the DEMATEL Model to Explore the Job Satisfaction of Research and Development Professionals in China's Photovoltaic Cell Industry. *Renew. Sustain. Energy Rev.* **2018**, *81*, 62–68. [CrossRef]

33. Lee, Y.-C.; Hsiao, Y.-C.; Peng, C.-F.; Tsai, S.-B.; Wu, C.-H.; Chen, Q. Using Mahalanobis-Taguchi System, Logistic Regression and Neural Network Method to Evaluate Purchasing Audit Quality. *Proc. Inst. Mech. Eng. B J. Eng. Manuf.* **2015**, *229*, 3–12. [CrossRef]

34. Lee, Y.-C.; Chen, C.-Y.; Tsai, S.-B.; Wang, C.-T. Discussing green environmental performance and competitive strategies. *Pensee* **2014**, *76*, 190–198.

35. Liu, B.; Li, T.; Tsai, S.-B. Low carbon strategy analysis of competing supply chains with different power structures. *Sustainability* **2017**, *9*, 835. [CrossRef]

36. Qu, Q.; Tsai, S.-B.; Tang, M.; Xu, C.; Dong, W. Marine ecological environment management based on ecological compensation mechanisms. *Sustainability* **2016**, *8*, 1267. [CrossRef]

37. Lee, Y.-C.; Wang, Y.-C.; Chien, C.-H.; Wu, C.-H.; Lu, S.-C.; Tsai, S.-B.; Dong, W. Applying revised gap analysis model in measuring hotel service quality. *SpringerPlus* **2016**, *5*, 1191. [CrossRef] [PubMed]

38. Wang, J.; Yang, J.; Chen, Q.; Tsai, S.-B. Collaborative Production Structure of Knowledge Sharing Behavior in Internet Communities. *Mob. Inf. Syst.* **2016**, *2016*, 8269474. [CrossRef]

39. Tsai, S.-B.; Lee, Y.-C.; Guo, J.-J. Using modified grey forecasting models to forecast the growth trends of green materials. *Proc. Inst. Mech. Eng. B J. Eng. Manuf.* **2014**, *228*, 931–940. [CrossRef]

40. Tsai, S.-B.; Zhou, J.; Gao, Y.; Wang, J.; Li, G.; Zheng, Y.; Ren, P.; Xu, W. Combining FMEA with DEMATEL Models to Solve Production Process Problems. *PLoS ONE* **2017**, *12*, e0183634. [CrossRef] [PubMed]

41. Ge, B.; Jiang, D.; Gao, Y.; Tsai, S.-B. The influence of legitimacy on a proactive green orientation and green performance: A study based on transitional economy scenarios in china. *Sustainability* **2016**, *8*, 1344. [CrossRef]

42. Lee, S.-C.; Su, J.-M.; Tsai, S.-B.; Lu, T.-L.; Dong, W. A comprehensive survey of government auditors' self-efficacy and professional Development for improving audit quality. *SpringerPlus* **2016**, *5*, 1263. [CrossRef] [PubMed]

43. Chen, H.-M.; Wu, C.-H.; Tsai, S.-B.; Yu, J.; Wang, J.; Zheng, Y. Exploring key factors in online shopping with a hybrid model. *SpringerPlus* **2016**, *5*, 2046. [CrossRef] [PubMed]

44. Lee, Y.-C.; Wang, Y.-C.; Lu, S.-C.; Hsieh, Y.-F.; Chien, C.-H.; Tsai, S.-B.; Dong, W. An empirical research on customer satisfaction study: A consideration of different levels of performance. *SpringerPlus* **2016**, *5*, 1577. [CrossRef] [PubMed]

45. Wang, J.; Yang, J.; Chen, Q.; Tsai, S.-B. Creating the sustainable conditions for knowledge information sharing in virtual community. *SpringerPlus* **2016**, *5*, 1019. [CrossRef] [PubMed]

Article

Generalized Correlation Measures of Causality and Forecasts of the VIX Using Non-Linear Models

David E Allen [1,2,3,*] and Vince Hooper [4]

1   School of Mathematics and Statistics, University of Sydney, Camperdown, NSW 2006, Australia
2   Department of Finance, Asia University, Taichung 41354, Taiwan
3   School of Business and Law, Edith Cowan University, Edith Cowan University, Joondalup 6027, Australia
4   School of Economics and Management, Xiamen University, 43900 Sepang, Selangor Darul Ehsan, Malaysia; hoovcomm@hotmail.com
*   Correspondence: profallen2007l@gmail.com

Abstract: This paper features an analysis of causal relations between the daily VIX, S&P500 and the daily realised volatility (RV) of the S&P500 sampled at 5 min intervals, plus the application of an Artificial Neural Network (ANN) model to forecast the future daily value of the VIX. Causal relations are analysed using the recently developed concept of general correlation Zheng et al. and Vinod. The neural network analysis is performed using the Group Method of Data Handling (GMDH) approach. The results suggest that causality runs from lagged daily RV and lagged continuously compounded daily return on the S&P500 index to the VIX. Sample tests suggest that an ANN model can successfully predict the daily VIX using lagged daily RV and lagged daily S&P500 Index continuously compounded returns as inputs.

1. Introduction

This paper features an analysis of the causal relationships between the daily value of the VIX and the volatility of the S&P500, as revealed by estimates of the realised volatility (RV) of the S&P500 index, sampled at 5 min intervals, to produce daily values, as calculated by the Oxford Man Institute of Quantitative Finance, utilising Reuter's high frequency market data and provided in their 'Realised Library.' The causal analysis features an application of generalised measures of correlation, as developed by Zheng et al. [1] and Vinod [2]. This metric permits a more refined measure of causal direction.

The concept of causality has been a central philosophical issue for millennia. Aristotle in 'Physics II 3 and Metaphysics V 2' offered a general account of his concept of the four causes. (See http://classics.mit.edu/Aristotle/physics.2.ii.html). His account was general in the sense that it applied to everything that required an explanation, including artistic production and human action. He mentioned the: material cause: that out of which it is made, the efficient cause: the source of the objects principle of change or stability, the formal cause: the essence of the object. And the final cause: the end/goal of the object, or what the object is good for.

This treatment is far more encompassing than the customary treatment of causality in economics and finance. The modern treatment has been reduced to an analysis of correlation and statistical modelling. The origins of which can be traced back to the Scottish Enlightenment philosopher and historian, David Hume, who explored the relationship of cause and effect. Hume is recognised as a thorough going exponent of philosophical naturalism and as a precursor of contemporary cognitive

science. Hume showed us that experience does not tell us much. Of two events, A and B, we say that A causes B when the two always occur together, that is, are constantly conjoined. Whenever we find A, we also find B and we have a certainty that this conjunction will continue to happen. This leads on to the concept of induction and a weak notion of necessity. (See: https://people.rit.edu/wlrgsh/HumeTreatise.pdf). It provides a backdrop to contemporary treatments of causality and statistical measures of association. The intricacies and difficulties involved in the concept of causality are further explored by Pearl [3].

In terms of statistical measures of association, or 'constant contiguity,' to adopt Hume's term, Carl Pearson developed the correlation coefficient in the 1890s [4]. Granger [5], introduced the time series linear concept of 'Granger' causality. Zheng et al. [1] point out that one of the limitations of the correlation coefficient is that it does not account for asymmetry in explained variance. They developed broader applicable correlation measures and proposed a pair of generalized measures of correlation (GMC) which deal with asymmetries in explained variances and linear or nonlinear relations between random variables. Vinod [2] has further applied these measures to applied economics issues and developed an R library package, 'generalCorr,' for the application of these metrics, used in the analysis in this paper.

In this paper, we explore the directional causality between the VIX and RV estimates of the S&P500 volatility applying non-linear (GMC) methods and then engage in a further non-linear volatility forecasting exercise using Artificial Neural Network (ANN) methods. We do this using the GMDH shell program (http:www.gmdhshell.com). This program is built around an approximation called the Group Method of Data Handling. This approach is used in such fields as data mining, prediction, complex systems modelling, optimization and pattern recognition. The algorithms feature an inductive procedure that performs a sifting and ordering of gradually complicated polynomial models and the selection of the best solution by external criterion.

The paper is divided into five sections; Section 2, which follows this introduction, discusses the previous literature, whilst Section 3 introduces the data and research methods applied, Section 4 presents the results and section five concludes.

2. Prior Literature

In response to concerns that the original VIX calculation methodology had several weaknesses which made the issuance of VIX-related derivatives difficult, changes were made in 2003 by the CBOE. The calculation methodology was redefined to use the prices of synthetic 30-day options on the S&P500 index. See the discussions in Carr and Wu [6] and Whaley [7].

The VIX index is the "risk-neutral" expected stock market variance for the US S&P500 contract and is computed from a panel of options prices. It is termed the 'fear index' (see Whaley [8]) and provides an indication of both stock market uncertainty and a variance risk premium, which is also the expected premium from selling stock market variance in a swap contract. The VIX is based on "model-free" implied variances which are computed from a collection of option prices without the use of a specific pricing model (see, for example, Carr and Madan [9]).

There are various approaches to empirical work on the VIX. Baba and Sekura [10] investigate the role of US macroeconomic variables as leading indicators of regime shifts in the VIX index using a regime-switching approach. They suggest there are three distinct regimes in the VIX index during the 1990 to 2010 period corresponding to: a tranquil regime with low volatility, a turmoil regime with high volatility and a crisis regime with extremely high volatility. Fernandes et al. [11] undertake an analysis of the relationship between the VIX index and financial and macroeconomic factors.

There has been a great deal of work on derivatives related to the VIX. This is not the concern of this paper but the relevant ground is covered in Alexander et al. [12]. The fact that the VIX provides an estimate of the variance risk premium has been used to explore its relationship with stock market returns. See, for example, Bollerslev et al. [13] and Baekart and Horova [14], who take a similar approach.

The variance premium is defined by Bollerslev at al. [13], as the difference between the VIX, an ex-ante risk-neutral expectation of the future return variation over the $[t, \ t+1]$ time interval (IV_t) and the ex post realized return variation over the $[t-1, \ t]$ time interval obtained from RV_t measures:

$$VarianceRiskPremium_t = VRP_t \equiv ImpliedVolatility_t - RealisedVolatility_t \tag{1}$$

Bollerslev et al. [13] use the difference between implied and realized variation, or the variance risk premium, to explain a nontrivial fraction of the time-series variation in post-1990 aggregate stock market returns, with high (low) premia predicting high (low) future returns. The direction of the presumed causality is motivated from the implications from a stylized self-contained general equilibrium model incorporating the effects of time-varying economic uncertainty.

The current paper is concerned with the relationship between the VIX, implied volatility and S&P500 index continuously compounded returns but the focus is on an investigation of the causal path. It seeks to explore whether there is a stronger causal link between the VIX, to RV and stock returns, or in the reverse direction, from RV and stock returns to the VIX. The GMC analysis used in the paper suggests that the latter is the stronger causal path.

3. Data and Research Methods

3.1. Data Sample

We analyse the relationship between the VIX, the S&P500 Index and the realised volatility of the S&P500 index sampled at 5 min intervals, using daily data from 3 January 2000 to 12 December 2017, a total, after data cleaning and synchronization, of 4504 observations. The data for the VIX and S&P500 are obtained from Yahoo finance, whilst the realised volatility estimates are from the Oxford Man Realised Library (see: https://realized.oxford-man.ox.ac.uk).

In this paper, unlike the literature that uses the variance risk premium to forecast returns, we reverse the assumed direction of causality, based on our GMC analysis and predict the VIX on the basis of market returns and realised volatility.

The approach taken by Bollerslev et al. [13] and Baekart and Horova [14], is constructed on theoretical grounds and is not subjected to any tests of causal direction. A key feature of the current paper is to test, in practice, whether the causal direction runs from the VIX to returns on the S&P500 and estimates of daily RV, or, as we will subsequently demonstrate, in the reverse direction.

Given that we will be using regression analysis we require that our data sets are stationary. We know that price levels are non-stationary and so we use the continuously compounded returns on the S&P500 index. The results of Augmented Dickey Fuller tests shown in Table 1, strongly reject the null of non-stationarity for both the VIX and RV5MIN series, so we can combine them with the continuously compounded returns for the S&P500 Index in regression analysis, without the worry of estimating spurious regression.

Table 1. Tests of Stationarity: VIX and RV5MIN.

	ADF Test with Constant	Probability	ADF Test with Constant and Trend	Probability
VIX	−3.86664	0.002306 *	−4.11796	0.005859 *
RV5MIN	−7.70084	0.000 *	−7.80963	0.0000 *

Note: * Indicates significant at 0.01 level.

Plots of basic series are shown in Figure 1. Figure 2 shows quantile plots of our base series. All series show strong departures from a normal distribution in both tails of their distributions. These departures from Gaussian distributions are confirmed by the summary descriptions of the series provided in Table 2. The summary statistics for our data sets in Table 2 confirm the results of the QQPlots and show that we have excess kurtosis in all three series and pronounced skewness in RV5MIN. We also undertook some

preliminary regression and quantile regression analysis of the relationships between our three-base series to explore whether or not the relationship between the three series is linear.

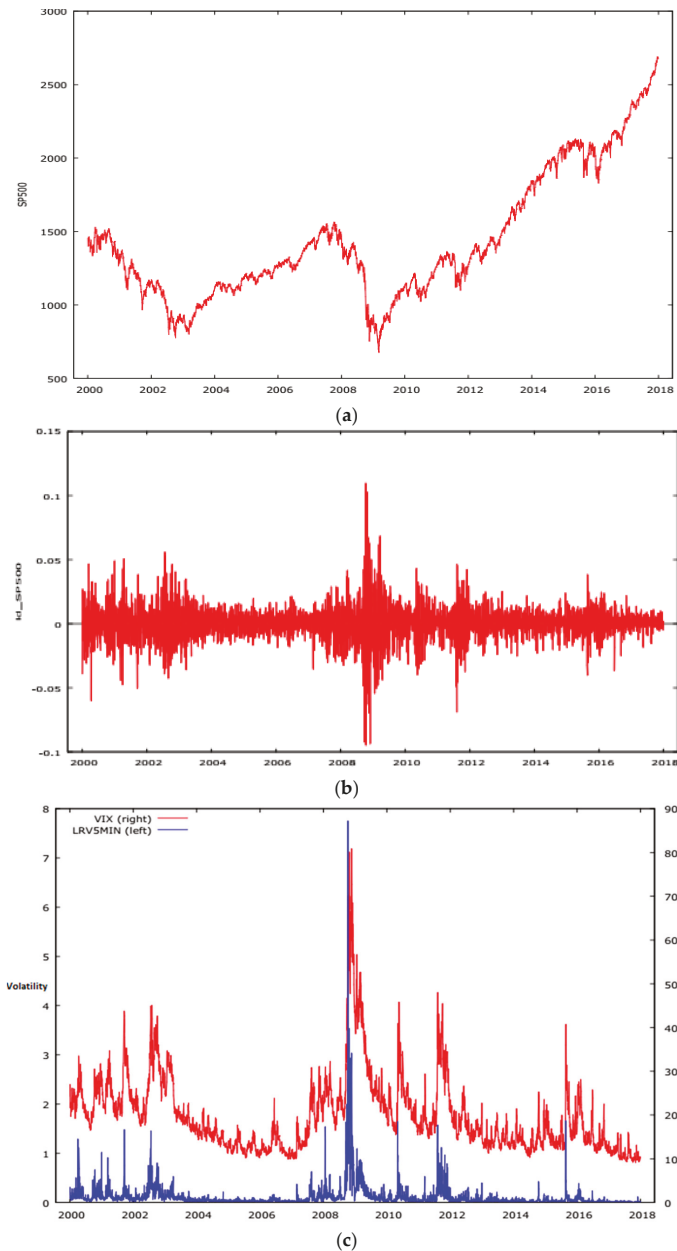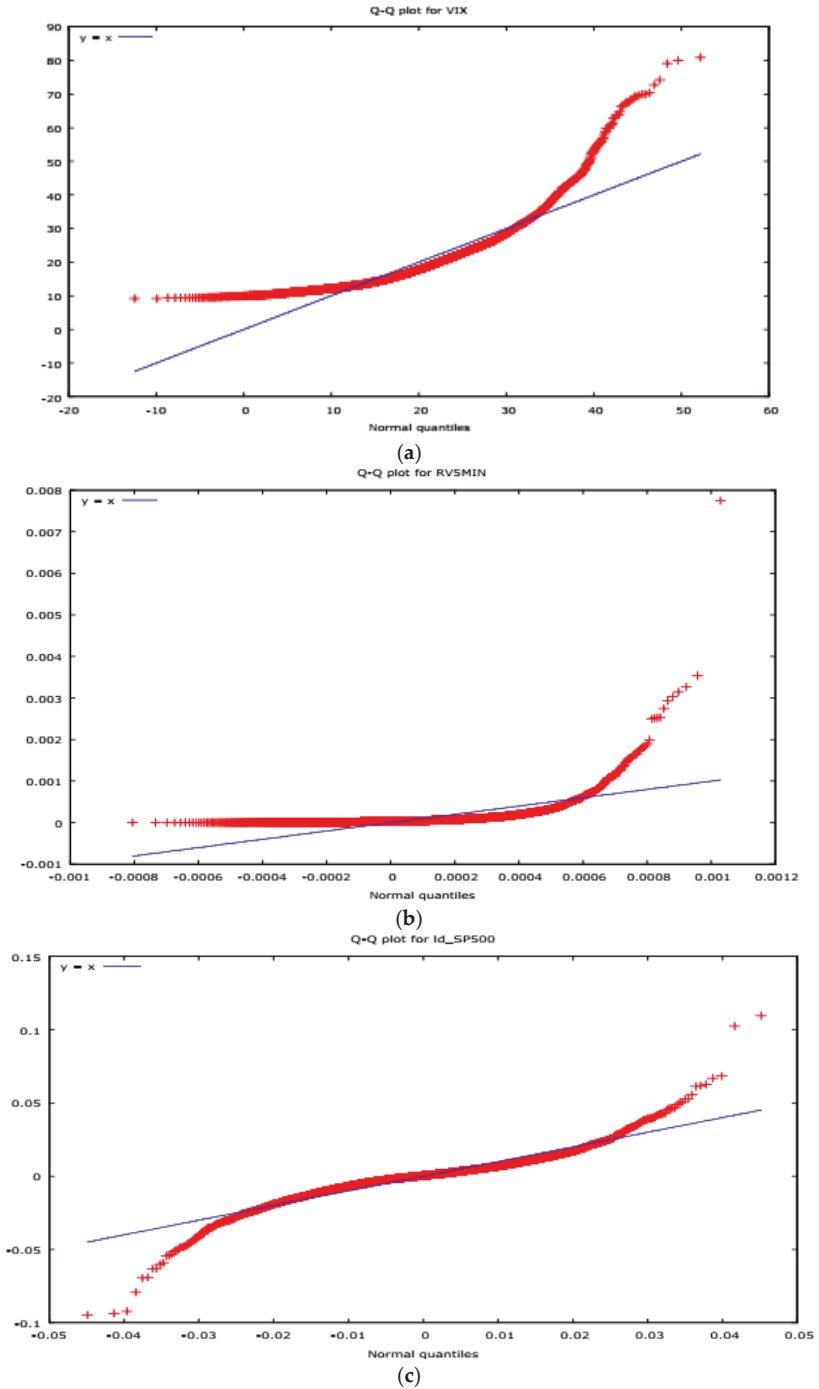**Figure 1.** Plots of Base Series. (**a**) S&P500 INDEX; (**b**) S&P500 INDEX CONTINUOUSLY COMPOUNDED RETURNS; (**c**) VIX and RV5MIN.

Figure 2. QQPlots of Base Series. (**a**) QQPLOT VIX; (**b**) QQPlot RV5MIN; (**c**) QQPLOT S&P500 RETURNS.

Table 2. Data Series: Summary Statistics 3 January 2000 to 29 December 2017.

	VIX	S&P500 Return	RV5MIN
Mean	19.8483	0.000135262	0.111837
Median	17.6700	0.000522156	0.0501000
Minimum	9.14000	−0.0946951	0.000878341
Maximum	80.8600	0.109572	7.74774
Standard Deviation	8.75231	0.0121920	0.248439
Coefficient of Variation	0.440961	90.1361	2.22143
Skewness	2.09648	−0.203423	11.4530
Excess Kurtosis	6.94902	8.65908	242.166

3.2. Preliminary Regression Analysis

We estimated an OLS regression of the VIX regressed on the continuously compounded S&P500 return 'SPRET. The results are shown in Table 3. The slope coefficient is insignificant and the R squared is a miniscule 0.000158. The Ramsey Reset test suggests that the relationship is non-linear and that the regression is miss-specified.

Table 3. OLS Regression of VIX on SPRET.

	Coefficient	*t*-Ratio	Probability Value
Constant	19.8485	43.35	0.00 ***
SPRET	−9.01551	−0.5215	0.6021
Adjusted R-squared			
F(1, 4495)	0.271949		*p*-value (F) 0.602053
Ramsey Reset Test			
Constant	−147551	−1.924	0.0544 *
SPRET	109932	2.105	0.0354 **
yhat^2	509.402	1.745	0.0811 *
yhat^3	−6.79270	−1.385	0.1662

Note: ***, **, * denotes significance at 1%, 5% and 10%.

A QQplot of the residuals from this regression shown in Figure 3 also suggests that a linear specification is inappropriate.

To further explore the relationship between the sample variables we employed quantile regression analysis. Quantile Regression is modelled as an extension of classical OLS (Koenker and Bassett, [15]), in quantile regression the estimation of conditional mean as estimated by OLS is extended to similar estimation of an ensemble of models of various conditional quantile functions for a data distribution. In this fashion quantile regression can better quantify the conditional distribution of $(Y|X)$. The central special case is the median regression estimator that minimizes a sum of absolute errors. We get the estimates of remaining conditional quantile functions by minimizing an asymmetrically weighted sum of absolute errors, here weights are the function of the quantile of interest. This makes quantile regression a robust technique even in presence of outliers. Taken together the ensemble of estimated conditional quantile functions of $(Y|X)$ offers a much more complete view of the effect of covariates on the location, scale and shape of the distribution of the response variable.

For parameter estimation in quantile regression, quantiles as proposed by Koenker and Bassett [15] can be defined through an optimization problem. To solve an OLS regression problem a sample mean is defined as the solution of the problem of minimising the sum of squared residuals, in the same way the median quantile (0.5%) in quantile regression is defined through the problem of minimising the sum of absolute residuals. The symmetrical piecewise linear absolute value function assures the same number of observations above and below the median of the distribution. The other quantile values can

be obtained by minimizing a sum of asymmetrically weighted absolute residuals, (giving different weights to positive and negative residuals). Solving:

$$min \xi \varepsilon R \sum \rho_\tau (y_i - \xi) \tag{2}$$

where $\rho_\tau(\cdot)$ is the tilted absolute value function as shown in Figure 4, which gives the τth sample quantile with its solution. Taking the directional derivatives of the objective function with respect to ξ (from left to right) shows that this problem yields the sample quantile as its solution.
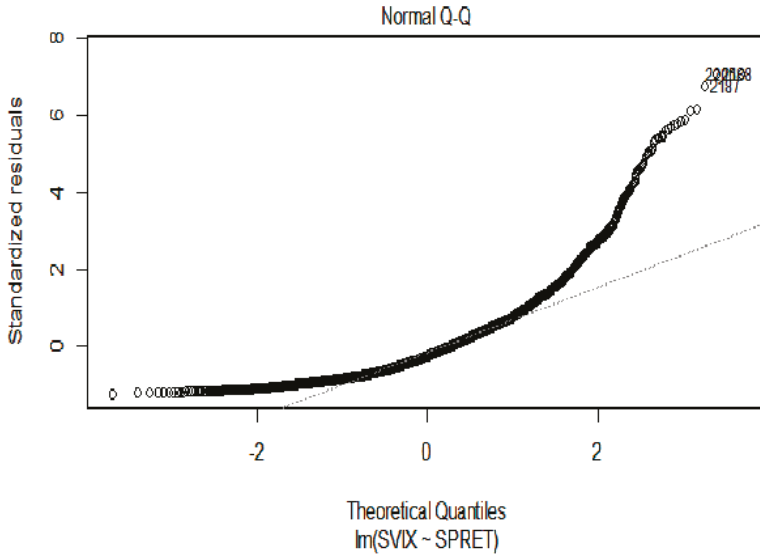
Figure 3. QQplot of residuals from OLS regression of VIX on SPRET.
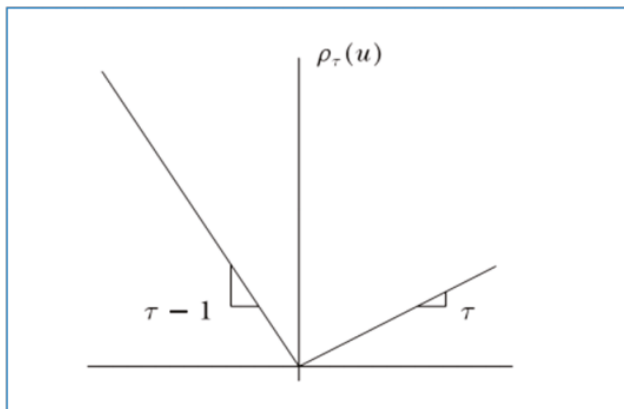
Figure 4. Quantile regression ρ function.

After defining the unconditional quantiles as an optimization problem, it is easy to define conditional quantiles similarly. Taking the least squares regression model as a base to proceed, for a random sample, $y_1, y_2, \ldots, y_n$, we solve:

$$min_{\mu \varepsilon R} \sum_{i=1}^{n} (y_i - \mu)^2, \tag{3}$$

Which gives the sample mean, an estimate of the unconditional population mean, EY. Replacing the scalar μ by a parametric function $\mu(x, \beta)$ and then solving:

$$min_{\mu \varepsilon R^p} \sum_{i=1}^{n} (y_i - \mu(x_i, \beta))^2 \tag{4}$$

gives an estimate of the conditional expectation function $E(Y \mid x)$.

Proceeding the same way for quantile regression, to obtain an estimate of the conditional median function, the scalar ξ in the first equation is replaced by the parametric function $\xi(x_t, \beta)$ and τ is set to 1/2. The estimates of the other conditional quantile functions are obtained by replacing absolute values by $\rho_\tau(\cdot)$ and solving:

$$min_{\mu \varepsilon R^p} \sum \rho_\tau (y_i - \xi(x_i, \beta)) \tag{5}$$

The resulting minimization problem, when $\xi(x, \beta)$ is formulated as a linear function of parameters and can be solved very efficiently by linear programming methods. Further insight into this robust regression technique can be obtained from Koenker and Bassett [15] and Koenker [16].

We used quantile regression to regress VIX on SPRET with the quantiles (tau), set at 0.05, 0.35, 0.5, 0.75 and 0.95 respectively. The results are shown in Table 4 and Figure 5.

Table 4. Quantile regression of VIX on SPRET (tau = 0.05, 0.25, 0.5, 0.75 and 0.95).

	Coefficient SPRET	*t* **Value**	**Probability**
tau = 0.05	−4.41832	−0.76987	0.44142
tau = 0.25	−2.79810	−0.43081	0.66663
tau = 0.50	−28.94626	−3.00561	0.00267 ***
tau = 0.75	−25.97296	−1.68811	0.09146 *
tau = 0.95	−29.40331	−0.57619	0.56452

Note: *** Significant at 1%, * Significant at 10%.

Figure 5. Quantile regression of VIX on SPRET, estimates and error bands.

These preliminary regression results suggest a non-linear relationship between the VIX and SPRET. The existence of this non-linear relationship is consistent with findings by Busson and Vakil [17]. The importance of non-linearity will be explored further when we apply the metric provided by the Generalised Measure of Correlation which we introduce in the next subsection.

3.3. Econometric Methods

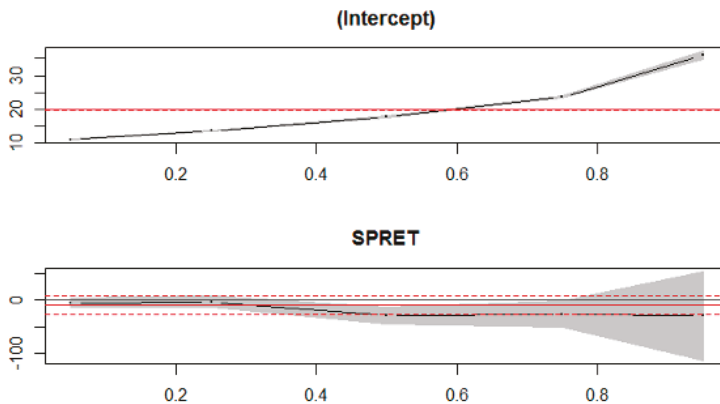Zeng et al. [1] point out that despite its ubiquity there are inherent limitations in the Pearson correlation coefficient when it is used as a measure of dependency. One limitation is that it does not account for asymmetry in explained variances which are often innate among nonlinearly dependent random variables. As a result, measures dealing with asymmetries are needed. To meet this requirement, they developed Generalized Measures of Correlation (GMC). They commence with the familiar linear regression model and the partitioning of the variance into explained and unexplained portions

$$Var(X) = Var(E(X \mid Y) + E(Var(X \mid Y)),$$ (6)

Whenever $E(Y^2) < \infty$ and $E(X^2) < \infty$. Note that $E(Var(X \mid Y))$ is the expected conditional variance of X given Y and therefore can be interpreted as the explained variance of X by Y. Thus, we can write:

$$\frac{E(Var(X \mid Y))}{Var(X)} = 1 - \frac{E(Var(X \mid Y))}{Var(X)} = 1 - \frac{E(\{X - E(X \mid Y)\}^2}{Var(X)}.$$

The explained variance of Y given X can similarly be defined. This leads Zheng et al. [1] to define a pair of generalised measures of correlation (GMC) as:

$$\{GMC(Y \mid X), GMC(X \mid Y)\} = \{1 - \frac{E(\{Y - E(Y \mid X)\}^2}{Var(Y)}, 1 - \frac{E(\{X - E(X \mid Y)\}^2}{Var(X)}\}.$$ (7)

This pair of GMC measures has some attractive properties. It should be noted that the two measures are identical when (X, Y) is a bivariate normal random vector.

Vinod [2] takes this measure in Expression (2) and reminds the reader that it can be viewed as kernel causality. The Naradaya Watson kernel regression is a non-parametric technique used in statistics to estimate the conditional expectation of a random variable. The objective is to find a non-linear relation between a pair of random variables X and Y. In any nonparametric regression, the conditional expectation of a variable Y relative to a variable X could be written $E(Y|X) = m(X)$ where m is an unknown function.

Naradaya [18] and Watson [19] proposed estimating m as a locally weighted average employing a kernel as a regression function.

$$\hat{m}_h(x) = \frac{\sum_{i=1}^{n} K_{h(x-x_i)y_i}}{\sum_{j=1}^{n} K_{h(x-x_j)}},$$

where K is a kernel with bandwidth h. The denominator is a weighting term that sums to 1.

$GMC(Y \mid X)$ is the coefficient of determination R^2 of the Nadaraya-Watson nonparametric Kernel regression:

$$y = g(X) + \epsilon = E(Y \mid X) + \epsilon,$$ (8)

where $g(X)$ is a nonparametric, unspecified (nonlinear) function. Interchanging X and Y, we obtain the other $GMC(X \mid Y)$ defined as the R^2 of the Kernel regression:

$$X = g'(Y) + \epsilon' = E(XY) + \epsilon'.$$ (9)

Vinod [2] defines $\delta = GMC(X \mid Y) - GMC(X \mid Y)$ as the difference of two population R^2 values. When $\delta < 0$, we know that X better predicts Y than vice versa. Hence, we define that X kernel causes Y provided the true unknown $\delta < 0$. Its estimate δ' can be readily computed by means of regression.

Zheng et al. [1] demonstrate that GMC can lead to a more refined version of the concept of Granger-causality. They assume an order one bivariate linear autoregressive model. Y_t Granger-causes X_t if:

$$E[\{X_t - E(X_t \mid X_{t-1})\}^2 > E[\{X_t - E(X_t \mid X_{t-1}, Y_{t-1})\}^2, \tag{10}$$

Which suggests that X_t can be better predicted using the histories of both X_t and Y_t than using the history of X_t alone. Similarly, we would say X_t Granger-causes Y_t if:

$$E[\{Y_t - E(Y_t \mid Y_{t-1})\}^2 > E[\{Y_t - E(Y_t \mid Yt-1, X_{t-1})\}^2. \tag{11}$$

They use the fact $E(Var(X_t \mid X_{t-1}) = E(\{X_t - E(X_t \mid X_{t-1}\}^2)$ and $E[\{E(X_t \mid X_{t-1}) - E(X_t \mid X_{t-1}, Y_{t-1})\}^2] = E[\{X_t - E(X_t \mid X_{t-1})\}^2 - E[\{X_t - E(X_t \mid X_{t-1}, Y_{t-1})\}^2]$. Which suggests that (5) is equivalent to:

$$1 - \frac{E[\{X_t - E(X_t \mid X_{t-1}, Y_{t-1})\}^2}{E(Var(X_t \mid X_{t-1}))} > 0. \tag{12}$$

In the same way (6) is equivalent to:

$$1 - \frac{E[\{Y_t - E(Y_t \mid Y_{t-1}, X_{t-1})\}^2}{E(Var(Y_t \mid Y_{t-1}))} > 0. \tag{13}$$

They add that when both (5) and (6) are true, there is a feedback system.

Suppose that $\{X_t, Y_t\}, Y_t > 0$ is a bivariate stationary time series. Zheng et al. [1] define Granger causality generalised measures of correlation as:

$$GcGMC = (X_t \mid \mathcal{F}_{t-1}) = 1 - \frac{E[\{X_t - \mid X_{t-1}, X_{t-1}, \ldots, Y_{t-1}, Y_{t-2}, \ldots,)\}^2]}{E(Var(X_t \mid X_{t-1}, X_{t-2}, \ldots))}, \tag{14}$$

$$GcGMC = (Y_t \mid \mathcal{F}_{t-1}) = 1 - \frac{E[\{Y_t - \mid Y_{t-1}, Y_{t-1}, \ldots, X_{t-1}, X_{t-2}, \ldots,)\}^2]}{E(Var(Y_t \mid Y_{t-1}, Y_{t-2}, \ldots))} \tag{15}$$

where $\mathcal{F}_{t-1} = \sigma(X_{t-1}, X_{t-2}, \ldots, Y_{t-1}, Y_{t-2}, \ldots)$.

Zheng et al. [1] suggest that if:

- $GcGMC = (X_t \mid \mathcal{F}_{t-1}) > 0$, they say Y Granger causes X.
- $GcGMC = (Y_t \mid \mathcal{F}_{t-1}) > 0$, they say X Granger causes Y.
- $GcGMC = (X_t \mid \mathcal{F}_{t-1}) > 0$ and $GcGMC = (Y_t \mid \mathcal{F}_{t-1}) > 0$, they say they have a feedback system.
- $GcGMC = (X_t \mid \mathcal{F}_{t-1}) > GcGMC = (Y_t \mid \mathcal{F}_{t-1})$, they say X is more influential than Y
- $GcGMC = (Y_t \mid \mathcal{F}_{t-1}) > GcGMC = (X_t \mid \mathcal{F}_{t-1})$, they say Y is more influential than X.

We explore the relationship between the VIX, the lagged continuously compounded return on the S&P500 Index, (LSPRET) and the lagged daily realised volatility on the S&P500, sampled at 5 min intervals within the day (LRV5MIN). Once we have established causal directions between these variables, we use them to construct our ANN model. The ANN model is discussed in the next section.

3.4. Artificial Neural Net Models

There are a variety of approaches to neural net modelling. A simple neural network model with linear input, D hidden units and activation function g, can be written as:

$$x_{t+s} = \beta_0 + \sum_{j=1}^{D} \beta_j g\left(\gamma_{0j} + \sum_{i=1}^{m} \gamma_{ij} x_{t-(i-1)d}\right). \tag{16}$$

However, we choose to apply a nonlinear neural net modelling approach, using the GMDH shell program (GMDH LLC 55 Broadway, 28th Floor New York, NY 10006) (http:www.gmdhshell.com). This program is built around an approximation called the 'Group Method of Data Handling.' This approach is used in such fields as data mining, prediction, complex systems modelling, optimization and pattern recognition. The algorithms feature an inductive procedure that performs a sifting and ordering of gradually complicated polynomial models and the selection of the best solution by external criterion.

A GMDH model with multiple inputs and one output is a subset of components of the base function:

$$Y(xi_1, \ldots, x_n) = a_0 + \sum_{i=1}^{m} a_i f_i, \tag{17}$$

where f are elementary functions dependent on different inputs, a are unknown coefficients and m is the number of base function components.

In general, the connection between input-output variables can be approximated by the Volterra functional series, the discrete analogue of which is the Kolmogorov-Gabor polynomial:

$$y = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} a_{ijk} x_i x_j x_k + \ldots, \tag{18}$$

where, $x = (x_i, x_2, \ldots, x_m)$, the input variables vector and $A = (a_0, a_1, a_2, \ldots, a_m)$ the vector of weights. The Kolmogorov-Gabor polynomial can approximate any stationary random sequence of observations and can be computed by either adaptive methods or a system of Gaussian normal equations. Ivakhnenko [20] developed the algorithm, 'The Group Method of Data Handling (GMDH)' by using a heuristic and perceptron type of approach. He demonstrated that a second-order polynomial (Ivakhnenko polynomial: $y = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2$) can reconstruct the entire Kolmogorov-Gabor polynomial using an iterative perceptron-type procedure.

4. Results

4.1. GMC Analysis

Vinod's (2017) R library package 'generalCorr' is used to assess the direction of the causal paths between the VIX and lagged values of the S&P500 continuously compounded return LSPRET and the lagged daily estimated realised volatility for the S&P500 index, LRV5MIN. The results of the analysis are shown in Table 5.

We use the R 'generalCorr' package to undertake the analysis shown in Table 5. The output matrix is seen to report the cause' along columns and 'response' along the rows. The value of 0.7821467 in the R.H.S. of the second row of Table 5 is larger than the value 0.608359 in the second column, third row of Table 5. These are our two generalised measures of correlation, when we first condition the VIX on LRV5MIN, in the second row of Table 5 and LRV5MIN on the VIX in the third row of Table 5. This suggests that causality runs from LRV5MIN, the lagged daily value of the realised volatility of the S&P500 index, sample at 5 min intervals.

We also test the significance of the difference between these two generalised measures of correlation. Vinod suggests a heuristic test of the difference between two dependent correlation values. Vinod [2] suggests a test based on a suggestion by Fisher [21], of a variance stabilizing and normalizing transformation for the correlation coefficient, r, defined by the formula: $r = tanh(z)$, involving a hyperbolic tangent:

$$z = tan^{-1} r = \frac{1}{2} log \frac{1+r}{1-r}. \tag{19}$$

The application of the above test suggests a highly significant difference between the values of the two correlation statistics in Table 5.

Table 5. GMC analysis of the relationship between the VIX and LRV5MIN.

	VIX	LRV5MIN
VIX	1.000	0.7821467
LRV5MIN	0.608359	1.000
Test of the difference between the two paired correlations		
$t = 21.26$	probability = 0.0	

We also analyse the relationship between the VIX and the lagged daily continuously compounded return on the S&P500 index, LSPRET. The results are shown in Table 6 and suggest that lagged value of the daily continuously compounded return on the S&P500 index, LSPRET, drives the VIX. This is because the generalised correlation measure of the VIX conditioned on LSPRET is 0.5519368, whilst the generalised correlation measure of LSPRET conditioned on the VIX is only 0.153411. Once again, these two measures are significantly different.

Regression analysis suggested that the relationship was non-linear. We proceed to an ANN model which will be used for forecasting the VIX. Given that the GMC analysis suggests a stronger direction of correlation running from LRV5MIN and LSPRET to the VIX, rather than vice-versa, we use these two lagged daily variables as the predictor variables in our ANN modelling and forecasting.

Table 6. GMC analysis of the relationship between the VIX and LSPRET.

	VIX	LSPRET
VIX	1.000	0.5519368
LSPRET	0.153411	1.000
Test of the difference between the two paired correlations		
$t = 24.07$	probability = 0.0	

4.2. ANN Model

Our neural network analysis is run on 80 per cent of the observations in our sample and then its out-of-sample forecasting performance is analysed on the remaining 20 per cent, of the total sample of 4504 observations. The idea of the GMDH-type algorithms used in the GMDH Shell program is to apply a generator using gradually more complicated models and select the set of models that show the highest forecasting accuracy when applied to a previously unseen data set, which in this case is the 20 per cent of the sample remaining, which is used as a validation set. The top-ranked model is claimed to be the optimally most-complex one.

GMDH-type neural networks which are also known as polynomial neural networks employ a combinatorial algorithm for the optimization of neuron connection. The algorithm iteratively creates layers of neurons with two or more inputs. The algorithm saves only a limited set of optimally-complex neurons that are denoted as the initial layer width. Every new layer is created using two or more neurons taken from any of the previous layers. Every neuron in the network applies a transfer function (usually with two variables) that allows an exhaustive combinatorial search to choose a transfer function that predicts outcomes on the testing data set most accurately. The transfer function usually has a quadratic or linear form but other forms can be specified. GMDH-type networks generate many layers but layer connections can be so sparse that their number may be as small as a few connections per layer.

Since every new layer can connect to previous layers the layer width grows constantly. If we take into account that only rarely the upper layers improve the population of models, we proceed by dividing the additional size of the next layer by two and generate only half of the neurons generated by the previous layer, that is, the number of neurons N at layer k is $N_K = 0.5 \times N_{k-1}$. This heuristic makes the algorithm quicker whilst the chance of reducing the model's quality is low. The generation

of new layers ceases when either a new layer does not show improved testing accuracy than previous layer, or in circumstances in which the error was reduced by less than 1%.

In the case of the model reported in this paper, we used a maximum of 33 layers and the initial layer width was a 1000, whilst the neuron function was given by $a + x_i + x_i x_j + x_i^2$. The ANN regression analysis produces a complex non-linear model which is shown in Table 7.

Table 7. ANN regression model—dependent variable the VIX.

Y1 = −22.5101 + N107(1.01249) − **N107**$^{0.00364084^2}$ + N87(1.67752) − **N87**$^{0.211077^2}$
N87 = −8.10876 + $LSPRET^{1919.7^2}$ + N99(1.66543) − N99$^{0.0120732^2}$
N99 = −18.9937 − LRV5MIN(669.032) + LRV5MIN(N100)(1297.44) − $LRV5MIN^{1.09098e+07^2}$ + N100(2.8838) − N100$^{0.0509041^2}$
N100 = 18.6936 + LRV5MIN(48378) − N107$^{0.00976245^2}$
N107 = 17.0884 + LRV5MIN(20457.2) − LSPRET(50.0534) + $LSPRET^{3277.01^2}$

A plot of the ANN model fit is shown in Figure 6. The model appears to be a good fit, within the estimation period and in the 20 per cent of the sample used as a hold-out forecast period. This is confirmed by the diagnostics for the ANN model, reported in Table 8. The mean absolute error is smaller in the forecasts with a value of 3.14658, than it is when the model is being fitted, with a value of 3.16466. Similarly, the R^2 is higher in the forecast hold out sample, with a value of 75 percent, than in the model fitting stage, in which it has a value of almost 74 percent.
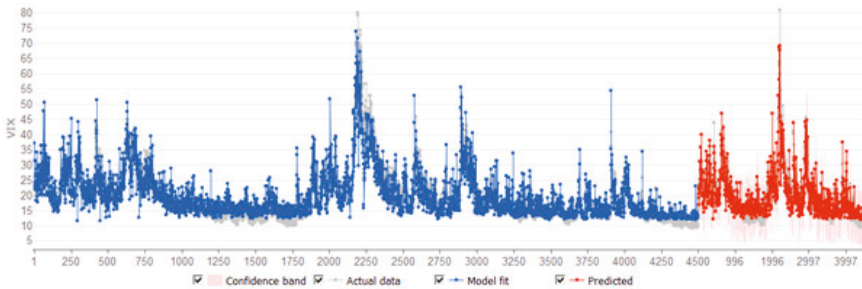
Figure 6. ANN regression model fit.

Table 8. ANN regression model diagnostics.

	Model Fit	Predictions
Mean Absolute Error	3.16466	3.14658
Root Mean Square Error	4.47083	4.36716
Standard Deviation of Residuals	4.47083	4.36697
Coefficient of Determination R^2	0.738519	0.752232

The diagnostic plots of the behaviour of the residuals, shown in Figure 7, also appears to show acceptable behaviour. Most of the residuals plot within the error bands, the residual histogram is approximately normal, though there is some evidence of persistence in the autocorrelations suggestive of ARCH effects.

As a further check on the mechanics of the model, we explored the effect on the root mean square errors in the forecasts if we replaced the two explanatory variable's observations with their means successively. LRV5MIN has the largest effect with an impact on RMSE of 10.5364% whilst LSPRET had an impact of 4.57003%. This is consistent with the previous GMC results which suggested that LRV5MIN had a relatively higher GMC with the VIX.
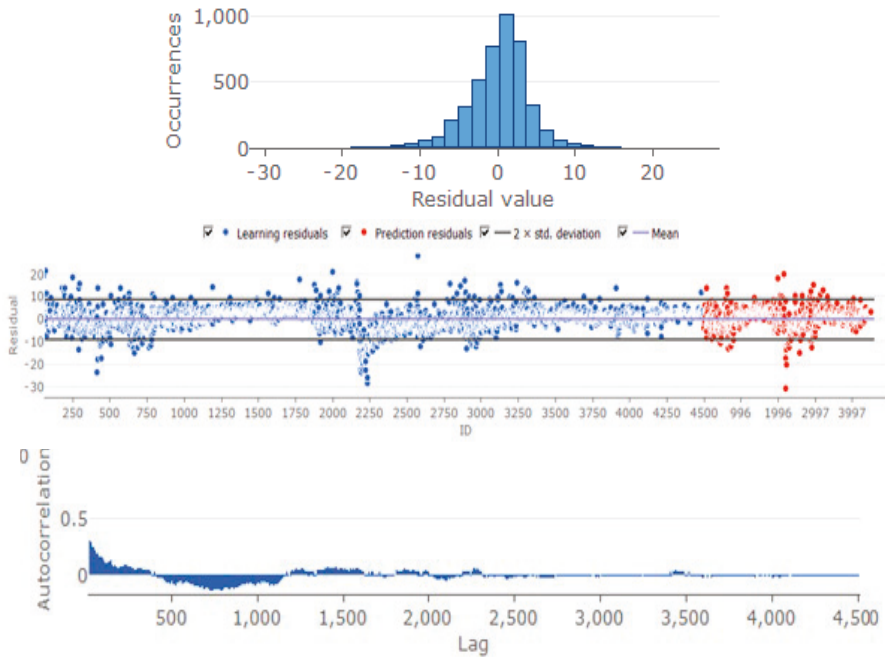
Figure 7. Residual diagnostic plots.

5. Conclusions

The paper featured an analysis of causal relations between the VIX and lagged continuously compounded returns on the S&P500, plus lagged realised volatility (RV) of the S&P500 sampled at 5 min intervals. Causal relations were analysed using the recently developed concept of general correlation Zheng et al. [1] and Vinod [2]. The results strongly suggested that causal paths ran from lagged returns on the S&P500 and lagged RV on the S&P500 to the VIX. The GMC analysis suggested that correlations running in this direction were stronger than those in the reverse direction. Statistical tests suggested that the pairs of correlated correlations analysed were significantly different.

An ANN model was then developed, based on the causal paths suggested, using the Group Method of Data Handling (GMDH) approach. The complex non-linear model developed performed well in both in and out of sample tests. The results suggest an ANN model can be used successfully to predict the daily VIX using lagged daily RV and lagged daily S&P500 Index continuously compounded returns as inputs.

References

1. Zheng, S.; Shi, N.-Z.; Zhang, Z. Generalized measures of correlation for asymmetry, nonlinearity, and beyond. *J. Am. Stat. Assoc.* **2012**, *107*, 1239–1252. [CrossRef]
2. Vinod, H.D. Generalized correlation and kernel causality with applications in development economics. *Commun. Stat. Simul. Comput.* **2017**, *46*, 4513–4534. [CrossRef]
3. Pearl, J. The foundations of causal inference. *Sociol. Methodol.* **2010**, *40*, 75149. [CrossRef]
4. Pearson, K. Notes on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* **1895**, *58*, 240–242. [CrossRef]
5. Granger, C. Investigating causal relations by econometric methods and cross-spectral methods. *Econometrica* **1969**, *34*, 424–438. [CrossRef]
6. Carr, P.; Wu, L. A tale of two indices. *J. Deriv.* **2006**, *13*, 13–29. [CrossRef]
7. Whaley, R. Understanding the VIX. *J. Portf. Manag.* **2006**, *35*, 98–105. [CrossRef]
8. Whaley, R.E. The investor fear gauge. *J. Portf. Manag.* **2000**, *26*, 12–17. [CrossRef]
9. Carr, P.; Madan, D. Towards a theory of volatility trading. In *Volatility: New Estimation Techniques for Pricing Derivatives*; Jarrow, R., Ed.; Risk Books: London, UK, 1998; Chapter 29; pp. 417–427.
10. Baba, N.; Sakurai, Y. Predicting regime switches in the VIX index with macroeconomic variables. *Appl. Econ. Lett.* **2011**, *18*, 1415–1419. [CrossRef]
11. Fernandes, M.; Medeiros, M.C.; Scharth, M. Modeling and predicting the CBOE market volatility index. *J. Bank. Financ.* **2014**, *40*, 1–10. [CrossRef]
12. Alexander, C.; Kapraun, J.; Korovilas, D. Trading and investing in volatility products. *J. Int. Money Financ.* **2015**, *24*, 313–347. [CrossRef]
13. Bollerslev, T.; Tauchen, G.; Zhou, H. Expected stock returns and variance risk premia. *Rev. Financ. Stud.* **2009**, *22*, 44634492. [CrossRef]
14. Bekaert, G.; Hoerova, M. The VIX, the variance premium and stock market volatility. *J. Econ.* **2014**, *183*, 181–192. [CrossRef]
15. Koenker, R.W.; Bassett, G. Regression quantiles. *Econometrica* **1978**, *46*, 33–50. [CrossRef]
16. Koenker, R. *Quantile Regression*; Cambridge University Press: Cambridge, UK, 2005.
17. Buson, M.G.; Vakil, A.F. On the non-linear relationship between the VIX and realized SP500 volatility. *Invest. Manag. Financ. Innov.* **2017**, *14*, 200–206.
18. Nadaraya, E.A. On estimating regression. *Theory Probab. Appl.* **1964**, *9*, 141–142. [CrossRef]
19. Watson, G.S. Smooth regression analysis. *Sankhyā Indian J. Stat. Ser. A* **1964**, *26*, 359–372.
20. Ivakhnenko, A.G. The group method of data handling—A rival of the method of stochastic approximation. *Sov. Autom. Control* **1968**, *1*, 43–55.
21. Fisher, R.A. On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond. A* **1922**, *222*, 309–368. [CrossRef]

Article

President Trump Tweets Supreme Leader Kim Jong-Un on Nuclear Weapons: A Comparison with Climate Change †

David E. Allen [1,2,3] **and Michael McAleer** [4,5,6,7]*

[1] School of Mathematics and Statistics, University of Sydney, Sydney, NSW 2006, Australia;
    profallen2007@gmail.com
[2] Department of Finance, Asia University, Taichung 413, Taiwan
[3] School of Business and Law, Edith Cowan University, Perth, WA 6027, Australia
[4] Discipline of Business Analytics, University of Sydney Business School, Sydney, NSW 2006, Australia
[5] Institute of Advanced Sciences, Yokohama National University, Yokohama 240-8501, Japan
[6] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam,
    3080 Rotterdam, The Netherlands
[7] Department of Economic Analysis and ICAE, Complutense University of Madrid, 28080 Madrid, Spain
*   Correspondence: michael.mcaleer@gmail.com
†   The analysis in the paper was undertaken with the R sentiment package.

Abstract: A set of 125 tweets about North Korea's Supreme Leader Kim Jong-Un by President Trump from 2013 to 2018 are analysed by means of the data mining technique, sentiment analysis. The intention is to explore the contents and sentiments of the messages contained, the degree to which they differ, and their implications about President Trump's understanding and approach to international diplomacy. The results suggest a predominantly positive emotion in relation to tweets about North Korea, despite the use of questionable nicknames such as "Little Rocket Man". A comparison is made between the tweets on North Korea and climate change, madefrom 2011–2015, as Trump has tweeted many times on both issues. It is interesting to find that Trump's tweets on North Korea have significantly higher positive polarity scores than his tweets on climate change.

Keywords: sentiment analysis; polarity; scientific verification; emotion; joy; sadness; climate change

JEL Classification: A1; C88; C44; Z0

---

1. Introduction

"Mentally Deranged U.S. Dotard" Tweets "Little Rocket Man"

A series of 125 tweets by President Trump on the topic of North Korea's Supreme Leader Kim Jong-Un are analysed by means of textual analysis using data mining techniques. The tweets date from 3 April 2018 to 10 March 2018. The analysis features the use of an R library package which facilitates sentiment analysis, 'sentiment'. The tweets were taken from an on-line sample available at https://twitter.com/search?q=donald%20Trump%20North%20Korea&src=typd&lang.

Data mining refers to the process of analysing data sets to reveal patterns, and usually involves methods that are drawn from statistics, machine learning, and database systems. There are two broad approaches to text mining and document analysis for extracting sentiment: the lexicon based approach and the text classification approach. The former involves using the semantic orientation of words or phrases in the document to calculate the orientation of the document. The latter approach could be described as a statistical or machine learning approach.

Text data mining involves the analysis of patterns in text data. Sentiment analysis is concerned with the emotional context of a text, and seeks to infer whether a section of text is positive or negative, or the nature of the emotions involved. There are a variety of methods and dictionaries that exist for undertaking the sentiment analysis of a piece of text.

Although sentiment is often framed in terms of being a binary distinction (positive versus negative), it can also be analysed in a more nuanced manner. We decided to apply the R package 'sentiment', which distinguishes between five different emotions, namely joy, sadness, anger, fear and surprise.

There are many different forms of sentiment analysis, but many use the same basic approach. They begin by constructing a list of words or dictionary associated with different emotions, count the number of positive and negative words in a given text, and then analyse the mix of positive and negative words to assess the general emotional tenor of the text. In our analysis we have used the inbuilt lexicon in the 'sentiment' package. This means we can compare our results with previous analyses we have undertaken using the same method. If we had so wished, we could have built our own lexicon, and this would have improved the accuracy of the fit, but would have been more data-set specific. We preferred to have consistency in the series of analyses we are undertaking using this method.

The purpose of the paper is not to compare different automated sentiment analysis packages, but to evaluate tweets using a specific and convenient Sentiment Analysis package. A comparison of the performance of alternative automated sentiment analysis packages will be considered in future research.

2. Research Method

The paper is a companion to other recent analyses of President Trump's tweets on the topic of climate change and his State of the Union Address 2018 [1,2]. It features the use of an R library package called 'sentiment'. The 'sentiment' package was written by De Vries (2012), is now archived from the current release of R, and can be loaded from 'Github.com'. The details of the seniment package are available in De Vries (2012) [3]. It is a dictionary-based method which calculates sentiment scores using affinity dictionaries. The program splits strings into words (by default at space), calculates an affinity score for each word, and returns the average, using a scale from +5 to −5.

The paper uses this package because it is more finely grained, categorizes five different sentiment emotions, namely joy, sadness, anger, fear and surprise, and reveals greater information about the emotional tenor of the text or string that is analysed.

The process of performing sentiment analysis requires textual input in a machine-readable format. Pre-processing is required to transform the text into single words, followed by what are common pre-processing steps: stopword removal, stemming, removal of punctuation, and conversion to lower case.

The limitations of the analysis should be borne in mind. The context of 'natural language processing', of which sentiment analysis is a component, is important. The use of sarcasm and other types of ironic language, including puns and backhanded compliments, are inherently problematic for machines to detect, when viewed in isolation. This is a potential issue, in particular, in the analysis of President Trump's tweets. Nevertheless, current methods are revealing, as will be seen in the next section which presents the results.

2.1. Results of the Analysis

The process commences with the results of the application of the sentiment package to President Trump's 125 tweets. The emotional content of these is shown in Figure 1 . Ignoring the 'unknown' category, the predominant emotion recognised in Figure 1 is 'joy', which accounts for 14.4 per cent of the total, followed by 'sadness' at 3.2 percent. 'Anger' and 'fear' both account for 1.6 percent,

and 'surprise' accounts for 0.8 per cent. 78.4 per cent of the tweets are not classified, but 14.4 per cent is classified as being 'joy', which is a positive emotion.
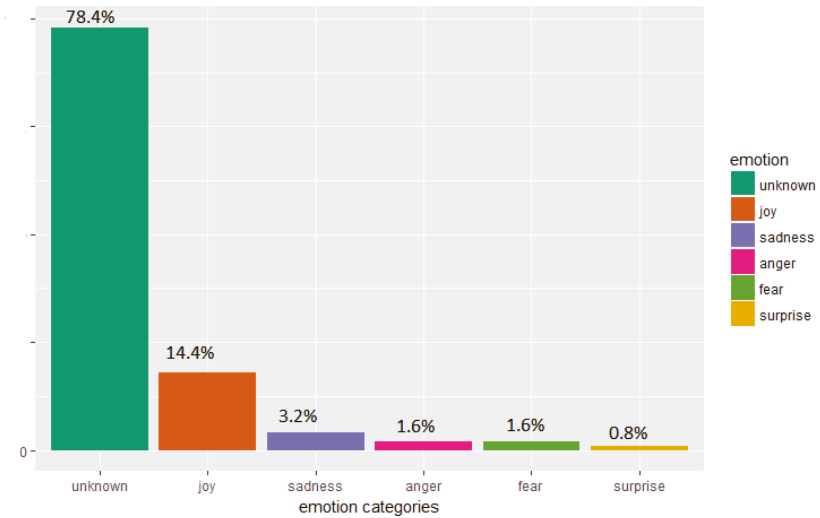
Figure 1. Trump's North Korea tweet sentiment.

Figure 2 classifies the tweets by President Trump according to whether they are negative, neutral or positive. The majority of the classifications in Figure 2 is positive, accounting for 72.8 per cent, while 16 per cent is negative.
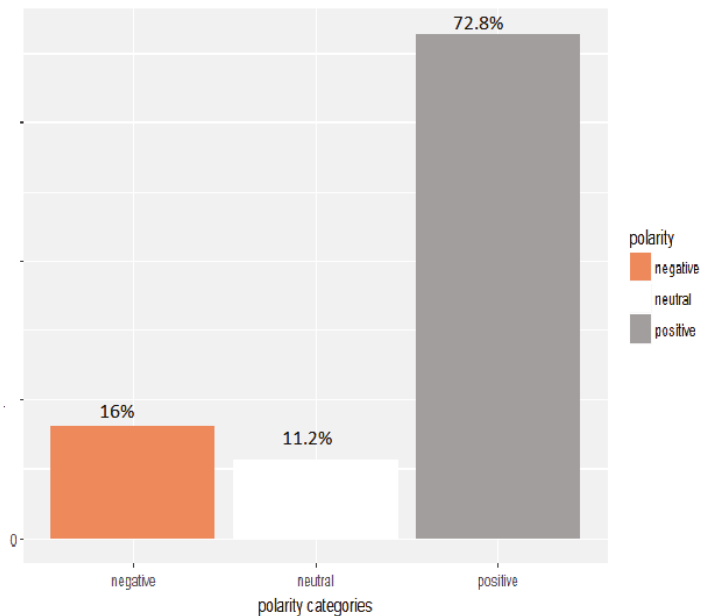
Figure 2. Trump's Korea tweet sentiment polarity.

Figure 3 shows a word cloud analysis of Trump's tweets. A word cloud is another form of visual representation of text data in which tags are single words, and their relative sizes and colours represent their weighting or importance in the context of the text considered.

Figure 3. Trump's North Korea tweets word cloud.

The most prominent words in the word cloud in Figure 3 are 'China', 'trade', 'US', 'working' and 'percent'. If we move around the cloud in an anti-clockwise manner, words in the 'joy' section include 'enthusiastic', 'freeze', 'fair', 'Syria', 'sanctions', 'korea', 'discussed', 'completed', 'deal', and so forth. In the 'sadness' section, we have 'unsuccessful', 'launched', 'moon', 'threaten', 'regime; 'military', 'gas', 'victim', 'hostage', 'immigration', 'brutality', and so on.

In the 'anger' section below, we have 'America', 'threats', 'respond', 'hostile', 'conducted', 'nuclear', 'test', 'dangerous','rogue', 'korean', and so on. In the 'fear' section, we see 'alert', 'attack', 'condemned', 'grossly', 'launch', 'impact', and so on. In the 'surprise' section, we have 'China', 'working', 'grew', 'US', 'per cent', 'trade', and so forth. The 'unknown' section has a diverse grouping of words, with 'news' and 'sadness', 'launches', 'talk', 'meet', 'press', and 'working', given prominence.

2.2. Bootstrapped t Tests

In a companion paper, Allen et al. [1] use sentiment analysis to analyse some of President Trump's tweets on the topic of climate change, takn from 2011 to 2015. On a global issues scale, it is worth comparing the sentiment scores of his tweets on North Korea and Supreme Leader Kim Jong-Un, with

his tweets on climate change. We use bootstrapped *t* tests because the number of tweets analysed are different, namely 125 in the North Korean sample and 115 in the climate change sample.

The results presented in Table 1 show that there is no significant difference in the bootstrapped *t* tests, using 1000 samples with replacement. Figure 4 shows that the bootstrapped *t* test vector QQ plot is consistent with a Gaussian distribution.

Table 1. Bootstrapped *t* test of the differences in the means in Trump's North Korea Tweets: emotion score 'joy' mean and Trump Climate Change Tweets emotion score 'joy'.

Trump's North Korea Tweets 'Joy' vs. Trump's Climate Change 'Joy'			
mean *t* = 0.467	mean *p*-value = 0.4521		
probability vect	1st Quartile	median	3rd Quartile
probability	0.2016	0.4215	0.6936
t vector	−0.2823	0.4941	1.1632
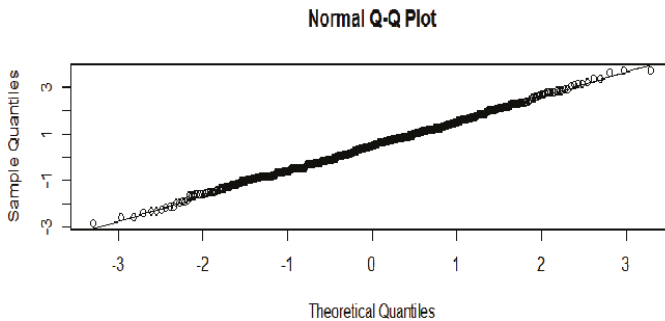
Figure 4. QQ plot of *t*: vector 'joy' comparisons.

A comparison was also made of the 'polarity' scores for President Trump's North Korean and Supreme Leader Kim Jong-Un tweets and his climate tweets. The mean polarity score for his North Korean Supreme Leader Kim Jong-Un tweets is 13.59, and 9.46 for his climate change tweets. The results are shown in Table 2. In terms of the results of the bootstrapped *t* tests, his tweets on North Korea have significantly higher positive polarity scores than his tweets on climate change.

Table 2. Bootstrapped *t* test of the differences in means in Trump's North Korea Tweets; positive polarity score mean and Trump's Climate Change Tweets positive polarity score.

Trump's North Korea Tweets; Positive Polarity score versus Trump's Climate Change Positive Polarity score			
mean *t* = 3.35	mean *p*-value = 0.0181		
probability vect	1st Quartile	median	3rd Quartile
probability	0	0.0001	0.008
t vector	2.67	3.34	3.996

3. Conclusions

This paper featured a sentiment analysis of 125 of President Trump's tweets on North Korea's Supreme Leader Kim Jong-Un. The sentiment analysis classified this series of tweets into five different emotional categories, with a large proportion of the total, 78 per cent, remaining unclassified. A comparison with the results in a companion paper featuring an analysis of President Trump's climate change tweets, using the 'joy' score reveals no significant differences in the scores for the two sets of tweets using bootstrapped *t* tests.

However, when the the coarser polarity classification is applied, using three categories, namely positive, negative and neutral, all the tweets can be classified. The results suggest that President Trump is far more positive in his tweets about North Korea than he is about climate change. This is confirmed by the results of the bootstrapped *t* tests.

In comparing the tweets on North Korea and climate change, on which Trump has tweeted many times, it is interesting to note that the tweets on North Korea have significantly higher positive polarity scores than his tweets on climate change.

It seems clear that President Trump likes to convey positive messages via his twitter feed, even when he faces thorny diplomatic challenges, as exemplified by North Korea's Supreme Leader Kim Jong-Un, or difficult scientific policy issues, such as climate change. It remains to be seen whether positive tweet sentiments assist or add to the complications faced in the resolution of these pressing and difficult global issues. "Stay tuned!"

References

1. Allen, D.E.; McAleer, M. *Fake News and Indifference to Scientific Fact: President Trump's Confused Tweets on Global Warming, Climate Change and Weather*; Econometric Institute Research Paper EI2018-24; Econometric Institute, Erasmus School of Economics (ESE), Erasmus University Rotterdam: Rotterdam, The Netherlands, 2018.
2. Allen, D.E.; McAleer, M.; Reid, D.M. *Fake News and Indifference to Truth: Dissecting Tweets and State of the Union Addresses by Presidents Obama and Trump*; Tinbergen Institute Discussion Paper, TI 2018-020/III; Tinbergen Institute: Amsterdam, The Netherlands; Rotterdam, The Netherlands, 2018.
3. De Vries, A. Sentiment Package. 2012. Available online: https://github.com/andrie/sentiment/blob/master/DESCRIPTION (accessed on 1 February 2018).

Article

Financial Risk Measurement and Prediction Modelling for Sustainable Development of Business Entities Using Regression Analysis

Katarina Valaskova *, Tomas Kliestik, Lucia Svabova and Peter Adamko

Faculty of Operation and Economics of Transport and Communications, University of Zilina, Univerzitna 1, 010 26 Zilina, Slovakia; tomas.kliestik@fpedas.uniza.sk (T.K.); lucia.svabova@fpedas.uniza.sk (L.S.); peter.adamko@fpedas.uniza.sk (P.A.)

* Correspondence: katarina.valaskova@fpedas.uniza.sk; Tel.: +421-910-454-102

Abstract: The issue of the debt, bankruptcy or non-bankruptcy of a company is presented in this article as one of the ways of conceiving risk management. We use the Amadeus database to obtain the financial and accounting data of Slovak enterprises from 2015 and 2016 to calculate the most important financial ratios that may affect the financial health of the company. The main aim of the article is to reveal financial risks of Slovak entities and to form a prediction model, which is done by the identification of significant predictors having an impact on the health of Slovak companies and their future prosperity. Realizing the multiple regression analysis, we identified the significant predictors in conditions of the specific economic environment to estimate the corporate prosperity and profitability. The results gained in the research are extra important for companies themselves, but also for their business partners, suppliers and creditors to eliminate financial and other corporate risks related to the unhealthy or unfavorable financial situation of the company.

Keywords: financial risk; bankruptcy; regression model; sustainable development; Slovak enterprises

1. Introduction

Financial risk is the possibility that shareholders will lose money when they invest in a company that has debt, if the corporate cash flow proves inadequate to meet its financial obligations. When a company uses debt financing, its creditors are repaid before its shareholders if the company becomes insolvent [1]. Financial risk is often perceived as the risk that a company may default on its debt payments. To eliminate potential financial risks and to be able to identify the level of the corporate financial health, predictions models are used, perceived as systems of timely warning of impending problems in the analyzed companies. Their task is to evaluate the financial health of the company based on selected financial indicators or other characteristics of the company or the environment in which they operate [2].

The originality of the research lies in the identification of crucial determinants in Slovak conditions than can predict either prosperity and profitability of Slovak companies or their default (bankruptcy), without regard to any sector, and thus eliminate potential financial risks threatening the company and its business partners. Determination of prosperity predictors in Slovak conditions can help form a complex Slovak multi-industry prediction model, which would be beneficial for all market subjects, as until the present time we only adopt the results of the models developed in foreign countries, the use of which in our conditions is disputatious.

The main aim of the paper is to extend the knowledge about identification and elimination of financial risks related to the unhealthy financial situation of the company, which is done by the formation of the regression model, results of which enable to estimate the profitability of the company.

The purpose of the paper is to measure financial risks considering national conditions. The research problem includes the formation of an econometric model of the corporate prosperity quantification, using the results of the regression analysis, based on the significant financial indicators identified in the multiple linear regression analysis. We consider the identification of the most significant predictors affecting the future prosperity and profitability of Slovak enterprises to be the main contribution of the paper; those are working capital, working capital to total assets ratio, current assets to total assets ratio, operating profit to total assets ratio, cash and cash equivalents to total assets ratio and current liabilities to total assets ratio.

The paper is divided into four main parts. Literature review highlights the current state of research in the field of prediction and bankruptcy models. Material and Methods depicts a brief description of business entities and financial indicators used as potential predictors in the research and specifies the methodology of the multiple linear regression. Chapter Results is focused on the description of all findings, resulting in the suggestion of the model, which estimates the corporate prosperity and profitability and thus eliminates financial risks. Discussion compares and analyses the studies and researches of other authors in the field of prediction models and emphasizes the various combinations of different financial indicators used as predictors in the models and compares the results of the realized study with results of other studies based on different calculation methods.

Literature Review

Financial risk measurement is a largely investigated research area; its relationship with imprecise probabilities has been mostly overlooked. Vicig [3] claims that risk measures can be viewed as instances of upper (or lower) previsions, thus letting us apply the theory of imprecise previsions to them. A complex approach to risk measurement in financial management is described in the work of Chobot [4,5]. Except for well-known risk measures, including value at risk [6] or coherent and convex risk measures [7], there are many others methods that authors use to measure financial risks. Su and Furman [8] apply a form of multivariate Pareto distribution to measure financial risks. Spatial financial time series models were introduced by Blasques et al. [9], Yang et al. [10] and Audrino and Barone-Adesi [11]. Kessler [12] presents an implementational systematic approach framework for risk, where the risk management target is to manage and mitigate the risk-around-loss causes. Campos et al. [13] underline the importance of innovative soft-computing techniques usage to classify correctly the default of a company by proper financial credit risk prediction. Chai and Xia [14] emphasize that to survive and develop in a drastically competitive market, business entities need to control possible financial risks and foresee their future financial development (using prediction models).

Since the first prediction model developed by Fitzpatrick [15], there have been numerous researches made and various predictors have been identified to predict the future situation of the business entities, e.g., Beaver model [16], Altman model [17], Springate model [18], Ohlson model [19], Taffler-Tisshawa model [20], Fulmer model [21], Zmijewski model [22], Horrigan model [23] etc. The accurate prediction of corporate bankruptcy for the companies in different industries is of a great concern to investors and creditors, as the reduction of creditors' risk can be possible [24]. The systematic review of bankruptcy prediction models is processed in the studies of Alaka et al. [25] or Peres and Antao [26]. The reviews show that there are two groups of popular and promising tools within the bankruptcy prediction models research area, i.e., statistical tools (multiple discriminant analysis and logistic regression) and artificial intelligence tools (decision trees, neural networks, etc.). In this study, we test the use of a quite simple classifier, linear regression approach (similar to Guo et al. [27]), for modelling the relationship between a scalar dependent variable and more explanatory variables (financial indicators) as it performs reasonably well in bankruptcy prediction, as proved by Jones et al. [28]. Regression analysis if often use for bankruptcy prediction, the realized analysis is supported by the study of Calabrese et al. [29] or latest researches in Romania [30] and Lithuania [31], which recommend regression models for bankruptcy prediction. A methodological framework of regression was used to construct predictive bankruptcy models for Asia, Europe and

America and the results verify the superiority of the global model compared to regional models [32]. Ben Jabeur [33] claims that regression model gives the opportunity to consider all the indicators in predicting financial distress. Hwang and Chu [34] propose a new procedure to estimate the loss given default using logistic regression. Li and Miu [35] establish a prediction model with dynamic loading on accounting ratio-based and market-based information using a regression approach.

In Slovak business environment, there are also a few representatives of prediction models. Chrastinová [36] and Gurcik [37] applied the methodology of financial health predictions to companies in the agricultural sector, Binkert [38] and Zalai [39] in commercial enterprises using multiple discriminant analysis. There is not any reputable prediction model in Slovakia, but several studies and researches have been developed. Kameníková [40] solved the limitations in the use of foreign models predicting the financial development of enterprises in conditions of the Slovak Republic. Lesaková [41] states that top management, based on predictions and forecasts, formulates the financial targets of the enterprise for the appropriate time horizon. Horvathova and Mokrisova [42] diagnosed business performance applying the modern financial performance assessment methods. Gundova [43] depicted the main reasons for not using foreign methods of predicting the financial situation in Slovak companies and underlined the importance of the formation of the national prediction model. The application of foreign prediction models and their modification in our conditions is searched by Adamko [44], Boda and Uradnicek [45], Hiadlovsky and Kral [46]. A method for logistic regression to assess the future corporate prosperity was in our national conditions firstly applied by Hurtošová [47]. Later, Delina and Packova [48] developed a new modified model in Slovak business environment while using regression analysis to get higher predictive performance of the model. Kovacova and Kliestik [49] intorduced a bankruptcy prediction model in the Slovak Republic using logistic regression and they proved significant classification accuracy of this model. Results of the last mentioned are significant but deeper research has to be done to develop a complex prediction model of the financial health of Slovak companies.

2. Materials and Methods

The aim is to form an econometric multi-industry model in Slovak environment to quantify the prosperity of the company in terms of the achieved economic result. For this reason, we used the Amadeus database; we chose the accounting and financial records of accounting entities operating in the territory of the Slovak Republic in the years 2015 and 2016. Companies included in the model were chosen considering the Nomenclature of Economic Activities in the European Community (NACE classification), representing a statistical classification based on a common statistical classification of economic activities in the European Union. We include the following economic categories in the model: A—agriculture, forestry and fishing; B—mining and quarrying; C—manufacturing; D—electricity, gas, steam and air conditioning supply; F—construction; G—wholesale and retail trade; H—transporting and storage; I—accommodation and food service activities; J—information and communication; N—administrative and support service activities; P—education; Q—human health and social work activities. The method of multiple linear regression was used to create the model; independent variables were calculated from the data of 2015, the dependent variable is from the records of 2016. Multiple linear regression consists of the following methodological steps:

1. Choosing a sufficiently large sample that accepts some of the rules for determining the appropriate sample size to perform the regression analysis. We used the Stepwise method, which does multiple regression several times, each time removing the weakest correlated variable. At the end, only those variables, that explain the distribution best, are left. The only requirements are that the data is normally distributed and that there is no correlation between the independent variables.

For this type of regression, at least 40 measurements should be added to each variable. We include 37 quantitative variables; the size of our sample from the database is more than 120,000 enterprises, so the sample size meets the necessary requirements.

2. The dependent variable was defined as the corporate prosperity and profitability measured by EBIT (marked as OPPL). We decided to choose the independent variables using the predictors, which are the most frequently used in the prediction models worldwide [50]. Identification of independent variables is summarized in Table 1.

Prosperity and profitability of the company in the future may be partly given by optimal values of the financial indicators [51]. Based on the calculated financial ratios we are able to classify the companies into two groups: default (unhealthy, non-prosperous) and non-default (healthy, prosperous) in the context of legislative adjustments. We consider three criteria, which have to be met simultaneously and which correspond with the default criteria determined by the Slovak legislation. If the value of the corporate solvency ratio is less than 0.4, current ratio is less than 1 and net income is negative, the company is not prosperous, if conditions are not met, the company is healthy and prosperous. Despite the fact, that the study identifies a set of explanatory variables that can help identifying the state of a company, we consider only two states of the corporate prosperity—default of non-default. We follow the Slovak Commercial Code defining the principles and economic criteria of the company in default, which were used to determine the dependent variable.

Table 1. Selected financial ratios.

Financial Ratios			
X1	Sales/Total assets	X20	Net income/Sales
X2	Current assets/Current liabilities	X21	Non-current liabilities/Total Assets
X3	Gross profit/Total assets	X22	Cash and cash equivalents/Current liabilities
X4	Net income/Shareholders equity	X23	Cash flow/Current liabilities
X5	EBITDA/sales	X24	Working capital/Sales
X6	(Non-current + current liabilities)/EBITDA	X25	Current ratio
X7	Net income/ Total assets	X26	Liquidity ratio
X8	Working capital/Total assets	X27	Return on assets
X9	Operating profit/Total assets	X28	Return on equity
X10	(Non-current + current liabilities)/total assets	X29	Shareholder liquidity ratio
X11	Current assets/Total assets	X30	Solvency ratio (liability-based)
X12	Cash & cash equivalents/Total assets	X31	Cash flow/Operating revenue
X13	Cash flow/Total assets	X32	Net assets turnover
X14	Cash flow/(Non-current + current liabilities)	X33	Interest paid
X15	Current liabilities/Total assets	X34	Gross margin
X16	Current assets/Sales	X35	Profit margin
X17	Operating profit/interest paid	X36	Net current assets
X18	Stock/Sales	X37	Working capital
X19	Cash flow/Sales		

3. Testing of Gauss-Markov assumptions: dependent and independent variables must be quantitative; the multi-collinearity condition must be complied; the outliers have to be removed; the variables must be in a linear relation (tested by Pearson correlation coefficient). We test the hypothesis of dependence between the individual independent variables and the dependent variable on the significance level of 0.05, which is compared to the critical p-value of the test of significance of Pearson correlation coefficient. Last assumption is to ensure normal distribution of model residuals that cannot be auto-correlated [52].

4. Realization of multiple linear regression and testing the significance of the individual independent variables in the model.

Multiple linear regression models the dependent variable as a linear combination of independent variables and an intercept [53]:

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j \cdot x_{ij} + u_i \qquad (1)$$

where:

y_i dependent variable
x_{ij} independent variable(s)
β_0, β_j unknown parameters of the model
u_i random variable

Parameters β_j are considered as unknown numerical constants, β_0 is an absolute number and, in general, β represents a slope (direction) of parameters. The parameter β_j explains the changes in the value of the dependent variable y_i, if the *j*-th independent variable x_{ij} changes of one unit, provided, that the values of other independent variable stay unchanged.

5. Testing the significance of the created model.

6. Write the equation of the regression model.

To provide the multiple regression analysis we used the software IBM SPSS Statistics v. 19. We consider all business entities in the database, accepting the selected sectors and their specificities, as we want to determine the general predictors to assess the future corporate prosperity of any company.

3. Results

Before the regression analysis itself, we test the mentioned Gauss- Markov assumptions. The regression analysis is very sensitive to outliers. To exclude all abnormal and extreme values, we used interquartile range, multiplied by the number 2.2, which is often used to detect outliers in the data. We modified the original database and used the remaining 105,708 enterprises in the regression model. One of the mentioned preliminary conditions is the character of dependent and independent variables, all of them are quantitative. However, it was not possible to calculate the values of some of the determined financial ratios due to missing or not available information in the Amadeus database, they had to be excluded from the regression. As a result, not 37 but 24 ratios are the proposed financial predictors. Their descriptive statistics (mean, standard deviation and coefficient of variation) are summarized in Table 2.

Table 2. Descriptive statistics of independent variables.

Independ. Variables	X1	X2	X4	X7	X8	X9	X10	X11
mean	1.92	4.41	0.14	0.09	0.15	0.14	0.51	0.76
std. dev.	3.92	7.76	1.05	0.18	0.29	0.21	0.31	0.28
var. coef.	2.03	1.76	7.50	2.00	1.93	1.50	0.61	0.37
	X12	**X15**	**X16**	**X18**	**X20**	**X21**	**X22**	**X24**
mean	0.37	0.45	9.45	5.73	−0.32	0.06	2.76	5.22
std. dev.	0.33	0.30	1560.19	1282.47	99.42	0.14	6.06	1040.72
var. coef.	0.89	0.67	165.10	223.82	−310.69	2.33	2.19	199.37
	X25	**X26**	**X27**	**X28**	**X30**	**X35**	**X36**	**X37**
mean	4.40	4.07	0.13	0.26	3.98	0.12	138.10	174.21
std. dev.	7.76	7.55	0.20	0.89	21.26	0.24	6419.88	4055.82
var. coef.	1.76	1.86	1.54	3.42	5.34	2.00	46.49	22.28

The assumption of the collinearity presents the high mutual correlation of variables. Multi-collinearity among the independent variables can cause the incorrect formulation of the model or could decrease the prediction ability of some variable. The simplest way to solve the existing multi-collinearity is to remove one of two independent variables with the mutual interdependence [54] and repeat the analysis. Table 3 shows the collinearity between the variables.

Table 3. Collinearity diagnosis.

Dimension	Eigenvalue	Condition Index	Variance Proportions						
			C	X37	X08	X11	X09	X12	X15
1	3.926	1.000	0.10	0.00	0.01	0.00	0.02	0.01	0.01
2	1.022	1.960	0.00	0.80	0.06	0.00	0.01	0.01	0.00
3	0.884	2.107	0.00	0.19	0.31	0.00	0.05	0.04	0.01
4	0.680	2.402	0.00	0.00	0.12	0.00	0.42	0.00	0.11
5	0.366	3.276	0.00	0.00	0.08	0.01	0.48	0.17	0.17
6	0.077	7.119	0.53	0.00	0.25	0.04	0.02	0.46	0.70
7	0.044	9.447	0.46	0.00	0.19	0.95	0.00	0.33	0.00

The collinearity diagnostics follows several important values to reveal the problems with multi collinearity-the eigenvalue, the condition index and the variance inflation factor. The resulting values of eigenvalues are different from 0 (and are not close to 0), indicating that the predictors are not intercorrelated.

The condition index is computed as the square root of the ratios of the largest eigenvalue to each successive eigenvalue. When two or more of the supposedly independent variables are correlated, the condition index for each will be above one. Values of one are independent; values of greater than 15 suggest there may be a problem, while values of above 30 indicate a serious problem. The resulting values of the condition index confirm that there are not any multi collinearity problems.

The variance inflation factor (VIF), calculated in Table 4, measures the impact of collinearity among the variables in a regression model. It is always greater than or equal to one. There is no formal VIF value for determining presence of multicollinearity; however, values that exceed 10 are often regarded as indicating multicollinearity. Based on the results in the model it can be concluded, that there is no multicollinearity symptom, as all values are between 1 to 10.

Table 4. Collinearity measured by VIF.

		Collinearity Statistics	
Model		Tolerance	VIF
6	Constant		
	X37	0.997	1.003
	X08	0.631	1.585
	X11	0.529	1.891
	X09	0.869	1.150
	X12	0.406	2.465
	X24	0.755	1.324

Gauss- Markov assumption of a liner relationship between variables claims that it is necessary to have individual independent variables in a linear relation to the dependent variable. The existence of linearity is determined by the Pearson correlation coefficient, Table 5. Indicative limits to determine the dependence by Pearson correlation coefficient in this study are (in both positive and negative relationships) [53]:

$0 < |r| \leq 0.3$ weak dependence
$0.3 < |r| \leq 0.8$ medium dependence
$0.8 < |r| \leq 1$ strong dependence

Table 5. Pearson correlation matrix.

	X1	X2	X4	X7	X8	X9	X10	X11	X12	X15	X16	X18
prosperity	−0.001	−0.008	0.005	0.008	−0.003	0.006	−0.001	−0.026	−0.024	−0.011	0.002	0.001
	X20	**X21**	**X22**	**X24**	**X25**	**X26**	**X27**	**X28**	**X30**	**X35**	**X36**	**X37**
	0.000	0.020	−0.011	0.002	−0.008	−0.009	0.007	0.005	−0.003	0.007	0.477	0.624

Values 0.000 means the figure is too small for three decimal place representation. It is clear, that there is a weak linear dependence between the independent variables and the dependent variable, except for X36 and X37 where their mutual relation with the dependent variable is described by the medium dependence.

We test the hypothesis of mutual dependence between the individual independent variables and the dependent variable on the significance level of 0.05, which is compared to the *p*-value of the test of significance of Pearson correlation coefficient, Table 6.

Table 6. *P*-value of Pearson correlation coefficient.

	X1	X2	X4	X7	X8	X9	X10	X11	X12	X15	X16	X18
prosperity	0.356	**0.013**	0.090	**0.021**	0.224	0.064	0.403	**0.000**	**0.000**	**0.002**	0.271	0.383
	X20	**X21**	**X22**	**X24**	**X25**	**X26**	**X27**	**X28**	**X30**	**X35**	**X36**	**X37**
	0.486	**0.000**	**0.002**	0.270	**0.013**	**0.009**	**0.042**	0.100	0.194	**0.027**	**0.000**	**0.000**

Based on the data shown in Table 4, we found that the p-value is higher than the significance level of some independence variables, so we claim that there is not any dependence between these independent variables and the dependent variable. However, Pearson correlation coefficient shows weak but existing linear dependence between these independent variables and the dependent variable, we decided to include these variables in the model of the corporate prosperity estimation. Considering the independent variables X2, X7, X11, X12, X15, X21, X22, X25, X26, X27, X35, X36 and X37, the *p*-value is lower than the level of significance, so we claim that there is a dependence between the individual independent variables and the dependent variable.

Gauss- Markov assumptions mentioned in the methodological part were fulfilled (the assumption of normal distribution and autocorrelation can be tested after the model formation) and the multiple linear regression can be performed.

Stepwise method of the regression analysis eliminates the multi-collinearity problems, constructs different models and shows statistics for each model, composed of different sets of variables. These models are the combinations of independent variables that best explain the dependent variable. Table 7 depicts the significant variables of the model.

Table 7. Variables entered/removed.

Model	Variables Entered	Variables Removed	Method
1	X37	.	Stepwise (Criteria: Probability-of-F-to-enter <= 0.050, Probability-of-F-to-remove >= 0.100).
2	X08	.	Stepwise (Criteria: Probability-of-F-to-enter <= 0.050, Probability-of-F-to-remove >= 0.100).
3	X11	.	Stepwise (Criteria: Probability-of-F-to-enter <= 0.050, Probability-of-F-to-remove >= 0.100).
4	X09	.	Stepwise (Criteria: Probability-of-F-to-enter <= 0.050, Probability-of-F-to-remove >= 0.100).
5	X12	.	Stepwise (Criteria: Probability-of-F-to-enter <= 0.050, Probability-of-F-to-remove >= 0.100).
6	X15	.	Stepwise (Criteria: Probability-of-F-to-enter <= 0.050, Probability-of-F-to-remove >= 0.100).
	Dependent Variable: OPPL		

The regression analysis reveals that the model includes six statistically significant independent variables, which best explains the variability of the dependent variable considering the order, in which they were added into the model. The multi-industry model of the corporate prosperity quantification in conditions of Slovak enterprises consists of these predictors: working capital, working capital to total assets ratio, current assets to total assets ratio, operating profit to total assets ratio, cash and cash equivalents to total assets ratio and current liabilities to total assets ratio.

It is interesting to compare the results of Pearson correlation coefficient with the results of the relevant independent variables according to the regression analysis. In most cases, both analysis provide the same results, i.e., if the results of Pearson correlation coefficient indicates to reject a significant relationship between the variables, the regression analysis often proves the same. The difference was only in the case of the independent variables X_8 and X_9, which the regression analysis considered to as significant attributes affecting the value of the corporate prosperity and profitability. The overall correlation between the variables left in the models (we consider six models) and the dependent variable is shown in Table 8, which portrays particular steps of addition or subtraction of variables from the set of explanatory variables based on some pre-specified criteria.

Table 8. Quality of the regression model (Model summary).

	Predictors in the Model	R	R Square	Adj R Square	Std. Error	Durbin-Watson
1	Constant (C), X37	0.624	0.389	0.389	2562.368	
2	C, X37, X8	0.625	0.390	0.390	2559.914	
3	C, X37, X8, X11	0.625	0.390	0.390	2559.385	
4	C, X37, X8, X11, X9	0.625	0.391	0.391	2558.734	
5	C, X37, X8, X11, X9, X12	0.625	0.391	0.391	2558.664	
6	C, X37, X8, X11, X9, X12, X15	0.625	0.391	0.391	2558.469	1.999813
	Dependent Variable: OPPL					

R squared presents the percentage of the variation in the dependent variable that is explained using the independent variables included in the model. The model 6, which includes all the relevant model predictors, explains 39.1% of the variation in the dependent variable. Adjusted R-squared indicates how well terms fit a curve or line, but adjusts for the number of terms in a model, in our case 39.1%.

Table 9 presents the linear regression equation coefficients for the various model variables.

Table 9. Coefficients of the models.

		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
6	Constant	265.557	33.305		7.973	0.000
	X37	0.505	0.002	0.625	211.843	0.000
	X08	−427.954	41.398	−0.038	−10.338	0.000
	X11	−138.346	47.360	−0.012	−2.921	0.003
	X09	284.326	50.352	0.018	5.647	0.000
	X12	−156.817	45.620	−0.016	−3.437	0.001
	X24	−125.806	36.777	−0.012	−3.421	0.001

The significance (Sig.) should be below the significance level 0.05 to consider all predictors significant for the model. All independent variables are below the determined significance level and thus may be used as relevant predictors in the multi-industry model. The statistical significance of the model is proved by the F-test (Table 10).

Table 10. Statistical significance of the final regression model (*F*-test).

Model	DF	Sum of Squares	Mean Squares	*F*	Sig.
Regression	6	2.845×10^{11}	4.914×10^{10}	7507.453	0.000
Residual	105,700	4.757×10^{11}	6,545,762.407		
Total	105,707	7.602×10^{11}			

The result value of the calculated F statistics is again compared to the significance level of 0.05, and as it is below the determined level, we can conclude that the model is statistically significant.

Gauss-Markov assumption of normal distribution and autocorrelation applies to model residues can be tested after the regression. Within the regression analysis, emphasis is given on the normality of residues. If the residues were not normally distributed, the results could be inaccurate. Central limit theorem guarantees that the violation of the normal distribution in large sample sets ($n > 100$) does not have critical consequences [55]. Autocorrelation was tested by Durbin-Watson test, its value 1.9998 (see Table 7) is compared with the critical value and thus we do not reject the null hypotheses that the residuals are not auto-correlated.

The final notation of the model of the prosperity quantification, based on the corporate profitability, in conditions of the Slovak enterprises is:

$$corporate\ prosperity = 265.557 + 0.505X_{37} - 427.954X_{08} - 138.346X_{11} - +284.326X_{09} - 156.817X_{12} - 125.806X_{15} \tag{2}$$

The multi-industry model of the corporate prosperity and profitability shows, that the value of the intercept is the limit value, which means, that if all financial ratios are zero and the company has the value of the corporate prosperity equal or less than the constant, the future prosperity and profitability of the company is bad, it is non-prosperous. In that case, its business partners have to consider their cooperation in the future or take measures to eliminate or prevent the financial risks. It the value of the corporate prosperity is higher than the constant the company is considered profitable in the future. Ceteris paribus, the value 0.505 X37 means that if the value of the working capital increases/decreases by one measure unit, the value of the corporate prosperity increases/decreases of 0.505 €. The value 427.964 X8 presents that if the value of the working capital to total assets ratio increases/ decreases by one measure unit, the value of the corporate prosperity increases/decreases of 427.954 €. The value 138.346 X11 determines that if the value of current assets to total assets ratio increases/ decreases by one measure unit, the value of the corporate prosperity increases/decreases of 138.346 €. The value 284.326 X09 means that if the value of the operating profit to total assets ratio increases/ decreases by one measure unit, the value of the corporate prosperity increases/decreases of 284.326 €. The value 156.817 X12 presents that if the value of the cash and cash equivalents to total assets ratio increases/ decreases by one measure unit, the value of the corporate prosperity increases/decreases of 156.817 €. In addition, the value 125.806 X15 states that if the value of the current liabilities to total assets ratio increases/decreases by one measure unit, the value of the corporate prosperity increases/decreases of 125.806 €.

Given that the coefficient of determination of our model is 39.1% we can describe only slightly more than 39% of changes in the value of corporate prosperity. The remaining changes in the prosperity value may be caused by other, and also non-measurable, factors that we were not able to quantify and measure or by other factors that may not be related to prosperity and profitability of the Slovak companies. Based on the results of the multiple linear regression analysis we can identify the financial predictors, which play a crucial role in the process of the corporate prosperity quantification and financial risks identification, those are: working capital, working capital to total assets ratio, current assets to total assets ratio, operating profit to total assets ratio, cash & cash equivalents to total assets ratio and current liabilities to total assets ratio.

To verify the prediction ability of the estimated model, we use Equation (2) to predict the future corporate prosperity, which was compared with the real values of the dependent variable OPPL(0 is for prosperous companies, 1 for the non-prosperous ones). The results are portrayed in Table 11.

Table 11. Prediction ability of the model (classification results).

			0	1	Total
Non-prosper. real	**0**	**count**	46,153	30,152	76,305
		%	60.5%	39.5%	100.0%
	1	**count**	13,284	16,119	29,403
		%	45.2%	54.8 %	100.0%
Total		**count**	59,437	46,271	105,708
			56.2%	43.8%	100.0%
58.91 % of original grouped cases correctly classified					

It is obvious that the formed model of the corporate prosperity identified correctly 60.5% of prosperous companies and 54.8% of non-prosperous companies, which corresponds to the weak level of the coefficient of determination. The total prediction ability of the model is 58.91%, which Hampel and Klepáč [56] classify as an acceptable prediction ability.

4. Discussion

The financial risk measurement and prediction modelling for sustainable development of business entities using regression analysis proved, that the predictors of the model are acceptable to be used to predict the future prosperity of Slovak business entities. However, its prediction ability is not sufficient, which is the consequence of the method used. The same database of companies was used to predict the future development of companies by multiple discriminant analysis, logistic regression. The overall classification ability of the model formed by the multiple discriminant analysis is 73%; however, the more important information is the correct classification of non-prosperous entities, which is at the level of 93% [57]. The results of the logistic regression model claim, that the overall percentage of correct classification is slightly above 79%, with more than 84% of non-prosperous companies correctly classified [58]. Significant results were proved also in the study of Rohacova and Kral [59], who used data envelopment analysis to predict the corporate failure.

The wide usage of the Altman model as a measure of a financial distress of strength in the economic and financial research points out that it is widely accepted as a reasonable, simple and consistent measure of the distressed entity at risk [60]. Thus, this model was tested in the conditions of Slovak business environment. In the research of Adamko and Svabova [61] Altman model was tested on the data of Slovak entities; the prediction ability of the model is 88.17%. Comparing the results of the studies realized in the Slovak business environment and based on different calculation method, it is clear, that the prediction ability of the latest Altman model slightly outperforms the other methods used. However, it has to be emphasized, that the informative value of some indicators of Altman model are significantly different in the economy with developed capital market and in the economy with less developed market, which is the case of Slovakia as the market does not reflect the expectations of the capital market.

Despite the fact that the companies in the database differ widely in their capital structure, firm size, access to external finance, management style, number of employees, the risk of financial failure can be modelled using the same set of independent variables for both prosperous and non-prosperous companies, which is confirmed by the study of Gupta et al. [62]. This knowledge leads to the identification of factors, which are significant enough to manage financial risks, and to affect the profitability and prosperity of the company. A similar research was conducted by Faltus [63], his research was aimed at finding the optimal default prediction model for Slovak companies using

the logistic regression, and Guo et al. [27], who used linear regression models and introduced a new parallel maximum likelihood estimator for multiple linear models fitted on the bankruptcy data.

Sharifabadi et al. [64] in their study of the impact of financial ratios on the prediction of bankruptcy of small and medium companies suppose the current assets to total assets ratio and operating profit to total assets to be the important indicators. Tian et al. [65] consider in their study 39 financial and market variables as candidate bankruptcy predictors, 85% of them are similar to independent variables used in our study. The most significant variables included in more than 5 models were recognized in the study of Bellovary et al. [66]. According to the results of this study, the predictors left in our model are significant variables included in many models worldwide. Current ratio appears in 51 prediction models, current assets to assets ratio in 10 and operating profit to total assets in 9, both working capital to total assets ratio and working capital in 7 models.

Ravi Kumar and Ravi [50] analyzed 62 prediction models and ranked most significant explanatory variables. Four out of six predictors used in the model are in the list of the most important explanatory variables; operating profit to total assets, ratio of current assets and total assets, current liabilities to total assets ratio and working capital to total assets.

The results of the study of 47 prediction models provided by Dimitras et al. [67] summarize the number of countries and number of models that include particular financial ratios. They identified 18 significant explanatory variables used in the prediction models worldwide. In the model, four of them are included: working capital ratio used in 5 countries and 16 models, current assets to total assets (6 countries and 12 models), operating profit to total assets (4 countries, 11 models) and net current liabilities to total asset (3 countries, 9 models).

Kliestik et al. [68] determined currently most commonly used explanatory variables and the number of studies in which they are included. Three ratios included in our model are from the list: current assets to total assets, operating profit to total assets and current liabilities to total assets. Moreover, the use of specific explanatory variables was revealed in the models of Visegrad countries [69].

In the study of Mihalovič [70], author focuses on the comparison of overall prediction performance of the two developed models, discriminant analysis and logistic regression, in conditions of the Slovak Republic and he reveals the most significant predictors net income to total assets, current ratio and current liabilities to total assets

Considering the studies on the most commonly used variables of the prediction models we can claim, that the statistically significant variables in the model of corporate prosperity belong to the group of variables, which are accepted by experts in this field. Mousavi et al. [71] conclude that the choice and design of independent variables and their nature affect the overall performance of the model. It is obvious that there are significant differences among variables used in various models and that for different countries with different type of economy should be developed a unique model with appropriate variables. The predictors identified in the study may be further applied in the formation of the complex prediction model in conditions of the Slovak Republic.

5. Conclusions

The bankruptcy prediction modelling helps predict the financial distress of companies. The importance of the area is underlined by the fact, that the information about the future corporate prosperity eliminates potential financial risks and enables to evaluate the financial health of the company based on selected financial indicators or other characteristics of the company or the environment in which they operate.

Realizing the multiple regression analysis, we identify the statistically significant determinants that affect the future financial development of the company and thus we form a regression model to estimate the corporate prosperity and profitability. As the statistically significant predictors were determined seven financial ratios: working capital, working capital to total assets ratio, current assets to total assets ratio, operating profit to total assets ratio, cash and cash equivalents to total assets ratio

and current liabilities to total assets ratio. These factors are significant enough to manage financial risks and to affect the profitability and prosperity of the company and can be later used in the model to predict the default of Slovak companies.

The multi-industry model of the corporate prosperity and profitability perceives the value of the intercept as the limit value, which means, that if the company has the value of the corporate prosperity equal or less than intercept value, there is a thread of financial problems in the future. Moreover, the corporate business partners have to consider their cooperation with the company in the future or take measures to eliminate or prevent the financial risks. The model has some limitation that is to be mentioned, and it is the low value of the R square (39.1%) which means, that there is a space for unknown and unmeasurable changes than can have some impact on the corporate prosperity and insufficient total prediction ability (58.91%). The choice of the method of linear regression may not be perceived positively, but despite that fact, we were able to identify crucial predictors to be used in the further research and also to quantify the future prosperity of the entities in the database. The further research with the same data revealed that it is more appropriate to use either the multiple discriminant analysis or the logistic regression to predict the future prosperity of any company.

The main aim of the paper was to extend the knowledge about identification and elimination of financial risks related to the unhealthy financial situation of the company. The results gained in the multi-industry model are extra important for companies themselves, but also for their business partners, suppliers and creditors to eliminate financial and other corporate risks related to the unhealthy or unfavorable financial situation of the company.

The formation of the complex prediction model in the economic conditions of the Slovak Republic is still missing, and thus the results of our research may be used to determine the financial ratios that can be, based on the future detailed research, used as the predictors of the Slovak prediction model.

References

1. Vinczeova, M.; Kascakova, A. How do Slovak small and medium-sized enterprises make decision on sources of finance? *Ekon.-Manazerske Spectr.* **2017**, *11*, 111–121.
2. Svabova, L.; Kral, P. Selection of Predictors in Bankruptcy Prediction Models for Slovak Companies. In Proceedings of the 10th International Days of Statistics and Economics, Praha, Czech Republic, 8–10 September 2016; Loster, T., Pavelka, T., Eds.; Melandrium: Slany, Czech Republic, 2016; pp. 1759–1768.
3. Vicig, P. Financial risk measurement with imprecise probabilities. *Int. J. Approx. Reason.* **2008**, *49*, 159–174. [CrossRef]
4. Chobot, M. Approaches to risk measurement in strategic and financial management 1. *Ekon. Cas.* **1993**, *41*, 751–758.
5. Chobot, M. Approaches to risk measurement in strategic and financial management 2. *Ekon. Cas.* **1994**, *42*, 40–54.
6. Bali, T.G. A generalized extreme value approach to financial risk management. *J. Money Credit Bank.* **2007**, *39*, 1613–1649. [CrossRef]
7. Drakos, A.; Kouretas, G.P. Bank ownership, financial segments and the measurement of systematic risk: An application of CoVaR. *Int. Rev. Econ. Financ.* **2015**, *40*, 127–140. [CrossRef]
8. Su, J.; Furman, E. A form of multivariate Pareto distribution with applications to financial risk measurement. *Astin Bull.* **2017**, *47*, 331–357. [CrossRef]
9. Blasques, F.; Koopman, S.J.; Lucas, A.; Schaumburg, J. Spillover dynamics for systemic risk measurement using spatial financial time series models. *J. Econom.* **2016**, *195*, 211–223. [CrossRef]

10. Yang, X.; Chatpatanasiri, R.; Sattayatham, P. Dynamic Risk Measurement of Financial Time Series with Heavy-Tailed: A new hybrid Approach. *Thai J. Math.* **2016**, *14*, 265–281.

11. Audrino, F.; Barone-Adesi, G. Functional gradient descent for financial time series with an application to the measurement of market risk. *J. Bank. Financ.* **2005**, *29*, 959–977. [CrossRef]

12. Kessler, A.M. A systematic approach to operational risk measurement in financial institutions. *J. Oper. Risk* **2007**, *2*, 27–68. [CrossRef]

13. Campos, R.; Riuz, F.J.; Agell, N.; Angulo, C. Financial credit risk measurement prediction using innovative soft-computing techniques. In *Computational Finance and its Applications*; Costantino, M., Brebbia, C.A., Eds.; WIT Press: Southampton, UK, 2004; Volume 38, pp. 57–66.

14. Chai, Y.; Xia, E. Financial risk measurement and prevention. In Proceedings of the International Conference on Risk and Reliability Management, Beijing, China, 10–12 November 2008; Li, J.L., Ed.; Universe Academic Press: Toronto, ON, Canada, 2008.

15. Fitzpatrik, P.J. A Comparison of the Ratios of Successful Industrial Enterprises with Those of Failed Companies. *Certif. Public Account.* **1931**, 598–605.

16. Beaver, W.H. Financial Ratios as Predictors of Failure. Empirical Research in Accounting: Selected Studies. *J. Account. Res.* **1966**, *4*, 71–111. [CrossRef]

17. Altman, E.I. Financial Ratios. Discriminant Analysis and Prediction of Corporate Bankruptcy. *J. Financ.* **1968**, *23*, 589–609. [CrossRef]

18. Springate, G.L.V. *Predicting the Possibiity of Failure in A Canadian Firm*; Unpublished M.B.A. Research Project; Simon Fraser University: Burnaby, BC, Canada, 1978.

19. Ohlson, J.A. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *J. Account. Res.* **1980**, *18*, 109–131. [CrossRef]

20. Taffler, R.J.; Tisshawa, H. The audit going–Concern in Practise. *Account. Mag.* **1984**, *88*, 263–269.

21. Fulmer, J.G.; Moon, J.E.; Gavin, T.A.; Erwin M., J. A Bankruptcy Classification Model for Small Firms. *J. Commer. Bank Lend.* **1984**, *66*, 25–37.

22. Zmijewski, M. Methodological issues related to the estimation of financial distress prediction models. *J. Account. Res.* **1984**, *22*, 59–82. [CrossRef]

23. Horrigan, J.O. The Determination of Long- Term Credit Standing with Financial Ratios. *J. Account. Res.* **1996**, *66*, 44–62. [CrossRef]

24. Lee, S.; Choi, W.S. A multi-industry bankruptcy prediction model using back-propagation neutral network and multivariate discriminant analysis. *Expert Syst. Appl.* **2013**, *40*, 2941–2946. [CrossRef]

25. Alaka, H.A.; Oyedele, L.O.; Owolabi, H.A.; Alaka, H.A.; Oyedele, L.O.; Owolabi, H.A.; Kumar, V.; Ajayi, S.O.; Akinade, O.O.; Bilal, M. Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Syst. Appl.* **2018**, *94*, 164–184. [CrossRef]

26. Peres, C.; Antao, M. The use of multivariate discriminant analysis to predict corporate bankruptcy: A review. *AESTIMATIO IEB Int. J. Financ.* **2016**, *14*, 108–131. [CrossRef]

27. Guo, G.; You, W.; Qian, G.; Shao, W. Parallel maximum likelihood estimator for multiple linear regression model. *J. Comput. Appl. Math.* **2015**, *273*, 251–263. [CrossRef]

28. Jones, S.; Johnstone, D.; Wilson, R. Predicting corporate Bankruptcy: An evaluation of Alternative Statistical Frameworks. *J. Bus. Financ. Account.* **2017**, *44*, 3–34. [CrossRef]

29. Calabrese, R.; Osmetti, S.A. Modelling small and medium enterprise loan defaults as rare events: The generalized extreme value regression model. *J. Appl. Stat.* **2016**, *40*, 1172–1188. [CrossRef]

30. Tudorache, I.C. Bankruptcy prediction models for Romanian restaurants. In Proceedings of the 16th International Conference on Informatics in Economy: Education, Research and Business Technologies, Bucharest, Romania, 4–7 May 2017; Boja, C., Doinea, P., Pocatilu, P., Eds.; Bucharest Univ Economic Studies: Bucharest, Romania, 2017; pp. 519–525.

31. Krusinskas, R.; Lakstutiene, A.; Stankeviciene, J. The research of reliability of bankruptcy prediction models in Lithuanian companies. *Transform. Bus. Econ.* **2014**, *13*, 102–123.

32. Alaminos, D.; del Castillo, A.; Angel, F.M. A global model for bankruptcy prediction. *PLoS ONE* **2016**, *11*. [CrossRef] [PubMed]

33. Ben Javeur, S. Bankruptcy prediction using Partial Least Squares Logistic Regression. *J. Retail. Consum. Serv.* **2017**, *36*, 197–202. [CrossRef]

34. Hwang, R.C.; Chu, C.K. A logistic regression point of view toward loss given default distribution estimation. *Quant. Financ.* **2018**, *18*, 419–435. [CrossRef]

35. Li, M.Y.L.; Miu, P. A hybrid bankruptcy prediction model with dynamic loadings on accounting-ratio-based and market-based information: A binary quantile regression approach. *J. Empir. Financ.* **2010**, *17*, 818–833. [CrossRef]

36. Chrastinová, Z. *Metódy Hodnotenia Ekonomickej Bonity A Predikcie Finančnej Situácie Poľnohospodárskych Podnikov*; Výskumný ústav ekonomiky poľnohospodárstva a potravinárstva: Bratislava, Slovakia, 1998; 34p, ISBN 8080580227.

37. Gurcik, Ľ. *Ekonomika Podnikov*, 1st ed.; Slovenská poľnohospodárska univerzita: Nitra, Slovakia, 2002; 184p, ISBN 80-8067-069-3.

38. Binkert, C.H. Fruherkennung von Unternehmenskrisen mit Hilfe geeigneter Methoden im deutschen und slowakischen Wirtschaftsraum. Ph.D. Thesis, University of Economics in Bratislava, Bratislava, Slovakia, 1999.

39. Zalai, K. Special Aspects of Forecasting the financial development of Slovak companies. *Biatec* **2000**, *8*, 12–14.

40. Kamenikova, K. Obmedzenia použitia modelov predikcie finančného vývoja podniku v podmienkach Slovenskej republiky. *Acta Montan. Slov.* **2005**, *10*, 337–343.

41. Lesakova, L. Evaluating Innovations in Small and Medium enterprises in Slovakia. *Procedia Soc. Behav. Sci.* **2014**, *24*, 74–81. [CrossRef]

42. Horvathova, J.; Mokrisova, M. Determination of Business performance applying modern methods of business performance evaluation. *Econ. Manag. Innov.* **2014**, *6*, 46–60.

43. Gundova, P. Verification of the selected prediction methods in Slovak Companies. *Acta Acad. Karviniensia* **2015**, *4*, 26–39.

44. Adamko, P. Analysis of the modifications of the Global Altman's model 2014. In Proceedings of the 16th International Scientific Conference on Globalization and Its Socio-Economic Consequences, Rajecke Teplice, Slovakia, 5–6 October 2016; Kliestik, T., Ed.; Univ Zilina: Zilina, Slovakia, 2016; pp. 1–6.

45. Boda, M.; Uradnicek, V. On usability of Altman's bankruptcy formula in the Slovak economic conditions. In Proceedings of the 10th International Scientific Conference on Financial Management of Firms and Financial Institutions, Ostrava, Czech Republic, 7–8 September 2015; Culik, M., Ed.; VSB-Tech Univ: Ostrava, Czech Republic, 2015; pp. 64–71.

46. Hiadlovsky, V.; Kral, P. A few notes to business financial health prediction. In Proceedings of the 7th International Scientific Conference on Managing and Modelling of Financial Risks, Ostrava, Czech Republic, 8–9 September 2014; Culik, M., Ed.; VSB TU Ostrava: Ostrava, Czech Republic, 2014; pp. 248–255.

47. Hurtosova, J. Konštrukcia ratingového modelu, nástroja hodnotenia úverovej spôsobilosti podniku. Ph.D. Thesis, University of Economics in Bratislava, Bratislava, Slovakia, 2009.

48. Delina, R.; Packová, M. Validácia predikčných bankrotných modelov v podmienkach SR. *Econ. Manag.* **2013**, *16*, 101–112.

49. Kovacova, M.; Kliestik, T. Logit and probit application for the prediction of bankruptcy in Slovak companies. *Equilib. Q. J. Econ. Econ. Policy* **2017**, *12*, 775–791. [CrossRef]

50. Ravi Kumar, P.; Ravi, V. Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *Eur. J. Oper. Res.* **2007**, *180*, 1–28. [CrossRef]

51. Bartosova, V.; Kral, P. Methodological framework of financial analysis results objectification in Slovak Republic. *J. Mod. Account. Audit.* **2017**, *13*, 394–400.

52. Verbeek, M. *A Guide to Modern Econometrics*, 2nd ed.; John Wiley & Sons: Chicago, IL, USA, 2008; p. 447. ISBN 0-470-85773-0.

53. Rimarčík, M. *Štatistika Pre Prax*; Vydavateľstvo Marián Rimarčík: Bratislava, Slovakia, 2007; 200p, ISBN 8096981311.

54. Cipra, T. *Finanční Ekonometrie*; Ekopress: Praha, Czech Republic, 2008; 538p, ISBN 9788086929439.

55. Meloun, M.; Militky, J. *Kompendium Statistického Zpracování Dát*; Karolinum: Praha, Czech Republic, 2013; 984p, ISBN 9788024621968.

56. Klepáč, V.; Hampel, D. Prediction of bankruptcy with SVM classifiers among retail business companies in EU. *Acta Univ. Agric. Silvic. Mendel. Brun.* **2016**, *64*, 627–634. [CrossRef]

57. Durica, M.; Svabova, L.; Valaskova, K. Comparison of Slovak company failure prediction models created using discriminant analysis and logistic regression. In Proceedings of the 5th International Scientific Conference on New Trends in Management and Production Engineering—Regional, Cross-Border and Global Perspectives, Brenna, Poland, 7–8 June 2018. in press.

58. Svabova, L.; Durica, M.; Podhorska, I. Prediction of default of small companies in the Slovak Republic. *Econ. Cult.* **2018**, *14*, in press.

59. Rohacova, V.; Kral, P. Corporate failure prediction using DEA: An application to companies the Slovak Republic. In Proceedings of the 18th International Scientific Conference on Applications of Mathematics and Statistics in Economics, Hradec Kralové, Czech Republic, 2–6 September 2016; Financ: Banska Bystrica, 2016.

60. Altman, E.I.; Iwanicz-Drozdowska, M.; Laitinen, E.K.; Suvas, A. Distresses firm and bankruptcy prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Soc. Sci. Electron. Publ.* **2014**. [CrossRef]

61. Adamko, P.; Svabova, L. Prediction of the risk of bankruptcy of Slovak companies. In Proceedings of the 8th International Scientific Conference on Managing and Modelling of Financial Risks, Ostrava, Czech Republic, 5–6 September 2016; VSB-Technical University: Ostrava, Czech republic, 2016; pp. 15–20.

62. Gupta, J.; Gregoriou, A.; Healey, J. Forecasting bankruptcy for SMEs using hazard function: To what extent does size matter? *Rev. Quant. Financ. Account.* **2015**, *45*, 845–869. [CrossRef]

63. Faltus, S. Firm Default Prediction Model for Slovak Companies. In Proceedings of the 11th International Conference on European Financial Systems, Lednice, Czech Republic, 12–13 June 2014; Deev, O., Kajurova, V., Krajicek, J., Eds.; Masarykova Univ: Brno, Czech Republic, 2014; pp. 173–177.

64. Sharifabadi, M.R.; Mirhaj, M.; Izadinia, N. The impact of financial ratios on the prediction of bankruptcy of small and medium companies. *Quid Investig. Cienca y Tecnol.* **2017**, 164–173.

65. Tian, S.; Yu, Y.; Guo, H. Variable selection and corporate bankruptcy forecasts. *J. Bank. Financ.* **2015**, *52*, 89–100. [CrossRef]

66. Bellovary, J.; Giacomino, D.E.; Akers, M.D. A review of bankruptcy prediction Studies: 1930 to Present. *J. Financ. Educ.* **2007**, *33*, 1–42.

67. Dimitras, A.I.; Zanakis, S.H.; Zopounidis, C. A survey of business failure with an emphasis on prediction method and industrial applications. *Eur. J. Oper. Res.* **1996**, *90*, 487–513. [CrossRef]

68. Kliestik, T.; Misankova, M.; Valaskova, K.; Svabova, L. Bankruptcy Prevention: New Effort to Reflect on Legal and Social Changes. *Sci. Eng. Ethics.* **2018**, *24*, 791–803. [CrossRef] [PubMed]

69. Kliestik, T.; Svabova, L. Some remarks on the regional disparities of prediction models constructed in the Visegrad countries. In Proceedings of the 8th International Conference the Economies of Balkan and Eastern Europe Countries in the Changing World, Split, Croatia, 6–8 May 2016; Karasavvoglou, A., Goic, S., Polychronidou, P., Eds.; Univ Split: Split, Croatia, 2016; pp. 89–96.

70. Mihalovic, M. Performance comparison of multiple discriminant analysis and logit models in bankruptcy prediction. *Econ. Soc.* **2016**, *9*, 101–118. [CrossRef]

71. Mousavi, M.M.; Ouenniche, J.; Xu, B. Performance evaluation of bankruptcy prediction models: An orientation-free super-efficiency DEA-based framework. *Int. Rev. Financ. Anal.* **2015**, *42*, 64–75. [CrossRef]

Article

Market Timing with Moving Averages

Jukka Ilomäki [1], Hannu Laurila [1,*] and Michael McAleer [2,3,4,5,6]

[1]   Faculty of Management, University of Tampere, FI-33014 Tampere, Finland; jukka.ilomaki@uta.fi
[2]   Department of Finance, Asia University, Taichung City 413, Taiwan; michaelmcaleer@gmail.com
[3]   Discipline of Business Analytics University of Sydney Business School, Sydney NSW 2006, Australia
[4]   Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam,
      3062 PA Rotterdam, The Netherlands
[5]   Department of Economic Analysis and ICAE, Complutense University of Madrid, 28040 Madrid, Spain
[6]   Institute of Advanced Sciences Yokohama National University, Yokohama 240-8501, Japan
*   Correspondence: hannu.laurila@uta.fi; Tel.: +358-50-318-5998

Abstract: Consider using the simple moving average (MA) rule of Gartley to determine when to buy stocks, and when to sell them and switch to the risk-free rate. In comparison, how might the performance be affected if the frequency is changed to the use of MA calculations? The empirical results show that, on average, the lower is the frequency, the higher are average daily returns, even though the volatility is virtually unchanged when the frequency is lower. The volatility from the highest to the lowest frequency is about 30% lower as compared with the buy-and-hold strategy volatility, but the average returns approach the buy-and-hold returns when frequency is lower. The 30% reduction in volatility appears if we invest randomly half the time in stock markets and half in the risk-free rate.

1. Introduction

According to the standard investing separation theorem of Tobin [1], investors allocate investments between risk-free and risky assets. If the risk-free rate is low (high), the investors shift their wealth to (from) the risky assets. Fama [2] divided forecasters into two categories, namely macro forecasters (or market timers) and micro forecasters (or security analysts), who try to forecast individual stock returns relative to the market returns.

Merton [3] defined a market timer to forecast when stocks will outperform (underperform) the risk-free asset, indicating that, when $r_t^m > r_t^f$ ($r_t^m < r_t^f$), where r_t^m is average stock market returns, r_t^f is the risk-free asset, $r_t^i = r_t^f + \beta^i(r_t^m - r_t^f) + \varepsilon_t^i$, r_t^i is the return for individual stock i included in the market portfolio m, β^i is a positive parameter, and $E[\varepsilon_t^i|r_t^m] = E[\varepsilon_t^i]$. That is, a market timer only forecasts the statistical properties of $r_t^m - r_t^f$, indicating that their forecasts contain only the differential performance among individual stocks arising from systematic risk in the markets.

Merton [3] showed theoretically that, when investors have heterogeneous beliefs and imperfect information, the value of a random market timing forecast is zero, and if the forecast variable is distributed independently or the forecast is based on public information, its value is zero, too. In fact, Merton showed that the maximum value of skilled market timing is the value of the protective put against buy-and-hold strategy.

Henriksson and Merton [4] presented an empirical procedure whereby correct forecasts can be analyzed statistically. However, if it is assumed that ε_t^i follows an approximate normal distribution, this leads to the Capital Asset Pricing Model (CAPM) of Sharpe [5], and Lintner [6].

The purpose of the paper is to detect whether the frequency used in calculating the MA affects the performance of the trading rule. We use a large sample with more than eight million observations for robustness of the empirical results, and a simple MA rule for the timing aspect for individual Dow Jones Industrial Average (DJIA) stocks with different frequencies. We use a simple MA rule for the timing aspect for individual Dow Jones Industrial Average (DJIA) stocks with different frequencies. Zhu and Zhou [7] showed analytically that MA trading rules, as a part of asset allocation rules, can outperform standard allocation rules when stock returns are partly forecastable. The standard rule means investing a fixed proportion of wealth in risky assets and the rest in risk-free assets, with the ratio determined by the risk tolerance of an investor. It is well known that MA is a widely used technical trading rule, which adds value for a risk averse investor if returns are predictable.

This is the well-known reward/risk (or mean-variance) principle in the spirit of Markowitz [8], Tobin [1], and Sharpe [5]. Zhu and Zhou [7] argued that the fixed allocation rule is not optimal if returns are forecastable by using the MA rule. Therefore, assuming that risk tolerance and the forecast performance of stock market returns are constant, the linear combination rule means that, when the MA rule suggests an uptrend (downtrend), the rule suggests that the total weight should be allocated to stock markets (the risk-free rate).

The empirical findings suggest a low volatility anomaly that might be explained by investors' affection to high volatility, as suggested by Baker et al. [9], and noted in Ang et al. [10]. On the other hand, the reported predictability of risk premia (see, for example, Cochrane [11], and Fama [12]) can explain why, for instance, MA rules forecast better than using random highs and lows in the stock market (as noted in Jagannathan and Korajczyck [13]). The topic is important, as Friesen and Sapp [14], among others, reported that mutual fund investors had negative outcomes, on average, in their timing to invest and withdraw cash from US mutual funds from 1991 to 2004. Munoz and Vicente [15] reported similar results with more recent data in US markets.

The remainder of the paper is organized as follows. Section 2 provides a literature review, and alternative model specifications are presented in Section 3. The empirical analysis is conducted in Section 4, while Section 5 gives some concluding comments.

2. Literature Review

In efficient markets, investors earn above average returns only by taking above average risks (Malkiel [16]). Samuelson [17] conformed with Fama [2] by noting that market efficiency can be divided into micro and macro efficiency. The former concerns the relative pricing of individual stocks, and the latter, for markets as a whole. The CAPM by Sharpe [5], and Lintner [6] argues that beta is a proper definition for systematic risk for stock *i*, if unexplained changes in risk adjusted returns for the stock follow approximately normal distribution with zero mean.

Black [18] stated that the slope of the security market line (SML) is flatter if there exist restrictions in borrowing, that is, leverage constraints in the model. Starting from Black et al. [19], many studies have reported that the security market line is too flat in US stocks compared with the SML suggested by the CAPM version of Sharpe and Lintner.

Ang et al. [10], Baker et al. [20], and Frazzini and Pedersen [21] found that low-beta stocks outperform high-beta stocks statistically significantly. In fact, Frazzini and Pedersen reported that significant excess profits in US stocks can be achieved by shorting high-beta stocks and buying low-beta stocks with leverage, but that leverage constraints make them disappear. Using Black [18], investors often have leverage constraints, thereby making them place too much weight on risky stocks, which results in lower required return for high-beta stocks than would be justified by the Sharpe–Lintner CAPM.

Markowitz [8] defined portfolio risk simply as the volatility of portfolio returns. Clarke et al. [22] found that the volatility of stock returns contains potentially an additional risk factor with respect to systematic risk that can be defined in the betas of CAPM by Sharpe and Lintner. Moreover,

Ang et al. [10] reported that the total volatility of international stock market returns is highly correlated with US stock returns, thereby suggesting a common risk factor for US stocks.

Baker et al. [9] suggested that the low-volatility anomaly is due to investor irrational behavior, mainly because an average fund manager seeks to beat the buy-and hold strategy by overinvesting in high-beta stocks. The explanations include preference for lotteries (Barberis and Huang [23]; Kumar [24]; Bali et al. [25]), overconfidence (Ben-David et al. [26]), and representativeness (Daniel and Titman [27]), which means that people assess the probability of a state of the world based on how typical of that state the evidence seems to be (Kahneman and Tversky [28]).

Baker et al. [9] argued that the anomaly is also related to the limits of arbitrage (see also Baker and Wurgler [29]). In fact, the extra costs of shorting prevent taking advantage of overpricing (Hong and Sraer [30]). More importantly, Li et al. [31] reported that the excess returns of low-beta portfolios are due to mispricing in US stocks, indicating that the low-volatility anomaly does not exist because of systematic risk by some rational, stock specific volatility risk factor. They tested the low-volatility anomaly with monthly data from January 1963 to December 2011 in NYSE, NASDAQ, and AMEX stocks.

Market timing is closely related to technical trading rules. Brown and Jennings [32] showed theoretically that using past prices (e.g., the MA rule of Gartley [33]) has value for investors, if equilibrium prices are not fully revealing, and signals from past prices have some forecasting qualities. More importantly, Zhu and Zhou [7] indicated that the MA rules are particularly useful for asset allocation purposes among risk averse investors, when markets are forecastable (quality of signal).

Moskowitz et al. [34] argued that there are significant time series momentum (TSM) effects in financial markets that are not related to the cross-sectional momentum effect (Jegadeesh and Titman [35]). However, TSM is closely related to MA rules, since it gives a buy (sell) signal according to some historical price reference points, whereas MA rules give a buy (sell) signal, when the current price moves above (below) the historical average of the chosen calculated rolling window measure.

Starting from LeRoy [36] and Lucas [37], the literature in financial economics states that financial markets returns in efficient markets are partly forecastable, when investors are risk averse. This leads to the time-varying risk premia of investors, as noted by Fama [12]. For example, Campbell and Cochrane [38] presented a consumption-based model, which indicates that when the markets are in recession (boom), risk averse investors require larger (smaller) risk premium for risky assets. More importantly, Cochrane [11] noted that the forecastability of excess returns may lead to successful market timing rules.

Brock et al. [39] tested different MA lag rules for US stock markets, and found that they gain profits compared with holding cash. On the other hand, Sullivan et al. [40] found that MA rules do not outperform the buy-and-hold strategy, if transaction costs are accounted for. Allen and Karjalainen [41] used a genetic algorithm to develop the best ex-ante technical trading rule model using US data, and found some evidence of outperforming the buy-and-hold strategy. Lo et al. [42] found that risk averse investors benefit from technical trading rules because they reduce volatility of the portfolio without giving up much returns when compared against the buy-and-hold strategy.

More recently, Neely et al. [43] used monthly data from January 1951 to December 2011, and reported that MA rules forecast the risk premia in US stock markets statistically significantly. Marshall et al. [44] found that MA rules give an earlier signal than TSM, suggesting better returns for MA rules, but they both work best with outside of large market value stocks.

Moskowitz et al. [34] used monthly data from January 1965 to December 2009, and reported that TSM provides significant positive excess returns in futures markets. However, Kim et al. [45] reported that these positive excess returns produced by TSM are due to the volatility scaling factor used by Moskowitz et al.

3. Model Specification

Consider an overlapping generation economy with a continuum of young and old investors $[0, 1]$. A young risk-averse investor j invests their initial wealth, w_t^j, in infinitely lived risky assets $i = 1, 2, \ldots I$, and in risk-free assets that produce the risk-free rate of return, r^f. A risky asset i pays dividend D_t^i, and has x_i^s outstanding. Assuming exogenous processes throughout, the aggregate dividend is D_t.

A young investor j maximizes their utility from old time consumption through optimal allocation of initial resources w_t^j, between risky and risk-free assets:

$$\max x_t^j \left(\frac{E_t(P_{t+1} + D_{t+1})}{P_t} - (1 + r^f) \right) - \frac{v^j}{2} x^{j^2} \sigma^2$$
$$s.t.$$
$$x_t^j P_t \leq w_t^j$$

where E_t is the expectations operator, P_t is the price of one share of aggregate stock, v^j is a constant risk-aversion parameter for investor j, σ^2 is the variance of returns for the aggregate stock, and x_t^j is the demand of risky assets for an investor j. The first-order condition is:

$$\frac{E_t(P_{t+1} + D_{t+1})}{P_t} - (1 + r^f) - v^j x_t^j \sigma^2 = 0,$$

which results in optimal demand for the risky assets:

$$x_t^j = \frac{E_t((P_{t+1} + D_{t+1})/P_t) - (1 + r^f)}{v^j \sigma^2} \tag{1}$$

Suppose that an investor j is a macro forecaster who allocates their initial wealth, w_t^j, between risky stocks and risk-free assets according to their forecast about the return of the risky alternative. Then, Equation (1) says that the investor invests in the risky stocks only if the numerator on the right hand side is positive.

4. Empirical Analysis

This section presents the empirical results from seven frequencies for the (MA) trend-chasing rules. The data consist of 29 companies included in the Dow Jones Industrial Average (DJIA) index in January 2018. The trading data (daily closing prices) cover 30 years from 1 January 1988 to 31 December 2017. Choosing the current DJIA companies for the last 30 years creates a "survivor bias" in the buy-and-hold results. However, this should not be an issue, as we intend to compare the performance of the alternative MA frequency rules.

The rolling window is 200 trading days. The first rule is to calculate MA in every trading day; the second frequency takes into account every 5th trading day (thereby providing a proxy for the weekly rule); the third frequency takes into account every 22th trading day (proxy for the monthly rule); the fourth rule is to calculate MA for every 44th trading day (proxy for every other month); the fifth rule takes into account every 66th trading day (proxy for every third month); the sixth rule takes into account every 88th trading day (proxy for every fourth month); and the seventh rule takes into account every 100th trading day (proxy for every fifth month).

For the 29 DJIA companies, 26 of them have daily stock data available from 27 March 1987, thereby giving 4 January 1988 as the first trading day. The data for Cisco are available from 12 February 1990, for Goldman Sachs from 4 May 1999, and for Visa from 19 March 2008. There are 217,569 observations of daily returns from DJIA stocks. Thus, there are 217,569 × 9 = 1,958,121 daily returns for the first three frequencies (rules), 217,569 × 4 = 870,276 daily returns for the fourth rule, 217,569 × 3 = 652,707 daily returns for the fifth rule, 217,569 × 2 = 435,138 daily returns for the sixth rule, and 217,569 daily returns for the seventh rule.

The trading rule for all cases is to use a simple crossover rule. When the trend-chasing MA turns lower (higher) than the current daily closing price, we invest the stock (three-month US Treasury Bills) at the closing price of the next trading day. Thus, the trading rule provides a market timing strategy where we invest all wealth either in stocks (separately, every stock included in DJIA), or to the risk-free asset (three-month U.S. Treasury bill), where the moving average rule advices the timing.

At the first frequency (every trading day), we calculate daily returns for MA200, MA180, MA160, MA140, MA120, MA100, MA80, MA60, and MA40. For example, MA200 is calculated as:

$$\left(\frac{P_{t-1} + P_{t-2} + \ldots + P_{t-200}}{200} \right) = X_{t-1}$$

At the lowest frequency, where every 100th daily observation is counted, MAC2 is calculated as:

$$\left(\frac{P_{t-1} + P_{t-100}}{2} \right) = X_{t-1}$$

If $X_{t-1} < P_{t-1}$, we buy the stock at the closing price, P_t, thereby giving daily returns as

$$R_{t+1} = \ln \left(\frac{P_{t+1}}{P_t} \right)$$

Tables 4–7 and A1–A3 in Appendix A show that the annualized average log returns of MA200−MA40 are **+0.053** *after* transaction costs (with 0.1% per change of position). Recall that there are 200 closing day prices in the rolling window MA200, whereas MA40 means that there are 44 closing day prices in the window. The respective log returns for MAW40−MAW8 (weekly) are **+0.063**; for MA10−MA2 (monthly) **+0.071**; for MAD5−MAD2 (every other month) **+0.078**; for MAT4−MAT2 (every third month) **+0.084**; for MAQ3−MAQ2 (every fourth month) **+0.094**; and for MAC2 (every fifth month) **+0.088** after transaction costs.

Tables 4–7 and A1–A3 show that, as the frequency decreases until every fourth month frequency (MAQ3−MAQ2), average returns tend to increase, and decrease thereafter. In comparison, the biased buy-and-hold strategy produces **+0.117** with equal weights among all DJIA stocks, and with **0.295** annual volatility. A random investment (half the time in the risk-free rate, and half in the equally weighted portfolio from 4 January 1988) produces $(0.117 \times 0.5 + 0.022 \times 0.5) =$ **+0.070** annually, on average, with $(1 - \sqrt{0.5} = 0.293) = 29.3\%$ reduction in volatility, indicating **0.209** annual volatility for that portfolio.

The data are dividend excluded, but the average annual dividend yield in DJIA stocks over the last thirty years has been +0.026, so that the biased buy and hold strategy produces +0.143 annually with equal weights among DJIA stocks before taxes. Thus, the random investment strategy produces +0.083 annually, with survivor bias.

Appendix A (namely the second column of Tables 4–7 and A1–A3) also reports the annualized average log returns calculated in the largest sample (full 200 observations) in every category: MA200 **+0.065**; MAW40 **+0.073**; MA10 **+0.079**; MAD5 **+0.083**; MAT4 **+0.089**; MAQ3 **+0.091**; and MAC2 **+0.088** after transaction costs and before dividends. Adding +0.013 produces after dividends and before taxes: MA200 **+0.078**; MAW40 **+0.086**; MA10 **+0.092**; MAD5 **+0.096**; MAT4 **+0.102**; MAQ3 **+0.104;** and MAC2 **+0.101**. These results imply that starting from every fifth trading day frequency, a macro forecaster beats the buy and hold strategy **in returns**.

Figure 1 illustrates the effects of frequency on the returns to volatility ratio (the second column in Appendix A, Tables 4–7 and A1–A3).
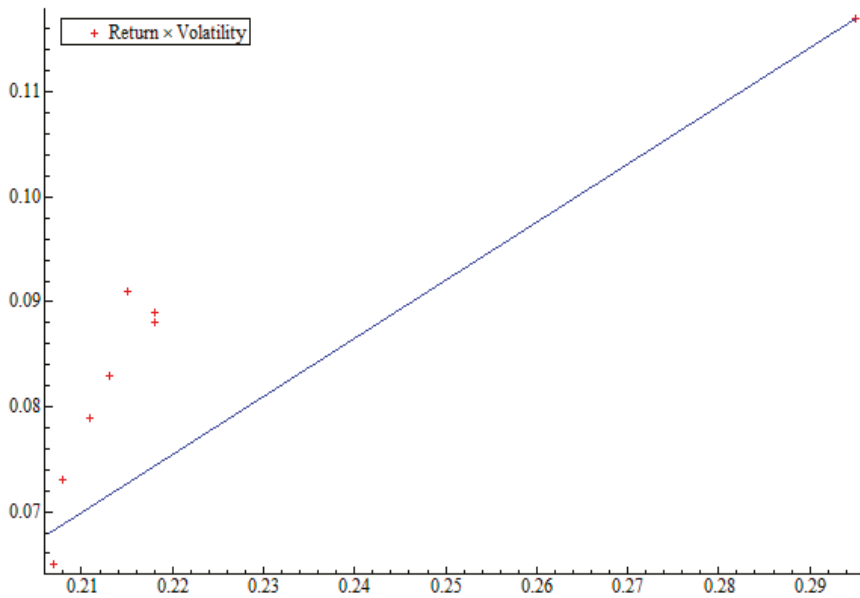
Figure 1. Returns to volatility ratio in MA200, MAW40, MA10, MAD5, MAT4, MAQ3, MAC2, and the theoretical random timing efficient SML.

In Figure 1, the straight line illustrates the return to volatility ratio of portfolios, where wealth is randomly invested in combinations of the three-month Treasury Bill (risk-free rate), with stocks included in the DJIA between 4 January 1988 and 31 December 2017. The red crosses represent the average return/volatility points calculated in the 200-day rolling window with the following frequencies: daily, every five days, every 22 days, every 44 days, every 66 days, every 88 days, and every 100 days (with only the most observations in each frequency giving 200, 40, 10, 5, 4, 3, and 2 observations). The red crosses plot a convex curve that deviates increasingly from the straight return to volatility ratio line, thereby symbolizing superior portfolio efficiency.

Tables 8–14 in Appendix B show that the annualized volatility of daily returns read, on average: MA200−MA40 **0.2044**; MAW40−MAW8 **0.205**; MA10−MA2 **0.2091**; MAD5−MAD2 **0.213**; MAT4−MAT2 **0.219**; MAQ3−MAQ2 **0.221**; and MAC2 **0.218**. Thus, there is virtually no difference between the MA frequencies, while the biased buy-and-hold strategy produces **0.295**.

Figure 1 presents the volatilities calculated in the largest sample (full 200 day rolling window in every category, the second column in Tables 8–14). They read MA200 **0.207**; MAW40 **0.208**; MA10 **0.211; MAD5 0.213**; MAT4 **0.218**; MAQ3 **0.215**; and MAC2 **0.218** after transaction costs. Investing randomly half of the time in the risk-free rate and the other half in the equally weighted portfolio, produces **0.209**. Thus, the difference between the annual volatilities produced in profitable market timing MA rules (MA10−MAC2) and random market timing (half and half) ranges from **0.009** to **0.002**.

In Figure 2, the straight line again presents the return to volatility ratio of portfolios with random investment in the risk-free rate and the stocks in DJIA between 4 January 1988 and 31 December 2017. The red crosses plot the average return to volatility ratios, calculated by using a 200-day rolling window, with the following frequencies: daily, every five days, every 22 days, every 44 days, every 66 days, every 88 days, and every 100 days. The averages of every lag are reported in Tables 4–14 and A1–A3, and. Thus, all daily returns from Tables 4–14 and A1–A3 are included.
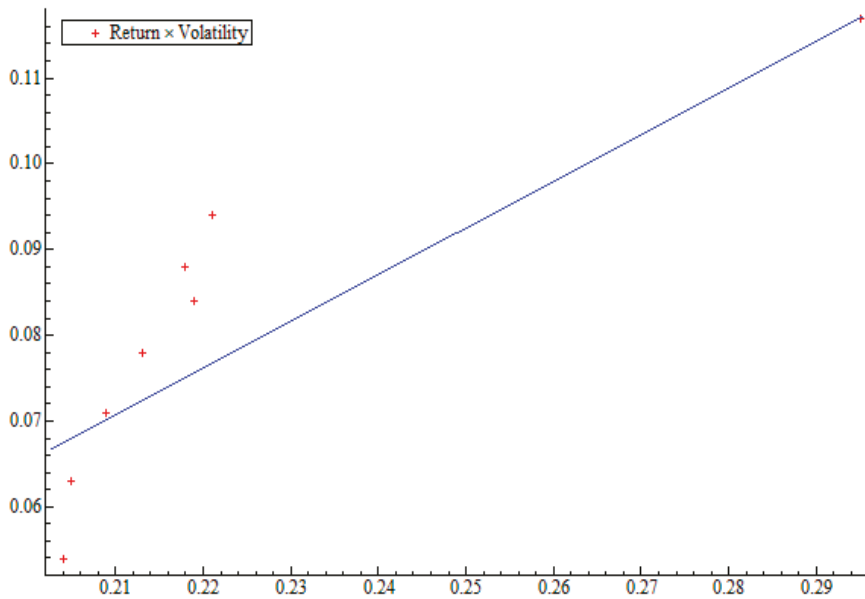
Figure 2. Returns to volatility ratio in MA200 − MA40, MAW40 − MAW8, MA10 − MA2, MAD5 − MAD2, MAT4 − MAT2, MAQ3 − MAQ2, MAC2, and the theoretical random timing efficient SML.

Comparing Figures 1 and 2, it is clear that using the whole 200 daily observation windows in the MA rules produces more efficient results in market timing. That is, comparing the products of shorter and longer MA rule rolling windows, e.g., the last two monthly observations compared with ten monthly observations, average realized returns drop from **+0.079** to **+0.059** before dividends, while volatility remains approximately unchanged (from 0.211 to 0.207). This suggests that, in both cases, about half and half is invested in the equally-weighted DJIA portfolios and in the risk-free rate, and the MA rules advise the timing. More importantly, Tables 8–14 in Appendix B show that the range in volatilities with all MA rules varies between 0.202 and 0.227 (with 0.02 difference), whereas Tables 4–7 and A1–A3 in Appendix A show that realized returns vary between 0.096 and 0.033 before dividends (with 0.063 difference).

These results indicate that a macro market timing with 200 days rolling window produces a reduction in volatility from **0.295** (the buy-and hold) to between 0.207 and 0.218, but the average annualized returns (dividends included) tend to rise as the MA frequency falls (+0.078 with all 200 observations to +0.104 with every fourth month observations). Thus, the results indicate that MA market timing finds long term stochastic trends more efficiently than short term stochastic trends.

The Sharpe ratio of random market timing (half and half) with dividends is **0.292**; for MA200 **0.271**; for MAW40 **0.308**; for MA10 **0.332**; for the MAD5 **0.347**; for MAT4 **0.370**; for MAQ3 **0.381**; and for MAC2 it is **0.362**.

Figure 3 shows that when the volatility changes 1% in the DJIA stocks, then the average returns change is 0.39%. Figures 1 and 2 suggest that the theoretical change should be such that, when the volatility changes 1%, the average returns change is 0.50%, suggesting a flatter SML line in the data. This suggests strongly that DJIA investors have overweight high-beta stocks in the last 30 years.
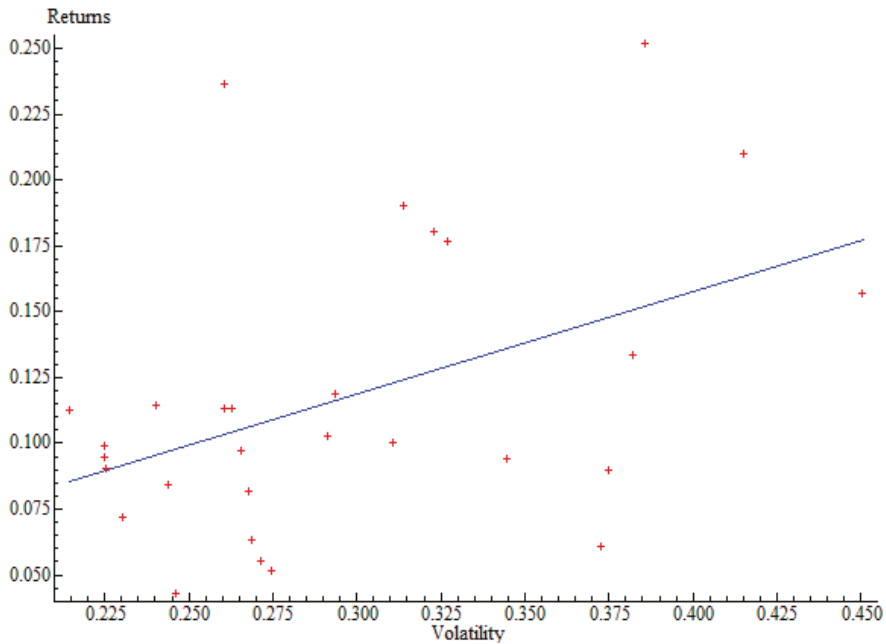
Figure 3. Returns to volatility ratio in current DJIA stocks, annual averages from 4 January 1988 to 31 December 2017.

It is obvious that transaction costs are crucial in MA performance. In the above calculations, the transaction costs are 0.1% per transaction from current wealth. Tables 15 and 16 in Appendix C report the transaction costs for the MA200−MA40 and MA10−MA2 rules. In the MA200−MA40 rules, the average annualized transaction costs are **0.0133**, such that the rules have about 13 changes in positions per year. Meanwhile, for the MA10−MA2 rules, the average annualized transaction costs are **0.0032,** suggesting about three changes in positions per year.

Allen and Karjalainen [41] gave reasons for using a cost of 0.2% per transaction in their sample, but since technological progress has reduced transaction costs since the mid-1990s, 0.1% per transaction should be fair, on average. Nevertheless, a trial with 0.2% transaction costs shows that, for example, the average annualized daily returns become 0.0403 for the MA200−MA40 rules, and 0.0674 for the MA10−MA2 rules. Note that the returns grow 67%, on average, for the MA10−MA2 rules (with about the same volatility) compared with costs of 0.1% per transaction.

Note that the model prohibits short selling since we only have long positions in stocks or investing in the risk-free rate. Then, the limits of arbitrage argument of Baker et al. [9] are consistent with our results.

5. Concluding Remarks

The analysis suggests that a macro forecaster can obtain higher returns with equal volatility (30% below that of the buy-and-hold strategy) by reducing the frequency used in MA rules. The return to volatility ratio for risk-averse investors with MA market timing significantly outperforms the random benchmark strategy, when the frequency in the MA rules is reduced. This indicates that the forecasts become more accurate as the time frame becomes longer.

The results suggest that a flatter SML in the CAPM can be followed by the irrational preference of investors in high-beta stocks, as suggested by Baker et al. (2011) and Li et al. (2016), since the

empirically efficient frontier of portfolios becomes flatter than the theoretically efficient SML (random timing) (see Figure 1). In other words, the empirical results suggests\that market timing with the few past observations (for example, every fourth month) in the past 200 rolling window daily prices, have produced significantly better returns to risk ratio for the portfolio of DJIA equally weighted stocks in the past 30 years than random timing. The finding points to the low-volatility anomaly.

One explanation for the results is that they are due to time-varying risk premiums. This is emphasized by Neely et al. (2014), who claimed that MA rules, in effect, forecast changes in the risk premium. If the results are rational products of time-varying risk premiums, the results suggest that investor sensitivity to risk must be extremely high, and their risk premium is larger (smaller) in downs (ups), as suggested by Campbell and Cochrane (1999). As volatility rises (decreases), usually in downs (ups), the results suggest that, when volatility is high, investors as a group tolerate significantly more risk (that is, volatility) than in calmer periods.

Consider the following numerical example: Assume that the risk premium is 0.08 in volatile downs, and 0.04 in calm ups, and the variance of returns is 0.09 in downs and 0.03 in ups. Then, the risk aversion coefficient must be 0.89 in volatile down periods, and 1.33 in calm up periods. As market timing with MA rules works better in longer periods with few observations, it seems to be more accurate in longer stochastic (up or down) trends.

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

Appendix A

Table A1. Annualized daily returns of MA40–MA200, average annualized returns.

	Buy and Hold	MA200	MA180	MA160	MA140	MA120	MA100	MA80	MA60	MA40
3M	0.090	0.042	0.034	0.017	0.015	0.019	0.014	0.006	−0.009	6×10^{-4}
American Express	0.094	0.035	0.037	0.039	0.055	0.039	0.042	0.043	0.041	0.008
Apple	0.157	0.147	0.145	0.147	0.142	0.156	0.149	0.150	0.146	0.164
Boeing	0.119	0.088	0.089	0.060	0.055	0.061	0.061	0.058	0.046	0.048
Caterpillar	0.100	0.075	0.079	0.058	0.058	0.049	0.034	0.028	0.039	0.025
Chevron	0.084	0.005	0.013	0.002	−0.000	−0.000	0.003	−0.01	−0.025	−0.05
Coca-Cola	0.099	0.058	0.055	0.030	0.035	0.039	0.027	0.023	0.009	0.003
Walt Disney	0.103	0.072	0.078	0.079	0.074	0.077	0.074	0.076	0.056	0.048
Exxon	0.072	−0.011	−0.010	−0.020	−0.030	−0.020	−0.025	−0.01	−0.044	−0.05
GE	0.052	0.072	0.071	0.058	0.039	0.039	0.033	0.018	0.013	9×10^{-4}
Home Depot	0.190	0.125	0.116	0.102	0.092	0.087	0.076	0.067	0.068	0.058
IBM	0.055	0.016	0.029	0.033	0.028	0.016	0.021	0.031	0.029	0.048
Intel	0.134	0.083	0.082	0.083	0.073	0.091	0.082	0.080	0.077	0.078
Johnson & Johnson	0.113	0.062	0.058	0.053	0.042	0.032	0.044	0.028	0.008	−0.00
JP Morgan	0.090	0.013	0.014	0.003	0.010	0.017	0.013	0.031	0.038	0.025
McDonalds	0.114	0.047	0.048	0.040	0.044	0.040	0.035	0.043	0.030	0.018
Merck	0.063	0.050	0.048	0.044	0.032	0.033	0.029	0.022	0.016	−0.02
Microsoft	0.180	0.117	0.128	0.105	0.102	0.104	0.095	0.090	0.070	0.062
Nike	0.177	0.087	0.093	0.085	0.102	0.108	0.107	0.119	0.133	0.112
Pfizer	0.097	0.059	0.056	0.043	0.042	0.052	0.044	0.040	0.024	0.009
Procter & Gamble	0.095	0.037	0.045	0.037	0.036	0.037	0.029	0.023	0.004	0.017
Travellers	0.082	0.036	0.035	0.038	0.029	0.008	−0.004	-9×10^{-4}	−0.001	0.006
United Technologies	0.113	0.051	0.057	0.046	0.059	0.057	0.049	0.049	0.041	0.017
United Health Group	0.252	0.181	0.182	0.157	0.147	0.136	0.130	0.118	0.125	0.076
Verizon	0.043	−0.017	−0.020	−0.010	−0.000	−0.020	−0.020	−0.02	−0.029	−0.02
Wal-Mart	0.113	0.019	0.016	0.010	0.012	0.012	0.016	0.012	0.020	0.024
Cisco	0.210	0.198	0.194	0.210	0.208	0.198	0.205	0.152	0.096	0.085
Goldman Sachs	0.061	0.038	0.029	0.033	0.038	0.050	0.057	0.078	0.076	0.063
Visa	0.236	0.112	0.118	0.129	0.141	0.128	0.132	0.120	0.094	0.085
Average	**0.117**	**0.065**	**0.066**	**0.059**	**0.058**	**0.057**	**0.053**	**0.05**	**0.041**	**0.033**

0.054

Table A2. Annualized daily (every fifth trading day) returns of MAW8–MAW40 (W = number of weeks), average annualized returns.

	Buy and Hold	MAW40	MAW36	MAW32	MAW28	MAW24	MAW20	MAW16	MAW12	MAW8
3M	0.090	0.035	0.033	0.020	0.021	0.019	0.012	0.019	0.032	0.026
American Express	0.094	0.058	0.053	0.062	0.063	0.047	0.046	0.035	0.034	0.015
Apple	0.157	0.130	0.137	0.143	0.131	0.134	0.131	0.188	0.174	0.144
Boeing	0.119	0.089	0.079	0.075	0.074	0.080	0.082	0.066	0.074	0.076
Caterpillar	0.100	0.057	0.062	0.058	0.058	0.061	0.054	0.049	0.043	0.023
Chevron	0.084	0.005	0.015	3×10^{-4}	0.004	0.008	0.009	0.004	0.004	−0.03
Coca-Cola	0.099	0.055	0.054	0.054	0.041	0.054	0.047	0.047	0.029	0.011
Walt Disney	0.103	0.071	0.073	0.062	0.080	0.076	0.080	0.078	0.065	0.051
Exxon	0.072	0.018	0.016	0.007	0.008	0.010	0.013	0.020	0.011	0.005
GE	0.052	0.061	0.046	0.047	0.047	0.045	0.023	0.018	0.031	0.023
Home Depot	0.190	0.135	0.133	0.124	0.112	0.110	0.088	0.076	0.096	0.077
IBM	0.055	0.020	0.037	0.044	0.040	0.051	0.027	0.028	0.008	0.016
Intel	0.134	0.088	0.091	0.075	0.061	0.075	0.073	0.070	0.076	0.085
Johnson & Johnson	0.113	0.074	0.079	0.071	0.059	0.050	0.050	0.048	0.042	0.027
JP Morgan	0.090	0.040	0.036	0.027	0.033	0.033	0.048	0.051	0.042	0.020
McDonalds	0.114	0.086	0.068	0.059	0.058	0.052	0.052	0.059	0.058	0.044
Merck	0.063	0.051	0.039	0.029	0.034	0.034	0.030	0.033	0.024	0.029
Microsoft	0.180	0.128	0.125	0.116	0.116	0.116	0.105	0.099	0.062	0.078
Nike	0.177	0.087	0.091	0.098	0.093	0.087	0.094	0.102	0.119	0.091
Pfizer	0.097	0.070	0.061	0.057	0.053	0.063	0.049	0.050	0.044	0.050
Procter & Gamble	0.095	0.050	0.044	0.050	0.051	0.040	0.043	0.042	0.031	0.033
Travellers	0.082	0.020	0.006	0.010	0.014	0.006	0.005	0.008	0.017	0.015
United Technologies	0.113	0.071	0.077	0.062	0.072	0.071	0.056	0.061	0.051	0.053
United Health Group	0.252	0.171	0.133	0.130	0.151	0.124	0.134	0.123	0.113	0.087
Verizon	0.043	−0.00	−0.01	0.002	0.006	−0.01	−0.01	−0.01	−0.009	−0.00
Wal-Mart	0.113	0.050	0.049	0.045	0.038	0.028	0.033	0.026	0.038	0.029
Cisco	0.210	0.209	0.211	0.219	0.222	0.219	0.204	0.164	0.120	0.094
Goldman Sachs	0.061	0.050	0.030	0.031	0.040	0.036	0.071	0.089	0.078	0.077
Visa	0.236	0.143	0.142	0.131	0.171	0.167	0.159	0.113	0.119	0.080
Average	**0.117**	**0.073**	**0.069**	**0.066**	**0.067**	**0.065**	**0.062**	**0.061**	**0.056**	**0.046**

0.063

Table A3. Annualized daily (every 22th trading day) returns of MA2–MA10, average annualized returns.

	Buy and Hold	MA10	MA9	MA8	MA7	MA6	MA5	MA4	MA3	MA2
3M	0.090	0.033	0.035	0.023	0.023	0.024	0.023	0.038	0.021	0.012
American Express	0.094	0.086	0.087	0.091	0.107	0.088	0.062	0.062	0.036	0.038
Apple	0.157	0.057	0.069	0.056	0.076	0.076	0.094	0.069	0.099	0.071
Boeing	0.119	0.122	0.122	0.102	0.099	0.115	0.110	0.100	0.091	0.077
Caterpillar	0.100	0.065	0.062	0.071	0.083	0.081	0.063	0.057	0.009	0.051
Chevron	0.084	0.022	0.021	0.025	0.026	0.019	0.032	0.032	0.013	0.005
Coca-Cola	0.099	0.083	0.072	0.087	0.071	0.073	0.072	0.069	0.046	0.026
Walt Disney	0.103	0.061	0.066	0.073	0.077	0.071	0.079	0.081	0.073	0.057
Exxon	0.072	0.040	0.038	0.028	0.028	0.034	0.020	0.027	0.025	0.026
GE	0.052	0.079	0.078	0.080	0.072	0.070	0.063	0.018	0.038	0.037
Home Depot	0.190	0.126	0.133	0.134	0.136	0.120	0.14	0.119	0.118	0.110
IBM	0.055	0.029	0.033	0.032	0.038	0.036	0.026	0.033	0.026	0.03
Intel	0.134	0.079	0.080	0.096	0.095	0.085	0.063	0.082	0.110	0.116
Johnson & Johnson	0.113	0.078	0.076	0.071	0.059	0.057	0.058	0.050	0.052	0.031
JP Morgan	0.090	0.057	0.051	0.051	0.063	0.046	0.070	0.079	0.067	0.067
McDonalds	0.114	0.077	0.077	0.057	0.055	0.045	0.056	0.042	0.045	0.033
Merck	0.063	0.069	0.069	0.054	0.059	0.05	0.045	0.027	0.011	3×10^{-4}
Microsoft	0.180	0.122	0.127	0.123	0.099	0.112	0.093	0.095	0.090	0.108
Nike	0.177	0.128	0.136	0.130	0.127	0.115	0.111	0.109	0.082	0.089
Pfizer	0.097	0.070	0.069	0.067	0.068	0.066	0.068	0.056	0.040	0.034
Procter & Gamble	0.095	0.057	0.060	0.055	0.042	0.043	0.021	0.024	0.038	0.039
Travellers	0.082	0.045	0.049	0.047	0.041	0.034	0.016	0.009	0.002	0.017
United Technologies	0.113	0.064	0.062	0.074	0.078	0.063	0.046	0.037	0.050	0.050
United Health Group	0.252	0.158	0.162	0.167	0.154	0.168	0.176	0.174	0.180	0.158
Verizon	0.043	0.002	9×10^{-4}	0.011	0.017	0.025	−0.00	0.01	−0.00	−0.02
Wal-Mart	0.113	0.046	0.046	0.040	0.044	0.032	0.041	0.037	0.023	0.038
Cisco	0.210	0.228	0.227	0.222	0.221	0.191	0.186	0.184	0.160	0.134
Goldman Sachs	0.061	0.029	0.030	0.020	0.052	0.067	0.065	0.070	0.041	0.068
Visa	0.236	0.171	0.161	0.162	0.149	0.122	0.113	0.115	0.142	0.097
Average	**0.117**	**0.079**	**0.079**	**0.078**	**0.078**	**0.073**	**0.069**	**0.066**	**0.059**	**0.055**

0.071

Table 4. Annualized daily (every other month) returns of MAD2–MAD5 (D = every other month, and 5, 4, 3, 2 are the numbers of observations in the rolling window), average annualized returns.

	Buy and Hold	MAD5	MAD4	MAD3	MAD2	
3M	0.090	0.062	0.063	0.042	0.049	
American Express	0.094	0.089	0.098	0.052	0.041	
Apple	0.157	0.040	0.042	0.030	0.085	
Boeing	0.119	0.112	0.110	0.102	0.110	
Caterpillar	0.100	0.079	0.09	0.089	0.084	
Chevron	0.084	0.033	0.036	0.026	0.028	
Coca-Cola	0.099	0.093	0.102	0.080	0.078	
Walt Disney	0.103	0.068	0.074	0.080	0.084	
Exxon	0.072	0.022	0.018	0.010	0.009	
GE	0.052	0.067	0.066	0.041	0.033	
Home Depot	0.190	0.174	0.175	0.156	0.160	
IBM	0.055	0.016	0.023	0.017	0.021	
Intel	0.134	0.093	0.098	0.089	0.112	
Johnson & Johnson	0.113	0.083	0.086	0.048	0.071	
JP Morgan	0.090	0.053	0.052	0.048	0.054	
McDonalds	0.114	0.094	0.098	0.071	0.070	
Merck	0.063	0.084	0.067	0.036	0.031	
Microsoft	0.180	0.138	0.136	0.106	0.088	
Nike	0.177	0.140	0.144	0.133	0.122	
Pfizer	0.097	0.062	0.051	0.061	0.059	
Procter & Gamble	0.095	0.048	0.054	0.048	0.034	
Travellers	0.082	0.018	0.015	0.018	2×10^{-4}	
United Technologies	0.113	0.066	0.073	0.096	0.060	
United Health Group	0.252	0.181	0.179	0.191	0.207	
Verizon	0.043	−0.018	−0.01	−0.02	−0.02	
Wal-Mart	0.113	0.067	0.065	0.050	0.061	
Cisco	0.210	0.217	0.226	0.207	0.196	
Goldman Sachs	0.061	0.041	0.059	0.060	0.039	
Visa	0.236	0.174	0.173	0.151	0.120	
Average	**0.117**	**0.083**	**0.085**	**0.073**	**0.072**	**0.078**

Table 5. Annualized daily (every third month) returns of MAT2–MAT4 (T = every third month, and 4, 3, 2 are the numbers of observations in the rolling window), average annualized returns.

	Buy and Hold	MAT4	MAT3	MAT2
3M	0.090	0.061	0.055	0.039
American Express	0.094	0.113	0.091	0.066
Apple	0.157	0.089	0.073	0.096
Boeing	0.119	0.127	0.131	0.114
Caterpillar	0.100	0.070	0.069	0.078
Chevron	0.084	0.047	0.053	0.037
Coca-Cola	0.099	0.077	0.078	0.072
Walt Disney	0.103	0.043	0.042	0.068
Exxon	0.072	0.055	0.049	0.037
GE	0.052	0.084	0.080	0.047
Home Depot	0.190	0.161	0.163	0.128
IBM	0.055	0.054	0.048	0.028
Intel	0.134	0.107	0.115	0.072
Johnson & Johnson	0.113	0.094	0.094	0.074
JP Morgan	0.090	0.058	0.076	0.007
McDonalds	0.114	0.080	0.082	0.069
Merck	0.063	0.062	0.062	0.049
Microsoft	0.180	0.127	0.128	0.080
Nike	0.177	0.146	0.151	0.099

Table 5. *Cont.*

	Buy and Hold	MAT4	MAT3	MAT2	
Pfizer	0.097	0.078	0.070	0.056	
Procter & Gamble	0.095	0.068	0.072	0.076	
Travellers	0.082	0.041	0.043	0.025	
United Technologies	0.113	0.077	0.089	0.079	
United Health Group	0.252	0.147	0.161	0.178	
Verizon	0.043	−0.00	−0.00	−0.02	
Wal-Mart	0.113	0.081	0.081	0.083	
Cisco	0.210	0.211	0.217	0.213	
Goldman Sachs	0.061	0.044	0.026	0.030	
Visa	0.236	0.183	0.199	0.177	
Average	**0.117**	**0.089**	**0.089**	**0.075**	**0.084**

Table 6. Annualized daily (every fourth month) returns of MAQ2–MAQ3 (Q = every fourth month, and 3 and 2 are the numbers of observations in the rolling window), average annualized returns.

	Buy and Hold	MAQ3	MAQ2	
3M	0.090	0.056	0.058	
American Express	0.094	0.089	0.094	
Apple	0.157	0.094	0.094	
Boeing	0.119	0.122	0.128	
Caterpillar	0.100	0.064	0.084	
Chevron	0.084	0.060	0.054	
Coca-Cola	0.099	0.083	0.093	
Walt Disney	0.103	0.061	0.062	
Exxon	0.072	0.056	0.064	
GE	0.052	0.069	0.081	
Home Depot	0.190	0.152	0.157	
IBM	0.055	0.048	0.031	
Intel	0.134	0.064	0.070	
Johnson & Johnson	0.113	0.080	0.079	
JP Morgan	0.090	0.085	0.091	
McDonalds	0.114	0.096	0.112	
Merck	0.063	0.056	0.061	
Microsoft	0.180	0.143	0.145	
Nike	0.177	0.181	0.199	
Pfizer	0.097	0.059	0.045	
Procter & Gamble	0.095	0.073	0.077	
Travellers	0.082	0.051	0.051	
United Technologies	0.113	0.080	0.077	
United Health Group	0.252	0.185	0.218	
Verizon	0.043	0.027	0.023	
Wal-Mart	0.113	0.087	0.076	
Cisco	0.210	0.195	0.180	
Goldman Sachs	0.061	0.042	0.056	
Visa	0.236	0.195	0.228	
Average	**0.117**	**0.091**	**0.096**	**0.094**

Table 7. Annualized daily (every fifth month) returns of MAC2 (C = every fifth month, and 2 = observations accounting in the rolling window), average annualized returns.

	Buy and Hold	MAC2
3M	0.090	0.076
American Express	0.094	0.088
Apple	0.157	0.132
Boeing	0.119	0.080
Caterpillar	0.100	0.094
Chevron	0.084	0.047
Coca-Cola	0.099	0.094
Walt Disney	0.103	0.044
Exxon	0.072	0.049
GE	0.052	0.048
Home Depot	0.190	0.143
IBM	0.055	0.032
Intel	0.133	0.057
Johnson & Johnson	0.113	0.081
JP Morgan	0.090	0.045
McDonalds	0.114	0.079
Merck	0.063	0.080
Microsoft	0.180	0.094
Nike	0.177	0.141
Pfizer	0.097	0.099
Procter & Gamble	0.095	0.039
Travellers	0.082	0.068
United Technologies	0.113	0.056
United Health Group	0.252	0.152
Verizon	0.043	0.048
Wal-Mart	0.113	0.093
Cisco	0.210	0.225
Goldman Sachs	0.061	0.053
Visa	0.236	0.217
Average	**0.117**	**0.088**

B.

Table 8. Annualized daily volatility of MA40–MA200, average annualized volatility.

	Buy and Hold	MA200	MA180	MA160	MA140	MA120	MA100	MA80	MA60	MA40
3M	0.225	0.164	0.165	0.161	0.161	0.159	0.159	0.158	0.158	0.157
American Express	0.345	0.227	0.228	0.221	0.225	0.224	0.225	0.224	0.228	0.229
Apple	0.451	0.317	0.321	0.315	0.315	0.313	0.315	0.315	0.310	0.305
Boeing	0.294	0.201	0.203	0.199	0.201	0.199	0.198	0.198	0.201	0.204
Caterpillar	0.311	0.216	0.218	0.216	0.216	0.214	0.215	0.214	0.213	0.215
Chevron	0.244	0.167	0.168	0.166	0.166	0.166	0.165	0.164	0.167	0.168
Coca-Cola	0.225	0.164	0.166	0.161	0.160	0.159	0.158	0.158	0.156	0.155
Walt Disney	0.291	0.196	0.201	0.199	0.200	0.199	0.198	0.203	0.204	0.203
Exxon	0.230	0.162	0.163	0.159	0.159	0.159	0.157	0.156	0.155	0.157
GE	0.275	0.174	0.175	0.172	0.173	0.173	0.171	0.168	0.168	0.168
Home Depot	0.314	0.226	0.228	0.223	0.221	0.221	0.219	0.217	0.217	0.214
IBM	0.271	0.187	0.189	0.185	0.184	0.181	0.179	0.177	0.176	0.174
Intel	0.382	0.273	0.275	0.267	0.265	0.263	0.260	0.257	0.256	0.254
Johnson & Johnson	0.215	0.163	0.164	0.161	0.159	0.157	0.155	0.153	0.152	0.149
JP Morgan	0.375	0.223	0.226	0.223	0.224	0.227	0.237	0.242	0.245	0.248
McDonalds	0.240	0.183	0.184	0.18	0.178	0.177	0.176	0.176	0.175	0.174
Merck	0.269	0.177	0.179	0.173	0.173	0.174	0.172	0.174	0.174	0.177
Microsoft	0.323	0.248	0.249	0.243	0.241	0.237	0.236	0.233	0.232	0.231
Nike	0.327	0.243	0.245	0.238	0.236	0.235	0.235	0.232	0.232	0.233
Pfizer	0.266	0.188	0.19	0.187	0.186	0.187	0.186	0.187	0.187	0.187
Procter & Gamble	0.225	0.169	0.169	0.164	0.163	0.161	0.158	0.157	0.156	0.156
Travellers	0.268	0.174	0.175	0.174	0.175	0.178	0.180	0.184	0.182	0.185
United Technologies	0.261	0.179	0.181	0.179	0.178	0.177	0.177	0.177	0.176	0.173
United Health Group	0.386	0.290	0.293	0.290	0.290	0.283	0.282	0.282	0.280	0.273
Verizon	0.246	0.163	0.165	0.164	0.163	0.163	0.163	0.161	0.161	0.163
Wal-Mart	0.263	0.203	0.204	0.200	0.198	0.195	0.191	0.19	0.189	0.191
Cisco	0.415	0.300	0.302	0.297	0.295	0.291	0.290	0.285	0.282	0.275
Goldman Sachs	0.373	0.222	0.226	0.22	0.222	0.223	0.228	0.230	0.227	0.229
Visa	0.260	0.209	0.212	0.209	0.208	0.212	0.208	0.206	0.205	0.197
Average	**0.295**	**0.207**	**0.209**	**0.205**	**0.205**	**0.204**	**0.203**	**0.203**	**0.202**	**0.202**
										0.204

Table 9. Annualized daily (every fifth trading day) volatility of MAW8–MAW40 (W = number of weeks), average annualized volatility.

	Buy and Hold	MAW40	MAW36	MAW32	MAW28	MAW24	MAW20	MAW16	MAW12	MAW8
3M	0.225	0.165	0.165	0.163	0.163	0.16	0.159	0.157	0.157	0.159
American Express	0.345	0.227	0.224	0.224	0.227	0.225	0.223	0.228	0.232	0.234
Apple	0.451	0.316	0.316	0.313	0.318	0.316	0.343	0.317	0.312	0.309
Boeing	0.294	0.204	0.203	0.204	0.204	0.203	0.203	0.201	0.201	0.206
Caterpillar	0.311	0.216	0.215	0.215	0.217	0.214	0.215	0.215	0.213	0.214
Chevron	0.244	0.169	0.168	0.169	0.168	0.168	0.167	0.166	0.168	0.172
Coca-Cola	0.225	0.165	0.165	0.164	0.162	0.160	0.159	0.159	0.157	0.155
Walt Disney	0.291	0.195	0.198	0.197	0.197	0.199	0.200	0.202	0.203	0.204
Exxon	0.230	0.163	0.161	0.160	0.161	0.160	0.157	0.156	0.153	0.158
GE	0.275	0.174	0.174	0.174	0.175	0.174	0.170	0.169	0.171	0.166
Home Depot	0.314	0.228	0.228	0.226	0.225	0.222	0.224	0.219	0.219	0.214
IBM	0.271	0.190	0.188	0.185	0.184	0.183	0.178	0.177	0.178	0.177
Intel	0.382	0.267	0.267	0.268	0.264	0.263	0.259	0.256	0.259	0.259
Johnson & Johnson	0.215	0.164	0.163	0.162	0.160	0.158	0.156	0.156	0.152	0.15
JP Morgan	0.375	0.222	0.225	0.224	0.230	0.236	0.239	0.243	0.241	0.252
McDonalds	0.240	0.185	0.182	0.181	0.179	0.177	0.177	0.176	0.174	0.171
Merck	0.269	0.179	0.175	0.174	0.173	0.173	0.172	0.175	0.176	0.175
Microsoft	0.323	0.250	0.247	0.245	0.244	0.24	0.236	0.236	0.230	0.232
Nike	0.327	0.244	0.241	0.239	0.240	0.241	0.238	0.235	0.232	0.232
Pfizer	0.266	0.189	0.187	0.186	0.187	0.188	0.190	0.189	0.189	0.184
Procter & Gamble	0.225	0.170	0.168	0.167	0.165	0.164	0.161	0.158	0.160	0.156
Travellers	0.268	0.175	0.175	0.175	0.178	0.177	0.177	0.184	0.184	0.185
United Technologies	0.261	0.181	0.179	0.178	0.177	0.177	0.177	0.177	0.176	0.172
United Health Group	0.386	0.292	0.291	0.292	0.291	0.290	0.289	0.287	0.282	0.278
Verizon	0.246	0.163	0.162	0.162	0.162	0.164	0.162	0.161	0.160	0.159
Wal-Mart	0.263	0.205	0.202	0.201	0.198	0.194	0.191	0.191	0.190	0.192
Cisco	0.415	0.307	0.305	0.300	0.296	0.292	0.293	0.288	0.285	0.281
Goldman Sachs	0.373	0.225	0.223	0.221	0.221	0.220	0.230	0.233	0.241	0.241
Visa	0.260	0.203	0.210	0.209	0.208	0.210	0.208	0.206	0.203	0.195
Average	**0.295**	**0.208**	**0.207**	**0.206**	**0.206**	**0.205**	**0.205**	**0.204**	**0.203**	**0.203**

0.205

Table 10. Annualized daily (rule in every 22th trading day) volatility of MA2–MA10, average annualized volatility.

	Buy and Hold	MA10	MA9	MA8	MA7	MA6	MA5	MA4	MA3	MA2
3M	0.225	0.167	0.169	0.162	0.163	0.161	0.161	0.157	0.156	0.156
American Express	0.345	0.232	0.235	0.222	0.218	0.22	0.219	0.22	0.243	0.235
Apple	0.451	0.343	0.347	0.342	0.339	0.339	0.338	0.342	0.335	0.331
Boeing	0.294	0.207	0.210	0.202	0.202	0.199	0.200	0.197	0.207	0.205
Caterpillar	0.311	0.216	0.220	0.217	0.215	0.214	0.217	0.218	0.221	0.224
Chevron	0.244	0.169	0.171	0.172	0.17	0.169	0.169	0.167	0.181	0.171
Coca-Cola	0.225	0.168	0.171	0.169	0.168	0.166	0.161	0.161	0.161	0.156
Walt Disney	0.291	0.203	0.207	0.202	0.203	0.203	0.210	0.212	0.215	0.211
Exxon	0.230	0.166	0.167	0.165	0.164	0.163	0.162	0.157	0.161	0.160
GE	0.275	0.177	0.177	0.175	0.175	0.175	0.172	0.169	0.172	0.180
Home Depot	0.314	0.234	0.235	0.228	0.221	0.230	0.228	0.233	0.225	0.219
IBM	0.271	0.194	0.196	0.196	0.196	0.196	0.19	0.194	0.195	0.190
Intel	0.382	0.273	0.277	0.272	0.272	0.268	0.266	0.266	0.264	0.259
Johnson & Johnson	0.215	0.168	0.169	0.167	0.167	0.162	0.158	0.158	0.154	0.150
JP Morgan	0.375	0.222	0.223	0.217	0.220	0.230	0.233	0.234	0.244	0.234
McDonalds	0.240	0.189	0.189	0.186	0.185	0.185	0.179	0.170	0.171	0.180
Merck	0.269	0.177	0.178	0.173	0.173	0.174	0.173	0.181	0.182	0.192
Microsoft	0.323	0.250	0.251	0.247	0.239	0.233	0.235	0.237	0.233	0.234
Nike	0.327	0.247	0.248	0.244	0.241	0.240	0.235	0.236	0.238	0.248
Pfizer	0.266	0.188	0.190	0.186	0.186	0.186	0.187	0.187	0.191	0.189
Procter & Gamble	0.225	0.173	0.174	0.171	0.167	0.165	0.163	0.164	0.158	0.155
Travellers	0.268	0.171	0.172	0.17	0.169	0.171	0.191	0.186	0.192	0.198
United Technologies	0.261	0.178	0.179	0.178	0.177	0.177	0.175	0.178	0.176	0.173
United Health Group	0.386	0.300	0.302	0.299	0.298	0.294	0.289	0.280	0.283	0.275
Verizon	0.246	0.167	0.167	0.164	0.162	0.160	0.164	0.157	0.160	0.163
Wal-Mart	0.263	0.208	0.210	0.205	0.199	0.196	0.197	0.198	0.198	0.189
Cisco	0.415	0.304	0.307	0.301	0.298	0.300	0.292	0.290	0.281	0.278
Goldman Sachs	0.373	0.230	0.232	0.225	0.232	0.245	0.239	0.253	0.268	0.256
Visa	0.260	0.204	0.203	0.212	0.225	0.221	0.219	0.217	0.217	0.196
Average	**0.295**	**0.211**	**0.213**	**0.209**	**0.208**	**0.208**	**0.208**	**0.208**	**0.210**	**0.207**

0.209

Table 11. Annualized daily (every other month) volatility of MAD2–MAD5 (D = every other month, and 5, 4, 3, 2 are the numbers of observations in the rolling window), average annualized volatility.

	Buy and Hold	MAD5	MAD4	MAD3	MAD2	
3M	0.225	0.168	0.169	0.162	0.159	
American Express	0.344	0.222	0.226	0.216	0.211	
Apple	0.450	0.351	0.363	0.357	0.338	
Boeing	0.294	0.210	0.216	0.211	0.208	
Caterpillar	0.311	0.218	0.229	0.215	0.211	
Chevron	0.244	0.168	0.175	0.166	0.165	
Coca-Cola	0.225	0.168	0.173	0.165	0.158	
Walt Disney	0.291	0.197	0.200	0.198	0.203	
Exxon	0.230	0.172	0.174	0.159	0.156	
GE	0.274	0.175	0.181	0.176	0.182	
Home Depot	0.314	0.229	0.230	0.221	0.237	
IBM	0.271	0.196	0.199	0.200	0.200	
Intel	0.382	0.274	0.286	0.267	0.265	
Johnson & Johnson	0.215	0.173	0.175	0.165	0.154	
JP Morgan	0.375	0.236	0.241	0.246	0.237	
McDonalds	0.240	0.182	0.186	0.178	0.169	
Merck	0.269	0.185	0.196	0.188	0.199	
Microsoft	0.323	0.245	0.249	0.238	0.250	
Nike	0.327	0.252	0.258	0.253	0.253	
Pfizer	0.266	0.199	0.203	0.191	0.189	
Procter & Gamble	0.225	0.173	0.177	0.169	0.166	
Travellers	0.268	0.176	0.178	0.183	0.191	
United Technologies	0.261	0.182	0.187	0.178	0.177	
United Health Group	0.386	0.313	0.313	0.299	0.305	
Verizon	0.246	0.163	0.171	0.165	0.153	
Wal-Mart	0.263	0.197	0.199	0.194	0.193	
Cisco	0.415	0.312	0.317	0.315	0.285	
Goldman Sachs	0.373	0.229	0.245	0.239	0.265	
Visa	0.260	0.215	0.215	0.225	0.222	
Average	**0.295**	**0.213**	**0.218**	**0.212**	**0.210**	**0.213**

Table 12. Annualized daily (every third month) volatility of MAT2–MAT4 (T = every third month, and 4, 3, 2 are the numbers of observations in the rolling window), average annualized volatility.

	Buy and Hold	MAT4	MAT3	MAT2	
3M	0.225	0.172	0.174	0.171	
American Express	0.344	0.230	0.237	0.206	
Apple	0.450	0.345	0.357	0.349	
Boeing	0.294	0.206	0.219	0.200	
Caterpillar	0.311	0.219	0.223	0.214	
Chevron	0.244	0.176	0.182	0.170	
Coca-Cola	0.225	0.177	0.179	0.181	
Walt Disney	0.291	0.220	0.228	0.205	
Exxon	0.230	0.168	0.176	0.158	
GE	0.274	0.178	0.185	0.177	
Home Depot	0.314	0.236	0.251	0.241	
IBM	0.271	0.205	0.209	0.193	
Intel	0.382	0.285	0.296	0.274	
Johnson & Johnson	0.215	0.185	0.188	0.165	
JP Morgan	0.375	0.242	0.248	0.240	
McDonalds	0.240	0.198	0.204	0.192	
Merck	0.269	0.191	0.191	0.180	
Microsoft	0.323	0.257	0.267	0.258	
Nike	0.327	0.264	0.265	0.258	

Table 12. *Cont.*

	Buy and Hold	MAT4	MAT3	MAT2	
Pfizer	0.266	0.195	0.206	0.208	
Procter & Gamble	0.225	0.177	0.181	0.168	
Travellers	0.268	0.187	0.188	0.198	
United Technologies	0.261	0.192	0.199	0.187	
United Health Group	0.386	0.300	0.308	0.315	
Verizon	0.246	0.176	0.176	0.160	
Wal-Mart	0.263	0.202	0.208	0.208	
Cisco	0.415	0.310	0.311	0.303	
Goldman Sachs	0.373	0.226	0.232	0.235	
Visa	0.260	0.204	0.215	0.208	
Average	**0.295**	**0.218**	**0.224**	**0.214**	**0.219**

Table 13. Annualized daily (every fourth month) volatility of MAQ2–MAQ3 (Q = every fourth month, 3 and 2 are the number of observations in the rolling window), average annualized volatility.

	Buy and Hold	MAQ3	MAQ3	
3M	0.225	0.168	0.176	
American Express	0.344	0.220	0.226	
Apple	0.450	0.360	0.373	
Boeing	0.294	0.213	0.224	
Caterpillar	0.311	0.222	0.239	
Chevron	0.244	0.167	0.177	
Coca-Cola	0.225	0.173	0.182	
Walt Disney	0.291	0.206	0.218	
Exxon	0.230	0.160	0.176	
GE	0.274	0.180	0.195	
Home Depot	0.314	0.237	0.242	
IBM	0.271	0.194	0.218	
Intel	0.382	0.274	0.293	
Johnson & Johnson	0.215	0.181	0.186	
JP Morgan	0.375	0.218	0.227	
McDonalds	0.240	0.177	0.193	
Merck	0.269	0.204	0.212	
Microsoft	0.323	0.248	0.260	
Nike	0.327	0.258	0.265	
Pfizer	0.266	0.198	0.207	
Procter & Gamble	0.225	0.173	0.174	
Travellers	0.268	0.182	0.192	
United Technologies	0.261	0.181	0.188	
United Health Group	0.386	0.299	0.314	
Verizon	0.246	0.167	0.177	
Wal-Mart	0.263	0.194	0.207	
Cisco	0.415	0.341	0.349	
Goldman Sachs	0.373	0.240	0.260	
Visa	0.260	0.212	0.225	
Average	**0.295**	**0.215**	**0.227**	**0.221**

Table 14. Annualized daily (every fifth month) volatility of MAC2 (C = every fifth month, 2 = observations in rolling window), average annualized volatility.

	Buy and Hold	MAC2
3M	0.225	0.176
American Express	0.344	0.226
Apple	0.450	0.323
Boeing	0.294	0.218
Caterpillar	0.311	0.227
Chevron	0.244	0.165
Coca-Cola	0.225	0.168
Walt Disney	0.291	0.206
Exxon	0.230	0.166
GE	0.274	0.187
Home Depot	0.314	0.242
IBM	0.271	0.202
Intel	0.382	0.296
Johnson & Johnson	0.215	0.187
JP Morgan	0.375	0.244
McDonalds	0.240	0.182
Merck	0.269	0.194
Microsoft	0.323	0.250
Nike	0.327	0.249
Pfizer	0.266	0.191
Procter & Gamble	0.225	0.187
Travellers	0.268	0.183
United Technologies	0.261	0.204
United Health Group	0.386	0.298
Verizon	0.246	0.170
Wal-Mart	0.263	0.223
Cisco	0.415	0.333
Goldman Sachs	0.373	0.218
Visa	0.260	0.220
Average	**0.295**	**0.218**

c.

Table 15. Transaction costs per year of MA40–MA200, with one transaction costing 0.1% of total wealth, average annualized transaction costs.

	MA40	MA60	MA80	MA100	MA120	MA140	MA160	MA180	MA200
3M	0.022	0.019	0.016	0.013	0.012	0.011	0.010	0.011	0.010
American Express	0.023	0.017	0.016	0.013	0.012	0.012	0.011	0.011	0.011
Apple	0.020	0.015	0.014	0.012	0.010	0.009	0.008	0.008	0.007
Boeing	0.020	0.015	0.015	0.012	0.011	0.011	0.010	0.009	0.008
Caterpillar	0.019	0.015	0.018	0.013	0.012	0.011	0.010	0.009	0.008
Chevron	0.024	0.020	0.015	0.016	0.014	0.013	0.012	0.012	0.011
Coca-Cola	0.022	0.018	0.013	0.012	0.011	0.011	0.011	0.010	0.009
Walt Disney	0.021	0.017	0.013	0.012	0.012	0.011	0.009	0.008	0.007
Exxon	0.028	0.023	0.019	0.018	0.017	0.017	0.016	0.013	0.011
GE	0.023	0.017	0.014	0.012	0.011	0.010	0.009	0.008	0.007
Home Depot	0.021	0.018	0.016	0.014	0.013	0.011	0.010	0.009	0.008
IBM	0.019	0.014	0.013	0.012	0.012	0.010	0.010	0.010	0.009
Intel	0.019	0.016	0.014	0.014	0.012	0.010	0.010	0.009	0.007
Johnson & Johnson	0.024	0.020	0.016	0.014	0.012	0.010	0.009	0.008	0.009
JP Morgan	0.020	0.016	0.015	0.014	0.012	0.012	0.011	0.010	0.010
McDonalds	0.023	0.018	0.016	0.014	0.012	0.013	0.011	0.011	0.010
Merck	0.022	0.017	0.015	0.013	0.011	0.011	0.009	0.009	0.008
Microsoft	0.020	0.015	0.015	0.013	0.011	0.010	0.010	0.009	0.008
Nike	0.019	0.014	0.013	0.012	0.011	0.010	0.010	0.009	0.009
Pfizer	0.021	0.017	0.014	0.012	0.011	0.011	0.010	0.010	0.008
Procter & Gamble	0.022	0.019	0.016	0.014	0.012	0.011	0.010	0.010	0.010
Travellers	0.024	0.018	0.016	0.015	0.013	0.012	0.012	0.011	0.010
United Technologies	0.021	0.018	0.015	0.014	0.012	0.011	0.011	0.010	0.009
United Health Group	0.021	0.017	0.014	0.012	0.011	0.010	0.010	0.008	0.008
Verizon	0.023	0.018	0.017	0.014	0.013	0.011	0.011	0.011	0.011
Wal-Mart	0.022	0.019	0.015	0.014	0.013	0.013	0.012	0.010	0.010
Cisco	0.023	0.017	0.014	0.010	0.009	0.010	0.008	0.006	0.006
Goldman Sachs	0.035	0.026	0.022	0.015	0.014	0.012	0.012	0.010	0.008
Visa	0.022	0.014	0.011	0.010	0.008	0.009	0.009	0.008	0.008
Average	**0.022**	**0.018**	**0.015**	**0.013**	**0.012**	**0.011**	**0.010**	**0.0010**	**0.009**

0.013

Table 16. Transaction costs per year of MA2–MA10, average annualized transaction costs.

	MA10	MA9	MA8	MA7	MA6	MA5	MA4	MA3	MA2	
3M	0.003	0.003	0.003	0.003	0.003	0.004	0.004	0.005	0.006	
American Express	0.002	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.006	
Apple	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.005	0.006	
Boeing	0.002	0.002	0.002	0.002	0.002	0.003	0.004	0.004	0.006	
Caterpillar	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.005	0.006	
Chevron	0.002	0.003	0.003	0.003	0.003	0.003	0.004	0.005	0.007	
Coca-Cola	0.002	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.006	
Walt Disney	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.004	0.006	
Exxon	0.002	0.002	0.003	0.003	0.003	0.004	0.004	0.005	0.006	
GE	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.004	0.006	
Home Depot	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.004	0.006	
IBM	0.003	0.002	0.003	0.002	0.003	0.003	0.004	0.004	0.006	
Intel	0.002	0.003	0.003	0.003	0.003	0.003	0.004	0.004	0.006	
Johnson & Johnson	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.005	0.006	
JP Morgan	0.002	0.003	0.003	0.003	0.003	0.003	0.003	0.004	0.006	
McDonalds	0.002	0.002	0.003	0.003	0.003	0.003	0.004	0.005	0.006	
Merck	0.002	0.002	0.002	0.003	0.003	0.003	0.004	0.005	0.006	
Microsoft	0.002	0.002	0.002	0.003	0.003	0.003	0.004	0.004	0.006	
Nike	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.004	0.006	
Pfizer	0.002	0.002	0.002	0.003	0.003	0.003	0.004	0.004	0.006	
Procter & Gamble	0.002	0.002	0.003	0.003	0.003	0.004	0.004	0.005	0.006	
Travellers	0.003	0.002	0.003	0.003	0.003	0.004	0.004	0.005	0.007	
United Technologies	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.004	0.006	
United Health Group	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.004	0.006	
Verizon	0.003	0.003	0.003	0.003	0.003	0.004	0.004	0.005	0.006	
Wal-Mart	0.003	0.003	0.003	0.003	0.003	0.004	0.004	0.005	0.006	
Cisco	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.005	0.006	
Goldman Sachs	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.004	0.005	
Visa	0.002	0.001	0.002	0.002	0.002	0.003	0.003	0.003	0.005	
Average	**0.002**	**0.002**	**0.002**	**0.003**	**0.003**	**0.003**	**0.004**	**0.004**	**0.006**	**0.003**

References

1. Tobin, J. Liquidity preference as behavior towards risk. *Rev. Econ. Stud.* **1958**, *67*, 65–86. [CrossRef]
2. Fama, E. Components on investment performance. *J. Financ.* **1972**, *27*, 551–568.
3. Merton, R. On market timing and investment performance. I. An equilibrium theory of value for market forecast. *J. Bus.* **1981**, *54*, 363–406. [CrossRef]
4. Henriksson, R.; Merton, R. On market timing and investment performance II: Statistical procedures for evaluating forecasting skills. *J. Bus.* **1981**, *54*, 513–533. [CrossRef]
5. Sharpe, W. Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Financ.* **1964**, *19*, 425–442.
6. Lintner, J. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Econ. Stat.* **1965**, *47*, 13–37. [CrossRef]
7. Zhu, Y.; Zhou, G. Technical analysis: An asset allocation perspective on the use of moving averages. *J. Financ. Econ.* **2009**, *9*, 519–544. [CrossRef]
8. Markowitz, H. Portfolio selection. *J. Financ.* **1952**, *7*, 77–91.
9. Baker, M.; Bradley, B.; Wurgler, J. Benchmark as limits to arbitrage: Understanding the low-volatility anomaly. *Financ. Anal. J.* **2011**, *67*, 1–15. [CrossRef]
10. Ang, A.; Hodrick, R.; Xing, Y.; Zhang, X. High idiosyncratic volatility and low returns: International and further U.S. evidence. *J. Financ. Econ.* **2009**, *61*, 1–23. [CrossRef]
11. Cochrane, J. The dog that did not bark: A defense of return predictability. *Rev. Financ. Stud.* **2008**, *21*, 1533–1573. [CrossRef]
12. Fama, E. Two pillars of asset pricing. *Am. Econ. Rev.* **2014**, *104*, 1467–1485. [CrossRef]

13. Jagannathan, R.; Korajczyk, R. Market timing. In *Portfolio Construction, Measurement and Efficiency*; Guerard, J., Ed.; Springer International Publishing: Cham, Switzerland, 2017; pp. 49–71. ISBN 978-3-319-33974-0.

14. Friesen, G.; Sapp, T. Mutual fund flows and investor returns: An empirical examination of fund investor timing ability. *J. Bank. Financ.* **2007**, *31*, 2796–2816, [CrossRef]

15. Munoz, F.; Vicente, R. Hindsight effect: What are actual cash flow timing skills of mutual fund investors? *J. Empir. Financ.* **2018**, *45*, 181–193. [CrossRef]

16. Malkiel, B. The efficient market hypothesis and its critics. *J. Econ. Perspect.* **2003**, *17*, 59–82. [CrossRef]

17. Samuelson, P. Summing up on business cycles: Opening address. In *Beyond Shocks; What Causes Business Cycles?* Fuhrer, J., Schuh, S., Eds.; Federal Reserve Bank of Boston: Boston, MA, USA, 1998.

18. Black, F. Capital market equilibrium with restricted borrowing. *J. Bus.* **1972**, *45*, 444–455. [CrossRef]

19. Black, F.; Jensen, M.; Scholes, M. The capital asset market model: Some empirical tests. In *Studies in the Theory of Capital Markets*; Jensen, M., Ed.; Praeger: New York, NY, USA, 1972; pp. 79–121.

20. Baker, M.; Bradley, B.; Taliaferro, R. The Low-risk anomaly: A decomposition into micro and macro effects. *Financ. Anal. J.* **2014**, *70*, 45–58. [CrossRef]

21. Frazzini, A.; Pedersen, L. Betting against betas. *J. Financ. Econ.* **2014**, *111*, 1–25. [CrossRef]

22. Clarke, R.; de Silva, H.; Thorley, S. Know your VMS exposure. *J. Portf. Manag.* **2010**, *36*, 52–59. [CrossRef]

23. Barberis, N.; Huang, M. Stocks as lotteries: The implications of probability weighting for security prices. *Am. Econ. Rev.* **2008**, *98*, 2066–2100. [CrossRef]

24. Kumar, A. Who gambles in the stock markets? *J. Financ.* **2009**, *64*, 1889–1933. [CrossRef]

25. Bali, T.; Cakici, N.; Whitelaw, R. Maxing out: Stock as lotteries and the cross-section of expected returns. *J. Financ. Econ.* **2011**, *99*, 427–446. [CrossRef]

26. Ben-David, I.; Graham, J.; Harvey, C. Managerial miscalibration. *Quart. J. Econ.* **2013**, *128*, 1547–1584. [CrossRef]

27. Daniel, K.; Titman, S. Market reactions to tangible and intangible information. *J. Financ.* **2006**, *61*, 1605–1643. [CrossRef]

28. Kahneman, D.; Tversky, A. Judgement under uncertainty: Heuristics and biases. *Science* **1974**, *185*, 1124–1131.

29. Baker, M.; Wurgler, J. Do strict requirements raise the cost of capital? Bank regulation, capital structure, and the low-risk anomaly. *Am. Econ. Rev. Pap. Proc.* **2015**, *105*, 315–320. [CrossRef]

30. Hong, H.; Sraer, D. Speculative betas. *J. Financ.* **2016**, *71*, 2095–2144. [CrossRef]

31. Li, X.; Sullivan, R.; Garcia-Feijoo, L. The low-volatility anomaly: Market evidence on systematic risk vs. mispricing. *Financ. Anal. J.* **2016**, *72*, 36–47. [CrossRef]

32. Brown, D.; Jennings, R. On technical analysis. *Rev. Financ. Stud.* **1989**, *2*, 527–551. [CrossRef]

33. Gartley, H. *Profits in the Stock Markets*; Lambert-Gann Publishing: Washington, DC, USA, 1935.

34. Moskowitz, T.; Ooi, Y.; Pedersen, L. Time series momentum. *J. Financ. Econ.* **2012**, *104*, 228–250. [CrossRef]

35. Jegadeesh, N.; Titman, S. Returns to buying winners selling losers: Implications for stock market efficiency. *J. Financ.* **1993**, *48*, 65–91. [CrossRef]

36. LeRoy, S. Risk aversion and the martingale property of stock prices. *Int. Econ. Rev.* **1973**, *14*, 436–446. [CrossRef]

37. Lucas, R. Asset prices in an exchange economy. *Econometrica* **1978**, *46*, 1429–1445. [CrossRef]

38. Campbell, J.; Cochrane, J. By force of habit: Consumption-based explanation of aggregate stock market behavior. *J. Political Econ.* **1999**, *107*, 205–251. [CrossRef]

39. Brock, W.; Lakonishok, J.; LeBaron, B. Simple technical trading rules and the stochastic properties of stock returns. *J. Financ.* **1992**, *47*, 1731–1764. [CrossRef]

40. Sullivan, R.; Timmermann, A.; White, H. Data-snooping, technical trading rule performance and the bootstrap. *J. Financ.* **1999**, *53*, 1647–1691. [CrossRef]

41. Allen, F.; Karjalainen, R. Using genetic algorithms to find technical trading rules. *J. Financ. Econ.* **1999**, *51*, 245–271. [CrossRef]

42. Lo, A.; Mamaysky, H.; Wang, J. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *J. Financ.* **2000**, *54*, 1705–1770. [CrossRef]

43. Neely, C.; Rapach, D.; Tu, J.; Zhou, G. Forecasting equity risk premium: The role of technical indicators. *Manag. Sci.* **2014**, *66*, 1772–1791. [CrossRef]

44. Marshall, B.; Nguyen, N.; Visaltanachoti, N. Time series momentum and moving average trading rules. *Quant. Financ.* **2017**, *17*, 405–421. [CrossRef]

45. Kim, A.; Tse, Y.; Wald, J. Time series momentum and volatility scaling. *J. Financ. Mark.* **2016**, *30*, 103–124. [CrossRef]

Article

An Improvement of Gain-Loss Price Bounds on Options Based on Binomial Tree and Market-Implied Risk-Neutral Distribution

Shi-jie Jiang [1],*, Mujun Lei [1] and Cheng-Huang Chung [2]

[1] College of Finance and Statistics, Hunan University, No.109, Shijiachong Road, Changsha 410-006, China; mujunlei@hnu.edu.cn

[2] Risk Management Department, Waterland Securities Co., Ltd., 5F., No.188, Sec. 5, Nanjing E. Rd., Songshan Dist., Taipei City 10571, Taiwan; d92723013@ntu.edu.tw

* Correspondence: actjiang@msn.com

Abstract: This paper investigates the approximated arbitrage bounds of option prices in an incomplete market setting and draws implications for option pricing and risk management. It gives consideration to periods of global financial crisis and European sovereign debt crisis. To this end, we employ the gain-loss ratio method combined with the market-implied risk-neutral distribution calculated by binomial tree to investigate the options price bounds. Our implied gain-loss bounds of option prices are preference-free and parametric-free to avoid the misspecification error of subjective choice on the benchmark model of gain-loss ratio, and consequently, greatly reduce model risk and market risk. The empirical results show that there are option prices breaking the gain-loss bounds, even after taking into account the market information. This means that a good risk management technique and good-deal investment opportunities exist if the implied binomial tree is used as a benchmark model in the gain-loss bounds.

---

1. Introduction

Asset pricing is an essential issue in financial economics. There are two main fundamental ideas that are explored in asset pricing: "equilibrium valuation" and "arbitrage-free valuation". If markets are complete, we can directly derive the equilibrium state prices by solving the individual agent's optimization problem, which maximizes the individual's utility function subjected to the wealth constraints. Following the equilibrium arguments, the Black–Scholes options pricing formula can be derived in a discrete-time economy [1,2], and the literature calls this an equilibrium valuation or a model-based pricing model. On the other hand, the arbitrage-free valuation, known as no-arbitrage pricing, says that a securities market price is arbitrage-free if there are no arbitrage opportunities [3]. A bundle of basis assets with given prices and an absence of arbitrage opportunity restricts the admissible set of pricing kernels. If the markets are complete, there is only one strictly positive vector of pricing kernel that correctly prices the basis assets. However, if the number of basis assets is fewer than the states of nature, the admissible set contains many pricing kernels that yield a range of asset prices [4]. For market participants, the main problem is that the prices obtained from the Black-Scholes model differ significantly from observed prices which have serious consequences for market participants measuring the market risk [5].

Two classical models, the good-deal bound model [6] and the gain–loss ratio model [7], propose a framework to unify the model-based model and no-arbitrage pricing model and tighten the pricing

bounds by introducing additional restrictions. The good-deal bounds model [6] is developed by imposing a restriction on Sharpe ratio of pricing kernel's variance and strengthened the option price "semi-arbitrage" bounds. The gain-loss ratio model [7] imposes a restriction on the pricing kernel with respect to the gain-loss ratio on the expected gains to expected losses to limit the possibility of the pricing kernel deviating from a reference benchmark pricing model. This method also has a more stable result in pricing the deep out-of-money options comparing to good-deal bound [7]. When restricting the variance of pricing kernel smaller than the level of Sharpe ratio, the good-deal bound model [6] is equal to the gain-loss ratio model [7]. In order to achieve robustness, the gain-loss ratio model is utilized in this paper. However, one limitation of the gain-loss ratio model is that the model builder has to subjectively choice the reference pricing kernel or benchmarked model to construct the gain-loss bounds. This exposes the price bounds to various types of model risk.

Many researchers use option prices to infer a state price vector for further applications in financial studies [8,9]. Breeden and Litzenberger [10] derived the prices of elementary contingent claims from options, and they call this an "inverse problem", which is the process of calculating from a set of observations to get the causal factors that produced them. By means of a state price vector, we can construct the risk-neutral probability to obtain the asset price as its expected present value which greatly eases systematic valuation errors and market risk. Several methods have been developed to extract the risk-neutral probability distribution form option prices [11–14]. In the literature, the inverse problem can be classified into two categories. One is a parametric method that uses a set of parameters for prior risk-neutral probability distribution and calculates option prices by varying the parameters to minimize the pricing error, such as expansion methods [15], generalized distribution methods [16–18], and mixture methods [19,20]. The other category is nonparametric methods, which search the risk-neutral probability distribution without a prior assumption of a specific distribution, such as maximum entropy methods [21,22], kernel methods [23] and curve-fitting methods [24–27]. The implied binomial tree of Rubinstein [11], hereinafter referred to as "IBT model", is a classical model of the nonparametric approach. As one of the basket options, the pricing and hedging of S&P 500 index option widely uses implied binomial trees [28,29]. Therefore, we follow this method.

In this article, we employ the IBT model to back out the risk-neutral probability and use the probability to replace the log-normal assumption (i.e., the normal distribution assumption for rate of returns) of the gain-loss bound. In other words, we use the market-implied distribution to calculate the option benchmark price and build up the gain-loss bounds, instead of using Black-Scholes log-normal assumption, which avoids the pricing errors resulting from the unrealistic assumption and gives a new idea of measuring market risk. To the best of our knowledge, it is the first attempt to employ the gain-loss ratio method combined with the market-implied risk neutral distribution which is calculated by a binomial tree setting. Using the S&P 500 index option for the period from January 2008 to December 2014, we estimate or update the market-implied distribution every 30 min to refresh the price bounds and review the trading opportunities. Specifically, the sampling period gives consideration to periods of global financial crisis and European sovereign debt crisis. Our empirical results show that there are option prices breaking the gain-loss bounds, even after taking into account the market information, which means that a good risk management technique and good-deal investment opportunities exist if the implied binomial tree is used as a benchmark model in the gain-loss bounds.

It is noteworthy that there are several particular features in our market-implied risk-neutral distribution. First, we avoid a subjective assumption on the underlying distribution in building up the gain-loss bounds. Second, with the modification of the benchmark model, the model risk of our gain-loss ratio bounds can be greatly reduced. Third, we have a better exploitation of the deep-out-of-money option prices by market-implied option prices that are not explained well in previous models [6,7]. Fourth, by employing the gain-loss ratio method combined with the market-implied risk neutral distribution calculated by binomial tree, we reduce the risk of distributional misspecification. According to Rubinstein's study [11], there are many violations of Black-Scholes

assumptions which are serious, difficult to remedy, and may destroy the arbitrage foundations of the Black-Scholes model. Our method provides an effective way to capture well the tail effects of the underlying asset distribution for both sides, which can provide market participants with a better model for option pricing and risk management.

The remainder of this paper is organized as follows: Section 2 presents the model based on a risk-neutral pricing kernel and sets the IBT method that can be employed to derive market-implied probability from the option data. Section 3 describes the selected S&P 500 index option data and shows the estimated market-implied gain-loss option price bounds. Section 4 offers several conclusive considerations and implications.

2. Methodology: Gain-Loss Pricing Bounds

2.1. The Properties of Gain–Loss Ratio

Gain-loss ratio is a very appealing, interesting model with many advantages [30]. We briefly review the result of the gain-loss ratio and its application for the valuation of uncertain payoffs that were developed by Bernardo and Ledoit [7]. Let $\tilde{z} = [\tilde{z}_1, \ldots, \tilde{z}_s] \in Z$ be the random payoffs and $\tilde{x} = \tilde{z} - (1 + r_f)\pi(\tilde{z})$ be the excess payoffs where $\pi(\tilde{z})$ is payoff function. We define $\tilde{x}^+ = \max(\tilde{x}, 0)$ as the positive part and $\tilde{x}^- = \max(-\tilde{x}, 0)$ as the negative part. $\tilde{b}$ is the basis assets portfolio which payoff comes close to $\tilde{z}$ from below and above, and B represents the space for basis assets. We define two subsets for B, A_1 and A_2, which can be expressed as

$$A_1 = \left\{ \tilde{b} : \frac{E^*\left[\left(\tilde{z} - \tilde{b}\right)^+\right]}{E^*\left[\left(\tilde{z} - \tilde{b}\right)^-\right]} \geq \overline{L} \right\}, A_2 = \left\{ \tilde{b} : \frac{E^*\left[\left(\tilde{b} - \tilde{z}\right)^+\right]}{E^*\left[\left(\tilde{b} - \tilde{z}\right)^-\right]} \geq \overline{L} \right\}, \tag{1}$$

where $E^*[\bullet]$ denotes the expectation value under the risk-adjusted probability measure. $E^*[\tilde{x}^+]/E^*[\tilde{x}^-]$ is the gain-loss ratio, hereinafter referred to as $\overline{L}$. And the gain-loss price bounds of a contingent claim are $\forall \tilde{z} \in Z$ subject to $\tilde{z} \notin B$

$$\max_{\tilde{b} \in B \cap A_1} \pi_B(\tilde{b}) \leq \pi(\tilde{z}) \leq \min_{\tilde{b} \in B \cap A_2} \pi_B(\tilde{b}). \tag{2}$$

As $\overline{L}$ increases, the bounds become wider; when $\overline{L}$ decreases, the bounds become narrower. In the limit, as $\overline{L}$ increases to infinity, the bounds converge to the no-arbitrage bounds; as $\overline{L}$ goes to unit, the bounds converge to the benchmark model price.

2.2. The Implied Risk-Neutral Probability and IBT Model

Rubenstein's IBT method [11] is introduced as follows. Firstly, assuming S_j for $j = 0, \ldots, n$ are the underlying asset prices at the end of the tree from lowest to highest. We denote the terminal ending nodal risk-neutral probabilities as P'_j and $\Sigma_j P'_j = 1$. P'_j can be inferred from the riskless interest rate, concurrent market prices of the underlying asset and its associated otherwise identical European options with different striking prices. Assuming the risk-neutral probability of one up-move over each period is p', then $P'_j = n! p'^j (1 - p')^{n-j} / j! (n - j)!$ can be obtained from the binomial distribution. As n increases, this probability distribution will approximately be log-normal. Let γ and δ be the riskless interest return and the underlying asset payout return over each period, respectively. Finally, assuming S^b to be the current bid price of the underlying asset adjusted with dividend payouts and S^a to be the current ask price of the underlying asset adjusted with dividend payouts. C_i^b for $i = 1, \ldots, m$ is the bid price at the end of the tree and C_i^a for $i = 1, \ldots, m$ is the ask price at the end of the tree. The implied posterior risk-neutral probabilities, P_j, can be obtained from the following quadratic minimizing problem:

$$\min_{P_j} \sum_j \left(P_j - P'_j\right)^2 \text{ subject to}$$

$$\sum_j P_j = 1 \text{ and } P_j \geq 0 \text{ for } j = 0, \ldots, n$$

$$S^b \leq S \leq S^a \text{ where } S = \sum_j P_j S_j / r^n \qquad (3)$$

$$C_i^b \leq C_i \leq C_i^a \text{ where } C_i = \sum_j P_j \max[0, S_j - K_i]/r^n \text{ for } i = 1, \ldots, m$$

where K_i is the strike price for the option traded in the market. The sum of posterior risk-neutral probabilities P_j is 1 and comes closest to a prior log-normal guessing. P_j makes the present values of all observed options fall between the bid and ask prices. Although we adopted a specific minimization function and a specific prior, the optimization method is quite flexible. As long as a solution exists and the number of probabilities n is greater than the number of options m, the solution will depend on the prior and minimization function chosen. The least squares form of the minimization function is just one of a number of candidates [11] (For instance, the Hellinger Distance is also a good candidate. The Hellinger Distance is defined in terms of the Hellinger integral, which is used to quantify the similarity between two probability distributions. In practice, the quadratic programming is much easier to perform than minimizing Hellinger Distance). Since our purpose is to employ the implied binomial trees model into the framework of gain-loss price bounds, we will remain on our focus. Besides, it is well known that none of the asset-pricing models can perfectly fit the real fluctuation of options, even the IBT model. A lot of factors cannot be adopted by the model, such as significant transactions costs or restrictions on short selling. However, in comparison to the BS model, IBT model has huge improvements in parameter risk control and market simulating, and it is the optimal benchmark model when you try to fit a basket option [28,29].

2.3. Our Methodology: Implied Gain-Loss Option Pricing Bounds

In the numerical example of Bernardo and Ledoit [7], the authors used the Black-Scholes risk-adjusted probability as the risk-neutral probability. Our model replaces the benchmark pricing kernel by an objective market-implied density abstracting from option transaction data. Here, we apply IBT's implied risk-neutral distribution. We define two subsets for B, B_1 and B_2, which can be expressed as

$$B_1 = \left[b_i : \frac{\sum_{i=1}^{I} P_i (b_i - C_i)^+}{\sum_{i=1}^{I} P_i (b_i - C_i)^-} \geq L \right], B_2 = \left[b_i : \frac{\sum_{i=1}^{I} P_i (C_i - b_i)^+}{\sum_{i=1}^{I} P_i (C_i - b_i)^-} \geq L \right], \qquad (4)$$

where b_i is the payoff in the ith simulation of the replicating portfolio of basis assets with weight w_0 on the risk-free bond and weight w_1 on the option on the traded asset. And the IBT gain-loss price bounds of a contingent claim are

$$\max_{\substack{w_0, w_1 \in R \\ b_i = w_0 + w_1 S_i \\ b_i \in B_1}} w_0 e^{-rt} + w_1 S \leq C \leq \min_{\substack{w_0, w_1 \in R \\ b_i = w_0 + w_1 S_i \\ b_i \in B_2}} w_0 e^{-rt} + w_1 S. \qquad (5)$$

And the benchmark call options price can be calculated as

$$C = \sum_{i=1}^{I} P_i (S_i - K)^+ e^{-rt}. \qquad (6)$$

As we can see, the study of Bernardo and Ledoit [7] proposed their price bounds under the Black-Scholes risk-adjusted probabilities, which involves many restricted assumptions that have been violated in the real world. Since the implied binomial trees model can provide a computationally effective way to value options even in the presence of various violations of assumptions [11], our modification incorporates more information which can greatly reduce the model risk of gain-loss pricing bounds. The advantages of this modification are (a) the implied gain-loss pricing bounds are

parameter-free that directly conform to market data and avoid an incorrect model specification and (b) the market price of options incorporates the investor's preference and sentiments on the current market. By applying a market-implied pricing kernel, we can capture the investor's risk-taking attitude more sensitively than the pricing kernel of model-based approaches, and the extreme case is more likely to occur in the market-implied distribution which greatly reduces the market risk.

It is noticeable that, even after using the market-implied distribution, there still exits a pricing gap between the benchmark prices and market prices. First, in the above programming program, the IBT method tries to obtain an "average probability" over all strike prices at one time. This means that, in Equation (3), P_j is an "average result" over various strike price K_i. The risk-neutral probability P_j has to coordinate the pricing errors of all the market prices with various strike prices. Therefore, when we use one market-implied probability to calculate option prices, there exists estimate errors or pricing biases across options with different strike prices. (This issue is similar to using a flat volatility curve to price all options, instead of using a volatility smile curve). In other words, we may have various risk-neutral probabilities, P_j, for options with different strike prices. Here we use one risk-neutral probability, P_j so that there exists some pricing gaps between real transaction prices and benchmark prices. Second, our model indeed suffers a small downside risk. Based on our model and the reality, the downside risk comes from the severe deviations from the mean-variance framework, which may result from market imperfections such as significant transactions costs, restrictions, and margins on short selling, taxes, non-competitive pricing, and other non-controllable events. Third, owing to the computational time, it is impossible to continuously calculate the market-implied density in practice. We therefore estimate the market-implied density every 30 min. There are two reasons that we update our market-implied distributions every 30 min. First, price fluctuations are quite normal in index option markets. If we update the distribution too frequently, it may incorporate too much redundant information. Second, nowadays, as we can gather the information from a more distant place than ever, the market prices of the index option may have different patterns even in a single day [31], and that is why Rubinstein [11] updated his model three times a day in his paper. To conduct analyses of S&P 500 index options markets, some researchers updated average volume every 30 min [32], some researchers reported autocorrelations for changes in the mean and the standard deviation of risk neutral density measured over different time intervals, from 1 min to 30 min [33], and some researchers studied the relationship between information and price change by applying the 30 min interval [34]. Considering the tradeoff between robustness and effectiveness, the update frequency has to be set as neither too short nor too long. Following the previous literature, we chose 30 min as our model update frequency. Distinguishing volatility estimating, our method uses past or lag market data to estimate the whole distribution instead of estimating only one parameter. Therefore, we have a priori estimate on the distribution, and we then process the out-of-sample test every 30 min. The details are described in Section 3.

3. Data and Estimated Risk-Neutral Distribution

This section describes how the implied risk-neutral probabilities are obtained from the market trading data. The market trading data of the S&P 500 index options for both the call and put options listed in the Chicago Board Options Exchange are collected from 2 January 2008 to 30 December 2014. Most of previous literature concerning gain-loss ratio used simulated data [35,36]. Few researchers adopt the real trading data which are low frequency and single variety [37]. Comparing to these studies, our data are quite rich and varied by using as many as a hundred simultaneously traded options on the S&P 500 index, all differing in strike price and time to expiration. We employ the MATLAB optimization procedure to perform the programming problem in Equation (3). We calculate the risk-neutral probabilities P_j of the S&P 500 index options for 200, 400, and 800 steps.

3.1. Market-Implied Distributions

The market-implied distributions are estimated every 30 min during the sample period. We plot two examples of the risk-neutral probabilities distribution of the S&P 500 index options on 5 August 2011 (maturing in 17 March 2012) and 5 August 2008 (maturing in 21 March 2009) in Figures 1 and 2, respectively. The distributions estimated from call options are depicted on the left side, and the distributions estimated from put options are depicted in the right side. These implied probabilities are plotted using 200, 400, and 800-step binomial trees. As we can see, in Figure 1, the choice of step of binomial tree does not greatly affect the distribution shapes. However, there is a sharp distinction between the risk-neutral initial guess (normal distribution) and the market-implied risk-neutral probability distribution. The risk-neutral implied distribution is more skewed and leptokurtic. Besides, the left tail is non-smooth, which means the extreme case is more likely to occur in market-implied distribution. Nowadays, black swan events appear more than ever, such as Snowden case, fiscal cliff, government shut down and so on. So, the left tail of the distribution tends to be thicker. Rubinstein [11] also found the same phenomenon too. He argued that the risk-neutral implied posterior distribution is slightly bimodal and more highly skewed and kurtotic and the bimodality coming from the lower tail ("crash-o-phobia") is quite common during the post-crash period. Recently, the market risk has become a focus of market participants due to spectacular bankruptcies like the Baring's Bank or the investment bank Lehman Brothers. As a result, the risk-neutral implied distribution has less market risk than the normal distribution. Figure 2 plots similar results for the S&P 500 options on 5 August 2008.
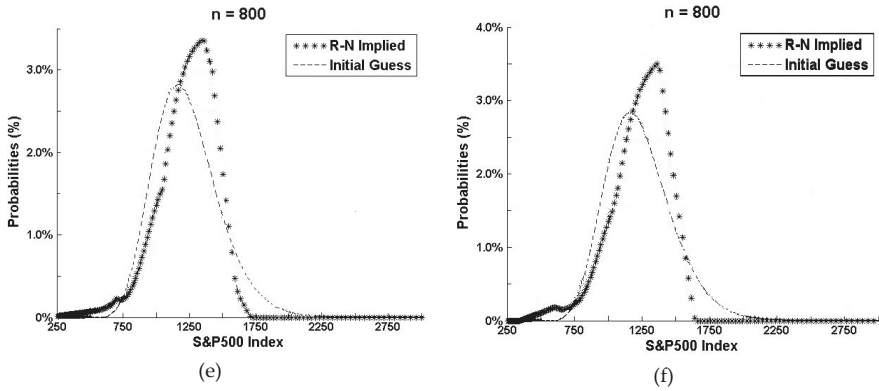

Figure 1. *Cont.*

(e)

(f)
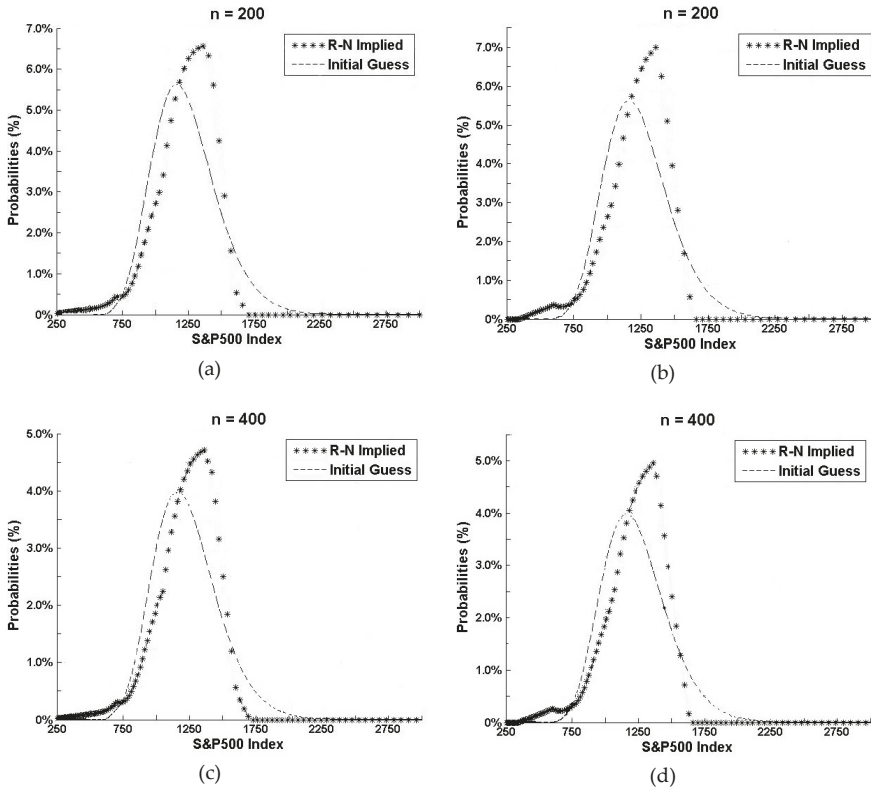
Figure 1. The risk-neutral probabilities of S&P 500 index options on 5 August 2011. The three figures on the left side (**a**,**c**,**e**) are probability distributions for call options using 200, 400, and 800-step binomial trees. The three figures on the right side (**b**,**d**,**f**) are probability distributions for put options using 200, 400, and 800-step binomial trees.
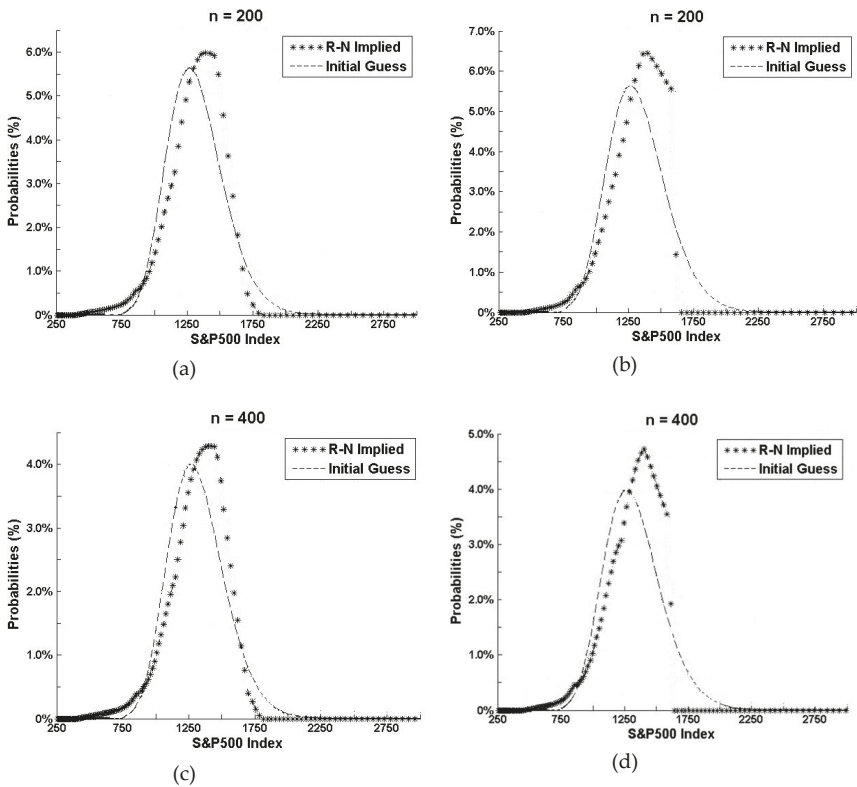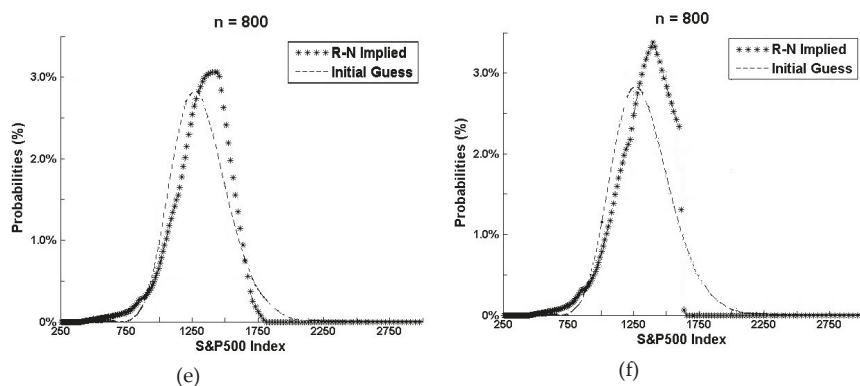

(a)

(b)

(c)

(d)

Figure 2. *Cont.*

Figure 2. The risk-neutral probabilities of S&P 500 index options on 5 August 2008. The three figures on the left side (**a**,**c**,**e**) are probability distributions for call options using 200, 400, and 800-step binomial trees. The three figures on the right side (**b**,**d**,**f**) are probability distributions for put options using 200, 400, and 800-step binomial trees.

3.2. Market-Implied Option Pricing Bounds

In Section 2, we propose the gain-loss price bounds. The option prices must lie between the bounds of Equation (5) to avoid a semi-arbitrage opportunity. This section shows how to find the option price bounds under given $\overline{L}$. Before calculating the gain-loss ratio bound, two subjective decisions have to be made. First is choosing a benchmark model to build up the theoretical prices. We employ both the Black-Scholes model (log-normal distribution) and the IBT model (market-implied distribution) to calculate benchmark prices. The other decision is to choose the value of gain-loss ratio, $\overline{L}$, to set up the gain-loss bounds under an incomplete market. We use $\overline{L}$= 3, 5, and 10, respectively.

Figure 3 plots the gain-loss price bounds of call and put options using the Black–Scholes model as the benchmark model under $\overline{L} = 1, 2, \ldots, 10$ and $\overline{L} = 200$, respectively. These options have a time to maturity of 225 days, a risk-free interest rate of 0.11%, a strike price of $1100, and a volatility of 25.38%. The values of price bounds are reported in Appendix A Table A1. Different from Figure 3, Figure 4 plots the price bounds of call and put options using the IBT model as a benchmark model. The values of price bounds are reported in Appendix A Table A2.

As shown in both figures, the smaller the $\overline{L}$, the tighter the price bounds. As $\overline{L}$ is 1, the upper bounds and lower bounds converge to the benchmark prices, and as $\overline{L}$ is 200, they converge to the no-arbitrage bounds. There are two main fundamental ideas that are explored in asset pricing: model-based pricing and no-arbitrage pricing. Model-based pricing can derive the exact price of the option and no-arbitrage pricing can derive the range of option price. The gain-loss price bounds propose a framework to unify the model-based pricing and no-arbitrage pricing by introducing additional restrictions on gain-loss ratio. It means that if $\overline{L}$ goes to one (its lower bound), the admissible set shrinks to model-based pricing containing only the benchmark pricing kernel. If $\overline{L}$ goes to infinity, the admissible set grows to no-arbitrage pricing including all pricing kernels consistent with the absence of arbitrage among the basis assets. Besides, the price bounds get wider for the at-the-money option because the at-the-money options are the least redundant ones. As shown in Figure 4, we choose options having a time to maturity of 225 days, which makes the analyses incorporate more price fluctuations, like deep-out-of-money price. As we can see in Figures 3 and 4, comparing to initial Black-Scholes gain-loss price bounds, the IBT gain-loss price bounds are narrower for call options in deep-out-of-money and wider for put options in deep-out-of-money, which shows different price fluctuations in market-implied distributions and gives investors a total different market trading choice to avoid risks and get returns. Taking the price of the put option as an example, when the option

prices are deep-out-of-money, investors will get a lower upper bound if they use Black-Scholes model as a benchmark model. Consequently, it makes them suffer a great risk generated from the rising prices of options due to the inaccurate estimate of the gain-loss price bounds. We also see that the price bounds in Figure 4 are wider than the bounds in Figure 3. The gain-loss price bounds using market-implied distribution are wider than the bounds of a log-normal distribution. This observation is consistent with our expectations. Because the market price of options reflects the traders' opinions and sentiments, it makes the implied distribution more sensitive than the Black–Scholes' log-normal assumption, which makes the implied distribution have less market risk. This phenomenon can also be explained by the shape differences between the market-implied distributions and the initial guess in Figures 1 and 2. The market-implied distribution is more skewed and leptokurtic than the normal distributions. Therefore, the extreme case is more likely to occur in market-implied distribution, and it results in wider gain-loss bounds.

(a)

(b)
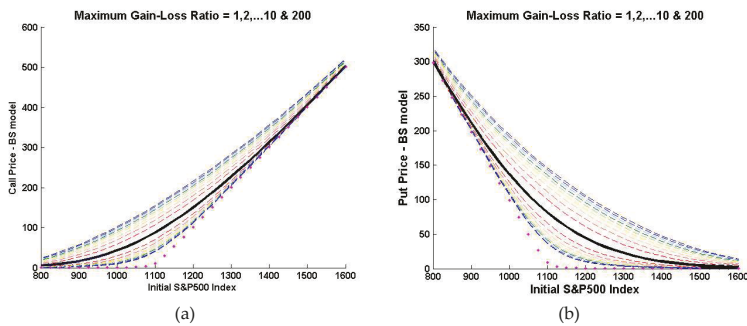
Figure 3. The gain-loss price bounds for both call options (**a**) and put options (**b**) by using Black-Scholes model as a benchmark model. The benchmark model is chosen to yield the Black-Scholes price given by the thick line in black. The thin lines in different colors represent the upper and lower bounds obtained by ruling out approximate arbitrage. For example, the red thin line represents the price bounds when gain-loss ratio equals to 2. The orange thin line represents the price bounds when gain-loss ratio equals to 3 and so on, for other seven thin lines. The dotted line represents the bound obtained by ruling out pure arbitrage (gain-loss ratio equals to 200).
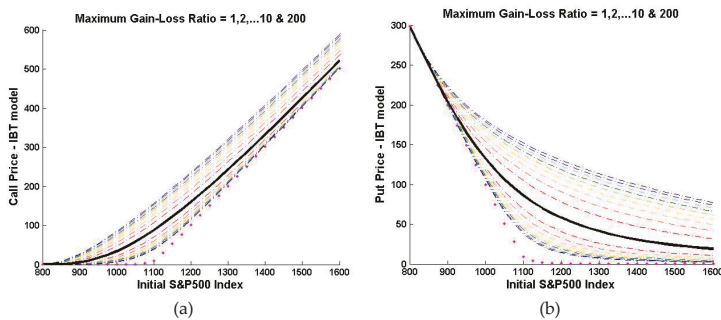

(a)

(b)

Figure 4. The gain-loss price bounds for both call options (**a**) and put options (**b**) by using IBT model as a benchmark model. The benchmark model is chosen to yield the IBT price given by the thick line in black. The thin lines in different colors represent the upper and lower bounds obtained by ruling out approximate arbitrage. For example, the red thin line represents the price bounds when gain-loss ratio equals to 2. The orange thin line represents the price bounds when gain-loss ratio equals to 3 and so on, for other seven thin lines. The dotted line represents the bound obtained by ruling out pure arbitrage (gain-loss ratio equals to 200).

3.3. The Good Deal Trading Strategy

Following the gain-loss bounds in Section 3.2, our trading strategies are selling the options whose prices are higher than the upper bounds and buying the options whose prices are smaller than the lower bounds with a complete buy and sell trade transaction cost of 1% [38]. The key parameter in the Black–Scholes formula, volatility, is updated every 30 min in each trading day. (We use three models to forecast daily volatility: GARCH, EGARCH, and GJR-GARCH. The choice of volatility model does not greatly affect the bounds and strategy performances, so we only show the results of GARCH. Other results can be provided upon request). The market-implied distribution is also updated or re-estimated at the same frequency. Given the market-implied distribution, the theoretical price of IBT model can be calculated by Equation (5), and P_j and S_j are obtained from the minimizing programming in Equation (3). The buying strategy is when the price smaller than the lower bound implied by $\overline{L}$ in Equation (5), we buy the option. The selling strategy is when the price is higher than the upper bound implied by $\overline{L}$ in Equation (5), we sell the option. The investment period is once the option purchased or sold, we hold the position to maturity day.

Table 1 shows the annual return of the selling strategy subjected to $\overline{L} = 3, 5$, and 10. The benchmark price is calculated by the Black–Scholes model and the IBT model, respectively. In Table 1, the positive returns mean that the selling strategy achieves positive returns for both strategies from 2008 to 2014. Bernardo and Ledoit [7] argued that the attractiveness of an investment opportunity is measured by the "gain-loss" ratio, which is the expectation of the investment's positive excess payoffs divided by the expectation of its negative excess payoffs. When applying our IBT price bounds to the real world, there is an important thing we need to consider: the trading opportunities. In theory, a bigger gain-loss ratio causes bigger rates of returns. But a bigger gain-loss ratio means the expectation of the investment's positive excess payoffs exceed the expectation of its negative excess payoffs far more, which appear rarely in the option markets. As a result, there is a tradeoff between the gain-loss ratio and trading opportunities. We therefore use cumulative returns to show the net returns during the whole investment period. The cumulative returns are shown in the last row in both Tables 1 and 2. Based on the empirical results, = 10 has the highest cumulative returns in average. Thus, it is the best in reality. It means that the IBT model incorporates more market information than the Black-Scholes model which results in having less market risk and better investment performance. These findings support the statement that the IBT gain-loss price bounds are more useful than the Black–Scholes gain-loss price bounds in risk management and return performance. As mentioned above, implied binomial trees model can provide a computationally effective way to value options even in the presence of violations of Black-Scholes model's assumptions. Therefore, our IBT gain-loss price bounds will provide a better model for option pricing than initial gain-loss pricing bounds.

Table 1. Annual return of the selling strategy subjected to different $\overline{L}$ using Black–Scholes and IBT as benchmark models.

Year	Annual Return under BS Bounds			Annual Return under IBT Bounds		
	$\overline{L} = 3$	$\overline{L} = 5$	$\overline{L} = 10$	$\overline{L} = 3$	$\overline{L} = 5$	$\overline{L} = 10$
2008	3.41%	4.52%	7.61%	10.06%	13.21%	14.88%
2009	1.45%	1.54%	1.68%	4.50%	4.75%	5.63%
2010	0.55%	1.35%	1.52%	2.66%	3.15%	4.27%
2011	3.75%	4.51%	6.69%	6.99%	6.77%	7.64%
2012	2.75%	3.95%	5.07%	6.71%	5.28%	10.58%
2013	0.19%	0.44%	0.67%	0.71%	0.97%	2.89%
2014	1.41%	2.08%	3.21%	1.71%	3.09%	4.34%
Cumulative returns	14.25%	19.81%	29.38%	38.08%	43.12%	61.68%

Figure 5 shows the return distributions of all selling opportunities under $\overline{L} = 3, 5$, and 10 from 2008 to 2014. In Figure 5, as we increase the gain-loss ratio, $\overline{L}$, the distribution shifts to the right,

the returns increase, and the frequency of extreme losses decreases. However, there still exists the result of some negative returns, which means that this strategy still suffers risk in extreme cases with very small probabilities.
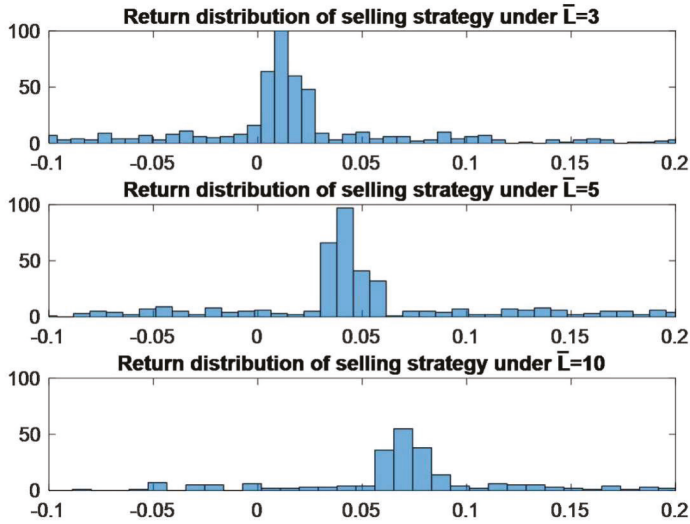

Figure 5. Return distribution of selling strategy under IBT gain–loss bounds.

Table 2 shows the annual returns of a buying strategy subjected to different $\overline{L}$. The benchmark prices are calculated by the Black–Scholes model and the IBT model, respectively. As shown in Table 2, there are many blanks in the annual return, which means no trading opportunities are available from 2008 to 2014. The trading opportunities only exist in the gain-loss bound with $\overline{L} = 3$ and $\overline{L} = 5$. No trading opportunity exists in $\overline{L} = 10$ for both methods.

Table 2. Annual return of the buying strategy subjected to different $\overline{L}$ using Black–Scholes and IBT as benchmark models.

	Annual Return under BS Bounds		Annual Return under IBT Bounds	
Year	$\overline{L} = 3$	$\overline{L} = 5$	$\overline{L} = 3$	$\overline{L} = 5$
2008	91.19%	51.86%	182.42%	196.10%
2009	−7.09%	-	0.58%	-
2010	−59.88%	-	−40.64%	-
2011	−44.23%	-	−23.62%	-
2012	−73.01%	-	−51.93%	-
2013	−62.72%	−65.38%	−55.48%	−39.39%
2014	−0.85%	-	−2.62%	-
Cumulative returns	−96.03%	−47.43%	−73.16%	79.47%

Note: - means no trading opportunity exists.

Figure 6 shows the daily returns of a buying strategy subjected to different gain-loss ratios from 2008 to 2014. In Figure 6, on increasing the gain-loss ratio, the distribution shifts to the right, and the returns increase. However, the trading opportunities decrease sharply. We can also observe that most of the returns are about −100%, which means that the investor loses all the option premiums in the buying strategy. In this vein, we cannot confirm that the buying strategy is a good-deal strategy as the selling strategy dose.
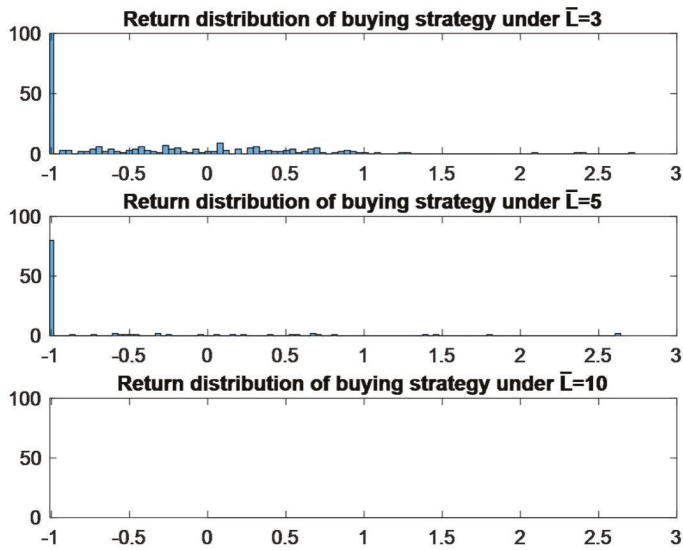
Figure 6. Return distribution of buying strategy under IBT gain-loss bounds.

In addition, there are some concerns when applying IBT price bounds in the real world. Firstly, the trading strategies in Table 1 are selling the options (put options or call options) whose prices are higher than the upper bounds and hold the position to the maturity. When applying our IBT price bounds to the real world, the trading opportunities need to be considered. For instance, if the gain-loss ratio is identical for various years, the year that the price fluctuates more frequently will have more trading opportunities, and thus, more returns. As a result, in the year 2008, the market was in more panic than ever, and the option prices fluctuated more than the following two years, which meant more trading opportunities and more returns. Besides, during the period of the Financial Crisis in the year 2008, due to the sheep-flock effect, the deviation of options prices from the benchmark prices became larger, which made the IBT gain-loss price bounds have more extra returns than the following two years. Secondly, the market prices of options reflect the traders' opinions and sentiments, and it makes the implied distribution more sensitive than the log-normal assumption of Black–Scholes model. Since IBT bounds are much wider than the BS bounds, some reasonably good option prices based on BS bounds may not be considered in that way in our IBT bounds. These somewhat good prices in BS bounds trigger a buying signal for the investors, consequently, causing them to lose money. For instance, due to more information has been considered in our model, it gives better buying signals than that of the BS bounds in the year 2009.

In sum, the selling strategy has a good performance on average return and the buying strategy has fewer trading opportunities but a poor performance on average return. After filtering by the gain-loss bounds of both strategies, the selling strategy provides a better annual return then the naïve buy-and-hold strategy. Finally, the gain-loss bounds of IBT method outperform the bounds of the Black-Scholes method in the sampling period.

4. Conclusions

In this paper, we compute the pricing bounds on the S&P 500 index option based on a market-implied distribution and a gain-loss restriction. We use the IBT model to refine the market-implied distribution, which can have less market risk for proceeding in efficient risk management. The estimated risk-neutral pricing kernel implicitly reflects the investor's preference and sentiment, which can reconcile the latest information in the market and reduce the market risk.

The empirical results show that when we increase the restriction on gain-loss ratio, the return of the strategies can approximately increase. We also confirm that the gain-loss bounds of the IBT method outperform the gain-loss bounds of the Black-Scholes model, which means the IBT model incorporates more market information than the Black-Scholes model; consequently, it results in having less market risk and better investment performance. Finally, the selling strategy has been evidenced to have a better performance than the buying strategy.

Several limitations to this study must be mentioned. First, the gain-loss bound we used might still suffer parameter and model risks. For example, the stochastic process in the stock index involves jumps that are not directly characterized in our model. We infer the jump probability by the option data and the implied risk-neutral probability, indirectly. However, the underlying assets in this study (i.e., S&P 500) are less likely to experience jumps than any of their component equities, and most other underlying assets such as commodities, currencies, and bonds [11]; how to design an IBT model to handle jumps; or stochastic volatility would be an interesting agenda and could be left for future studies. Second, there exists a trade-off between the trading opportunity and rate of return. When we raise $\overline{L}$, the return increases; however, the trading opportunity decreases sharply. When we decrease $\overline{L}$, the return decreases; however, the trading opportunity increases. The optimal trade-off strategy between absolute return and trading frequency is not explored in this article. Third, as shown in Figure 5, a small downside risk in the selling strategy still suffers, which cannot be mitigated or eliminated by the IBT gain-loss bound. To sum up, the IBT model still suffers from misspecification error but is superior to the BS model, while the no-arbitrary pricing model is robust in asset pricing but is too imprecise to be economically interesting. In this paper, the IBT gain-loss price bounds can greatly reduce the weaknesses of IBT model and no-arbitrary pricing model.

On the suggestion for further works, we propose that the objective function in Equation (3) could be revised. For example, we can minimize the difference of an option's implied volatility of Black-Scholes model instead of minimizing the distance of the probability distributions. The implied volatility can avoid the disturbances of size effects of option prices resulting from different exercise prices and time to maturities. Finally, a subjective method to decide the ceiling value of $\overline{L}$ is also worthwhile for further research. The problem due to the arbitrary chosen ceiling value of $\overline{L}$ could be left for future studies.

Author Contributions: Shi-Jie Jiang and Mujun Lei conceived the empirical study, contributed analysis tools and finalized the paper; Cheng-Huang Chung designed and performed the simulation study and analyzed the data. All authors read and approved the final manuscript.

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

Appendix

The appendix illustrates the detailed number of price bound plotted in Figures 3 and 4. We only show the call option price bounds, and the tables of put option can be provided upon request.

Table A1. Call Option Price Bounds on 5 August 2011 using Black-Scholes as benchmark model.

Upper Bound

$\bar{L}$	S_0 800	825	850	875	900	925	950	975	1000	1025	1050	1075	1100	1125	1150	1175	1200
1	4.37	6.42	9.12	12.43	16.51	21.50	27.52	34.66	42.93	52.40	63.09	74.96	87.93	102.00	117.13	133.25	150.36
2	7.92	11.25	15.46	20.47	26.29	33.04	40.75	49.46	59.14	69.80	81.41	93.93	107.28	121.46	136.43	152.16	168.65
3	10.91	15.15	20.32	26.36	33.27	41.00	49.62	59.13	69.49	80.70	92.74	105.55	119.08	133.31	148.22	163.78	180.00
8	21.38	28.06	35.59	43.89	52.95	62.75	73.25	84.41	96.18	108.54	121.45	134.90	148.84	163.26	178.15	193.50	209.31
9	22.94	29.92	37.72	46.28	55.54	65.54	76.22	87.53	99.44	111.90	124.92	138.44	152.42	166.88	181.79	197.13	212.92
10	24.39	31.62	39.68	48.45	57.92	68.08	78.90	90.35	102.36	114.93	128.02	141.60	155.64	170.12	185.05	200.39	216.17
200	81.97	94.18	106.82	119.75	132.98	146.53	160.39	174.53	188.96	203.64	218.57	233.74	249.13	264.74	280.56	296.58	312.80
Inf	800.00	825.00	850.00	875.00	900.00	925.00	950.00	975.00	1000.00	1025.00	1050.00	1075.00	1100.00	1125.00	1150.00	1175.00	1200.00

Lower Bound

$\bar{L}$	S_0 800	825	850	875	900	925	950	975	1000	1025	1050	1075	1100	1125	1150	1175	1200
1	4.37	6.42	9.12	12.43	16.51	21.50	27.52	34.66	42.93	52.40	63.09	74.96	87.93	102.00	117.13	133.25	150.36
2	2.32	3.49	5.07	7.08	9.70	13.07	17.39	22.84	29.52	37.61	47.21	58.36	70.99	85.11	100.66	117.49	135.58
3	1.58	2.40	3.51	4.96	6.90	9.45	12.85	17.29	22.94	30.08	38.92	49.54	61.92	76.09	91.95	109.31	128.07
8	0.61	0.94	1.39	1.99	2.84	4.00	5.67	8.02	11.30	16.01	22.62	31.62	43.29	57.77	74.81	93.87	114.54
9	0.54	0.84	1.24	1.78	2.54	3.58	5.11	7.25	10.29	14.71	21.03	29.80	41.37	55.91	73.13	92.42	113.35
10	0.49	0.75	1.12	1.61	2.30	3.25	4.64	6.62	9.45	13.62	19.68	28.23	39.71	54.31	71.70	91.20	112.35
200	0.02	0.04	0.06	0.08	0.12	0.17	0.26	0.38	0.58	0.95	1.71	3.63	10.90	29.28	52.60	76.85	101.49
-Inf	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.75	25.75	50.75	75.75	100.75

Upper Bound

$\bar{L}$	S_0 1225	1250	1275	1300	1325	1350	1375	1400	1425	1450	1475	1500	1525	1550	1575	1600
1	168.37	187.22	206.80	227.06	247.90	269.26	291.15	313.49	336.25	359.35	382.77	406.45	430.34	454.40	478.64	503.01
2	185.85	203.74	222.26	241.38	261.05	281.23	301.96	323.18	344.87	366.99	389.51	412.38	435.54	458.96	482.65	506.56
3	196.84	214.28	232.28	250.83	269.88	289.42	309.48	330.03	351.07	372.55	394.46	416.73	439.34	462.28	485.54	509.07
8	225.57	242.27	259.39	276.95	294.90	313.25	332.03	351.23	370.80	390.77	411.16	431.98	453.19	474.78	496.75	519.08
9	229.14	245.80	262.87	280.34	298.19	316.44	335.10	354.14	373.56	393.38	413.62	434.28	455.33	476.76	498.57	520.72
10	232.37	248.99	266.00	283.40	301.19	319.34	337.89	356.79	376.09	395.78	415.90	436.42	457.33	478.62	500.28	522.27
200	329.21	345.83	362.62	379.59	396.73	414.04	431.55	449.21	467.00	484.96	503.08	521.35	539.78	558.32	577.02	595.99
Inf	1225.00	1250.00	1275.00	1300.00	1325.00	1350.00	1375.00	1400.00	1425.00	1450.00	1475.00	1500.00	1525.00	1550.00	1575.00	1600.00

Lower Bound

$\bar{L}$	S_0 1225	1250	1275	1300	1325	1350	1375	1400	1425	1450	1475	1500	1525	1550	1575	1600
1	168.37	187.22	206.80	227.06	247.90	269.26	291.15	313.49	336.25	359.35	382.77	406.45	430.34	454.40	478.64	503.01
2	154.74	174.85	195.73	217.30	239.42	261.96	284.95	308.29	331.93	355.81	379.90	404.16	428.56	453.06	477.65	502.32
3	147.98	168.89	190.51	212.80	235.59	258.72	282.22	306.00	330.02	354.23	378.59	403.07	427.66	452.31	477.03	501.81
8	136.35	159.04	182.27	205.91	229.88	254.03	278.36	302.84	327.43	352.11	376.85	401.64	426.48	451.34	476.23	501.15
9	135.37	158.23	181.62	205.38	229.45	253.68	278.08	302.62	327.25	351.96	376.73	401.54	426.40	451.28	476.18	501.10
10	134.55	157.57	181.09	204.95	229.10	253.40	277.86	302.44	327.10	351.84	376.63	401.47	426.33	451.22	476.14	501.07
200	126.27	151.13	176.04	200.97	225.92	250.88	275.86	300.83	325.81	350.80	375.79	400.78	425.78	450.77	475.76	500.76
-Inf	125.75	150.75	175.75	200.75	225.75	250.75	275.75	300.75	325.75	350.75	375.75	400.75	425.75	450.75	475.75	500.75

* The no-arbitrage upper/lower bounds are $\tilde{S}_0$ and $\max(\tilde{S}_0 - e^{-rt}K, 0)$ for call option when threshold $\bar{L} = \infty$.

Table A2. Call Option Price Bounds on 5 August 2011 using IBT as benchmark model.

Upper Bound

$\overline{L}$ \ S_0	800	825	850	875	900	925	950	975	1000	1025	1050	1075	1100	1125	1150	1175	1200
1	0.15	0.44	1.20	2.84	5.74	10.21	16.43	24.41	34.09	45.35	58.20	72.38	87.78	104.43	122.03	140.60	159.92
2	0.30	0.86	2.29	5.20	9.96	16.68	25.33	35.74	47.58	60.80	75.23	90.68	107.06	124.40	142.46	161.25	180.63
3	0.45	1.26	3.29	7.18	13.33	21.55	31.64	43.30	56.41	70.64	85.92	102.09	119.07	136.86	155.28	174.33	193.90
8	1.15	3.09	7.23	14.48	24.19	35.88	49.25	63.87	79.62	96.10	113.33	131.20	149.64	168.65	188.00	207.98	228.19
9	1.29	3.42	7.90	15.56	25.70	37.72	51.60	66.40	82.44	99.25	116.75	134.83	153.44	172.59	192.16	212.14	232.46
10	1.42	3.74	8.56	16.54	27.05	39.52	53.66	68.83	84.93	101.99	119.72	138.00	156.77	176.05	195.71	215.77	236.14
200	16.08	26.26	40.19	56.78	74.11	93.21	112.69	132.54	153.10	174.27	195.70	217.33	239.13	261.11	283.23	305.48	327.84
Inf	800.00	825.00	850.00	875.00	900.00	925.00	950.00	975.00	1000.00	1025.00	1050.00	1075.00	1100.00	1125.00	1150.00	1175.00	1200.00

Lower Bound

$\overline{L}$ \ S_0	800	825	850	875	900	925	950	975	1000	1025	1050	1075	1100	1125	1150	1175	1200
1	0.15	0.44	1.20	2.84	5.74	10.21	16.43	24.41	34.09	45.35	58.20	72.38	87.78	104.43	122.03	140.60	159.92
2	0.08	0.22	0.61	1.49	3.13	5.86	9.90	15.56	22.92	32.06	43.16	56.00	70.44	86.59	104.02	122.70	142.35
3	0.05	0.15	0.41	1.01	2.15	4.11	7.16	11.60	17.60	25.46	35.38	47.38	61.29	77.21	94.66	113.59	133.61
8	0.02	0.06	0.16	0.39	0.84	1.65	3.03	5.20	8.41	13.12	20.21	29.90	42.54	58.22	76.23	96.26	117.55
9	0.02	0.05	0.14	0.34	0.75	1.47	2.71	4.70	7.61	12.00	18.71	28.22	40.56	56.36	74.42	94.64	116.11
10	0.02	0.04	0.12	0.31	0.67	1.33	2.45	4.30	6.96	11.07	17.46	26.67	38.90	54.67	72.91	93.31	114.93
200	0.00	0.00	0.01	0.02	0.03	0.07	0.13	0.24	0.43	0.68	1.40	3.61	10.10	29.62	52.58	77.08	101.64
-Inf	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.75	25.75	50.75	75.75	100.75

Upper Bound

$\overline{L}$ \ S_0	1225	1250	1275	1300	1325	1350	1375	1400	1425	1450	1475	1500	1525	1550	1575	1600
1	179.95	200.52	221.61	243.11	265.02	287.28	309.84	332.68	355.78	379.07	402.53	426.14	449.89	473.74	497.69	521.72
2	200.56	220.95	241.78	262.97	284.50	306.36	328.24	350.89	373.53	396.36	419.35	442.50	465.80	489.21	512.72	536.31
3	213.93	234.35	255.16	276.30	297.76	319.51	341.50	363.71	386.14	408.75	431.53	454.45	477.51	500.69	523.94	547.28
8	248.81	269.64	290.73	312.00	333.50	355.20	377.04	399.14	421.39	443.78	466.30	488.93	511.65	534.45	557.35	580.32
9	253.08	273.93	295.04	316.36	337.90	359.64	381.57	403.67	425.94	448.33	470.83	493.40	516.12	538.90	561.78	584.73
10	256.81	277.73	298.90	320.28	341.88	363.67	385.64	407.77	430.04	452.40	474.87	497.46	520.15	542.93	565.80	588.74
200	350.31	372.84	395.46	418.13	440.86	463.64	486.46	509.33	532.23	555.16	578.12	601.11	624.14	647.19	670.27	693.37
Inf	1225.00	1250.00	1275.00	1300.00	1325.00	1350.00	1375.00	1400.00	1425.00	1450.00	1475.00	1500.00	1525.00	1550.00	1575.00	1600.00

Lower Bound

$\overline{L}$ \ S_0	1225	1250	1275	1300	1325	1350	1375	1400	1425	1450	1475	1500	1525	1550	1575	1600
1	179.95	200.52	221.61	243.11	265.02	287.28	309.84	332.68	355.78	379.07	402.53	426.14	449.89	473.74	497.69	521.72
2	162.86	183.98	205.69	227.81	250.35	273.24	296.40	319.82	343.46	367.26	391.21	415.28	439.47	463.74	488.08	512.48
3	154.55	176.14	198.31	220.88	243.85	267.14	290.66	314.41	338.36	362.44	386.64	410.94	435.34	459.81	484.32	508.89
8	139.86	162.74	186.13	209.75	233.66	257.78	282.05	306.46	330.96	355.53	380.16	404.85	429.59	454.36	479.16	503.97
9	138.58	161.60	185.12	208.87	232.87	257.07	281.39	305.86	330.42	355.03	379.69	404.41	429.18	453.97	478.79	503.62
10	137.53	160.67	184.28	208.14	232.22	256.48	280.86	305.37	329.98	354.62	379.31	404.05	428.85	453.66	478.50	503.35
200	126.48	151.31	176.23	201.17	226.10	251.06	276.02	300.99	325.97	350.95	375.93	400.92	425.90	450.90	475.89	500.88
-Inf	125.75	150.75	175.75	200.75	225.75	250.75	275.75	300.75	325.75	350.75	375.75	400.75	425.75	450.75	475.75	500.75

* The no-arbitrage upper/lower bounds are $\tilde{S}_0$ and $\max(\tilde{S}_0 - e^{-rt}K, 0)$ for call option when threshold $\overline{L} = \infty$.

References

1. Rubinstein, M. The valuation of uncertain income streams and the pricing of options. *Bell J. Econ.* **1976**, *7*, 407–425. [CrossRef]
2. Brennan, M.J. The pricing of contingent claims in discrete time models. *J. Financ.* **1979**, *34*, 53–68. [CrossRef]
3. Ross, S. Mutual fund separation in financial theory: The separation distributions. *J. Econ. Theory* **1978**, *17*, 254–286. [CrossRef]
4. Cochrane, J. *Asset Pricing: Revised Edition*; Princeton University Press: Princeton, NJ, USA, 2005; ISBN 9780691121376.
5. Lehar, A.; Scheicher, M.; Schittenkopf, C. GARCH vs. stochastic volatility: Option pricing and risk management. *J. Bank Financ.* **2002**, *26*, 323–345. [CrossRef]
6. Cochrane, J.; Saa-Requejo, J. Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *J. Polit. Econ.* **2000**, *108*, 79–119. [CrossRef]
7. Bernardo, A.; Ledoit, O. Gain, loss and asset pricing. *J. Polit. Econ.* **2000**, *108*, 173–206. [CrossRef]
8. Corrado, C.; Su, T. Implied volatility skews and stock index skewness and kurtosis implied by S&P 500 index option prices. *J. Deriv.* **1997**, *4*, 8–19.
9. Jackwerth, J.C.; Rubinstein, M. Recovering probability distributions from contemporary security prices. *J. Financ.* **1996**, *51*, 1611–1631. [CrossRef]
10. Breeden, D.; Litzenberger, R. Prices of state-contingent claims implicit in options prices. *J. Bus.* **1978**, *51*, 621–651. [CrossRef]
11. Rubinstein, M. Implied binomial trees. *J. Financ.* **1994**, *49*, 771–818. [CrossRef]
12. Aït-Sahalia, Y.; Lo, A. Nonparametric estimation of state-price densities implicit in financial asset prices. *J. Financ.* **1998**, *53*, 499–548. [CrossRef]
13. Aït-Sahalia, Y.; Lo, A. Nonparametric risk management and implied risk aversion. *J. Econom.* **2000**, *94*, 9–51. [CrossRef]
14. Jackwerth, J.C. Recovering risk aversion from option prices and realized returns. *Rev. Financ. Stud.* **2000**, *13*, 433–451. [CrossRef]
15. Longstaff, F. Option pricing and the martingale restriction. *Rev. Financ. Stud.* **1995**, *8*, 1091–1124. [CrossRef]
16. Aparicio, S.; Hodges, S. Implied risk-neutral distribution: A comparison of estimation methods. In *FORC Preprint*; Working Paper; Warwick University: Coventry, UK, 1998.
17. Corrado, C. Option pricing based on the generalized lambda distribution. *J. Futures Mark.* **2001**, *21*, 213–236. [CrossRef]
18. De Jong, C.; Huisman, R. From Skews to a Skewed-t: Modelling Option-Implied Returns by a Skewed Student-t. In Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering, New York, NY, USA, 28 March 2000.
19. Ritchey, R. Call option valuation for discrete normal mixtures. *J. Financ. Res.* **1990**, *13*, 285–295. [CrossRef]
20. Melick, W.; Thomas, C. Recovering an asset's implied PDF from option prices: An application to crude oil during the Gulf crisis. *J. Financ. Quant. Anal.* **1997**, *32*, 91–115. [CrossRef]
21. Buchen, P.; Kelly, M. The maximum entropy distribution of an asset inferred from option prices. *J. Financ. Quant. Anal.* **1996**, *31*, 143–159. [CrossRef]
22. Rockinger, M.; Jondeau, E. Entropy densities with an application to autoregressive conditional skewness and kurtosis. *J. Econom.* **2002**, *106*, 119–142. [CrossRef]
23. Härdle, W.; Yatchew, A. Dynamic nonparametric state price density estimation using constrained least squares and the bootstrap. In *Working Paper*; Humboldt University Berlin: Berlin, Germany, 2002.
24. Mayhew, S. On estimating the risk-neutral probability distribution implied by option prices. In *Working Paper*; Purdue University: West Lafayette, IN, USA, 1995.
25. Andersen, A.; Wagener, T. Extracting risk neutral probability densities by fitting implied volatility smiles: Some methodological points and an application to the 3M EURIBOR futures option prices. In *Working Paper*; European Central Bank: Frankfurt, Germany, 2002.
26. Rosenberg, J. Nonparametric pricing of multivariate contingent claims. *J. Deriv.* **2003**, *10*, 9–26. [CrossRef]
27. Wang, Y.; Zhao, M.; Han, Y.; Zhou, J. A fuzzy expression way for air quality index with more comprehensive information. *Sustainability* **2017**, *9*. [CrossRef]

28. Wan, H. Pricing American-style basket options by implied binomial tree. In *Working Paper*; Haas School of Business: Berkeley, CA, USA, 2002.

29. Cox, J.; Ross, S.; Rubinstein, M. Option pricing: A simplified approach. *J. Financ. Econ.* **1979**, *7*, 229–263. [CrossRef]

30. Longarela, I. Gain, loss, and asset pricing: It is much easier. A note. In *Working Paper*; Stockholm School of Economics: Stockholm, Sweden, 2008.

31. Stephan, J.; Whaley, R. Intraday price change and trading volume relations in the stock and stock option markets. *J. Financ.* **1990**, *45*, 191–220. [CrossRef]

32. Bergsma, K.; Fodor, A.; Singal, V.; Tayal, J. Intraday option to stock volume ratios and stock return predictability. Social Science Electronic Publishing. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3095239 (accessed on 16 May 2018).

33. Birru, J.; Figlewski, S. Anatomy of meltdown: The risk neutral density for the S&P 500 in the fall of 2008. *J. Financ. Mark.* **2012**, *15*, 151–180. [CrossRef]

34. Stoll, H.; Whaley, R. Programme trading and expiration-day effects. *Financ. Anal. J.* **1987**, *43*, 16–28. [CrossRef]

35. Longarela, I. A Simple linear programming approach to gain, loss and asset pricing. *Top. Theor. Econ.* **2015**, *2*, 1064. [CrossRef]

36. Pinar, M. Gain-loss pricing under ambiguity of measure. *ESAIM Control Optim. Calc. Var.* **2010**, *16*, 132–147. [CrossRef]

37. Voelzke, J.; Mentemeier, S. Computing the substantial-gain-loss-ratio. In *CQE Working Papers*; Center for Quantitative Economics: Münster, Germany, 2017.

38. Mugwagwa, T.; Ramiah, V.; Moosa, I. The profitability of option-based contrarian strategies: An empirical analysis. *Int. Rev. Financ.* **2015**, *15*, 1–26. [CrossRef]

Bayesian Approach for Estimating the Probability of Cartel Penalization under the Leniency Program

Jihyun Park [1], Juhyun Lee [1] and Suneung Ahn [2],*

[1]   Department of Industrial and Management Engineering, Hanyang University, Seoul 04763, Korea;
      pjh3226@hanyang.ac.kr (J.P.); ljh812@hanyang.ac.kr (J.L.)
[2]   Department of Industrial and Management Engineering, Hanyang University ERICA, Ansan 15588, Korea
*   Correspondence: sunahn@hanyang.ac.kr; Tel.: +82-031-400-5267

Abstract: Cartels cause tremendous damage to the market economy and disadvantage consumers by creating higher prices and lower-quality goods; moreover, they are difficult to detect. We need to prevent them through scientific analysis, which includes the determination of an indicator to explain antitrust enforcement. In particular, the probability of cartel penalization is a useful indicator for evaluating competition enforcement. This study estimates the probability of cartel penalization using a Bayesian approach. In the empirical study, the probability of cartel penalization is estimated by a Bayesian approach from the cartel data of the Department of Justice in the United States between 1970 and 2009. The probability of cartel penalization is seen as sensitive to changes in competition law, and the results have implications for market efficiency and the antitrust authority's efforts against cartel formation and demise. The result of policy simulation shows the effectiveness of the leniency program. Antitrust enforcement is evaluated from the estimation results, and can therefore be improved.

Keywords: Bayesian approach; conjugate prior; cartel; leniency program; policy simulation

1. Introduction

Cartels cause tremendous damage to perfect competition markets and consumers by effectually applying upward pressure on prices and downward pressure on quality; moreover, cartels are difficult to detect because of their tacit nature. In this way, cartels mitigate against perfect competition under which consumers are offered the best goods and services at the lowest possible prices. Antitrust authorities have continuously sought to maintain a free-market system against cartels, but with only partial and limited success.

In previous research, the probability of cartel detection was a key indicator for measuring the effectiveness of antitrust policies. Detection is the state in which unobserved cartels are caught by the antitrust authority. After introducing a leniency program as a new antitrust policy, both the number of cartel investigations and the probability of cartel detection increase. The higher the probability of cartel detection, the greater the expected penalties, and therefore, the likelihood of cartel formation will decrease. On this principle, it is possible to measure the deterrence effect according to the change in antitrust policy. This study uses the probability of cartel penalization as a key indicator.

The Markov transition process and the birth and death process model were widely used. Bryant and Eckard [1] constructed the birth and death process model to empirically analyze cartel data provided by the United States (US) Department of Justice, and estimated the probability of cartel detection in the US in 1961–1988 as between 13–17%. Using the same method, Combe et al. [2] estimated European Commission (EC) cartel detection probabilities of 12.9–13.2% for 1969–2007. When the birth and death model has two states of competition and collusion, the lifetimes and inter-arrival times between the births of cartels were independent and had exponential distributions with means of λ^{-1}

and θ^{-1}. The number of cartels at a particular time *t* follows a Poisson distribution with a mean of $\theta_T = (\theta/\lambda)\{1 - e^{-\lambda T}\}^2$. Both Bryant and Eckard [1] and Combe et al. [2] assumed that every cartel would eventually be caught and prosecuted. However, this assumption is not realistic, because some cases are not penalized, despite having been detected.

Further, Bryant and Eckard [1] and Combe et al. [2] do not take account of the unobservable cartel population. J. E. Harrington and Chang [3] sought to estimate the unobservable population by developing the birth and death model from that noted above. They concluded that cartel duration could be a good indicator of whether new competition law had a significant cartel-dissolution effect. Using Harrington and Chang [3]'s model, Zhou [4] analyzed the EC cartel data for 1985–2012, and concluded that the EU's new leniency program in 2002 had the effect of deterring cartels.

In the research of Bryant and Eckard [1], Combe et al. [2], Harrington and Chang [3], and Zhou [4], the probability of cartel detection—as derived from cartel duration—entailed the determination of the time-average probability from continuous variables. On the other hand, there is research indicating that the probability of cartel detection represents the ensemble-average probability obtained from discrete variables such as caseloads. The time-average probability is the average of a stochastic process that is obtained by selecting a sample path randomly, and taking the average of a period in a particular state on that sample path over the observation period. The ensemble-average probability is that mean of a quantity at time *t* that is estimated by the average of the ensemble of possible states of total sample paths in stochastic process theory [5,6].

Miller [7] formulated a cartel behavior model using the Markov process, and used the number of cartel cases as discrete variables. The model assumed that the cartel transition process is in a non-absorbing and first-order Markov chain in contrast with previous Markov models, and showed the change of the number of cartel detections before and after a leniency program. He concluded that the introduction of this leniency program in 1993 increased the detection and deterrence capabilities of competition enforcement. The previous research above [1–4,7] used Markov process models; this research had two notable points.

First, the duration of cartels and inter-arrival times between cartels follow exponential distributions. Verifying this assumption requires a hypothesis testing of the null hypothesis that "the distribution is exponential". The cumulative distribution function $\hat{F}(x)$ of durations and inter-arrival times is given by:

$$\hat{F}(x) = \frac{number\ of\ observations \leq x}{total\ number\ of\ observations}$$

Under the exponential distribution, $log(1 - \hat{F}(x))$ should be approximately linear in *x*. The result of these previous works indicates that the cartels' duration and inter-arrival times between cartels follow the exponential distribution; therefore, models can be applied to the Markov process [7].

Second, this research assumed that the cartel process was stationary for adopting the Markov process, and that the values could be analyzed when the cartel process attained a steady state; this is also unrealistic. In the research of Bryant and Eckard [1] and Combe et al. [2], the probability is the resultant value when it reaches a steady state. This kind of probability is called a time-independent probability. Otherwise, the form of estimators needs to be a time-dependent rather than time-independent, because the purpose of estimating the probability of cartel detection is evaluating the effects of various competition policies [8]. Thus, Hinloopen [8]'s research was an theoretical literature review for analyzing a subgame of collusion.

A new mathematical methodology has emerged recently in the form of a non-Markov process. Ormosi [9] estimated the annual probability of cartel detection by employing capture–recapture methods based on EC information in the period between 1981–2001. The methods of Ormosi [9], which are frequently used in ecology, reflect that transition parameters are not steady state, and that detection and survival rates are time-independent. However, there are two unreasonable assumptions. First, capture–recapture methods assume that temporary migrations between the two states (compete–collude) do not exist; thus, they are regarded as robust design methods. The antitrust

policy tends to vary broadly according to governmental power or social issues. Second, Ormosi [9] deduced a result from moving average methods, specifically in the moving average of three or five years. If the probability is used on the basis of a single year, the accuracy of the probability may decrease due to data insufficiency. The industry reacts immediately to changes in competition law; therefore, the probability needs to be estimated for the smallest unit of time.

This paper seeks to estimate the probability of cartel penalization using a Bayesian approach and evaluate the impact of the leniency program as an antitrust policy. This study uses the conjugate family of the beta-binomial in that the cartel occurs in binomial events. The posterior mean of the beta distribution is the probability of cartel penalization in a year. This shows the trend of the probability of cartel penalization, and can then improve the antitrust policy using the measured impact of the leniency program. In this light, the present research makes three contributions.

First, this paper estimates the probability of cartel penalization for analyzing cartels in contrast to the probability of cartel detection as treated in previous research. The probability of cartel detection means the probability that unobserved cartels will be investigated, prosecuted, and penalized. However, the probability of cartel penalization means the penalized likelihood of investigated cartels through sufficient investigation. This is used as an indicator with which to evaluate the impact of the leniency program and the capability of antitrust authorities.

Second, the methodology of this paper makes up for the weak points of previous probability estimation methods. Previous methods have many unrealistic assumptions such as the analyzed cases being eventually caught/detected cases, the time-average probability, etc. We can improve on these assumptions by estimating the time-dependent ensemble-average probability based on the discrete data of caseloads, which is more practical than the time-average probability for the sensitive estimation of probability.

Third, this study shows that the Bayesian approach could play a practical role in modeling and analyzing the cartel situation. Although the Markov process model, which was commonly used in previous research, is an essential consideration "in steady-state probability", it is difficult to assume "in steady-state probability", because cartel cases continuously vary over time. The probability of cartel penalization estimated using the Bayesian approach does not need to consider "steady-state probability". The Bayesian approach for estimating probability can contain significant uncertainty, but has good predictive performance in itself [10]. The bias between the estimation probability and the actual value could be solved from the update procedure of the Bayesian approach. Therefore, we present reliable results using the non-informative prior and conjugate prior distribution when prior information is insufficient.

The paper is organized as follows: Section 2 defines the penalization probability and Bayesian probabilistic model; Section 3 presents an empirical study based on US cartel data; and Section 4 draws conclusions.

2. Bayesian Probabilistic Model

When faced with suspected cartel cases, a competition authority carries out an initial investigation to determine whether there are sufficient grounds to prosecute. Prosecuted cartels are penalized in the form of fines through a trial. Eventually, the three states of cartel cases are investigation, prosecution, and penalization [11]. The estimated probability of this study is based on investigation and penalization states. The probability of cartel penalization (ρ_t) is described as the proportion of the numbers of penalized cases to investigated cases for year t ($t = 1, 2, \cdots$).

The estimation of the penalization probability using the Bayesian approach involves two assumptions. First, the unit of case is an industry. Accordingly, the research of Bryant and Eckard [1] and Miller [7] is based on the analysis unit of the industry. Bos and Harrington [12] argued that firm-based analysis is more realistic; nonetheless, this study was analyzed based on the analysis unit of industry for easy analysis. In practice, cartels can participate in all firms of an industry. Second, a cartel only arises as one event during a year. Every cartel is transferred to the competition as a result

of punishment by the authorities. This is called the "Grim trigger strategy" [13,14]. Thereafter, if some player deviates from the cartel, the game cannot be colluded indefinitely.

This study constructed a Bayesian probabilistic model to estimate the probability of cartel penalization. The probability of cartel penalization is the posterior mean calculated from the posterior distribution. Inferring a posterior distribution requires determining the proper prior distribution. A Bayesian probabilistic model is comprised of a prior distribution to induce a posterior distribution, hyperparameters, and a likelihood function. A Bayesian sequential analysis of the dynamic Bayesian model can be used to reflect the latest trends of time-series data [15,16].

Two things should be considered to induce a posterior distribution from a prior distribution: the likelihood function and the parameters in the prior distribution, which are known as hyperparameters [17]. The natural conjugate priors are generally recommended in the Bayesian approach, because its functional form is similar to the likelihood distribution [18,19]. Therefore, we have to obtain the appropriate likelihood function to adopt the notion of natural conjugacy. Consider the following notations for the Bayesian probabilistic model.

ρ_t: The probability of cartel penalization cases in year t

n_t: The number of cartel investigation cases by the competition authority in year t

k_t: The number of cartel penalization cases by the competition authority in year t

When the investigated industry participating in a cartel is n, Figure 1 shows a binomial tree to demonstrate the process of cartel formation and demise in year t.
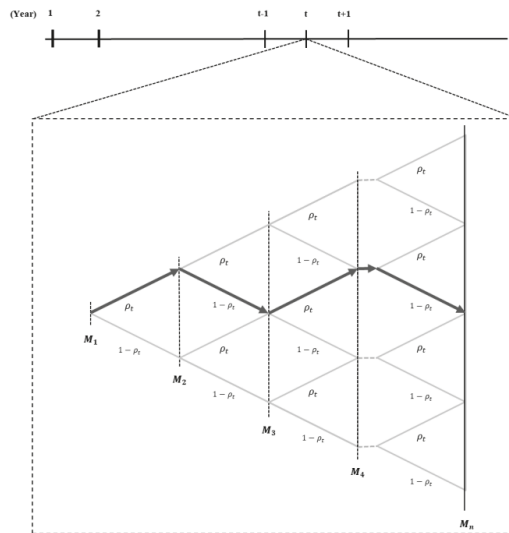
Figure 1. Estimating the probability of cartel penalization through a binomial tree.

In Figure 1, M_1, M_2, $\cdots$, M_n is the industry of investigated cartels in year t. Arrows in the path show whether the investigated cartels were finally penalized. When a route contains an arrow pointing to the right, this cartel will be finally penalized; otherwise, it is not penalized. For example, the industry M_2 is in the left direction; this means that industry M_2 will be not finally penalized as the probability ρ_t. This study wants to infer the probability of industry $n + 1$ penalization in path G; this probability is estimating the likelihood function based on the data from industry 1 to n, and the prior distribution while inferring a posterior distribution from the Bayesian approach [13]. The expectation of a posterior distribution indicates the probability of cartel penalization.

2.1. Likelihood Function and Prior Distribution

The variable n_t is the number of cartel cases investigated in year t, and each case follows the Bernoulli process with an independent and identical distribution. Therefore, the Bernoulli random variable X_i with one case shown is given by:

$$X_i = \begin{cases} 1 & \text{if penalizing with probability } \rho_t \\ 0 & \text{if non} - \text{penalizing with probability } 1 - \rho_t, \end{cases}$$

where i is the number of cartel firm ($i = 1, \cdots, n_t$) and $0 < \rho_t < 1$. The probability mass function of the random variable, which is known as the Bernoulli probability, is given by:

$$f(x_i | \rho_t) = \rho_t^{x_i} (1 - \rho_t)^{1 - x_i}. \tag{1}$$

Once the number of cases n_t is investigated, and k_t is penalized in year t, the joint probability mass function of cartel cases is given by:

$$\begin{aligned} L(\rho_t | x_1, \cdots, x_{n_t}) = f(x_1, \cdots, x_{n_t} | \rho_t) \quad &= \prod_{i=1}^{n_t} f(x_i | \rho_t) \\ &= \prod_{i=1}^{n_t} \rho_t^{x_i} (1 - \rho_t)^{1 - x_i} \\ &= \rho_t^{\sum x_i} (1 - \rho_t)^{n_t - \sum x_i} \\ &= \rho_t^{k_t} (1 - \rho_t)^{n_t - k_t}. \end{aligned} \tag{2}$$

The probability of cartel penalization has a value between 0 and 1. In Equation (2), $f(\rho_t)$ is a binomial form as the prior distribution, because there are only two final states of a cartel: whether it has been penalized or not. Thus, we use the beta distribution as a prior distribution based on the natural conjugacy [17,20]. The prior distribution $f(\rho_t)$ is the beta distribution with hyperparameters α and β; thus, the probability density function is given by:

$$f(\rho_t) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \rho_t^{\alpha - 1} \cdot (1 - \rho_t)^{\beta - 1}, \tag{3}$$

where $\alpha > 0$ and $\beta > 0$ are the hyperparameters. The function $\Gamma(\cdot)$ is a gamma function, which is defined as:

$$\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha - 1} dx. \tag{4}$$

Note that when α is a positive integer, $\Gamma(\alpha) = (\alpha - 1)!$.

2.2. Bayesian Estimation

In the Bayesian approach, the posterior distribution is given by:

$$f(\rho_t | x_1, \cdots, x_{n_t}) = \frac{f(x_1, \cdots, x_{n_t}, \rho_t)}{f(x_1, \cdots, x_{n_t})}. \tag{5}$$

The joint probability distribution $f(x_1, \cdots, x_{n_t}, \rho_t)$ in Equation (5), which reflects the multiplicative laws of probability in Equations (2) and (3), is:

$$\begin{aligned} f(x_1, \cdots, x_{n_t}, \rho_t) &= f(x_1, \cdots, x_{n_t} | \rho_t) \cdot f(\rho_t) \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \rho_t^{k_t + \alpha - 1} \cdot (1 - \rho_t)^{n_t - k_t + \beta - 1}. \end{aligned} \tag{6}$$

The marginal probability distribution $f(x_1, \cdots, x_{n_t})$, which is calculated by the law of total probability, is given by:

$$
\begin{aligned}
f(x_1, \cdots, x_{n_t}) &= \int_0^1 f(x_1, \cdots, x_{n_t}, \rho_t) d\rho_t \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \rho_t^{k_t+\alpha-1} \cdot (1-\rho_t)^{n_t-k_t+\beta-1} d\rho_t \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \int_0^1 \rho_t^{k_t+\alpha-1} \cdot (1-\rho_t)^{n_t-k_t+\beta-1} d\rho_t,
\end{aligned}
\tag{7}
$$

where $\int_0^1 \rho_t^{k_t+\alpha-1} \cdot (1-\rho_t)^{n_t-k_t+\beta-1} d\rho_t = \frac{\Gamma(\alpha+k_t) \cdot \Gamma(n_t-k_t+\beta)}{\Gamma(\alpha+\beta+n_t)}$.

Suppose that the initial probability (ρ_t) is 0.5 meaning whether the investigated or the non-investigated case for eliminating the dependence on the prior information. The hyperparameters α and β are 1 as a non-informative prior. Therefore, the posterior distribution is a beta distribution with the parameters $\alpha + k_t$ and $\beta + n_t - k_t$. The posterior distribution of Equation (5) is represented by:

$$
\begin{aligned}
f(\rho_t | x_1, \cdots, x_{n_t}) &= \frac{f(x_1, \cdots, x_{n_t}, \rho_t)}{f(x_1, \cdots, x_{n_t})} \\
&= \frac{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \rho_t^{k_t+\alpha-1} \cdot (1-\rho_t)^{n_t-k_t+\beta-1}}{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+k_t) \cdot \Gamma(n_t-k_t+\beta)}{\Gamma(\alpha+\beta+n_t)}} \\
&= \frac{\Gamma(\alpha+\beta+n_t)}{\Gamma(\alpha+k_t)\Gamma(\beta+n_t-k_t)} \cdot \rho_t^{k_t+\alpha-1} \cdot (1-\rho_t)^{n_t-k_t+\beta-1}.
\end{aligned}
\tag{8}
$$

The posterior mean $E[\rho_t | x_1, \cdots, x_{n_t}]$ from Equation (8) is:

$$
\begin{aligned}
E[\rho_t | x_1, \cdots, x_{n_t}] &= \int_0^1 \rho_t \cdot f(\rho_t | x_1, \cdots, x_{n_t}) \, d\rho_t \\
&= \int_0^1 \rho_t \frac{\Gamma(\alpha+\beta+n_t)}{\Gamma(\alpha+k_t)\Gamma(\beta+n_t-k_t)} \cdot \rho_t^{k_t+\alpha-1} \cdot (1-\rho_t)^{n_t-k_t+\beta-1} d\rho_t \\
&= \frac{\Gamma(\alpha+\beta+n_t)}{\Gamma(\alpha+k_t)\Gamma(\beta+n_t-k_t)} \cdot \int_0^1 \rho_t^{(k_t+\alpha+1)-1} \cdot (1-\rho_t)^{(n_t-k_t+\beta)-1} d\rho_t \\
&= \frac{\Gamma(\alpha+\beta+n_t)}{\Gamma(\alpha+k_t)\Gamma(\beta+n_t-k_t)} \cdot \frac{\Gamma(\alpha+k_t+1) \cdot \Gamma(n_t-k_t+\beta)}{\Gamma(\alpha+\beta+n_t+1)} \\
&= \frac{\alpha+k_t}{\alpha+\beta+n_t}.
\end{aligned}
\tag{9}
$$

3. Empirical Study

3.1. Data

This study uses data from the Workload statistics published by the Antitrust Division of the Department of Justice (DOJ) for the period between 1970–2009 [21]. The information is shown in Table 1. It contains the annual statistics of penalized cases and investigated cases by the criminal enforcement and civil enforcement of district courts, with respect to the laws of Sherman §1-Restraint of Trade, Sherman §2-Monopoly, and Clayton §7-Mergers. The antitrust division prosecutes in the form of criminal enforcement cases if the cartels, which are known as "hardcore cartels," are determined by preliminary examination to have an especially injurious impact on the industry; otherwise, it prosecutes in the form of civil enforcement cases. This study does not consider the appellate cases and the cases of contemporary criminal–civil enforcement at the same time, due to a few of applicable cases.

Table 1. Statistic of cartel data by the Department of Justice (DOJ) between 1970 and 2009.

Years	1970–1979	1980–1989	1990–1999	2000–2009
Number of investigation cases	4155	2505	3145	1881
Total penalization cases	544	859	780	535
(Civil cases)	351	210	161	129
(Criminal cases)	193	649	619	406
Total fines ($thousand)	47,712	187,548	1,612,993	4,222,407

3.2. Time-Series Analysis

Prior to the model application, a time-series analysis was implemented to eliminate spurious relations. This study, alternatively, employed the augmented Dickey–Fuller (ADF) unit root test to confirm the stability of the time-series data (details are provided in Appendix A).

If the result shows that the time-series data is unstable, the difference stationary process is needed. The representative method for stabilizing time-series data is order difference or log order difference. However, using order difference, it is possible that the meaning of original data will be lost, leading to different conclusions in the economy [22]. Economic variables such as price, currency, and stock index cannot be used to verify the stability of time-series data, because they are commonly non-stationary data [23].

3.3. Results

The empirical study, using the model defined in Section 2, drew an annual beta distribution for the probability of cartel penalization. The results are summarized in Table 2, and Figure 2 illustrates the distribution for every year.

Figure 2 shows that the probability distributions tend to increase over time. Beta distributions converge on a specific range with Bayesian updating [17]. Indeed, the result shows the convergence of the present distribution on the specific range at around 0.22. We were able to calculate the posterior mean by Equation (9). Figure 3, accordingly, illustrates the annual expected probability of cartel penalization.

Table 2. The probability of cartel penalization through Bayesian sequential analysis.

Years (*t*)	Investigation Cases (n_t)	Penalization Cases (k_t)	Prior α	Prior β	Posterior α	Posterior β	The Expected Probability of Cartel Penalization ($æ_t$)
1970 (1)	473	53	1	1	54	421	
1971 (2)	593	51	54	421	105	963	0.11368
1972 (3)	465	53	105	963	158	1375	0.09831
1973 (4)	538	61	158	1375	219	1852	0.10307
1974 (5)	338	57	219	1852	276	2133	0.10575
1975 (6)	381	29	276	2133	305	2485	0.11457
1976 (7)	374	64	305	2485	369	2795	0.10932
1977 (8)	484	46	369	2795	415	3233	0.11662
1978 (9)	290	68	415	3233	483	3455	0.11376
1979 (10)	407	62	483	3455	545	3800	0.12265
1980 (11)	377	89	545	3800	634	4088	0.12543
1981 (12)	255	93	634	4088	727	4250	0.13427
1982 (13)	262	109	727	4250	836	4403	0.14607
1983 (14)	245	99	836	4403	935	4549	0.15957
1984 (15)	257	80	935	4549	1015	4726	0.17050
1985 (16)	254	77	1015	4726	1092	4903	0.17680
1986 (17)	307	98	1092	4903	1190	5112	0.18215
1987 (18)	270	27	1190	5112	1217	5355	0.18883
1988 (19)	216	55	1217	5355	1272	5516	0.18518
1989 (20)	220	132	1272	5516	1404	5604	0.18739
1990 (21)	178	77	1404	5604	1481	5705	0.20034
1991 (22)	178	81	1481	5705	1562	5802	0.20610
1992 (23)	176	113	1562	5802	1675	5865	0.21211
1993 (24)	224	84	1675	5865	1759	6005	0.22215
1994 (25)	269	58	1759	6005	1817	6216	0.22656
1995 (26)	249	86	1817	6216	1903	6379	0.22619
1996 (27)	436	59	1903	6379	1962	6756	0.22978
1997 (28)	454	64	1962	6756	2026	7146	0.22505
1998 (29)	408	89	2026	7146	2115	7465	0.22089
1999 (30)	373	69	2115	7465	2184	7769	0.22077
2000 (31)	261	64	2184	7769	2248	7966	0.21943
2001 (32)	225	61	2248	7966	2309	8130	0.22009
2002 (33)	192	50	2309	8130	2359	8272	0.22119
2003 (34)	218	43	2359	8272	2402	8447	0.22190
2004 (35)	171	46	2402	8447	2448	8572	0.22140
2005 (36)	217	42	2448	8572	2490	8747	0.22214
2006 (37)	204	48	2490	8747	2538	8903	0.22159
2007 (38)	186	40	2538	8903	2578	9049	0.22183
2008 (39)	172	58	2578	9049	2636	9163	0.22173
2009 (40)	164	83	2636	9163	2719	9244	0.22341
2010 (41)							0.22728

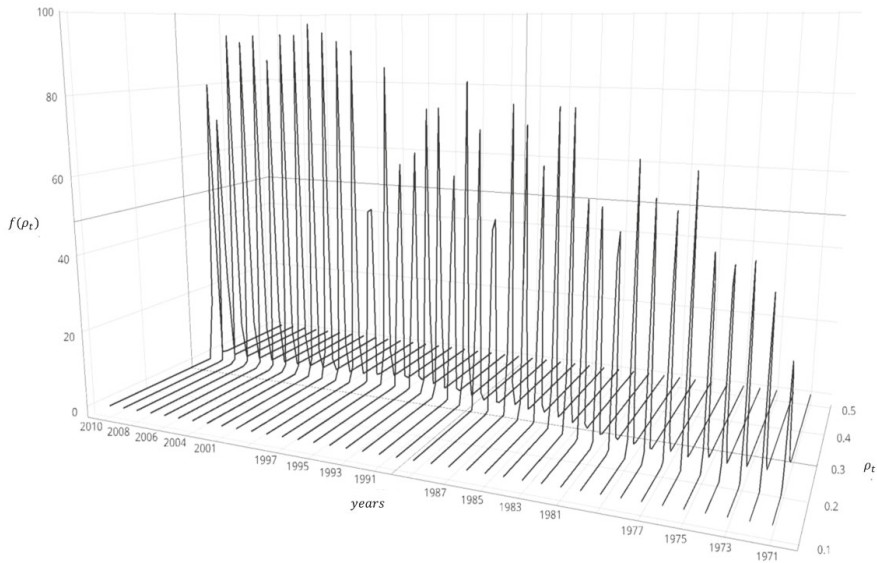Source: Workload Statistics, Department of Justice in the United States (US).

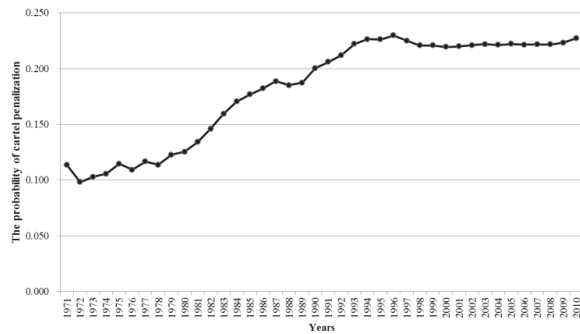**Figure 2.** Annual beta distributions of the probability of cartel penalization.

Figure 3. The annual probability of cartel penalization.

In the late 19th century, the United States was confronted with a very significant change: large-scale manufacturing interests emerged, in great numbers, and enjoyed excessive economic power. In response, the Interstate Commerce Act in 1887 began a shift towards federal rather than state regulation of big business. This was followed by the Sherman Antitrust Act in 1890, which is the basis of US competition laws. Later, the Clayton Antitrust Act in 1914 was enacted to prohibit price discrimination, corporate mergers, and interlocking directorates.

We can now show how the change of probability of cartel penalization impacted upon the antitrust laws in the analysis periods. The Antitrust Penalty and Procedure Act in 1974, which was known as the Tunney Act, required that prospective mergers and acquisitions obtain approval from the DOJ. In 1976, the Hart–Scott–Rodino Antitrust Improvements Act was passed, and in 1978, the leniency program was instituted. At this notable time, the probability of cartel penalization was increasing. At the peak of cartel penalization probability, in 1994, the DOJ reformed the leniency program. The reformed version of the program included an additional amnesty for those who cooperate with investigations.

Figure 3 indicates that the probability after 1994 has been steady and stable. The reform of competition laws clearly had an impact on the industry.

3.4. Model Comparison

Chang and Harrington [24] constructed a Markov process model to consider the stochastic formation and demise of cartels. By numerical analysis, they estimated the impact of the leniency program on the steady-state rate. Figure 4, in the form of the analysis results, plots the change in the rate of penalized cartels according to the proportion of prosecuted cases.
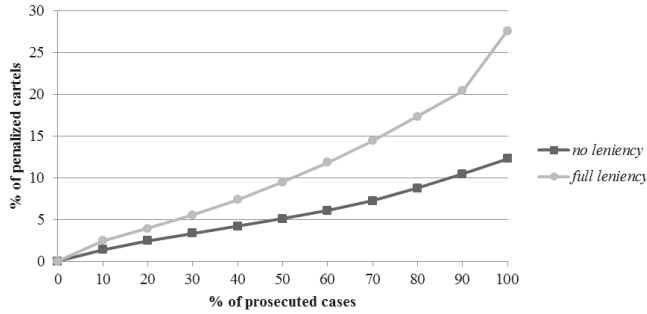
Figure 4. Effects of the proportions of penalized cartels according to the probability of prosecuted cases.

The proportion of probable prosecution cases, as reflects the 1970–2009 Workload statistics, was about 20~40%. In this value, the rate of penalized cartels is estimated about 5~10%.

The estimated probability of cartel penalization of this study and Bryant and Eckard [1]'s results are similar in their proportion of penalization to investigation. However, the present approach is the ensemble-average probability using discrete data, whereas that of Bryant and Eckard [1] is the time-average probability using continuous data. Cartel analysis is more commensurate with discrete data than with continuous data, because the form of Workload statistics data, as announced annually by the DOJ, is discrete. With our similar definition of probability, we could draw a box plot in the overlapped analysis period 1962–1988.

Figure 5 shows that the Bayesian probabilistic model estimates 0.114 for the top 25th percentile, and 0.1737 for the top 75th percentile, which are statistically significant. These are close to Bryant and Eckard [1]'s estimates, which fell between 0.128 and 0.174.
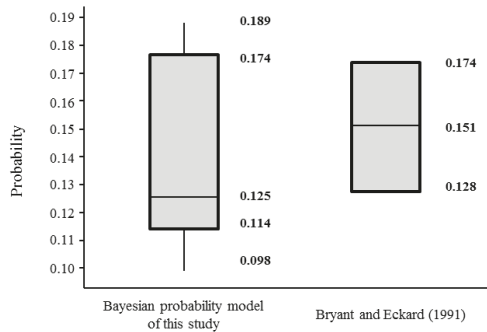
Figure 5. Box plots of the Bayesian probabilistic model and Bryant and Eckard's (1991) model.

3.5. Impact of Leniency Program

This study utilized a policy simulation to analyze the impact of competition policies [25,26]. In policy evaluation research, the impact of policy implementation is indicated as value-added. In other words, the impact is described as the difference of outcomes between implementing the policy and otherwise. The leniency program has been deemed an effective antitrust policy for detecting and deterring cartels in many countries. In general, the leniency program provides partial or total exemption for penalty to a cartel member who voluntarily reports information or agreements that prove helpful to the antitrust authorities. Under the leniency program, a firm or individual in a cartel is bound to first confess involvement for avoiding conviction or fines. The optimal policy is found by evaluating the impact of the leniency program. It is given by:

$$\frac{BX_{1992} - AX_{1992}}{AX_{1992}} \times 100 \, . \tag{10}$$

The impact of the leniency program (%) is the difference between the penalization probability under both it and non-leniency. The leniency program was originally launched in 1978 in the US, and was reformed in 1993. In Equation (10), BX_{1992} is the 1992 penalization probability estimated on the basis of the leniency program's implementation in 1978, and AX_{1992} is the penalization probability in 1992 estimated on the basis of the leniency program's non-implementation. The estimated probability BX_{1992} was calculated as 0.21211 by the Bayesian probabilistic model, and AX_{1992} was calculated as 0.1328 by the ordinary least squares estimation method of regression. The impact of the leniency program by the policy simulation, finally, is 65.39%. This can be seen in Figure 6.
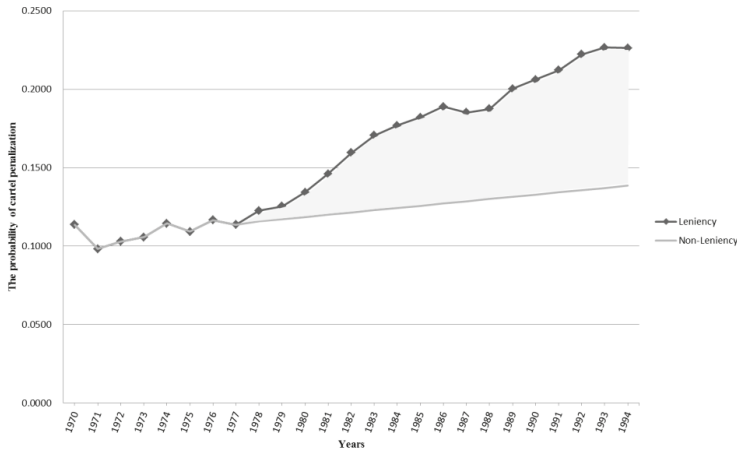
Figure 6. The increment of the probability of cartel penalization in US.

There has been much research that has analyzed the effectiveness and efficiency of the leniency program (i.e., Miller [7], Chang and Harrington [24], and Brenner [27]). The result of this study is similar to those of the research of Chang and Harrington [24] and Miller [7], which is based on US data; the implication was that the leniency program is a very effective policy. Chang and Harrington [24] argue that the occurrence of cartels decreased by about 70%, and the deterrence capability of the antitrust authority increased by about 60% after introducing the leniency program. Miller [7], through Poisson regression analysis, estimated the impact of the leniency program every half year using US data for the years 1985 to 2005. In the results, the detection capability increased by about 60%, and the deterrence capability improved by about 40%.

4. Conclusions

This study attempted to estimate the probability of cartel penalization using a Bayesian approach. Bryant and Eckard [1], Combe et al. [2], Harrington and Chang [3], and Zhou [4] estimated the probability of cartel detection in the form of the time-average probability from continuous data. However, the probability of cartel penalization of this study was estimated in the form of the ensemble-average probability from Workload statistics. Bryant and Eckard [1], Combe et al. [2], Harrington and Chang [3], Zhou [4], and Miller [7] all assumed that the duration of cartels and the inter-arrival times between cartels follow exponential distributions, and that the stochastic process for cartel cases is stationary. However, we built a Bayesian probabilistic model, as it did not need to consider a stationary process. This study made two assumptions: an industry-based analysis, and the grim trigger strategy. On the basis of the 1970–2009 Workload statistics from the US Department of Justice, the determined probability of cartel penalization reflected a sensitive response according to the change of antitrust policy. The result of the policy simulation of the impact of the leniency program was about 65%. The results are similar with the results of Chang and Harrington [24] and Miller [7], and similar to that of Bryant and Eckard [1]; indeed, the common finding among all of the studies, including the current study, was that the leniency program is a very effective policy.

This study evaluated the impact of antitrust policy and, therefrom estimated the probability of cartel penalization. From the antitrust authority standpoint, it provides an improved optimal policy, and from the corporate standpoint, it provides more effective decision-making. Certainly, the present paper has several limitations. First, further studies on realistic situations in specific countries and industries are needed. New antitrust policies recently have been introduced, such as for example, Amnesty Plus, punitive damage, class action, and consent order. These were also considered in further study.

Author Contributions: J.P., J.L. and S.A. conceived and designed the research; J.P. and J.L. conducted the research and drafted the manuscript. S.A. supervised the overall work. All authors read and approved the final manuscript.

Appendix A

An ADF unit root test of maximum time lag 10 based on the Schwarz information criterion is performed using E-Views software. The regression of the time series for the test is

$$y_t = \delta y_{t-1} + u_t, \tag{A1}$$

where u_t is the white noise error term, following the normal distribution of mean 0 and variance σ^2.

The case of $\delta = 1$ in Equation (A1) indicates that the model has a unit root with a random walk. Time lags usually account for one-third of the total time series [22]. Accordingly, in the ADF unit root test, the time series is 30, and so the maximum time lag is 10. In any ADF unit root test, the procedure is important [28,29]. Such procedures are the model including the constant and time trend ($y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + u_t$), the model including the constant ($y_t = \beta_1 t + \delta y_{t-1} + u_t$), and the model including nothing ($y_t = \delta y_{t-1} + u_t$).

There are information criteria for ADF unit root tests: the AIC (Akaike information criterion), and the above-noted SIC (Schwarz information criterion). SIC, which supplements the AIC with the Bayesian view, is mainly used in empirical analysis, and is also known as the Bayesian information criterion [30].

$$AIC = e^{2k/n} \frac{RSS}{n}, SIC = n^{k/n} \frac{RSS}{n}, \tag{A2}$$

where k is the number of regressors, n is the number of observations and RSS (residual sum of squares) is the sum of square error between the data. The null hypothesis for the ADF unit root test is "including a unit root ($\delta = 1$)." Initially, the present study used the ADF unit root test with the model including the constant and time trend based on the detection cases data. The results are provided in Table A1.

Table A1. ADF unit root test with the model including constant and time trend based on the detection cases data.

			t-Statistic	Prob.
Augmented Dickey-Fuller test statistic			−2.981691	0.1501
Test critical values:	1% level		−4.211868	
	5% level		−3.529758	
	10% level		−3.196411	
Variable	**Coefficient**	**Std. Error**	***t*-Statistic**	**Prob.**
Detection cases (−1)	−0.391532	0.131312	−2.981691	0.0051
Constant	156.6549	60.87809	2.573255	0.0143
@TREND (1970)	−2.307219	1.308371	−1.763428	0.0863

Table A1 shows that the *p*-value of the ADF test statistic, 0.1501, is greater than the significance level (0.05). This means that the null hypothesis cannot be rejected (the detection cases data has a unit root). Testing of the constant and time trend can show variable *Constant* and *@TREND* in the below of Table A2. The *p*-value of the constant is about 0.0143, smaller than the significance level (0.05). That is, the null hypothesis "no constant ($\beta_0 = 0$)" can be rejected. The *p*-value of the trend is 0.0863, again greater than the significance level (0.05). That is, the null hypothesis "no time trend ($\beta_1 = 0$)" also cannot be rejected. The time series data on the detection cases includes the unit root as well as the. Because of the lack of any time trend, we progress to the next step, which is the ADF unit root test with the model including only the constant. The results of this test are summarized in Table A2.

Table A2. ADF unit root test with the model including constant based on the detection cases data.

			t-Statistic	Prob.
Augmented Dickey-Fuller test statistic			−2.343469	0.1641
Test critical values:	1% level		−3.610453	
	5% level		−2.938987	
	10% level		−2.607932	
Variable	**Coefficient**	**Std. Error**	***t*-Statistic**	**Prob.**
Detection cases (−1)	−0.245224	0.104641	−2.343469	0.0246
Constant	66.25393	33.75776	1.962628	0.0572

Table A2 shows that the *p*-value of the ADF test statistic is 0.1641, greater than the significance level (0.05). This result means that the data has a unit root. The *p*-value for constant is 0.0572, again greater than significance level (0.05). That is, the null hypothesis ($\beta_0 = 0$) cannot be rejected. The time series data on the detection cases includes the unit root. Because of no constant, we progress to the final step, which is the ADF unit root test with the model including nothing. The results of the ADF root test are summarized in Table A3.

Table A3. ADF unit root test with the model including nothing based on the detection cases data.

			t-Statistic	Prob.
Augmented Dickey-Fuller test statistic			−1.396253	0.1487
Test critical values:	1% level		−2.625606	
	5% level		−1.949609	
	10% level		−1.611593	
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
Detection cases (−1)	−0.052658	0.037714	−1.396253	0.1707
R-squared	0.038594	Mean dependent var		−7.923077
Adjusted R-squared	0.038594	S.D. dependent var		77.49139
S.E. of regression	75.98132	Akaike info criterion		11.52416
Sum squared resid	219380.1	Schwarz criterion		11.56681
Log likelihood	−223.7211	Hannan-Quinn criter.		11.53946
Durbin-Watson stat	2.689882			

Table A3 shows that the Durbin-Watson statistic is 2.689882 where $k = 1$ and $n = 30$. The significance level (0.05) of these variables sets up as $d_L = 1.352$, $d_U = 1.489$. The null hypothesis "serially uncorrelated" can be rejected, because DW statistics (d) is included between $4 - d_L$ and 4. The data on detection cases presents an eventually negative correlation. *p*-value of the ADF test statistic is 0.1487, greater than the significance level (0.05). This result means that the data has a unit root. In conclusion, the time series data on the detection cases includes the unit root and does not include constant and time trend. In the sequence analysis, we also use an ADF unit root test with the model including the constant and time trend based on the penalization cases data. The results are summarized in Table A4.

Table A4. ADF unit root test with the model including constant and time trend based on the penalization cases data.

			t-Statistic	Prob.
Augmented Dickey-Fuller test statistic			−2.189536	0.4808
Test critical values:	1% level		−4.234972	
	5% level		−3.540328	
	10% level		−3.202445	
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
Penalization cases (−1)	−0.472117	0.215624	−2.189536	0.0365
D (Penalization cases (−1))	−0.158801	0.238570	−0.665637	0.5107
D (Penalization cases (−2))	−0.242738	0.210821	−1.151395	0.2587
D (Penalization cases (−3))	0.243527	0.182994	1.330796	0.1933
Constant	38.11916	18.03466	2.113660	0.0430
@TREND (1970)	−0.233903	0.352704	−0.663171	0.5123

Table A4 shows that the *p*-value of the ADF test statistic, 0.4808, which is very much greater than the significance level (0.05). This means that the null hypothesis cannot be rejected (the penalization cases data has a unit root). The *p*-value of the constant is about 0.0043, smaller than the significance level (0.05). The *p*-value of the trend is 0.5123, greater than the significance level (0.05). The time series data on the penalization cases includes the unit root as well as the constant with the model including the constant and time trend. Because of the lack of any time trend, we progress to the next step, which is the ADF unit root test with the model including only the constant. The results of this test are summarized in Table A5.

Table A5. ADF unit root test with the model including constant based on the penalization cases data.

			t-Statistic	Prob.
Augmented Dickey-Fuller test statistic			−2.131969	0.2339
Test critical values:	1% level		−3.626784	
	5% level		−2.945842	
	10% level		-2.611531	
Variable	**Coefficient**	**Std. Error**	***t*-Statistic**	**Prob.**
Penalization cases (−1)	−0.450100	0.211119	−2.131969	0.0410
D (Penalization cases (−1))	−0.154786	0.236329	−0.654959	0.5173
D (Penalization cases (−2))	−0.229258	0.207934	−1.102552	0.2787
D (Penalization cases (−3))	0.260693	0.179510	1.452247	0.1565
Constant	31.58062	14.96390	2.110454	0.0430
R-squared	0.474430	Mean dependent var		0.611111
Adjusted R-squared	0.406615	S.D. dependent var		27.22633
S.E. of regression	20.97285	Akaike info criterion		9.052581
Sum squared resid	13635.67	Schwarz criterion		9.272514
Log likelihood	−157.9465	Hannan-Quinn criter.		9.129343
F-statistic	6.995902	Durbin-Watson		2.098929
Prob (F-statistic)	0.000391			

Table A5 shows that the Durbin-Watson statistic is 2.098929 where $k = 1$ and $n = 30$. The significance level (0.05) of these variables sets up as $d_L = 1.352$, $d_U = 1.489$. The null hypothesis "serially uncorrelated" cannot be rejected, because DW statistics (d) is included between d_U and $4 - d_U$. The data on penalization cases eventually resulted in no correlation. It shows that the *p*-value of the ADF test statistic is 0.2339 greater than the significance level (0.05). This result means that the data has a unit root. The *p*-value for constant is 0.043, greater than the significance level (0.05). That is, null hypothesis ($\beta_0 = 0$) can be rejected. Therefore, we finish the steps. The time series data about penalization cases includes unit root and constant.

References

1. Bryant, P.G.; Eckard, E.W. Price Fixing—The Probability of Getting Caught. *Rev. Econ. Stat.* **1991**, *73*, 531–536. [CrossRef]
2. Combe, E.; Monnier, C.; Legal, R. Cartels: the Probability of Getting Caught in the European Union. Bruges European Economic Research (BEER) Papers 12/March 2008. Available online: http://aei.pitt.edu/58577/ (accessed on 8 June 2018).
3. Harrington, J.E.; Chang, M.H. Modeling the Birth and Death of Cartels with an Application to Evaluating Competition Policy. *J. Eur. Econ. Assoc.* **2009**, *7*, 1400–1435. [CrossRef]
4. Zhou, J. *Evaluating Leniency with Missing Information on Undetected Cartels: Exploring Time-Varying Policy Impacts on Cartel Duration*; TILEC Discussion Paper; Tilburg University—Tilburg Law and Economics Center (TILEC): Tilburg, The Netherlands, 2011.
5. Gallager, R.G. *Stochastic Processes: Theory for Applications*; Cambridge University Press: Cambridge, UK, 2013.
6. Li, X.R. *Probability, Random Signals, and Statistics*; CRC Press: Boca Raton, FL, USA, 1999.
7. Miller, N.H. Strategic Leniency and Cartel Enforcement. *Am. Econ. Rev.* **2009**, *99*, 750–768. [CrossRef]
8. Hinloopen, J. Internal cartel stability with time-dependent detection probabilities. *Int. J. Ind. Organ.* **2006**, *24*, 1213–1229. [CrossRef]
9. Ormosi, P.L. A Tip of the Iceberg? The Probability of Catching Cartels. *J. Appl. Econ.* **2014**, *29*, 549–566. [CrossRef]
10. Armstrong, J.S. Judgmental bootstrapping: Inferring experts' rules for forecasting. In *Principles of Forecasting*; Springer: Berlin, Germany, 2001; pp. 171–192.
11. Harrington, J.E. Behavioral screening and the detection of cartels. In *European Competition Law Annual, 2006*; Hart Publishing: Oxford, UK, 2006; pp. 51–69.

12. Bos, I.; Harrington, J.E. Endogenous cartel formation with heterogeneous firms. *Rand J. Econ.* **2010**, *41*, 92–117. [CrossRef]
13. Friedman, J.W. A non-cooperative equilibrium for supergames. *Rev. Econ. Stud.* **1971**, *38*, 1–12. [CrossRef]
14. Harrington, J.E. Optimal corporate leniency programs. *J. Ind. Econ.* **2008**, *56*, 215–246. [CrossRef]
15. Berger, J.O. *Statistical Decision Theory and Bayesian Analysis*; Springer: Berlin, Germany, 1985.
16. Da-Silva, C.Q.; Migon, H.S.; Correia, L.T. Dynamic Bayesian beta models. *Comput. Stat. Data Anal.* **2011**, *55*, 2074–2089. [CrossRef]
17. Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*; CRC Press: Boca Raton, FL, USA, 2013.
18. Ahn, S.E.; Park, C.S.; Kim, H.M. Hazard rate estimation of a mixture model with censored lifetimes. *Stoch. Environ. Res. Risk Assess.* **2006**, *21*, 711–716. [CrossRef]
19. Koop, G.; Poirier, D.J. Bayesian-Analysis of Logit-Models Using Natural Conjugate Priors. *J. Econ.* **1993**, *56*, 323–340. [CrossRef]
20. Guérin, F.; Dumon, B.; Usureau, E. Reliability estimation by Bayesian method: Definition of prior distribution using dependability study. *Reliab. Eng. Syst. Saf.* **2003**, *82*, 299–306. [CrossRef]
21. Antitrust Division. *10-Year Workload Statistics*; Department of Justice: Washington, DC, USA, 1970–2009.
22. Gujarati, D. *Basic Econometrics*; MeGraw-Hill: New York, NY, USA, 2008.
23. Nelson, C.R.; Plosser, C.R. Trends and random walks in macroeconmic time series: Some evidence and implications. *J. Monet. Econ.* **1982**, *10*, 139–162. [CrossRef]
24. Chang, M.-H.; Harrington, J.E. *The Impact of a Corporate Leniency Program on Antitrust Enforcement and Cartelization*; Department of Economics, The Johns Hopkins University: Baltimore, MD, USA, 2008.
25. Pindyck, R.S.; Rubinfeld, D.L. *Econometric Models and Economic Forecasts*; McGraw-Hill: New York, NY, USA, 1981.
26. Sims, C.A. Policy Analysis with Econometric-Models. *Brook. Pap. Econ. Act.* **1982**, *1982*, 107–164. [CrossRef]
27. Brenner, S. An empirical study of the European corporate leniency program. *Int. J. Ind. Organ.* **2009**, *27*, 639–645. [CrossRef]
28. Dickey, D.A.; Rossana, R.J. PRACTITIONERS' CORNER Cointegrated Time Series: A Guide to Estimation and Hypothesis Testing. *Oxf. Bull. Econ. Stat.* **2009**, *56*, 325–353. [CrossRef]
29. Dolado, J.J.; Jenkinson, T.; Sosvilla-Rivero, S. Cointegration and Unit Roots. *J. Econ. Surv.* **1990**, *4*, 249–273. [CrossRef]
30. Koehler, A.B.; Murphree, E.S. A Comparison of the Akaike and Schwarz Criteria for Selecting Model Order. *Appl. Stat. J. R. Stat. Soc. Ser. C* **1988**, *37*, 187–195. [CrossRef]

Toward a More Resilient Financial System: Should Banks Be Diversified?

Laura Baselga-Pascual [1,*], Olga del Orden-Olasagasti [1] and Antonio Trujillo-Ponce [2]

[1]   Department of Finance and Accounting, University of Deusto, Mundaitz Kalea, 50,
      20012 Donostia, Spain; olgadelorden@deusto.es
[2]   Department of Financial Economics and Accounting, Universidad Pablo de Olavide,
      Ctra. de Utrera km. 1, 41013 Seville, Spain; atrujillo@upo.es
[*]   Correspondence: lbaselga@deusto.es; Tel.: +34-943-326-600

Abstract: This article empirically analyzes the effects of revenue diversification on the profitability and risk of a large sample of Eurozone banks over the period from 2000 to 2012. We use the generalized method of moments (GMM) estimator, which is also referred to as the system-GMM estimator. We conclude that higher income diversification favors bank profitability. However, our study does not find a significant relationship between revenue diversification and bank risk, even when considering a crisis period. Our results suggest that establishing restrictions in the universal banking model could damage the resilience of the financial system, and thus affect the sustainability of the uneven economic recovery in Europe.

Keywords: diversification; bank profitability; bank risk; dynamic panel; European banking system; sustainability of economic recovery

---

1. Introduction

In the past, bankers believed that they could reduce earnings volatility by diversifying into activities that are imperfectly correlated with traditional banking income [1]. In this vein, studies from Johnson and Meinster [2], Heggestad [3], Wall and Eisenbeis [4], and Litan [5], among others, concluded that banks offering nonbanking products significantly reduced their risk without decreasing their expected returns.

This "conventional wisdom", which is supported by the early literature, initiated a deregulation process in the banking industry in the late 1980s by progressively revoking the Banking Act, known as the Glass-Steagall Act (GSA), from 1933. The deregulation process enabled banks to diversify their asset portfolios and to significantly increase their profits by fostering the emergence of numerous alternative sources of income, such as underwriting, securities trading, brokerage and investment banking, as well as other untraditional banking activities [6,7], leading to the so-called "global financial supermarkets" and "global banks", the most representative company of which was Citigroup. This deregulated scenario, along with an expansionary monetary policy (with low interest rates and easy loans), aimed at alleviating an economic collapse from the burst of the Dotcom bubble at the beginning of the 21st century, is considered one of the main causes of the subprime mortgage and financial crisis (2007–2009) that started in the United States (U.S.) and rapidly spread to Europe [8].

Recommendations for avoiding a repeat of the financial and economic crisis face the challenge of how to meaningfully contribute to a sustainable economic recovery with the help of new banking business models. In response to the global financial crisis, the High-Level Expert Group on reforming the structure of the European Union (EU) banking sector issued a report in 2012 (the so-called Liikanen report), which recommended that proprietary trading and market-making activities be separated from other banking activities [9]. The idea of isolating certain types of activities that are considered especially

important to the real economy from other riskier, but less important, activities have also been shared by other recent proposals, such as the "Volcker rule" in the United States and the Vickers Commission report in the United Kingdom. These proposals imply restrictions on the universal banking model, in which banks offer a full range of financial services, and thus could lead to less diversified banks [10].

This article complements the existing literature in various ways. First, this work contributes to the ongoing debate surrounding bank diversification by focusing on the Eurozone as an interesting case study. Countries in the Eurozone have become increasingly integrated and set apart from other parts of the EU by their economic management since 1999, following the establishment of the euro. Moreover, the sovereign debt crisis in 2010–2011 highlighted the greater interdependence of countries in the monetary union and emphasized the need to create an integrated financial framework to restore confidence in banks and the euro. The Banking Union, which was initiated in 2012 with the agreement on the establishment of a Single Supervisory Mechanism with the involvement of the European Central Bank (ECB), aims to deliver an integrated financial safety net for these countries. Second, the selected time span, from 2000 to 2012, considers the impact of both the 2008 financial crisis and the sovereign debt crisis starting in 2010 in the European banking sector. As the impact of the financial crisis on the real economy has urged policy makers and regulators to drastically change the "rules of the game" by proposing limits to banking activities, it is important to provide further insights into the effect of revenue diversification on bank performance. Most studies have provided evidence of the effect of revenue diversification under normal economic conditions ([11–13] among others). Third, we use a proxy for revenue diversification that reflects the balance of different types of income (interest, net commissions, trades, and other operating income). Finally, we use the generalized method of moments (GMM) estimator, which was developed for dynamic panel models by Arellano and Bover [14] and Blundell and Bond [15], which is also known as the system-GMM estimator. The system-GMM estimator for dynamic panel data models combines moment conditions for the model in first differences with moment conditions for the model in levels. This method has been shown to improve on the GMM estimator in the first-differenced model in terms of bias and root mean squared error [16]. The system-GMM estimator has also less bias and greater efficiency than either ordinary least squares (OLS) regressions or static panel data models (fixed effects or random effects).

The article is structured as follows: after this introduction, Section 2 reviews the literature; Section 3 describes the data and methodology employed in the empirical research; Section 4 presents and discusses the results; and, Section 5 summarizes and concludes the article.

2. Literature Review

Theory provides conflicting predictions about the impact of revenue diversification on the performance of banks. According to Diamond [17] and Stein [18], banks acquire relevant information from their clients in the process of making loans, facilitating the efficient provision of other less traditional services, including underwriting, securities trading, or insurance. Similarly, nontraditional banking activities can produce information that improves lending [19]. Moreover, the sharing of inputs, such as labor, technology, and information across multiple outputs also constitutes a source of potential cost savings for diversified banks [11]. Therefore, part of the literature suggests that revenue diversification provides economies of scope that enhance the profitability of financial institutions. Nevertheless, diversification may reduce bank profits due to agency problems [20] (not only between managers and shareholders, but also between the divisions of the bank and between the bank and its customers in the form of conflicts of interest [11]); regulatory costs that are associated with multiple supervision [11]; or an inefficient resource allocation between different business segments due to a malfunctioning of internal capital markets [21]. Similarly, diversified activities have different theoretical implications for bank risk. Although, according to the portfolio theory, banks may get risk diversification benefits if noninterest income streams are uncorrelated with interest income, diversification can expose banks to new forms of risks (in addition to credit risk), such as market, liquidity, and operational risk [22]. Furthermore, diversified banks may operate with lower capital

ratios and pursue riskier activities because many fee-based activities can be performed while holding little or no regulatory capital [23].

The empirical literature has also reported the opposing effects of revenue diversification on bank performance. Chiorazzo et al. [24] concluded that noninterest income increases risk-adjusted returns for a sample of Italian banks, with this relationship stronger for larger banks. Köhler [25] concluded that retail-oriented banks, such as savings, cooperative, and other banks that focus on lending and deposit services, become significantly more stable (in the sense of having higher Z-scores) if they increase their share of noninterest income. Köhler [26] analyzed the impacts of business models on bank stability in 15 EU countries between 2002 and 2011, indicating that substantial benefits can be gained from income diversification. Brighi and Venturelli [27] reported evidence suggesting that greater diversification among different fee and commission components decreases bank risk and increases risk-adjusted profitability. Elsas et al. [28] showed that diversification increases bank profitability and market valuation. Lee et al. [29] concluded that bank performance can be improved through diversification, confirming the hypothesis of the portfolio diversification effect for the Asia-Pacific banking industry. Sanya and Wolfe [30] and Meslier et al. [9] focused on banks in emerging economies and provided empirical evidence that an observed shift toward noninterest income-generating activities has a positive effect on bank performance and it decreases the risk of insolvency. Finally, the benefits of revenue diversification have also been confirmed in some African countries [31,32].

Nevertheless, there is also a large body of empirical studies concluding that diversification has detrimental effects on bank performance. Mercieca et al. [12] found an inverse association between noninterest income and performance for small banks. Lepetit et al. [13] showed that expansion into noninterest income activities, particularly into activities with increased commissions and fees and increases insolvency risk. De Jonghe [33] stated that noninterest income-generating activities increased the systemic risk of Eurozone banks over the period of 1992–2007. DeYoung and Rice [34] and Stiroh and Rumble [35] showed that increased noninterest income is associated with poorer risk-return tradeoffs in U.S. banks. Saona [36] stated that noninterest income activities are negatively and significantly correlated with profitability in Latin American banks. Laeven and Levine [19] examined an international sample of financial conglomerates and found that their respective market values reflect a diversification discount; i.e., the market values are lower than if those financial conglomerates were broken up into financial intermediaries that specialize in the individual activities. Finally, Demirgüç-Kunt and Huizinga [37] and Baele et al. [11] reported negative effects of diversification on bank risk, although they found positive effects on asset returns and franchise values, respectively.

Table 1 provides a summary of the main findings on revenue diversification in the empirical literature.

Table 1. Summary of the findings on the impacts of revenue diversification on bank profitability and bank risk.

References	Sample	Period	Empirical Findings
[7]	Philippines	1999–2005	A shift toward noninterest activities increases bank profits and risk-adjusted profits, particularly when banks are more involved in trading in government securities.
[11]	15 EU countries + Norway and Switzerland	1989–2004	A higher share of noninterest income among total income affects banks' franchise values positively. Diversification of revenue streams from distinct financial activities increases the systematic risk of banks, while the effect on the idiosyncratic risk component is nonlinear and predominantly downward sloping.
[12]	Europe (small banks)	1997–2003	A shift from interest income activities to noninterest income activities results in lower performance of small European banks, caused by weak monitoring of activities that lie outside of their traditional lending business.
[13]	Europe	1996–2002	Banks expanding into noninterest income activities present greater risk and greater insolvency risk than banks that mainly supply loans. The positive link with risk is mostly accurate for small banks and is essentially driven by commissions and fee activities.
[19]	International (43 countries)	1998–2002	The market values of financial conglomerates that engage in multiple activities (i.e., lending and non-lending financial services) are lower than if those financial conglomerates were broken up into financial intermediaries that specialize in individual activities.
[24]	Italy	1993–2003	Income diversification increases risk-adjusted returns, with this relationship stronger with larger banks. Small banks can profit from increasing noninterest income but only when they have very little noninterest income share in the beginning.
[25]	Germany	2002–2012	The impact of noninterest income on bank risk differs between retail- and investment-oriented banks: while retail-oriented banks become significantly more stable (in the sense of having a higher Z-score), if they increase their share of noninterest income, investment-oriented banks become significantly riskier.
[26]	15 EU countries	2002–2011	Banks are significantly more stable and profitable if they increase their share of noninterest income. Such benefits are particularly large for savings and cooperative banks. Investment banks, in contrast, become significantly riskier.
[27]	Italy	2006–2012	Greater revenue diversification decreases bank risk and increases risk-adjusted profitability, especially after crisis periods.
[28]	U.S., Canada, Australia and Europe	1996–2008	Diversification increases bank profitability and, as a consequence, market valuations. This evidence against a conglomerate discount in banking also remained robust during the subprime mortgage crisis.
[29]	Asia-Pacific	1995–2009	Bank performance can be improved through diversification in bank-based groups. Under different financial systems, the relationships among revenue diversity, financial reforms and bank performances are multidimensional.
[30]	226 listed banks across 11 emerging economies	2000–2007	Revenue diversification decreases insolvency risk and enhances profitability.
[31]	Africa	1996–2014	There is a positive relationship between bank market power and income from nontraditional activities.
[32]	Africa	2002–2013	Cross-border banking and revenue diversification within banks in Africa can create value.
[33]	15 EU countries	1992–2007	Noninterest-generating activities increase banks' tail beta.
[34]	U.S.	1989–2001	Small, well-managed banks generate less noninterest income. Marginal increases in noninterest income are associated with higher and more variable profits and with worsening of the risk-return trade-off for the average commercial bank.
[35]	U.S.	1997–2002	Diversification benefits exist between Financial Holding Companies (FHCs), but these gains are offset by increased exposure to noninterest activities, which are much more volatile but not necessarily more profitable than interest-generating activities.
[36]	Latin America	1995–2012	There is a negative relationship between bank profitability and revenue diversification.
[37]	International (101 countries)	1995–2007	Expansion into noninterest income-generating activities, such as trading, increases the rate of return on assets, and it could offer some risk diversification benefits at very low levels. Overall, banking strategies that rely prominently on generating noninterest income are very risky.

3. Data and Methodological Aspects

3.1. Sample

We use a sample consisting of an unbalanced panel with 4268 observations that were obtained from commercial banks, savings banks, and credit cooperatives operating in 14 European countries over the period from 2000 to 2012. We restrict the analysis to countries that adopted the euro during the sample period and thus have a common monetary policy: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Slovakia, Slovenia, and Spain (see Table 2). Because banks must have information available for all of the variables analyzed for at least five consecutive years to test the second-order serial correlation when the system-GMM methodology is used [38], we exclude some countries (e.g., Estonia). Unlike previous studies that focused only on listed banks (e.g., [8]), our sample includes all (listed and unlisted) banks in the euro area (commercial, savings, and cooperative banks), which affords us a more representative picture of the European banking sector, as unlisted banks account for the majority of banks in this area [26].

Table 2. Observations in the sample by country.

Country	Observations
Austria	57
Belgium	29
Finland	39
France	700
Germany	1156
Greece	50
Ireland	51
Italy	1642
Luxembourg	32
Netherlands	77
Portugal	97
Slovakia	40
Slovenia	90
Spain	208
Total Eurozone	4268

Bank balance sheet data and income statements are obtained from the BankScope database that was maintained by Bureau Van Dijk (now Orbis Bank Focus). Indicators are calculated on a nonconsolidated basis, indicating that banking subsidiaries and foreign branches are considered to be separate credit institutions. This assumption reduces the possibility of introducing aggregation bias into the results [39]. Merged banks are considered as separate entities before the merger and a single entity subsequently. All of the ratios capturing bank-specific characteristics are calculated based on the standardized global accounting format. Entities that present abnormal ratios or extreme values are eliminated from the sample as outliers in order to ensure that the analysis is not affected by potential measurement errors and misreporting. The frequency of the data is annual.

We obtain the data on industry concentration from the Banking Structural Financial Indicators database of the European Central Bank (ECB), whereas macroeconomic data are obtained from Eurostat.

3.2. Methodology

Bank-specific factors determining bank profitability and risk can be endogenous. For example, more profitable banks can have more resources to increase their equity; such banks might also find it easier to increase their customer base through successful advertising, and thereby enhance their profitability [40]. Similarly, banks could have incentives to increase their stock of liquid assets if they become riskier to protect themselves against premature withdrawals of funds [26]. In addition to endogeneity, some characteristics that affect bank performance are difficult to measure or to identify

in an equation (so-called unobserved heterogeneity). If the influence of such characteristics is not considered, then one could observe correlations between some of the coefficients of the explanatory variables and the error terms that bias these coefficients. Finally, the persistence of profitability and risk has been well documented in the literature. To address these concerns, we use the GMM estimator that was developed for dynamic panel models by Arellano and Bover [14] and Blundell and Bond [15], also referred to as the system-GMM estimator. This methodology is a better alternative than traditional panel data estimators, which can produce biased and inconsistent estimates when being applied to our equation. Following Windmeijer [41], we use the two-step estimation procedure with finite sample corrected standard errors, providing less biased coefficient estimates and more accurate standard errors.

Our baseline equation is as follows:

$$Y_{i,j,t} = \alpha + \delta \cdots Y_{i,j,t-1} + \beta \cdots DIV_{i,j,t} + \lambda \cdots BS_{i,j,t} + \gamma \cdots M_{j,t} + \theta \cdots R_{j,t-1} + \eta \cdots D_{i,j,t} + \varepsilon_{i,j,t} \quad (1)$$

where Y denotes the variable used to measure either the profitability or the risk of bank i in country j in year t; $Y_{i,t-1}$ represent their lagged values; δ measures the speed of adjustment (a value of δ close to 0 implies that bank profitability/risk is characterized by a high speed of adjustment, whereas a value that is close to 1 indicates that the adjustment is very slow); DIV represents a proxy of revenue diversification; $BS_{i,j,t}$ and $M_{j,t}$ denote, respectively, the bank-specific and the industry and macroeconomic variables that were considered in our study; $R_{j,t-1}$ refers to several bank regulation and supervision control variables; and, $D_{i,j,t}$ represents dummy variables controlling for the public status of the bank and the bank type. β, λ, γ, θ, and η are vectors of coefficient estimates. Finally, $\varepsilon_{i,t}$ is the disturbance term that contains the unobserved bank-specific effect (η_i) and the idiosyncratic error ($\nu_{i,t}$).

3.3. Measures of Profitability and Risk

More than 90% of the banks in our sample are unlisted, providing us with a broad representation of the Eurozone banking system, but requiring us to use accounting-based, instead of market-based, indicators. However, we acknowledge that accounting metrics have limitations. For instance, managers could use some timing discretion over these metrics to minimize regulatory costs. They are also backward looking [42].

We consider the return on average assets (ROA) as our proxy for bank profitability, which is the single most important ratio for comparing the efficiency and operational performance of banks [40]. ROA is computed as pretax profits that are divided by total assets. This ratio considers the returns that are generated from the assets that a bank finances; it is primarily an indicator of managerial efficiency, although it can be misleading as a result of off-balance-sheet activities. Our measure of bank risk is the Z-score, defined as the number of standard deviations that a bank's return on assets must fall below the mean for the bank to become insolvent [43]. Therefore, this index can be interpreted as an inverse measure of the probability of insolvency; i.e., a higher Z-score indicates that a bank incurs fewer risks and is more stable [26]. The Z-score is considered to be a better measure of bank risk than the nonperforming loan ratio (NPLr) because nonperforming loans are traditionally backward looking and highly procyclical [44,45]. In addition, the Z-score is an overall measure of bank risk that captures more than credit risk alone [46]. The Z-score is calculated as the sum of ROA and equity-to-assets ratio (Eq/TA), divided by the three-year standard deviation of ROA (SD_ROA):

$$\text{Z-score}_t = \frac{\text{ROA}_t + \text{Eq/TA}_t}{\text{SD}_{\text{ROA}_t}} \quad (2)$$

Finally, we calculate natural logarithms to control for the skewness that was exhibited by the original metric.

3.4. Revenue Diversification

Following Elsas et al. [28], we proxy revenue diversification using a Herfindahl–Hirschman index (HHIRD) that reflects the balance of different types of income, such as interest, net commissions, trading, and other operating income. HHIRD is calculated, as follows:

$$\text{HHIRD} = 1 - \left[\left(\frac{\text{INT}}{\text{TOR}} \right)^2 + \left(\frac{\text{COM}}{\text{TOR}} \right)^2 + \left(\frac{\text{TRAD}}{\text{TOR}} \right)^2 + \left(\frac{\text{OTH}}{\text{TOR}} \right)^2 \right] \tag{3}$$

where INT denotes gross interest income; COM denotes gross commissions and fee revenue; TRAD denotes trading revenue; and, OTH denotes all other gross operating income. TOR represents total operating revenue and is equal to the sum of the absolute values of INT, COM, TRAD, and OTH. HHIRD can take values between 0 (no revenue diversification) and 75 (indicating a bank that generates a fully balanced revenue mix from all four business areas).

3.5. Control Variables

We include in our regression several bank-specific variables, the influences of which on bank profitability and risk have been widely contrasted in the literature (see [26,40,43,47]). We control for the bank's asset structure using the ratio of loans to total assets (Loan/TA), as the literature suggests that this ratio is positively correlated with bank returns [7] and risk [12,43]. The effect of capitalization on bank performance is controlled for using Eq/TA. We expect a positive effect of capitalization on ROA. There appears to be a consensus in the previous literature that more capital (and therefore, better solvency) reduces the costs of external debt, compensating for the higher costs of one's own funds [40]. Consistent with previous studies (e.g., [23,35]), a negative relationship between bank capitalization and risk can be expected due to the potential danger of leverage. The recent financial crisis has clearly exposed the risks of a bank's excessive reliance on non-deposit funding; thus, following Laeven and Levine [19] and Köhler [26], among others, we include in our equation the ratio of non-deposit funds to total liabilities (NonDep/TL). We control for operational efficiency using the cost-to-income ratio (CIR), which can be negatively related to bank returns [23] but positively related to risk [47]. The natural logarithm of bank assets accounts for the effect of size on bank performance (Size) The effect of size on bank returns could be nonlinear, with profitability initially increasing with size and then declining for bureaucratic and other reasons [48]. However, the effect of size on bank risk-taking remains unclear: although larger banks might incur more risk due to a moral hazard problem [26,33], they might be less prone to risk because of their managerial capacity and efficiency [12]. We also include the NPLr to account for credit risk in the profitability equation [40]; and the ROA to account for profitability in the risk equation [43].

We additionally control for industry concentration, which is measured in terms of the Herfindahl–Hirschman index (HHIIC), which is calculated as the sum of the squares of all credit institutions' market shares within a country in terms of total assets (in percentages). We further include selected macroeconomic variables, following previous studies: (i) the annual growth rate of the real gross domestic product of the country (GDP), which controls for the effect of economic growth in our regressions [40,43]; (ii) the consumer price index annual average rate of change (Inflation), as inflation could affect both bank profitability and bank risk [26,40]; (iii) the annual average unemployment rate (UR), which could negatively affect the returns and the risks of banks [43,47]; and, (iv) the interest rate of the main refinancing operations of the European Central Bank (Interest). An environment of low interest rates can exert pressure on the operating margin and negatively affect banks' profitability [49,50]. It could also affect bank risk [43].

Moreover, as there still might be some differences in the regulatory and supervisory environments between each country's banking system inside the Eurozone [43], we use four indices from the World Bank database on "Bank Regulation and Supervision", as developed by Barth et al. [51] as regulation controls in our equation specification (see Appendix A). This database is based on four

surveys conducted by the World Bank (Survey I was released in 2001, and for most of the countries, the information corresponds to 1999; Survey II describes the regulatory situation at the end of 2002; Survey III describes the regulatory environment in 2005–2006; Survey IV provides information about bank regulation and supervision in 125 countries for 2011 (with some corrections in 2012) [52]). The capital stringency index evaluates the regulatory approach to assessing and verifying the degree of capital at risk in a bank; the supervisory power metric reflects the degree to which the country's bank supervisory agency has the authority to undertake specific actions (e.g., force a bank to change its internal organizational structure); the private monitory index shows the degree to which banks are forced to disclose accurate information to the public and whether there are incentives to increase market discipline; and finally, the activity restrictions index measures the degree to which banks face regulatory restrictions on their activities in securities markets, insurance, and real estate, as well as on owning shares in nonfinancial firms, with higher values indicating greater restrictions [36]. We use the first lag of these indices, as regulative initiatives are unlikely to affect bank profitability and/or bank risk immediately [39,46].

Finally, we control for the public status of the bank—as market discipline exerted by the stock market might influence bank performance [53]—and for the bank type—as commercial banks, savings banks, and credit cooperatives have different business models, objectives, and ownership structures [26].

Table 3 summarizes the variables considered in the current study.

Table 3. Variables considered in the study.

Classification	Variable	Notation	Data Source
Profitability	Return on Assets (%)	ROA	BankScope
Bank risk	Z-score (in logarithmic form)	Z-score	BankScope/Own elaboration
Revenue diversification	Herfindahl–Hirschman Index	HHIRD	BankScope/Own elaboration
Asset structure	Loans/Total Assets (%)	Loan/TA	BankScope
Capitalization	Equity/Total Assets (%)	Eq/TA	BankScope
Non-deposit funding	Non-deposit Funds/Total Liabilities (%)	NonDep/TL	BankScope
Efficiency	Cost-to-Income Ratio (%)	CIR	BankScope
Size	Total Assets (in logarithmic form)	Size	BankScope
Credit risk	Nonperforming Loan Ratio (%)	NPLr	BankScope
Industry concentration	Herfindahl-Hirschman Index	HHIIC	ECB
Economic growth	Annual Real GDP Growth Rate (%)	GDP	Eurostat
Inflation	Annual Average Rate Change in CPI (%)	Inflation	Eurostat
Unemployment	Unemployment Rate (%)	UR	Eurostat
Interest rates	Interest Rate on the MRO of the ECB (%)	Interest	ECB
Regulation and supervision	Capital Stringency Index, Supervisory Power Index, Private Monitoring Index, Activity Restrictions Index	Regulatory indices	[51,52,54]
Listed	Dummy variable, taking the value of 1 for listed bank and 0 otherwise	Listed dummy	BankScope/Own elaboration
Bank type	Two dummy variables, taking the values of 1 for commercial banks (savings banks) and 0 otherwise	Bank type dummies	BankScope/Own elaboration

3.6. Endogeneity

We acknowledge that there is a problem of endogeneity in our regression. To address it, the system-GMM estimator uses suitable instruments. In line with Arellano and Bover [14] and Blundell and Bond [15], we employ lagged first differences of the bank-specific variables as instruments for the equation in levels and the lagged values of these variables in levels as instruments for the equation in differences. Regulatory and supervisory indicators are also treated as endogenous, as we assume that the regulators may change banking rules to prevent financial turbulence if they observe low profitability and/or excess risk taking [39,43]. The public status of the bank is considered to be endogenous in our equation, as banks may choose to become listed or not, based on the expected future changes in profitability and/or risk [43]. The industry concentration, macroeconomic variables,

and bank type dummies are treated as exogenous. We verify that the instruments are statistically valid using Hansen's *J*-test of over identifying restrictions.

4. Empirical Results

4.1. Results from the Baseline Model

Table 4 presents descriptive statistics of the variables that are considered in this paper. The average ROA for banks in the sample is 0.48% with a standard deviation of 0.82%. As reported by Baselga-Pascual et al. [43], profitability levels have remained low since the onset of the crisis in 2008 and they have been characterized by a high degree of heterogeneity. Banks have, on average, a Z-score (calculated in logarithmic form) of 3.58, with a standard deviation of 1.01. The index of revenue diversification has a mean value of 35.89, indicating that Eurozone banks present a certain degree of diversification in the origin of their income (interest, commissions and fees, trading, and other operating income). Most of the control variables have the expected values, demonstrating in the macroeconomic variables the deterioration of the economic situation in the Eurozone.

Table 4. Descriptive statistics.

Variable	Observations	Mean	Standard Deviation	Minimum	Median	Maximum
ROA	4268	0.48	0.82	−9.03	0.48	9.24
Z-score	3291	3.58	1.01	−1.58	3.60	9.20
HHIRD	4268	35.89	11.32	4.74	35.50	74.17
Loan/TA	4268	65.61	16.56	0.29	69.16	97.72
Eq/TA	4268	9.92	4.90	0.25	9.23	73.98
NonDep/TL	4268	43.10	17.11	1.02	44.16	100.00
CIR	4268	66.08	14.66	6.50	65.45	307.14
Size	4268	14.15	2.30	10.00	13.60	21.50
NPLr	4268	7.35	5.58	0.00	6.11	62.33
HHIIC	4268	485.52	361.33	151	407	3700
Inflation	4268	2.13	1.01	−4.48	2.07	5.65
GDP	4268	−0.25	2.71	−8.54	0.48	6.87
UR	4268	8.38	2.72	3.10	8.40	24.80
Interest	4268	1.84	1.19	0.75	1.00	4.75

Table 5 reports the dynamic panel data regressions for measures of both bank profitability (ROA) and bank risk (Z-score) using the system-GMM estimator. The high statistical significance in the lagged dependent variables confirms the dynamic character of the model specification.

We show a direct and significant relationship between bank profitability and our proxy of bank diversification. This result is consistent with the part of the literature considering that increasing the share of noninterest activities in banks could be beneficial for returns (e.g., [23,29]). However, we do not find evidence to conclude that the effect of income diversification on bank stability is negative. Therefore, greater diversification could lead banks to obtain greater profits without necessarily implying an increase in their risk.

Regarding control variables in the profitability equation, we find that the effect of capitalization on ROA is positive and highly significant. Athanasoglou et al. [48] stated that this positive impact could be the result of capital acting as a safety net in the cases of adverse developments. This relationship would help banks to finance their assets at more favorable interest rates, increasing expected profitability, and offsetting the cost of equity. Our results also show that non-deposit funding reduces bank profitability (as reported by Laeven and Levine [19] and Saona [36]). The negative sign of the CIR variable and its high statistical significance confirm that improvements in efficiency are translated into improvements in profitability.

Table 5. The effect of diversification on the profitability and risk of Eurozone banks (2000–2012).

Variables	ROA	Z-Score
Dep. Var.$_{t-1}$	0.421 **	0.477 ***
	(0.203)	(0.034)
HHIRD	**0.008 *****	**−0.006**
	(0.002)	**(0.004)**
Loan/TA	0.002	−0.006 **
	(0.003)	(0.003)
Eq/TA	0.037 ***	0.029 ***
	(0.013)	(0.008)
NonDep/TL	−0.006 **	−0.008 **
	(0.003)	(0.004)
CIR	−0.025 ***	−0.001
	(0.002)	(0.003)
Size	−0.074 **	0.028
	(0.030)	(0.057)
NPLr	−0.076 ***	–
	(0.011)	
ROA	–	0.284 ***
		(0.057)
HHIIC	0.000 ***	0.000 *
	(0.000)	(0.000)
Inflation	0.039 *	−0.035 *
	(0.021)	(0.020)
GDP	0.011 *	0.061 ***
	(0.006)	(0.008)
UR	−0.039 ***	−0.019 *
	(0.011)	(0.011)
Interest	0.073 ***	0.035 *
	(0.019)	(0.020)
Regulatory Indexes	Yes	Yes
Listed dummy	Yes	Yes
Bank type dummies	Yes	Yes
Constant	3.725 ***	1.664 *
	(0.653)	(0.877)
z_1	1269.62 (20)	802.58 (20)
m_1	−4.07	−6.13
m_2	−0.81	−1.05
Hansen	444.96 (444)	443.17 (444)

Notes: This table reports the determinants of profitability and risk of Eurozone banks over the period of 2000 to 2012 using the system-generalized method of moments (GMM) estimator. The sample comprises 4268 observations (3291 in the Z-score equation). See Table 3 for a description of the variables. With the exception of industry concentration, macroeconomic variables and bank type dummies, all of the independent variables in our model are considered endogenous. We report heteroskedasticity-consistent asymptotic standard errors in parentheses, and significance levels are indicated as follows: *** = significant at the 1% level; ** = significant at the 5% level; and * = significant at the 10% level. z_1 is a Wald test of the joint significance of the reported coefficients, asymptotically distributed as χ^2 under the null of no significance, with degrees of freedom in parentheses. m_i is a serial correlation test of order i using residuals in first differences, asymptotically distributed as $N(0,1)$ under the null of no serial correlation. Hansen is a test of over-identifying restrictions, asymptotically distributed as χ^2 under the null of no correlation between the instruments and the error term, with degrees of freedom in parentheses.

We report that the effect of size on bank profitability is negative. In principle, one would expect that larger banks experience more significant increases in profitability through economies of scale. However, beyond a certain threshold of size, diseconomies of scale can arise, rendering the size of a bank detrimental to its profitability [40]. The effect of NPLr on profitability is negative (see, e.g., [48]). There appears to be a consensus that an increase in doubtful assets, which do not accrue income, requires a bank to allocate a significant portion of its gross margin to provisions to cover expected credit losses; thus, profitability will be lower. Our results suggest a positive relationship between bank concentration and profitability in the Eurozone. As could be expected, bank profitability is directly

related to GDP growth and is inversely related to unemployment rates. Inflation positively affects bank profitability, indicating that managers might anticipate inflation expectations and adjust interest rates to achieve greater profits [40]. Finally, bank profitability shows a positive correlation with the interest rate on the MRO of the ECB.

We also find significant relationships between the control variables and the Z-score in the risk equation. We report a positive relationship between the relative percentage of loans in the assets of a bank and its risk, based on the literature finding that laxity in lending can be a source of banking problems (e.g., [55]). Our results confirm that the higher the capitalization, the lower the bank risk is. We report that a greater dependence on wholesale funding can increase bank risk. The financial crisis has clearly exposed the risk of wholesale funding, especially after the bankruptcy of Lehman Brothers in 2008 [43]. We also find an inverse relationship between bank profitability and risk (as reported by Baselga-Pascual et al. [43]). With regard to the exogenous variables, we show that the more concentrated that the banking sector is, the lower that the risk tends to be. This result is in line with the empirical evidence that was provided by Beck et al. [56] and Caprio et al. [57]. Among the macroeconomic variables, GDP and UR show negative and positive relationships, respectively, with risk. These findings confirm the abundant support in the literature for the view that bank risk exhibits a clear cyclical behavior [58,59]. We finally reveal a positive relationship between inflation and bank risk, and a negative relationship between interest rates and bank risk.

4.2. Robustness Checks

To further confirm the aforementioned findings, we conduct a number of robustness checks (see Table 6).

Table 6. Robustness checks.

(A) Considering alternative proxies of profitability and risk		
	ROE	**NPLr**
HHIRD	0.132 ***	0.003
	(0.045)	(0.005)
z_1	1043.94 (20)	1059.06 (20)
m_1	−4.99	−6.18
m_2	−0.77	−1.07
Hansen	447.08 (444)	441.60 (444)
(B) Considering NonINT instead of HHIRD		
	ROA	**Z-Score**
NonINT	0.014 ***	−0.001
	(0.003)	(0.004)
z_1	1332.29 (20)	599.33 (20)
m_1	−4.05	−6.13
m_2	−0.83	−1.05
Hansen	444.85 (444)	443.17 (444)
(C) Including year and country dummies instead of macroeconomic variables		
	ROA	**Z-score**
HHIRD	0.008 ***	−0.004
	(0.003)	(0.006)
z_1	510.98 (38)	1424.22 (38)
m_1	−4.21	−5.63
m_2	−0.66	−1.01
Hansen	435.10 (441)	438.82 (441)

Notes: This table presents the results after changing some of the variables included in the baseline regression. The models in section (A) consider alternative proxies of bank profitability and risk. The models in section (B) consider NonINT in place of Herfindahl–Hirschman index (HHIRD). The models in section (C) consider year and country dummies in place of the macroeconomic variables. We also include control variables (not reported) in all of the regressions. With the exception of industry concentration, macroeconomic variables, year, country and bank type dummies, all of the independent variables in our models are considered endogenous. We report heteroskedasticity-consistent asymptotic standard errors in parentheses, and significance levels are indicated as follows: *** = significant at the 1% level; ** = significant at the 5% level; and, * = significant at the 10% level. z_1 is a Wald test of the joint significance of the reported coefficients, asymptotically distributed as χ^2 under the null of no significance, with degrees of freedom in parentheses. m_i is a serial correlation test of order *i* using residuals in first differences, asymptotically distributed as $N(0,1)$ under the null of no serial correlation. Hansen is a test of over-identifying restrictions, asymptotically distributed as χ^2 under the null of no correlation between the instruments and the error term, with degrees of freedom in parentheses.

First, we re-estimate our baseline equation when considering the return on equity (ROE) as a proxy for profitability and the NPLr as a proxy for risk. Second, we use noninterest income as a percentage of total operating revenue (NonINT), instead of HHIRD as our indicator of revenue diversification. Similarly, we re-estimate the baseline equation to consider year and country dummies in the place of macroeconomic variables. The result does not differ from that obtained previously.

5. Summary and Conclusions

This paper empirically analyzes the effect of revenue diversification on bank profitability and bank risk in the Eurozone. We consider a revenue diversification index that reflects the balance of different types of income, such as interest, net commissions, trading, and other operating income. Our sample comprises an unbalanced panel data set of 4268 observations from 2000 to 2012, allowing for us to consider the impact of the recent financial and economic crisis on the Eurozone banking system. Because previous studies have suggested that our regression could be affected by endogeneity, unobserved heterogeneity and the persistence of dependent variables, we employ a dynamic panel data model with the system-GMM estimator.

We provide evidence that higher income diversification favors bank profitability. This result is consistent with the theory that suggests that the potential benefits of diversification arising from economies of scope are larger than its costs (agency problems, regulatory costs, or inefficient internal resource allocation). However, our study does not find a significant relationship between revenue diversification and bank risk, even when considering a crisis period. Therefore, we cannot confirm theoretical predictions about the effect of diversification on either a reduction (according to the portfolio theory) or an increase (due to new forms of risk or lower capital ratios) of bank risk.

Our findings have important policy implications and could have broader significance for supervisors concerned about benchmarking and validation issues that are related to banking regulation. The recent recommendations of policy makers to isolate certain types of activities that would imply restrictions on the universal banking model could reduce the profitability of banks. This outcome, together with an environment of low interest rates, can damage the resilience of the banking system, and thus affect the sustainability of the uneven economic recovery in the Eurozone.

Appendix A

Table A1. Description of regulatory and supervisory control variables.

Variable	Description
Capital Stringency Index	The capital stringency index is built by adding two measures of capital stringency: overall and initial capital stringency. Overall capital stringency indicates whether risk elements and value losses are considered while calculating the regulatory capital. It is based on the following questions: (i) Is the minimum capital-asset ratio requirement risk weighted in line with the Basel guidelines? (ii) Does the minimum ratio vary as a function of credit risk? (iii) Does the minimum ratio vary as a function of market risk? (iv) Are market values of loan losses not realized in accounting books deducted from capital? (v) Are unrealized losses in securities portfolios deducted from capital? (vi) Are unrealized foreign exchange losses deducted from capital? (vii) What fraction of revaluation gains is allowed as part of capital? Initial capital stringency refers to whether certain funds may be used to initially capitalize a bank and whether they are officially verified. It is based on the following questions: (viii) Are the sources of funds to be used as capital verified by the regulatory or supervisory authorities? (ix) Can the initial disbursement or subsequent injections of capital be performed with assets other than cash or government securities? (x) Can the initial disbursement of capital be performed with borrowed funds? We assign a value of 1 if the answer to questions (i), (ii), (iii), (iv), (v), (vi) and (viii) is yes and 0 otherwise, while the opposite holds in the case of questions (ix) and (x). In addition, we assign a value of 1 if the fraction of revaluation gains that is allowed to count as regulatory capital (question (vii)) is less than 0.75. Otherwise, we assign a value of 0. By adding all these values together, we create the variable capital stringency index, which ranges in value from 0 to 10, with higher values indicating greater stringency.
Supervisory Power Index	This index is a measure of the power of supervisory agencies, indicating the extent to which these authorities can take specific actions against bank management and directors, shareholders and bank auditors. This variable is determined by adding 1 if the answer is yes and 0 otherwise for each of the following 14 questions: (i) Does the supervisory agency have the right to meet with external auditors to discuss their report without the approval of the bank? (ii) Are auditors legally required to communicate directly to the supervisory agency any presumed involvement of bank directors or senior managers in illicit activities, fraud, or insider abuse? (iii) Can supervisors take legal action against external auditors for negligence? (iv) Can the supervisory authorities force a bank to change its internal organizational structure? (v) Does the institution disclose off-balance-sheet items to supervisors? (vi) Can the supervisory agency order the bank's directors or management to constitute provisions to cover actual or potential losses? (vii) Can the supervisory agency suspend directors' decisions to distribute dividends? (viii) Can the supervisory agency suspend directors' decisions to distribute bonuses? (ix) Can the supervisory agency suspend directors' decisions to distribute management fees? (x) Can the supervisory agency supersede bank shareholder rights and declare the bank insolvent? (xi) Does banking law allow a supervisory agency or any other government agency (other than a court) to suspend some or all ownership rights at a problem bank? (xii) Regarding bank restructuring and reorganization, can the supervisory agency or any other government agency (other than a court) supersede shareholder rights? (xiii) Regarding bank restructuring and reorganization, can the supervisory agency or any other government agency (other than a court) remove and replace management? (xiv) Regarding bank restructuring and reorganization, can the supervisory agency or any other government agency (other than a court) remove and replace directors? This index takes values between 0 and 14, with higher values indicating more supervisory power.
Private Monitoring Index	To control for market discipline, we use the private monitoring index. This index examines the degree to which regulatory and supervisory policies encourage the private monitoring of banks, and it is composed of information on (i) whether bank directors and officials are legally liable for the accuracy of information disclosed to the public; (ii) whether banks must publish consolidated accounts; (iii) whether banks must be audited by certified international auditors; (iv) whether 100 percent of the largest 10 banks are rated by international rating agencies; (v) whether off-balance sheet items are disclosed to the public; (vi) whether banks must disclose their risk management procedures to the public; (vii) whether accrued, though unpaid, interest/principal enters the income statement while the loan is still non-performing; (viii) whether subordinated debt is allowable as part of capital and (ix) whether there is no explicit deposit insurance system and no insurance was paid the last time a bank failed. The maximum value of the private monitoring index is 12, and the minimum value is 0, where larger values indicate greater regulatory empowerment of the monitoring of banks by private investors.
Activity Restrictions Index	The score for this variable is determined based on the level of regulatory restrictiveness for bank participation in: (i) securities activities; (ii) insurance activities; (iii) real estate activities and (iv) bank ownership of non-financial firms. These activities can be unrestricted, permitted, restricted, or prohibited and can receive values of 1, 2, 3, or 4, respectively. We create an overall index by calculating the summation value of the four categories. This index ranges from 4 to 16, with higher values indicating higher restrictions.

References

1. DeYoung, R.; Roland, K.P. Product mix and earnings volatility at commercial banks: Evidence from a degree of total leverage model. *J. Financ. Intermed.* **2001**, *10*, 54–84. [CrossRef]

2. Johnson, R.D.; Meinster, D.R. Bank holding companies: Diversification opportunities in nonbank activities. *East. Econ. J.* **1974**, *1*, 316–323.

3. Heggestad, A. Riskiness of investments in nonbank activities by bank holding companies. *J. Econ. Bus.* **1975**, *27*, 219–223.

4. Wall, L.D.; Eisenbeis, R.A. Risk considerations in deregulating bank activities. *Fed. Reserv. Bank Atl. Econ. Rev.* **1984**, *69*, 6–19.

5. Litan, R.E. Evaluating and controlling the risks of financial product deregulation. *Yale J. Regul.* **1985**, *3*, 1–52.

6. Wilmarth, A.E. The Dodd-Frank Act: A Flawed and Inadequate Response to the Too-Big-to-Fail Problem. *Or. Law Rev.* **2011**, *89*, 951–1058.

7. Meslier, C.; Tacneng, R.; Tarazi, A. Is bank income diversification beneficial? Evidence from an emerging economy. *J. Int. Financ. Mark. Inst. Money* **2014**, *31*, 97–126. [CrossRef]

8. Altunbas, Y.; Manganelli, S.; Marques-Ibanez, D. *Bank Risk during the Financial Crisis: Do Business Models Matter?* Working Paper 1394; European Central Bank: Frankfurt, Germany, 2011.

9. Liikanen, E. The High-Level Expert Group on Reforming the Structure of the EU Banking Sector: Final Report. 2012. Available online: http://ec.europa.eu/internal_market/bank/docs/high-level_expert_group/report_en.pdf (accessed on 1 April 2014).

10. Gambacorta, L.; Van Rixtel, A. Structural bank regulation initiatives: Approaches and implications. *BANCARIA Bancaria Editrice* **2013**, *6*, 14–27.

11. Baele, L.; De Jonghe, O.; Vander Vennet, R. Does the stock market value bank diversification? *J. Bank. Financ.* **2007**, *31*, 1999–2023. [CrossRef]

12. Mercieca, S.; Schaeck, K.; Wolfe, S. Small European banks: Benefits from diversification? *J. Bank. Financ.* **2007**, *31*, 1975–1998. [CrossRef]

13. Lepetit, L.; Nys, E.; Rous, P.; Tarazi, A. Bank income structure and risk: An empirical analysis of European banks. *J. Bank. Financ.* **2008**, *32*, 1452–1467. [CrossRef]

14. Arellano, M.; Bover, O. Another look at the instrumental-variable estimation of error-components models. *J. Econ.* **1995**, *68*, 29–52. [CrossRef]

15. Blundell, R.W.; Bond, S.R. Initial conditions and moment restrictions in dynamic panel data models. *J. Econ.* **1998**, *87*, 115–143. [CrossRef]

16. Bun, M.J.G.; Windmeijer, F. The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econ. J.* **2010**, *13*, 95–126. [CrossRef]

17. Diamond, D.W. Monitoring and reputation: The choice between bank loans and directly placed debt. *J. Polit. Econ.* **1991**, *99*, 689–721. [CrossRef]

18. Stein, J.C. Information production and capital allocation: Decentralized versus hierarchical firms. *J. Financ.* **2002**, *57*, 1891–1921. [CrossRef]

19. Laeven, L.; Levine, R. Is there a diversification discount in financial conglomerates? *J. Financ. Econ.* **2007**, *85*, 331–367. [CrossRef]

20. Jensen, M.C.; Meckling, W.H. Theory of the firm: Managerial behavior, agency costs and ownership structure. *J. Financ. Econ.* **1976**, *3*, 305–360. [CrossRef]

21. Rajan, R.; Servaes, H.; Zingales, L. The cost of diversity: The diversification discount and inefficient investment. *J. Financ.* **2000**, *55*, 35–80. [CrossRef]

22. Moudud-Ul-Huq, S.; Nadeem Ashraf, B.; Das Gupta, A.; Zheng, C. Does bank diversification heterogeneously affect performance and risk-taking in ASEAN emerging economies? *Res. Int. Bus. Financ.* **2018**. [CrossRef]

23. Brighi, P.; Venturelli, V. How do income diversification, firm size and capital ratio affect performance? Evidence for bank holding companies. *Appl. Financ. Econ.* **2014**, *24*, 1375–1392. [CrossRef]

24. Chiorazzo, V.; Milani, C.; Salvini, F. Income diversification and bank performance: Evidence from Italian banks. *J. Financ. Serv. Res.* **2008**, *33*, 181–203. [CrossRef]

25. Köhler, M. Does non-interest income make banks more risky? Retail-versus investment-oriented banks. *Rev. Financ. Econ.* **2014**, *23*, 182–193. [CrossRef]

26. Köhler, M. Which banks are more risky? The impact of business model on bank stability. *J. Financ. Stab.* **2015**, *16*, 195–212. [CrossRef]

27. Brighi, P.; Venturelli, V. How functional and geographic diversification affect bank profitability during the crisis. *Financ. Res. Lett.* **2016**, *16*, 1–10. [CrossRef]

28. Elsas, R.; Hackethal, A.; Holzhäuser, M. The anatomy of bank diversification. *J. Bank. Financ.* **2010**, *34*, 1274–1287. [CrossRef]

29. Lee, C.; Hsieh, M.F.; Yang, S.J. The relationship between revenue diversification and bank performance: Do financial structures and financial reforms matter? *Jpn. World Econ.* **2014**, *29*, 18–35. [CrossRef]

30. Sanya, S.; Wolfe, S. Can Banks in Emerging Economies Benefit from Revenue Diversification? *J. Financ. Serv. Res.* **2011**, *40*, 79–101. [CrossRef]

31. Nguyen, M.; Perera, S.; Skully, M. Bank market power, ownership, regional presence and revenue diversification: Evidence from Africa. *Emerg. Mark. Rev.* **2016**, *27*, 36–62. [CrossRef]

32. Sissy, A.M.; Amidu, M.; Abor, J.Y. The effects of revenue diversification and cross border banking on risk and return of banks in Africa. *Res. Int. Bus. Financ.* **2017**, *40*, 1–18. [CrossRef]

33. De Jonghe, O. Back to the basics in banking? A micro-analysis of banking system stability. *J. Financ. Intermed.* **2010**, *19*, 387–417. [CrossRef]

34. DeYoung, R.; Rice, T. Noninterest income and financial performance at U.S. commercial banks. *Financ. Rev.* **2004**, *39*, 101–127. [CrossRef]

35. Stiroh, K.J.; Rumble, A. The dark side of diversification: The case of US financial holding companies. *J. Bank. Financ.* **2006**, *30*, 2131–2161. [CrossRef]

36. Saona, P. Intra and extra bank determinants of Latin American Banks profitability. *Int. Rev. Econ. Financ.* **2016**, *45*, 197–214. [CrossRef]

37. Demirgüç-Kunt, A.; Huizinga, H. Bank activity and funding strategies: The impact on risk and returns. *J. Financ. Econ.* **2010**, *98*, 626–650. [CrossRef]

38. Arellano, M.; Bond, S.R. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev. Econ. Stud.* **1991**, *58*, 277–297. [CrossRef]

39. Delis, M.D.; Staikouras, P.K. Supervisory effectiveness and bank risk. *Rev. Financ.* **2011**, *15*, 511–543. [CrossRef]

40. Trujillo-Ponce, A. What determines the profitability of banks? Evidence from Spain. *Account. Financ.* **2013**, *53*, 561–586. [CrossRef]

41. Windmeijer, F. A finite sample correction for the variance of linear efficient GMM estimators. *J. Econ.* **2005**, *126*, 25–51. [CrossRef]

42. Altunbas, Y.; Carbó, S.; Gardener, E.P.M.; Molyneux, P. Examining the relationships between capital, risk and efficiency in European banking. *Eur. Financ. Manag.* **2007**, *13*, 49–70. [CrossRef]

43. Baselga-Pascual, L.; Trujillo-Ponce, A.; Cardone-Riportella, C. Factors influencing bank risk in Europe: Evidence from the financial crisis. *N. Am. J. Econ. Financ.* **2015**, *34*, 138–166. [CrossRef]

44. Laeven, L.; Majnoni, G. Loan loss provisioning and economic slowdowns: Too much, too late? *J. Financ. Intermed.* **2003**, *12*, 178–197. [CrossRef]

45. Bikker, J.A.; Metzemakers, P. Bank provisioning behavior and procyclicality. *J. Financ. Mark. Inst. Money* **2005**, *15*, 141–157. [CrossRef]

46. Agoraki, M.-E.K.; Delis, M.D.; Pasiouras, F. Regulations, competition and bank risk-taking in transition countries. *J. Financ. Stab.* **2011**, *7*, 38–48. [CrossRef]

47. Louzis, D.; Vouldis, A.T.; Metaxas, V. Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios. *J. Bank. Financ.* **2012**, *36*, 1012–1027. [CrossRef]

48. Athanasoglou, P.P.; Brissimis, S.N.; Delis, M.D. Bank-specific, industry-specific and macroeconomic determinants of bank profitability. *J. Int. Financ. Mark. Inst. Money* **2008**, *18*, 121–136. [CrossRef]

49. Claeys, S.; Vander Vennet, R. Determinants of bank interest margins in Central and Eastern Europe: A comparison with the West. *Econ. Syst.* **2008**, *32*, 197–216. [CrossRef]

50. García-Herrero, A.; Gavilá, S.; Santabárbara, D. What explains the low profitability of Chinese banks? *J. Bank. Financ.* **2009**, *33*, 2080–2092. [CrossRef]

51. Barth, J.R.; Caprio, G., Jr.; Levine, R. *The Regulation and Supervision of Bank around the World: A New Database*; Policy Research Working Paper 2588; World Bank: Washington, DC, USA, 2001.

52. Barth, J.R.; Caprio, G., Jr.; Levine, R. Bank regulation and supervision in 180 countries from 1999 to 2011. *J. Financ. Econ. Policy* **2013**, *5*, 111–219. [CrossRef]

53. Nichols, D.; Wahlen, J.; Wieland, M. Publicly-traded versus privately-held: Implications for conditional conservatism in bank accounting. *Rev. Account. Stud.* **2009**, *14*, 88–122. [CrossRef]

54. Barth, J.R.; Caprio, G., Jr.; Levine, R. Bank regulation and supervision: What works best? *J. Financ. Intermed.* **2004**, *13*, 205–248. [CrossRef]

55. Festić, M.; Kavkler, A.; Repina, S. The macroeconomic sources of systemic risk in the banking sectors of five new EU member states. *J. Bank. Financ.* **2011**, *35*, 310–322. [CrossRef]

56. Beck, T.; Demirgüç-Kunt, A.; Levine, R. Bank concentration, competition, and crises: First results. *J. Bank. Financ.* **2006**, *30*, 1581–1603. [CrossRef]

57. Caprio, G., Jr.; D'Apice, V.; Ferri, G.; Puopolo, G.W. Macro-financial determinants of the great financial crisis: Implications for financial regulation. *J. Bank. Financ.* **2014**, *44*, 114–129. [CrossRef]

58. Marcucci, J.; Quagliariello, M. Is bank portfolio riskiness procyclical? Evidence from Italy using a vector autoregression. *J. Int. Financ. Mark. Inst. Money* **2008**, *18*, 46–63. [CrossRef]

59. Uhde, A.; Heimeshoff, U. Consolidation in banking and financial stability in Europe: Empirical evidence. *J. Bank. Financ.* **2009**, *33*, 1299–1311. [CrossRef]

Article

Is Liquidity Risk Priced? Theory and Evidence

Seok-Kyun Hur [1], Chune Young Chung [1],* and Chang Liu [2]

[1] School of Business Administration, College of Business and Economics, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Korea; shur@cau.ac.kr

[2] College of Business, Hawaii Pacific University, 900 Fort Street Mall, Honolulu, HI 96813, USA; cliu@hpu.edu

* Correspondence: bizfinance@cau.ac.kr; Tel.: +82-2-820-5544

Abstract: This study studies a recently proposed measure of liquidity premium (or discount). Specifically, the liquidity premium we utilize is defined as a function of a time discount factor, a relative risk aversion parameter, and the expected return and volatility of the asset, given the risk-free rate. Using U.S. stock market data, our empirical results confirm that the proposed liquidity premium measure is largely comparable to that commonly used in existing studies. Our results also imply that a risk factor based on the liquidity premium measure not only explains cross-sectional stock returns, but also time-series excess returns on portfolios sorted on the commonly used liquidity measure. In addition, our study suggests that better understanding the liquidity risk leads to sustainable trading for investors.

Keywords: liquidity premium; uncertainty termination; investment horizon; Amihud's illiquidity ratio; factor models

1. Introduction

A large body of literature attempts to measure various types of risks in financial assets. For example, Allen et al. [1] develop a new measure of risk based on the application of regular vine copulas and apply it to the assessment of composite financial risk. In addition, Yan et al. [2] propose the new empirical method by combining generalized autoregressive score functions and a copula model with high-frequency data to model the conditional time-varying joint distribution of the government bond yields.

An application of regular vine copulas, which are a novel and recently developed statistical and mathematical tool which can be applied in the assessment of composite financial risk. However, liquidity (or illiquidity) risk is not a readily measurable characteristic of financial assets, yet understanding the implications of liquidity on investments results is crucial for the sustainability of investors who face increasingly more investment alternatives that are illiquid (e.g., hedge funds, private equity, and real estate). Existing studies in the literature generally employ an asset's order flow, transaction volume, and the corresponding price impact to measure illiquidity [3–6]. This arises from the conventional wisdom that a transaction's impact on the asset price captures the liquidity premium (illiquidity discount) that a buyer (seller) is willing to pay (offer) to fulfill an order.

In this paper, we study a novel liquidity premium measure based on the equilibrium derived from a dynamic model in Hur and Chung [7], and apply the measure to the US stock market. Their model implies that the liquidity premium of an asset is a function of a time discount factor, the relative risk aversion of the investor, and the expected return and volatility of the asset, given the risk-free rate. Based on reasonable specifications of the parameters, our empirical findings show that the proposed liquidity premium is highly related to Amihud's [8] illiquidity ratio, a measure commonly used in existing studies. More specifically, we find that the cross-sectional variation of the new liquidity

premium is significantly explained by Amihud's illiquidity ratio, which supports the validity of the measure.

The model implies that a highly liquid asset should command a price premium while illiquid assets must be offered at price discounts. To examine this implication, we perform two sets of tests based on U.S. data: the first is a test of the cross-sectional relationship between the new liquidity premium and stock returns; and the second is a test of the time-series relationship between a risk factor based on the new liquidity premium and the expected excess returns on portfolios sorted on the Amihud's illiquidity ratio. As predicted, we find a negative relationship in both settings. Overall, the empirical findings corroborate that the liquidity risk factor based on our measure of the liquidity premium is priced in stock/portfolio returns, which further validates the robustness of the proposed liquidity measure.

2. Liquidity Premium

In this section, as in Hur and Chung [7] (who attempted to apply the model's implications to the Korean stock market, and their empirical results are qualitatively similar to ours in this study, hence, we believe that our new measure has compatibility in the global markets.), we derive a closed-form representation of liquidity premium based on a continuous time model.

2.1. Model Setup

We begin by defining the following:

Definition 1. *A random investment horizon is the first time a pre-determined investment goal is attained.*

Definition 2. *An investment goal is the targeted rate of return.*

Definition 3. *An asset's liquidity premium (discount) is the maximum willingness-to-pay that makes an investor indifferent between two consumption options—cashing out and consuming all positions now in the absence of any future liquidation shock, or waiting until the investment horizon terminates before consuming all wealth—while maintaining the same investment strategy with a random horizon.*

The first consumption option in Definition 3 measures utility from holding a financial asset when its liquidity is perfect to the level of money. The second option measures utility when the asset's liquidity is limited. Therefore, the certainty equivalent variation that make these two utilities equal represents the asset's liquidity premium. For brevity and consistency, we focus on locked-in (or locked-out) strategies as these are more convenient when assessing the contributions of holding a designated asset. In contrast, it is challenging to disentangle the contributions of holding an asset from others under a cross-sectional diversification strategy.

We assume that stock price X_t follows a log-normal Brownian motion with drift, as follows:

$$dX_t = \mu X_t dt + \sigma X_t dB_t.$$

A solution to this stochastic differential equation is easily obtained by applying Ito's lemma:

$$X_t = X_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B_t\right]$$

This solution for X_t has a drift of $\left(\mu - \frac{1}{2}\sigma^2\right)$, implying that it would grow continuously at the rate of $\left(\mu - \frac{1}{2}\sigma^2\right)$. As a benchmark, we define $\widetilde{X}_t(b)$ to grow at the rate of the risk-free rate r. Note that it is always greater (smaller) than $X_0 \exp[rt]$ by $\exp[b]$ for $b > 0$ ($b < 0$), as follows:

$$\widetilde{X}_t(b) \equiv X_0 \exp[rt + b].$$

More specifically, consider an investor who purchases one share of a stock at time 0. Their investment goal is to outperform a risk-free asset by $\exp[b]$ times. The investor would like to know when this goal will be attained. Such an investment strategy is known by different names, such as buy-and-hold, stop-gains, or locked-in. As these names suggest, the essence of the strategy is to survive in order to achieve the investment goal.

We define time T_b to a level of $b \in R$ and $b > 0$ as follows:

$$T_b(\omega) = inf\{t \geq 0; X_t(\omega) = \widetilde{X}_t(b)\}$$

$$T_b(\omega) = inf\{t \geq 0; X_t(\omega) = \widetilde{X}_t(b)\} = inf\{t \geq 0; B_t = \frac{b}{\sigma} - \frac{1}{\sigma}\left(\mu - \frac{1}{2}\sigma^2 - r\right)t\},$$

where a new Brownian motion with drift, $\widetilde{B}_t$, has a drift of $\frac{1}{\sigma}\left(\mu - \frac{1}{2}\sigma^2 - r\right)$ as:

$$\widetilde{B}_t \equiv B_t + \frac{1}{\sigma}\left(\mu - \frac{1}{2}\sigma^2 - r\right)t.$$

From Karatzas and Shreve [9], we have:

$$P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))}[T_b \in dt] = \frac{b}{\sqrt{2\pi\sigma^2 t^3}}\exp\left[-\frac{\left(b - \left(\mu - \frac{1}{2}\sigma^2 - r\right)t\right)^2}{2\sigma^2 t}\right]dt, \ t > 0$$

$$P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))}[T_b \leq t] = \int_0^t \exp\left[\frac{b}{\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right) - \frac{1}{2\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right)^2 s\right]P[T_b \in dt]$$

$$P[T_b \in dt] = \frac{b}{\sqrt{2\pi\sigma^2 t^3}}\exp\left[-\frac{b^2}{2\sigma^2 t}\right]dt, \ t > 0.$$

Applying $Ee^{-\alpha T_b} = e^{-\frac{b}{\sigma}\sqrt{2\alpha}}$, we can calculate the value of $\lim_{t\to\infty}P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))}[T_b \leq t]$:

$$P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))}[T_b < \infty] = e^{\frac{b}{\sigma^2}(\mu - \frac{1}{2}\sigma^2 - r)}E\left[\exp\left(-\frac{1}{2\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right)^2 T_b\right)\right]$$
$$= \exp\left[\frac{b}{\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right) - \frac{b}{\sigma^2}\left|\mu - \frac{1}{2}\sigma^2 - r\right|\right].$$

Assumption 1. $\mu - \frac{1}{2}\sigma^2 > r$

Proposition 1. In an infinite investment horizon, an investor is certain to receive a stochastic cash flow, the present value of which exceeds the present price of an asset by simply holding the asset until time T_b.

Proof. By construction, at $X_t(\omega) = \widetilde{X}_t(b)$ the investor can sell the asset at a price that has continuously grown faster than the speed of r. In addition, T_b is known to be reached in finite time with a probability of 1; that is, $P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))}[T_b < \infty] = 1$, based on Assumption 1.

As b is arbitrary, $P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))}[T_b < \infty] = 1$ implies that the investor will achieve any targeted return in the infinite investment horizon, although the timing is still uncertain. Investors differ in their motives for investment, life cycles, and economic abilities, and these sources of heterogeneity compel them to invest in different time horizons. Thus, $P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))}[T_b < \infty] = 1$ should not be regarded as a sign of arbitrage opportunities.

Proposition 2. The distribution of T_b exhibits first-order stochastic dominance (FOSD) with respect to (b, μ, r), as follows:

$$\frac{\partial}{\partial b} P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))} [T_b \leq t] \leq 0 \tag{1}$$

$$\frac{\partial}{\partial \mu} P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))} [T_b \leq t] \geq 0 \tag{2}$$

$$\frac{\partial}{\partial r} P^{(\frac{1}{\sigma}(\mu - \frac{1}{2}\sigma^2 - r))} [T_b \leq t] \leq 0. \tag{3}$$

Proof.

(1) As $b < b'$ and $\widetilde{X}_t(b) < \widetilde{X}_t(b')$, X_t should touch $\widetilde{X}_t(b)$ before it touches $\widetilde{X}_t(b')$. Hence, the distribution of $T_{b'}$ exhibits FOSD over the distribution of T_b.

(2) As $\frac{\partial}{\partial \mu} X_t > 0$, the distribution of T_b shifts leftward as μ increases.

(3) As $\frac{\partial}{\partial r} \widetilde{X}_t(b) > 0$, the distribution of T_b shifts rightward as r rises. □

In contrast, the effect of an increase in σ^2 is ambiguous in that it raises the volatility of X_t, while lowering the drift. The lower drift implies that the distribution of T_b before the change in σ^2 exhibits FOSD over the distribution after the change. However, the increased volatility implies that the distribution of T_b before the change in σ^2 exhibits second-order stochastic dominance (SOSD) over the distribution after the change. Summarizing these two effects, we cannot characterize the shift in the distribution of T_b as either FOSD or SOSD.

The validity of the buy-and-hold (locked-in) strategy is confirmed as a certain investment horizon is guaranteed. The buy-and-hold is a long-term strategy that becomes more favorable as the investment horizon increases. Thus, it appears rational for an investor to constrain risky asset holding to a certain level. The investor can extend the investment horizon effectively by increasing the probability of the investment position, maintaining a positive balance until it reaches the targeted return.

2.2. Explicit Representation of Liquidity Premium

In this section, we derive a probability distribution function for the first passage of time after which any locked-in strategy will attain a given investment goal in continuous time. From Karatzas and Shreve [9], we confirm that any locked-in strategy targeting a positive excess gain will be attained eventually with a probability of 1.

Uncertainty of termination differs from the early resolution of uncertainty. The latter is related to the timing of future uncertainty being revealed. Realization of the future affects an agent's economic interests. Agents can better prepare if they are informed earlier of future uncertainty, in which case their ex ante expected utility may improve.

In contrast, this first passage of time is not related to the timing of uncertainty resolution or to the agent's preference. It only indicates that the effective investment horizon can be random, based on exogenous factors, such as mortality, or on the choice of a locked-in or locked-out investment strategy.

Assumption 2. $b > 0$

Assumption 3. All investors hold a constant relative risk aversion utility $U(W) = \frac{W^{1-\gamma}}{1-\gamma}$, $\gamma > 0$

The previous subsection shows that a locked-in strategy of T_b will be achieved with a probability of 1. In other words, an investor with W_0 will eventually receive $W_0 \exp[b]$ by adopting a locked-in strategy. Hence, in principle, the present value of his/her wealth is equal to the value of a bond that pays $W_0 \exp[rt + b]$ at a randomly chosen maturity.

In a financial market without any frictions and liquidation risks, the present value of the agent's wealth will be equal to $W_0 \exp[b]$. To avoid arbitrage opportunities or violating the law

of one price, it should be decreased to at least W_0, which implies that frictions and liquidation risks are present in the current market environment. Hence, we measure the expected utilities based on these two cases.

However, we minimize the investment strategy's liquidity by requiring that an investor liquidate his/her position at a stopping time T_b and consume all the proceeds. The expected utility of this case is compared with that of the previous two, and their differences are measured as the lower and upper bounds respectively, of the liquidity premium (or discount).

First, in the case of perfect liquidity, the current wealth of W_0 increases to $W_0 \exp[rt + b]$, and the corresponding utility is:

$$V^P(W_0) \equiv \frac{W_0{}^{1-\gamma} \exp[b(1-\gamma)]}{1-\gamma}.$$

$$V^L(W_0) \equiv \frac{W_0{}^{1-\gamma}}{1-\gamma} < V^P(W_0).$$

Third, in the case of no liquidity, the utility of the investor with the locked-in strategy of T_b is calculated as follows:

$$V(W_0, b) \equiv \frac{W_0{}^{1-\gamma}}{1-\gamma} \chi_0$$

$$\chi_0 \equiv \int_0^\infty \frac{b}{\sqrt{2\pi\sigma^2 t^3}} \exp\left[-\frac{\left(b-(\mu-\frac{1}{2}\sigma^2-r)t\right)^2}{2\sigma^2 t} + (1-\gamma)(rt+b) - \rho t\right] dt$$

$$= \int_0^\infty \frac{b}{\sqrt{2\pi\sigma^2 t^3}} \exp\left[-\frac{\left(b-(\mu-\frac{1}{2}\sigma^2-r)t\right)^2}{2\sigma^2 t}\right] \exp[(1-\gamma)(rt+b) - \rho t] dt$$

$$= e^{(1-\gamma)b} E_0[\exp[((1-\gamma)r - \rho)t]].$$

Considering the properties of the first Brownian motion with drift of μ, we know that $E^{(\mu)} e^{-\alpha T_b} = \exp[\mu b - |b|\sqrt{\mu^2 + 2\alpha}, \alpha > 0]$. Hence:

$$\chi_0 = \exp\left[\frac{b}{\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right) - \frac{|b|}{\sigma}\sqrt{\frac{1}{\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right)^2 + 2(\rho - (1-\gamma)r)} + (1-\gamma)b\right].$$

The upper and lower bounds of the financial asset's liquidity premium are calculated by equating the following inequalities ($0 \le \varphi_L < \varphi_L$ if $b > 0$, and $\varphi_L > \varphi_L \ge 0$ if $b < 0$.). For $0 < \gamma < 1$:

$$V^P(W_0(1 - \varphi_P)) = V(W_0, b)$$

$$V^L(W_0(1 - \varphi_L)) = V(W_0, b)$$

$$\varphi_P = 1 - \chi_0^{\frac{1}{(1-\gamma)}} \exp[-b], \quad \varphi_L = 1 - \chi_0^{\frac{1}{(1-\gamma)}}.$$

As $\varphi_P > \varphi_L$ ($b > 0$), φ_L and φ_P are regarded as the lower and upper bounds respectively, of the liquidity premium. They are functions of b, μ, σ^2, ρ, γ, and r

For $\gamma > 1$:

$$V^P(W_0(1 + \phi_P)) = V(W_0, b)$$

$$V^L(W_0(1 + \phi_L)) = V(W_0, b)$$

$$\phi_P = \chi_0^{\frac{1}{(1-\gamma)}} \exp[-b] - 1, \quad \phi_L = \chi_0^{\frac{1}{(1-\gamma)}} - 1.$$

As $\phi_P < \phi_L$ ($b > 0$), ϕ_L and ϕ_P are perceived to be the upper and lower bounds, respectively, of the liquidity discount. They are functions of b, μ, σ^2, ρ, γ, and r.

Proposition 3. The liquidity premium measures the increase in φ_L and φ_P with respect to (b, μ, ρ, γ), but the decrease with respect to (σ^2, r).

Proposition 3 implies that an asset's liquidity premium is inversely related to volatility and the risk-free rate, but is positively linked to the expected return and relative risk-aversion parameter. This result is consistent with the mean–variance trade-off.

Proposition 4. The liquidity discount measures the decrease in ϕ_L and ϕ_P with respect to (b, μ, ρ, γ), but the increase with respect to (σ^2, r).

Proposition 4 implies that an asset's liquidity discount is positively related to volatility and the risk-free rate, but is inversely linked to the expected return and relative risk-aversion parameter. This result is also consistent with the mean–variance trade-off.

The risk premium and discount measures tend to have different values, depending on b. In order to make them independent of b, we devise the following variants by applying L'Hopital's rule in order to obtain explicit representations of the liquidity premium and discount (note that liquidity premium and liquidity discount have the same magnitude, but opposite signs):

$$\varphi_L^* \equiv \lim_{b \to 0} \frac{\varphi_L}{b} = \frac{1}{(1-\gamma)\sigma} \sqrt{\frac{1}{\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right)^2 + 2(\rho - (1-\gamma)r)} - \frac{1}{(1-\gamma)\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right) - 1$$

$$\phi_L^* \equiv \lim_{b \to 0} \frac{\phi_L}{b} = \frac{1}{(1-\gamma)\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right) - \frac{1}{(1-\gamma)\sigma} \sqrt{\frac{1}{\sigma^2}\left(\mu - \frac{1}{2}\sigma^2 - r\right)^2 + 2(\rho - (1-\gamma)r)} + 1$$

$$\varphi_P^* \equiv \lim_{b \to 0} \frac{\varphi_P}{b} = \varphi_L^* + 1, \phi_P^* \equiv \lim_{b \to 0} \frac{\phi_P}{b} = \phi_L^* + 1.$$

Note that the liquidity premium is simply defined as a function of the mean and variance of an asset's return, together with the risk-free rate, the time discount factor, and the relative risk aversion. In addition, the premium does not depend on the transactional characteristics of the market.

3. Empirical Results

In this section, we empirically estimate the liquidity premiums of stocks. Further, we examine whether the liquidity premium explains the cross-sectional variation in stock returns and whether the expected excess returns on portfolios sorted on the liquidity premium are at least partially represented by the factor loading of the liquidity premium over time. The model implies that a stock with high (low) liquidity should command a high (low) liquidity premium; thus, we not only expect a negative relationship between the liquidity premiums and cross-sectional stock returns, but also a negative relation between the factor based on the liquidity premium and expected excess returns on portfolios sorted on the liquidity premium in the time series.

First, we estimate the liquidity premiums for the sample of all stocks traded on the NYSE, AMEX, and NASDAQ in the U.S. stock market during the period 1980 to 2014. We collect data for daily stock returns from the Center for Research in Security Prices (CRSP), and compute the monthly means and variances of the stock returns. We consider the daily return of the one-month treasury bill rate as the daily risk-free rate. In addition, we set the time discount factor and the relative risk aversion to 0.99 and 0.5 (1.5), respectively, consistent with previous studies on asset pricing (e.g., [10–12]) (the empirical results are not sensitive to the choices of the time discount factor and relative risk aversion, as suggested in the literature. Though unreported for brevity, the results are available upon request). This allows us to calculate the daily liquidity premiums based on the model solution above when the relative risk aversion either falls between 0 and 1 or is greater than 1. Further, we compute the monthly liquidity premium (LIQ) using relative risk aversions of 0.5 and 1.5 as the average daily liquidity premium over a month for each stock, and convert it to the logarithmic form in the empirical analysis to alleviate the outlier effect.

For comparison, we also measure individual stocks' degrees of illiquidity on each day using Amihud's illiquidity ratio [8], which is a widely used measure of a stock's liquidity in the literature.

Amihud's illiquidity ratio is defined as the ratio of the absolute stock return to the dollar value of the trading volume. It captures the price impact of trading and is widely considered a good proxy for the illiquidity of stocks. Amihud [8], Acharya and Pedersen [13], Chordia et al. [14], and Brennan et al. [15] show that the ratio is significant in explaining a cross-section of stock returns. Using daily stock returns and trading volumes obtained from the CRSP, we measure the monthly average of Amihud's daily illiquidity ratio (unlike our model-based measure, Amihud's ratio does not require to consider a specific degree of investor's risk aversion). Further, we calculate the logarithm of this value as follows:

$$\log(Amihud_i) = \log\left(\frac{\sum_{d=1}^{t} \frac{|r_{i,d}|}{Vol_{i,d}}}{D_i}\right),$$

where $r_{i,d}$ is the return of stock i on day d, $Vol_{i,d}$ is the trading volume of stock i on day d (in USD), and D_i is the number of days in the month. A high value of the ratio indicates the stock is less liquid over the month.

4. Results

We report in Table 1 the descriptive statistics for the time-series average of the monthly cross-sectional liquidity premium and Amihud's illiquidity ratio for the period 1980 to 2014. We find that both the premium and illiquidity ratio vary significantly across firms. In addition, the variation of the liquidity premium is consistently evident, regardless of the parameter values of the relative risk aversion. This suggests that the newly proposed measure can appropriately explain the cross-sectional difference in stocks' liquidity.

Table 1. Descriptive statistics.

Variable	Mean	Std. Dev.	Min.	Median	Max.
Liquidity premium ($\gamma = 0.5$; $\rho = 0.99$)	4.76344	0.79846	−2.43841	4.7633	37.14187
Liquidity premium ($\gamma = 1.5$; $\rho = 0.99$)	4.78653	0.78017	0.73592	4.7803	37.14187
Amihud's illiquidity	−2.41855	3.37081	−36.78423	−2.4165	14.59196

Table 2 contains the correlations between the monthly liquidity premium and Amihud's illiquidity measure over the sample period. The results show a significantly negative relationship between the two measures, which is evidence that a higher liquidity premium is required for a more liquid stock. This further confirms that the proposed measure is a viable alternative as a measure of the degree of stock liquidity.

Table 2. Pearson correlation coefficients.

($\gamma = 0.5$; $\rho = 0.99$)	Liquidity Premium	Amihud's Illiquidity
Liquidity premium	1.000	
Amihud's illiquidity	−0.422 (0.000) ***	1.000
($\gamma = 1.5$; $\rho = 0.99$)		
Liquidity premium	1.000	
Amihud's illiquidity	−0.420 (0.000) ***	1.000

Notes: *p*-values are provided in parentheses. *** indicates statistical significance at the 1% level.

Next, we run simple cross-sectional regressions to examine the extent to which a stock's illiquidity (or liquidity) affects its liquidity premium. We use Fama and Macbeth's [16] approach to estimate the model, and compute the coefficients as the time-series averages from the monthly cross-sectional regressions. Table 3 reports the coefficient estimates of regressions of the liquidity premium on Amihud's illiquidity for different values of relative risk aversion. The liquidity premium is negatively

and significantly associated with Amihud's illiquidity ratio, suggesting that the liquidity premium reflects the stock's liquidity. Overall, our empirical results validate our measure of liquidity.

Table 3. Effect of stock illiquidity on liquidity premium.

	Intercept		Amihud's Illiquidity		Adj. R^2
	Coefficient	***t*-Stat**	**Coefficient**	***t*-Stat**	
($\gamma = 0.5$; $\rho = 0.99$)	4.5818 ***	206.52	−0.0951 ***	−29.47	0.1682
($\gamma = 1.5$; $\rho = 0.99$)	4.6093 ***	214.03	−0.0925 ***	−29.66	0.1671

Notes: The *t*-statistics are adjusted for Newey–West autocorrelations with three lags and are reported in parentheses. *** indicates statistical significance at the 1% level.

Following Amihud [5], we consider a monthly cross-sectional regression model that relates the liquidity measure to stock returns. In particular, we estimate the model following the Fama and Macbeth [13] method. In each month of year t, stock returns are regressed cross-sectionally on the liquidity measures and on the stock characteristics obtained at the end of year $t − 1$. The liquidity measures, including the LIQ and Amihud's illiquidity, are computed as monthly averages over the year (LIQ in Tables 4 and 5 is estimated based on $\gamma = 0.5$, but the results are qualitatively similar to those based on $\gamma = 1.5$. The results with $\gamma = 1.5$ are available upon request). Following Amihud [5], we also control for various stock characteristics. BETA is computed using the Scholes and Williams [17] method. R100 is the buy-and-hold return over the last 100 days of the year, and R100YR is the buy-and-hold return from the beginning of the year to 100 days before its end. SIZE is the logarithm of the market capitalization at year-end. BM is the book-to-market ratio of equity, computed as the book value for the fiscal year ended before the most recent June 30 divided by the market capitalization at year-end. SDRET is the standard deviation of daily returns during the year. DIVYLD is the dividend yield, computed as cash dividends for the fiscal year ended before the most recent June 30 divided by the market capitalization at year-end. The sample period is 198,101–201,412 for the liquidity measures and the stock characteristics, and 198,101–201,512 for the corresponding stock returns. Following Amihud [5], we impose the following data filters on the sample: (1) the stock has return data for more than 200 trading days during the year $t − 1$; (2) the stock price is greater than USD 5 at the end of year $t − 1$; and (3) the stock has market capitalization data available at the end of year $t − 1$ in CRSP (between 2498 and 5345 stocks are included in the cross-sectional regression). Table 4 presents the estimation results and reports the means of the coefficients from the monthly cross-sectional regressions of the stock returns on different liquidity measures and stock characteristics. We find that the LIQs and stock returns are negatively and significantly related, implying that when a stock's liquidity and premium are low, investors require a higher stock return as compensation. Consistent with the findings of Amihud [5], we also find that Amihud's illiquidity is positively related to the cross-sectional stock returns.

Table 4. Cross-sectional regressions of stock return on liquidity and other stock characteristics.

Variable	(1)	(2)	(3)
Amihud	0.090 *** (3.22)		
LIQ ($\gamma = 0.5$)		−0.440 *** (−2.74)	
LIQ ($\gamma = 1.5$)			−0.442 *** (−2.74)
BETA	0.027 (0.22)	−0.084 (−0.69)	−0.084 (−0.69)

Table 4. *Cont.*

R100	0.531 **	0.601 ***	0.602 ***
	(2.53)	(2.90)	(2.90)
R100YR	0.111	0.118	0.118
	(1.22)	(1.32)	(1.32)
SIZE	0.051	−0.050 *	−0.050 *
	(1.16)	(−1.84)	(−1.84)
BM	0.001	0.002	0.002
	(0.31)	(1.36)	(1.36)
SDRET	−0.273 ***	−0.386 ***	−0.385 ***
	(−4.79)	(−7.18)	(−7.17)
DIVYLD	−0.096	−0.105 *	−0.105 *
	(−1.03)	(−1.94)	(−1.94)
R^2	4.87%	4.86%	4.86%

Notes: *t*-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively (we believe that R^2 could increase by controlling additional factors affecting stock returns).

Given that LIQ explains the cross-sectional stock returns, similar to Amihud's illiquidity, we next investigate the effect of liquidity on excess returns on the portfolios sorted on LIQ and/or Amihud's illiquidity in the time-series. This allows us to directly compare the predictive power of the liquidity factor loading based on LIQ and Amihud's illiquidity in the excess returns. Table 5 presents the estimation results, and reports the mean returns and factor loadings of the portfolios formed on different liquidity measures. At the beginning of each year, stocks are sorted into 10 decile portfolios based on their previous year's liquidity measure. Stocks with the lowest liquidity values are included in decile 1, and those with the highest values are included in decile 10. LIQF is the liquidity factor, constructed as the value-weighted return of the stocks with the lowest 20% liquidity lagged by one month minus the value-weighted return of the stocks with the highest 20% liquidity lagged by one month. Further, MKT, SMB, and HML denote the three factors of Fama and French [18], and MOM denotes the momentum factor defined in the same work. The factor loadings are obtained by regressing the value-weighted portfolio returns on the market factors. Alpha is the intercept from the regression.

In particular, the results in Panel A of Table 5 are based on the portfolios sorted by Amihud's illiquidity and the liquidity factor loadings are formed on Amihud's illiquidity. We find that LIQF is significant, suggesting that the liquidity factor based on Amihud's illiquidity explains the return spread due to Amihud's illiquidity. Alpha remains significant in the five-factor setup, implying that the liquidity factor cannot fully explain the return spread. In Panel B, when the portfolios are sorted by LIQ and the liquidity factor loadings are formed on LIQ, the results show that LIQF is significant, suggesting that the liquidity factor based on LIQ explains the return spread due to LIQ. In addition, we find that alpha is insignificant in the five-factor setup, which suggests that the liquidity factor based on LIQ may fully explain the return spread. Panel C provides the estimation results for the portfolios sorted by Amihud's illiquidity and the liquidity factor loadings formed on LIQ. The results reveal a non-significant role of LIQF, showing that the liquidity factor based on LIQ has some explanatory power on the return spread due to Amihud's illiquidity. However, when the portfolios are sorted by LIQ and the liquidity factor loadings are formed on Amihud's illiquidity in Panel D, we find that the factor loading for LIQF is non-significant. This implies that the liquidity factor based on Amihud's illiquidity cannot explain the return spread due to LIQ. Considered together, our findings highlight that the liquidity factor based on LIQ is priced in the asset pricing test and captures the liquidity risk that Amihud's illiquidity cannot explain. Overall, the empirical results not only validate the robustness of the liquidity premium motivated by liquidity risk, but it also suggests that it plays a complementary role in the widely used Amihud's illiquidity ratio.

Table 5. Portfolio of stocks formed on liquidity lagged by one year.

(A) Scoring Variable = Amihud's illiquidity, Factor = Amihud's illiquidity

		Four-Factor Model					Five-Factor Model					
Portfolio	Mean Return (%)	Alpha (%)	MKT	SMB	HML	MOM	Alpha (%)	LIQF	MKT	SMB	HML	MOM
1 (Low)	2.633	1.632	0.644	0.881	0.057	0.194	0.987	−0.362	0.728	0.585	−0.039	0.295
2	2.527	1.442	0.776	0.706	0.279	0.092	0.905	−0.301	0.846	0.459	0.199	0.176
3	2.435	1.347	0.795	0.804	0.235	0.085	1.108	−0.134	0.826	0.695	0.199	0.123
4	2.418	1.308	0.888	0.808	0.127	0.089	1.177	−0.073	0.905	0.748	0.107	0.110
5	2.302	1.216	0.880	0.783	0.119	0.063	1.121	−0.053	0.893	0.740	0.105	0.078
6	2.226	1.148	0.923	0.599	0.155	0.019	1.062	−0.048	0.934	0.559	0.143	0.033
7	2.039	0.961	0.966	0.522	0.153	−0.012	0.902	−0.033	0.973	0.495	0.145	−0.003
8	2.051	0.965	0.976	0.534	0.110	0.019	1.023	0.033	0.968	0.560	0.119	0.009
9	1.916	0.810	1.042	0.332	0.142	0.002	0.826	0.009	1.040	0.339	0.145	−0.001
10 (high)	1.492	0.569	0.968	−0.149	−0.052	−0.014	0.590	0.012	0.965	−0.139	−0.049	−0.018
10 − 1 spread	−1.141 ***	−1.063 ***	0.324 ***	−1.030 ***	−0.109 ***	−0.209 ***	−0.397 **	0.373 ***	0.238 ***	−0.725 ***	−0.010	−0.313 ***
	(−5.17)	(−7.02)	(9.06)	(−20.39)	(−2.43)	(−6.18)	(−2.43)	(8.08)	(6.80)	(−12.01)	(−0.20)	(−9.22)

(B) Scoring Variable = LIQ, Factor = LIQ

		Four-Factor Model					Five-Factor Model					
Portfolio	Mean Return (%)	Alpha (%)	MKT	SMB	HML	MOM	Alpha (%)	LIQF	MKT	SMB	HML	MOM
1 (Low)	2.949	1.956	1.481	0.960	−0.979	−0.100	1.082	0.587	1.152	0.191	−0.519	0.091
2	2.520	1.530	1.355	0.614	−0.636	−0.113	0.999	0.357	1.155	0.146	−0.356	0.004
3	2.242	1.223	1.278	0.497	−0.418	−0.089	0.828	0.265	1.130	0.150	−0.210	−0.003
4	1.975	0.932	1.196	0.313	−0.202	−0.056	0.706	0.152	1.111	0.114	−0.082	−0.007
5	1.750	0.637	1.151	0.102	0.031	0.015	0.564	0.049	1.123	0.038	0.070	0.031
6	1.706	0.658	1.057	−0.062	0.145	−0.033	0.735	−0.052	1.086	0.006	0.104	−0.050
7	1.581	0.518	0.991	−0.182	0.249	0.026	0.658	−0.094	1.043	−0.060	0.176	−0.004
8	1.485	0.487	0.884	−0.165	0.269	0.010	0.679	−0.129	0.956	0.003	0.168	−0.032
9	1.371	0.470	0.724	−0.175	0.309	−0.010	0.696	−0.152	0.809	0.024	0.190	−0.059
10 (high)	1.306	0.505	0.526	−0.074	0.282	0.028	0.747	−0.163	0.617	0.139	0.154	−0.025
10 − 1 spread	−1.643 ***	−1.452 ***	−0.955 ***	−1.034 ***	1.261 ***	0.128 **	−0.335	−0.750 ***	−0.535 ***	−0.053	0.672 ***	−0.116 **
	(−3.36)	(−5.11)	(−14.24)	(−10.92)	(12.38)	(2.02)	(−1.63)	(−20.45)	(−10.42)	(−0.65)	(8.74)	(−2.53)

(C) Scoring Variable = Amihud's illiquidity, Factor = LIQ

		Four-Factor Model					Five-Factor Model					
Portfolio	Mean Return (%)	Alpha (%)	MKT	SMB	HML	MOM	Alpha (%)	LIQF	MKT	SMB	HML	MOM
1 (Low)	2.633	1.632	0.644	0.881	0.057	0.194	1.446	0.125	0.574	0.718	0.155	0.235
2	2.527	1.442	0.776	0.706	0.279	0.092	1.441	0.001	0.776	0.705	0.280	0.092
3	2.435	1.347	0.795	0.804	0.235	0.085	1.354	−0.005	0.798	0.811	0.231	0.084
4	2.418	1.308	0.888	0.808	0.127	0.089	1.268	0.027	0.873	0.773	0.148	0.098
5	2.302	1.216	0.880	0.783	0.119	0.063	1.194	0.015	0.872	0.764	0.131	0.068
6	2.226	1.148	0.923	0.599	0.155	0.019	1.165	−0.012	0.929	0.614	0.146	0.016
7	2.039	0.961	0.966	0.522	0.153	−0.012	0.956	0.004	0.964	0.517	0.156	−0.011
8	2.051	0.965	0.976	0.534	0.110	0.019	0.912	0.035	0.956	0.487	0.138	0.030
9	1.916	0.810	1.042	0.332	0.142	0.002	0.728	0.055	1.012	0.260	0.185	0.019
10 (high)	1.492	0.569	0.968	−0.149	−0.052	−0.014	0.521	0.032	0.950	−0.191	−0.027	−0.004
10 − 1 spread	−1.141 ***	−1.063 ***	0.324 ***	−1.030 ***	−0.109 ***	−0.209 ***	−0.925 ***	−0.093 ***	0.376 ***	−0.909 ***	−0.182 ***	−0.239 ***
	(−5.17)	(−7.02)	(9.06)	(−20.39)	(−2.01)	(6.18)	(−5.96)	(−3.36)	(9.75)	(−14.77)	(−3.14)	(−6.91)

Table 5. *Cont.*

		(D) Scoring Variable = LIQ, Factor = Amihud's illiquidity										
		Four-Factor Model					Five-Factor Model					
Portfolio	Mean Return (%)	Alpha (%)	MKT	SMB	HML	MOM	Alpha (%)	LIQF	MKT	SMB	HML	MOM
1 (Low)	2.949	1.956	1.481	0.960	-0.979	-0.100	1.858	-0.055	1.494	0.915	-0.994	-0.085
2	2.520	1.530	1.355	0.614	-0.636	-0.113	1.350	-0.101	1.378	0.531	-0.663	-0.084
3	2.242	1.223	1.278	0.497	-0.418	-0.089	1.145	-0.044	1.288	0.461	-0.430	-0.077
4	1.975	0.932	1.196	0.313	-0.202	-0.056	0.795	-0.077	1.214	0.250	-0.222	-0.035
5	1.750	0.637	1.151	0.102	0.031	0.015	0.551	-0.048	1.162	0.063	0.019	0.028
6	1.706	0.658	1.057	-0.062	0.145	-0.033	0.573	-0.048	1.068	-0.101	0.132	-0.020
7	1.581	0.518	0.991	-0.182	0.249	0.026	0.437	-0.046	1.001	-0.220	0.237	0.039
8	1.485	0.487	0.884	-0.165	0.269	0.010	0.543	0.032	0.877	-0.139	0.277	0.001
9	1.371	0.470	0.724	-0.175	0.309	-0.010	0.427	-0.024	0.730	-0.195	0.303	-0.003
10 (high)	1.306	0.505	0.526	-0.074	0.282	0.028	0.459	-0.025	0.532	-0.095	0.275	0.035
10 − 1 spread	-1.643 ***	-1.452 ***	-0.955 ***	-1.034 ***	1.261 ***	0.128 **	-1.399 ***	0.030	-0.962 ***	-1.010 ***	1.269 ***	0.119 *
	(-3.36)	(-5.11)	(-14.24)	(-10.92)	(12.38)	(2.02)	(-4.25)	(0.32)	(-13.63)	(-8.29)	(12.10)	(1.74)

Notes: *t*-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

5. Conclusions

This paper highlights the empirical application of an explicit representation of liquidity premium (or discount) of financial assets by parameterizing the return process in Hur and Chung [3]. They suggest that the derived measure of liquidity is confirmed to be a function of a time discount factor, a relative risk aversion parameter, and the expected return and volatility of the asset, given the risk-free rate.

Our empirical analysis particularly based on the US data suggests that the proposed measure of liquidity premium is highly comparable to the existing measures in the literature. This implies that the proposed liquidity premium can be used as an alternative to conventional measures of liquidity (e.g., it can be used in empirical asset pricing studies to investigate the liquidity betas of stocks). We also find a negative relationship between the liquidity premium and cross-sectional stock returns, as well as a negative relationship between the risk factor based on the liquidity premium and the expected excess returns on portfolios sorted on the liquidity premium in the time series. This suggests that the risk factor captured by the proposed liquidity premium measure is priced in stock and portfolio returns, which further corroborates the theoretical foundation of the new measure in Hur and Chung [3].

In future research, we would like to introduce an additional coefficient that governs the liquidity premium (or discount) in order to alleviate the burden on the risk-aversion parameter. In a related study, Epstein and Zin [19] adopt the power utility version of Kreps and Porteus [20] and separate risk aversion from the intertemporal elasticity of substitution in a similar manner.

Author Contributions: The authors contributed equally to this work.

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

References

1. Allen, D.E.; McAleer, M.; Singh, A.K. Risk measurement and risk modelling using applications of vine copulas. *Sustainability* **2017**, *9*, 1762. [CrossRef]
2. Yang, L.; Ma, J.Z.; Hamori, S. Dependence structures and systemic risk of government securities markets in central and eastern europe: A CoVaR-Copula approach. *Sustainability* **2018**, *10*, 324. [CrossRef]
3. Amihud, Y.; Mendelson, H. Dealership market: Market-making with inventory. *J. Financ. Econ.* **1980**, *8*, 31–53. [CrossRef]
4. Easley, D.; O'Hara, M. Price, trade size, and information in securities markets. *J. Financ. Econ.* **1987**, *19*, 69–90. [CrossRef]
5. Keim, D.B.; Madhavan, A. The upstairs market for large-block transactions: Analysis and measurement of price effects. *Rev. Financ. Stud.* **1996**, *9*, 1–36. [CrossRef]
6. Kyle, A.S. Continuous auctions and insider trading. *Econometrica* **1985**, *53*, 1315–1335. [CrossRef]
7. Hur, S.K.; Chung, C.Y. A novel measure of liquidity premium: Application to the Korean stock market. *Appl. Econ. Lett.* **2018**, *25*, 211–215. [CrossRef]
8. Amihud, Y. Illiquidity and stock returns: Cross-section and time-series effects. *J. Financ. Mark.* **2002**, *5*, 31–56. [CrossRef]
9. Karatzas, I.; Shreve, S. *Brownian Motion and Stochastic Calculus*, 2nd ed.; Springer: Berlin, Germany, 1991.
10. Guvenen, F. A parsimonious macroeconomic model for asset pricing. *Econometrica* **2009**, *77*, 1711–1750.
11. Mankiw, N.G.; Zeldes, S.P. The consumption of stockholders and nonstockholders. *J. Financ. Econ.* **1991**, *29*, 97–112. [CrossRef]
12. Mehra, R.; Prescott, E.C. the equity premium: A puzzle. *J. Monetary Econ.* **1985**, *15*, 145–161. [CrossRef]
13. Acharya, V.V.; Pedersen, L.H. Asset pricing with liquidity risk. *J. Financ. Econ.* **2005**, *77*, 375–410. [CrossRef]
14. Chordia, T.; Goyal, A.; Sadka, G.; Sadka, R.; Shivakumar, L. Liquidity and the post-earnings-announcement drift. *Financ. Anal. J.* **2009**, *65*, 18–32. [CrossRef]
15. Brennan, M.; Huh, S.W.; Subrahmanyam, A. An analysis of the Amihud illiquidity premium. *Rev. Asset Pricing Stud.* **2013**, *3*, 133–176. [CrossRef]

16. Fama, E.F.; MacBeth, J.D. Risk, return, and equilibrium: Empirical tests. *J. Polit. Econ.* **1973**, *81*, 607–636. [CrossRef]
17. Scholes, M.; Williams, J. Estimating betas from nonsynchronous data. *J. Financ. Econ.* **1977**, *5*, 309–327. [CrossRef]
18. Fama, E.F.; French, K.R. Common risk factors in the returns on stocks and bonds. *J. Financ. Econ.* **1993**, *33*, 3–56. [CrossRef]
19. Epstein, L.G.; Zin, S. Substitution, risk aversion and the temporal behavior of consumption and asset returns: An empirical framework. *J. Polit. Econ.* **1991**, *99*, 263–286. [CrossRef]
20. Kreps, D.M.; Proteus, E.L. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* **1978**, *46*, 185–200. [CrossRef]

Article

Financial Security and Optimal Scale of Foreign Exchange Reserve in China

Guangyou Zhou [1,†], **Xiaoxuan Yan** [1,†] **and Sumei Luo** [2,*,†]

[1]   School of Economics, Fudan University, Shanghai 200433, China; zgy@fudan.edu.cn (G.Z.);
       17210680201@fudan.edu.cn (X.Y.)
[2]   School of Finance, Shanghai University of Finance and Economics, Shanghai 200433, China
*   Correspondence: luosumei@shufe.edu.cn; Tel.: +86-21-6590-8396
†   These authors contributed equally to this work.

Abstract: The study of how foreign exchange reserves maintain financial security is of vital significance. This paper provides simulations and estimations of the optimal scale of foreign exchange reserves under the background of possible shocks to China's economy due to the further opening of China's financial market and the sudden stop of capital inflows. Focused on the perspective of financial security, this article tentatively constructs an optimal scale analysis framework that is based on a utility maximization of the foreign exchange reserve, and selects relevant data to simulate the optimal scale of China's foreign exchange reserves. The results show that: (1) the main reason for the fast growth of the Chinese foreign exchange reserve scale is the structural trouble of its double international payment surplus, which creates long-term appreciation expectations for the exchange rate that make it difficult for international capital inflows and excess foreign exchange reserves to enter the real economic growth mechanism under the model of China's export-driven economy growth; (2) the average optimal scale of the foreign exchange reserve in case of the sudden stop of capital inflows was calculated through parameter estimation and numerical simulation to be 13.53% of China's gross domestic product (GDP) between 1994 and 2017; (3) with the function of the foreign exchange reserves changing from meeting basic transaction demands to meeting financial security demands, the effect of the foreign exchange reserve maintaining the state's financial security is becoming more and more obvious. Therefore, the structure of foreign exchange reserve assets should be optimized in China, and we will give full play to the special role of foreign exchange reserve in safeguarding a country's financial security.

Keywords: the sudden stop of capital inflow; financial security; the optimal scale of foreign exchange reserve; utility maximization; finance risk

1. Introduction

In recent years, China's foreign exchange reserve has entered a sustained and rapid growth stage, and in 2006, it surpassed Japan as the largest foreign exchange reserve of the world. According to the statistics of the State Administration of Foreign Exchanges (SAFE), as of March 2018, the scale of China's foreign exchange reserve reached $3.14 trillion. Although in the past two years it had obvious fluctuations, it still ran at a high level. The foreign exchange reserve is an important part of the international reserve assets that are held by government; this part of assets can meet the demand of import and export trade, pay back the foreign debt, keep balances of payments, guarantee the stability of the exchange rate, safeguard national financial security, and play an irreplaceable role in the national economy.

Since the 1960s, researches on the moderate scale of the foreign exchange reserve has increased both at home and abroad. Research studies on the optimal scale of the foreign exchange reserve have

used both qualitative and quantitative methods; "ratio analysis", "cost-benefit analysis", "reserve function analysis", "the qualitative analysis" have all achieved fruitful results. With the global financial crisis of recent years and its growing infectivity, concern about the impact of the financial crisis around the world is unprecedented. The foreign exchange reserve has an important role to guard against and defuse financial risk; its importance is being further understood around the world, and it has been given a new connotation as a result. At the same time, how to make use of the foreign exchange reserve in ways that guard against financial risks, and research the optimal scale of the foreign exchange reserve for financial security, has become common concerns within theoretical circles both at home and abroad. Generally speaking, foreign exchange reserve management includes three aspects: scale management, the choice of currency structure, and asset structure optimization. Among the three, defining the optimal scale reserve reasonably is the basis and premises for the management of the foreign exchange reserve. Therefore, holding a huge foreign exchange reserve in China will bring greater risks. Under the background of the function of the foreign exchange reserve gradually changing from meeting basic transaction demands to meeting financial security demands, research on an optimal scale for foreign exchange reserve has important theoretical value and practical significance. Therefore, this article will define the foreign exchange reserve as the national financial assets. From the perspective of financial security, an analysis framework will be tentatively built based on the theory of utility maximization of the optimal scale of the foreign exchange reserve, and choose China's actual data to estimate and determine the important parameters of the theoretical model. The optimal scale of the foreign exchange reserve will be simulated and calculated based on financial security. Then, countermeasures and suggestions with strong maneuverability will be put forward.

The main differences between this article and the existing research are as follows. (1) The perspective of this study is novel. In the context of meeting basic trade and financial security demands, the foreign exchange reserves are regarded as national financial assets from the perspective of financial security. The optimal scale of China's foreign exchange reserve is presented under circumstances that take full account of the special functions of foreign exchange reserves. To a certain extent, this goes beyond most of the existing studies on foreign exchange reserves, which start from the perspective of demand. (2) The combination of the theoretical analysis and numerical simulation. In order to fully consider the external shocks caused by the sudden stop of capital inflow, this article introduces the utility maximization analysis method in order to construct a theoretical analysis framework for the foreign exchange reserve based on financial security through establishing a cross-term consumption model. It also measures the optimal scale of China's foreign exchange reserves and overcomes the shortcomings of most studies, which either focus only on theory or carrys out tests. At the same time, through these measurements, this article obtains a more intuitive and optimal scale for the foreign exchange reserve that is easy to control, and thus provides a strong and targeted suggestion for the foreign exchange reserve management department. (3) A variety of calculation methods are considered comprehensively. This study is not only an extension of the proportional analysis method; it is also an improvement of the cost–benefit analysis method, or a specific application of the reserve function analysis method. Therefore, to a certain extent, it makes up for the existing research that only measures the optimal scale of foreign exchange reserve in a limited way; as a result, the calculation of foreign exchange reserve is more reasonable, and the measurement results are more reliable.

The following sections of this article mainly include: a literature review in Part 2, a theoretical model in Part 3, a simulation and test of the optimal scale in Part 4, and the conclusion and enlightenment in Part 5.

2. Literature Review

Internationally, the study of foreign exchange reserve scale has had a long history. In very early times, the main function of the foreign exchange reserve was maintaining a stable domestic money supply, because under the gold standard, foreign exchange reserve assets were gold and sterling. The open trade and external economic capital investment factors have not been included in the study

of the international reserve. After the First World War, as the typical gold standard gradually collapsed, countries restricted the free export of gold, which resulted in the fluctuation of the exchange rate and an acceleration of the development of international trade. In the early 1930s, Keynes introduced foreign trade and investment fluctuations into the analysis of the international reserve. Then, the study of the international reserve scale had a new breakthrough. The study of the moderate scale of international reserve was mainly concentrated in the 1960s and 1970s; its main theory measured the optimal scale of foreign exchange reserve from quantitative or qualitative perspectives. With the development of the international financial field in the 21st century, including international investment and financing activities, and an increase in the frequency of international financial crises, the important role of foreign exchange reserve in the prevention of external capital impact was gradually incorporated into the theory of the optimal scale of the foreign exchange reserve. Among them, the representative measuring methods for a moderate scale of foreign exchange reserve can be summarized as follows:

The first method is ratio analysis, which refers to establishing a model to calculate the moderate scope through exchanging the foreign exchange reserve into one of the important indexes under the open economy, such as the import–export volume, debts, or the ratio of foreign output. Triffin (1960) [1] insisted that the ratio between the foreign exchange reserve and the import–export volume of a country's trade should not below a certain limit, which was 20% in their study. Another important standard is that the scale of a country's foreign exchange reserve should meet three months of its import volume. The "Triffin ratio" has become an international general index for estimating the adequacy of the foreign exchange reserve. The generalized ratio analysis method has become a basic thought in the theory of the foreign exchange reserve, and has been incorporated in the relevant theory of later scholars. In "Greenspan—Guidotti law", the foreign exchange reserve has been defined by its ratio with a country's short-term foreign debts. For this measurement, the foreign exchange reserve should be greater than the country's short-term foreign debts, and the sum of all of its long-term debt that is due within one year. In recent years, there has been a certain breakthrough for ratio analysis in academic research. According to Jeanne and Ranciere (2006) [2], the optimal scale of a country's foreign exchange reserve can be calculated by studying the ratio of the foreign exchange reserve to the country's gross domestic product (GDP) through use of a utility maximization model. The advantages of ratio analysis are that the optimal solution is a relative ratio, the calculation is relatively simple, and empirical data and analysis can be obtained for a single country. The drawback is that only one economic variable is considered, which will underestimate the effects of other factors on the foreign exchange reserve to a certain extent.

The second method is a cost–benefit analysis, which was first put forward by Heller (1960) [3]. Its main idea is to determine the optimal scale of foreign exchange reserves by maximizing the marginal revenue of a country's income in relation to its marginal cost. This method changes the orientation from thinking of the lowest line to seeking an optimal value or range. Heller considered three variables: the cost of holding foreign exchange reserves, the cost of the adjustment of external imbalances, and the probability of foreign exchange reserves requirements. The three variables can be used to estimate whether a country's foreign exchange reserves are excessive or insufficient. Agarwal (1971) [4] changed Heller's model by considering the economic and institutional differences between developed and developing countries. He argued that developing countries needed to use foreign currency to keep their balance of payments and buy foreign resources, and made the moderate scale model of foreign exchange reserves suitable for developing countries. Chinese scholars often use this model as a reference in their study of China's foreign exchange reserves at the moderate scale.

The third method is reserve function analysis, which is also called the regression analysis method. It is more intuitive than ratio analysis and cost–benefit analysis. This model considers many factors influencing the foreign exchange reserve requirements, establishes a model for the related parameters for the regression-influencing factors, uses significant variables to construct a function for foreign exchange reserve demand, and determines the moderate scale of the foreign exchange reserve. The reserve function analysis method was first proposed by Flanders (1971) [5], and further

improved by Frenkel (1974) [6], Iyoha (1976) [7], and other economists. Frenkel constructed a logarithm model that considered three variables—the balance of payments, the foreign exchange trading size, and average import trends—and then used statistical data for parameter estimation. Iyoha's research set up the dynamic demand function for developing countries and introduced the expected spending expected exports and the first-order and second-order lag dynamic variables of the change of import rate, to determine a country's foreign exchange reserve scale. In doing so, Iyoha research discovered the opportunity cost of variable salience. The cost of every 10% increase in the holding reserve would lead to a country's foreign exchange reserves decreasing by 9%.

The fourth method is qualitative analysis, which uses descriptions and the quality analysis method to analyze the scale of a country's foreign exchange reserve. The qualitative method determines a country's foreign exchange reserve according to the changes in a country's macroeconomic variables and the degree of the macroeconomic policies influence on macroeconomic variables. The main premise is that an ideal economic policy is moderate: if a country's foreign exchange reserve scale is moderate and its macroeconomic policy is reasonable, the changes in the economic variables indicators must be normal. Carbaugh and Fan (1976) [8] proposed that the quality of a country's reserves, the degree of its cooperation with foreign economic policy, the effectiveness of its international balance of payments adjustment mechanism, its government policy, international solvency stability, changes in the direction of the balance of payments, and a country's economic conditions should be put into the framework of qualitative analysis. The qualitative analysis method has the advantage of conforming to the actual situation of a country's foreign economic operation and being analyzed from an intuitive angle. It has guiden significance on the preliminary judgment of a macroeconomic policy. The downside is that it is difficult to quantify these variables. Since the variables do not have clear associations with a moderate scale of foreign exchange reserves and lack theoretical model support, the quantitative analysis and research of a country's foreign exchange reserve scale cannot be realized.

Since entering the 21st century, with the increase of the frequency of international financial crises, the risk of external capital flows involved measuring the optimal size of international foreign exchange reserves. Guaranteeing external capital flows and maintaining financial stability became the main target of foreign exchange reserves for individual countries, especially developing countries. After the Asian financial crisis, the foreign exchange reserves of developing countries grew at more than 60% a year. Mendoza (2004) [9] studied the policy implications of the self-insurance motive of holding excess foreign exchange reserves in 65 developing countries. Similarly, Aizenman and Lee (2007) [10] thought that East Asian countries such as China, Japan, and South Korea holding high foreign exchange reserves could be a monetary manifestation of mercantilism. That is to say, the cause of the rapid growth of China's foreign exchange reserves mainly lies in the East Asian countries maintaining exchange rate stability, steady trade, and financial system stability. In recent years, foreign scholars' researches on moderate scales of foreign exchange reserves became more innovative. Jeanne and Ranciere (2011) [11] added the "self-insurance" mechanism into the original utility maximization model, and further analysed the surge of foreign exchange reserves in emerging market countries after 1998. Their research suggested that the risk-aversion coefficient of emerging market countries, particularly East Asian countries, increased significantly, and that foreign exchange reserves increased as well to cushion the risks of crisis. These developments came about due to past crises suddenly stopping capital inflows and affecting domestic output and investment, which were kind of "mercantilism" thoughts. Aizenmant and Hutchison's (2012) [12] research showed the "absorbing" role of foreign exchange reserves in the foreign exchange market during the 2008–2009 global financial crisis. Aizenman et al. (2012) [13] argued that foreign exchange reserves could reduce output cost during the 2008–2010 financial crisis. Goncalo Pina (2014) [14] thought that although large foreign exchange reserves in developing countries would have a negative effect on the economy, the accumulation of moderate foreign exchange reserve by a central bank could share the costs that were associated with inflation over a period of time. The research of Pietro Cova et al. (2016) [15] showed that the diversification of foreign exchange reserves and "exorbitant privilege" both had a significant impact on global macroeconomic development.

Goncalo (2017) [16] studied the relationship between foreign exchange reserves and global interest rates, and argued that the movement of the exchange rate would affect foreign exchange reserves.

Researches on the moderate scale of China's foreign exchange reserves normally applies overseas models and theories to China's open economy. This article divides the research literature that is related to China into lack scale theory, moderate scale theory, and excess scale theory. Lack scale theory argues that China's foreign exchange reserves are insufficient. Few scholars have drawn this conclusion in recent years (liu Bin, 2003; Li Shikai, 2006) [17,18], and with the growth of foreign exchange reserves, the voice of insufficient reserves is gradually weakened. Scholars who believe in the theory of moderate foreign exchange reserve scale argue that China's current foreign exchange reserve scale is appropriate (Wang Qunlin, 2008; Li Wei, Zhang Zhichao, 2009; Deng Changchun, 2016) [19–21]. As China's foreign exchange reserves grew in recent years, the research conclusion of China holding an excess scale gained the upper hand (Zhou Guangyou, Luo Sumei, 2011; Yang Yi, Tao Yongcheng, 2011; Wang Wei, 2016 etc.) [22–24].

In recent years, Chinese scholars' researches on the appropriate scale of foreign exchange reserves tend to consider precautionary demands for foreign exchange reserves. Xiao Wen et al. 2012 [25] adjusted the Agarwal model, divided China's demand for foreign exchange reserves into six categories, focused on estimations of the preventive demand, and concluded that China has exceeded the optimal scale of foreign exchange reserves since 2004. This excess reserve originates from changes in demand for China's foreign exchange reserves, which were accumulated in order to maintain market stability and exchange rate stability. Man Xiangyu et al. focused on BRICS countries (Brazil, Russia, India, China, South Africa) including China, and found that China, Russia, and Brazil's foreign exchange reserves are deviating from optimal values, while South Africa is in a state of inadequacy, and India is now maintaining a moderate size. On this basis, emerging market countries should be classified according to the different needs and functions of their foreign exchange reserves. Jiang Boke and Ren Fei (2013) [26] took the exchange rate as the core variable, and studied the optimal scale of a country's long-term foreign exchange reserves by introducing the double equilibrium model. Gong Jian et al. (2017) [27] hold that among the macrovariables affecting the growth rate of the foreign exchange reserve, the effect of the real effective exchange rate on the growth rate of the foreign exchange reserve shows significant asymmetric and nonlinear characteristics that then affect the scale and structure of the foreign exchange reserve. Lu Lei et al. (2017) [28] estimated China's optimal foreign exchange reserves using the open conditional DSGE(Dynamic Stochastic General Equilibrium) model with China's economic characteristics. The results show that China's foreign exchange reserves have exceeded the optimal scale of foreign payment and prudent precautionary demand since 2004.

In summary, the research on the optimal scale of foreign exchange reserves at home and abroad is extremely rich, and valuable achievements also are emerging in an endless stream, which form the basis of this study. However, the existing studies abroad mainly focus on the same kind of countries (such as emerging market countries, although some studies also include China). Although such studies can reveal the common determinants of the optimal scale of foreign exchange reserves for similar countries, such analyses often overlook their different personality traits. As a result, it is difficult to measure the optimal scale of foreign exchange reserves in different countries by using common determinants. Especially for China, it is even less persuasive. Although China has carried out thorough researches on the optimal scale of China's foreign exchange reserves, most of them focus on the measurement of the optimal scale based on the demand of foreign exchange reserves, and seldom study the optimal scale of foreign exchange reserves from the perspective of financial security and financial risk. In fact, with the evolution of the function of foreign exchange reserves, the most important function of foreign exchange reserves has gradually changed from meeting transaction needs to meeting financial security needs, and the relevant research is relatively deficient. Therefore, in the context of the impact of the sudden halt of capital inflow on China, this paper, from the perspective of financial security, tries to introduce the optimal scale theory of foreign exchange reserves based on utility maximization. On the

basis of modifying the model, this paper selects the relevant data to simulate the optimal scale of China's foreign exchange reserves, and puts forward more targeted policy recommendations.

3. Theoretical Model

To judge whether a country's foreign exchange reserve is appropriate or not, it is necessary to take the optimal scale of foreign exchange reserve as the basis, and measure it with empirical data. This paper draws lessons from the model of Jeanne and Rancière (2011). The model uses the idea of utility maximization and the three-period model to simulate the foreign exchange reserve as a buffer mechanism that buffers the change of domestic absorption and reduces changes in the balance of payments when the capital suddenly stops. In terms of constraints, under the framework of maximum utility, the cost of holding foreign exchange reserves mainly lies in the cost of holding external liabilities when countries hold large amounts of foreign exchange reserves, which are lower than the interest rate gains. Under cost constraints and utility functions, the ratio of the foreign exchange reserve scale to GDP is used as a function of seven measurable variables: the probability of capital halt, the economic growth rate, the risk-free interest rate, the risk aversion coefficient, the time premium, the output loss rate, and the ratio of short-term foreign debt to output, which is used to measure the optimal scale of foreign exchange reserves in different emerging market countries according to their actual economic development.

3.1. Hypothesis and Derivation of the Model

Considering the small open economy in an emerging market, output Y_t can be expressed as the sum of domestic absorption A_t and trade account balances TB_t. Thus, domestic absorption A_t can be expressed as:

$$A_t = Y_t - TB_t \tag{1}$$

Under the international balance of payments, the balance TB_t of the trade account can be expressed as the reverse variable capital and financial account balance KA_t and foreign income and transfer payments IT, as well as the sum of change amount ΔR_t of the current foreign exchange reserves.

$$TB_t = -KA_t - IT_t + \Delta R_t \tag{2}$$

Through the simultaneous calculation of Equations (1) and (2), international absorption can be expressed as a function of total output, capital and financial account balances, income and transfer payments from abroad, and changes in foreign exchange reserves in the current period, that is:

$$A_t = Y_t + KA_t + IT_t - \Delta R_t \tag{3}$$

Equation (3) is the change mechanism of relevant variables in the normal flow of capital under the open economy. Then, we consider the change of the variable mechanism under the crisis situation, and assume that when capital inflows suddenly stop, and the KA account balance plummets, domestic absorption will decline accordingly. Since output Y and capital and financial account KA are also changing in the same direction, the domestic absorption due to the impact of capital halt will be amplified by the output effect. At this point, the government's strategy will be to use the reduction of foreign exchange reserves to compensate for the enlargement influence of the sudden halt of capital inflows on the domestic absorption, namely, adjusting ΔR_t to a negative value and consuming foreign exchange reserves. In reality, it can be understood that the government uses foreign exchange reserves to make up for the foreign debt that is difficult to pay because of the sudden halt of capital.

The ratio of financial account balances to the country's current GDP output is more than 5% lower than that of the $t - 1$ period, in which the sudden halt of capital inflow is defined as the capital of a country in the period of t. Namely, in defining $k_t = KA_t/Y_t$, moreover $k_{t-1} - k_t > 5\%$, and a sudden halt of capital inflows is considered to have occurred in the t period.

Continuing, we consider a small open economy, in a discrete period of $t = 0, 1, 2, \ldots$ where a commodity is consumed both at home and abroad. Without taking into account the real exchange rate movements, the only foreign shock to an economy is the risk of a sudden halt in capital inflows, without which the economy will continue to develop healthily along the path of output growth. The domestic economy consists of two parts of the private sector and the government sector. There is a representative consumer in the private sector, whose budget constraints are as follows:

$$C_t = Y_t + L_t - (1+r)L_{t-1} + Z_t \tag{4}$$

Among them, C_t is the current consumption, L_t is the current foreign debt, L_{t-1} represents the previous foreign debt, and Z_t is the transfer payment from the government, which can be understood as a contract signed between the government and consumers to help consumers in the event that they are unable to pay their foreign debts to reduce the foreign exchange reserve account, subsidize consumers, ensure the level of consumption, and pay off certain foreign debts. The short-term interest rate r is defined as a constant value. Therefore, when the consumer does not default on foreign debts, the current consumption budget is equal to the total output of the current period minus the remaining capital after repaying the current and last period of the foreign debt $Y_t + L_t - (1+r)L_{t-1}$, and plus the subsidy of the government's reserve contracts Z_t.

It is assumed that the two sectors of the economy, the private sector and the government sector, are growing at a constant rate of growth g, provided that capital inflows are normal. This growth will stop when the capital inflow is suddenly stopped. In the event of a sudden halt in capital inflows, there is a risk that foreign debt will not be repaid in the current period as a result of a decline in total output. In other words, there are two situations when capital inflow stops: one is that the representative consumer is unable to rollover the current foreign debt, and the other is that output Y has decreased at a γ rate relative to its long-term growth trajectory.

Suppose that the foreign debt of consumers is all short-term. When the capital suddenly stops, consumers cannot borrow from outside. The current external debt income L is reduced to 0, and the output is also out of the original growth trend and has decreased the γ ratio. After the collapse of the crisis of capital halt, the foreign debt income is still 0, and the output Y comes back to the original long-term growth path. It is assumed that the probability of each period of capital halt is π. After the capital halt, all of the uncertainties were removed, and the economy grew at a rate g less than the short-term risk-free rate r.

In order to simplify, assume that the crisis occurs only once, and b, d, and a are defined as three periods before, at, and after the occurrence of a sudden capital halt. λ represents the ratio of foreign debt to total output before the crisis, namely $\lambda = L_t/Y_t$. Therefore:

Before the crisis, $Y_t^b = (1+g)^t Y_0$; $L_t^b = \lambda(1+g)^t Y_0$
At the time of the crisis, $Y_t^d = (1-\gamma)(1+g)^t Y_0$; $L_t^d = 0$
After the crisis, $Y_t^a = (1+g)^t Y_0$; $L_t^a = 0$

Next, we consider the situation of government sector. Unlike the private sector, which can borrow only short-term foreign debt, governments can issue a long-term bond that does not require immediate repayment in the event of a capital standstill. The government-issued bonds pay a unit of the country's merchandise to bondholders as compensation until a capital halt occurs, and after a sudden capital halt, the bonds cease to yield. The term of government bonds tends to be very long, because the probability of sudden capital halt π is very small, and in order to be able to ensure that the term is long enough to cover the non-payment of the short-term foreign debt of the private sector, its term $1/\pi$ would be a relatively large value. For example, it is equal to 0.1, which means that government bonds should have a lifespan of 10 years.

Before the sudden stop of capital, the price of the government bond should be equal to the discount value of a unit commodity that it needs to pay in the next period, plus the present value of

the expected value of the market value of the bond. When calculating a unit of merchandise to be paid in each period, whether or not they are stopped, each period of payment will occur, that is, $\frac{1}{1+r+\delta}$. When calculating the expected average value of the bond market value, we must make sure that the price of the long-term bond is constant before the capital halt occurs, and it will be reduced to 0 when the capital suddenly stops, so the expected value should consider the probability π of capital sudden halt, that is:

$$\frac{1}{1+r+\delta}[(1-\pi)\cdot P]$$

Therefore:

$$\frac{1}{1+r+\delta}[1+(1-\pi)\cdot P]$$

And the following was solved:

$$P = \frac{1}{r+\delta+\pi}$$

Assuming that the interest rate level that was used to calculate the present value of long-term bonds is higher than the short-term interest rate level r, then the difference δ between the long-term and short-term interest rates exists as a time premium δ in the formula.

The government issued the long-term bonds to finance foreign exchange reserves because the government bonds cannot be issued at the time of capital arrest; then, foreign exchange reserves must rely on long-term bonds to accumulate foreign exchange reserves to a certain extent before the capital halt. Supposing that N_t is the number of long-term bonds issued by the government in the period of t, then the accumulated foreign exchange reserves are as follows:

$$R_t = PN_t; R_{t-1} = PN_{t-1} \tag{5}$$

Before the capital halt, with the government budget constraints, it means that government revenue and expenditure are equal, namely:

$$Z_t + R_t + N_{t-1} = P(N_t - N_{t-1}) + (1+r)R_{t-1} \tag{6}$$

The left side of Equation (6) is the sum of the total government expenditure in the current period, including the transfer payment to representative consumers, the value of the goods repaid in the previous period, and the necessary foreign exchange reserves for the current period. The right side of the Equation (9) is the total revenue of the current government, namely, the net income from the repayment of the principal of the previous long-term bond and the current period of borrowing, plus the present value of the foreign exchange reserves held for the $t-1$ period in the current period.

Taking advantage of Equation (5) as well as replacing and in Equation (6), in order to solve the expression of the transfer payment Z that the government subsidizes to the representative consumer in order to guarantee the level of consumption before the sudden halt occurs:

$$Z_t^b = -(\frac{1}{p} - r)R_{t-1} = -(\delta + \pi)R_{t-1} \tag{7}$$

As can be understood from Equation (7), prior to the occurrence of a capital halt, the transfer payment is a negative value, which is a tax levied by the government on the representative consumer to offset the cost to the government holding the reserve without investment, which is expressed as a proportion of the reserve, namely, the sum of the time premium δ and the probability of capital halt π.

When capital halt occurs, the government, while taxing, will transfer the entire net foreign exchange reserves of the previous period to subsidize the consumer and help him or her repay his or her short-term foreign debt, which cannot be postponed. Then, the transfer payment is:

$$Z_t^d = (1-\delta-\pi)R_{t-1} \tag{8}$$

Assuming $\delta + \pi < 1$, in the event of a capital halt, the transfer payments are positive values, so that the government subsidizes consumers.

After a sudden halt of capital, the transfer of the government stops, at which time the foreign exchange reserves R_t, transfer payments Z_t, and the number of long-term bonds N are all reduced to zero.

Then, we take advantage of Equations (7) and (8), as well as replace the transfer payments Z_t of Equation (4), so as to solve out the domestic consumer budget constraints before, during, and after the capital halt occurs.

$$C_t^b = Y_t^b + L_t^b - (1+r)L_{t-1}^b - (\delta + \pi)R_{t-1} \tag{9}$$

$$C_t^d = (1-\gamma)Y_t^b - (1+r)L_{t-1}^b + (1-\delta-\pi)R_{t-1} \tag{10}$$

$$C_t^a = Y_t^a \tag{11}$$

Equations (9) and (10) can well describe two aspects of trade-offs in the choice of the optimal scale of foreign exchange reserves: increasing the previous period of foreign exchange reserves R_{t-1} can increase domestic consumption C at the time of capital halt in this period, but it will also reduce domestic consumption (taxes that consumers have to pay to reduce the cost of holding excess foreign exchange reserves) when the current period of capital halt does not occur. In fact, the accumulation of foreign exchange reserves could be equivalent to an insurance measure that would transfer a portion of the purchasing power under the state of a steady capital flow to the state of capital halt to compensate for reduced domestic consumption.

In order to further close the model and obtain the closed solution of the optimal foreign exchange reserve, we need to introduce the constraint condition, that is, the government's objective effect function. Following the general social welfare theory, we assume that the government's goal is to optimize the welfare of this representative consumer. After the welfare function is added to the t period, every capital sudden halt may have the consumption utility function, and discounting:

$$U = \sum_{S=0,\ldots,+\infty} (1+r)^{-S} u(C_{t+s}) \tag{12}$$

Among them, the consumption utility function contains a constant relative risk aversion coefficient; the higher the degree of consumer risk aversion, the higher the welfare utility due to consumption.

$$u(C) = \frac{C^{1-\sigma} - 1}{1 - \sigma} \tag{13}$$

At this point, the government's strategy is to find out the scale of a foreign exchange reserve R_t in order to maximize the greatest utility of this representative consumer obtained in the t period before every capital sudden halt may occur.

Combining the budget constraints of the representative consumer and the government budget constraints, namely, Equations (4) and (6), the following equation could be obtained:

$$C_t = Y_t + (L_t - PN_t) - (1+r)(L_{t-1} - PN_{t-1}) + PN_t - (1+r+\delta+\pi)PN_{t-1} \tag{14}$$

Equation (14) shows that the amount of foreign exchange reserve R is equivalent to replacing consumers' non-renewable short-term debt, L, with the government's long-term debt, *PN*, in all of the foreign debt of a country. Under the constraints of the overall budget, holding foreign exchange reserves is equivalent to the government using the issuance of long-term bonds to repay the short-term foreign debt that a representative consumer cannot repay in the event of a sudden halt. Although long-term foreign debt reduces the risk that short-term foreign debt cannot be repaid, it brings higher holding costs.

3.2. Model Solution

At this time, the model is solved by the closed method. The optimal scale of foreign exchange reserves chosen by the government is the scale that maximizes consumer utility at time t before each sudden capital halt (which may or may not occur). According to the consumption utility function *u*, we can conclude that the optimum of R_t is only related to the consumption level of the $t + 1$ period. The optimal scale of the foreign exchange reserve in the t period maximizes the expected value of the utility function of the consumption level in the $t + 1$ period.

$$R_t = \text{argmax}[(1 - \pi)u(C_{t+1}^b) + \pi u(C_{t+1}^d)] \tag{15}$$

Among them, (C_{t+1}^b) and (C_{t+1}^d) are defined by the Equations (7) and (8) of the $t + 1$ period.

Then, the first condition is that the expected function has a derivative R_t of 0 to the first order, that is:

$$(1 - \pi)(\delta + \pi)u'(C_{t+1}^b) = \pi(1 - \delta - \pi)u'(C_{t+1}^d) \tag{16}$$

The left side of Equation (16) is the marginal cost of the probability without capital halt multiplied by the holding foreign exchange reserves without capital halt, and the right side of the equation is the marginal utility of the probability with capital halt multiplied by the consuming foreign exchange reserves in the event of capital halt. This condition can produce a closed solution to the optimal scale of foreign exchange reserves. Defining P_t is identically equal to the marginal substitution rate of consumption in cases of sudden halt and non-sudden halt, that is:

$$P \equiv \frac{u'(c_t^d)}{u'(c_t^b)} = \frac{(1 - \pi)}{\pi} \frac{\delta + \pi}{1 - \delta - \pi} \tag{17}$$

Substituting utility function *u(C)* can also be obtained:

$$(c_t^d)^{-\sigma} = p(c_t^b)^{-\sigma} \tag{18}$$

Considering a situation in which the time premium δ is zero, that is, the holding costs of long-term and short-term foreign debt are the same, the cost of using foreign exchange reserves to cope with the capital crisis is equal to the cost of servicing short-term foreign debt. There is no additional cost of holding foreign exchange reserves, and the consumption substitution rate *p* is identically equal to 1, which fully subsidizes the income budget and consumption that domestic consumers will lose because of sudden capital halt. If the time premium δ is positive, then $p > 1$, which means that domestic consumption would decrease when the capital stops abruptly.

In order to characterize the optimal scale of foreign exchange reserves and solve it conveniently, it is defined that under the premise of normal capital flow, the optimal scale of foreign exchange reserves is a constant proportion of output ρ in the next period, namely:

$$R_t = \rho Y_{t+1}^b \tag{19}$$

The expressions of $R_{t-1} = \rho Y_t^b$ as well as the output and short-term liabilities under the two period conditions of b and d are replaced by the two-period consumer budget Equations (9) and (10), and the consumer budget constraints under the two-period conditions are re-expressed. Then, simultaneously with Equation (18), the ratio of the optimal foreign exchange reserve to the output level is solved, which obtains the following equation:

$$\rho = \lambda + \gamma - \frac{p^{1/\sigma} - 1}{1 + (p^{1/\sigma} - 1)(1 - \delta - \pi)} \left(1 - \frac{r - g}{1 + g}\lambda - (\delta + \pi)(\lambda + \gamma)\right) \tag{20}$$

Then, Equation (20) is the measurement model of the optimal foreign exchange reserve scale based on utility maximization.

Under the framework of utility maximization analysis, we can use seven macroeconomic variables to measure the optimal scale of foreign exchange reserves. The optimal scale of foreign exchange reserves ρ, which are measured by output, are proportional to the probability of sudden capital halt π, time premium δ, economic growth rate g, and the risk aversion coefficient of consumer σ. Through a certain mathematical verification, the optimal scale of foreign exchange reserves ρ is also proportional to the rate of decline in output γ due to sudden capital halt and the ratio of short-term foreign debt to output λ.

It should be noted that this model does not fully consider the impact of real exchange rate movements on the scale of foreign exchange reserves. In fact, the fluctuation of the exchange rate is also one of the important factors affecting the fluctuation of capital flows and foreign exchange reserves. When the public expects that the exchange rate of a country has started to appreciate, the capital will flow into the country to gain the income of the foreign exchange, achieve the goal of its appreciation, and further strengthen the appreciation of the exchange rate. The process of capital inflow is also the process of the accumulation of foreign exchange reserves. When a country's exchange rate is expected to devalue, the international capital, especially hot money, will quickly flow into foreign countries, assuming the sudden stop and reversal of capital inflow, which not only causes the fluctuation of the country's exchange rate, but also strengthens the expected depreciation of the exchange rate of this country, thereby reducing foreign exchange reserves. In order to facilitate the study, this article assumes that the exchange rate is the same when constructing the optimal model of foreign exchange reserves, based on utility maximization. Of course, this does not mean that the optimal scale model of foreign exchange reserves based on utility maximization fails to consider the effect of exchange rate. In fact, when choosing the parameters, variables, and data, we will take full account of the impact of exchange rate fluctuations on foreign exchange reserves.

4. The Simulation and Measurement of the Optimal Scale

So far, on the basis of establishing the measurement model of the optimal scale of foreign exchange reserves based on utility maximization, the parameters are set in combination with the actual situation in China, the relevant data of 1994–2017 are selected, and the optimal foreign exchange reserve scale is calculated by using Formula (20).

4.1. Parameter Setting

4.1.1. The Probability of the Occurrence of the Sudden Arrest of Capital π

The JR model(the model of Jeanne and Rancière (2011))defines the sudden arrest of the capital inflow as the ratio of the capital and balance of the financial account 't' of a country to the current output GDP of the country being more than 5% lower than $t-1$. That is to say, if $k_t = KA_t/Y_t$ and $k_{t-1} - k_t > 5\%$, it will be considered that there is a sudden arrest of capital inflow in the t period of the country. To measure this probability, Jeanne and Rancière calculated the average probability of the occurrence of capital arrest in 34 middle-income countries in 1975–2003, and the result was 10%. In view of the situation of the capital arrest in China, this paper first calculates the number of the occurrence of capital arrest since the development of open economy in China. Since the relative value of π in the target period, 20 years, is slightly shorter in the calculation of this parameter, we choose the change trend of the proportion of the difference of Chinese capital and the financial account to the output in the 30 years between 1988 and 2017.

From Figure 1, we can see that in the 30-year period between 1988 and 2017, there is not a single year in which the proportion of the balance of capital and financial account to the output is consistent with the condition of the sudden arrest of capital inflow over the critical value of 5% under the JR model. This does not mean that the condition is unreasonable, but the situation in China is very special,

and does not conform to the capital and financial account not being fully open. In fact, when Jeanne and Rancière (2011) defined the condition for the sudden arrest of the capital inflow, they directly referred to the related research (Guidotti et al., 2004) [29] of the existing scholars on the impact of capital arrest. According to Guidotti's definition of "sudden arrest", the decline in the balance of the capital and the financial account is measured by taking the next standard deviation relative to the mean value as the critical value; then, the critical value of 5% can be understood as the mean value of the threshold value of the sudden arrest of the capital inflow in global countries.
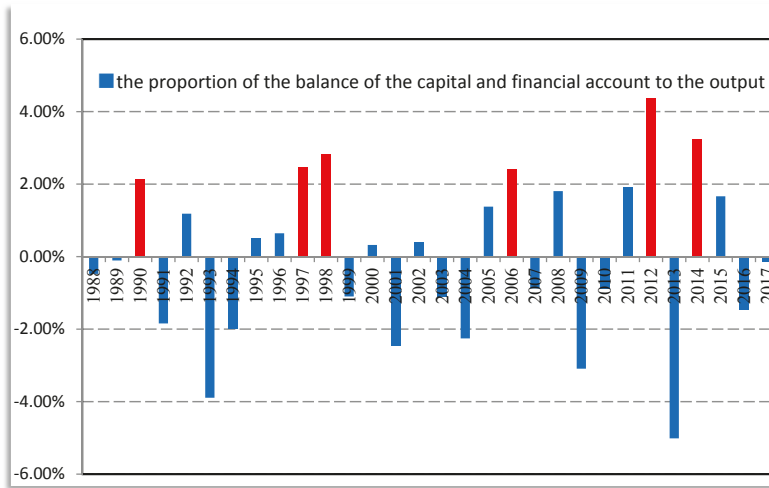
Figure 1. The proportion of the balance of the capital and financial account to the output. Data sources: China State Administration of Foreign Exchange and China National Bureau of Statistics.

However, because the situation is different in different countries, it is difficult to accurately measure the situation of a single country with the same threshold value. So, we need to recalculate the threshold value with the same method especially for China. According to the sample of the 30 years between 1988 and 2017, the mean value of reducing the degree of, $k_{t-1} - k_t$, the capital inflow ratio, k, is about 0.14%, and the standard deviation is about 2.24%. In combination with the reality of China, after fully considering that Chinese capital account are not fully open, there is still more strict control, the number of international capital imports and exports to China has been continuously increasing in recent years, and the impact on the economy and finance has become stronger and stronger, we set the critical value to 2%.

According to the critical value of 2%, in the 30 years between 1988 and 2017, the critical value of five years has exceeded 2%, and includes 1990, 1997, 1998, 2006, 2012 and 2015 (see the red part of the Figure 1). They all conform to the condition of the sudden arrest of Chinese capital inflow. Among them, the sudden arrest of the capital inflow in 1997 and 1998 was mainly caused by the Asian financial crisis, while the deficit in capital and financial accounts in 2006 was the result of the relaxation of the domestic capital outflows of countries. Of course, it also includes the reasons in the policy level of encouraging domestic enterprises to invest in foreign countries, increasing QDII (Qualified Domestic Institutional Investor), and so on. It is worth noticing that 2012 was the first year with the deficit in the annual capital and financial account since the 1998 crisis in China. Then, because of the capital flight caused by the United States (US) increasing the interest rate, the devaluation of the Chinese exchange rate, and other factors, there is the deficit again in 2015. However, no matter in which case, it will affect the consumption of the foreign exchange reserves of the government to offset the decline of net capital inflows. Therefore, considering the increasing frequency of the financial crisis in recent years,

the increasing intensity of the contagion, the increasing uncertainty, the increasing possibility of large capital outflows, the further deepening of Chinese financial opening, and the increasing probability of the sudden arrest of capital inflows, the probability of the sudden arrest of capital inflows in the JR model is relatively conservative, and the probability of the sudden arrest of capital inflows in China is identified as $\pi = 0.15$ in this paper.

4.1.2. The Calculation of the Economic Growth Rate, g, and the Risk Aversion Coefficient, σ

This paper selects the mean of the annual rate of growth of GDP from 1994 to 2017 and gets the economic growth rate, $g = 9.39\%$ (National Bureau of Statistics http://www.stats.gov.cn/tjsj/).

The risk aversion coefficient, σ, is the only parameter that cannot be directly measured in the seven parameters. However, we can make reference to the existing research and international experience to measure it. Previous studies have shown that the risk aversion coefficient of representative consumers is valued between two and six (Yang Yi, Tao Yongcheng, 2011), and the greater the value, the higher the degree of risk aversion. The risk aversion coefficient in the JR model is calculated with the data of the emerging Asian market and set to two. Studies also show that the risk aversion coefficient of developing countries is generally higher than that of developed countries (Donadelli, Prosperi, 2012). Since Chinese investors are relatively conservative and the degree of risk aversion is higher, this paper takes $\sigma = 5$ as the risk aversion coefficient to calculate.

4.1.3. The Calculation of Short-Term Risk-Free Interest Rate, r, and Time Premium, δ

About 80% of Chinese external debt balance comes from international commercial loans, while the bonds of the foreign government account for a small proportion. In terms of currency composition, the ratio of US dollar debt is about 70%, which means that in the demand of the foreign exchange, the US dollar is still a relatively general and stable currency. In existing studies, the risk-free interest rate of short-term foreign debt is usually replaced by the interest rate of short-term US Treasury bonds. However, as the current short-term foreign debt is mainly composed of international trade credit financing and interbank credit, compared with the interest rate of US Treasury bonds, the LIBOR (London Inter-Bank Offer Rate) is more reasonable for calculating the lowest cost of the foreign debt that is in China's possession. Here, this paper uses the weighted average $r = 0.03$ (StockQ database http://www.stockq.org/economy/libor.php) of the three-month LIBOR, and interest rates on three-month US Treasury bills within the last 20 years to calculate the interest rate level of holding short-term foreign debt.

The time premium, δ, represents the interest rate difference between long-term foreign debt and short-term foreign debt. The holding costs of Chinese long-term foreign debt are consistent with the US long-term treasury bonds. When the value of π is 0.1, it can be calculated that the average interest rate on 10-year US Treasury notes for nearly 20 years is 5% (US Treasury http://www.treasury.gov/resource-center/), so the time premium $\delta = 0.02$.

4.1.4. The Calculation of the Ratio of Short-Term Foreign Debt, λ, and the Rate of Output Loss, γ

According to the definition of the ratio of short-term foreign debt based on the GR model, we can calculate the ratio of short-term foreign debt to GDP for the 20 years between 1998 and 2017, and its average value $\lambda = 6.24\%$ (State Administration of Foreign Exchange http://www.safe.gov.cnzmodel_sarezindex.html). However, this only takes into account the demand of short-term foreign debts for foreign exchange reserves. In fact, China has medium and long-term foreign debts. It also has two ways of repayment. One is a one-time repayment at maturity. In this way, we should consider the maturity of the term within one year as a short-term bill. Another way is the annual amortization of medium and long-term foreign debts, and the part that is paid annually should also be regarded as a short-term bill. The two ways of repayment will make a large portion of the long-term foreign debts short-term, which will form short-term demand for foreign exchange reserves, while the GR model only considers the ratio of short-term foreign debts and estimates $\lambda = 6.24\%$, significantly

underestimating the effect of short-term foreign debts on foreign exchange reserves. Here, according to the condition of the medium and long-term external debt in China over the years, we assume that the average year is 10 years, allocate the total amount to the first year, consider the part whose time limit of the one-time repayment in the long-term foreign debt is within one year, and estimate the ratio of annual balance to GDP as about 3.5%. Therefore, combining short-term foreign debt and the short term of the medium and long-term debt, we determine the ratio of short-term foreign debt to be 8%.

The output loss rate, γ, is the loss when the sudden arrest of capital inflow occurs, and the output deviates from the original growth trajectory. When estimating the output loss rate in this paper, we will give full consideration to the impact of the sudden arrest of capital inflows and large amount of international capital outflows on China. On the one hand, from Figure 1, it can be seen that the sudden arrest of capital inflows in China in the last 30 years has occurred five times. Although the number is relatively low, in the context of the frequent international financial crisis of recent years, the probability of the sudden arrest of capital inflow will increase, and it will also accompany a large amount of international capital outflows at the same time. On the other hand, Chinese exports will be affected during the international financial crisis. For China, which has long pursued an export-oriented economic growth model, the financial crisis will inevitably affect the output. Based on these two considerations, we calculate the average γ of the decreasing amount of the output growth rate of 34 developing countries after the capital crisis calculated by the JR model as 6.5%, which is lower than the Chinese reality. Therefore, the output loss rate is set to 8.5% in this paper.

Based on the above analysis, the calculation results of each parameter can be summarized as follows (Table 1).

From Table 1, we can see that in combination with the reality of China, this paper calculates the original value of the JR model accordingly, which makes the model more convincing.

Table 1. Original Values and Calculated Values of Related Parameters.

Related Parameters	Original Values of JR Model	Calculated Values of This Paper
Probability of Capital Sudden Arrest π	0.1	0.15
Economic Growth Rate g	0.066	0.94
Consumer Risk Aversion Coefficient σ	2	5
Short-term Risk-free Interest Rate r	0.05	0.03
Time Premium δ	0.015	0.02
Ratio of Short-term Foreign Debt λ	0.107	0.08
Output Loss Rate γ	0.065	0.085

4.2. Solution of the Optimal Scale

Through the estimation of seven parameters, this paper will conduct the numerical simulation of the scale of Chinese optimal foreign exchange reserves under the framework of utility maximization, calculate the optimal scale of foreign exchange reserves, and compare it with the actual foreign exchange reserves.

First, we calculate the representative consumers' consumption level in the period that the capital normally flows and compare it to the consumption marginal substitution rate P in the period of capital arrest:

$$P \equiv \frac{u'(c_t^d)}{u'(c_t^b)} = \frac{(1-\pi)}{\pi} \frac{\delta + \pi}{1 - \delta - \pi} = 1.1606$$

Then, using Equation (20), $\rho = \lambda + \gamma - \frac{p^{1/\sigma}-1}{1+(p^{1/\sigma}-1)(1-\delta-\pi)}\left(1 - \frac{r-g}{1+g}\lambda - (\delta + \pi)(\lambda + \gamma)\right)$. It can be calculated $\rho = 13.53\%$.

Finally, take the ratio of ρ as the ratio of the scale of the optimal foreign exchange reserves and the output, and calculate the optimal scale of the foreign exchange reserves. The results are as follows.

In order to be more intuitive, it can also be illustrated as shown in Figure 2.

It can be seen from Table 2 and Figure 2, that first of all, from the perspective of the scale of foreign exchange reserves, due to China's reform of the foreign exchange system in 2005, and loosening the restrictions on capital flows, the scale of China's foreign exchange reserves has also become the world's largest foreign exchange reserves after breaking through the $1 trillion mark in 2006 and surpassing Japan, after which China has entered into a fast rising channel and reached a maximum value of $3 trillion and 840 billion in 2014, adding that the RMB(Renminbi) exchange rate continued to appreciate after 2005, forming a "double surplus" pattern under current account, capitaland financial account. Although it declined in 2015 and 2016, it rose again in 2017. At present, the scale of foreign exchange reserves has remained at around $3 trillion, but it is still operating at a high level.

Table 2. Comparison of Actual Scale of Foreign Exchange Reserves and Optimal Scale (Unit: 100 million dollars).

Year	Actual Scale	Optimal Scale	Excess Scale
1994	516.2	756.63	−240.43
1995	735.97	984.96	−248.99
1996	1050.29	1158.28	−107.99
1997	1398.90	1288.94	109.96
1998	1449.59	1379.33	70.26
1999	1546.75	1465.68	81.07
2000	1655.74	1621.54	34.20
2001	2121.65	1792.48	329.17
2002	2864.07	1967.02	897.05
2003	4032.51	2220.23	1812.28
2004	6099.32	2613.51	3485.81
2005	8188.72	3054.56	5134.16
2006	10,663.44	3671.36	6992.08
2007	15,282.49	4729.63	10,552.86
2008	19,460.30	6118.03	13,342.27
2009	23,991.52	6752.18	17,239.34
2010	28,473.38	7954.37	20,519.01
2011	31,811.48	10,111.34	21,700.14
2012	33,115.89	11,305.52	21,810.37
2013	38,213.15	12,229.31	25,983.84
2014	38,430.18	14,044.65	24,385.53
2015	33,303.62	14,859.77	18,443.85
2016	30,105.17	16,579.63	13,525.54
2017	31,399.49	17,823.86	13,575.63

Secondly, from the perspective of the optimal scale of the foreign exchange reserves, the optimal scale of China's foreign exchange reserves has increased synchronously with the actual scale of foreign exchange reserves. However, its rise is relatively slow, and has does not changed significantly with the short-term changes in the actual scale of foreign exchange reserves. The optimal scale of foreign exchange reserves based on financial security has been greatly affected by external shocks; especially since the US financial crisis in 2008, the growth rate of the optimal scale of China's foreign exchange reserves has been accelerating. However, the optimal scale of China's foreign exchange reserves has not been reduced because of the decline in the scale of China's actual foreign exchange reserves in 2015 and 2016. The major developed countries in the West entered the interest rate cycle in 2015; As a result, the fluctuations in the RMB exchange rate intensified and the signs of RMB depreciation emerged, leading to a massive outflow of the international capital from China and increasing financial risks. Therefore, from the perspective of the functional evolution of the foreign exchange reserves, the scale of foreign exchange reserves based on financial security will increase as the financial risks increase. The role of the foreign exchange reserves will become more and more obvious in maintaining the financial security of a country, which is consistent with the theoretical analysis of the previous article.
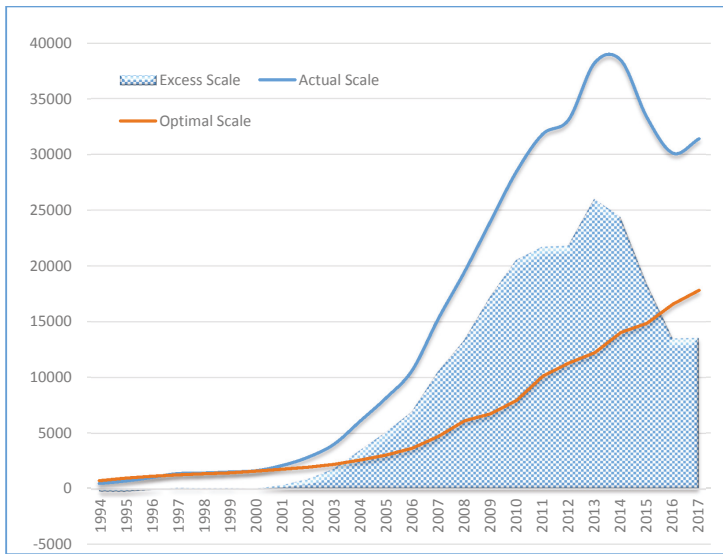
Figure 2. The Actual Scale and Optimal Scale of Chinese Foreign Exchange Reserves between 1994 and 2017. Data source: China National Administration of Foreign Exchange.

Finally, from the scale of excess foreign exchange reserves, before 1996, the actual scale of China's foreign exchange reserves was absent. After 1996, with the rapid rise in the scale of China's foreign exchange reserves, a large amount of excess foreign exchange reserves had been formed and fluctuated synchronously with the actual scale of foreign exchange reserves. At present, China's foreign exchange reserves had obviously surpassed the most optimum output ratio, and thus, it fails to achieve the maximization of the social welfare utility. In other words, on the premise of the goal of financial stability and the maximization of the social welfare of the government, China's actual foreign exchange reserves have gone beyond the optimal reserve scale since 2001. In 2017, the actual foreign exchange reserve of China was $3139.9 billion dollars. According to the calculation of this paper, the optimal foreign exchange reserve is $1782.4 billion dollars, while the excess foreign exchange reserve scale is $1357.6 billion dollars. Although the scale of Chinese foreign exchange reserves has fallen sharply since 2014, it still exceeded Chinese demand for foreign exchange reserves.

5. Conclusions and Enlightenment

To sum up, we can draw the following basic conclusions and enlightenment.

(1) With the rapid growth and the gathering of foreign exchange reserves in emerging markets, especially East Asian countries including China, the distribution of foreign exchange reserves has become increasingly uneven around the world. At the same time, with the increasing frequency of financial crises in recent years and more and more strong contagions, the special role of foreign exchange reserves in preventing financial risks and safeguarding national financial security has been fully recognized by most countries. The function of foreign exchange reserves has also shifted from simply meeting the needs of daily transactions to mainly meeting financial security demand. In China, the risks of foreign exchange reserves mainly lie in two aspects. On the one hand, the vast majority of China's foreign exchange reserves are held in the form of foreign government bonds. Such bonds are relatively low in yields, and will lead to a depreciation of China's foreign exchange reserve assets due to the devaluation of the currency of the issuer. This makes China's foreign exchange reserves face higher risk and opportunity cost, thus making it trapped in the "double shrinkage" situation. On the

other hand, a large number of foreign exchange reserves have forced the central bank to passively increase their supply of money, thus increasing the pressure of inflation and making it difficult for the central bank to implement monetary policy. Therefore, in this special historical period in which China holds huge foreign exchange reserves, we should give full play to the role of foreign exchange reserves in maintaining financial safety.

(2) This paper tries introducing the JR model. From the perspective of financial security, a theoretical analysis framework based on the maximization of utility is constructed on the basis of its correction. The framework takes the three-phase model of the external capital impact into the measurement process of the optimal scale of foreign exchange reserves, and introduces the social welfare function. Under the limit of the maximum of social welfare, we took a closed solution of the ratio of the optimal scale of foreign exchange reserves to output. This model is in line with the actual situation in which China is an open economy and an emerging market. It can integrate the risk factors of China's financial instability into the estimation of the opportunity cost and welfare benefits of holding foreign exchange reserves, and the seven related important parameters can be depicted with the operation index of the Chinese open economy. Finally, it can also calculate the optimal ratio of Chinese foreign exchange reserves to GDP.

(3) This paper takes the change trend of the scale of Chinese foreign exchange reserves from 1994 to 2017 as the research object. The stages of the scale changes of Chinese foreign exchange reserve are divided. The direct reason for the rapid growth of the scale of Chinese foreign exchange reserves is mainly the inflow of interregional capital, which was caused by the structural plight of the double surplus of international balance of payments and the long-term appreciation expectation of the exchange rate. The deeper reason is that under the economic growth mode driven by Chinese exports, the excess foreign exchange cannot enter a mechanism of real economic growth. Since 2014, the reduction of foreign exchange reserves is mainly due to the economic recovery of the main developed countries in the West, especially in relation to the US entering the cycle of increasing the interest rate. In addition, the devaluation of Chinese RMB leads the international capital flow out from China and the exchange rate fluctuation aggravates, which threatens the financial security of China. Therefore, under the circumstance of the abnormal flow of international capital, we should give full play to the role of foreign exchange reserves in maintaining financial safety and preventing financial risks.

(4) The impact of the sudden arrest of capital inflows on the finance of a country has aroused wide attention from countries all over the world. Questions regarding how to deal with the impact of the sudden arrest of capital inflows and how much foreign exchange reserves needed are appropriate. The study of this paper gives a preliminary answer. Through parameter estimation and numerical simulation, the average optimal foreign exchange reserves scale of China between 1994 and 2017 was 13.53% of GDP. With this ratio as the standard, the Chinese foreign exchange reserve has shown a significant surplus since 2001. On the one hand, the holding of excess foreign exchange reserves and the rapid growth will inevitably cause the rising holding cost of China's foreign exchange reserves, thus triggering the waste of resources and idle funds. On the other hand, the rapid growth of foreign exchange reserves aggravates the pressure of the appreciation of the RMB, which inevitably affects the international competitiveness of China's export commodities. This shows that there are excess foreign exchange reserves in China, but foreign exchange reserves are not a case of the more, the better. Too much foreign exchange reserves will not only bring risks to China, they will also increase the holding cost. Therefore, how to manage foreign exchange reserves scientifically is still a difficult problem to be solved at present.

Author Contributions: G.Z. performed the theory analysis and contributed to drafting this manuscript. S.L. conceived and empirical analysis. X.Y. analyzed the data.

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

References

1. Triffin, R. *Gold and the Dollar Crisis: The Future of Convertibility*; Yale University Press: New Haven, CT, USA, 1960.
2. Jeanne, O.; Rancière, R. The Optimal Level of International Reserves for Emerging Market Countries: Formulas and Applications. IMF Working Paper, No. 6. 2006. Available online: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiQ-fz77JXbAhVLerwKHYCSB8gQFggnMAA&url=https%3A%2F%2Fwww.imf.org%2Fexternal%2Fpubs%2Fft%2Fwp%2F2006%2Fwp06229.pdf&usg=AOvVaw3v7PcYfo6XVXeWSngL3a_d (accessed on 18 April 2018).
3. Heller, H.R. Optimal international reserves. *Econ. J.* **1966**, *302*, 296–311. [CrossRef]
4. Agarwal, J.P. Optimal monetary reserves for developing countries. *Weltwirtschaftliches Arch.* **1971**, *107*, 76–91. [CrossRef]
5. Flanders, M.J. *The Demand for International Reserves*; Princeton University: Princeton, NJ, USA, 1971.
6. Frenkel, J.A. The demand for international reserves by developed and less-developed countries. *Economica* **1974**, *161*, 14–24. [CrossRef]
7. Iyoha, M.A. Demand for international reserves in less developed countries: A distributed lag specification. *Rev. Econ. Stat.* **1976**, *3*, 351–355. [CrossRef]
8. Carbaugh, R.J.; Fan, L.S. *The International Monetary System: History, Institutions, and Analyses*; University Press of Kansas: Lawrence, KS, USA, 1976.
9. Mendoza, R.U. International reserve-holding in the developing world: Self-insurance in a crisis-prone era? *Emerg. Mark. Rev.* **2004**, *5*, 61–82. [CrossRef]
10. Aizenman, J.; Lee, J. International reserves: Precautionary versus mercantilist views, theory and evidence. *Open Econ. Rev.* **2007**, *2*, 191–214. [CrossRef]
11. Jeanne, O.; Rancière, R. The Optimal Level of International Reserves for Emerging Market Countries: A New Formula and Some Applications. *Econ. J.* **2011**, *555*, 905–930. [CrossRef]
12. Aizenman, J.; Edwards, S.; Riera-Crichton, D. Adjustment patterns to commodity terms of trade shocks: The role of exchange rate and international reserves policies. *J. Int. Money Financ.* **2012**, *8*, 1990–2016. [CrossRef]
13. Aizenman, J.; Hutchison, M.M. Exchange market pressure and absorption by international reserves: Emerging markets and fear of reserve loss during the 2008–2009 crisis. *J. Int. Money Financ.* **2012**, *31*, 1076–1091. [CrossRef]
14. Pina, G. The recent growth of international reserves in developing economies: A monetary perspective. *J. Int. Money Financ.* **2015**, *58*, 172–190. [CrossRef]
15. Cova, P.; Pagano, P.; Pisani, M. Foreign exchange reserve diversification and the exorbitant privilege: Global macroeconomic effects. *J. Int. Money Financ.* **2016**, *67*, 82–101. [CrossRef]
16. Pina, G. International reserves and global interest rates. *J. Int. Money Financ.* **2017**, *74*, 371–385. [CrossRef]
17. Liu, B. Empirical Analysis of the Change of Foreign Exchange Reserves. *Econ. Rev.* **2003**, *2*, 114–117.
18. Li, S. Measurement and Analysis of the Relative Scale of Chinese Foreign Trade Surplus: 1994–2004. *Financ. Trade Econ.* **2006**, *1*, 70–75.
19. Wang, Q. Empirical Analysis of Moderate Scale of Chinese Foreign Exchange Reserves. *Int. Financ. Res.* **2008**, *9*, 73–79.
20. Li, W.; Zhang, Z. A Framework for the analysis of foreign exchange reserves based on financial stability—On the moderate scale of chinese foreign exchange reserves. *Econ. Res.* **2009**, *8*, 27–36.
21. Deng, C. Analysis of effective management of Chinese foreign exchange reserves. *Manag. World* **2016**, *5*, 170–171.
22. Zhou, G.; Luo, S. Dynamic decision on the optimal scale of foreign exchange reserves—Analysis framework based on multilevel substitution effect. *Financ. Res.* **2011**, *5*, 29–41.
23. Yang, Y.; Tao, Y. Measurement of appropriate scale of Chinese international reserves: 1994–2009. *Int. Financ. Res.* **2011**, *6*, 4–13.
24. Wang, W.; Yang, J.; Wang, X.; Zhu, L. Development stage, exchange arrangement and the scale of Chinese high foreign exchange reserve. *World Econ.* **2016**, *2*, 23–47.

25. Xiao, W.; Liu, L.; Liu, Y. Study on the moderate scale and the demand structure of Chinese foreign exchange reserves—Based on the revised Agarwal model. *Financ. Trade Econ.* **2012**, *3*, 46–52.

26. Jiang, B.; Ren, F. A new exploration of the theory of the optimal foreign exchange reserve scale. *Fudan J. (Soc. Sci. Ed.)* **2013**, *4*, 10–16.

27. Gong, J.; Gao, T.; Zhang, Z. The Asymmetric transfer effect of exchange rate fluctuations on chinese foreign exchange reserves—Based on the nonlinear LSTARX-GARCH model. *Financ. Res.* **2017**, *2*, 84–100.

28. Lu, L.; Li, H.; Su, N. The optimal foreign exchange reserves and finance Open. *Ref. Financ. Trade Econ.* **2017**, *12*, 19–34.

29. Guidotti, P.E. On the consequences of sudden stops. *Economia* **2004**, *2*, 171–214. [CrossRef]

Article

Multi-Step Inflation Prediction with Functional Coefficient Autoregressive Model

Man Wang [1], Kun Chen [2,*], Qin Luo [3] and Chao Cheng [4]

1   Department of Finance, Donghua University, Shanghai 200051, China; wangman@dhu.edu.cn
2   School of Statistics, Southwestern University of Finance and Economics, Chengdu 611130, China
3   Guangxi Xijiang Venure Investment Co. Ltd., Nanning 530022, China; Luoqin__1123@163.com
4   Department of Mathmatical Sciences, Tsinghua University, Beijing 100084, China; cqplus@126.com
*   Correspondence: chenkun@swufe.edu.cn; Tel.: +86-028-8709-2207

Abstract: Forecasting inflation rate is one of the most important topics in finance and economics. In recent years, China has stepped into a "New Normal" stage of economic development, with a different state from the fast growth period during the past few decades. Hence, forecasting the inflation rate of China with a time-varying model may give high accuracy. In this paper, we investigate the problem of forecasting the inflation rate with a functional coefficient autoregressive (FAR) model, which allows the coefficient to change over time. We compare the FAR model based on the B-splines estimation method with the autoregressive moving average (ARMA) model by extensive simulation studies. In addition, with the monthly CPI data of China, we conduct both in-sample analysis and out-of-sample forecasting. The forecasting result shows that the FAR model based on the B-splines estimation method has a better performance than the ARMA model.

Keywords: B-splines; inflation forecast; monthly CPI data; out-of-sample forecast

1. Introduction

The inflation rate is a key index which is closely related to the economic stability and general well-being of a country. It guides policy-makers to formulate the country's macroeconomic and monetary policies. In addition, the households and businesses can make well-informed decisions based on future prices, and investors can construct long-run portfolios based on inflation rate (Bampinas and Panagiotidis, 2016 [1]). Due to these reasons, inflation forecasting attracts much interests from various fields.

The main methods to predict inflation rate include the Phillips curve model, the vector autoregressive-type (VAR) model, and the univariate linear autoregressive moving average (ARMA) model. In the past few decades, although the Phillips curve model has been widely adopted, many research results show that this model cannot provide satisfactory inflation forecasting for countries like China (Stock and Waston, 1999 [2]; Atkeson and Ohanian, 2001 [3]; Mcnelis and Mcadam, 2004 [4]; Matheson, 2006 [5]). On the other hand, the VAR-type model depends on some exogenous factors, for example the real GDP, unemployment, industrial production, manufacturing production, and capacity utilization. Based on this model, inflation forecasting is highly affected by the selection of exogenous factors (see Sekine (2001) [6], Ramakrishman and Vamvakidis (2002) [7] and Ang et al. (2007) [8]). Besides, if the dynamics of the inflation is non-linear, its prediction can be conducted in the framework of non-linear multivariate models. For instance, the regime-switching smooth transition vector autoregressive model used in Lekkos et al. (2007) [9] and the non-linear (asymmetric and polynomial) error correction models used in Milas et al. (2004) [10]. With the development of time series, univariate linear ARMA models without exogenous factors are widely adopted (Bos et al., 2001 [11]; Ang et al., 2007 [8]). In particular, Stock and Waston (1999) [2] found that

the univariate linear ARMA models perform better than the Phillips curve model and the VAR-type model in forecasting inflation rate.

However, the univariate linear ARMA model implies that the dynamic mechanism of the underlying process is time-invariant, which is not satisfied by many real time series data (Tong, 1990 [12]). In fact, many empirical studies showed that the dynamics of the inflation rate is time-varying (Chen et al., 2016 [13]).

If a linear ARMA model is used to fit a non-linear data generating process (DGP), the order of the ARMA model is always quite large, resulting in difficult and inaccurate estimation. Therefore, many non-linear time series models are proposed to fit this type of data, for example the threshold autoregressive model of Tong (1990) [12], the bilinear model of Granger and Andersen (1978) [14] and the exponential autoregressive model of Haggan and Ozaki (1981) [15]. These well-known non-linear models belong to parametric models, meaning that one needs to specify the formulation in advance, which is difficult and questionable in many real applications. To overcome this hurdle, some non-parametric time series models are proposed. For example, the non-parametric autoregressive conditional heteroscedastic model and the non-parametric autoregressive model (see Fan and Yao, 2003 [16]).

Although the non-parametric approach is appealing, its application usually requires an unrealistically large sample size when more variables are introduced into the model. This problem is called "curse of dimensionality". In order to avoid this problem and preserve the appreciable flexibility, semi-parametric models are proposed by imposing parametric structures to part of the non-parametric model. Chen and Tsay (1993) [17] proposed the functional-coefficient autoregressive (FAR) model for time series. Similarly, Hastie and Tibshirani (1993) [18] proposed the varying-coefficient model to increase the flexibility of ordinary linear regression model and improve the out-of-sample prediction. Chen and Hong (2012) [19] constructed a test of the smooth transition autoregressive model versus the FAR model. More recently, Chen et al. (2016) [13] established a functional coefficient moving average (FMA) model, with the coefficients estimated by using the local linear estimation technique. Application is made to the monthly CPI data of China. In this paper, we forecast the inflation rate in China by the FAR model, which belongs to the semi-parametric non-linear time series models. The FAR model has three main advantages. First, its formulation is analog to the linear ARMA model and preserve the satisfactory forecasting power. Second, the FAR model is flexible in coefficient specification and easy to interpret. Third, the semi-parametric formulation avoids the possible model mis-specification problem.

A key step to use the FAR model is estimating the time-varying coefficient. The main estimation methods include the kernel estimation method, the local polynomial estimation method, the spline estimation method and the wavelet method. The first two estimation methods belong to local estimation methods, while the last two methods belong to the global estimation method. Based on an iteration algorithm, Chen and Tsay (1993) [17] adopted the arranged local regression estimation and achieved good fitting results. Cai et al. (2000) [20] proposed the local linear estimators and investigated the bandwidth selection problem. Chen and Liu (2001) [21] mainly focused on the local polynomial estimators and associated hypothesis test. Huang and Shen (2004) [22] extend the FAR model to the general functional coefficient regression model and adopted the polynomial spline estimator to estimate the coefficient functions. They showed the consistency of the spline estimator and forecasted the Dutch guilder-US dollar exchange rate based on the estimated model. In this paper, we adopt the B-spline estimators (belonging to the polynomial spline estimator) to estimate the coefficient functions. Compared to other polynomial splines, the B-spline method is numerically stable and the calculation can be obtained recursively, thus significantly reducing the computational burden.

The rest of the paper is organized as follows. Section 2 introduces the FAR model and the spline-based nonparametric estimation method, and provides some implementation details. Section 3 conducts some simulation studies to compare the performance of the FAR model and the ARMA

model for fitting data generated from different DGPs. In Section 4, we apply the estimated FAR model to forecast the inflation rate of China. Finally, in Section 5, we give some conclusions.

2. Functional Coefficient Autoregressive Model

2.1. Model Specification

The FAR(p) model proposed by Chen and Tsay (1993) [17] extends the autoregressive (AR) model with the form

$$x_t = \theta_1(U_t)x_{t-1} + \theta_2(U_t)x_{t-2} + \cdots + \theta_p(U_t)x_{t-p} + \varepsilon_t, \ t = 1, \ldots, T, \tag{1}$$

where p is a positive integer representing the lag order, $\{\varepsilon_t\}_{t=1}^{T}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean 0 and variance σ^2, $\text{Cov}(x_s, \varepsilon_t) = 0$ for $s < t$, and $\theta_i(\cdot)$ $i = 1, 2, \ldots, p$ are unknown measurable functions. Here, x_t is called the significant variable and U_t is called the threshold variable which can be an exogenous or endogenous variable. In the time series setting, U_t is usually set to be the lagged value of x_t, i.e., $U_t = x_{t-d}$ with $1 \leq d \leq p$. Then the resulting model (denoted as FAR(p, d)) becomes

$$x_t = \theta_1(x_{t-d})x_{t-1} + \theta_2(x_{t-d})x_{t-2} + \cdots + \theta_p(x_{t-d})x_{t-p} + \varepsilon_t, \ t = 1, \ldots, T. \tag{2}$$

Throughout this paper, we impose some conditions on model (1).

Regularity Conditions (RC)

(a) The density function of U_t is nonzero and bounded.
(b) Let $X_t = (x_{t-1}, x_{t-2}, \ldots, x_{t-p})'$, then for any $u \in \mathbb{R}$, the eigenvalues of $\text{E}(X_t X_t' | U_t = u)$ are nonzero and bounded.
(c) The density function of ε_t is positive everywhere.

If **RC** is satisfied, the FAR(p) model (1) is geometric ergodic (Chen and Tsay, 1993 [17]).

2.2. Model Estimation and Prediction Based on B-Spline

We first discuss the identification of the coefficient functions $\theta_i(\cdot), i = 1, \ldots, p$ for model (1). The coefficient functions are said to be identifiable if $\sum_{j=1}^{p} \theta_j^{(1)}(U_t)x_{t-j} \equiv \sum_{j=1}^{p} \theta_j^{(2)}(U_t)x_{t-j}$ implies that $\theta_j^{(1)}(U_t) = \theta_j^{(2)}(U_t), j = 1, 2, \ldots, p$. We can prove that the coefficients in model (1) are identifiable under Condition (b). In fact, if we denote $\boldsymbol{\theta}(U_t) = (\theta_1(U_t), \ldots, \theta_p(U_t))'$, then

$$\text{E}\left[\left\{\sum_{j=1}^{p} \theta_j(U_t)x_{t-j}\right\}^2 \mid U_t = u\right] = \boldsymbol{\theta}(U_t)' \text{E}(X_t X_t' | U_t = u)\boldsymbol{\theta}(U_t). \tag{3}$$

If $\sum_{j=1}^{p} \left(\theta_j^{(1)}(U_t) - \theta_j^{(2)}(U_t)\right) x_{t-j} \equiv 0$, then $\text{E}\left[\left\{\sum_{j=1}^{p} \left(\theta_j^{(1)}(U_t) - \theta_j^{(2)}(U_t)\right) x_{t-j}\right\}^2\right] = 0$.

From this, we can immediately obtain $\text{E}\left[\left\{\sum_{j=1}^{p} \left(\theta_j^{(1)}(U_t) - \theta_j^{(2)}(U_t)\right) x_{t-j}\right\}^2 \mid U_t = u\right] = 0$ for any u. By applying (3) and Condition (b), we further get that $\theta_j^{(1)}(U_t) - \theta_j^{(2)}(U_t) = 0$ almost surely for $j = 1, \ldots, p$.

Since the coefficient functions in model (1) are identifiable, we can now estimate the model. In the literature, the main estimation methods include the kernel estimation method, the local polynomial method, the wavelet method and the spline method, see Fan and Yao (2003) [16] for more details. In particular, as a global smoothing method, the spline method outperforms the kernel estimation method and the local polynomial method for multi-step-ahead forecasting. Moreover, the spline-based method is computationally feasible and the estimated model has a parsimonious explicit expression.

Therefore, we can easily produce multi-step-ahead forecasts by iteratively generating one-step-ahead forecast based on the previous forecasts. On the contrary, if the functional coefficients are estimated based on the local polynomial method, then it is computationally intensive to conduct multi-step-ahead forecasting. Pointed out by Huang and Shen (2004) [22], one needs extra effort to relieve the computational burden for forecasting based on the local polynomial method.

With the aforementioned advantages, the spline-based estimation method is adopted in this paper to estimate the FAR model. Among many different types of splines functions, we adopt the polynomial splines, which are piecewise smooth. This means that the polynomial pieces join together smoothly at a set of interior knot points. Specifically, a polynomial spline of degree $K \geq 0$ on interval $[a, b]$ with knots $a = \xi_0 < \xi_1 < \cdots < \xi_{M+1} = b$ is a polynomial function of degree K on each of the intervals $[\xi_i, \xi_{i+1}), 0 \leq i \leq M - 1$ and $[\xi_M, \xi_{M+1}]$, and has $K - 1$ continuous derivatives on $[a, b]$.

When $K = 0, 1, 2, 3$, the polynomial splines are called constant spline, linear spline, quadratic spline and cubic spline, respectively. For a given sequence of knots and degree K, the corresponding collection of spline functions form a linear function space. Discussed in de Boor (1972) [23], there are different basis for this space, among which the B-spline basis is most widely used. The B-splines are numerically more stable than other polynomial splines and can be obtained recursively (see Fan and Yao, 2003 [16]). Therefore, we use B-spline basis in this paper for its good numerical property.

In the following, we display the calculation of B-splines. Any spline function of degree K with interior knot sequence $\{\xi_i\}_{i=1}^{M}$ can be expressed as a linear combination of the corresponding B-spline basis $B_{i,K}(x)$, given by

$$B(x) = \sum_{i=0}^{M+K} \beta_i B_{i,K}(x), \tag{4}$$

where $B(x)$ is the spline function and $B_{i,K}(x)$, $i = 0, \ldots, M + K$ are the B-splines. To calculate $B_{i,K}(x)$, we first define $\xi_{-K} = \cdots = \xi_0 < \xi_1 < \cdots < \xi_{M+1}$ as the new augmented sequence of knots, and relabel it as $\eta_0 = \cdots = \eta_K < \eta_{K+1} < \cdots < \eta_{M+K+1}$. The B-splines basis $B_{i,j}(x)$ for $j = 0, 1, \ldots, K$, $i = 0, 1, 2, \ldots, M + K$ are as follows:

$$B_{i,0}(x) = \begin{cases} 1, & \eta_i \leq x \leq \eta_{i+1}, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

and

$$B_{i,j+1}(x) = \frac{x - \eta_i}{\eta_{i+j+1} - \eta_i} B_{i,j}(x) + \frac{\eta_{i+j+2} - x}{\eta_{i+j+2} - \eta_{i+1}} B_{i+1,j}(x). \tag{6}$$

The recursive Formula (6) shows that the B-spline basis relies on the knot sequence and the degree K.

Now, we give an example to demonstrate the calculation process for the B-spline function. Assume the interval is $[0, 1]$ with the interior knot sequence $\{0.25, 0.5, 0.75\}$ and $K = 3$, then the augmented knots vector is $(0, 0, 0, 0, 0.25, 0.5, 0.75, 1)$. When $x = 0.1$, the values of $B_{i,j}(x)$, $i = 0, 1, 2, 3, j = 0, 1, 2, 3$ are shown in Table 1.

Table 1. An example of the calculation process of B-spline.

$j = 3$	$B_{0,3}(0.1)$	$B_{1,3}(0.1)$	$B_{2,3}(0.1)$	$B_{3,3}(0.1)$	$B_{4,3}(0.1)$	$B_{5,3}(0.1)$	$B_{6,3}(0.1)$
	0.216	0.592	0.1813	0.0107	0	0	0
$j = 2$	$B_{0,2}(0.1)$	$B_{1,2}(0.1)$	$B_{2,2}(0.1)$	$B_{3,2}(0.1)$	$B_{4,2}(0.1)$	$B_{5,2}(0.1)$	$B_{6,2}(0.1)$
	0	0.36	0.56	0.08	0	0	0
$j = 1$	$B_{0,1}(0.1)$	$B_{1,1}(0.1)$	$B_{2,1}(0.1)$	$B_{3,1}(0.1)$	$B_{4,1}(0.1)$	$B_{5,1}(0.1)$	$B_{6,1}(0.1)$
	0	0	0.6	0.4	0	0	0
$j = 0$	$B_{0,0}(0.1)$	$B_{1,0}(0.1)$	$B_{2,0}(0.1)$	$B_{3,0}(0.1)$	$B_{4,0}(0.1)$	$B_{5,0}(0.1)$	$B_{6,0}(0.1)$
	0	0	0	1	0	0	0

The B-spline is then used to approximate the coefficient functions $\theta_i(U_t)$, $i = 1, \ldots, p$. According to de Boor (1978) [23] and Schumaker (2007) [24], if $\theta_i(U_t)$ is assumed to be smooth, then as the number of knots tends to infinity, $\theta_i(U_t)$ will be well approximated by a linear combination of the corresponding K-degree B-spline basis. That is, there exist a vector of constants $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{ip})'$ and spline function $\theta_i^{\star}(U_t)$, such that

$$\theta_i(U_t) \approx \theta_i^{\star}(U_t) = \sum_{j=0}^{M+K} \beta_{ij} B_{j,K}(U_t), \; i = 1, \ldots, p. \tag{7}$$

Then we approximate model (1) by

$$x_t \approx \sum_{i=1}^{p} \left(\sum_{j=0}^{M+K} \beta_{ij} B_{j,K}(U_t) \right) x_{t-i} + \varepsilon_t. \tag{8}$$

Based on (8), the estimation of approximation model (8) is equivalent to the estimation of the vector of parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_p')'$. We estimate $\boldsymbol{\beta}$ by the method of ordinary least squares (OLS), i.e.,

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\mathrm{argmin}} \sum_{t=p+1}^{T} \left(x_t - \sum_{i=1}^{p} \sum_{j=0}^{M+K} \beta_{ij} B_{j,K}(U_t) x_{t-i} \right)^2. \tag{9}$$

Once $\hat{\boldsymbol{\beta}}$ is obtained, the OLS estimate of $\theta_i(U_t)$ is given by

$$\hat{\theta}_i^{\star}(U_t) = \sum_{j=1}^{M+K} \hat{\beta}_{ij} B_{j,K}(U_t), \text{ for } i = 1, \ldots, p. \tag{10}$$

In particular for the FAR(p, d) model (2), the estimated model becomes

$$x_t = \sum_{i=1}^{p} \sum_{j=0}^{M+K} \hat{\beta}_{ij} B_{j,K}(x_{t-d}) + \epsilon_t. \tag{11}$$

where ϵ_t are residuals. Based on (11), we can construct multi-step-ahead forecasts for x_t. Let $\hat{x}_t(l)$ be the minimum mean square error prediction of x_{t+l}, then it can be carried out by iteratively implementing one-step-ahead prediction as

$$\hat{x}_{t+1} = \sum_{i=1}^{p} \left(\sum_{j=0}^{M+K} \hat{\beta}_{ij} B_{j,K}(\hat{x}_{t+1-d}) \right) \hat{x}_{t+1-i},$$

$$\vdots$$

$$\hat{x}_{t+l} = \sum_{i=1}^{p} \left(\sum_{j=0}^{M+K} \hat{\beta}_{ij} B_{j,K}(\hat{x}_{t+l-d}) \right) \hat{x}_{t+l-i}, \tag{12}$$

where $\hat{x}_{t+l-d}$ equals x_{t+l-d} if $l - d \leq 0$.

2.3. Selection of Threshold Variable and Significant Variables

For ARMA(p, q) model, the impact of the past value x_{t-i}, $i = 1, \ldots, p$ on the current value x_t is direct and linear. However, in the real world, the impact may be related to another past value x_{t-d}. Compared to the ARMA(p, q) model, the FAR(p, d) model (2) is more flexible, where the impact of x_{t-i} on x_t can be related to the threshold variable x_{t-d}. The flexibility of the FAR(p, d) model can alleviate the model mis-specification problem and reduce forecasting error caused by choosing wrong models. In addition, the FAR(p, d) model does not involve the selection of exogenous variables. Due to these advantages, the FAR(p, d) model is widely used in empirical studies. Chen and Tsay (1993) [17] used the FAR(p, d) model to fit the monthly records of cases of chickenpox in New York City and the Wolf's

annual sunspot numbers data set. Harvill and Ray (2005) [25] applied the FAR(p, d) model to forecast the U.S.GNP and the unemployment rate. Recently, the FAR(p, d) model has been extended and used in many fields, including survival analysis and risk management. For example, Xie et al. (2014) [26] proposed the varying-coefficient expectile (VCE) model to estimate the value at risk. In particular, by using the closing bid prices of the Euro in terms of the U.S. dollar, they applied the VCE model to fit the expectile of the weekly exchange return.

In this paper, we use the FAR(p, d) model to fit and forecast the inflation rate of China, and in the following we discuss some issues arising in the implementation of the FAR(p, d) model. In modelling and estimating the FAR(p, d) model, a key issue is choosing an appropriate threshold variable x_{t-d} and a set of significant variables $\{x_{t-i}\}_{i=1}^{p}$. This is equivalent to selecting the threshold lag d and the significant lag p with $d \leq p$. In practice, this can be achieved by a two-stage procedure. At the first stage, the significant lag p can be chosen based on subjective determination or by objective data driven methods such as information criteria. At the second stage, based on the chosen p, one decides d by minimizing the information criterion such as AIC (Akaike, 1974 [27]) or BIC (Schwarz, 1978 [28]) when p is fixed.

2.4. Selection of B-spline Basis Related Quantities

A crucial step in the B-spline estimation is to determine the B-spline basis, which is equivalent to determining the degree K, the number of knots M and the knots locations. The most commonly used degrees of the B-splines are 2 and 3, corresponding to the quadratic spline and the cubic spline. The determination of K can be conducted by the method of information criterion. Essentially, M is a smoothing parameter. As M increases, the spline function becomes less smooth leading to a more complicated model, while as M decreases, the spline function become more smooth with worse model fitting. In practice, M is chosen to balance the smoothness and model fitness. Huang and Shen (2004) [22] showed that the AIC outperforms other criterions such as the BIC and the modified cross-validation (Cai et al., 2000 [20]) in choosing M. Therefore in this paper we use the AIC to determine M.

For a given M, there are two popular ways to arrange the interior knots: $a < \xi_1 < \cdots < \xi_M < b$. One method gives the equally spaced knots, meaning that the distance between two adjacent knots is the same. The other one is the quantile knots, meaning that the knots locate at the $i/(M+1)$ sample quantiles $(i = 1, \ldots, M)$ of the threshold variable. If the distribution of the threshold variable is not flat, the quantile knots are preferable (Huang and Shen, 2004 [22]).

3. Simulation Study

In this section, we check the fitting performance of the FAR model for different DGPs. Intuitively, if the underlying DGP follows a constant coefficient process, the linear ARMA will provide a better fit than the non-linear FAR model. On the other hand, if the underlying DGP follows time-varying coefficient process, then the FAR model can outperform than the ARMA model. Specifically, we compare the performance of the FAR and autoregressive (AR) model for different DGPs. The first case is a constant AR(2) model:

$$\text{Case I:} \quad \left\{ \begin{array}{l} x_t = \theta_1^0 x_{t-1} + \theta_2^0 x_{t-2} + \varepsilon_t, \\ \theta_1^0 = 0.4, \quad \theta_2^0 = 0.5, \end{array} \right. \tag{13}$$

with $\varepsilon_t \overset{i.i.d.}{\sim} N(0,1)$. The second case is the exponential AR model in Haggan and Ozaki (1981) [15] and Huang and Shen (2004) [22]:

$$\text{Case II:} \quad \begin{cases} x_t = \theta_1^0(x_{t-1})x_{t-1} + \theta_2^0(x_{t-1})x_{t-2} + \varepsilon_t, \\ \theta_1^0(x_{t-1}) = 0.138 + (0.316 + 0.982x_{t-1})e^{-3.89x_{t-1}^2}, \\ \theta_2^0(x_{t-1}) = -0.437 - (0.659 + 1.26x_{t-1})e^{-3.89x_{t-1}^2}. \end{cases} \quad (14)$$

In each simulation, a series of length $T = 100, 400, 1000$ and 2000 are drawn and the experiment is replicated 1000 times. For each case, we fit the generated data by both FAR(2,1) and AR(2) models. The coefficients of FAR(2,1) model and AR(2) model are estimated by the B-spline method and the OLS method, respectively. For each run s, the resulting estimates are denoted as $\hat{\theta}_{i,s}^\star(x_{t-d})$ and $\hat{\theta}_{i,s}$, respectively, $i = 1, 2, s = 1, \ldots, 1000$. All the simulations are run in R 3.4.3. To estimate the ARMA(p, q) model, we use the "arima" function built in the "stats" package.

To evaluate the fitting performance of both the FAR(2,1) model and the AR(2) model in the Cases I&II, we adopt the bias (BIAS) and the root mean square error (RMSE) of the corresponding estimates. Since the coefficients of the FAR model are functions of the threshold variable, the definitions of BIAS and RMSE of FAR models are different from those of AR models. Specifically, the BIAS and RMSE for Case I are defined as

$$\text{BIAS}_i = \begin{cases} \frac{1}{1000}\sum_{s=1}^{1000}(\hat{\theta}_{i,s} - \theta_i^0), & \text{AR(2)}, \\ \frac{1}{1000}\sum_{s=1}^{1000}(\frac{1}{T-2}\sum_{t=3}^{T}\hat{\theta}_{i,s}^\star(x_{t-1}) - \theta_i^0), & \text{FAR(2,1)}, \end{cases}$$

and

$$\text{RMSE}_i = \begin{cases} \sqrt{\frac{1}{1000}\sum_{s=1}^{1000}(\hat{\theta}_{i,s} - \theta_i^0)^2}, & \text{AR(2)}, \\ \sqrt{\frac{1}{1000}\sum_{s=1}^{1000}\frac{1}{T-2}\sum_{t=3}^{T}(\hat{\theta}_{i,s}^\star(x_{t-1}) - \theta_i^0)^2}, & \text{FAR(2,1)}, \end{cases}$$

respectively, $i = 1, 2$. The BIAS and RMSE for case II are defined as

$$\text{BIAS}_i = \begin{cases} \frac{1}{1000}\sum_{s=1}^{1000}(\frac{1}{T-2}\sum_{t=3}^{T}\hat{\theta}_{i,s} - \theta_i^0(x_{t-1})), & \text{AR(2)}, \\ \frac{1}{1000}\sum_{s=1}^{1000}(\frac{1}{T-2}\sum_{t=3}^{T}\hat{\theta}_{i,s}^\star(x_{t-1}) - \theta_i^0(x_{t-1})), & \text{FAR(2,1)}, \end{cases}$$

and

$$\text{RMSE}_i = \begin{cases} \sqrt{\frac{1}{1000}\sum_{s=1}^{1000}\frac{1}{T-2}\sum_{t=3}^{T}(\hat{\theta}_{i,s}^\star - \theta_i^0(x_{t-1}))^2}, & \text{AR(2)}, \\ \sqrt{\frac{1}{1000}\sum_{s=1}^{1000}\frac{1}{T-2}\sum_{t=3}^{T}(\hat{\theta}_{i,s}^\star(x_{t-1}) - \theta_i^0(x_{t-1}))^2}, & \text{FAR(2,1)}, \end{cases}$$

respectively, $i = 1, 2$. The results of BIAS and RMSE are reported in Tables 2 and 3.

Table 2. BIAS and RMSE of AR and FAR modelds for Case I.

T	Index	θ_1		θ_2	
		AR(2)	FAR(2,1)	AR(2)	FAR(2,1)
100	BIAS	−0.0332	−0.0486	−0.0454	−0.0051
400	BIAS	−0.0063	−0.0094	−0.0117	0.0042
1000	BIAS	−0.0025	−0.0036	−0.0053	0.0028
2000	BIAS	−0.001	−0.0015	−0.0026	0.0015
100	RMSE	0.0971	2.2412	0.1058	1.8270
400	RMSE	0.0449	1.8076	0.0455	1.5866
1000	RMSE	0.0278	1.7246	0.0281	1.5442
2000	RMSE	0.0191	1.6309	0.0194	1.5191

Table 3. BIAS and RMSE of the two models for case II.

T	Index	θ_1		θ_2	
		AR(2)	**FAR(2,1)**	**AR(2)**	**FAR(2,1)**
100	BIAS	0.0735	0.0003	0.0050	−0.0001
400	BIAS	0.0628	0.0001	0.0150	−0.0001
1000	BIAS	0.0609	0.0001	0.0177	0
2000	BIAS	0.0622	0.0001	0.0161	0
100	RMSE	1.8801	0.0038	2.8682	0.0060
400	RMSE	3.4010	0.0038	5.4510	0.0057
1000	RMSE	5.2670	0.0053	8.5164	0.0096
2000	RMSE	7.4267	0.0074	11.9953	0.0145

The results in Tables 2 and 3 show that, in terms of bias, when the DGP follows Case I, the performances for the FAR(2,1) and the AR(2) are similar, while the performance for FAR(2,1) is better than the AR(2) when the DGP follows Case II. Note that the RMSE for the FAR(2,1) specification in Case I is relatively large. This fact is reasonable because the squared sum of the difference between the constant true coefficients and the fitted time-varying coefficients becomes large as the time varies. Similarly, the RMSE for the AR(2) fitting for Case II are larger than that for the FAR(2,1). In terms of the RMSE, the difference between the performance of FAR(2,1) and AR(2) in Case II are larger than that in Case I. This result is due to the complicated time-varying coefficients of Case II. This fact implies that the FAR model works well when the underlying DGP is a constant coefficient, but the AR model works poor when the DGP is a time-varying coefficient model. In conclusion, the FAR model can reduce the risk of model mis-specification and thus can be widely used in empirical study when one has no prior information on the true model.

4. Empirical Analysis

4.1. Data Preprocessing

In this paper, inflation is measured in terms of the Consumer Price Index (CPI). The growth rate of CPI can be regarded as a proxy for the inflation rate. In particular, we use the monthly CPI data of China from Jan. 1995 to December 2017, with a total of 276 observations denoted as $\{x_t\}_{t=1}^{276}$. The data is displayed in Figure 1, showing that the CPI is quite large in the beginning of this period and drops down slowly. The inflation rate has been relatively stable at a level around 2% since 2012, which indicates that the development of the economy of China has stepped into the stage of "New Normal". Intuitively, the underlying inflation process may change since the beginning of the "New Normal" period.
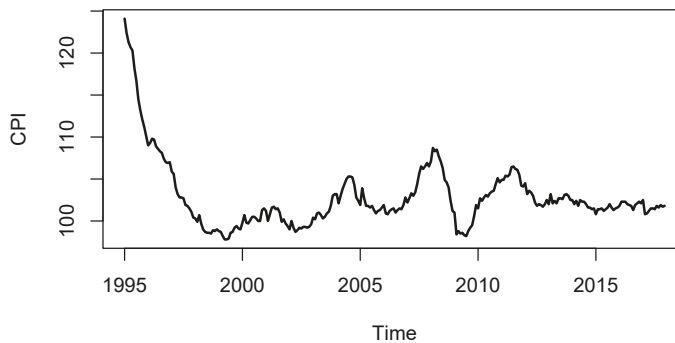

Figure 1. Monthly CPI data of China from January 1995 to December 2017.

In practice, we take the logarithm of the raw CPI data, and check the stationarity of the log-CPI data, at significance level 0.05. The result of the Phillips-Perron unit root test shows that the log-CPI data is nonstationary (with *p*-value 0.1494), but the first order difference of the log-CPI is stationary (with *p*-value 0). Therefore, the following analysis is based on the first order difference of the log-CPI denoted as

$$y_t = \log(x_t) - \log(x_{t-1}).$$ (15)

The plot of y_t is shown in Figure 2. Parameter instability is also observed in our analysis of the data. We build an AR(1) model $y_t = \theta y_{t-1} + \epsilon_t$ and estimate the AR coefficient θ on an expanding window basis and rolling window basis with a 60 window-width. These estimates are plotted in Figure 3. It can be seen that the estimates of θ are quite variable. In conclusion, a non-linear FAR(p, d) model with time-varying coefficients is more reasonable and flexible than the linear ARMA model.
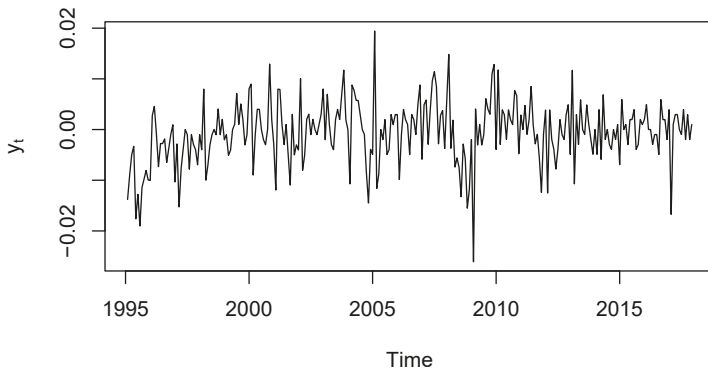
Figure 2. The first order difference of the monthly log-CPI.
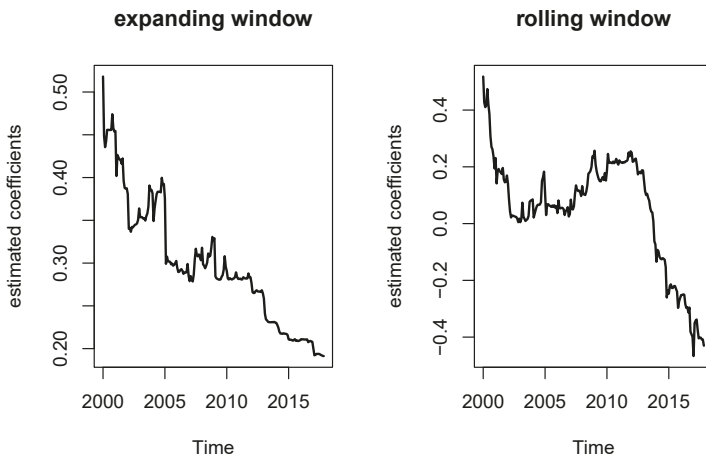
Figure 3. Estimates of θ. **Left:** expanding window; **right:** rolling window.

4.2. In-Sample Fit Analysis

We first use the FAR(p, d) model (2) to fit y_t given in (15). Based on the AIC values, we have $p = 4$. Then $d \leq p = 4$. For a given threshold lag d, the FAR(4, d) model is estimated by the B-splines method. To construct the B-spline basis, the degree of the B-spline basis, the number and locations of the knots

need to be determined. We consider different choices of the degree K and the number of interior knots M, i.e., $K = 2, 3$ and $M = 1, 2, 3, 4, 5$. For different values of (d, K, M), the locations of the knots ξ_i, $i = 1, \ldots, M$ are set to be the $i/(M + 1)$ sample quantiles $(i = 1, \ldots, M)$ of x_{t-d}. The B-spline basis can be calculated according to (5) and (6), and is then used to estimate the FAR$(4, d)$ model by (9). We use the AIC criterion to determine the value of (d, K, M).

Table 4 reports the AIC values for different combinations of (d, K, M), showing that $(d, K, M) = (2, 2, 1)$ leads to the smallest value of AIC. Therefore, we use y_{t-2} as the threshold variable. $K = 2$ and $M = 1$ implies that the computational complexity is not large. The only internal knot is the median of $\{y_{t-2}\}_{t=3}^T$, while the boundary knots are $\max_{3 \leq t \leq T}\{y_{t-2}\}$ and $\min_{3 \leq t \leq T}\{y_{t-2}\}$, respectively. Thus, the resulting augmented vector of knots is $(-0.02608, -0.02608, -0.02608, -0.02608, -0.0009, 0.0194)$.

Table 4. AIC values of FAR models with different (d, K, M).

AIC		$d = 1$	$d = 2$	$d = 3$	$d = 4$
	$M = 1$	−2002.20	−2008.77	−2006.09	−2001.36
	$M = 2$	−2002.38	−2002.27	−2003.53	−1994.18
	$M = 3$	−2005.80	−1996.11	−1997.89	−1989.72
$K = 2$	$M = 4$	−2002.97	−1989.58	−1991.41	−1993.06
	$M = 5$	−2007.87	−1985.52	−1983.82	−1996.12
	$M = 1$	−2006.04	−2003.56	−2003.16	−1994.44
	$M = 2$	−2006.58	−1996.92	−1996.13	−1995.36
	$M = 3$	−2002.43	−1990.42	−1989.67	−1990.75
$K = 3$	$M = 4$	−2006.26	−1987.78	−1985.55	−1985.35
	$M = 5$	−2002.85	−1982.75	−1983.23	−1987.87

The estimated FAR(4,2) model is

$$\hat{y}_t = -0.0003 + \hat{\theta}_1^\star(y_{t-2})y_{t-1} + \hat{\theta}_2^\star(y_{t-2})y_{t-2} + \hat{\theta}_3^\star(y_{t-2})y_{t-3} + \hat{\theta}_4^\star(y_{t-2})y_{t-4}, \tag{16}$$

where the estimated coefficients are $\hat{\theta}_i^\star(y_{t-2}) = \sum_{j=0}^3 \hat{\beta}_{ij}B_{j,2}(y_{t-2})$, where $\hat{\beta}_{ij}$ are given in Table 5.

Based on the B-spline estimation results in Table 5, we plot the estimated functional coefficients of the FAR(4,2) model (16) in Figure 4. The fitted first order differenced monthly log-CPI $\hat{y}_t$ can be recursively obtained by (16). The Ljung-Box test shows that the residuals $\{\hat{\epsilon}_t = y_t - \hat{y}_t\}_{t=5}^T$ is a white noise sequence, indicating that the FAR(4,2) model provides a satisfactory model fitting. By substituting $\hat{y}_t$ into (15), we obtain the fitted CPI $\hat{x}_t$.

We also fit y_t by an ARMA(p, q) model, where $p, q \leq 6$. The order p and q are determined by the AIC criterion, which gives $p = q = 5$. The estimated model is

$$\hat{y}_t = -0.0008 + 0.6556y_{t-1} + 0.9923y_{t-2} - 0.7343y_{t-3} - 0.6319y_{t-4} + 0.5494y_{t-5}$$
$$- 0.5173\epsilon_{t-1} - 0.1533\epsilon_{t-2} + 0.6032\epsilon_{t-3} + 0.9862\epsilon_{t-4} - 0.5607\epsilon_{t-5}. \tag{17}$$

where $\epsilon_{t-1}, \ldots, \epsilon_{t-4}$ are residuals. Also, the residuals of model (17) are white noise, implying a satisfactory model fitting.

Table 5. B-spline estimation results of the FAR(4,2) model.

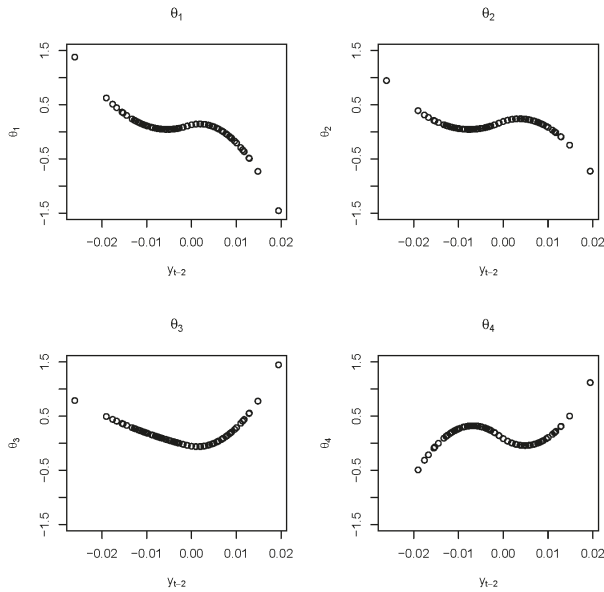| $\hat{\beta}_{ij}$	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$i = 1$	1.3807	−0.2424	0.3918	−1.4556
$i = 2$	0.9471	−0.2776	0.5190	−0.7262
$i = 3$	0.7882	0.2199	−0.2503	1.4473
$i = 4$	−1.6952	0.9303	−0.5024	1.1203

Figure 4. Estimation results of the coefficient functions.

Figures 5 and 6 display $\hat{y}_t$ and $\hat{x}_t$ given by the estimated FAR(4,2) and ARMA(5,5) models, respectively. From Figure 6, both the FAR(4,2) and the ARMA(5,5) describe the main characteristic of data quite well. In particular, they capture the three falling and rising processes of the inflation since 1995. Intuitively, the fitted ARMA(5,5) is more fluctuated, meaning that the model overestimates the peak and underestimates the trough, while the fitted FAR(4,2) model is more smoothing, i.e., the FAR(4,2) model underestimates the peak and overestimates the trough. Thus, it is expected that the ARMA(5,5) model fit the CPI data better during the fluctuating period, while the FAR(4,2) model performs better during the stable period. From Figure 6, the inflation rate during 1995 to 2011 fluctuates heavily, and the fitted CPI by using the ARMA(5,5) model is closer to the real CPI than that given by the FAR(4,2) model. However, for the period from 2012 to 2017, the inflation rate is quite stable around the level of 2%, and the FAR(4,2) model performs better.
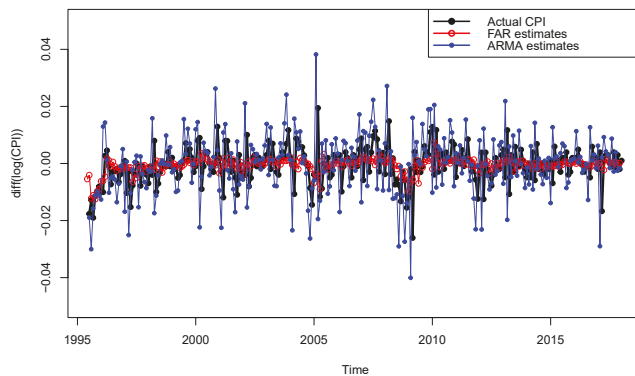
Figure 5. Estimated first order difference of the log-CPI given by the FAR(4,2) model and the ARMA(5,5) model.
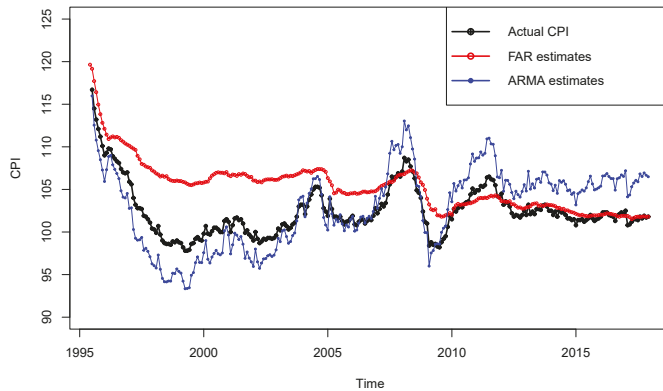
Figure 6. Estimated CPI given by the FAR(4,2) model and the ARMA(5,5) model.

We compare the fitting performance of the FAR(4,2) model and the ARMA(5,5) model for x_t by using mean absolute errors (MAE) and RMSE defined as

$$\text{MAE} = \frac{1}{T} \sum_{i=1}^{T} |x_t - \hat{x}_t|, \tag{18}$$

and

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{T} (x_t - \hat{x}_t)^2}{T}}. \tag{19}$$

The definitions of MAE and RMSE for y_t are similar and thus are omitted. Furthermore, to compare predictive accuracy, we employ the Diebold and Mariano test (denoted as DM test hereafter) proposed in Diebold and Mariano (1995) [29], based on the corresponding MAE and RMSE. We only introduce the DM test based on the MAE for simplicity. Define the forecasting error of x_t as $e_t = x_t - \hat{x}_t$ and the loss function $L(e_t) = |e_t|$. Based on $L(e_t)$, the forecasting loss difference between ARMA and FAR is $d_t = L(e_t^A) - L(e_t^F)$. Here, the null hypothesis is that the ARMA model has equal predictive accuracy as the FAR model, which is equivalent to $E(d_t) = 0$. Let $\bar{d} = \frac{1}{T} \sum_{t=1}^{T} d_t$, then the DM statistic is given by

$$DM = \frac{\bar{d}}{\hat{\sigma}_{\bar{d}}},$$

where $\hat{\sigma}_{\bar{d}}$ is a consistent estimate of the standard deviation of $\bar{d}$. The alternative can be set as the forecast of FAR is more accurate than that of ARMA. The procedure proceeds as follows. First, we calculate DM test statistic and the two-sided p-value based on the limiting standard normal distribution. Then if the value of the test statistic is positive (negative), the one-sided p-value is just one half of the two-sided p-value (one minus one half of the two-sided p-value). Similarly, if the alternative is that the forecast of ARMA is more accurate than that of FAR, and the value of the test statistic is positive (negative), the one-sided p-value is one minus a half of the two-sided p-value (one half of the two-sided p-value).

In addition, we divide the in-sample forecasts into two halves: the first half ranges from January 1995 to September 2006 and the second half ranges from October 2006 to December 2017. The two-sided testing results are given in Table 6. We first discuss the forecast of y_t. With the full sample, the MAE and RMSE of ARMA(5,5) model are almost the same as those of FAR(4,2) model for fitting y_t. The scenario is similar for the first period data and the second period data. This fact is in line with the DM test results, which shows that in most cases, the fitting accuracies of the two models are similar under the 5% significance level.

Then we consider the results for x_t. For the full sample, the MAE of ARMA(5,5) is 15.8% smaller than those of FAR(4,2) model for fitting x_t. This improvement increases to 58.19% with the first period data, which can be also verified by the larger absolute values of the DM statistic. The DM test result implies that the ARMA(5,5) model provides better fitting accuracies than the FAR(4,2) model. However, for the more stable second period data, the MAE of ARMA(5,5) is three times of that of FAR(4,2) model. This superiority can be also verified by the DM test. Therefore, the FAR(4,2) model has better fitting accuracy for more stable data. Similar conclusions can be obtained based on RMSE.

Combining the previous analysis, it is expected that the FAR model will provide more accurate prediction for x_t when China are stepping into the "New Normal" stage.

Table 6. Comparion of MAE and RMSE in ARMA(5,5) and FAR(4,2) model.

		$\hat{y}_t$				$\hat{x}_t$			
		FAR	ARMA	DM Test	*p*-Value	FAR	ARMA	DM Test	*p*-Value
Full sample	MAE	0.0039	0.0041	−1.2065	0.2287	3.0783	2.5918	−2.6366	0.0083
	RMSE	0.0052	0.0055	−1.9427	0.0531	3.9805	2.9115	−6.0828	0.0000
First half	MAE	0.0041	0.0041	−0.3154	0.7527	5.1302	2.1452	−38.4356	0.0000
	RMSE	0.0055	0.0052	−1.4703	0.1427	5.4395	2.5039	−27.1525	0.0000
Second half	MAE	0.0042	0.0039	−1.4839	0.1402	1.0337	3.0451	**11.3261**	0.0000
	RMSE	0.0056	0.0051	−1.6909	0.0932	1.4738	3.2722	**10.6661**	0.0000

"Full sample" means the loss function and the DM test are calculated based on the full sample period ranging from June 1995 to December 2017. "First half" and "Second half" means the results are calculated by using the first half samples and the second half samples, respectively. The calculated DM statistic (in bold) and the corresponding two-sided *p*-value are reported.

4.3. Out of Sample Forecast

From the in-sample analysis, the predictive performance of each model is not the same for periods before and in the stage of "New Normal". Compared to the in-sample fitting, the out-of-sample forecasting of a model is more important, since precise forecasts of the inflation rate are crucial for economic agents (e.g., investors, consumers) as well as for economic policy decision makers. In particular, we are interested in the inflation rate for the period of "New Normal". Thus, we divide the whole sample into two parts: the first part covers data from January 1995 to October 2013, and the second part covers data from November 2013 to December 2017. We use the rolling window method to obtain the future CPI value. That is, when the forecast proceeds, the estimation window rolls forward by adding one new data and dropping the most distant data. In this way, the size of the estimation window remains the same. The one-step-ahead forecast is implemented based on the estimated FAR(4,2) model (16) and the ARMA(5,5) model (17). To get an *h*-step-ahead forecast ($h > 1$), we can iteratively implement one-step-ahead forecast *h* times as given in (12). In the following, we consider different values of forecast horizons $h \in \{1, 3, 6, 9, 12, 15, 18, 21, 24\}$. The forecasting results are shown in Figure 7.

The figure shows that the forecasts based on the FAR(4,2) model are more smoothing, while the forecasts based on the ARMA(5,5) model are more fluctuated. Since the true CPI data in the forecasting period are stable, the FAR(4,2) model provide more accurate forecasts. It is also shown that as *h* increases, the forecasts for both models become more fluctuated and less accurate.

To evaluate the forecasting accuracy, we use the forecasting MAE and RMSE defined as

$$\text{MAE}^h = \frac{1}{N} \sum_{t=1}^{N} |x_t - \hat{x}_t^h|, \tag{20}$$

and

$$\text{RMSE}^h = \sqrt{\frac{\sum_{t=1}^{N}(x_t - \hat{x}_t^h)^2}{N}},\qquad(21)$$

respectively, where x_t is the true CPI, and $\hat{x}_t^h$ is the h-step-ahead forecast of CPI, N is the total number of forecasts. We also conduct the DM test for comparison. In Table 7, we report the MAE, the RMSE, the DM statistic and the p-value with alternative hypothesis that the two models have different predictive accuracies. The forecast horizon is fixed at $h \in \{1, 3, 6, 9, 12, 15, 18, 21\}$. It can be observed that at short horizon levels (i.e., $h \in \{1, 3\}$), the MAE and RMSE of the two models are comparable. When the horizon level becomes large (i.e., $h \in \{6, 9, 12, 15, 18, 21\}$), the MAE and RMSE of FAR(4,2) model are lower than those of ARMA(5,5) model. Moreover, as the horizon level h increases, the improvement of FAR(4,2) model becomes larger. This phenomenon is also detected by the DM test. Specifically, at 5% significance level, the FAR(4,2) model outperforms the ARMA(5,5) model for $h \in \{6, 9, 12, 15, 18, 21\}$, while the predictive accuracies of the two models are similar for $h \in \{1, 3\}$. This fact implies that the FAR model is better for moderate and long-term inflation rate forecasting and is comparable to ARMA model for short-term inflation rate forecasting.
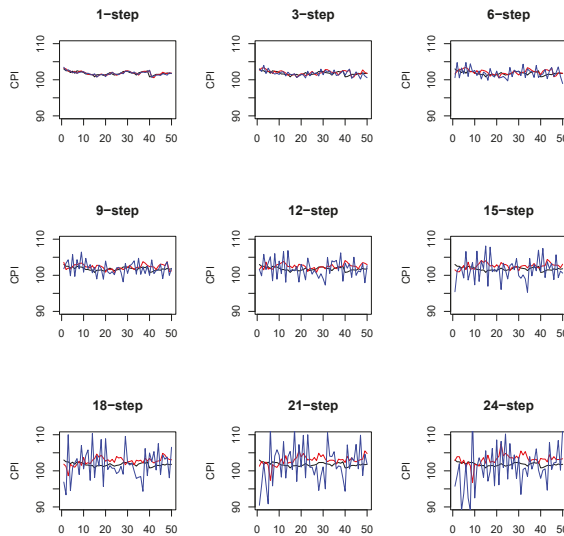
Figure 7. Multi-step-ahead forecasts of the monthly CPI data given by the FAR model and the ARMA model (black line: true CPI; blue line: ARMA; red line: FAR).

Table 7. MAE and RMSE of multi-step-ahead forecasts.

	h	1	3	6	9	12	15	18	21
MAE	ARMA(5,5)	0.3057	0.6052	1.1990	1.4782	2.1926	2.5480	3.3967	4.2165
	FAR(4,2)	0.2969	0.5278	0.6196	0.7560	0.9500	1.1295	1.3590	1.5832
	DM test	**0.3244**	**1.3658**	**4.9881**	**2.9283**	**4.9559**	**4.8971**	**3.8484**	**2.7157**
	p-value	0.7470	0.1782	0.0000	0.0052	0.0000	0.0000	0.0003	0.0091
MSE	ARMA(5,5)	0.4246	0.7447	1.4027	1.7951	2.5945	3.2504	4.2081	5.2963
	FAR(4,2)	0.4102	0.6656	0.8286	0.8883	1.2054	1.3406	1.6212	1.8738
	DM test	**0.6118**	**1.2482**	**3.4193**	**3.4778**	**5.3781**	**5.4008**	**2.7253**	**2.2040**
	p-value	0.5435	0.2179	0.0013	0.0011	0.0000	0.0000	0.0089	0.0323

The calculated DM statistic (in bold) and the corresponding two-sided p-value are reported.

5. Conclusions

The inflation rate is a critical quantity for both policy-makers and economic researchers. Hence, forecasting the inflation rate has long attracted the interests from various fields. In this paper, we apply the FAR model to forecast the inflation rate after the economy of China stepping into a new stage of "New Normal". The FAR model belongs to the semi-parametric non-linear time series model, which has three main advantages. First, compared to the traditional linear time series model, this model can describe the non-linear dynamics of the underlying process, which is common for many real time series data. Second, compared to the fully parametric non-linear time series model for example the threshold autoregressive model, the FAR model is more flexible and avoids the problem of model mis-specification. Third, the FAR model attains a satisfactory forecasting power, particularly for data without much fluctuations. The last advantage matches the characteristic of the inflation rate of China during the "New Normal" period. To estimate the functional coefficients of the FAR model, we adopt the B-spline method, which is numerically stable and can be obtained recursively. Thus, this estimation method largely reduces the computational burden and can be applied to handle huge amount of data.

Code for this illustration was written in R (R Development Core Team, 2010) and is available upon request from the authors.

Author Contributions: Q.L. and C.C. designed and performed the simualtion study and analyzed the data; M.W. and K.C. designed the empirical study, contributed analysis tools and wrote the paper.

References

1. Bampinas, G.; Panagiotidis, T. Hedging inflation with individual US stocks: A long-run portfolio analysis. *N. Am. J. Econ. Financ.* **2016**, *37*, 374–392. [CrossRef]
2. Stock, J.H.; Watson, M.W. Forecasting inflation. *J. Monet. Econ.* **1999**, *44*, 293–335. [CrossRef]
3. Atkeson, A.; Ohanian, L.E. Are Phillips Curve useful for forecasting inflation? *Fed. Reserv. Bank Minneap. Q. Rev.* **2001**, *25*, 2–11.
4. Mcadam, P.; Mcnelis, P. Forecasting inflation with thick models and neural networks. *Econ. Model.* **2005**, *22*, 846–867. [CrossRef]
5. Matheson, T.D. Phillips curve forecasting in a small open economy. *Econ. Lett.* **2006**, *98*, 161–166. [CrossRef]
6. Sekine, T. *Modeling and Forecasting Inflation in Japan*; IMF Working Paper No.01/82; International Monetary Fund: Washington, DC, USA, 2001.
7. Ramakrishman, U.; Vamvakidis. A. *Forecasting Inflation in Indonesia*; IMF Working Paper No.02/111; International Monetary Fund: Washington, DC, USA, 2002.
8. Ang, A.; Bekaert, G.; Wei, M. Do macro variables, asset markets, or surveys forecast inflation better? *J. Monet. Econ.* **2007**, *54*, 1163–1212. [CrossRef]
9. Lekkos, I.; Milas, C.; Panagiotidis, T. Forecasting interest rate swap spreads using domestic and international risk factors: Evidence from linear and non-linear models. *J. Forecast.* **2007**, *26*, 601–619. [CrossRef]
10. Milas, C.; Otero, J.; Panagiotidis, T. Forecasting the spot prices of various coffee types using linear and non-linear error correction models. *Int. J. Financ. Econ.* **2004**, *9*, 277–288. [CrossRef]
11. Bos, C.S.; Franses, P.H.; Ooms, M. Inflation forecast intervals and long memory regressions models. *Int. J. Forcast.* **2000**, *18*, 243–264. [CrossRef]
12. Tong, H. *Nonlinear Time Series: A Dynamical System Approach*; Oxford University Press: Oxford, UK, 1990.
13. Chen, S.X.; Lei, L.; T,Y. Functional coefficient moving average model with applications to forecasting chinese CPI. *Stat. Sin.* **2016**, *26*. 1649–1672. [CrossRef]
14. Granger, C.W.; Andersen, A. On the invertibility of time series models. *Stoch. Proc. Appl.* **1978**, *8*, 87–92. [CrossRef]

15. Haggan, V.; Ozaki, T. Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika* **1981**, *68*, 189–196. [CrossRef]
16. Fan, J.Q.; Yao, Q.W. *Nonlinear Time Series: Nonparametric and Parametric Methods*; Springer: New York, NY, USA, 2003.
17. Chen, R.; Tsay, R.S. Functional coefficient autoregressive model. *J. Am. Stat. Assoc.* **1993**, *88*, 298–308.
18. Hastie, T.; Tibshirani, R. Varying-coefficient models. *J. R. Stat. Soc. B* **1993**, *55*, 757–796.
19. Chen, B.; Hong, Y. Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica* **2012**, *80*, 1157–1183.
20. Cai, Z.W.; Fan, J.Q.; Yao, Q.W. Functional-coefficient regression models for nonlinear time series. *J. Am. Stat. Assoc.* **2000**, *95*, 941–956. [CrossRef]
21. Chen, R.; Liu, L.M. Functional coefficient autoregressive models: Estimation and tests of hypotheses. *J. Time Ser. Anal.* **2001**, *22*, 151–173. [CrossRef]
22. Huang, J.Z.; Shen, H.P. Functional coefficient regression models for non-linear time series: A polynomial spline approach. *Scand. J. Stat.* **2004**, *31*, 515–534. [CrossRef]
23. De Boor, C. On Calculating with B-Spline. *J. Approx. Theory* **1972**, *6*, 50–62. [CrossRef]
24. Schumaker, L. *Spline Functions: Basic Theory*; Cambridge University Press: Cambridge, UK, 2007.
25. Harvill, J.L.; Ray, B.K. A note on multi-step forecasting with functional coefficient autoregressive models. *Int. J. Forecast.* **2005**, *21*, 717–727. [CrossRef]
26. Xie, S.; Zhou, Y.; Wan, A.T.K. A Varying-Coefficient Expectile Model for Estimating Value at Risk. *J. Bus. Econ. Stat.* **2014**, *32*, 576–592. [CrossRef]
27. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
28. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
29. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263.

Article

Financial Hazard Map: Financial Vulnerability Predicted by a Random Forests Classification Model

Katsuyuki Tanaka, Takuji Kinkyo and Shigeyuki Hamori *

Graduate School of Economics, Kobe University, 2-1, Rokkodai, Nada-Ku, Kobe 657-8501, Japan;
katsutanaka@econ.kobe-u.ac.jp (K.T.); kinkyo@econ.kobe-u.ac.jp (T.K.)
* Correspondence: hamori@econ.kobe-u.ac.jp; Tel.: +81-78-803-6832

Abstract: This study develops a systematic framework for assessing a country's financial vulnerability using a predictive classification model of random forests. We introduce a new indicator that quantifies the potential loss in bank assets and measures a country's overall vulnerability by aggregating these indicators across the banking sector. We also visualize the degree of vulnerability by creating a Financial Hazard Map that highlights countries and regions with underlying risks in their banking sectors.

1. Introduction

The severe economic consequences of the global financial crisis of 2008–2009 highlighted the importance of crisis prevention and sparked a renewed interest in early warning systems (EWSs). An EWS aims to detect potential vulnerabilities in a financial system that could trigger a system-wide crisis. A reliable EWS provides useful guidance for policy-makers to activate macro-prudential policy in an effective and timely manner. The International Monetary Fund (IMF) [1] and the Committee on the Global Financial System [2] provide a comprehensive discussion of the operational aspects of the macro-prudential policy. After Frankel and Rose [3] and Kaminsky et al.'s [4] early contributions, researchers have made considerable efforts to develop a consistently useful EWS for various types of crises.

Kaminsky et al. [4] proposed the popular signaling approach, which Alessi and Detken [5] recently used. This approach seeks to identify the threshold values for individual indicators that signal crises, and thus trigger an early warning when the pre-defined threshold for the pre-selected indicator is breached. A popular indicator common in these studies is the credit-to-GDP (Gross Domestic Product) ratio, which is a key indicator signaling credit booms. However, this signaling approach has a shortcoming given that, as a univariate approach, the decision would rely on only a single factor, which can send a misleading signal.

Another conventional approach from the EWS literature is estimating the multivariate probit and logistic regressions, which relate the probability of a crisis to a set of explanatory variables, such as current account balance, real exchange rates, credit growth, and fiscal balance [6–10]. Despite its popularity, the conventional approach has certain limitations. For one, researchers must pre-select explanatory variables from a wide range of economic indicators based on some prior information. For another, the logistic regression does not readily allow for non-linear or threshold effects of explanatory variables. More generally, linear regressions often perform poorly in terms of prediction performance relative to newer machine learning models [11]. Linear regressions may work well for small datasets but they are not readily scalable to larger datasets.

Ghosh and Ghosh [12] and Frankel and Wei [13] employed a decision tree method that uses a sequence of splitting rules to segment the space of explanatory variables. Hastie et al. [11] and

James et al. [14] provide details of tree methods, including decision trees and random forests. At each node of a tree, the sample is split into two sub-branches according to the threshold value of an explanatory variable. For classification trees, either the Gini index or the cross entropy is used to evaluate the quality of a split. A smaller value of these indices indicates that the node is purer, and thus contains more observations from a single class. The process is repeated until a stopping criterion is reached, such as the minimum number of observations at each node. Each terminal node at the bottom of the tree provides a class prediction for a given observation. Whereas linear logistic regressions models require a handcrafted selection of explanatory variables to obtain reasonable early warning performance, the decision tree systematically learns important variables, performs better in early warning, and allows for non-linear effects. Although a decision tree is simple and provides explanatory and intuitive decision rules, it suffers from high variance (i.e., a small change in the data can cause a large change in the financial tree), so is likely to suffer over-fitting problems. This is largely owing to the fact that the values of the thresholds depend heavily on the values of the training observations.

With an increased opportunity to gain access to larger datasets, exploring the significant scope for economic modeling and analysis for a more flexible approach has become popular with data scientists [15,16]. In this study, we take advantage of the advancements in predictive modeling techniques of machine learning to build an EWS, and develop a systematic framework to assess and visualize a country's financial vulnerability. The main contributions of our study are three-fold. First, our study differs from previous ones in that we used a novel machine-learning technique known as random forests to construct an EWS to predict bank failures (random forests EWS). Random forests are a variant of decision trees that significantly improve prediction accuracy by combining a large number of trees using random input selection [17]. Second, we introduce a new indicator that quantifies the expected potential loss in bank assets computed using the prediction of the random forests EWS. To assess a country's overall financial vulnerability, we aggregate individual banks' expected potential asset losses across the domestic banking sector. Finally, we visualize the degree of a country's financial vulnerability by creating a Financial Hazard Map that highlights countries and regions with significant risks in their underlying banking sectors. Our work is similar to that of Tanaka et al. [18], but differs by a few points. Our paper provides a financial analysis of the finance sector, whereas the interest of Tanaka et al. focused on the industrial sector. Furthermore, we propose a novel indicator to assess the overall financial vulnerability of each country.

We chose random forests (RF) for three reasons. First, RF can significantly improve prediction accuracy by building a large number of decision trees on bootstrapped training samples—a technique known as ensemble learning. Random forests also circumvent the over-fitting problem by adding randomness to the tree building process, and thus reducing correlations among trees; hence, it performs well with out-of-sample data. Second, random forests can better handle a large dataset as multiple trees can be trained in parallel efficiently with a very simple hyper-parameter setting. The model can be built by merely setting the number of trees. Finally, RF provide the importance measurement, which can be used for certain levels of causality inference. Whereas various application areas use random forests, including computer vision and bioinformatics, its application to economics remains limited. Tanaka et al. [19] used random forests to predict bank failure in OECD member countries.

Another important feature of our study is the use of bank-level financial statements to predict bank failure using the random forests EWS built from a large dataset of more than 15,000 banks globally. As previous studies typically used macroeconomic indicators to predict currency and financial crises, the recent literature indicates that the state of bank financial statements can explain differences in performance across banks during financial crises [20,21]. Moreover, previous studies often defined a crisis as an event in which the values of preselected indicators exceed predetermined thresholds. Consequently, the prediction performance significantly depends on the choice of threshold. We define the event of a bank failure as the change in a bank's status from active to inactive (i.e., bankrupt, in liquidation, or dissolved) based on the information provided by the Bureau Van Dijk Bankscope. By doing so, we wanted to minimize arbitrariness, and thus reduce the possible bias in prediction.

The remainder of the paper proceeds as follows. Section 2 describes the methodology and data of building the random forests EWS. Section 3 introduces a new indicator that quantifies the expected potential losses in bank assets. We present the assessment of a country's financial vulnerability and visualize it by creating a Financial Hazard Map. Section 4 provides our conclusions.

2. Materials and Methods

In this study, we considered the task of building an EWS as a classification problem to identify a bank's status (i.e., active or inactive) based on the underlying financial conditions. Drawing on insights from the extensive literature on corporate bankruptcy predictions, we used information about individual banks' financial statements as predictors to build models. Altman [22] provides an early contribution to the literature. In contrast to existing studies that are more concerned with identifying the key predictors of bankruptcy, we prioritized improving the prediction accuracy. To this end, we used random forests that tend to perform better in terms of prediction accuracy than conventional methods, such as logistic regressions, which have been widely used in previous studies.

2.1. Major Features of Random Forests

Random forests are a variant of decision trees, which overcome the over-fitting problem by building multiple trees and combing the results of these trees [17], effectively forming forests. Each tree in a random forest is built using randomly selected data samples and/or randomly selected input variables from the original data to split each node. After generating a large number of trees, the model votes for the most popular class. A single-tree classifier tends to have only marginally better accuracy than a random choice of class. However, by combining a large number of trees using random input selection, random forests can produce a powerful model.

Breiman et al. [17] constructed such trees using the Gini index criterion, which measures the best split criterion based on the impurity of each node. The algorithm aims to select the optimal splitting variable and the corresponding threshold value by making each node as pure as possible. Suppose M_n is the number of pieces of information reaching node n and M_n^i is the number of data points belonging to class C_i, the Gini index, GI_n, of node n is obtained using Equation (1):

$$GI_n = 1 - \sum_{i=1}^{K} \left(p_n^i \right)^2, \text{ where } p_n^i = \frac{M_n^i}{M_n} \tag{1}$$

A smaller Gini index value for node n represents greater purity, which implies that the node contains more observations from a single class. Hence, a decreasing Gini index is an important criterion when splitting a node.

In comparison to single-tree modeling, random forests have several desirable features [17,23]. First, random forests perform better in terms of classification accuracy by building a large number of trees instead of only a single tree. Each tree is built using randomly selected data samples and randomly selected input variables from the original data to split each node. After generating a large number of trees, they vote for the most popular class. A single-tree classifier tends to have only marginally better accuracy than a random choice of class. However, by combining a large number of trees using random input selection, random forests can improve accuracy. Second, random forests provide better generalization abilities and are robust to over-fitting. Hence, RF may have better out-of-sample accuracy when using a random selection of input variables to split each node and combining the results of multiple trees yields error rates that compare favorably to alternative methods and are more robust with respect to noise. Third, random forests can better handle large datasets as multiple trees can be efficiently trained in parallel. Finally, random forests provide a measure for the relative contribution of each variable to generate a prediction. These variable importance measures help identify the variables that are important for distinguishing between active and inactive banks, and thus for predicting bank failure.

2.2. Data

We sourced our data for bank financial statement indicators from Bankscope. The advantage of using this data source is that it provides a broad coverage of banks with standardized data formats across countries. We used 48 indicators derived from the Summary Analytics category classified into four groups: profitability ratio, capitalization, loan quality, and funding (Appendix B). Our sample included 23,455 commercial banks, saving banks, and cooperatives incorporated in 198 countries and regions. The training set included annual observations of the latest available financial statements for each bank up to 2014. We defined a bank failure event as the change in a bank's status from active to inactive (i.e., bankrupt, in liquidation, or dissolved) as reported by Bankscope. We assumed that the latest available financial statements for active banks had sound financial status and inactive banks had unsound financial status. We then systematically identified patterns distinguishing the differences by random forests.

As there were fewer inactive banks (7294 banks), we selected the largest 7294 active banks in terms of total assets to match the number of inactive banks. We also selected the smallest 7294 active banks to build a more flexible model to prevent a bias toward larger banks. To avoid model bias created by an imbalanced training set, we evened out the sample sizes of active and inactive banks by doubling the sample size of inactive banks by duplicating each observation. In addition, we eliminated variables if more than 50% of its values were missing (7294 × 2 biggest and smallest active banks + 7294 inactive banks). Thus, we eliminated 6 variables and used 42 variables for experiments. We used the random forests and caret packages in the R software package to train and evaluate our models. Figure 1 illustrates the model building process for the random forests EWS. Appendix C reports the classification accuracy of the random forests EWS.
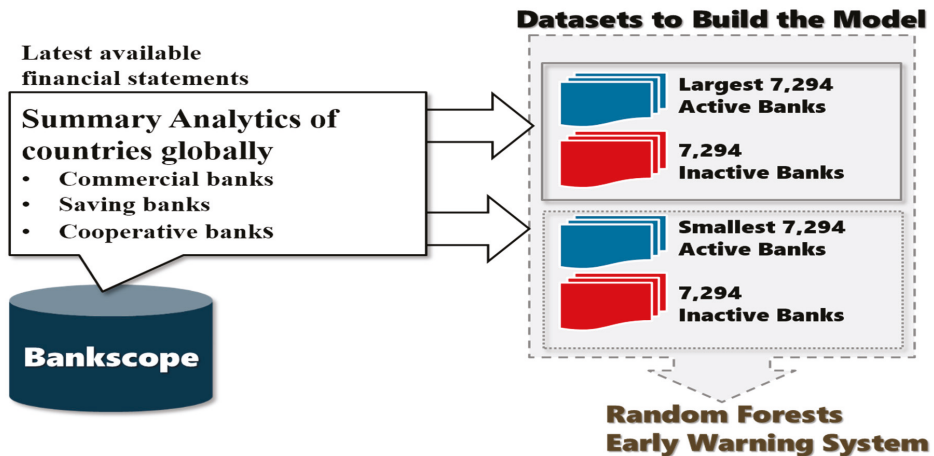

Figure 1. Model building process of random forests early warning system (EWS).

3. Results

3.1. Variable Importance Measures

A useful property of random forests is that it provides variable importance measures that help identify the most important variables for distinguishing between active and inactive banks. Hence, RF should provide some clues to the underlying causes of bank failures. For classification trees, we obtained the variable importance measures from each variable's contribution to the reduction in the Gini index. The Gini index is a common measure of the degree of inequality in income distribution. The smaller the value of the index, the more equal the society.

In our random forests algorithm, the Gini index is the measure for the purity of each node. A smaller value of the Gini index represents a purer node, which implies that the node contains more observations from a single class. The goal of the algorithm was to make each node as pure as possible by selecting the optimal splitting variable and the corresponding threshold value. Therefore, we calculated variable importance by summing the total reduction in the Gini index by splits over a given variable, averaged over all bagged trees.

Figure 2 illustrates the variable importance measures as the mean decrease in the Gini index for each variable. Considering this model, we identified the following indicators as the top four predictors: interest expense/average interest-bearing liabilities, interest income on loan/average gross loans, interest expense on customer deposits/average customer deposits, and interest income/average earning assets. The importance measure for the first indicator was by far the largest. These top four indicators fall into the profitability ratio category. In contrast, the importance measures for the other categories of indicators, that is, capitalization, loan quality, and liquidity are much smaller. The results indicate that bank profitability has the most important impact on the probability of bank failure.
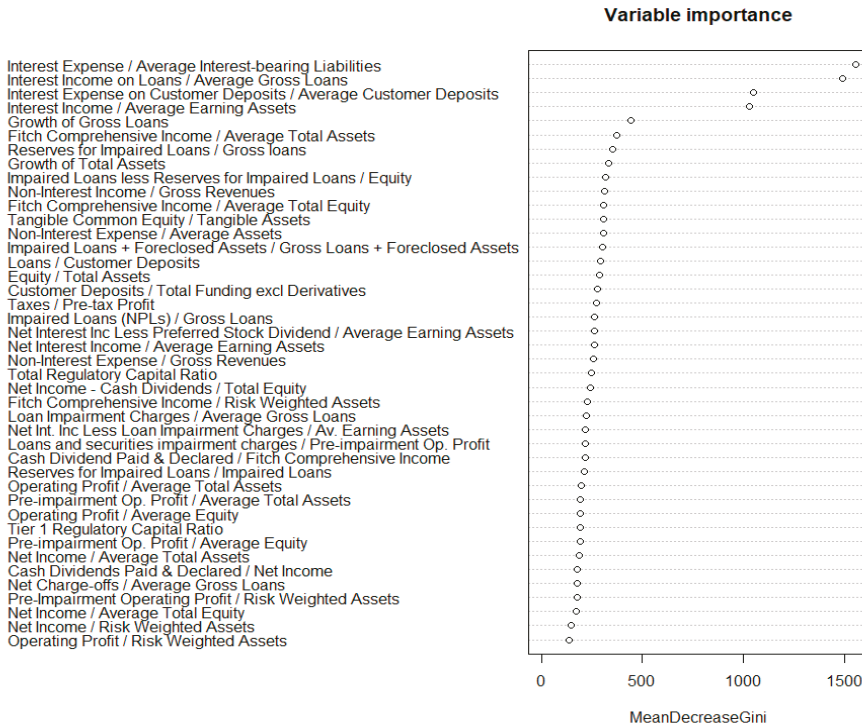

Figure 2. Variable importance measures of random forests.

We show the experimental result of a single decision tree in Figure 3 for comparison. Though a single tree selects similar criteria to distinguish between active and inactive banks, it does not perform as well as random forests. This is due to the fact that the random forests model produces a more flexible model as it produces multiple trees to analyze different patterns in the data, whereas a single tree only produces one set of rules for a classification decision.
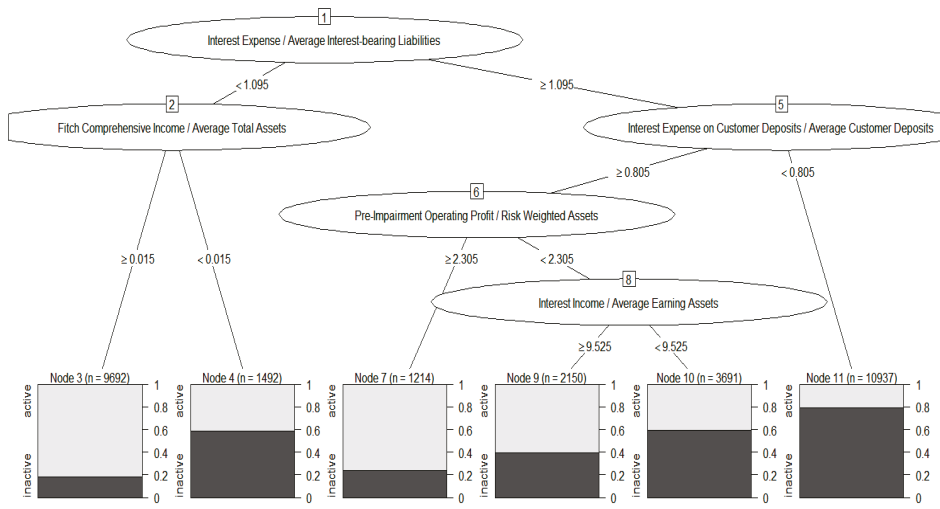
Figure 3. Plot of a tree model.

3.2. New Indicator for the Expected Potential Asset Loss

We introduce a new indicator to assess the degree of financial vulnerability using the prediction of the random forests EWS. We used the 2014 financial statement data to predict the probability of bank failure. We define the expected potential asset loss of a bank as follows:

$$\text{EPAL}_{i,j} = P_{i,j} \times \text{Total Assets}_{i,j}, \tag{2}$$

where $\text{EPAL}_{i,j}$ denotes the expected potential loss in bank i in country j, $P_{i,j}$ denotes the probability of failure for bank i in country j given by the random forests EWS prediction, and Total Assets$_{i,j}$ denotes the value of total assets of bank i in country j. To measure a country's overall financial vulnerability, we aggregate the value of $\text{EPAL}_{i,j}$ across the domestic banking sector. Given that we used consolidated financial statement data, all the expected potential loss of multinational banks was counted as losses in the country where the headquarters of these banks were located. We acknowledge that this is the limitation of our work and consider overcoming this limitation as our future task. Hence, the country-level expected potential loss in the domestic banking sector denoted by EPAL_j is given by:

$$\text{EPAL}_j = \sum_i \text{EPAL}_{i,j}. \tag{3}$$

To gauge the impact of the expected potential asset loss on the domestic banking sector and economic activities, we calculated the share of EPAL_j in the total assets of the domestic banking sector and in nominal GDP. Table 1 summarizes the results.

The left column of the table ranks 50 countries in terms of their share in banking sector assets. The ranking indicates that Suriname, Grenada, Denmark, Gabo, and Guatemala are the five most vulnerable countries in the sense that the impact of the expected potential loss on the domestic banking sector can be relatively large. Thus, these countries have a relatively high risk of a system-wide banking crisis.

The right column of the table ranks countries in terms of the share in nominal GDP. The ranking indicates that the Palestinian Territories, Luxembourg, Cyprus, Denmark, and France are the five most vulnerable countries in the sense that the impact of the expected potential asset loss on domestic

economic activities could be relatively large. This is particularly the case if the assets of banks with high probability of failure consist primarily of domestic loans and investments.

Table 1. Top 50 countries in terms of the shares of expected potential asset loss based on 2014 financial statements.

Share of Expected Potential Asset Loss in Domestic Banking Sector		Share of Expected Potential Asset Loss in Nominal GDP	
SURINAME	37.92%	PALESTINIAN TERRITORIES	137.53%
GRENADA	31.00%	LUXEMBOURG	135.54%
DENMARK	28.14%	CYPRUS	100.92%
GABON	28.06%	DENMARK	81.80%
GUATEMALA	27.33%	FRANCE	71.26%
VENEZUELA	25.65%	VENEZUELA	67.12%
SENEGAL	25.30%	PORTUGAL	58.03%
NEPAL	24.41%	LEBANON	56.36%
UZBEKISTAN	24.41%	JORDAN	56.25%
KYRGYZSTAN	24.31%	BAHRAIN	52.90%
CAMEROON	24.04%	SPAIN	50.81%
LESOTHO	23.54%	MAURITIUS	50.56%
EL SALVADOR	23.25%	UNITED KINGDOM	41.92%
DOMINICA	23.00%	SWITZERLAND	39.83%
ROMANIA	22.44%	AUSTRIA	38.16%
THAILAND	21.48%	HONG KONG	37.43%
HUNGARY	20.99%	GERMANY	37.28%
MONTENEGRO	20.74%	NETHERLANDS	36.94%
ETHIOPIA	20.51%	BAHAMAS	34.16%
ARGENTINA	20.38%	SAINT KITTS AND NEVIS	32.75%
LIBERIA	19.95%	THAILAND	31.86%
TUNISIA	19.81%	BELGIUM	31.19%
LIECHTENSTEIN	19.72%	FINLAND	30.61%
BOSNIA AND HERZEGOVINA	19.66%	SAN MARINO	30.59%
COTE D'IVOIRE	19.01%	NEW ZEALAND	28.96%
PARAGUAY	18.75%	PANAMA	28.10%
SERBIA	18.50%	GRENADA	27.02%
MADAGASCAR	18.22%	ITALY	25.76%
BOLIVIA	18.17%	GREECE	23.50%
PORTUGAL	18.17%	MOROCCO	21.80%
REPUBLIC OF KOREA	18.09%	AUSTRALIA	21.56%
SAINT KITTS AND NEVIS	17.99%	CHILE	20.94%
HONDURAS	17.96%	IRELAND	20.55%
BELIZE	17.81%	NEPAL	19.99%
PERU	17.30%	CROATIA	19.43%
AUSTRIA	17.30%	ICELAND	18.82%
JORDAN	17.27%	CAPE VERDE	18.66%
MAURITIUS	17.13%	GUATEMALA	17.61%
UKRAINE	17.05%	REPUBLIC OF KOREA	17.19%
CAPE VERDE	17.03%	EL SALVADOR	17.05%
TURKMENISTAN	16.96%	ANTIGUA AND BARBUDA	17.03%
NEW ZEALAND	16.70%	SWEDEN	16.97%
MALI	16.15%	CANADA	16.74%
CROATIA	15.97%	HUNGARY	16.69%
CHILE	15.60%	DOMINICA	16.46%
BRUNEI DARUSSALAM	15.33%	HONDURAS	16.04%
ARMENIA	15.30%	BARBADOS	15.89%
DOMINICAN REPUBLIC	15.04%	SURINAME	15.52%
BAHRAIN	14.89%	VIETNAM	15.34%
ECUADOR	14.84%	TUNISIA	15.23%

Interestingly, many Organisation for Economic Co-operation and Development (OECD) countries, including European countries, at the top of the list. This may indicate that these countries have not recovered fully from the major financial crises, notably, the global financial crisis of 2008–2009 and the European debt crisis of 2010–2013, or new financial risks may be looming. Given the relatively large size of their domestic banking sectors, these countries can be the epicenter of cross-border financial spillovers by withdrawing overseas loans and investments in the face of financial difficulties.

In Figure 4, we create a scatter plot of the vulnerability measures reported in Table 1. The combination of higher values of these measures in a particular country implies greater financial vulnerability. The figure clearly indicates that Denmark and Venezuela stand out in terms of both measures, signaling significant risks. The bold horizontal and vertical lines indicate the medians of these measures; the shadows indicate the first quartiles. The medians of the share of $EPAL_j$ in the banking sector assets and nominal GDP are 11.60% and 7.85%, respectively. The level of these

medians indicates the overall vulnerability of the global banking sector, with a significant increase in signaling financial risks. Notably, our vulnerability measure raises a red flag for potential trouble, but it does not identify the causes or the likely outcomes of the trouble. However, we believe that these measures are useful for spotting vulnerabilities, and thus encouraging regulators and investors to take preemptive actions.
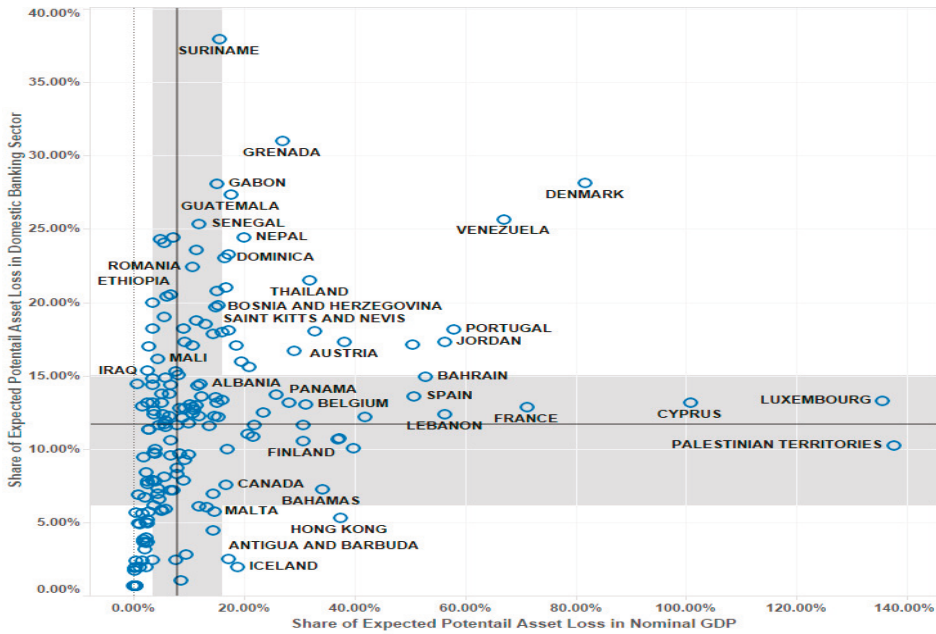

Figure 4. Scatter plot of country vulnerabilities.

In Appendix A, we summarize the predicted bank failures for each country for 2014. The table shows the number of banks and the sum of assets held by the banks for each category of predicted probability of failures with a 10-percentage-point interval.

3.3. Financial Hazard Map

Finally, we visualized the degree of a country's financial vulnerability by creating a Financial Hazard Map. Corresponding to each definition of vulnerability in Table 1, we present two types of maps. Figure 5 shows the share of $EPAL_j$ in the assets of domestic banking sectors. The areas that are darker red indicate a higher degree of vulnerability in terms of the impact of the expected potential asset loss on the domestic banking sector. Figure 6 shows the share of $EPAL_j$ in nominal GDP. Darker red areas indicate a higher degree of vulnerability in terms of the impact of the expected potential asset loss on domestic economic activities.

The Financial Hazard Map highlights countries and regions with significant vulnerability in their underlying banking sector. The darker red areas correspond to the top 50 countries listed in Table 1. The map provides a clear and understandable assessment of financial vulnerability in particular countries and regions. It also shows the geographical distribution of financial risk and the danger of potential contagion for neighbors of high vulnerability areas.
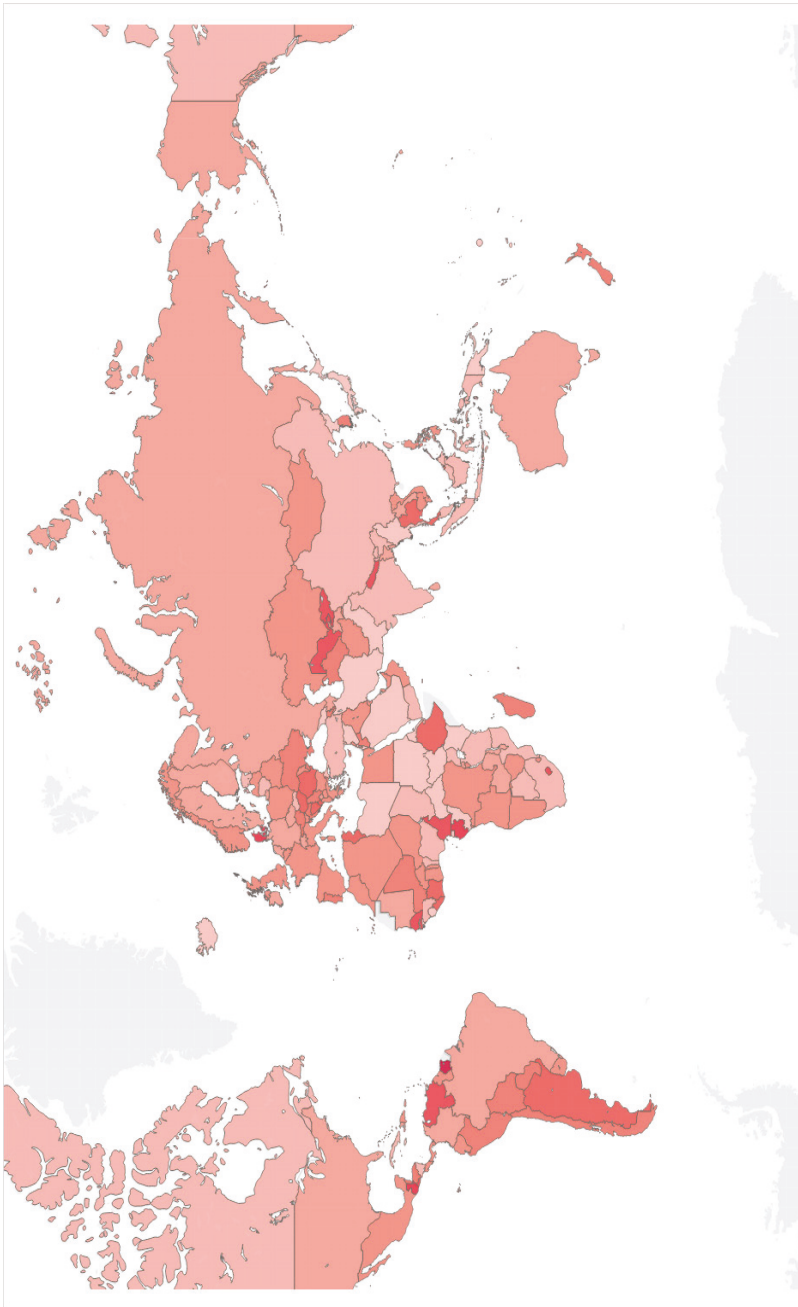
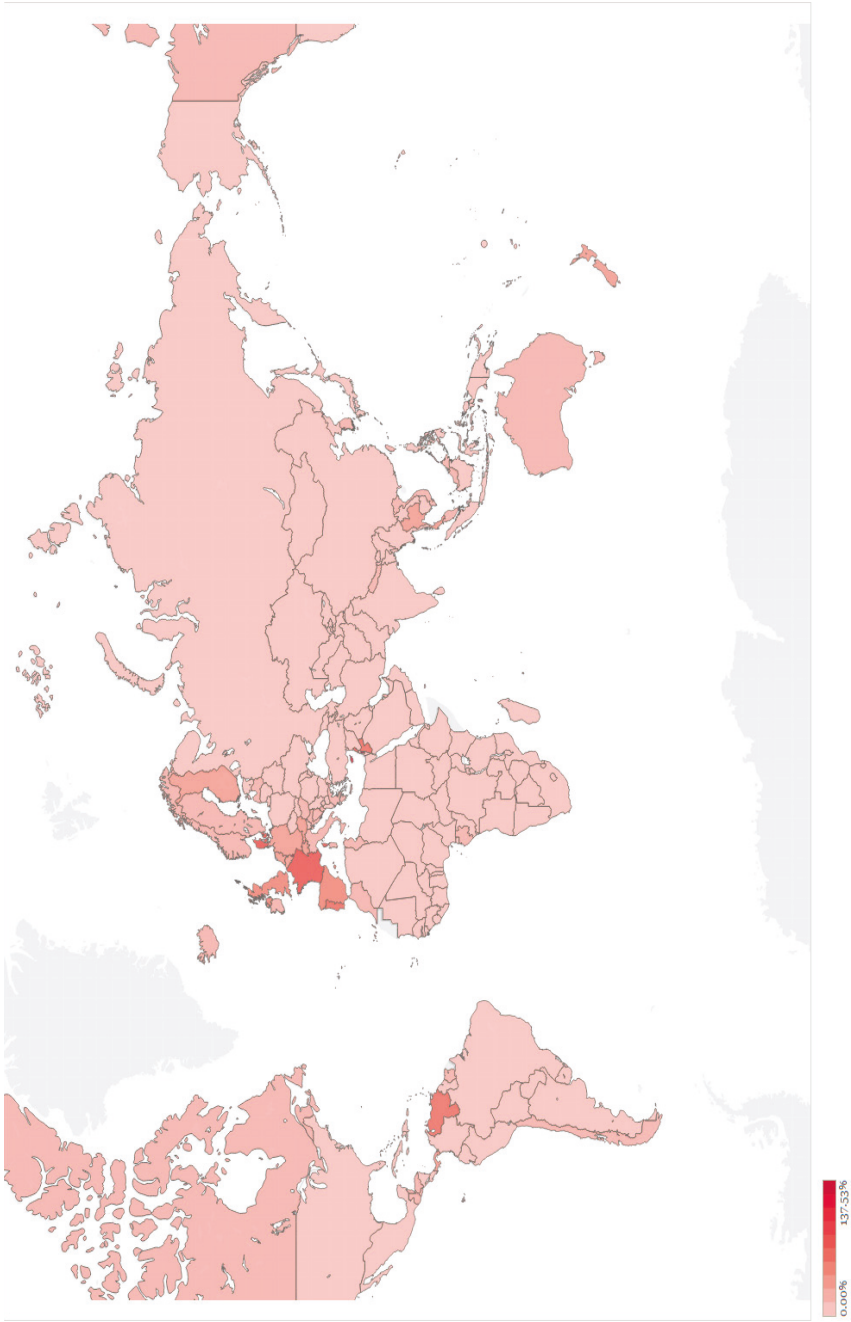**Figure 5.** Impact of expected potential asset loss on the banking sector.

Figure 6. Impact of expected potential asset loss on economic activities.

4. Conclusions

We developed a systematic framework for assessing and visualizing a country's financial vulnerability. We employed a novel machine-learning approach known as random forests to construct an EWS to predict bank failures and introduced a new indicator that quantifies the expected potential loss in bank assets computed based on the random forests EWS prediction. We assessed the financial vulnerability of each country by aggregating individual banks' indicators across the banking sector. To gauge the impact of expected potential asset loss, we calculated the shares in the banking sector assets and nominal GDP. We identified countries and regions with high vulnerability in terms of these shares. Furthermore, we visualized the degree of a country's financial vulnerability by creating a Financial Hazard Map. We demonstrated the usefulness of the Financial Hazard Map in spotting vulnerable countries and regions and understanding the geographical distribution of risk.

We hope that the Financial Hazard Map will prove useful for both policy-makers and private investors in detecting potential risk, and thereby prompting precautionary actions. Our framework of assessing financial vulnerability is simple, and therefore readily applicable to other types of risk analysis. A future task may be to develop a dynamic framework that allows for an assessment of contagion risks between banks and countries potentially in trouble, taking account of country-specific macroeconomic and institutional factors.

Author Contributions: S.H. and T.K. conceived and designed the experiments; K.T. performed the experiments; S.H., T.K. and K.T. analyzed the data; K.T. contributed reagents/materials/analysis tools; K.T. wrote the paper.

Appendix A

Numbers and assets of banks for each category of bank failure probabilities.

Country	NB 0–10%	NB 10–20%	NB 20–30%	NB 30–40%	NB 40–50%	NB 50–60%	NB 60–70%	NB 70–80%	NB 80–90%	NB Total	TA 0–10%	TA 10–20%	TA 20–30%	TA 30–40%	TA 40–50%	TA 50–60%	TA 60–70%	TA 70–80%	TA 80–90%	TA Total	Nominal GDP (billions of US dollar)
AFGHANISTAN	3	2	2	3	1					11	1,832	1,485	467	572	341					4,697	20.44
ALBANIA	2	8	2		1					13	1,361	8,371	932		385					11,049	13.30
ALGERIA	5	4	6	1	1					17	61,900	6,990	43,402	278	413					112,983	213.52
ANDORRA	3	1								4	11,106	1,566								12,672	
ANGOLA	8	2	4	1						15	31,675	8,428	12,464	1,965						54,532	126.78
ANGUILLA	2									2	644									644	
ANTIGUA AND BARBUDA	8		1	1						10	8,012		113	416						8,541	1.25
ARGENTINA	23	25	12	5						65	15,688	70,364	68,799	6,694						161,545	544.73
ARMENIA	8	3	2	1	2					16	3,531	731	716	206	671					5,855	11.64
ARUBA	1	1								2	580	1,037								1,617	
AUSTRALIA	16	7	14	1	1	1				40	1,309,679	1,357,747	127,694	16,161	1,749	47,393				2,860,423	1441.95
AUSTRIA	184	43	29	21	11	1				289	283,139	239,865	344,623	58,462	35,246	3,782				965,117	437.58
AZERBAIJAN	21	5	1		1					28	10,969	14,605	1,623		367					27,564	75.25
BAHAMAS	23	1	3	3						30	31,715	160	3,656	4,655						40,186	8.51
BAHRAIN	4	5		2	1					12	14,232	104,028		729	1,248					120,237	33.84
BANGLADESH	24	13		3						40	59,439	25,723		3,566						88,728	183.82
BARBADOS	2	1	1	2						6	3,020	257	200	1,713						5,190	4.35
BELARUS	17	2	7							26	10,101	15,299	10,705							36,105	76.14
BELGIUM	19	9	9	1	4					42	652,514	329,070	273,204	3,069	15,149					1,273,006	532.39
BELIZE	4	1			1					6	621	386			358					1,365	1.70
BENIN	8	1								9	2,765	1,431								4,196	9.59
BERMUDA	2		2							4	21,317		1,775							23,092	
BHUTAN	3									3	1,202									1,202	1.99
BOLIVIA	1	8	2	1						12	53	12,050	3,128	1,382						16,613	33.24
BOSNIA AND HERZEGOVINA	4	11	4	2		5				26	1,831	6,158	3,118	1,225						13,988	18.52
BOTSWANA	8	1	2							11	5,830	418	1,168							7,416	15.88
BRAZIL	83	36	17	5						141	1,376,576	599,875	29,506	74,468						2,080,425	2417.16
BRUNEI DARUSSALAM	1									1	2,706									2,706	17.10
BULGARIA	10	10	2	1						23	22,808	21,205	1,468	2,729						48,210	56.72
BURKINA FASO	5	3	1							9	1,158	3,029	611							4,798	12.48
BURUNDI	4	1	1							6	259	224	187							670	2.90
CAMBODIA	5	11	4	1	2					23	4,297	7,969	873	99	332					13,923	16.78
CAMEROON	5	1	3	3						12	959	433	2,957	2,936						7,285	31.63
CANADA	53	18	10	3		1				85	2,846,839	1,051,419	39,746	3,701		353				3,941,715	1783.78
CAPE VERDE	1	2	2		1					6	259	998	788		10					2,045	1.87
CAYMAN ISLANDS	18	2								20	22,196	12,376								34,572	
CENTRAL AFRICAN REPUBLIC	2									2	258									258	1.73
CHAD	4	1								5	814	319								1,133	13.94
CHILE	11	6	10	6	2					35	73,641	192,501	54,875	5,914	20,391					347,322	258.68
CHINA	111	34	13	29	3					190	20,605,051	725,637	217,344	1,504,762	25,481					23,078,275	10430.71
COLOMBIA	10	10	4	3						27	123,640	94,505	15,177	551						233,873	377.87
COMOROS	1									1	0									0	0.70
CONGO	5									5	3,147									3,147	13.55
COSTA RICA	25	13	5	3						46	17,819	28,841	1,065	213						47,938	49.60
COTE D'IVOIRE	7	3	6	1						17	1,640	2,658	5,176	376						9,850	33.74
CROATIA	8	10	11	3	1					33	5,984	46,496	15,954	749	366					69,549	57.17
CUBA	5	1								6	10,490	5,073								15,563	

Country	Number of Banks 0%-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	Country Total	Total Asset (millions of US dollar) 0%-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	Country Total	Nominal GDP (billions of US dollar)
CURACAO	13	1		2						16	11,647	3,703		1,463						16,813	
CYPRUS	11	6	3	1						21	60,348	62,969	53,753	529						177,599	23.11
CZECH REPUBLIC	9	8	2	1						20	93,674	107,793	3,958	271						205,696	205.27
DEMOCRATIC PEOPLE'S REPUB	1									1	8,955									8,955	
DEMOCRATIC REPUBLIC OF CON	5	8	1							14	913	3,109	131							4,153	35.92
DENMARK	34	25	10	8						77	55,066	99,672	134,603	716,900						1,006,241	346.12
DJIBOUTI	4	1								5	972	410								1,382	1.59
DOMINICA			1							1			375							375	0.52
DOMINICAN REPUBLIC	28	17	14	2						61	10,873	16,973	5,942	119						33,907	64.06
ECUADOR	6	9	4	2						21	6,285	24,036	5,940	1,936						38,197	100.92
EGYPT	16	4	1	2						23	155,170	11,839	1,634	47,083						215,726	301.39
EL SALVADOR	2	5	3	2	2					14	836	9,680	2,651	2,702	2,585					18,454	25.16
EQUATORIAL GUINEA	3									3	1,763									1,763	15.53
ERITREA	2									2	1,854									1,854	4.05
ESTONIA	5	1	1	1						8	22,246	233	311	311						23,101	26.51
ETHIOPIA	1	7	1	1						10	207	4,923	12,424	399						17,953	55.51
FEDERATED STATES OF MICRO	1									1	129									129	0.32
FIJI	1									1	405									405	4.53
FINLAND	38	3	2	2						45	574,425	152,017	349	65,636						792,427	272.77
FRANCE	128	51	28	18	2	3				230	6,953,615	4,961,192	3,770,293	51,662	785	1,484				15,739,031	2833.69
GABON	1	2	2	1	1					7	131	1,489	5,965	270	1,928					9,783	18.21
GAMBIA	7	1								8	346	104								450	0.82
GEORGIA	14	1								15	10,845	270								11,115	16.52
GERMANY	1,280	239	66	31	12					1,628	8,755,293	1,235,557	3,092,779	309,316	65,984					13,458,929	3874.44
GHANA	14	6	2	2						24	5,258	3,647	1,117	1,322						11,344	38.62
GIBRALTAR	2									2	5,474									5,474	
GREECE	4	5								9	10,781	433,608								444,389	235.95
GRENADA	1			1	1					3	204			278	313					795	0.91
GUATEMALA	8	2		10	3	1				24	11,344	799		22,113	3,469	173				37,898	58.83
GUINEA	6	1								7	653	410								1,063	6.70
GUINEA BISSAU	1									1	114									114	1.11
GUYANA	2		1							3	1,152		270							1,422	3.08
HAITI	4	1								5	3,655	109								3,764	8.71
HONDURAS	7	6	3	1		1	1			19	7,548	2,522	5,150	1,526		372	306			17,424	19.51
HONG KONG	27	5	3							35	1,895,556	158,181	3,129							2,056,866	291.23
HUNGARY	8	5	4	5	4	1				27	3,295	73,216	9,724	20,421	2,624	695				109,975	138.35
ICELAND	14	2	1							17	159,422	249	8,070							167,741	17.16
INDIA	54	15	8	3		1				81	1,615,932	355,212	19,492	6,558		3,413				2,000,607	2042.56
INDONESIA	60	18	3			1				82	374,962	59,895	2,662			329				437,848	890.60
IRAQ	6	1	4	1	1					13	4,219	412	3,114	268	479					8,492	223.51
IRELAND	10	3	1			1				15	225,999	209,285	15,146			16,649				467,079	250.81
ISLAMIC REPUBLIC OF IRAN	1									1	7,332									7,332	416.49
ISRAEL	10	2								12	364,103	6,268								370,371	305.67
ITALY	399	93	21	9	1		1		1	525	2,007,276	1,315,729	354,360	346,926	269		274		1,265	4,026,099	2141.94
JAMAICA	6	1								7	9,088	256								9,344	13.89
JAPAN	555	33	5							593	15,230,263	193,546	25,979							15,449,788	4596.16
JORDAN	8	2		1						11	29,743	50,691		36,421						116,855	35.88
KAZAKHSTAN	18	9	6	1						34	44,640	46,100	17,985	372						109,097	217.87
KENYA	32	3		1						36	37,487	365		289						38,141	60.94

Country	Number of Banks										Total Asset (millions of US dollar)										Nominal GDP (billions of US dollar)
	0%–10%	10%–20%	20%–30%	30%–40%	40%–50%	50%–60%	60%–70%	70%–80%	80%–90%	Country Total	0%–10%	10%–20%	20%–30%	30%–40%	40%–50%	50%–60%	60%–70%	70%–80%	80%–90%	Country Total	
KIRIBATI	1	1								2	27	122								149	0.19
KOSOVO	3	2								5	1,517	773								2,290	7.40
KUWAIT	3	1	1		1					6	104,380	14,363	17,917		13,303					149,963	171.96
KYRGYZSTAN	5		2	2	1					10	432		483	330	223					1,468	7.47
LAO PEOPLE'S DEMOCRATIC RE	5	2	2							9	4,233	496	304							5,033	11.69
LATVIA	7	9	2	1		1				20	9,643	23,203	781	304		278				34,209	31.34
LEBANON	34	8	5	1						48	117,629	72,592	37,338	206						227,765	49.94
LESOTHO	2	1	1							4	148	311	604							1,063	2.22
LIBERIA	3		1							4	73		279							352	2.01
LIBYA	6	4								10	56,416	7,620								64,036	44.42
LIECHTENSTEIN		1	1							2		669	361							1,030	
LITHUANIA	4	3		2	1					10	24,641	2,587		2,146	300					29,674	48.47
LUXEMBOURG	25	23	15	6	2					71	209,156	363,201	82,838	3,923	3,959					663,077	64.98
MACAO	6	2	1	1						10	25,935	31,416	112	150						57,613	55.52
MACEDONIA (FYROM)	8	5	2	1						16	3,641	2,533	927	318						7,419	11.33
MADAGASCAR	3		2		1					6	950		628		401					1,979	10.67
MALAWI	4	1	2		1					8	1,018	222	137		103					1,480	6.06
MALAYSIA	20	14	2							36	526,240	174,995	789							702,024	338.11
MALDIVES	1									1	1,179									1,179	3.06
MALI	6	3		1						10	916	2,374		644						3,934	14.45
MALTA	9	3	1							13	24,699	2,021	536							27,256	10.74
MAURITANIA	6	2	1	1						10	1,127	236	163	140						1,666	5.30
MAURITIUS	5	6	1	5						17	15,305	10,605	3,231	8,132						37,273	12.63
MEXICO	24	18	11	8	3	1				65	145,916	252,836	34,409	36,873	391	359				470,784	1297.85
MONGOLIA	7	1		1						9	1,648	2,150		781						4,579	12.20
MONTENEGRO	3	1	2	2		1				9	1,231	706	472	630		285				3,324	4.60
MOROCCO	8	3	4		1					16	84,532	95,839	23,250		3,049					206,670	110.01
MOZAMBIQUE	8	3	1	1						13	8,633	824	319	375						10,151	16.86
MYANMAR	11	1		1						13	551,525	179		5,522						557,226	65.75
NAMIBIA	3	2		1						6	3,523	4,881		1,883						10,287	13.19
NEPAL	8	7	4	7	3	1				30	3,599	3,834	2,311	4,889	1,069	477				16,179	19.76
NETHERLANDS	14	15	4	1						34	1,879,558	1,122,609	51,090	766						3,054,023	880.72
NEW ZEALAND	3	9	7	3	2					24	1,527	264,776	74,001	2,868	159					343,331	197.93
NICARAGUA	6									6	5,813									5,813	11.81
NIGER	5	1		1						7	691	431		429						1,551	8.26
NIGERIA	12	6	1	1						20	127,573	23,153	675	1,483						152,884	574.00
NORWAY	70	31	19	9	1					130	404,326	95,215	35,785	61,185	193					596,704	500.52
OMAN	2	2	2							6	37,348	11,462	17,832							66,642	77.77
PAKISTAN	18	3				1				22	110,300	3,922				9				114,231	243.38
PALESTINIAN TERRITORIES	1	1			1					3	2,425	669			279					3,373	0.25
PANAMA	52	11	11	4						78	41,083	48,724	11,305	3,995						105,107	49.17
PAPUA NEW GUINEA	4									4	10,752									10,752	16.65
PARAGUAY	3	5	5	5						18	3,696	7,057	3,807	4,037						18,597	30.88
PERU	6	4	1	9						20	18,411	66,660	690	22,673						108,434	202.90
PHILIPPINES	33	11	4	9						57	128,842	45,889	50,488	19,347						244,566	284.78
POLAND	21	12	5	2	5	2				47	169,073	172,667	28,596	4,302	7,693	5,467				387,798	544.86
PORTUGAL	62	36	5	10	2					115	147,780	301,485	15,495	191,662	79,598					736,020	230.48
QATAR	5	2								7	60,552	181,884								242,436	210.11
REPUBLIC OF KOREA	8	6	4	1	5					24	402,300	382,133	532,250	1,031	22,478					1,340,192	1410.38
REPUBLIC OF MOLDOVA	7	3	1	2						13	2,648	311	26	875						3,860	7.98

Country	Number of Banks 0%–10%	10%–20%	20%–30%	30%–40%	40%–50%	50%–60%	60%–70%	70%–80%	80%–90%	Country Total	Total Asset (millions of US dollar) 0%–10%	10%–20%	20%–30%	30%–40%	40%–50%	50%–60%	60%–70%	70%–80%	80%–90%	Country Total	Nominal GDP (billions of US dollar)
ROMANIA	6	9	7	3	2	1				28	14,088	33,901	26,593	13,572	4,948	334				93,436	199.37
RUSSIAN FEDERATION	132	45	344	283	24	3				831	1,029,291	177,322	124,970	58,073	10,672	238				1,400,566	2029.62
RWANDA	5	1		2						8	1,257	285		409						1,951	7.89
SAINT KITTS AND NEVIS	1	1								2	195	1,356								1,551	0.85
SAINT LUCIA	4	1								5	1,323	1,379								2,702	1.40
SAINT VINCENT AND THE GRENA	1									1	301									301	0.73
SAMOA	3									3	309									309	0.82
SAN MARINO	3	2								5	309	4,542								4,851	1.85
SAO TOME AND PRINCIPE	3									3	37									37	0.34
SAUDI ARABIA	8									8	432,362									432,362	753.83
SENEGAL	3	2	5	1	1					12	820	778	3,461	699	1,365					7,123	15.36
SERBIA	12	7	6	3	2	1				31	10,296	5,159	11,822	1,549	1,883	319				31,028	44.21
SEYCHELLES	4	2								6	646	580								1,226	1.35
SIERRA LEONE	10	1								11	709	71								780	4.75
SINGAPORE	17	3	1							21	928,556	30,223	310							959,089	306.36
SINT MAARTEN	1									1	623									623	
SLOVAKIA	9	3	2							14	45,180	19,143	4,102							68,425	100.33
SLOVENIA	6	10	1		2					19	6,542	36,068	1,300		860					44,770	49.57
SOUTH AFRICA	10	4	1		1					16	315,712	649	3,651		27					320,039	350.14
SOUTH SUDAN	5									5	568									568	13.90
SPAIN	80	32	19	6	4			1		142	1,970,542	1,219,637	1,939,048	12,757	12,150			13,194		5,167,328	1383.54
SRI LANKA	10	4		4						18	39,683	10,147		1,533						51,363	75.05
SUDAN	8		1							9	2,808		182							2,990	74.36
SURINAME	1			1	2					4	90			1,237	805					2,132	5.21
SWAZILAND	1	3								4	310	840								1,150	4.42
SWEDEN	74	7	1	3						85	661,030	304,789	563	5,211						971,593	571.10
SWITZERLAND	237	67	11	11		1				327	1,584,496	1,135,554	56,579	10,637		271				2,787,537	701.23
SYRIAN ARAB REPUBLIC	10	3								13	27,997	1,780								29,777	
TAIWAN	39	4	1	1		1				46	1,451,115	227,072	25,829	4,330		382				1,708,728	530.04
TAJIKISTAN	5	1			1					7	1,395	60			340					1,795	9.24
THAILAND	5	6	9	2	1					23	32,781	285,736	284,832	7,877	8,435					599,661	404.32
TOGO	3	2		1						6	509	1,436		194						2,139	4.61
TONGA	1									1	109									109	0.43
TRINIDAD AND TOBAGO	7	1	2							10	21,232	3,900	6,166							31,298	27.27
TUNISIA	6	6	3		2					17	6,111	13,238	11,328		5,922					36,599	47.61
TURKEY	23	6	4	2						35	630,271	167,915	6,137	146						804,469	798.33
TURKMENISTAN	1			1						2	5,002			2,167						7,169	46.22
TUVALU	1									1	20									20	0.04
UGANDA	16	1	2							19	5,463	266	194							5,923	27.52
UKRAINE	19	65	58	6	3					151	15,676	28,102	38,077	1,086	65					83,006	132.34
UNITED ARAB EMIRATES	8	7	4							19	215,505	183,895	73,201							472,601	399.45
UNITED KINGDOM	63	34	25	18	6					146	6,330,619	353,267	2,458,011	1,174,218	1,734					10,317,849	2991.69
UNITED REPUBLIC OF TANZANIA	25	2	3	2	1					33	13,290	504	733	255	69					14,851	48.09
UNITED STATES OF AMERICA	5,857	778	208	76	24	2	6	1	1	6,953	12,991,678	1,654,967	267,778	1,374,887	83,455	1,808	8,177	690	113	16,383,553	17348.08
URUGUAY	9	10	4							27	24,220	11,196	3,709							39,125	57.47
UZBEKISTAN	9	4	3	2	1					19	4,568	3,898	4,236	493	5,248					18,443	63.10
VANUATU	1	1								2	252	141								393	0.82
VENEZUELA	15	1	20	2						38	37,641	2,310	523,950	90,935						654,836	250.28
VIETNAM	16	21	4	3	3	1				48	66,397	151,559	4,899	5,606	5,326	123				233,910	185.90
YEMEN	5	1								6	3,441	2,227								5,668	43.23
ZAMBIA	9	5	1	1						16	4,183	2,050	360	289						6,882	27.14
ZIMBABWE	9	1	3	1						14	2,733	159	1,019	287						4,198	14.20

Appendix B

List of variables used to build the random forests EWS.

List of incorporated bank-level variables

- Interest Income on Loans/Average Gross Loans
- Interest Expense on Customer Deposits/Average Customer Deposits
- Interest Income/Average Earning Assets
- Interest Expense/Average Interest-bearing Liabilities
- Net Interest Income/Average Earning Assets
- Net Int. Increase Less Loan Impairment Charges/Average Earning Assets
- Net Interest Increase Less Preferred Stock Dividend/Average Earning Assets
- Non-Interest Income/Gross Revenues
- Non-Interest Expense/Gross Revenues
- Non-Interest Expense/Average Assets
- Pre-impairment Operating Profit/Average Equity
- Pre-impairment Operating Profit/Average Total Assets
- Loans and securities impairment charges/Pre-impairment Operating Profit
- Operating Profit/Average Equity
- Operating Profit/Average Total Assets
- Taxes/Pre-tax Profit
- Pre-Impairment Operating Profit/Risk Weighted Assets
- Operating Profit/Risk Weighted Assets
- Net Income/Average Total Equity
- Net Income/Average Total Assets
- Fitch Comprehensive Income/Average Total Equity
- Fitch Comprehensive Income/Average Total Assets
- Net Income/Risk Weighted Assets
- Fitch Comprehensive Income/Risk Weighted Assets
- Tangible Common Equity/Tangible Assets
- Tier 1 Regulatory Capital Ratio
- Total Regulatory Capital Ratio
- Equity/Total Assets
- Cash Dividends Paid and Declared/Net Income
- Cash Dividend Paid and Declared/Fitch Comprehensive Income
- Net Income–Cash Dividends/Total Equity
- Growth of Total Assets
- Growth of Gross Loans
- Impaired Loans (NPLs)/Gross Loans
- Reserves for Impaired Loans/Gross loans
- Reserves for Impaired Loans/Impaired Loans
- Impaired Loans less Reserves for Impaired Loans/Equity
- Loan Impairment Charges/Average Gross Loans
- Net Charge-offs/Average Gross Loans
- Impaired Loans + Foreclosed Assets/Gross Loans + Foreclosed Assets
- Loans/Customer Deposits
- Customer Deposits/Total Funding excluding Derivatives

List of eliminated bank-level variables (variables with missing values more than 50% of observations)

- Net Income/Average Total Assets + Average Managed Securitized Assets
- Fitch Core Capital/Weighted Risks
- Fitch Eligible Capital/Weighted Risks
- Core Tier 1 Regulatory Capital Ratio
- Cash Dividends and Share Repurchase/Net Income
- Interbank Assets/Interbank Liabilities

Appendix C

Evaluating classification accuracy

To evaluate the classification accuracy of the random forests EWS, we used K-fold cross-validation, which is a standard resampling technique used for estimating model performance. The basic idea involves using parts of the sample data to fit the model (training set) and the remaining part to estimate the prediction error of the model (validation set). First, we randomly split the observations into K folds of roughly equal size. Then, we treated one of the K folds as a validation set and fit the model on the remaining K–1 folds. We calculated the prediction error of the observations in the validation set. We repeated the process K times to obtain K different estimates of the prediction error. The K-fold cross-validation estimate of the prediction error is the average of these values. Since the typical choice of K is either 5 or 10, we used a 10-fold cross-validation. Using the same setup, we compared the performance of the random forests EWS with that of conventional EWSs based on a logistic regression and a decision tree. The random forests model produced an accuracy rate of 93.64%, whereas the logistic regression and decision tree produced accuracy rates of 65.73% and 73.75%, respectively. The result clearly indicates that random forests can build more reliable EWSs than conventional methods.

We also conducted a historical back test to evaluate the classification accuracy. More specifically, we predicted the bank status (active or inactive) in 2013 and 2014 based on banks' financial statements in 2013, excluding those that became inactive before 2013. We evaluated performance in terms of accuracy by comparing the predicted bank status with the actual status in 2013 and 2014. The random forests model produced an accuracy rate of 85.27%, whereas the logistic regression and decision tree produce accuracy rates of 53.44% and 56.43%, respectively. Once again, the result clearly indicates that the random forest EWS outperforms the conventional EWSs.

References

1. International Monetary Fund (IMF). *Macroprudential Policy: An Organizing Framework*; International Monetary Fund: Washington, DC, USA, March 2011.
2. Committee on the Global Financial System. *Operationalising the Selection and Application of Macroprudential Instruments*; CGFS Paper No. 48; Bank for International Settlements: Basel, Switzerland, December 2012.
3. Frankel, J.A.; Rose, A.K. Currency crashes in emerging markets: An empirical treatment. *J. Int. Econ.* **1996**, *41*, 351–366. [CrossRef]
4. Kaminsky, G.L.; Lizondo, S.; Reinhart, C.M. The leading indicators of currency crises. *IMF Staff Pap.* **1998**, *45*, 1–48. [CrossRef]
5. Alessi, L.; Detken, C. *Identifying Excessive Credit Growth and Leverage*; ECB Working Paper No. 1723; European Central Bank: Frankfurt, Germany, August 2014.
6. Bussiere, M.; Fratzscher, M. Towards a new early warning system of financial crises. *J. Int. Money Financ.* **2006**, *25*, 953–973. [CrossRef]
7. Rose, A.K.; Spiegel, M.M. Cross-country causes and consequences of the crisis: An update. *Eur. Econ. Rev.* **2011**, *55*, 309–324. [CrossRef]
8. Frankel, J.; Saravelos, G. Can leading indicators assess country vulnerability? Evidence from the 2008–2009 global financial crisis. *J. Int. Econ.* **2012**, *87*, 216–231. [CrossRef]
9. Duca, M.L.; Peltonen, T.A. Assessing systemic risks and predicting systemic events. *J. Bank. Financ.* **2013**, *37*, 2183–2195. [CrossRef]

10. Mathias, D.; Mikael, J. Evaluating early warning indicators of banking crisis: Satisfying policy requirements. *Int. J. Forecast.* **2014**, *30*, 759–780.
11. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer: New York, NY, USA, 2008.
12. Ghosh, S.R.; Ghosh, A.R. Structural vulnerabilities and currency crises. *IMF Staff Pap.* **2002**, *50*, 481–506. [CrossRef]
13. Frankel, J.; Wei, S.-J. Managing macroeconomic crises: Policy lessons. In *Managing Economic Volatility and Crises: A Practitioner's Guide*; Aizenman, J., Pinto, B., Eds.; Cambridge University Press: Cambridge, UK, 2005; pp. 315–405.
14. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: New York, NY, USA, 2013.
15. Varian, H.R. Big data: New tricks for econometrics. *J. Econ. Perspect.* **2014**, *28*, 3–28. [CrossRef]
16. Einav, L.; Levin, J. The data revolution and economic analysis. *Innov. Policy Econ.* **2014**, *14*, 1–24. [CrossRef]
17. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
18. Tanaka, K.; Higashide, T.; Kinkyo, T.; Hamori, S. Forecasting the Vulnerability of Industrial Economic Activities: Predicting the Bankruptcy of Companies. *J. Manag. Inf. Decis. Sci.* **2017**, *20*, 1–24.
19. Tanaka, K.; Kinkyo, T.; Hamori, S. Random forests-based early warning system for bank failures. *Econ. Lett.* **2016**, *148*, 118–121. [CrossRef]
20. Berger, A.N.; Bouwman, C.H.S. How does capital affect bank performance during financial crises? *J. Financ. Econ.* **2013**, *109*, 146–176. [CrossRef]
21. Vazquez, F.; Federico, P. Bank funding structures and risk: Evidence from the global financial crisis. *J. Bank. Financ.* **2015**, *61*, 1–14. [CrossRef]
22. Altman, E.I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J. Financ.* **1968**, *23*, 189–209. [CrossRef]
23. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013.

Article

Balancing Project Financing and Mezzanine Project Financing with Option Value to Mitigate Sponsor's Risks for Overseas Investment Projects

Jae-Il Yoo [1], Eul-Bum Lee [2,*] and Jin-Woo Choi [2]

[1]   Pohang Iron and Steel Company (POSCO), Donghaean-Ro 6261, Nam-Ku, Pohang 37859, Korea;
       jaeilu@postech.ac.kr
[2]   Graduate Institute of Ferrous Technology & Graduate School of Engineering Mastership,
       Pohang University of Science and Technology (POSTECH), 77 Cheongam-Ro, Nam-Ku,
       Pohang 37673, Korea; jonah@postech.ac.kr
*   Correspondence: dreblee@postech.ac.kr

Abstract: Major steel-making companies in Korea have recently been trying to advance into international markets for better profitability and new market shares. Even with strategic partnerships with local organizations, the Korean steel companies are facing and incurring significant risks which impact their ability to achieve a sustainable profit. The objective of this research is to determine an optimum combination of financial models, specifically Project (PF) and Mezzanine Financing (MF) with an option (convertible bond and bond with warrant). The results of the proposed model can lower interest rates of financing, thereby increasing the profitability of the project investors. To analyze the MF method's effectiveness and proper use, the following three steps are applied: (1) Monte-Carlo Simulations (MCS) using Excel and @Risk software are performed for the Net Present Value (NPV) of the project and its volatility; (2) the Black-Scholes model (BSM) is applied to evaluate MF based on project value; and (3) interest rate of MF is calculated from its option value and is reapplied back to the NPV calculation of the project to determine the effects of MF. Assuming a 50% debt/equity ratio, these simulations were performed on five cases (50% senior debt, 0% MF for a base case then increasing MF and decreasing senior debt by 10% four times). Through this process, using the 10%, MF lowered the borrowing size by 20% and using MF continued to lower the borrowing size up to 40% borrowing when using 40% MF. Based on this result, the researchers support the use of MF to optimize Korean steel international financial models. The resultant data will serve as an effective method to increase net cash flow in overseas steel-plant project investments. This research was performed for a steel plant located in Iran as a case-study, but this optimized financing method using MF with an option product can be applied sustainably not only for overseas investment of steel plants but also any other business, such as oil & gas, power generation, and transportation industries.

Keywords: Project Financing; Mezzanine Financing; option value; Monte Carlo Simulations; probabilistic cash flow; optimizing financial model; risks mitigation; investment profitability

---

1. Introduction

The Korean steel industry as a whole has been recently trying to secure new overseas markets. The companies choosing to look overseas for work are taking on significant risks in this venture due to the uncertainty of international investments, much larger than that of domestic investments. Many of them have suffered major losses from poor investments caused by entering international markets hastily. To mitigate these risks, companies prefer to organize a consortium consisting of local companies rather than make a direct, sole overseas investment. When several companies as investment

sponsors (or developers) establish a Special Purpose Company (SPC), the required capital for resultant project(s) must be financed through Project Financing (PF). However, PF has higher interest rates than Corporate Financing (CF) because PF is a type of credit loan without recourse to the sponsors, whereas CF is financed by a company's own assets with collateral securities, typically used on domestic projects. For sponsors to be successful internationally, a method is needed to lower project interest rates, ensuring sustainable profits. Moreover, as major collateral securities for the investment sponsors for PF, offtake-agreements (as purchasing agreements so-called long-term forward contracts with the recourse buyers) have not been arranged historically for the steel production industry. Instead, the trade of the production steel resources has been done at steel spot markets as a commodity like LME (London Metal Exchange). Therefore, PF has not been fully utilized in the steel industry and has big risks when used in investments abroad. For these reasons, the authors will consider the feasibility of sponsors using MF, which is a type of PF that could provide a comparatively lower interest rate.

This study investigates the feasibility of using option-based MF to supplement the high interest rates of PF. To set this discussion, the paper begins with an overview of previous literature on project capital procurement methods followed by a description of the basic concepts and characteristics of PF and MF. The possibility of improving the profitability through combination of PF and MF is analyzed by conducting an MCS of the profitability. These findings are validated through a case study [1,2].

2. Literature Review

A literature review for previous studies was carried out to understand how PF is implemented to improve profitability of the sponsors. Milton Haris and Artur Raviv [3] proposed the capital structure theories based on agency costs, asymmetric information, product/input market interactions, and cooperate control consideration. They tried to prove the capital structures based on the mathematical models and showed the relationship between leverage and exogenous factors like profitability and characteristics of the product market. However, as they applied the traditional CF for financing modeling, the increase in the project's profitability was not found.

Alternatively, Sandalkhan Bakatjan [4] proposed an interest scheme in the Build/Operate/Transfer (BOT) for infrastructure projects that were rapidly growing. They presented a simplified model to determine the optimum equity level for the decision makers and sponsors at the beginning of the evaluation stage of BOT hydro-electric power plants. Similar to the study of this paper, they tried to make a 'combination' between a financial model and a linear programming model to incorporate the objective of maximizing the benefit from the equity. Through research, they found different equity levels for optimal capital structure in BOT projects which supports the project sponsors to ensure the required equity level. Although they proposed a meaningful idea to merge a financing model and a programming model, they could not evaluate and prove the increase in profitability.

Finally, James A. Milers and John R. Ezzell [5] found that for financial management to make properly informed decisions concerning maximizing capital budgeting, a correct determination is required for a project's levered cash flows. However, although many good solutions based on existing CF or PF have been suggested in the paper, a new concept such as the combination of PF and MF has not been suggested. Therefore, they could not propose lower interest rates as long as the rate was fixed according to the model of CF or PF, and also could not expect high profitability in comparison to present financing methods.

Contrary to the previous studies by others, in this paper, the result of the MF study shows that it is the optimal capital structure when considering different factors, which directly affects sustainable profitability based on the low interest rate when compared to PF depending on the type of capital used by the company. In addition, either method can be used to evaluate the firm value because the result is the same when evaluating the firm value through capital costs, which are the Weighted Average Cost of Capital (WACC) and Adjusted Present Value (APV). A review of studies related to PF now follows [5].

Even though many studies have been done for PF and CF respectively, no proper alternative types of financing have been suggested. Therefore, this paper could change the fundamental idea for historical financing methodologies.

3. Definition for Financing Terminologies

As the publication of this paper is not a financial one, the readership of this paper may be of varying financial competency. To ensure the reader has a basic understanding of the terminology used, below is a basic definition of the fundamental financing terms used in this paper.

3.1. Classification of Financing

Financing is needed to start a business and ramp it up to profitability. There are several sources to consider when looking for start-up financing. First one needs to consider how much and at what time money is needed. The financial needs of a business will vary according to the type and size of the business. For example, processing businesses are usually capital intensive, requiring large amounts of capital. Retail businesses usually require less capital [6].

Debt and equity are the two major sources of financing. Government grants to finance certain aspects of a business may be an option. Also, incentives may be available in certain communities to encourage activities in particular industries. These are described below.

3.1.1. Equity Financing

Equity financing is the exchanging of a portion of the business ownership for a financial investment. Said ownership stake resulting from an equity investment allows the investor to share in the company's profits. Equity involves a permanent investment in a company and is not repaid by the company at a later date. The investment is also properly defined in a formally created business entity. An equity stake in a company can be in the form of membership units, as in the case of a limited liability company, or in the form of common or preferred stock as in a corporation. Companies may establish different classes of stock to control voting rights among shareholders. Similarly, companies may use different types of preferred stock. For example, common stockholders can vote while preferred stockholders generally cannot. Common stockholders are last in line for the company's assets in case of default or bankruptcy. Preferred stockholders receive a predetermined dividend before common stockholders. Equity financing incudes personal savings, friends and relatives, venture capital, angel investors, governmental grants, equity offerings, public offerings and warrants [7].

3.1.2. Debt Financing

Debt financing involves borrowing funds from creditors with the stipulation of repaying the borrowed funds plus interest at a specified future time. For the creditors (those lending the funds to the business), the reward for providing the debt financing is the interest on the amount lent to the borrower. Debt financing may be secured or unsecured. Secured debt has collateral (a valuable asset which the lender can attach to satisfy the loan in case of default by the borrower). Conversely, unsecured debt does not have collateral and places the lender in a less secure position relative to repayment in case of default. Debt financing (loans) may be short term or long term in their repayment schedules. Generally, short-term debt is used to finance current activities, such as operations, while long-term debt is used to finance assets such as buildings and equipment. Debt financing includes friends and relatives, banks and other commercial lenders, commercial finance companies, governmental programs and bonds [7].

There are two types of loans for debt financing, senior and junior loans. Senior loan is debt that takes priority over other unsecured loans and has greater seniority in the issuer's capital structure than subordinated debt, as shown in Figure 1. In the event the issuer goes bankrupt, senior debt theoretically must be repaid before other creditors receive any payment. Senior debt is often secured by collateral on which the lender has put in place a first lien. Usually, this covers all the assets of a corporation and is often used for revolving credit lines. On the other hand, in the event of insolvency,

junior debt is prioritized lower than other classes of debt. The most common kind of junior debt is an unsecured loan, which has no collateral. Another kind of junior debt is a secured loan in which another loan has priority on the collateral; a second mortgage is an example of a secured junior debt. This class of debt carries higher risk but also pays higher interest than other classes [7]. Figure 1 shows the classification of debt financing:
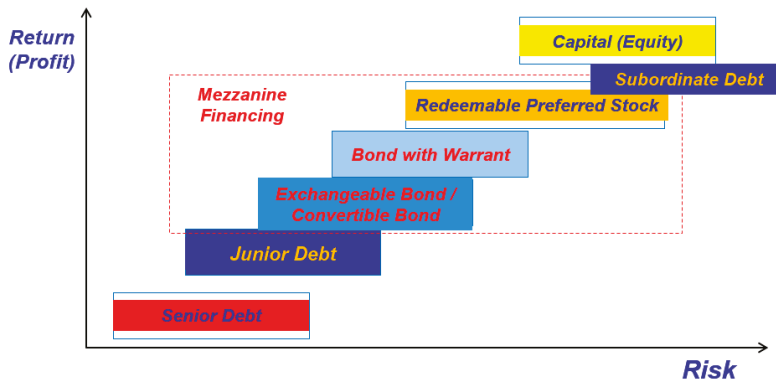
Figure 1. Classification of debt financing [7].

3.1.3. Option

One of the benefits of MF is that the lender can provide the appropriate option (call or put option), thereby allowing the sponsor to borrow at a lower interest rate than PF or CF. An option is a financing derivative that allows you to exercise your rights, which is trading underlying assets at a specified price. In contrast to futures and forward trading, the option is based on whether the option owner exercises the right to choose. That is, the option is not the obligation, but the right to trade. There are two types of options by the property of trading, one is a call option that allows you to buy an underlying asset at a set price according to the nature of the rights option, and another is a put option that can be sold at a fixed price. Also, the option can be classified as per the time of exercise. One is the American option, which allows you to exercise your rights at any time prior to the maturity of options, and another is the European option, which allows you to exercise your rights only at the time of maturity [7].

3.1.4. Lease

A lease is a method of obtaining the use of assets for the business without using debt or equity financing. It is a legal agreement between two parties that specifies the terms and conditions for the rental use of a tangible resource such as a building and equipment. Lease payments are often due annually. The agreement is usually between the company and a leasing or financing organization and not directly between the company and the organization providing the assets. When the lease ends, the asset is either returned to the owner, the lease renewed, or the asset is purchased [7].

3.2. Financing Methods

3.2.1. Corporate Financing and Project Financing

In organizations where CF is practiced, the objective is to maximize the wealth of the shareholders. CF mainly deals with the sources of funds and how the optimum capital structure is achieved. For example, a hypothetical company, ABC Ltd., acquires 50% of their funds from creditors with an assurance to give back 15% within 5 years. The remaining amount is sourced from their equity

shareholders. In this example, ABC Ltd. will pay a dividend cost of 10% on the profit. The 15% payback and 10% profit sharing are their cost of capital which they want to reduce by any means. This is done by optimizing the debt-equity ratio (50:50 in the current example) to reduce the cost of capital. At the same time, if they can reduce their total cost of capital (debt and equity included), they are able to keep better profits or think of re-investing the profit into the business. CF allows for flexibility to discover an optimum solution.

In cases where finance is required for a large industrial or renewable energy project, PF is used. In PF, the full required financing is not invested upfront. With PF, the financial institution finances the project based on projected cash flow versus the balance sheet upfront. The institution will invest in the project if the cash flow is beneficial to the financial institution. For example, X project contacts a bank or a financial institution to request 10% of the required financing, outlying the projected future cash flow. The bank or financial institution has complete discretion on whether or not to invest in X project. If the choice is to invest, there is usually a number of equity investors who invest as sponsors. The loans given are typically non-recourse (secured) loans, given against project property. The loans are paid completely from the project cash flow. If the parties default to pay back the loan, then the project properties are seized. To conduct the process properly, an SPC entity is created for the entire project.

Both CF and PF are commonly used on steel mill projects. As a means of comparison and description, Table 1 shows the basic elements of each. PF is a financing method that collates project assets and has the main repayment source through future cash-flows. SPC manages the project, and the recourse is limited when there is a problem to repay borrowings. Figure 2 depicts a standard organization of PF.
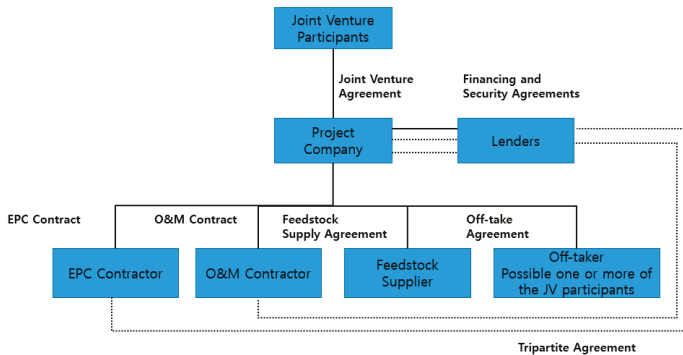
Figure 2. Stakeholders Relationships and Contracts Arrangement for **Project Financing** (PF) [8].

Table 1. Comparison between PF and CF.

Item	Project Financing	Corporate Financing
Loan Structure		
Mortgage	Project Asset and Cash Flow	Mortgage Provision of Sponsor
Accounting	Off Balance Sheet	Reflection on Business Statements of Sponsor
Borrowing Period	Long-Term	Short-Term
Responsibility	No/Limited Recourse	Full Recourse
Interest Rate	(Relatively)High	Low (Credit rating of Sponsor, Mortgage)

3.2.2. Benefits of Project Financing

PF has been expanded into various industries, and related research has been conducted. Salman [9] summarized the reasons to use PF and explained it in relation to the optimal capital structure. Afterwards, John Teresa [10] conducted a PF optimization study to conduct a comparison with CF, and Benjamin [11] focused on its economic advantages. These studies indicate that PF has many advantages in that it lowers the project risk, uncertainty, and agency costs while also resolving overdrafts. Enzo Scannella [12] found advantages of PF as a new financial model, especially in the energy industry [13].

3.2.3. Disadvantages of PF

In spite of these advantages, PF is quite complex and costly to assemble. The cost of capital arranged through this route is high in comparison to capital arranged through conventional routes. The complexity of PF deals is due to the need to structure a set of contracts that must be negotiated by all of the parties to the project. This also leads to higher transaction costs on account of the legal expenses involved in designing the project structure, dealing with project-related tax and legal issues, and the preparation of necessary project ownership, loan documentation, and other contracts [13]. MF can potentially mitigate PF's disadvantages due to its ability to lower the interest, which is described in greater detail below.

3.2.4. Mezzanine Financing

MF refers to derivatives in the middle stage between stocks and bonds. MF can be flexibly funded because it can combine equities and debt components to respond appropriately to a given situation. It provides incentives to investors to finance unsecured funds because securing senior debt is not appropriate when business risk is high, or the credit rate of the lender is low. Investors receive lower interest rates than regular bonds but have a higher return on investment than general bonds because they have stable interest income as well as performance-based options. If a company is liquidated due to bankruptcy, it is subordinated to senior debt, which is borrowed money that a company must repay first if it goes out of business. This debt is more senior to equity, so the expected profit and risk are intermediate between equity and debt [14–17].

MF is an offshoot of traditional financial instruments that combine the characteristics of stocks and bonds to pursue both profitability and stability. In addition to fixed interest rates, investors can exercise additional rights as these provide a high risk and high return and in the worst case, the investor can obtain a fixed interest income. MF can be classified into two categories: option portion and priority portion. These are shown in Table 2. The convertible bond (CB) and bond with warrant (BW) are optional products with additional options by agreement. An option is a product that has the right to buy or sell an underlying asset at a specific time or within a set period of time. In the position of an issuer who raises funds, they have the advantage of lowering the capital cost through options and relatively financing. In addition, debts are debited at the time of the initial issuance, but the property changes to Equity is applied at the time the option is exercised, thus improving the financial state of the enterprise. On the other hand, investors who lend money have an opportunity to earn stable, high profits at the same time, so there is merit as an investment product. Table 2 shows the types and characteristics of MF [18,19].

As stated, due to the consortium of companies developed for overseas steel projects, PF is the required financing model. However, this model has high interest rates, and a method is needed to mitigate these interest rates and improve a project's profitability. The author proposes utilizing an option-based MF for this purpose, and its use is defended in the sections below. As such, the features of PF and MF are discussed in the following section.

Table 2. Types and Characteristics of Mezzanine.

Item	Characteristic
Convertible bond	Additional options are provided by the general bond product agreement. At the time of conversion, the function of the bond is terminated.
Bond with warrant	Additional options are provided by the general bond product agreement. Maintains the function of bonds even when issuing warrants.
Junior loan	Reimbursed after the repayment of senior loans
Preferred stock	Unlike ordinary shareholders, there are no voting rights and preferred dividends

4. Research Methods

For initiating a project, the project owner needs to raise funds though PF or CF. However, as mentioned above, PF needs a high interest rate in return for low risks in the case of project failure, while CF entails a significantly high level of liability to the owner in the case of failure. In order to find a better way in which the owner can obtain lower interest rates while keeping a low level of failure impact, MF is reconsidered. This study investigates the feasibility of using option-based MF to supplement and mitigate the high interest rates of PF. This process includes the following research methodologies: (1) literature review; (2) Monte-Carlo simulation modeling; (3) a case study.

4.1. Literature Review

A literature review was conducted on the basics in project capital procurement methods and the three financing methods within this paper: CF, PF, and MF. Much of these findings are discussed in the introduction section as background for the problem. The literature used for this research includes previous discussions on methods to test project profitability and how to convert cash-flow findings into option pricing.

4.2. Monte-Carlo Simulation Modelling

From literature findings, the authors have developed a three-step process to calculate the optimal ratio of MF, shown below in Figure 3 and is as follows: (1) calculate project cash-flow and volatility through two @Risk Monte-Carlo analyses; (2) determine the adjusted interest rate of MF with consideration of the option value, calculated by inserting results from step 1 into the BSM formula; and (3) apply the adjusted interest rate from step 2 to the Net Present Value (NPV) calculations to determine the effects of MF on the project with a final output of the optimal MF ratio.

4.3. Implementation of Case Study

The project chosen for the case study is an Iranian Public Knowledge Project (PKP) integrated steel mill project with an annual production capacity of 1.5 million tons that utilizes FINEX and CEM (Continuous Endless Milling). FINEX is an iron making technology developed by Siemens VAI and POSCO. Molten iron is produced directly using iron ore fines and non-coking coal rather than traditional blast furnace methods through sintering and reduction with coke. Elimination of preliminary processing is claimed to make the plant for FINEX less expensive to build than a blast furnace facility of the same scale. Additionally, a 10–15% reduction in production costs is expected/claimed through cheaper raw materials, reduction of facility cost, pollutant exhaustion, maintenance staff and production time. The process is claimed to produce less pollutants such as SO_x, NO_x, and carbon dioxide than traditional methods. This process is essentially a combination of FINMET's Fluidized Bed and COREX's Melter Gasifier, hence its name "FINEX". Both methods are unique to the domestic company P and is to be constructed in the Chabahar economic zone on the coast of Oman, southern Iran. The total investment (CAPEX) is about 1.8 billion USD.
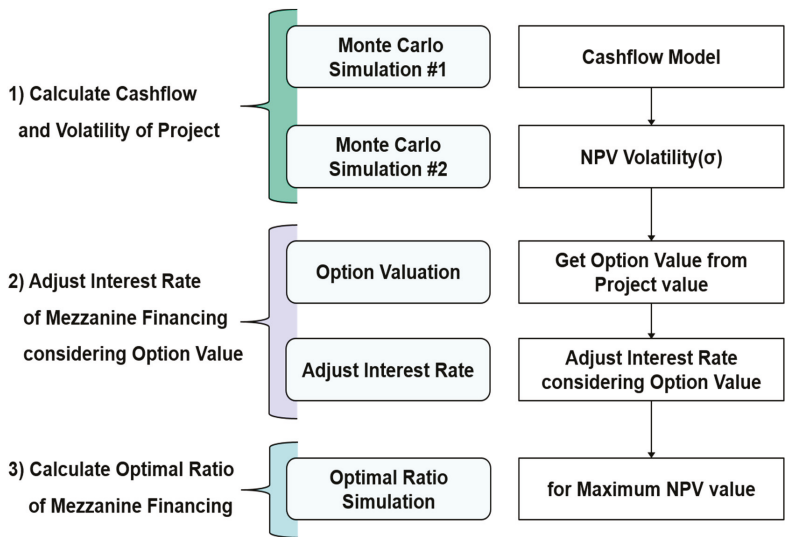
Figure 3. Framework of Net Present Value (NPV) Simulation Model.

The main conditions are as shown in Table 3, and the production capacity, the investment cost, construction period, and production products are obtained from the official announcement data when the Memorandum of Understanding (MOU) was concluded. Tax, market risk, beta are assumed with the reference paper. The risk-free rate uses the last 4-year average of 10-year US Treasury yields, and the consequences of a change in the country risk premium are simulated through a sensitivity analysis after applying the same numerical value as that of Brazil, Indonesia, and India, since the country risk of premium of Iran is not defined [20–22].

Table 3. Iran Public Knowledge Project (PKP) conditions.

Category	Data
Capacity	1.5 MT/Year
Investment Cost	1.8 Billion USD
Construction Period	3 years
Tax	30%
Product	HR Coil
Risk-Free Rate	2.48%
Market Risk	10%
Beta	1.5
Country Risk	3.4%

Concerning the financing structure for this project, the total investment cost of 2 billion USD is raised with 50% equity and 50% debt. The joint venture between Iran PKP and domestic company P will invest 1 billion USD in equity, and the SPC will raise the remaining 1 billion USD from the lender. Debt at this time is procured as a general senior debt, and the interest rate is later defined according to the weighting.

5. Literature Review Findings

The conventional engineering economic analysis with discounted cash-flow (DCF), which typically yields IRR (internal rate of Return) and payback period has been traditionally used as

an industry practice to assess the profitability of overseas investment projects, at least for major Korean steel companies. However, in this study, the authors suggested to use a more advanced methodology such as Mezzanine Financing with Call or Put option values in order to reinforce the profitability of the investors by lowering the interest rate of financing.

Along with a better understanding of the problem, the literature review shed light on how to calculate and compare different options. Kim, Yong-gu calculated the volatility from the project profitability using the outputs to calculate the option value based on the measured volatility [23]. Because this study has the same goals, it also uses this methodology. The basic structure is to calculate the option value of the MF based on the volatility of the Project profitability, reflect this in the interest rate, and borrow at this low interest rate [24]. Lee, Cheuk Wing studied the feasibility of applying a hybrid bond, which is a type of MF for Renewable Energy, but analyzed it in terms of only reducing risk, not in deriving the economic value [18]. This paper applies the option value to the MF reviewed by Lee, Cheuk Wing [18] based on the volatility of the project profitability used by Kim, Yong-gu [24] and Jung, Young Ki [25] to calculate the interest rate of PF. These results are used to find the optimal ratio of the MF while re-calculating the capital procurement cost and project profitability.

Next, the authors needed a way to convert the NPV volatility into an adjusted interest rate. The total value of the option mezzanine products, which is the subject of this study on MF, can be divided into the value as a general bond and as the value of an option. In other words, when compared to junior loans, which is a priority order mezzanine product under the same conditions, the interest value can lower only the remaining interest profit after subtracting the option value from the total value that needs to be obtained.

This paper attempts to evaluate the value of CB among option MF and assumes that the CB is a European call option product that can be converted only at maturity, with the option value measured using BSM. An option can be regarded as a means to avoid the volatility of the underlying assets, and studies have been conducted to evaluate the value of the options as trading becomes more active in the market.

Black and Scholes were awarded the Nobel Prize by presenting a systematic option pricing model for the first time, and their model is widely used although it is based on unrealistic assumptions. BSM is a European option that allows investors to borrow under risk-free rate and the full capital market, and the execution is allowable only on the maturity day and assumes no basic dividend. The pricing model of the European call option presented under this assumption can be expressed as Equations (1) and (2) [20,24,26].

$$C = S \times N(d_1) - K \times e^{-R_f \times T} \times N(d_2) \tag{1}$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(R_f + \frac{1}{2}\sigma^2\right) \times T}{\sigma\sqrt{T}}, \; d_2 = d_1 - \sigma\sqrt{T} \tag{2}$$

where,

- C: Call premium
- S: Current stock price
- N: Cumulative standard normal distribution
- K: Option striking price
- R: Risk-free interest rate
- T: Time until option exercise
- e: Exponential term
- ln: Natural log

6. Simulation Modeling for Optimizing MF

As shown in sub-sections below, the project NPV Simulation process is divided into three stages. The first step is developing a cash-flow model through two MCS to obtain the volatility of the NPV.

The option value of the MF and adjusted interest rate are then calculated using the Black-Scholes equations and the results from the first step. Since the option value varies depending on the portion of the MF borrowing, this work should be repeated for each borrowing weight section. In the final step, the project NPV is simulated according to the ratio of the senior debt and MF, and the optimal Mezzanine interest ratio is obtained. The profitability of the project is expressed as NPV, discounting the future cash flow (inflow and outflow) as the present value. The project NPV Simulation Modeling is described in detail below [22].

6.1. STEP 1a: Calculate Cash-Flow Modeling (1st Monte Carlo Simulation)

The first MCS is run with the @Risk Program to create a cash-flow diagram and calculate the project NPV. This step begins by developing a project cash-flow model in Excel. Considering that the cash inflow and outflow are the same as analyzing with the existing DCF (Discounted Cash-Flow) method, a valuation method used to estimate the attractiveness of an investment opportunity (the free cash flow needed during the construction period and operation period) is derived. DCF analyses use future free cash flow projections and discounts them with a required annual rate to arrive at present value estimates. A distribution model of the input data based on the past historical data is obtained for the simulation. In addition, if there is a correlation between the input variables, it is necessary to set the correlation to reduce unrealistic cases and improve the reliability of the results. A correlation function (formula) can be defined in the @Risk program, and it has a function to reflect the execution of the simulation (simulation). Therefore, the correlation of related variables should be defined to improve the reliability of the result. Once the probability distribution of the input variables and the correlation are defined, the simulation is ready to be run. To obtain reliable results, the number of repetitions should be high, generally 10,000 times. When the simulation is run, the program will output the results in a short time and will display these as a probability distribution graph for project NPV as shown in Figure 4a [27]. The advantage of the MCS using @Risk software is to easily perform a sensitivity analysis, which represents the relative impact and consequently, the sensitivity of input parameters to the project NPV on a so-called "Tornado-chart", as illustrated in Figure 4b.

Figure 4. *Cont.*

1000 USD
Inputs Ranked By Effect on Output Mean

Sales Price	331,058.51	749,049.67
Exchange Rate	423,594.69	761,817.89
Raw Material	411,141.77	673,895.81
Risk-Free Interest Rate	555,284.70	622,677.32
Investment Cost/FS	550,039.06	599,460.29

Baseline = 583,006.30

1000 USD

(**b**)
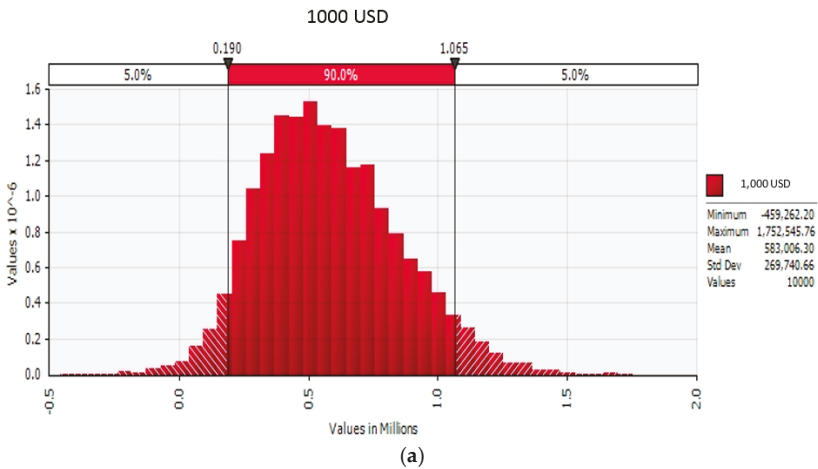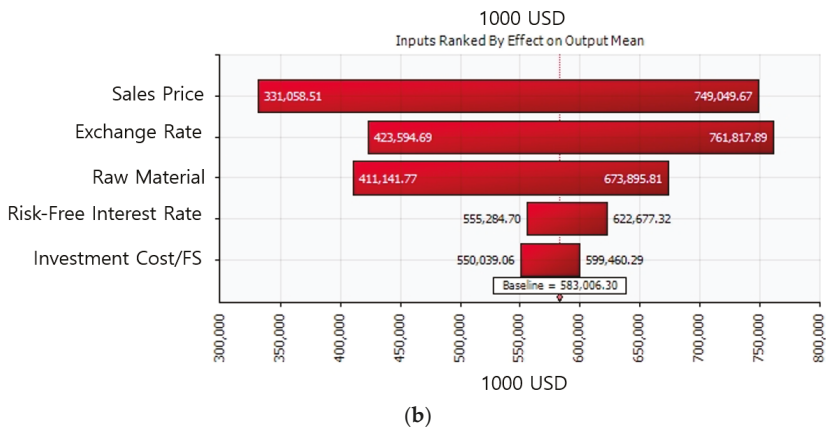
Figure 4. Screenshots of Mote Carlo simulations for cash-flow modeling. (**a**) Probability Distribution of Project NPV; (**b**) Sensitivity chart (so-called Tornado-chart).

6.2. STEP 1b: Project NPV Volatility (2nd Monte Carlo Simulation)

Since the NPV is calculated as the basis of the volatility analysis, the volatility for basic assets that can be used to evaluate the option value is obtained through an analysis. MCS is run once more to obtain the volatility of the NPV. The MCS is the process to standardize the normal distribution of the mean and the standard deviation of the NPV obtained previously using an Excel Macro. First, 10,000 random numbers (mean variance) with a mean of 0 and a standard deviation of 1 are generated through a Gaussian distribution in Excel. By multiplying the random variables by the standard deviation of the NPV and adding the mean, the NPV can be a standardized normal distribution. The NPV volatility is obtained by obtaining the log method of 10,000 generated results, and the standard deviation of the log method. This is called the Wiener Process as defined by Black-Scholes and must be satisfied to utilize the BSM [20,24].

6.3. STEP 2: Adjustment of MF Interest Rate Reflecting the Option Value

As mentioned earlier, CB periodically receives interest, and at the time of conversion at a certain point, the status as a general bond ends and holds stock. Therefore, the total value of the CBs can be divided into bond value and convertible value. In other words, the value as a general bond held until the conversion right is exercised is the expected profit due to the interest rate, and the total value can be obtained by adding the value of the conversion right obtained using the BSM. When compared to the junior loan, which is a priority order mezzanine product under the same condition, the interest value can be reduced because it needs to obtain only the remaining interest profit, which is the subtracted option value from the total value. The first step is to calculate the option value of the CB by substituting the NPV and the NPV volatility of the project into the BSM, and the second step is to deduct it from the value of the junior loan to obtain the adjusted interest rate. Since the option value varies depending on the size of issuance of the CB, the previous two steps must be repeated for each size, as shown in Figure 5 [28].

The option value (C) is calculated according to the value (S), the conversion price (K), the maturity (T), and the volatility (σ) of the underlying asset after defining BSM formula in Excel. Since the underlying asset is the NPV of the project, the calculated NPV value can be substituted. However, since the effect on the NPV varies depending on the loan, the NPV value calculated by multiplying the total NPV by the portion of the loan is used as the value of the underlying asset. Since the CB receives the principal of the loan at maturity or exercises the conversion right, the exercise price is the loan

itself. Since the maturity (T) of the CB is the same as the maturity of the bonds at issue, it is necessary to include the maturity of the bond, which is set to 10 years as the general loan period. Finally, the volatility is calculated by substituting values derived from the two MCS of the previous section.
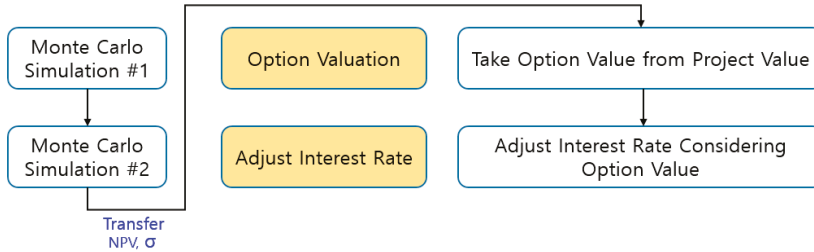
Figure 5. Option value calculation and Mezzanine Financing (MF) interest rate adjustment process.

For companies issuing CB, the investment profitability can be expected to increase as much as the interest rate paid annually is decreased. The interest rate for each CB should be assessed per issue, as the convertible value varies according to the −issuing CB volume with its lender's different expectations and its portion amongst the total loan as well. In addition, since CB is a substitute for senior debt while keeping the same debt rate, not a substitute for existing equity, the interest rate of the senior debt should also be changed. To conduct a comparison, the reference interest rate is required. The authors used data from Bond Capital, which has analyzed and studied MF since the early 2000s.

The expected return rate of the senior debt and MF is defined to consist of 50% Equity and 50% Debt based on the above data in the following manner, as shown in Figure 6.

(1) Interest rate based on senior debt: 12% (maximum loan portion: 50%)
(2) Interest rate variation due to a 10% decrease in borrowing weight: −0.5%
(3) Interest rate based on MF: 13% (minimum borrowing portion: 10%)
(4) Interest rate variation due to a 10% increase in borrowing weight: +1.0%.
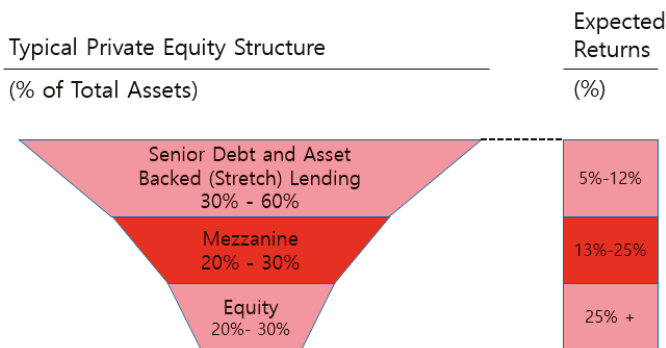
Figure 6. Expected return rate by means of capital procurement [29].

6.4. STEP 3: Deriving Mezzanine Optimum Utilization Ratio

Based on the adjusted interest rate of the MF calculated in the previous section and the interest rate of the senior debt, the change in the NPV is simulated according to the ratio of borrowing for both capitals. The optimal capital structure for which the value of WACC is minimum and NPV is

maximum is also calculated. The simulation is conducted by changing the ratio of the MF to the base case for which equity is 50% and the debts are 50%, with senior debts only. A total of four cases are compared where the Equity is fixed at 50% and senior debt is replaced by MF in increments of 10%, from 10 to 40%. The case of no senior debt is excluded in this study because the MF itself has a senior position and the basic interest rate is lowered, so there is no meaning in a comparison. Based on the simulation model set up here, the next chapter simulates the NPV according to the utilization ratio of the MF applied to the overseas steel project, which company P in Korea is promoting together with PKP of Iran, taking into account the optimal ratio of MF.

7. MF Case Study Project with Monte-Carlo Simulation

As stated in Section 4.3, the project chosen for the project case study was an Iranian PKP integrated steel mill project. The above steps are executed for the project case study obtained data below.

7.1. STEP 1a: Project NPV Calculation (1st Monte Carlo Simulation)

For the same probabilistic analysis as the DCF, a deterministic analytical method, the cash-flow model needs to be implemented in Excel. The net cash flow needs to be found during the construction period and the operation period considering the cash inflow and cash outflow. Unlike the DCF method, it is necessary to set the distribution of the input variables into the simulation. Using the distribution fitting function in the @Risk program with the data from the last four years (2013–2016), as shown in Table 4, a distribution model is obtained for each variable.

Table 4. Distribution of Input Variables for Economic Analysis of Project.

Item	2013	2014	2015	2016	Source
Ore Iron ($/Ton)	158.5	131.3	102.1	114.3	
Coking Coal ($/Ton)	136.4	97.1	54.9	57.5	Steel data
Price ($/Ton)	750.8	676.7	585.0	582.5	
US Treasury Bond Rates (10 years, %)	3.03	2.17	2.27	2.44	Bank of Korea—Economic
Won/Dollar exchange rate	1095.04	1053.22	1131.49	1160.5	Statistics System

The distribution model is obtained using monthly data over four years, and the figure is produced with a total of 48 data values for each factor. Once the distribution model of the input variables is obtained from historical data, a Cash-flow model using it as a boundary condition needs to be implemented. The revenue is comprised of revenue from product sales, intermediate goods, and by-product sales. Expenses are comprised of selling and administrative expenses and maintenance expenses based on cost of materials, labor, expenses, and others. The ratio of each item is adjusted to the present situation by referring to the paper that examined the existing PKP. The income statement is made based on the annual income and expenses, thereby the tax and after-tax profit can be sequentially obtained. Based on this, the cash outflow, inflow, and free cash flow are obtained during the construction period and the operation period. The sum of discounted free cash flow is calculated reflecting the present value, that is the Project's NPV using the DCF method. Now that one representative project NPV has been obtained using the DCF method, the @Risk Program is now used to reflect the correlation of the input variables to implement the MCS. The product price and raw materials consisting of iron ore and coking coal showed a positive correlation of 0.9 using the correlation function of Excel. @Risk's Define Correlations function can be used to define the correlation between the above variables, and a value of 0.9 is entered. To obtain reliable results, a simulation is performed by selecting 10,000 cycles, which is the maximum number of simulation cycles that can be selected in the program [15].

The average NPV of the PKP project financial model was 583 million USD and the standard deviation was 270 million USD. The second MCS is run based on these values to determine the volatility of the NPV.

7.2. STEP 1b: Project NPV Volatility (2nd Monte Carlo Simulation)

The mean and standard deviations of the project NPV that are the basis of the volatility analysis are obtained, and the volatility that satisfies the Black-Scholes' Wiener process [7] can be obtained. The results of the MCS using the Excel macro shows that the volatility of the NPV is 87.06%. An option product is traded as a risk hedge against such volatility. Generally, the greater the volatility, the greater the option value. Based on the volatility of the NPV and NPV calculated above, the value of option-based MF is calculated, and based on this value, the process of adjusting the appropriate interest rate reflecting MF is shown.

7.3. STEP 2: Adjustment of MF Interest Rate Reflecting the Option Value

A standard interest rate should exist to calculate the appropriate interest rate by reflecting the option value to the interest rate of the option MF. Based on the data of the bond capital as defined above, the following amendment to the SPC is applied in accordance with the cost of equity procurement in the project. The interest rate for senior debt, which was recalculated based on the equity fund procurement cost of the PKP project, is 2.8–6.7%, and the interest rate of the MF is 7–14%. Based on the data, the standard interest rate to be applied to the actual simulation should be defined by the borrowing weight. Iran has the highest risk rate of 6.7% when the debt ratio of senior debt is 50% because of the country risk premium (3.4%), which is higher than that of developed countries. The interest rate of the MF rate is defined as 7.2% when the ratio of borrowing is at least 10%, and it is defined according to the ratio of borrowing by the following criteria.

(1) Interest rate based on senior debt: 6.7% (maximum loan portion, 50%)
(2) Interest rate variation due to a 10% decrease in borrowing weight: -0.5%
(3) Interest rate based on MF: 7.2% (minimum borrowing portion, 10%)
(4) Interest rate variation with a 10% increase in borrowing portion: +1%
(5) Interest rate variation according to the increase and decrease in the borrowing ratio of the MF is a factor that can influence the result in the future as the basis to derive the optimal utilization ratio, and the simulation is performed by changing it in the sensitivity analysis and will compare the results.

Since all data required to calculate the option value of the MF is calculated, the option value is calculated according to the following process, and the interest rate is adjusted by reflecting this value. Figure 7 shows the process of the mezzanine option value and the interest rate adjustment.
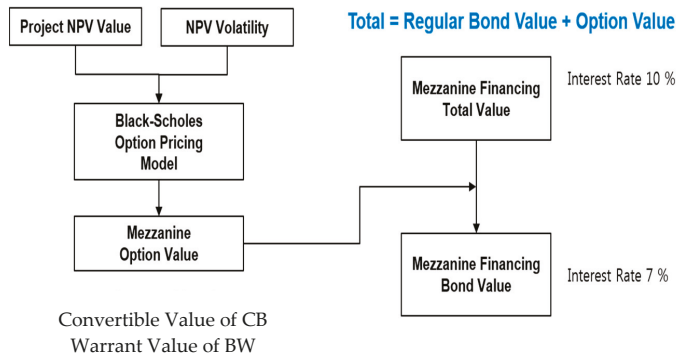
Figure 7. Mezzanine option value calculation and interest rate adjustment process.

Amongst the MF options, only the value of CB is evaluated, and the option value is calculated using the BSM as setting a European call option with a 10-year maturity product that is converted only

upon maturity. Based on the calculated option value, the interest rate should be adjusted according to the interest rate based on the MF defined in the previous section. The base case is where the ratio of equity and debt is 50%, the WACC for capital procurement is 9.29%, and the target rate (hurdle rate) is 12.69, which is the sum of the WACC and country risk premium 3.4%.

In the case of an international investment, the hurdle rate is the discount rate, and the NPV calculated based on the base case of the Iranian PKP Project is 580 million USD. A total of 4 cases are compared while the Equity remains fixed at 50% and senior debt is replaced by 10% MF increments from 10 to 40%. The option value for Case 1 can be calculated, where the loan amount of the MF is 10% and the loan period is 10 years out of the total investment cost of 2 billion USD. Among the total investment of 2 billion USD, 10% of the borrowing portion is worth 208.7 million USD, and the interest rate when borrowing from a subordinated loan (junior loan) is 7.23%, and the present value of the revenue expected by the mezzanine lender is 30.22 million USD. The total NPV of the project is 580 million USD, and the portion of the MF is 10%. Therefore, the NPV of the CB is 58 million USD. The volatility of the NPV is 87.06%, on applying this to the BSM, the conversion value of the CB is obtained as 42.3 million USD. Therefore, the lender needs to earn only 259.9 million USD, which is the profit from 302.2 million USD obtained from the subordinated loan (junior loan) minus the value of 42.3 million USD which is the value of conversion right, and the interest rate at that time is 4.95%. Based on the calculated interest rate, the WACC is lowered from 9.29 to 9.03% of the original base case, and the discount rate decreases from 12.69 to 12.43%.

When the figure is converted to the present value with the condition of no cash-flow change in the Iran PKP project's financing model, if NPV is recalculated reflecting 12.43% as the modified discount rate, the value is 620 million USD, which is 40 million higher compared to 580 million initially. The simulation is executed for cases 2, 3, 4, in the same manner as in case 1, and the obtained figures are as shown in Table 5.

Table 5. Iran PKP Project NPV Simulation result.

Category			Base	Case 1	Case 2	Case 3	Case 4
Capital Structure	Equity		50	50	50	50	50
	Debt	Senior	50	40	30	20	10
		Mezzanine	0	10	20	30	40
IRR (%)			17	17	17	17	17
WACC (%)			9.29	9.03	8.98	9.13	9.50
Discount Rate (%)			12.69	12.43	12.38	12.53	12.90
NPV (USD)			583	622	629	606	551

7.4. STEP 3: Deriving Mezzanine Optimum Utilization Ratio

According to the results of the case study, the maximum project NPV was obtained by minimum WACC of the capital cost and the minimum discount rate when the CB (senior debt) was 20% of the total investment. As can be seen, the WACC was reduced as the senior debt is replaced by MF up to a 20% MF. However, when the MF reached 30% and 40%, the WACC increased. The reason for this result is that MF is more dependent on the interest rate than senior debt, and the interest rate rises as the amount of the loan increases. If the borrowing amount is large, borrowing with senior debt can be more advantageous even if an option value is considered.

In summary, the WACC decreased from the base case (consisting only of senior debt) at 10% and 20% MF but then increased at 30% and 40% MF as shown in Figure 8a (the red-circle indicates the optimal low-interest point). As illustrated on Figure 8b there is an inverse relationship between WACC and NPV; i.e., when the WACC decreases, the NPV increases (the red-circle indicates the highest NPV point).
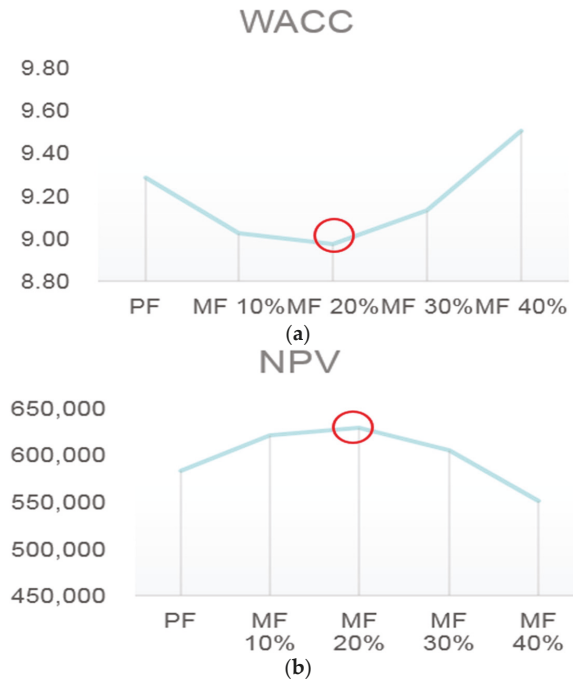
Figure 8. Result of Simulations Sensitivity. (**a**) Weighted Average Cost of Capital (WACC) versus PF and MF portions; (**b**) NPV versus PF and MF portions.

8. Sensitivity Analysis for the Case Study Project

Although not discussed in the NPV simulation discussion, a sensitivity analysis was conducted on the project to understand the three factors that have the greatest impact on the results of the simulation. These were determined to be the NPV's volatility, interest rate increase and decrease with MF, and country risk premium. These variables are discussed in greater detail below.

8.1. NPV Volatility

First, the results of the simulation change are analyzed according to the NPV volatility. The NPV volatility is one of the values that represents both the project cash inflow and cash outflow, and it is a key factor that determines the option value of the MF through the BSM. The greater the volatility of the project NPV, the more likely it is that the project becomes larger or smaller than the average value, which means the option value of the MF is higher. The NPV volatility is simulated by dividing it into five categories by changing the volatility calculated from the Iranian PKP Project (87.06%) to 30%, 60%, 120% and 150%, as shown in Figure 9.

As the volatility of the NPV increases, the option value of the MF increases, resulting in a decrease in the capital cost and an increase in the NPV. In addition, as the value of the MF increases, the optimal utilization ratio also increases, moving to the right side of the graph. On the other hand, as the NPV variability becomes smaller, the value of the option becomes smaller, so the value of the utilization becomes lower, and the optimum utilization ratio becomes smaller.
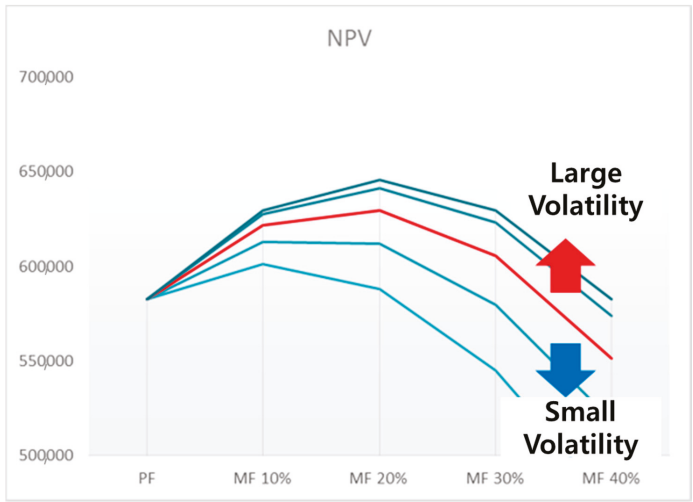
Figure 9. Simulation result according to the NPV Volatility.

8.2. Increase/Decrease in the Interest Rate by the MF Borrowing Size

The second sensitivity analysis simulates how the Project NPV changes according to the increase/decrease in the interest rate by the amount of MF borrowed. The increase in interest rate with the MF was divided into five categories, and a simulation was performed with 0%, 0.5%, 1.5% and 2% based on the 1% used in the project.

The smaller the increase in the interest rate per CB, the greater the increase in ratio utilization of the maximum NPV, therefore the graph moves to the right. In an unrealistic case, the NPV is maximized in Case 4 due to the option value if the interest rate is the same. In the case of an interest rate increase of 0.5%, the NPV for Cases 2 and 3 is almost the same, and the NPV improves significantly compared to that of 1%. In contrast, if the increase in interest rate is large, the cost of capital procurement increases, so the value of the MF utilization becomes smaller, and the optimal utilization point moves to the left in the graph as shown in Figure 10.
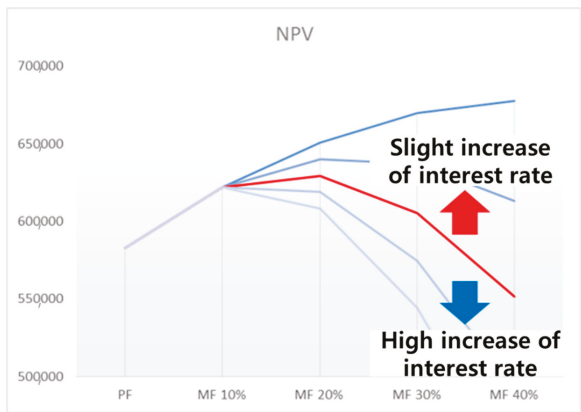
Figure 10. Simulation result according to increase/decrease in the interest rate by the MF borrowing size.

8.3. Country Risk Premium

Finally, the changes in the results are simulated according to the country risk. A simulation was conducted with two cases where the low country risk premium is 1% and 2%, and the high-country risk premium is 4% and 5% based on Iran's country risk premium (3.4%).

As a result of simulating the change in NPV according to the country risk premium, the change in NPV is shown to change in the base case, unlike in the previous two cases. Since the country risk premium is different from the cases when the project is executed in Iran and USA, respectively, the discount rate is also different. Therefore, the project NPV changes. The results of the simulation show that a smaller country risk premium results in a better overall NPV and higher MF value. This is because as the NPV increases, the option value increases for the same loan amount. As the country risk premium increases, the NPV is lowered overall, and the value of the MF is also lowered.

The sensitivity analysis result for above three factors (i.e., (1) NPV volatility; (2) MF interest rate; and (3) county risk premium) shows that the overall effect of NPV improvement and the adjustment of optimal ratio accordingly, as the factors change. However, overall the use of MF improves the NPV, and the optimal ratio is found to exist as shown in Figure 11.
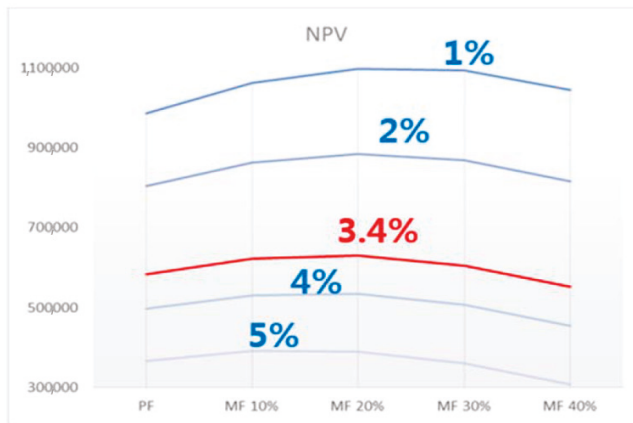
Figure 11. Simulation result according to country risk premium.

9. Conclusions

As the domestic steel industry is saturated within the domestic market, Korean steel companies have been actively trying to advance overseas to secure new markets and improve profitability. To minimize risk, the Korean companies often strategically contract with local companies. This results in the necessity of project financing (PF) whose interest rate is higher than the traditional corporate financing (CF) used domestically increasing capital procurement, decreasing profit. As such, this paper sought to lower the interest rate, maximizing profit. Based on literature findings, the mezzanine financing (MF) method can procure funds with lower interest rate, and was investigated within this paper for its efficiency.

To investigate said efficiency, this paper used multiple Monte Carlo Simulations to achieve NPV and project volatility values, the Black-Scholes model to convert these values into adjusted interest rates, and inputting the findings into four cases (base case with 50% PF and cases 1–4 decreasing PF and increasing MF at intervals of 10%). In performing these steps on a case study Iran PKP project, it was found that the project NPV is maximum (WACC of capital cost and discount rates minimum) when the MF (using the option of a convertible bond) was 20% of total investment cost. However,

when MF was increased to 30% and 40%, the NPV decreased (WACC of capital cost and discount rates increased).

These results can be expected to vary depending on the conditions under which the Iranian PKP project is based and the results obtained under a specific environment. To investigate the impacts said environment can have, a sensitivity analysis was performed with the following having the most significant impact on NPV: volatility, increase/decrease in the interest rate by the MF borrowing size, and country risk premium. The results of the sensitivity analysis for these three factors show that the NPV improvement effect and optimal utilization ratio changed depending on the variation of the factors for utilizing MF. However, it is clear that usage of MF is beneficial for improving MF and the existance of an optimal point. Therefore, it can be concluded that the proper use of MF can improve project profitability by considering various conditions rather than financing only senior debt to carry out the project, which could affect a number of sponsors and stakeholders to initiate more projects easily in the future. Though this paper was used for a Korean steel company's international investment cases, the overall results can be applicable with some relevant local adjustments for other types of international investment mega-projects in which a special purpose company (comprised of multiple investors and corporations) needs to be created.

In the course of this study, the results were calculated with a widely used program based on proven theory and data, but there are some limitations. First, the data from the last four years is similar to that used in a deterministic method to calculate the volatility of the NPV. However, considering that the project period is 15 years, this is a short period for a probabilistic analysis. Therefore, the accuracy and reliability of the probabilistic analysis can increase based on sufficiently long data. In the NPV simulation, the ratio of borrowing by MF is divided by 10%, and the total of four cases are analyzed. Therefore, the optimum utilization ratio is also calculated in units of 10%. However, since the optimal point needs to be found through a more accurate and continuous analysis in order to apply it to real projects, the ratio of the MF should be divided into smaller increments, for example, 5, 3, or 1%. In addition, applying MF was only examined from an economic perspective, without including the additional time needed to increase the number of contractors and the complex characteristics of the MF. The sensitivity analysis methodology introduced in this case study can be applied to cover a wider range of various "what-if scenarios".

The actual project is based on a very limited amount of time, and sometimes it is necessary to maintain the timeline, even if the cost is higher, so the funding plan should be thoroughly reviewed in advance. If funds are raised without planning, money will likely be borrowed at a higher interest rate. Finally, the convertible bond was analyzed as being issued at the beginning of the business with a fixed 10-year maturity. However, the application to an actual project is higher if the time of issuance and maturity are considered to vary.

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

Abbreviations

APV	Adjusted Present Value
BW	Bond with Warrant
BSM	Black & Scholes Model
CAPEX	Capital Expenditure
CB	Convertible Bond

CF	Corporate Financing
DCF	Discounted Cash Flow
MCS	Monte-Carlo Simulation
MF	Mezzanine Financing
NPV	Net Present Value
PF	Project Financing
PKP	Public Knowledge Project
SPC	Special Purpose Company
WACC	Weighted Average Cost of Capital

References

1. Raymond, V. International Investment and International Trade in the Product Cycle. *Q. J. Econ.* **1966**, *80*, 190–207.
2. D'Costa, A.P. *The Global Restricting of the Steel Industry*; Routledge: Abingdon, UK, 1999.
3. Milton, H.; Artur, R. The theory of capital structure. *J. Financ.* **1991**, *46*, 297–355.
4. Sandalkhan, B. Optimal capital structure model for BOT Power Projects in Turkey. *J. Constr. Eng. Manag.* **2003**, *129*. [CrossRef]
5. Miles, J.A.; Ezzell, J.R. The Weighted Average Cost of Capital, Perfect Capital Markets, and Project Life: A Clarification. *J. Financ. Q. Anal.* **1980**, *15*, 719–730. [CrossRef]
6. Don, H. *Types and Sources of Financing for Start-Up Business.* Available online: https://www.extension.iastate.edu/agdm/wholefarm/html/c5-92.html (accessed on 3 May 2018).
7. Credit Suisse Group AG. *Rating Report*; Credit Suisse Group AG: Zürich, Switzerland, 2017.
8. McNair, D. *EPC Contracts in the Process Plant Sector*; DLA Piper: Sydney, Australia, 2012.
9. Salman, S. *Optimal Capital Structure and PF*; Elsevier: Amsterdam, The Netherlands, 1987.
10. John, T. *Optimality of PF: Theory and Empirical Implications in Finance and Accounting*; Springer: New York, NY, USA, 1991.
11. Esty, B.C. *An Overview of Project Finance and Infrastructure Finance*; Harvard Business School: Boston, MA, USA, 2014.
12. Enzo, S. PF in the energy industry: new-debt based financing models. *Int. Bus. Res.* **2012**, *5*. [CrossRef]
13. Inselbag, I.; Kaufold, H. Two DCF approaches for valuing companies under alternative financing strategies (and how to choose between them). *J. Appl. Corp. Financ.* **1997**, *10*, 114–122. [CrossRef]
14. Esty, B.C. *The Economic Motivations for Using Project Finance*; Harvard Business School: Boston, MA, USA, 2002.
15. Bean, L. MF: Is It for You? *J. Corp. Acc. Financ.* **2008**, *19*, 33–35. [CrossRef]
16. Kent, G.; Amin, A. Flexible Financing with Mezzanine Debt. *J. Bus. Strateg.* **1990**, *11*, 54–56.
17. Management Institute. *Understanding and Utilization of MF KB Daily Knowledge Vitamin*; Management Institute: London, UK, 2012.
18. Lee, C.W.; Zhong, J. Financing and risk management of renewable energy projects with a hybrid bond. *Renew. Energy* **2015**, *75*, 779–787. [CrossRef]
19. Song, J.-S. A Study on the Search for a New Business Structure in the Housing Development Financial Market Domestic. Master's Thesis, Konkuk University Graduate School of Real Estate, Seoul, Korea, 2007.
20. Kang, J.-H. *Corporate Value Evaluation Practice*; Nanyang Technological University: Singapore, 2016.
21. Kim, D.-H. The Alternative Project Model for Steel. Master's Thesis, POSTECH, Pohang, Korea, 2016.
22. Dong, H.L. Development of Profitability Prediction Model of Steel Plant Project Using NPV at Risk (Net Present Value-Risk). Master's Thesis, Graduate School of Engineering, POSTECH, Pohang, Korea, 2015.
23. Kim, Y.-G.; Lee, E.-B. A Probabilistic Alternative Approach to Optimal Project Profitability Based on the Value-at-Risk. *Sustainability* **2018**, *10*, 747. [CrossRef]
24. Kim, Y.-G.; Shin, G.-Y.; Ahn, J.S.; Lee, E.-B. Probabilistic Cash Flow-Based Optimal Investment Timing Using Two-Color Rainbow Options Valuation for Economic Sustainability Appraisement. *Sustainability* **2017**, *9*, 1781. [CrossRef]
25. Jung, Y.-K. Capitalization of Borrowing Costs and International Convergence. Available online: http://www.riss.kr/search/detail/DetailView.do?p_mat_type=1a0202e37d52c72d&control_no=530448677e221782ffe0bdc3ef48d419 (accessed on 17 April 2018).

26. Song, H.-N. An Empirical Analysis of Domestic Public Offering Convertible Bonds Using Option Pricing Models. Master's Thesis, Ajou University, Suwon, Korea, 2017.
27. Park, K.; Wang, S. An Empirical Study on the Valuation of Korean Convertible Bonds Using the Option Pricing Model. Master's Thesis, Sogang University Graduate School, Seoul, Korea, 2000.
28. Black, F.; Scholes, M. The Pricing of Options and Corporate Liabilities. *J. Political Econ.* **1973**, *81*, 637–654. [CrossRef]
29. Silbernagel, C. "MF", Bond Capital, Canada. 2012. Available online: https://www.bondcapital.ca/wp-content/uploads/pdfs/what-is-mezzanine-finance.pdf (accessed on 30 March 2018).

Efficiency and Risk in Sustaining China's Food Production and Security: Evidence from Micro-Level Panel Data Analysis of Japonica Rice Production

Chengjun Wang [1,2,*], Zhaoyong Zhang [3,*] and Ximin Fei [2]

[1] Faculty of Economics and Management, Zhejiang A&F University, Lin'an, Zhejiang 311300, China
[2] Faculty of Management, Zhejiang University, Hangzhou, Zhejiang 310082, China; helenfxm@sina.com
[3] School of Business & Law, Edith Cowan University, Joondalup WA6027, Australia; zhaoyong.zhang@ecu.edu.au
[*] Correspondence: cjwang_77@126.com (C.W.); zhaoyong.zhang@ecu.edu.au (Z.Z.)

Abstract: Sustainable food production and food security are always challenging issues in China. This paper constructs a multi-element two-level constant-elasticity-of-substitution (CES) model to assess technological progress in, and its contribution to, japonica rice production in China. The results show that the speed of technological progress in the production of japonica rice on average was 0.44% per annum in 1985–2013, and technological progress has contributed significantly to the growth of japonica rice production in China. Robustness checks show that the results appear to be sensitive to which sub-sample is used. Labour and some other inputs are found to be significant but negative, especially during the middle sampling period of 1994–2006 and in eastern and western regions. This has important policy implications on the impact of rural-to-urban migration and farmers' human development.

Keywords: China's food policy; sustainable food security system; japonica rice production; two-level CES function; technological progress

1. Introduction

In recent decades rapid population growth and urbanization have made food security one of the most important global issues. According to Alexandratos and Bruinsma [1] and the United Nations (UN) [2], the world population is predicted to grow from 6.9 billion in 2010 to 8.6 billion in 2030, 9.8 billion in 2050 and 11.2 billion in 2100, while food demand is predicted to increase by 50% by 2030 and 70% by 2050. The main challenge facing the agricultural sector is how to sustain the global food production system to ensure food security and meet the projected increase in food demand, given rising resource constraints for agricultural production, yields slowing down and climate change (see Popp et al. [3,4]). Ultimately, global food production capacity will be constrained by the amount of farmland and water resources available and suitable for crop production and by biophysical limits on crop growth (Van Ittersuma et al. [5]). For instance, Popp et al. [4] maintain that there are growing opportunities and demands for the use of biomass to provide additional renewables, energy for heat, power and fuel, pharmaceuticals and green chemical feedstock. Burchi and De Muro [6] propose a capability-based analysis of food security by highlighting the importance of factors such as participation in household decision making and empowerment, and distinguishing between the capability to be food secure and functioning of food security. Studies have shown that average farm yields have reached 75–90% of the yield potential ceiling (Cassman et al. [7]; Grassini et al. [8]), and it becomes very difficult to further raise yields and have significant breakthroughs in the

genetic improvement of photosynthesis or drought tolerance (Fischer and Edmeades [9]; Hall and Richards [10]).

In China, sustainable food production and food security have always been listed as a top policy priority. It is widely believed that food security is related to national stability, independence and social stability. Hence, achieving food security and safety and maintaining the stability of domestic food production have been the major focus of Chinese agricultural policy (See Peng et al. [11]; Gautam and Yu [12]). Feeding one fifth of the world's population with rising incomes from less than a tenth of its arable land and freshwater is posing significant challenges to Chinese policy makers, while China's domestic food production and food security status will have large effects on the global food stability and security. This issue becomes even more prominent when considering recent rapid urbanization and industrialization, decline in arable farmland, a rapidly ageing urban population and other resource constraints in China. Several studies have shown that rice and maize yields appear to have plateaued or become stagnant in China and other major grain production areas (Brisson et al. [13]); Cassman et al. [7]); Van Wart et al. [14]), and farmland in China declined by about 11% between 1978 and 2006 (Fleming [15]). Some suggest that China should import more land-intensive food to reduce pressure on its already strained land and water resources, and others express fear over the fact that China's long-term dependency on foreign exports will fuel food-price increase and worsen the food insecurity status in many resource-poor countries (Wang et al. [16]; Liang et al. [17]; Ghose [18]). All these factors cast doubt on the possibility of continuing to rely on the traditional way of farming, such as increasing inputs of labour and expanding cultivated land and irrigation. Many advocate the need for "sustainable intensification" of agricultural production focusing on increasing production efficiency while minimizing economic and environmental costs (Godfray et al. [19]; Garnett et al. [20]). Thus, it has become increasingly important to link agricultural production efficiency and productivity with the sustainability of agriculture and the food security system, especially in the case of China.

Rice is one of the most important food staples in China, accounting for about 20% of the total crop area harvested (Chen et al. [21]) and about 65% of total staples consumption (Peng et al. [11]). Rice can be categorized into two main types, indica and japonica. With the rapid rise in incomes and private wealth as well as a higher standard of living in China, rice consumers have become increasingly concerned about the quality of the rice they consume. As compared to indica and other rice varieties, japonica rice is mostly preferred and considered as premium quality rice in China, and demand for japonica rice has been rising rapidly in recent years. However, continuous expansion of japonica rice production has been constrained by declining arable land, labor and capital as well as other natural resources, and it will increasingly become difficult to continue relying on resource inputs to expand production. This implies that, in order for China to maintain its sustainable farming system and ensure the stability and security of the food supply, technological progress will be the major driving force for increasing agricultural productivity and promoting agriculture development. It is therefore essential to ask the question of to what extent China's japonica rice production growth is due to technical efficiency and technological progress. According to the European Union (EU) [22], total factor productivity (TFP) is the main indicator to measure changes in productivity and TFP growth is defined as the ratio between the change in production volumes over a considered period and the corresponding change in inputs (or factors) used to produce them, and hence measures the growth in productivity over a given time span. An increase in TFP reflects a gain in output quantity which is not originating from an increase in input use.

The purpose of this paper is to adopt an improved multi-element two-level constant-elasticity-of-substitution (CES) production function to measure the technical efficiency and technological progress and its contribution to japonica rice production using cross-provincial panel data from China from 1985 to 2013. In particular, we intend to assess the rate of scientific and technological progress and to measure its contribution rate in japonica rice production in China by using panel data at provincial level. We will also shed light on the impact of the relevant agricultural policies on technological progress and contribution rate across the regions, and draw

policy implications for how to sustain japonica rice production in China. This paper is among the first to assess the contribution of science and technological advances to japonica rice production in China using an improved multi-factor two-level CES production function and provincial panel data across China. This study has important implications for China's long-term agricultural policy and development strategy.

The rest of the paper is structured as follows: Section 2 provides a brief literature review. Section 3 discusses the analytical framework and data sets employed in this study. In Section 4 we analyze the estimation results. Finally, Section 5 concludes.

2. Literature Review

Numerous studies empirically examine the effects of science and technological progress on agricultural production, and many advocate that the sustainable intensification of agricultural production focusing on technological progress and increasing production efficiency is the key to ensuring the sustainability of agriculture and food security (Tilman [23]; Godfray et al. [19]; Garnett et al. [20]; EU [20]). Shankar and Thirtle [24] estimate a damage-control specification and a conventional Cobb–Douglas production function to assess pesticide productivity and transgenic cotton technology in South Africa, and find evidence that the main potential contribution of the new technology is to enable farmers to realise lost productivity resulting from under-use. Popp et al. [3] provide a thorough review of pesticide-related productivity. Crost et al. [25] estimate a production function using a fixed-effects model to control for selection bias, and report that efficient farmers in India adopt Bt cotton at a higher rate than their less-efficient peers. Neumann et al. [26] explore the yield gap of global grain production by combining an econometric approach with spatial analysis, and report that the rapid increase in global food supplies over the past 50 years was mainly due to enhanced agricultural intensification and introduction of new technologies. Most recently, Ma et al. [27] estimate a standard production model and a damage-control model to assess the effects of Bt cotton use on productivity in Pakistan, and report that the yield-enhancing inputs (fertiliser and labour) have a strong effect on cotton productivity and all Bt variables reduce yield losses alongside insecticide use. In addition, Binswanger-Mkhize and Savastano [28] report that rapid urbanization and population growth have put farming systems under stress in Sub-Saharan African countries during the past two decades, and fallow areas have disappeared, but cropping intensities remain very low. They also find that the use of organic and chemical fertilizers is too low to maintain soil fertility and the process of intensification across these countries appears to have been weak. There has been little evidence of Boserupian agricultural intensification with respect to cropping intensity, area farmed, or irrigation. Sheahan and Barrett [29] also report that, although the use of inorganic fertilizer and agro-chemicals remains relatively low on average in Sub-Saharan Africa, the use rates are actually quite high in some countries and regions within countries, which may relate to the fact that input use is no higher on cash-crop plots than on those cultivated mainly with staple cereals. Villano et al. [30] employ a stochastic production frontier framework to examine the impact of modern rice technologies on farm productivity while disentangling technology gaps from managerial gaps in the Philippines, and find that the adoption of certified seeds has a significant and positive impact on productivity, efficiency and net income in rice farming. On the other hand, using a new micro-level dataset from four African countries, McCullough [31] finds that individuals participating in agriculture tend to devote fewer hours to agriculture, on average, than individuals participating in other sectors, and the returns to an hour of labor supplied outside of agriculture are about 1.4 times as high as returns to an hour of agricultural labor on average in these countries.

Recently, there have also been some country studies of several Asian countries that find similar results. Gautam and Yu [12] study comparatively agricultural TFP growth in China and India, and report that this is mainly propelled by technological advances in these two countries but efficiency has been stagnant or has even deteriorated. They conclude that, faced with similar challenges of limited resources and growing demand, improving productivity will be the only way to meet long-term food

security in China and India. Using a front line demonstration on potatoes in India, Mishra et al. [32] report that technological progress and the adoption of improved agricultural technologies are essential to reverse the trend of wide extension gaps for potato output. Iliyasu et al. [33] employ the Malmquist productivity index method to analyze the efficiency of technological change in Malaysia, and find that technological change and efficiency changes are important contributors to TFP growth in aquaculture production. Gao and Song [34] use DEA, Moran's I and the Theil index to measure technical efficiency in China's food production using provincial panel data spanning from 1978 to 2012, and find that the improvement of efficiency in grain production is mainly driven by the rising contribution rate of agricultural science and technology. They further confirm that since 2000, technological progress has gradually replaced the physical inputs of human capital and resources in driving China's food production growth (Gao and Song [34]). Similar findings are also reported in several recent studies. For instance, Liu et al. [35] use the logarithmic mean weight divisive method (LMDI) to explore the factors that have dominant impacts on grain production using data from 347 counties in China's Huang-Huai-Hai region, and conclude that technological progress, cultivated land balance, and the optimization of the regional planting structure should be enhanced to sustain the growth trend of grain production in this region. Yang et al. [36] apply the Malmquist–DEA model to study China's TFP growth in grain production, and report that there was a continuing decline in the TFP index at an annual rate of 0.7% even during the so called "seven consecutive years harvest period", which is largely due to the significant drop in the average annual index of technological progress. Using provincial panel data in 1985–2010 from China with an improved CES production function, Jiang et al. [37] report that the contribution made by science and technology to grain production accounts for about 52% of China's grain production increase during the period in 1985–2010, a rate closer to that in the developed countries. Liu et al. [38] report similar findings using data from Hebei province. EU [22] reports that productivity in the EU has increased over time, with a growth rate surpassed by 1% per year between 1995 and 2005, and around 0.8% between 2005 and 2015.

There are a few studies focusing on technical efficiency in rice production, especially japonica rice production in China given its importance in sustaining the country's food security. These studies (for instance, Chen et al. [39]; Lin et al. [40]; Xu et al. [41]) mostly focus more on the development trends, seed-breeding skills and technical aspects of rice production, and a few on the technical efficiency and contribution of scientific and technological progress in rice production (Tian et al. [42]; Liu et al. [43]). For instance, Chen et al. [39] provide a comprehensive review of the breeding research progress for China's super high-yield japonica rice, and conclude that cultivating and promoting super high-yield hybrid rice is still considered to be one of the most effective strategies for improving rice yield in northern China. Lin et al. [40] compare the yield and main quality traits of japonica rice with that of other rice varieties including hybrid japonica rice, and find that japonica rice has some obvious advantages in terms of yield and quality. Using Solow's growth model, Tian et al. [42] estimate the contribution of technological progress to China's rice production and report that technological progress contributed over 20% to China's rice production growth during the period in 1978–2008. Liu et al. [43] confirm that technological progress plays a much more important role than China's climate policy and other factors in driving rice production. This paper will be among the first few to assess the contribution of science and technological advances to japonica rice production in China using an improved multi-factor two-level CES production function and provincial panel data.

3. Model Specification and Data Sources

3.1. Model Specification

The process of japonica rice production involves the adoption of different levels of farming technology, equipment and inputs. In order to appropriately measure technological change and its contribution to japonica rice production, we consider three-factor inputs including capital goods, labor and others such as fertilizer, seeds etc., and develop an improved two-level nested CES

production function a la Sato [44] and Henningsen and Henningsen [45] in this study. Recently, León-Ledesma et al. [46] investigate if a simultaneous identification of the capital–labor substitution elasticity and the direction of technical change is feasible, and find that jointly modeling the production function and first-order conditions is superior to single-equation approaches. In order to reduce the complexity of measuring for excessive variables, all the variables in this study are measured per unit of land. Let Z_{it} be the first level production function value of japonica rice production in province i in year t, and Y_{it} denotes the yield of japonica rice produced by three inputs production in province i in year t. The two-level CES function is then specified as follows,

$$Z_{it} = [\omega K_{it}^{-\theta_1} + (1 - \omega) L_{it}^{-\theta_1}]^{\frac{1}{\theta_1}}, \tag{1}$$

$$Y_{it} = Ae^{rt}[\psi Z_{it}^{-\theta} + (1 - \psi) F_{it}^{-\theta}]^{-\frac{m}{\theta}}. \tag{2}$$

where K_{it} denotes the input of capital equipment and mechanical power in japonica rice production in province i in year t; L_{it} is the input of labor in province i in year t; F_{it} denotes the input of fertilizer, seed and other material inputs in province i in year t; ω and ψ denote the distribution parameter related to input i; θ and θ_1 is the substitution between inputs; γ is the efficiency parameter representing the speed of annual technical progress of japonica rice production; and A denotes comprehensive benefit index, including the impact on the yield of japonica rice due to the improvement of the input quality and the management level. Mechanical elements mainly include machinery and livestock, and labor elements included self-casting workers and employees, indicating the amount of labor in the production per unit of planting acreage. Hence, $Ae^{\gamma t}$ denotes the multiple of output increase due to the improvement of technology level, and m denotes the scale factor, indicating non-constant return when $m \neq 1$.

As Y_{it} is a CES function in Z_{it} and Z_{it} in turn is a CES function of K and L, following Sato [44] we call Y_{it} a two-level CES function in K and L. Now we take the natural log of Equation (2) and use the second-order Taylor series expansions at $\theta = 0$ to obtain:

$$\ln Y_{it} = \ln A + \gamma t + \psi m \ln Z_{it} + (1 - \psi) m \ln F_{it} - \frac{1}{2} \theta m \psi (1 - \psi) \left(\ln \frac{Z_{it}}{F_{it}} \right)^2. \tag{3}$$

Similarly, for Equation (1) we have the Taylor expanded form as follows,

$$\ln Z_{it} = \omega \ln K_{it} + (1 - \omega) \ln L_{it} - \frac{1}{2} \theta_1 \omega (1 - \omega) \left(\ln \frac{K_{it}}{L_{it}} \right)^2. \tag{4}$$

Combining Equations (3) and (4), we obtain Equation (5),

$$\begin{aligned} \ln Y_{it} = {} & \ln A + \gamma t + m\omega\psi \ln K_{it} + m(1 - \omega)(1 - \psi) \ln L_{it} + m(1 - \psi) \ln F_{it} \\ & - \frac{1}{2} m \theta_1 \omega (1 - \omega) \psi \left(\ln \frac{K_{it}}{L_{it}} \right)^2 - \frac{1}{2} m \theta \psi (1 - \psi) \left(\ln \frac{K_{it}}{F_{it}} \right)^2. \end{aligned} \tag{5}$$

We use a_j as the composite parameter to represent the different combinations of distribution and substitution parameters and add an error term to Equation (5). We then have the two-level or nested CES production function for japonica rice specified as follows,

$$\ln Y_{it} = \alpha_0 + \alpha_1 t + \alpha_2 \ln K_{it} + \alpha_3 \ln L_{it} + \alpha_4 \ln F_{it} + \alpha_5 \left(\ln \frac{K_{it}}{L_{it}} \right)^2 + \alpha_6 \left(\ln \frac{K_{it}}{F_{it}} \right)^2 + \varepsilon_{it}, \tag{6}$$

From Equation (6) we have 7 parameters, a_j (j = 0,1, . . . 6), for 7 primitives, A, m, γ, θ, $\theta_1\psi$, ω, from Equation (5). The contribution rate of technical progress to the yield growth of japonica rice can be calculated based on the following equation,

$$\boldsymbol{R}_{\text{it}} = \frac{\gamma_{it}}{y_{it}} \times 100\% = \frac{\gamma_{it}}{\sqrt{\frac{Y_{it}}{Y_{ib}} - 1}} \times 100\%. \tag{7}$$

In formula (7), $\boldsymbol{R}_{it}$ is the contribution rate of technological progress in province i at year t; γ_{it} denotes the speed of technology progress in province i at year t; y_{it} denotes the average growth rate of japonica rice yield in province i at year t; Y_{it} denotes the japonica rice yield during the reporting period from province i; and Y_{ib} represents the japonica rice yield during the based period from province i.

3.2. Data Description

All data used in this study were obtained from China's National Bureau of Statistics: Statistics on Agricultural Costs and Returns, various issues; Sixty Year Agricultural Statistics; Statistical Yearbook of China, various issues; Yearbook of Agricultural Statistics, various issues; as well as several official government websites. All datasets were collected over the period from 1985 to 2013. There are several reasons for choosing this sample period and, among others, data availability is one of the most important considerations. This period also witnesses the rapid growth of japonica rice production in China. This study focuses on 12 provinces in China, including Heilongjiang, Jilin, Liaoning, Hebei, Jiangsu, Zhejiang, Anhui, Shandong, Henan, Hubei, Yunnan, Ningxia, and so on. These regions are not only the major farming areas for China's grain production, but also for japonica rice production, with the latter accounting for more than 60% of China's total japonica rice production. In order to reduce the unnecessary complexity of the measurement issue, all the input and output variables in this study are measured in terms of per unit of farming land. Finally, we take 2000 as the base year for those variables proxied with the form of index numbers. Table 1 presents the key descriptive statistics of our sample. To ensure all the series are stationary, we have conducted the panel unit root test by applying the Levin–Lin–Chu (LLC) test and the Fisher–augmented Dickey–Fuller (ADF) test. The results (available upon request from the authors) confirm that all the series are stationary.

Table 1. Descriptive statistical information.

Variable	Average	Maximum Value	Minimum Value	Standard Deviation	Observations
$ln(Y_{it})$	6.114	6.425	5.606	0.169	348
$ln(K_{it})$	3.045	5.465	−0.117	1.228	348
$ln(L_{it})$	2.638	3.740	1.284	0.509	348
$ln(F_{it})$	3.285	4.697	2.209	0.487	348
$ln^2(K_{it}/L_{it})$	2.833	17.266	2.12×10^{-6}	3.350	348
$ln^2(K_{it}/F_{it})$	2.180	19.661	9.37×10^{-6}	2.959	348

Note: The unit of machinery, livestock and other mechanical power cost (C) is (yuan/mu); the unit of the amount of labor (L) is (one working day/mu); the unit of seeds, fertilizer and other material cost (F) is (yuan/mu); the unit of Japonica rice yield (G) is (kg/mu).

4. Empirical Results and Analysis

To investigate technological change and its contribution to japonica rice production, we estimate a range of panel data models for different sub-sample periods and regions from 1985 to 2013. The models are estimated with panel least square (PLS), mixed effects (ME), fixed effects (FE) with both province and year, and random effects (RE). Using a panel FE model is advantageous because FE can control for unobserved time invariant region-specific effects. We also estimate random effects to capture the influence of unobserved factors that may produce heterogeneity across the countries. We conduct the

F-test (Wald test) and the Hausman test to determine the significance and choice of the models. Table 2 reports the estimation results of the random effect, fixed effect and mixed effect model.

Table 2. Estimation results of the constant-elasticity-of-substitution (CES) production function for japonica rice.

Coefficient	RE	FE	ME
a_0	5.740 ***	5.829 ***	5.079 ***
	(50.815)	(49.381)	(68.708)
a_1	0.004 **	0.004	0.003 **
	(2.055)	(1.622)	(2.076)
a_2	0.061 ***	0.061 ***	0.108 ***
	(4.522)	(4.374)	(7.737)
a_3	−0.011 ***	−0.031 ***	−0.117 ***
	(−3.73)	(−9.830)	(−6.052)
a_4	0.048 **	0.040 *	0.109 ***
	(2.373)	(1.944)	(5.297)
a_5	−0.002 ***	−0.003 ***	−0.004 ***
	(−7.010)	(−9.860)	(−5.897)
a_6	0.0004 *	0.002	0.008
	(1.971)	(0.359)	(1.303)
Observations	348	348	348
Adjusted R^2	0.591	0.780	0.571
F	84.470 ***	73.316 ***	77.931 ***

Note: RE refers to random effect, FE to fixed effect and ME to mixed effect. ***, **, * indicate the level of significance at the 1%, 5% and 10%, respectively.

As it can be seen in Table 2, the coefficients of capital equipment and mechanical power, fertilizer and the squared ratio of capital and fertilizer are all positive and significant, with the exception of the squared capital fertilizer ratio in fixed-effect and mixed-effects estimations. The result suggests that a 1% increase in mechanical power input will lead to a 0.061% increase in japonica yield, while with a 1% increase in fertilizer input the japonica yield rises by 0.048%. It is interesting to note that the labor input coefficient is negative and significant at the 1% level in all the estimations, implying that an increase in labor input will reduce the yield of japonica rice. This finding seems contrary to our causal observation. One possible explanation is that the production of japonica rice in China is mostly managed by small-scale family farming units. Due to its scarce arable land and the abundance of manpower, farming in China has always been very labor-intensive. Hence, it is not surprising to note the declining labor productivity in the production of japonica rice. Another possible explanation is that, as small-scale family farmers are more risk averse, they tend to be reluctant to adopt quality seeds and new agricultural technologies in the production of japonica rice. Although intensive farming practices can generate relatively high yield, over-reliance on increasing inputs of labor in farming can be counter-productive. The finding is consistent with Mishra et al. [32] and Liu et al. [43]. This, together with the low estimate of the capital terms coefficient, seems to reaffirm Solow's [47] conclusion that accumulation of capital and an increase in the labor participation rate had a relatively minor effect on growth, and cannot explain all the growth of japonica rice in the past few decades. It also suggests that traditional farming practices with a sole reliance on increasing inputs of capital and labor will not be sustainable in the long term, and emphasis should be given to continuous improvement in seed quality, farmer skill, technological progress and the adoption of technologies for sustainable farming systems.

Using the estimates in Table 2, we will be able to identify the primitives in the multiple element two-level CES function and determine technological progress in the production of japonica rice. Table 3 reports the results. The results show that the speed of technological progress in the production of japonica rice on average is 0.44% per annum in the sample period in 1985–2013. According to Chinese statistics, the average yield of japonica rice per unit of farming land in the 12 provinces concerned was 297.91 kg in 1985, and rose to 436.76 kg in 2013, with an average annual growth rate of about 1.33%.

This study finds that technical progress has contributed significantly to the growth of japonica rice production in China, with a rate accounting for 33.13% from 1985 to 2013.

Table 3. Identified distribution and substitution parameters and technical progress in China's japonica rice production.

Primitive Parameters	A	γ	ω	θ_1	ψ	θ	m
Values	312.19	0.004	1.22	−0.30	0.51	−0.03	0.10
Contribution rate of technological progress						R_{it}: 33.13%	

Note: γ_{it} refers to the technical progress speed; R_{it} refer to the contribution rate of technical progress.

4.1. The Dynamics of Japonica Rice Output and Technological Progress

Since the late 1970s, China's agricultural sector and farming system have undergone a series of reforms, aiming at improving the efficiency of resources allocation and promoting agricultural and rural production technology productivity. To assess the dynamic impacts of various reform measures and policy changes on the growth of japonica rice production, we divide the whole sample period into three: the first period covers 1985 to 1993 to catch the effects of the primary reform in agriculture during the early reform era; the second period spans from 1994 to 2006 to reflect the impact of implementing the new policy measures on current agricultural and rural economic development on 5 November 1993; and the third sub-sample period ranges from 2007 to 2013 to allow for the effects of adopting various fiscal subsidies measures in the agricultural sector since 2006. We re-estimate the models using the new sub-sample data to measure the technological progress and its contribution rate over the different sub-sample periods. Table 4 reports the estimation results.

Table 4. Estimation results for different sub-sample periods.

Coefficient	RE 1985–1993	RE 1994–2006	FE 2007–2013
a_0	4.917 ***	5.199 ***	5.946 ***
	(5.567)	(4.254)	(19.787)
a_1	0.007 ***	0.002 ***	0.013 ***
	(6.262)	(8.345)	(2.977)
a_2	0.263 ***	0.101 ***	0.055 **
	(7.983)	(4.587)	(2.145)
a_3	0.080 *	0.071 **	−0.038 ***
	(1.700)	(2.409)	(−3.470)
a_4	0.057 ***	0.124 ***	−0.044 ***
	(4.196)	(3.477)	(−3.934)
a_5	0.035 ***	0.008 ***	−0.007 ***
	(2.737)	(6.341)	(−3.945)
a_6	0.012 ***	−0.031 ***	−0.008 **
	(5.032)	(−3.451)	(−2.811)
Obs.	108	156	84
R^2	0.424	0.442	0.890
Adjusted R^2	0.389	0.420	0.862
F	12.366 ***	19.677 ***	31.383 ***
Fixed effect F test	18.993 ***	19.795 ***	12.161 ***
Hausman test	0.000	4.238	20.902 ***

Note: ***, ** and * indicate the level of significance at the 1%, 5% and 10%, respectively.

It can be seen from Table 4 that all the coefficients are statistically significant over the three sub-periods. The estimated coefficient of capital equipment and mechanical power has a positive sign for all the sub-sample periods while all of the rest of the variables have positive signs only for the first two sub-periods except the capital–fertilizer ratio. By contrast with the results for the whole

sample period, the coefficients of labour and fertilizer inputs as well as the squared capital-labour and capital-fertilizer ratios become negative during the third stage of development from 2007 to 2013. The findings reaffirm that over-reliance on increasing physical inputs in japonica rice production cannot be sustained in the long term, and can be counter-productive.

Using the estimates in Table 4, we then identify the primitives in the multiple element two-level CES function and determine technological progress in the production of japonica rice during each of the sub-sample periods. We report the results in Table 5. As can be seen in Table 5, the speed of technological progress in the production of japonica rice varies across the sub-periods, with a V-shape over time. The period 2007–2013 witnessed the fastest progress in technology in China's japonica rice production, followed by the 1985–1993 period, while the period 1994–2006 was the worst when China experienced negative growth in technology. The finding is consistent with that in Zhang [48] for the national economy. There are several reasons to explain this pattern of technological progress. First, China's agricultural reform started in the late 1970s. The initial success of China's agricultural reform was remarkable and led to a rapid increase in agricultural productivity and growth through the early 1990s, a result of which has been the rapid progress in technology in China's japonica rice production during the period in 1985–1993. Second, with the increasing demand for quality and variety of agricultural products and rapid urbanization in the 1990s, the Chinese agricultural sector experienced a sharp decline in both arable farmland and rural labour force. The popularity among and interest of farmers in adopting new rice farming technology was also affected, a result of which has been the decrease both in rice output and technological progress in japonica rice production in the period 1994–2006. Finally, since the early 2000s, China has implemented its New Blueprint for Rural Reform and Development, including a series of new agricultural policies and favourable measures, to build a new countryside. These policies and measures have stimulated farmers to adopt new agricultural technologies and improve efficiency, and also reduced the costs of using quality rice seeds and production. This led to the rapid expansion of rice output and technological progress in japonica rice production in the period 2007–2013.

Table 5. Identified distribution and substitution parameters and technical progress.

Primitive Parameters	Period	A	γ	ω	θ_1	ψ	θ	m
	1985–1993	137.43	0.007	0.76	−0.97	0.84	−0.40	−0.4
Values	1994–2006	181.87	−0.003	0.59	−0.49	0.58	0.87	0.29
	2007–2013	385.20	0.013	3.00	−0.12	−1.00	0.40	−0.02
	1985–1993	34.31%						
Contribution rate of technological progress R_{it}	1994–2006	26.9%						
	2007–2013	147.16%						

4.2. Technological Progress Across the Regions

Given the diverse natural and socio-economic environment across China, it is believed that region-specific factors would also play an important role in japonica rice production. To further assess the region-specific impacts on the technological progress and japonica rice production nexus, we divide the 12 sample major rice-production provinces into three regions, i.e., the eastern, central and western regions. The eastern region includes the provinces of Liaoning, Hebei, Jiangsu, Zhejiang, and Shandong; the central includes Heilongjiang, Jilin, Anhui, Henan, and Hubei; and the western region includes Yunnan and Ningxia. It is worth noting that after re-grouping our sample, the panel dataset has been transformed into one with a long-time series and short cross-sectional data, i.e., panel data with large T and small N. The typical complications of TSCS panel data are heteroskedasticity and autocorrelation. The conventional way to deal with heteroskedasticity and correlated errors is to use the feasible generalized least squares (FGLS) technique, but FGLS has been found to be less efficient and tends to underestimate standard errors (Beck and Katz [49]). In this study we use the panel-corrected standard error (PCSE) to address panel heteroskedasticity and include a lagged

dependent variable in the model to address serial correlation. Table 6 reports the estimation results for the three sub-regions. We further identify the primitives in the multiple element two-level CES function and determine the technical progress in the production of japonica rice for each sub-region. We report the results in Table 7.

Table 6. Estimation results at the regional level.

Coefficient	Eastern Region	Central Region	Western Region
a_0	5.881 ***	5.538 ***	7.066 ***
	(34.290)	(4.164)	(24.480)
a_1	0.007	0.003 **	0.0003 *
	(1.762)	(2.270)	(1.871)
a_2	0.051 *	0.097 ***	0.017 ***
	(1.953)	(5.12)	(3.417)
a_3	−0.079 *	0.008 **	−0.249 ***
	(−1.67)	(2.051)	(−4.31)
a_4	0.086 **	0.049 *	−0.015 **
	(2.05)	(1.920)	(−2.237)
a_5	−0.014 ***	−0.005 *	−0.028 ***
	(−3.35)	(−1.830)	(−3.630)
a_6	−0.009 *	0.002 **	0.021 **
	(−1.806)	(2.234)	(2.320)
Observations	145	145	58
R^2	0.926	0.924	0.936
wald	112.562	119.951	107.915
rho	0.385	0.358	0.249

Note: Eastern including Liaoning, Hebei, Jiangsu, Zhejiang, Shandong; Central including Heilongjiang, Jilin, Anhui, Henan, Hubei; Western including Yunnan, Ningxia. ***, **, * represents the significant level of 1%, 5% and 10% respectively.

As can be seen in Table 7 and also from Table 6, the eastern region has the fastest growth of technological progress in japonica rice production during the whole sample period from 1985 to 2013, with an average growth rate of 0.7% for technological progress and 43.75% for the technological contribution. This is followed by the central region, where the rate of technological progress and the contribution rate are 0.3% and 24%, respectively. The western region has the lowest technical progress rate and contribution rate with the former being 0.03% and the latter 3.37%, respectively. There are several possible reasons for these regional differences in technological progress and contribution rate. Among others, the well-established social and economic environment and favourable natural farming conditions in the eastern region, and to some extent also in the central region, are the major factors explaining why these regions have higher technological progress and contribution rates. By contrast, the economy of the western region is poorly developed largely due to its backward infrastructure and poor natural endowments. It is financially difficult for the farmers in the western region to adopt new agricultural technologies and use quality rice seeds and production, a result of which has been the lowest technological progress in japonica rice production.

Table 7. Identified distribution and substitution parameters and technical progress at regional level.

Primitive Parameters	Region	A	γ	ω	θ_1	ψ	θ	m
	Eastern	359.14	0.7%	−1.67	0.15	−0.50	−0.4	0.06
Values	Central	255.57	0.3%	0.93	1.5	0.68	−0.12	0.16
	Western	1181.41	0.03%	−0.09	2.80	0.89	2.3	−0.25
	Eastern				43.75%			
Contribution rate of technological progress R_{it}	Central				24%			
	Western				3.37%			

The results in Table 6 show that the input of capital equipment and mechanical power has a positive impact on japonica yields in the central region, and has a negative impact in other regions, but the effect of labour, the capital–labor ratio and the capital–fertilizer ratio all are negative in the eastern region, and labour, fertilizer and the capital–labour ratio in the western region are also negative. This can be explained by the decline in arable farmland and rising rural surplus labour due to rapid urbanization and industrialization, as well as inappropriate use of chemical fertilizers in farming. Accompanying the process of rapid urbanization and industrialization has been massive work-related rural-to-urban migration in China over the past few decades. As a result, there has been a shortage of capable farm labour in rural areas and farm work has to be undertaken by those left-behind, mostly the elderly, women and children. They will have to increase the time spent on farm work with low efficiency. One study reports that elderly people in rural households with rural-to-urban migrants tend to work 100–200 more hours per year on farm work than those households without migrants (Chang et al. [50]). On the other hand, those left behind in rural areas normally have low education, and lack modern farming techniques and knowledge. Driven by the desire for higher yields, it has become a common practice for farmers to overuse chemical fertilizers in farming. However, over-fertilization or the use of imbalanced fertilization does not always increase crop output and contribute to high rice yield. Instead, over-fertilization may deteriorate the soil conditions and cause a series of economic and environmental problems (Ju et al. [51]). This has clearly been an issue among policy makers. In contrast, our results show that the estimated coefficient of fertilizer for the central region is positive, which seems to suggest that over-fertilization does not exist in farming in the central region. The central region is China's main grain production base where farmers are well trained in how to utilize chemical fertilizer in farming. That explains the positive coefficient of fertilizer in the regression for japonica rice production.

5. Concluding Remarks

Sustainable food production and food security have been China's top policy priority. Due to recent rapid urbanization and industrialization, decline in arable farmland, a rapidly ageing urban population and other resource constraints in Chinese agriculture, technological progress will have to be the major driving force for increasing agricultural productivity and promoting agricultural development. In this study, we have constructed an improved multi-element two-level CES production model to assess the rate of technological progress and its contribution rate in japonica rice production in China by using cross-provincial panel data from China from 1985 to 2013. The results from the whole sample estimations show that the labor input coefficient is statistically significant but negative, suggesting that increases in labor input will reduce the yield of japonica rice. Furthermore, the results from the sub-sample estimations further show that the coefficients of labor and fertilizer inputs as well as the squared capital–labour and capital–fertilizer ratios become negative only during the third stage of development from 2007. The sub-regional studies further confirm that the coefficients of labour, capital–labor ratio and capital–fertilizer ratio all are negative in the eastern region, and the coefficients of labour, fertilizer and capital–labour ratio in the western region are also negative. Finally, the results show that the speed of technological progress in the production of japonica rice on average was 0.44% per annum in the sample period in 1985–2013, and technological progress has contributed significantly to the growth of japonica rice production in China, with a rate accounting for 33.13% from 1985 to 2013. The dynamics of the speed of technological progress in the production of japonica rice show a V-shape over time, with the period in 2007–2013 being the fastest and the period in 1994–2006 the worst. It is found that the eastern region had the fastest growth of technology progress in japonica rice production during the whole sample period, followed by the central region. The western region had the lowest technical progress and contribution rates.

The empirical findings have several policy implications. First, due to rapid urbanization and industrialization there has been massive work-related rural-to-urban migration in China over the past several decades. The production of japonica rice in China is mostly managed by small-scale family

farming units, and by those left-behind in rural areas, mostly the elderly, women and children, who are incapable of farm labour and have to devote increasing labour hours to farming. Hence it is not surprising to note declining labor productivity in the production of japonica rice with low efficiency. Another possible explanation is that as small-scale family farmers are more risk averse, they tend to be reluctant to adopt quality seeds and new agricultural technologies in the production of japonica rice. Then, although intensive farming practices can generate relatively high yields, an over-reliance on increasing input of labor in farming can be counter-productive. The over-fertilization or the use of imbalanced fertilization does not always increase crop output and contribute to high rice yields. Instead, over-fertilization may deteriorate the soil conditions and cause a series of economic and environmental problems. This, together with the low estimate of the capital terms coefficient, seems to reaffirm Solow's conclusion that an accumulation of capital and an increase in the labor participation rate had a relatively minor effect on growth, and cannot explain all the growth of japonica rice in the past few decades. It also suggests that traditional farming practices with a sole reliance on increasing inputs of capital and labor will not be sustainable in the long term, and an emphasis should be placed on continuous improvement in seed quality, farmer skill, technological progress and the adoption of technologies for sustainable farming systems.

Author Contributions: Wang, Zhang and Fei, conceived the paper, Fei analyzed the data; Wang and Zhang wrote and finalized the manuscript. All authors read and approved the final manuscript.

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

References

1. Alexandratos, N.; Bruinsma, J. *World Agriculture towards 2030/2050: The 2012 Revision*; ESA Working Paper No. 12-03; FAO: Rome, Italy, 2012.
2. United Nations. *The World Population Prospects: The 2017 Revision*; UN Department of Economic and Social Affairs: New York, NY, USA, 2017.
3. Popp, J.; Pető, K.; Nagy, J. Pesticide productivity and food security. A review. *Agron. Sustain. Dev.* **2013**, *33*, 243–255. [CrossRef]
4. Popp, J.; Lakner, Z.; Harangi-Rákos, M.; Fári, M. The effect of bioenergy expansion: Food, energy, and environment. *Renew. Sustain. Energy Rev.* **2014**, *32*, 559–578. [CrossRef]
5. Van Ittersuma, M.K.; Cassmanb, K.G.; Grassinib, P.; Wolfa, J.; Tittonell, P.; Hochmand, Z. Yield gap analysis with local to global relevance—A review. *Field Crops Res.* **2013**, *143*, 4–17. [CrossRef]
6. Burchi, F.; De Muro, P. From food availability to nutritional capabilities: Advancing food security analysis. *Food Policy* **2016**, *60*, 10–19. [CrossRef]
7. Cassman, K.G.; Dobermann, A.; Walters, D.T.; Yang, H.S. Meeting cereal demand while protecting natural resources and improving environmental quality. *Ann. Rev. Environ. Resour.* **3003**, *28*, 315–358. [CrossRef]
8. Grassini, P.; Yang, H.; Irmak, S.; Thorburn, J.; Burr, C.; Cassman, K.G. Highyield irrigated maize in the Western U.S. Corn Belt: II. Irrigation management and crop water productivity. *Field Crops Res.* **2011**, *120*, 133–144. [CrossRef]
9. Fischer, R.A.; Edmeades, G.O. Breeding and cereal yield progress. *Crop Sci.* **2010**, *50*, 85–98. [CrossRef]
10. Hall, A.J.; Richards, R.A. Prognosis for genetic improvement of yield potential and water-limited yield of major grain crops. *Field Crops Res.* **2013**, *143*, 18–33. [CrossRef]
11. Peng, S.B.; Tang, Q.Y.; Zou, Y.B. Current Status and Challenges of Rice Production in China. *Plant Prod. Sci.* **2009**, *12*, 3–8. [CrossRef]
12. Gautam, M.; Yu, B.X. Agricultural productivity growth and drivers: A comparative study of China and India. *China Agric. Econ. Rev.* **2015**, *7*, 573–600. [CrossRef]

13. Brisson, N.; Gate, P.; Couache, D.; Charmet, G.; Oury, F.X.; Huard, F. Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. *Field Crops Res.* **2010**, *119*, 201–212. [CrossRef]

14. Van Wart, J.; Kersebaum, K.C.; Peng, S.B.; Milner, M.; Cassman, K.G. Estimating crop yield potential at regional to national scales. *Field Crops Res.* **2013**, *143*, 34–43. [CrossRef]

15. Fleming, C. Food security, urbanization and social stability in China. *J. Agrar. Chang.* **2009**, *9*, 548–575.

16. Wang, H.X.; Zhang, M.H.; Cai, Y. Problems, Challenges, and Strategic Options of Grain Security in China. *Adv. Agron.* **2009**, *103*, 101–147.

17. Liang, F.; Li, Y.L.; Zhang, G.L.; Tan, M.G.; Lin, J.; Liu, W.; Li, Y.; Lu, W.W. Total and special arsenic levels in rice from China. *Food Addit. Contam. Part A* **2010**, *27*, 810–816. [CrossRef] [PubMed]

18. Ghose, B. Food security and food self- sufficiency in China: From past to 2050. *Food Energy Secur.* **2014**, *3*, 86–95. [CrossRef]

19. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food Security: The Challenge of Feeding 9 Billion People. *Science* **2010**, *327*, 812–818. [CrossRef] [PubMed]

20. Garnett, T.; Appleby, M.C.; Balmford, A.; Bateman, I.J.; Benton, T.G.; Bloomer, P.; Burlingame, B.; Dawkins, M.; Dolan, L.; Fraser, D.; et al. Sustainable intensification in agriculture: Premises and policies. *Science* **2013**, *341*, 33–34. [CrossRef] [PubMed]

21. Chen, M.; Shelton, A.; Ye, G.Y. Insect-Resistant Genetically Modified Rice in China: From Research to Commercialization. *Ann. Rev. Entomol.* **2011**, *56*, 81–101. [CrossRef] [PubMed]

22. European Commission. Productivity in EU agriculture—Slowly but steadily growing. Available online: https://ec.europa.eu/agriculture/sites/agriculture/files/markets-and-prices/market-briefs/pdf/10_en.pdf (accessed on 10 January 2018).

23. Tilman, D. Agricultural sustainability and intensive production practices. *Nature* **2002**, *418*, 671–677. [CrossRef] [PubMed]

24. Shankar, R.; Thirtle, C. Pesticide productivity and transgenic cotton technology: The South African smallholder case. *J. Agric. Econ.* **2005**, *56*, 97–116. [CrossRef]

25. Crost, B.; Shankar, B.; Bennett, R.; Morse, S. Bias from Farmer Self-Selection in Genetically Modified Crop Productivity Estimates: Evidence from Indian Data. *J. Agric. Econ.* **2007**, *58*, 24–36. [CrossRef]

26. Neumann, K.; Verburg, P.H.; Stehfest, E.; Müller, C. The yield gap of global grain production: A spatial analysis. *Agric. Syst.* **2010**, *45*, 316–326. [CrossRef]

27. Ma, X.L.; Smale, M.; Spielman, D.J.; Zambrano, P.; Nazli, H.; Zaidi, F. A Question of Integrity: Variants of Bt Cotton, Pesticides and Productivity in Pakistan. *J. Agric. Econ.* **2017**, *58*, 366–385. [CrossRef]

28. Binswanger-Mkhize, H.P.; Savastano, S. Agricultural intensification: The status in six African countries. *Food Policy* **2017**, *67*, 26–40. [CrossRef] [PubMed]

29. Sheahan, M.; Barrett, C.B. Ten striking facts about agricultural input use in Sub-Saharan Africa. *Food Policy* **2017**, *67*, 12–25. [CrossRef] [PubMed]

30. Villano, R.; Bravo-Ureta, B.; Solís, D.; Fleming, E. Modern Rice Technologies and Productivity in the Philippines: Disentangling Technology from Managerial Gaps. *J. Agric. Econ.* **2015**, *66*, 129–154. [CrossRef]

31. McCullough, E.B. Labor productivity and employment gaps in Sub-Saharan Africa. *Food Policy* **2017**, *67*, 133–152. [CrossRef] [PubMed]

32. Mishra, D.K.; Tailor, R.S.; Pathak, G.; Deshwal, A. Yield Gap Analysis of Blight Disease Management in Potato through Front Line Demonstration. *Indian Res. J. Ext. Edu.* **2007**, *48*, 82–84.

33. Iliyasu, A.; Mohamed, Z.; Hashim, M. Productivity growth, technical change and efficiency change of the Malaysian cage fish farming: An application of Malmquist Productivity Index approach. *Aquac. Int.* **2015**, *23*, 1013–1024. [CrossRef]

34. Gao, M.; Song, H.Y. Productivity under the perspective of China's grain growth factor analysis. *Chin. J. Popul. Sci.* **2015**, *35*, 59–69.

35. Liu, Y.; Gao, B.B.; Pan, Y.C.; Ren, X.H. Influencing factor decomposition of grain production at county level in Huang-Huai-Hai region based on LMDI. *Trans. Chin. Soc. Agric. Eng.* **2013**, *29*, 1–10.

36. Yang, J.Y.; Han, X.N.; Fang, X.M. An Empirical Study of China's grain production efficiency. *Econ. Inf.* **2013**, *9*, 47–53.

37. Jiang, S.; Wang, Z.; Huang, Q.H.; Zhou, Z.B.; Chan, X.D. Technological advances in food production speed and contribution to the study. *J. Agrotech. Econ.* **2012**, *10*, 40–51.

38. Liu, Y.P.; Li, T.; Zhao, H.F. Food technology progress contribution rate is calculated based on the micro perspective. *Chin. Agric. Sci. Bull.* **2012**, *28*, 114–117.

39. Chen, W.F.; Xu, Z.J.; Zhang, B.L.; Zhang, W.Z.; Ma, D.R. Northern japonica rice breeding for super high yield theory and practice. *Sci. Agric. Sin.* **2007**, *40*, 869–874.

40. Lin, H.; Pang, Q.L.; Ruan, L.Q.; Wang, Z.G. Analysis by nearly a decade of yield and quality of rice varieties validated. *China Rice* **2011**, *17*, 1–5.

41. Xu, Q.; Yin, R.L.; Zhang, H. Economies of scale, returns to scale and Agricultural Scale Management: Based on Empirical Study about China's grain production. *Econ. Res. J.* **2011**, *46*, 59–71.

42. Tian, Y.; Li, B.; Zhang, J.B. Rice Technological Progress Measure of Our Country. *Stat. Obs.* **2012**, *2*, 93–95.

43. Liu, Z.; Huang, F.; Li, B.G. Analysis on characteristics and influential factors of grain yield fluctuation in China based on empirical mode decomposition. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 7–13.

44. Sato, K. A Two-Level Constant-Elasticity-of-Substitution Production Function. *Rev. Econ. Stud.* **1967**, *34*, 201–218. [CrossRef]

45. Henningsen, A.; Henningsen, G. *Econometric Estimation of the "Constant Elasticity of Substitution" Function in R: Package micEconCES*; FOI Working Paper; University of Copenhagen: København, Denmark, 2011.

46. León-Ledesma, M.A.; Mcadam, P.; Willman, A. Identifying the Elasticity of Substitution with Biased Technical Change. *Am. Econ. Rev.* **2010**, *104*, 1330–1357. [CrossRef]

47. Solow, R.M. A Contribution to the Theory of Economic Growth. *Q. J. Econ.* **1956**, *70*, 65–94. [CrossRef]

48. Zhang, Z.Y. Productivity and economic growth: An empirical assessment of the contribution of FDI to the Chinese economy. *J. Econ. Dev.* **2002**, *27*, 81–94.

49. Beck, N.; Katz, J.N. Throwing Out the Baby with the Bath Water: A Comment on Green, Kim and Yoon. *Int. Organ.* **2001**, *55*, 487–495. [CrossRef]

50. Chang, H.Q.; Dong, X.Y.; Fiona, M.P. Labor Migration and Time Use Patterns of the Left-behind Children and Elderly in Rural China. *World Dev.* **2011**, *39*, 2199–2210. [CrossRef]

51. Ju, X.; Kou, C.; Zhang, F.; Christie, P. Nitrogen balance and groundwater nitrate contamination: Comparison among three intensive cropping systems on the North China Plain. *Environ. Pollut.* **2006**, *143*, 117–125. [CrossRef] [PubMed]

Article

Risk Profile Indicators and Spanish Banks' Probability of Default from a Regulatory Approach

Pilar Gómez-Fernández-Aguado, Purificación Parrado-Martínez and Antonio Partal-Ureña *

Department of Financial Economics and Accounting, University of Jaén, 23071 Jaén, Spain;
pigomez@ujaen.es (P.G.-F.-A.); pparrado@ujaen.es (P.P.-M.)
*   Correspondence: apartal@ujaen.es

Abstract: This paper analyses the relationships between the traditional bank risk profile indicators and a new measure of banks' probability of default that considers the Basel regulatory framework. First, based on the SYstemic Model of Bank Originated Losses (SYMBOL), we calculated the individual probabilities of default (PD) of a representative sample of Spanish credit institutions during the period of 2008–2016. Then, panel data regressions were estimated to explore the influence of the risk indicators on the PD. Our findings on the Spanish banking system could be important to regulatory and supervisory authorities. First, the PD based on the SYMBOL model could be used to analyse bank risk from a regulatory approach. Second, the results might be useful for designing new regulations focused on the key factors that affect the banks' probability of default. Third, our findings reveal that the emphasis on regulation and supervision should differ by type of entity.

Keywords: probability of default; bank risk; banking regulation; SYMBOL; financial stability

1. Introduction

The review of the international financial regulatory framework has set two objectives for financial stability: first, to reduce the probability of the bankruptcy of financial institutions by increasing solvency; and, second, to reduce public costs in the event of a bank crash and improve the framework for resolution. Regulatory reform in the European Union (EU) has been conducted by introducing the policies agreed upon by the G-20 using three directives that are the pillars of the banking union: (1) Directive IV (Directive 2013/36/EU of the European Parliament and of the Council of 26 June 2013 on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms, amending Directive 2002/87/EC and repealing Directives 2006/48/EC and 2006/49/EC) and Capital Requirements Regulation (CRD IV and CRR, 2013), which incorporates the third Basel agreement on micro-prudential regulation [1] (Basel III), significantly increasing the capital and liquidity requirements of financial institutions to manage unexpected losses; (2) the Directive on Banking Resolution and Recovery (Directive 2014/59/EU of the European Parliament and of the Council of 15 May 2014 establishing a framework for the recovery and resolution of credit institutions and investment firms), which establishes a set of rules and resolution tools, such as the sale of the company or shares of the entity subject to resolution, the creation of a bridge entity, the separation of productive assets from impaired assets or underperforming assets of an entity, and the rescue of shareholders and creditors in difficulty; and (3) a directive on deposit guarantee schemes (Directive 2014/49/EU of the European Parliament and of the Council of 16 April 2014 on deposit guarantee schemes) to preserve retail deposits at all times.

The importance of analysing the effects that regulatory standards have on financial stability requires the development of models and risk measures that can be adjusted to the actual regulatory framework and can provide a basis for stress tests or quantitative impact studies. The European

Union has recently started to use the SYMBOL model to assess the effectiveness of bank regulation and to measure the quantitative impact of a series of legislative proposals promoted by the European Commission [2,3]. SYMBOL is a model of micro-simulation based on the Basel risk assessment framework that estimates the distribution of bank losses originating in the system beginning with the balance and regulatory capital of individual banks. The methodological basis of the model was developed by the Commission's Joint Research Centre (JRC), the Directorate General Internal Market and Services and banking regulation experts (the original methodology was developed by De Lisa et al. [4], but, in this paper, the abbreviation SYMBOL was not employed).

While the SYMBOL model has been used to quantify systemic losses, it also provides a proxy for a bank's default probability as a tail risk [5] that reflects bank capital strength and asset quality [4]. The probability of default of a bank is estimated as the probability that its obligor's loss exceeds actual capital, given by the sum of its minimum capital requirement plus the bank's excess capital. In the Basel framework, each bank must satisfy a capital requirement that provides a buffer against unexpected losses at a specific level of statistical confidence, set by regulators at 99.9% [6]. From the regulatory perspective, the probability that banks default can be seen as the probability that banks' losses fall in the tail of their loss distribution. This "tail risk" is equal to 0.1% whenever banks set their amount of capital at a level equal to their regulatory minimum or lower if they hold capital in excess of the regulatory minimum. This risk occurs with a low probability, but it generates serious losses that endanger financial stability.

However, the recent development of this model indicates that there is currently no empirical evidence relating to the measurement of the likelihood of default to the indicators traditionally used for bank risk assessment. Our work extends the previous research on the determinants of bank risk, analysing the relationships between the risk indicators proposed by the European Banking Authority (EBA) [7] and the individual default probabilities of Spanish credit institutions during the period of 2008–2016, based on a new measure supported by banking regulation. We aim to determine whether these two sets of measures provide the same information of whether they are supplementary for bank risk analysis. After the last restructuring of the Spanish banking system, the regulation and supervision of bank risk have become crucial. In this sense, the analysis of new proxies from a regulatory approach could be decisive to avoid the financial instability. Indeed, this instability is a key obstacle to long-term investments which are required to achieve the sustainability in the financial system.

Our study is structured as follows. Section 2 presents a review of the literature. Section 3 describes the data and methodology employed. Section 4 presents and discusses the results. Finally, Section 5 concludes the study.

2. Related Literature

The literature on measuring the credit risk of banks and calculating the probability of default (PD) has been a focus in the use of accounting and market information. Studies that have used accounting information have basically applied two proxies for bank risk: NPL ratio and Z-score. The NPL ratio (non-performing loans (impaired loans) to total gross loans) or changes in the ratio have been used as a measurement of the strength of a bank [8–15] since they reflect the quality of a loan portfolio. A higher value of this ratio indicates a greater probability that the bank will default. Similarly, the Z-score has been widely used to measure bank risk [16–20]. Z-scores equal the return on assets plus the capital-to-asset ratio divided by the standard deviation of asset returns. A higher Z-score indicates that a bank is farther from default [21].

At the same time, research using market information as a complement to accounting indicators has been based on Merton's [22] approach to modelling the risk of credit defaults. Some studies have used credit risk spread CDs [23–31] or credit ratings [32–34].

More recently, De Lisa et al. [4] introduced a methodology for the estimation of the probability of the default of a bank based on the framework of capital requirements for credit risk from the Basel agreements. Based on public balance sheet data and the regulatory capital of the entities, the probability

of the default of a bank (PD Bank) is defined as the probability that unexpected losses associated with the portfolio's debtors (implied obligor probability of default, IOPD) exceed bank capital (given by the sum of regulatory capital requirements and any excess capital).

The methodology developed by De Lisa et al. [4] is the basis of the SYMBOL model. It has been used to evaluate the implications of new Basel III definitions (Risk weighted assets and computable resources), recent European banking regulation (CRD IV, CRR, Bank Recovery and Resolution Directive and Deposit Guarantee Schemes Directive), stress scenarios and quantitative valuations [2,3,35–39].

From a regulatory point of view, bank risk has traditionally been evaluated using indicators associated with different risk categories, such as capital adequacy, asset quality, liquidity and business model and management. EBA [7] proposed a series of indicators to determine bank risk profiles as the basis for contributions to deposit guarantee schemes. It also established the expected relationships (negative/positive) between indicators and bank risk.

This paper contributes to the existing literature by analysing the relationships between the traditionally used risk indicators and the new measure of a bank's PD based on the SYMBOL model.

3. Data and Methodological Aspects

3.1. Sample

Our sample consisted of Spanish credit institutions that have available information for all the analysed variables during the period of 2008–2016. Data were obtained from the banks' public documents (audited annual reports and information of prudential relevance). We used consolidated data since some risk indicators are not available on a solo basis. According to EBA (2015) guidelines, for each member institution, the values of risk indicators should be calculated on a solo basis. However, the value of risk indicators should be calculated at a consolidated level, at which the member state exercises the option given in Article 13(1) of DGSD to allow the central body and all credit institutions permanently affiliated with the central body, as referred to in Article 10(1) of Regulation (EU) 575/2013, to be subject as a whole to the risk weight determined for the central body and its affiliated institutions on a consolidated basis. Where a member institution has received a waiver from meeting capital and/or liquidity requirements on a solo basis pursuant to Articles 7, 8 or 21 of Regulation (EU) 575/2013, the corresponding capital/liquidity indicators should be calculated at the consolidated or semi-consolidated level. The entities are classified into two groups: commercial banks and credit cooperatives. Savings banks are included in the first group because, after the restructuring process that occurred in the Spanish financial sector during the last financial crisis, the majority of these entities were transformed into commercial banks. There are notable differences between these two types of institutions regarding not only their objectives but also their sizes. The total assets of commercial banks are more than 30 times greater than the total assets of credit unions (Table 1).

After eliminating extreme values considered outliers, we finally had an unbalanced panel with 359 observations related to 70 entities during the period of 2008–2016. The panel represented 87.85% of the total assets in the Spanish banking system. Table 1 presents the number of observations that constitute the sample, organised by year and type of entity, as well as their representativeness.

Table 1. Number of observations and representativeness of the sample.

	2008	2009	2010	2011	2012	2013	2014	2015	2016
Number of commercial Banks	43	46	30	30	18	20	19	20	20
Number of cooperative credit unions	11	12	12	14	13	13	11	14	13
Total credit institutions	54	58	42	44	31	33	30	34	33

Table 1. *Cont.*

	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total assets in commercial banks (billions of euros)									
Population	3537.0	3637.6	3691.9	3793.9	3754.2	3358.6	3476.0	3555.5	3492.3
Sample	2848.8	3002.1	3304.5	3406.6	2850.9	3028.9	3174.8	3347.0	3305.1
Representativeness of the sample (%)	80.5	82.5	89.5	89.8	75.9	90.2	91.3	94.1	94.6
Total assets in cooperative credit unions (billions of euros)									
Population	71.4	77.5	91.5	107.0	113.1	114.3	103.1	110.4	110.5
Sample	69.3	75.8	86.9	100.9	106.7	108.5	99.3	110.4	108.5
Representativeness of the sample (%)	97.1	97.9	95.0	94.3	94.4	94.9	96.4	100.0	98.2
Total assets in credit institutions (billions of euros)									
Population	3608.4	3715.1	3783.5	3901.0	3867.3	3472.9	3579.0	3665.9	3602.8
Sample	2918.1	3078.0	3391.5	3507.5	2957.7	3137.4	3274.1	3457.4	3413.6
Representativeness of the sample (%)	80.9	82.9	89.6	89.9	76.5	90.3	91.5	94.3	94.8

3.2. Dependent Variable: Probability of Default (PD)

SYMBOL simulates the distribution of losses in a banking system (usually a country) by aggregating individual banks' losses. Individual banks' losses are generated via Monte Carlo simulation using the Basel FIRB loss distribution function [40]. Simulated losses are based on an estimation of the average default probability of the portfolio of assets of any individual bank. Default of a bank is determined by the size of simulated losses and the regulatory capital available to absorb unexpected shocks. Banks are expected to cover their expected losses with provisions. Unexpected losses, in contrast, relate to potentially large losses that occur rather seldom. According to this concept, capital would be needed only for absorbing unexpected loss. The model thus assumes that [3]: (1) the Basel III regulatory model for credit risk is correct; (2) banks report risks accurately that are in agreement with this model; and (3) all risks in the bank can be represented as a single portfolio of credit risks (this representation does not indicate that other risks are not considered—simply that they can be "mapped" in credit risk terms and modelled using the same framework). The model is processed in four methodological steps [4,38,41,42]. The first, second and third stages are necessary to estimate the banks' individual probability of default (PD bank), and the last stage determines the aggregate loss distribution of the system as a whole. Below, we describe the methodological considerations to estimate a bank's PD.

3.2.1. Estimation of the Implied Obligor Probability of Default of the Portfolio of Each Individual Bank

The main parameter of the model is the average implied obligor probability of default of a bank ($IOPD_i$). This variable is obtained from the Basel IRB formula to establish the minimum capital requirements for credit risk (*FIRB approach*). For each exposure l of bank i, the IRB formula establishes capital requirement $CR_{i,l}$ to cover the unexpected losses in the time horizon of one year for a 99.9% confidence level, as follows:

$$CR_{i,l}(PD_{i,l}) = \left[LGD \cdot N\left(\sqrt{\frac{1}{1-R(PD_{i,l})}} \cdot N^{-1}(PD_{i,l}) + \sqrt{\frac{R(PD_{i,l})}{1-R(PD_{i,l})}} \cdot N^{-1}(0.999) \right) - PD_{i,l} \cdot LGD \right] \cdot M(PD_{i,l}) \quad (1)$$

where $PD_{i,l}$ is the default probability of exposure l; and R is the correlation among the exposures in the portfolio, defined as:

$$R(PD_{i,l}) = 0.12 \cdot \frac{1-e^{-50 \cdot PD_{i,l}}}{1-e^{-50}} + 0.24 \cdot \left(1 - \frac{1-e^{-50 \cdot PD_{i,l}}}{1-e^{-50}} \right) - 0.04 \cdot \left(1 - \frac{s-5}{45} \right)$$

where obligor size s is equal to 50; LGD is the loss given default (considered to be 45% in the FIRB approach); and $M(PD_{i,l})$ is an adjustment term, defined as:

$$M(\text{PD}_{i,\,l}) = \frac{(1 + (M - 2.5)\cdot b_{i,l})\cdot 1.06}{1 - 1.5\cdot b_{i,l}}$$

where M is the time to maturity (considered to be 2.5 years in the FIRB approach), and b is the maturity adjustment, computed as $b_{i,l} = (0.11856 - 0.05478\cdot \ln(\text{PD}_{i,l}))^2$.

The minimum capital requirement of the bank (MCR_i) is equal to the sum of the capital requirements of all the exposures:

$$\text{MCR}_i = \sum_l \text{CR}_{i,l}\cdot A_{i,l} \tag{2}$$

where $A_{i,l}$ is the amount of exposure l.

Because there is no available public information about the different banking exposures, the model considers only one debtor that is equivalent to the total portfolio, and it estimates IOPD_i by solving the following equation:

$$\text{CR}(\text{IOPD}_i)\cdot \sum_l A_{i,l} = \text{MCR}_i \tag{3}$$

where $\text{CR}(\text{IOPD}_i)$ is the minimum capital requirement based on the Basel regulation (equal to 8% of the risk-weighted assets), and $\sum_l A_{i,l}$ is the total assets of the bank.

3.2.2. Simulation of Correlated Losses for the Banks in the System

Given the estimated IOPD, the SYMBOL model uses Monte Carlo simulation to generate the bank loss distribution, using the same IRB formula and imposing a correlation structure among banks (with a correlation coefficient set to $\rho = 50\%$). SYMBOL is often run imposing an equal correlation factor of 50% among all banks. A discussion and a sensitivity check of this assumption can be found in De Lisa et al. [4] and Benczur et al. [37]. This correlation exists as a consequence of the banks' common exposure, either to the same borrower or, more generally, to a particular common influence of the business cycle. In each simulation run j, losses for bank i are simulated as:

$$L_{i,j} = \text{LGD}\cdot N\left[\sqrt{\frac{1}{1 - R(\text{IOPD}_i)}}\cdot N^{-1}(\text{IOPD}_i) + \sqrt{\frac{R(\text{IOPD}_i)}{1 - R(\text{IOPD}_i)}}\cdot N^{-1}(\alpha_{i,\,j})\right] \tag{4}$$

where N is the normal distribution function, and $N^{-1}(\alpha_{n,i})$ are correlated normal random shocks. IOPD_i is the implied obligor probability of default estimated for each bank in Step 1, and LGD is the loss given default, set as in the Basel regulation equal to 45%.

3.2.3. Determination of the Failure Event and Estimation of PD Bank

In the SYMBOL model, a bank i is considered in default when simulated losses (L_{ij}) exceed the sum of the expected losses (EL_i) and the total actual capital (K_i) given by the sum of its minimum capital requirement plus the bank's excess capital:

$$L_{i,j} - EL_i - K_i > 0 \rightarrow \text{Bank } i \text{ defaults} \tag{5}$$

The probability of the default of a bank (PD bank) is calculated as the number of times that the bank defaults over the total number of simulations (100,000).

3.3. Independent Variables: Risk Profile Indicators

As independent variables, we use the core risk indicators established by the EBA in 2015 [7] (Table 2). They include the following: (i) concerning capital, the leverage ratio (C1) and the capital coverage ratio, calculated as actual own funds divided by required own funds (C3) and the CET1 ratio (C4); (ii) concerning liquidity and funding, the LCR (L1), the NSFR (L2) and the liquidity ratio (L3); (iii) concerning asset quality, the non-performing loans (NPL) ratio (AQ1); and (iv) concerning business

model and management, the risk-weighted asset to total assets ratio (G1) and the ROA (G2). Because we do not have information for some of the liquidity indicators proposed, we use the loans-to-deposits (LTD) ratio and the stable funding/stable assets (ST/SA) ratio as a proxy for the LCR and the NSFR, respectively. EBA [7] stated that competent authorities may exclude a core indicator upon justification that it is unavailable because of the legal characteristics or supervisory regimes of some institutions. In these cases, the most appropriate proxy should be used for the removed indicator.

Table 2. Core indicators of likelihood of failure.

Category	Indicator	Description	Notation	Expected Sign on Bank Risk
Capital	Leverage ratio	Tier 1 capital/Total assets	C1	Negative
	Capital coverage ratio	Actual common equity Tier 1 ratio/Required common equity Tier 1 ratio or actual own funds/Required own fund	C2/C3	Negative
	Common equity Tier 1 (CET1) ratio	Common equity Tier 1 capital/Risk-weighted assets	C4	Negative
Liquidity and Funding	Liquidity coverage ratio (LCR)	LCR ratio as defined in Regulation (EU) No. 575/2013 once it becomes fully operational	L1	Negative
	Net stable funding ratio (NSFR)	NSFR ratio as defined in Regulation (EU) No. 575/2013 once it becomes fully operational	L2	Negative
	Liquidity ratio	Liquid assets/Total assets	L3	Negative
Asset Quality	Non-performing loans (NPL) ratio	Non-performing loans/Total loans and debt instruments	AQ1	Positive
Business model and management	Risk-weighted assets to total assets ratio (RWA/TA)	Risk-weighted assets/Total assets	G1	Positive
	Return on assets (ROA)	Net income/Total assets	G2	Positive/Negative

Source: EBA (2015).

The choice of explanatory variables is driven by availability considerations and multicollinearity issues. Correlation analyses and collinearity diagnostics were performed to assess the extent of multicollinearity among independent variables. These problems were detected using the correlation matrix and the variance inflation factor (VIF). As a result, the initially considered CET1 ratio (C4) and RWA/TA ratio (G1) were finally excluded.

3.4. Methodology

The panel data regression technique is the most appropriate method when the sample comprises cross-sectional and time-series data. The main benefit of using panel data is that they overcome the unobservable, constant and heterogeneous characteristics of each bank in the sample [43]. In panel data models, the stochastic part (ε_{it}) consists of two components: u_i and e_{it} The unobserved heterogeneity (u_i) represents the bank-specific error term, capturing all the variation at the bank level that is not controlled for by the independent variables in the model. The idiosyncratic error (e_{it}) captures all peculiarities, apart from the effects that are already controlled for in the model, that affect the dependent variable for each bank at each point in time [44]. The Hausman test indicates in all cases that fixed effects models are preferred. These models assume that u_i is correlated with the variables. Our baseline equation is the following:

$$\begin{aligned} \mathrm{PD}_{it} = \quad & \alpha + \beta_1 \times C1_{i,t} + \beta_2 \times C3_{i,t} + \beta_3 \times L1_{i,t} + \beta_4 \times L2_{i,t} \\ & + \beta_5 \times L3_{i,t} + \beta_6 \times AQ1_{i,t} + \beta_7 \times G2_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{6}$$

where PD_{it} represents the probability of default of the bank i in year t; the notations of the explanatory variables are described in Table 2; the subscripts i and t represent banks and years, respectively; α is the constant term; β_k are the coefficient estimates; and ε_{it} is the disturbance.

4. Results

4.1. Relationship between EBA Bank Risk Indicators and PD

First, we report the results of the econometric models for the entire sample of credit institutions (Table 3). The Hausman test shows that the best model is a fixed effects (within) one. This model recognises the existence of individual differences at the level of banks (unobserved heterogeneity u_i) that are not controlled for by the explanatory variables so that the PD of each bank depends on individual factors that cannot be observed.

Table 3. Relationships between risk indicators and PD for Spanish banks.

	(1)	(2)	(3)
C1	−0.01920 *** (0.00315)	−0.01834 *** (0.00332)	−0.01904 *** (0.00329)
C3	−0.00049 *** (0.00006)	−0.00050 *** (0.00006)	−0.00044 *** (0.00006)
L1	−0.00026 * (0.00013)	−0.00024 * (0.00013)	−0.00021 (0.00014)
L2	−0.00029 (0.00020)	−0.00030 (0.00020)	−0.00017 (0.00021)
L3	−0.00081 ** (0.00038)	−0.00080 ** (0.00038)	−0.00086 ** (0.00038)
AQ1	−0.00713*** (0.00113)	−0.00732 *** (0.00116)	−0.00419 *** (0.00149)
G2	0.00423 (0.00460)	0.00456 (0.00462)	0.00196 (0.00513)
Constant	0.52557 *** (0.02951)	0.40203 *** (0.15251)	0.48428 *** (0.03537)
Size	-	0.00737 (0.00892)	-
Year dummies	-	-	Yes
GDP	-	-	−0.00300 * (0.00175)
Inflation	-	-	0.00916 *** (0.00259)
Risk Premium	-	-	−0.00009 ** (0.00004)
Number of obs.	359	359	359
F(*df*)	45.50 (7, 282)	39.85 (8, 281)	23.00 (15, 274)
Prob > F	0.0000	0.0000	0.0000
R^2 (within)	0.5304	0.5315	0.5573
Hausman	27.01 (7)	29.16 (8)	29.42 (12)

Notes: This table reports the results of the fixed effects (within) regressions for the entire sample (70 financial institutions). Model (1) includes only the main indicators. The Model (2) includes the variable size of the entity. Model (3) includes macroeconomic variables and time dummies. The standard errors are reported in parentheses. R^2 is the proportion of variation in the dependent variable explained by the model. The Hausman test compares the fixed versus random effects, asymptotically distributed as χ^2 under the hypothesis that the individual effects are uncorrelated with the other regressors in the model, with degrees of freedom in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

The results of the baseline Model (1) show significant coefficients for the capital adequacy ratios, indicating the relevant impact of these variables on the banks' probability of default. The C1 and C3 ratios have significantly negative impacts on PD such that an increase in the capitalisation of banks increases the ability to absorb sudden losses and, therefore, to reduce the probability of default. Our findings support previous research, such as Čihák and Schaeck [45], Baselga-Pascual et al. [15] and Leung et al. [46].

Regarding the influence of liquidity indicators on bank risk, the literature has not produced conclusive findings. Ahmand and Arrif [11] found evidence for both a positive and a negative relationship between liquidity indicators and credit risk, depending on the analysed country. For some countries, the authors suggested that banks with more liquid assets tend to have higher credit risk. They believed that this contrasting result might be due to short-term assets or loans that banks own to hold more liquid assets to defend against rapidly changing interest rate markets. Corroborating this belief, Khan et al. [20] found that banks with lower fund liquidity risk tend to take more risks. In other countries, Ahmand and Arrif [11] found that liquidity is negatively related to credit risk. Veloz and Benou de Gomez [10] and Baselga-Pascual et al. [15] supported the premise that banks with higher levels of liquidity are less likely to experience an increase in the default rate. Our research also supports this last premise: the liquidity indicators (L1 and L3) appear to be statistically significant, such that the higher the liquidity is, the lower the probability of default of the commercial banks is.

The asset quality of banks also influences our measure of PD. There is a general consensus in the literature that an increase in the percentage of non-performing loans (NPLs) can reduce the quality of the banking sector's assets and increase the probability of default (e.g., [13,15,47–49]). However, following the financial crisis, higher capital requirements can generate the opposite effect, encouraging banks to increase risk taking, although this increase in risk taking is offset by the beneficial impact of increased loss absorbing arising from higher capital requirements, leading to more stable banks [18,50–52]. Our findings agree with this last argument. An increase in the delinquency rate is offset by a reinforcement of capital, and this higher buffer of capital will lead to lower unexpected losses and therefore lower PD. Finally, the remaining variable related to profitability does not prove to be significant.

In Model (2), we introduced the logarithm of the total assets as a proxy for the size of the bank, but it does not have any impact on PD. This result supports the "one-size-fits-all" financial regulation for the Spanish banking system.

In Model (3), we control for the macroeconomic variables and we find a relevant impact of GDP growth, inflation and risk premium on the probability of default of Spanish banks. These results indicate that the traditional bank risk indicators are not the only determinants of the new measure of PD from a regulatory approach. Therefore, the PD based on the SYMBOL model should be used jointly with the traditional indicators for a more complete analysis of the bank risk.

4.2. Differences between Commercial Banks and Credit Cooperatives

We contrast the existence of significant differences between groups of entities (commercial banks and credit cooperatives) using analysis of variance (ANOVA). Table 4 shows the descriptive statistics (mean and standard deviation) and the results for the Welch test (Because the majority of variables do not meet the assumption of homoscedasticity, we employ Welch's test, which is more robust if this assumption fails. Levene's test is employed to contrast the homogeneity of variances).

The results show notable differences in the risk profiles of commercial banks and credit cooperatives. There are significant differences in the level of solvency by type of entity. The mean probability of default for commercial banks is higher than the mean PD for credit cooperatives. Furthermore, credit cooperatives have higher mean values of capital adequacy (C1 and C3) and liquidity indicators (L1, L2, and L3). From this initial outline, we can deduce the existence of different risk profiles among Spanish banks. It seems that credit cooperatives have contributed to financial

stability to a greater extent than commercial banks, perhaps because the cooperative business model of credit cooperatives in Spain is more conservative about taking risks.

Table 4. Summary statistics for commercial banks and credit cooperatives.

	Commercial Banks	Cooperative Credit Unions	Total	Welch's Test
PD	0.211 (0.083)	0.177 (0.083)	0.201 (0.085)	12.805 ***
C1	5.984 (1.363)	6.828 (1.432)	6.250 (1.438)	27.698 ***
C3	173.561 (61.765)	214.895 (51.245)	186.571 (61.659)	44.094 ***
L1	84.659 (26.227)	91.563 (18.943)	86.832 (24.358)	7.983 ***
L2	109.424 (26.350)	128.120 (34.130)	115.309 (30.252)	26.620 ***
L3	20.061 (12.187)	26.908 (13.134)	22.216 (12.875)	22.010 ***
AQ1	5.790 (3.423)	5.784 (3.854)	5.788 (3.559)	0.002
G2	0.277 (0.662)	0.341 (0.530)	0.297 (0.623)	0.963

Notes: This table reports the means and standard deviations (in parentheses) for the commercial banks, the credit cooperatives and the entire sample during the period of 2008–2016. We also report the results of the test for differences in means to analyse the statistically significant differences. *** Significant at the 1% level.

The previous literature found some evidence concerning the influences of the ownership structure and the sizes of banks on their risk. Klomp and Haan [53] found that the effects of banking regulation and supervision on bank risk depends not only on the level of risk of the bank but also on its ownership structure and size as well. In this sense, Shehzad et al. [54] stated that risk-taking behaviour depends on the bank ownership concentration. Finally, Vázquez and Federico [55] concluded that smaller banks are more vulnerable to liquidity problems, while larger banking groups usually have a lack of capital buffers. Based on these findings, we introduce interaction dummies into our econometric models to explore the impact of the risk indicators on PD by type of entity (see Table 5).

The statistical significance of the interaction dummies in the specification (1) indicates that the effects of the risk indicators on the PD differ by type of entity. In fact, for commercial banks (2), the capital adequacy (C1 and C3) and the asset quality (AQ1) ratios appear to be significant, whereas for credit cooperatives (3), the capital (C3) and the liquidity (L3) indicators seem to be the most important determinants of the PD. This result shows notable differences regarding key factors in the risk profiles of commercial banks and credit cooperatives, and our findings agree with research previously mentioned [53–55].

4.3. Differences between Crisis and Post-Crisis Periods

Finally, we contrasted the existence of significant differences between the crisis period (2008–2012) and the post-crisis period (2013–2016) using analysis of the Variance (ANOVA). We considered the post-crisis period from 2013, when the first signs of economic recovery were seen (some signs are the reduction in 10-year Germany bond spreads and the increases in real GDP and domestic demand) (Bank of Spain, 2017).

Table 6 shows the descriptive statistics (mean and standard deviation), and the results of the Welch test.

The results show notable differences in the risk profiles of banks during and after the most recent financial crisis. The mean probability of default after the crisis is much lower than the mean PD for the crisis period. Furthermore, after the crisis, Spanish banks presented higher mean values of capital adequacy and liquidity indicators. Finally, the delinquency rate experienced an increase after the crisis. From this initial outline, we can deduce that Spanish banking institutions presented a better financial situation after the crisis, likely because of the regulatory actions adopted.

Table 5. Differences in the determinants of the probability of default by type of entity.

	(1) Interaction Dummies	(2) Commercial Banks	(3) Credit Cooperatives
C1	−0.00567 (0.00528)	−0.02849 *** (0.00400)	−0.00567 (0.00528)
C3	−0.00088 *** (0.00014)	−0.00033 *** (0.00007)	−0.00088 *** (0.00014)
L1	−0.00042 (0.00027)	−0.00026 * (0.00015)	−0.00042 (0.00027)
L2	0.00034 (0.00033)	−0.00004 (0.00029)	0.00034 (0.00033)
L3	−0.00214 ** (0.00086)	−0.00060 (0.00044)	−0.00214 ** (0.00086)
AQ1	−0.00299 (0.00193)	−0.00959 *** (0.00139)	−0.00299 (0.00193)
G2	−0.00417 (0.00908)	0.00475 (0.00524)	−0.00417 (0.00908)
Constant	0.51424 *** (0.03094)	0.51424 *** (0.03094)	0.51424 *** (0.03094)
Number of obs.	359	246	113
C1 x dummy type	−0.02282 *** (0.00662)		
C3 x dummy type	0.00055 *** (0.00016)		
L1 x dummy type	0.00016 (0.00031)		
L2 x dummy type	−0.00038 (0.00044)		
L3 x dummy type	0.00153 (0.00096)		
AQ1 x dummy type	−0.00660 *** (0.00238)		
G2 x dummy type	0.00892 (0.01049)		
$F_{(df)}$	26.56 (14, 275)		
Prob > F	0.0000		
R^2 (within)	0.5748		
Hausman	26.28 (14)		

Notes: This table reports the results of fixed effects (within) regressions. The model in Column (1) includes interaction dummies. Considering the differences in the determinants of PD by type of entity, we report the coefficients of the regression for commercial banks (2) and for credit cooperatives (3). The standard errors are reported in parentheses. R^2 is the proportion of variation in the dependent variable explained by the model. The Hausman test compares the fixed versus random effects, asymptotically distributed as χ^2 under the hypothesis that the individual effects are uncorrelated with the other regressors in the model, with degrees of freedom in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table 6. Summary statistics for periods of time.

	2008–2012	2013–2016	Total	Welch's Test
PD	0.226 *(0.084)*	0.156 *(0.065)*	0.201 *(0.085)*	77.129 ***
C1	6.320 *(1.436)*	6.127 *(1.437)*	6.250 *(1.438)*	1.499
C3	169.311 *(37.324)*	216.976 *(81.414)*	186.571 *(61.659)*	39.810 ***
L1	84.732 *(22.267)*	90.532 *(27.359)*	86.832 *(24.358)*	4.246 **
L2	108.732 *(20.131)*	126.893 *(40.144)*	115.309 *(30.252)*	23.282 ***
L3	18.848 *(9.085)*	28.150 *(16.078)*	22.216 *(12.875)*	36.839 ***
AQ1	4.368 *(2.249)*	8.289 *(4.040)*	5.788 *(3.559)*	104.145 ***
G2	0.269 *(0.681)*	0.346 *(0.505)*	0.297 *(0.623)*	1.483

Following these arguments, we introduce interaction dummies into our econometric models to explore the impact of the risk indicators on PD by period of time (see Table 7).

Table 7. Differences in the determinants of the probability of default by periods of time.

	Interaction Dummies	(1) Crisis Period	(2) Post-Crisis Period
C1	−0.01251 *** *(0.00366)*	0.01267 ** *(0.00504)*	−0.01251 *** *(0.00366)*
C3	−0.00032 *** *(0.00006)*	−0.00211 *** *(0.00021)*	−0.00032 *** *(0.00006)*
L1	0.00003 *(0.00014)*	−0.00065 * *(0.00036)*	0.00003 *(0.00014)*
L2	−0.00035 * *(0.00020)*	0.00034 *(0.00042)*	−0.00035* *(0.00020)*
L3	−0.00076 ** *(0.00038)*	0.00004 *(0.00055)*	−0.00076 ** *(0.00038)*
AQ1	−0.00129 *(0.00149)*	−0.00037 *(0.00197)*	−0.00129 *(0.00149)*
G2	0.01204 *(0.00864)*	−0.00217 *(0.00534)*	0.01204 *(0.00864)*
Constant	0.36656 *** *(0.03891)*	0.52600 *** *(0.03601)*	0.36656 *** *(0.03891)*
Dummy crisis	0.15944 *** *(0.04199)*		
Number of obs.	359	229	130
C1 x dummy crisis	0.02518 *** *(0.00514)*		
C3 x dummy crisis	−0.00179 ***		

<div align="center">**Table 7.** *Cont.*</div>

	Interaction Dummies	(1) Crisis Period	(2) Post-Crisis Period
	(*0.00022*)		
L1 x dummy crisis	−0.00068 * (*0.00037*)		
L2 x dummy crisis	0.00069 (*0.00042*)		
L3 x dummy crisis	0.00081 (*0.00059*)		
AQ1 x dummy crisis	0.00092 (*0.00193*)		
G2 x dummy crisis	−0.01422 (*0.01018*)		
F(*df*)	32.94 (*15, 274*)		
Prob > F	0.0000		
R^2 (within)	0.6433		
Hausman	76.08 (*15*)		

Notes: This table reports the results of fixed effects (within) regressions. The model in Column (1) includes interaction dummies. Considering the differences in the determinants of PD by period of time, we report the coefficients of the regression for the crisis period (2) and for the post-crisis period (3). The standard errors are reported in parentheses. R^2 is the proportion of variation in the dependent variable explained by the model. The Hausman test compares the fixed versus random effects, asymptotically distributed as χ^2 under the hypothesis that the individual effects are uncorrelated with the other regressors in the model, with degrees of freedom in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

The statistical significance of the interaction dummies in specification (1) indicates that the effects of the risk indicators on the PD differ by period of time. In fact, during the crisis period (2), the main determinants of PD were the capital adequacy (C1 and C3) ratios, whereas after the financial crisis (3), the liquidity (L2 and L3) indicators seem to also be important determinants of the PD. Our findings agree with Vázquez and Federico [55], who concluded that banks with structural liquidity difficulties and higher leverage during a pre-crisis period are more prone to experience distress in a post-crisis period. These results show notable differences regarding key factors in the risk profiles of Spanish banks during and after the crisis.

5. Conclusions

Given the importance of stress exercises as an element of analysis and a policy tool, the EU started to use a new model, called SYMBOL, to measure the quantitative impact of different European Commission proposals. The model can also be used to determine the probability of bank default under the Basel regulatory approach [5]. Our work extends the previous research on the determinants of bank risk, analysing the relationships between the risk indicators proposed by the EBA [7] and the individual default probabilities of Spanish credit institutions during the period of 2008–2016, based on a new measure supported by banking regulation.

Our results for Spanish banking credit institutions reveal that capital adequacy, liquidity and asset quality are the main determinants of the banks' probability of default from a regulatory approach. The macroeconomic scenario appears to be significant as well. In addition, the explanatory power of the models indicates that the traditionally used indicators are not the only determinants of the new measure of banks' PD. Thus, this new proxy could be supplementary to the traditional risk indicators in bank risk analysis.

When we contrast the existence of significant differences between groups of entities, we find different risk profiles between commercial banks and credit cooperatives. The mean probability of

default of commercial banks is higher than the mean PD of credit cooperatives. Furthermore, the credit cooperatives present greater capital adequacy and liquidity than commercial banks. It seems that Spanish credit cooperatives have contributed to financial stability to a greater extent than commercial banks. In addition, the introduction of interaction dummies into the models reveals that the influence of each indicator in the PD differs by the type of entity. For commercial banks, the probability of default is mainly influenced by the capital adequacy (C1 and C3) and asset quality (AQ1) ratios, while, for credit cooperatives, the capital (C3) and liquidity (L3) indicators seem to be the most important determinants of the PD. These findings support previous research [53–55], and they reveal notable differences regarding the key factors in the risk management of commercial banks and credit cooperatives.

Regarding the influence of the most recent financial crisis, we can deduce that Spanish banking institutions present a better financial situation after the crisis, likely because of the regulatory actions adopted. The results show that the mean probability of default after the crisis is much lower than the mean PD for the crisis period. Furthermore, after the crisis, Spanish banks present higher mean values of capital adequacy and liquidity indicators. The statistical significance of the interaction dummies indicates that the effects of the risk indicators on the PD also differ by period of time. In fact, during the crisis period, the main determinants of PD were the capital adequacy (C1 and C3) ratios, whereas after the financial crisis, the liquidity (L2 and L3) indicators also seem to be important determinants of the PD. Our findings agree with Vázquez and Federico [55].

In conclusion, our findings could be important to regulatory and supervisory authorities for several reasons. First, the PD based on the SYMBOL model could be considered to analyse bank risk from a regulatory approach. It could be used jointly with the traditional indicators for a more complete analysis of unexpected bank losses. Second, given the standardised calculation of this measurement, it could enable comparisons across different credit institutions [56], thus improving market discipline. Finally, our results might be useful to the design of new regulations focused on the key factors that affect the banks' probability of default. We have shown the importance of considering the differences between types of entities from the point of view of financial stability, indicating that the emphasis of regulation and supervision in Spain should differ by type of entity to control for different risk factors.

References

1. Basel Committee on Banking Supervision. *Basel III: International Regulatory Framework for Banks*; Basel Committee on Banking Supervision: New York, NY, USA, 2011.
2. European Commission. Directorate-General for Economic and Financial Affairs. *Fiscal Sustainability Report*. 2012. Available online: http://ec.europa.eu/economy_finance/publications/european_economy/2012/pdf/ee-2012-8_en.pdf (accessed on 12 August 2016).
3. European Commission. Effects Analysis (EA) on the European Deposit Insurance Scheme (EDIS). 2016. Available online: https://ec.europa.eu/info/publications/effects-analysis-european-deposit-insurance-scheme-edis_en (accessed on 27 September 2017).
4. De Lisa, R.; Zedda, S.; Vallascas, F.; Campolongo, F.; Marchesi, M. Modelling deposit insurance scheme losses in a Basel 2 framework. *J. Financ. Serv. Res.* **2011**, *40*, 123–141. [CrossRef]
5. Repullo, R.; Suarez, J. Loan pricing under Basel capital requirements. *J. Financ. Intermed.* **2004**, *13*, 496–521. [CrossRef]

6. Basel Committee on Banking Supervision. International Convergence of Capital Measurement and Capital Standards. A Revised Framework. 2004. Available online: https://www.bis.org/publ/bcbs107.pdf (accessed on 10 July 2016).

7. European Banking Authority (EBA). Guidelines on Methods for Calculating Contributions to Deposit Guarantee Schemes; EBA/GL/2015/10. 2015. Available online: https://www.eba.europa.eu/documents/10180/1089322/EBA-GL-2015-10+GL+on+methods+for+calculating+contributions+to+DGS.pdf (accessed on 1 August 2016).

8. Berger, A.N.; DeYoung, R. Problem loans and cost efficiency in commercial banks. *J. Bank. Financ.* **1997**, *21*, 849–870. [CrossRef]

9. Salas, V.; Saurina, J. Credit Risk in Two Institutional Regimes: Spanish Commercial and Savings Banks. *J. Financ. Serv. Res.* **2002**, *22*, 203–224. [CrossRef]

10. Veloz, A.; Benou de Gómez, G. Determinantes de fragilidad del sistema bancario en la República Dominicana: Una aplicación micro-macro de modelos de alerta temprana. *Sci. Soc.* **2007**, *32*, 69–87. [CrossRef]

11. Ahmad, N.H.; Arrif, M. Multi-country study of bank credit risk determinants. *Int. J. Bank. Financ.* **2007**, *5*, 135–152.

12. Delis, M.D.; Kouretas, G.P. Interest rates and bank risk-taking. *J. Bank. Financ.* **2011**, *35*, 840–855. [CrossRef]

13. Festic, M.; Kavkler, A.; Repina, S. The macroeconomic sources of systemic risk in the banking sectors of five new EU member states. *J. Bank. Financ.* **2011**, *35*, 310–322. [CrossRef]

14. Climent-Serrano, S.; Pavía, J.M. An analysis of loan default determinants: the Spanish case. *Banks Bank Syst.* **2014**, *9*, 116–123.

15. Baselga-Pascual, L.; Trujillo-Ponce, A.; Cardone-Riportella, C. Factors influencing bank risk in Europe: Evidence from the financial crisis. *N. Am. J. Econ. Financ.* **2015**, *34*, 138–166. [CrossRef]

16. Agoraki, M.-E.K.; Delis, M.D.; Pasiouras, F. Regulations, competition and bank risk-taking in transition countries. *J. Financ. Stab.* **2011**, *7*, 38–48. [CrossRef]

17. Demirgüç-Kunt, A.; Huizinga, H. Bank activity and funding strategies: The impact on risk and returns. *J. Financ. Econ.* **2010**, *98*, 626–650. [CrossRef]

18. Laeven, L.; Levine, R. Bank governance, regulation and risk taking. *J. Financ. Econ.* **2009**, *93*, 259–275. [CrossRef]

19. Lepetit, L.; Strobel, F. Bank insolvency risk and Z-score measures: A refinement. *Financ. Res. Lett.* **2015**, *13*, 214–224. [CrossRef]

20. Khan, M.S.; Scheule, H.; Wu, E. Funding liquidity and bank risk taking. *J. Bank. Financ.* **2017**, *82*, 203–216. [CrossRef]

21. Delis, M.D.; Staikouras, P.K. Supervisory effectiveness and bank risk. *Rev. Financ.* **2011**, *15*, 511–543. [CrossRef]

22. Merton, R.C. On the pricing of corporate debt: The risk structure of interest rates. *J. Financ.* **1974**, *29*, 449–470.

23. Kool, C.J.M. Financial Stability in European Banking: The Role of Common Factors. *Open Econ. Rev.* **2006**, *17*, 525–540. [CrossRef]

24. Huang, X.; Zhou, H.; Zhu, H. A Framework for Assessing the Systemic Risk of Major Financial Institutions. Available online: https://www.federalreserve.gov/pubs/feds/2009/200937/200937pap.pdf (accessed on 9 April 2018).

25. Hart, O.; Zingales, L. *A New Capital Regulation for Large Financial Institutions*; CEPR Discussion Paper No. DP7298; NBER & CEPR; University of Chicago: Chicago, IL, USA, 2009.

26. Alter, A.; Schüler, Y.S. Credit spread interdependencies of European states and banks during the financial crisis. *J. Bank. Financ.* **2012**, *36*, 3444–3468. [CrossRef]

27. Chiaramonte, L.; Casu, B. The determinants of bank CDS spreads: evidence from the financial crisis. *Eur. J. Financ.* **2013**, *19*, 861–887. [CrossRef]

28. Annaert, J.; De Ceuster, M.; Van Roy, P.; Vespro, C. What determines Euro area bank CDS spreads? *J. Int. Money Finan.* **2013**, *32*, 444–461. [CrossRef]

29. Hasan, I.; Liu, L.; Zhang, G. The Determinants of Global Bank Credit-Default-Swap Spreads. Available online: https://link.springer.com/article/10.1007/s10693-015-0232-z (accessed on 9 April 2018).

30. Samaniego-Medina, R.; Trujillo-Ponce, A.; Parrado-Martínez, P.; di Pietro, F. Determinants of bank CDS spreads in Europe. *J. Econ. Bus.* **2016**, *86*, 1–15. [CrossRef]

31. Drago, D.; Di Tommaso, C.; Thornton, J. What determines bank CDS spreads? Evidence from European and US banks. *Financ. Res. Lett.* **2017**, *22*, 140–145. [CrossRef]

32. Calomiris, C. Financial innovation, regulation and reform. *Cato J.* **2009**, *29*, 65–91.

33. Hau, H.; Langfield, S.; Marques-Ibanez, D. *Bank Ratings. What Determines Their Quality?* Working Paper Series, No. 1484; European Central Bank: Frankfurt, Germany, 2012.

34. Wang, L. Bank Rating Gaps as Proxies for Systemic Risk. University of Alberta, 2017. Available online: papers.ssrn.com/sol3/papers.cfm?abstract_id=2966413 (accessed on 9 April 2018).

35. Marchesi, M.; Petracco Giudici, M.; Cariboni, J.; Zedda, S.; Campolongo, F. *Macroeconomic Cost-benefit Analysis of Basel III Minimum Capital Requirements and of Introducing Deposit Guarantee Schemes and Resolution Funds*; European Commission JRC Scientific and Policy Report, 24603; EUR: Brussels, Belgium, 2012.

36. Cannas, G.; Cariboni, J.; Forys, M.; Joensson, H.; Langedijk, S.; Marchesi, M.; Nda-cyayisenga, N.; Pagano, A.; Petracco-Giudici, M. *Quantitative Estimation of a Part of the Costs and Benefits of Bank Structural Separation*; JRC Scientific and Technical Report; EUR: Brussels, Belgium, 2013.

37. Benczur, P.; Berti, K.; Cariboni, J.; Di Girolamo, F.E.; Langedijk, S.; Pagano, A.; Petracco Giudici, M. *Banking Stress Scenarios for Public Debt Projections*; European Economy Economic Papers; Economic Papers 548; Publications Office of the European Union: Luxembourg, 2015.

38. Benczur, P.; Cannas, G.; Cariboni, J.; Di Girolamo, F.; Maccaferri, S.; Petracco Giudici, M. Evaluating the effectiveness of the new EU bank regulatory framework: A farewell to bail-out? *J. Financ. Stab.* **2016**, *33*, 207–223. [CrossRef]

39. Gómez, P.; Partal, A.; Trujillo, A. A note on the adequacy of the EU scheme for bank recovery, resolution and deposit insurance in Spain. *J. Bank Regul.* **2016**, *17*, 332–337. [CrossRef]

40. Basel Committee on Banking Supervision. International Convergence of Capital Measurement and Capital Standards. A Revised Framework. 2006. Available online: https://www.bis.org/publ/bcbs128.pdf (accessed on 10 August 2016).

41. Cariboni, J.; Fontana, A.; Langedijk, S.; Maccaferri, S.; Pagano, A.; Petracco-Giudici, M.; Rancan, M.; Schich, S. Reducing and sharing the burden of bank failures. *OECD J. Financ. Mark. Trends* **2015**, *2*, 29–61. [CrossRef]

42. Muresano, R.; Pagano, A. Adapting and Optimizing the Systemic Model of Banking Originated Losses (SYMBOL) Tool to the Multi-core Architecture. *Comput. Econ.* **2016**, *48*, 253–280. [CrossRef]

43. Saona, P. Determinants of the Profitability of the US Banking Industry. *Int. J. Bus. Soc. Sci.* **2011**, *2*, 255–269.

44. Andreß, H.-J.; Golsch, K.; Schmidt, A.W. *Applied Panel Data Analysis for Economic and Social Surveys*; Springer: Berlin/Heidelberg, Germany, 2013.

45. Čihák, M.; Schaeck, K. How Well Do Aggregate Prudential Ratios Identify Banking System Problems? *J. Financ. Stab.* **2010**, *6*, 130–144. [CrossRef]

46. Leung, W.S.; Taylor, N.; Evans, K.P. The determinants of bank risks: Evidence from the recent financial crisis. *J. Int. Financ. Mark. Inst.* **2015**, *34*, 277–293. [CrossRef]

47. De Nicoló, G.; Geadah, S.; Rozhkov, D. Bridging the great divide: Poorly developed financial systems in the CIS-7 countries may jeopardize their sustained growth. *Financ. Dev.* **2003**, *40*, 42–45.

48. Blasco, M.; Sinkey, J. Bank asset structure, real-estate lending, and risk-taking. *Q. Rev. Econ. Financ.* **2006**, *46*, 53–81. [CrossRef]

49. Männasoo, K.; Mayes, D.G. Explaining bank distress in Eastern European transition economies. *J. Bank. Financ.* **2009**, *33*, 244–253. [CrossRef]

50. Jokipii, T.; Milne, A. Bank capital buffer and risk adjustment decisions. *J. Financ. Stab.* **2011**, *7*, 165–178. [CrossRef]

51. Ghosh, A. Banking-industry specific and regional economic determinants of non-performing loans: Evidence from US states. *J. Financ. Stab.* **2015**, *20*, 93–104. [CrossRef]

52. Acosta, J.; Grill, M.; Hannes, J. *The Leverage Ratio, Risk-Taking and Bank Stability*; Working Paper Series, No. 2079; European Central Bank: Frankfurt, Germany, 2017.

53. Klomp, J.; Haan, J. Banking risk and regulation: Does one size fit all? *J. Bank. Financ.* **2012**, *36*, 3197–3212. [CrossRef]

54. Shehzad, C.T.; Haan, J.; Scholtens, B. The impact of bank ownership concentration on impaired loans and capital adequacy. *J. Bank. Financ.* **2010**, *34*, 399–408. [CrossRef]

55. Vazquez, F.; Federico, P. Bank funding structures and risk: Evidence from the global financial crisis. *J. Bank. Financ.* **2015**, *61*, 1–14. [CrossRef]

56. Schmieder, C.; Schumacher, L. Introduction to the balance sheet-based approach to stress testing. In *A Guide to IMF Stress Testing: Methods and Models*; Ong, L.L., Ed.; International Monetary Fund: Washington, DC, USA, 2014; pp. 13–16.

Article

Maslow Portfolio Selection for Individuals with Low Financial Sustainability

Zongxin Li [1], Xinge Li [1], Yongchang Hui [1,*] and Wing-Keung Wong [2,3,4]

[1] School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China; lzx@stu.xjtu.edu.cn (Z.L.); xinge_li@hotmail.com (X.L.)
[2] Department of Finance, Fintech Center, and Big Data Research Center, Asia University, Taichung City 413, Taiwan; wong@asia.edu.tw
[3] Department of Economics and Finance, Hang Seng Management College, Hong Kong, China
[4] Department of Economics, Lingnan University, Hong Kong, China
* Correspondence: huiyc180@xjtu.edu.cn

Abstract: In this paper, we extend Maslow's need hierarchy theory and the two-level optimization approach by developing the framework of the Maslow portfolio selection model (MPSM) by solving the two optimization problems to meet the need of individuals with low financial sustainability who prefer to satisfy their lower-level (safety) need first, and, thereafter, look for higher-level (self-actualization) need to maximize the optimal return. We illustrate our proposed model with real American stock data from the S&P index and conduct the out-of-sample analysis to compare the performance of our proposed Variance-CVaR (conditional value-at-risk) MPSM with both traditional mean-variance and mean-CVaR models. Our empirical analysis shows that our proposed Variance-CVaR MPSM is not only sustainable, but also obtains the best out-of-sample performance in the sense that the optimal portfolios obtained by using our proposed Variance-CVaR MPSM obtain the highest cumulative returns in the out-of-sample period among the models used in our paper. We note that our proposed model is not only suitable to individuals with low financial sustainability, but also suitable to institutions or investors with high financial sustainability.

Keywords: portfolio selection; need hierarchy theory; two-level optimization; variance; coherent risk measures

JEL Classification: G11; C61; C44

---

1. Introduction

There is an increasing number of organizations, especially governments, who take financial sustainability into account. Traditionally, financial sustainability can be defined as the ability of government to finance the provision of public services at present without compromising the ability to do so in the future [1–3]. In the latest decade, empirical research has been conducted on the financial sustainability of Microfinance institutions [4–6]. Nurmakhanova et al. [7] claim that sustainable Microfinance institutions are the ones that operate profitably and do not require subsidies to succeed. They show that focusing on financial sustainability does not necessarily hurt the depth and breadth of outreach (social mission).

However, financial sustainability is ignored by conventional portfolio selection models. The mean-risk model is the most widely used framework in the portfolio selection. In the pioneer work of the mean-variance (MV) model, Markowitz [8] uses the variance of a portfolio return as risk function. As a measure of dispersion about the expected mean, variance is desirable risk measure for investors to evaluate their risk, especially for individuals with low financial sustainability. However, variance

has a serious shortcoming that it equally penalizes returns on both sides of the distribution, limiting its application to jointly elliptical distributions [9]. To circumvent the limitation, Artzner et al. [10] introduce the coherent risk measure via an axiomatic method and discuss that a good risk measure should satisfy monotonicity, translation invariance, positive homogeneity, and subadditivity.

Recently, the seemingly large daily price movements in high-tech stocks have further generated investigation into the market risk. Extreme price movements in the financial markets are rare, but extraordinarily important for investors and financial institutions with both high and low financial sustainability. Value-at-Risk (VaR) [11] is one well-known tool to take care of the left tails of distributions (extremely unfavorable outcomes). Nonetheless, VaR is not a coherent risk measure. To overcome this limitation, conditional VaR (CVaR) becomes popular because CVaR possesses some attractive theoretical properties including being coherent, consistent with second-order stochastic dominance [12], and easy to operate [13]. Dobrovolskiene and Tamosiuniene [14] develop a sustainability-oriented model of financial resource allocation in a project portfolio by integrating a composite sustainability index of a project into Markowitz's classical mean-variance model.

Investors could consider different kinds of risk in the portfolio selection according to different sustainability. This idea is consistent with the need hierarchy theory proposed in [15]. De Brouwer [16] offers an alternative formulation of the behavioural portfolio theory via the theory of needs, and Colson et al. [17] develop the approach for the two-level optimization estimation. In this paper, we extend their theories by developing the framework of the Maslow portfolio selection model (MPSM) and solving the two optimization problems to meet the need of individuals with low financial sustainability who prefer to satisfy their lower-level (safety) need first, and thereafter, look for higher-level (self-actualization) need to maximize the optimal return. In addition, we develop a special MPSM call Variance-CVaR MPSM.

In order to check whether our proposed model is superior to the traditional MV model in [8] and the mean-CVaR (M-CVaR) model in [13], we illustrate our proposed model with real American stock data from the S&P index and conduct the out-of-sample analysis to compare the performance of our proposed Variance-CVaR MPSM with both MV and M-CVaR models. Our empirical analysis shows that our proposed Variance-CVaR MPSM is not only sustainable, but also obtains the best out-of-sample performance in the sense that the optimal portfolios obtained by using our proposed Variance-CVaR MPSM obtain the highest cumulative returns in the out-of-sample period among the models used in our paper.

Our findings support our conjecture that individuals with low financial sustainability who prefer to satisfy their lower-level (safety) needs first, and, thereafter, look for higher-level (self-actualization) needs do prefer our proposed models to both MV and M-CVaR models. We also find that institutions or investors with high financial sustainability who prefer to look for higher-level (self-actualization) needs first, and, thereafter, satisfy their lower-level (safety) needs also prefer our proposed models to both MV and M-CVaR models in our illustration.

The rest of the paper is organized as follows. In Section 2, we do a quick review on the coherent risk measure. The Variance-Coherent Maslow portfolio selection model are proposed in Section 3. Empirical study on the American stock market is carried out in Section 4. Section 5 draws inferences and conclusions from the findings in this paper.

2. Coherent Risk Measures

Artzner et al. [10] define the coherent risk measure as follows:

Definition 1. *A risk measure $\rho : \mathcal{L}_\infty \to \mathbb{R}$ is called* **coherent** *if it satisfies the following four axioms:*

1. **Monotonicity:** *for all $X, Y \in \mathcal{L}_\infty$ with $X \leq Y$, $\rho(Y) \leq \rho(X)$;*
2. **Translation invariance:** *for all $X \in \mathcal{L}_\infty$ and $\alpha \in \mathbb{R}$, $\rho(X + \alpha) = \rho(X) - \alpha$;*
3. **Positive homogeneity:** *for all $X \in \mathcal{L}_\infty$ and $\lambda \geq 0$, $\rho(\lambda X) = \lambda \rho(X)$; and*
4. **Subadditivity:** *for all $X, Y \in \mathcal{L}_\infty$, $\rho(X + Y) \leq \rho(X) + \rho(Y)$.*

Applying the Separation Theorem for convex sets, we obtain the following general representation for all coherent risk measures in "generalized scenarios":

Proposition 1. *A risk measure* $\rho : \mathcal{L}_\infty \to \mathbb{R}$ *is called* **coherent** *if and only if there exists a family* $\mathcal{P}$ *of probability measures on the set of states of nature such that*

$$\rho(X) = \sup_{\mathbb{P} \in \mathcal{P}} \left\{ E_{\mathbb{P}}(-X) \right\}. \tag{1}$$

Using Definition 1 for the coherent risk measure, some specific risk measures have been proposed [18,19]. One of the popular measures is the conditional value-at-risk (CVaR) developed in [13,20]. To be specific, for any $\alpha \in (0,1]$, we have

$$CVaR_\alpha(X) = \inf_{t \in \mathbb{R}} \{ t + \frac{1}{1-\alpha} E[-X-t]_+ \}. \tag{2}$$

This coherent risk measure not only has a clear interpretation, but also obtains some good properties. For example, the ranking of assets by using the mean return and CVaR is equivalent to that by using the second-order stochastic dominance [12]. Furthermore, CVaR is the best convex approximation to VaR [21]. In this paper, in order to describe investors' subjective risk preference better, we propose to use the **spectral risk measure**, which is defined as follows:

Definition 2. *A* **spectral risk measure** $\rho_\phi : \mathcal{L}_\infty \to \mathbb{R}$ *is defined as*

$$\rho_\phi(X) = - \int_0^1 \phi(p) F_X^{-1}(p) dp, \tag{3}$$

in which $\phi : [0,1] \to [0,1]$ *is called a* **spectral function** *satisfying*

1. $\phi(p) \geq 0$ *for all* $p \in (0,1]$;
2. $\phi(p_1) \geq \phi(p_2)$ *for all* $p_1, p_2 \in (0,1]$ *with* $p_1 < p_2$; *and*
3. $\int_0^1 \phi(p) dp = 1$.

3. Variance-Coherent Maslow Portfolio Selection Model

In this section, we first recall the framework of the Maslow portfolio selection model, and, thereafter, set the specific model with coherent risk measures. Assume that there are n risky assets $S = (s_1, ..., s_n)'$ whose random returns are denoted by $R = (R_1, ..., R_n)'$ with mean $\mu = (\mu_1, ..., \mu_n)'$ and covariance matrix $\Sigma = (\sigma_{ij})_{n \times n}$. Denote the portfolio held by an investor as $x = (x_1, ..., x_n)'$ with $\sum_{i=1}^n x_i = 1$, and the random return rate of her/his portfolio is $\sum_{i=1}^n R_i x_i$ with mean $\sum_{i=1}^n \mu_i x_i$ and variance $x'\Sigma x$.

3.1. Framework of MPSM

According to the hierarchy theory proposed in [15], the safety and self-actualization needed are closely related to investment activity such that investors would like to ensure the safety of their initial wealth first, and, thereafter, gain as much profit as possible. Maslow suggests that the safety needs (lower-level needs) must be fulfilled before the individuals desire for self-actualization needs (higher-level needs). This hierarchy can be modelled by using the two-level optimization in [17], in which one optimization problem is embedded within another optimization problem.

In this paper, we extend the theory developed in [15] and the two-level optimization approach developed in [17] to obtain the framework of MPSM by solving the following two optimization

problems to meet the need of individuals with low financial sustainability who prefer to satisfy their lower-level (safety) need first, and, thereafter, look for higher-level (self-actualization) need:

$$\textbf{upper-level:} \quad \max \quad \sum_{i=1}^{n} \mu_i x_i, \tag{4}$$

$$\text{s.t.} \quad r_u(x) \leq L, \tag{5}$$

$$r_l(x) \leq f^* + \delta, \tag{6}$$

$$\sum_{i=1}^{n} x_i = 1, \tag{7}$$

$$x_i \geq 0, \ i = 1, \cdots, n, \tag{8}$$

where f^* is the optimal value obtained from solving the following optimization problem:

$$\textbf{lower-level:} \quad \min \quad r_l(y), \tag{9}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} \mu_i y_i \geq G, \tag{10}$$

$$\sum_{i=1}^{n} y_i = 1, \tag{11}$$

$$y_i \geq 0, \ i = 1, \cdots, n. \tag{12}$$

Here, two risk measures, $r_l(x)$ and $r_u(x)$, depend on portfolio composition x and describe the risks corresponding to the lower-level (safety) needs and higher-level (self-actualization) needs, and the two preset parameters, G and L, are the minimum return and maximum loss investors could accept, respectively. Short-selling is restricted in the MPSM to stabilize the portfolio problem [22].

Constraint (6) links the lower-level (safety) needs to the higher-level (self-actualization) needs by f^*. δ is viewed as a tolerance parameter for the safety requirement. The greater the δ, the lower the requirement for safety. Constraint (6) is designed to fit the needs of individuals with low financial sustainability who prefer to satisfy their lower-level (safety) needs first, and thereafter, look for higher-level (self-actualization) needs.

Remark 1. *The two-level optimization problem stated in (4)–(12) can be viewed as the $\delta-$constraint method for solving bi-objective programming [23].*

Remark 2. *Since Equations in (4)–(12) do not give any restriction on the forms of both $r_l(x)$ and $r_u(x)$, investors have all the freedom and flexibility to construct different MPSMs.*

3.2. Variance-Coherent MPSM

Since our proposed model allows investors to have all the freedom and flexibility to construct different MPSMs, in this section, we propose the following particular MPSM: when setting $r_l(x)$ and $r_u(x)$ as variance and coherent risk measures, respectively, we obtain the concrete portfolio selection model, which is called Variance-Coherent MPSM under the framework of MPSM as stated in the following:

$$\textbf{higher-level:} \quad \max \quad \sum_{i=1}^{n} \mu_i x_i, \tag{13}$$

$$\text{s.t.} \quad \rho(x) \leq L, \tag{14}$$

$$x' \Sigma x \leq f^* + \delta, \tag{15}$$

$$\sum_{i=1}^{n} x_i = 1, \tag{16}$$

$$x_i \geq 0, \ i = 1, \cdots, n, \tag{17}$$

where f^* is the optimal value obtained from solving the following optimization problem:

$$\textbf{lower-level:} \quad \min \quad y' \Sigma y, \tag{18}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} \mu_i y_i \geq G, \tag{19}$$

$$\sum_{i=1}^{n} y_i = 1, \tag{20}$$

$$y_i \geq 0, \ i = 1, \cdots, n. \tag{21}$$

There are several reasons we set the two risk functions in the above. Firstly, when pursuing higher-level needs, investors would like to use more advanced risk control methods or financial instruments as shown from Equations (13) to (17). Secondly, variance and coherent risk measures are good alternatives for investors to meet different levels of needs. On one hand, variance measures the overall volatility, and, thus, it helps to ensure the safety of the initial wealth. On the other hand, coherent risk measures depict the tail risks so as to provide investors more chance to obtain high payoffs. Thirdly, from the viewpoint of empirical comparability, setting $r_l(x)$ and $r_u(x)$ as variance and coherent risk measures makes it easier for academics and practitioners to compare our proposed model with the classical models in [8,13] through empirical analyses since they are closely related. To some extent, this chronological order provides us a criterion for choosing risk functions corresponding to different levels of needs.

4. Empirical Analyses

To illustrate the effectiveness of our proposed model stated in Equations (13)–(21) of Section 3.2, we investigate the superiority of using our proposed Variance-Coherent MPSM empirically. Here, we take the coherent risk measure as CVaR; that is to say, the spectral function is taken as $\phi(p) = \frac{1}{1-\alpha} 1_{\{0 \leq p \leq 1-\alpha\}}$. This investigation is based on the historical data of stocks from all components of the S&P 500 Index in the American stock market.

For comparison purposes, we also analyse the performance of a few other closely related models, including the classical MV and M-CVaR models in [8,13]. In essence, they depict both investors' safety needs and self-actualization needs. For convenience, we use "MV", "M-CVaR", "Maslow (0.008)", and "Maslow (0.010)" to stand for the optimal portfolios obtained under the classical MV model, M-CVaR model and our proposed model (13)–(21) with $\delta = 0.008$ and $\delta = 0.010$, respectively.

4.1. Data

To ensure decentralization, we choose 30 stocks classified into different industries from the components of the S&P 500 Index. The stocks shown in Table 1 are chosen because they have either the 30 highest market capitalizations or the 30 largest trading volumes, and, thus, they are the leading stocks in the corresponding industries.

Historical weekly returns of the aforementioned 30 stocks in six years, from 6 October 2011 to 5 October 2017, are downloaded from the Yahoo Finance (http://finance.yahoo.com/) to form the data set. Moreover, the data from 6 October 2011 to October 5 2016 are used to obtain the optimal portfolios, and the remaining data are used for evaluating the out-of-sample performance.

Table 1. Stock pool (in alphabetical order).

No.	Name	No.	Name	No.	Name
1	AAPL	11	DUK	21	MCD
2	AEP	12	F	22	MDT
3	AIG	13	FDX	23	MSFT
4	AMGN	14	GE	24	NKE
5	AXP	15	HD	25	PG
6	BA	16	HON	26	SLB
7	BAC	17	INTC	27	T
8	CAT	18	JNJ	28	UNH
9	CMCSA	19	JPM	29	WMT
10	CVX	20	KO	30	XOM

Note: These 30 stocks are numbered in alphabetical order. They are the leading ones in their corresponding industries. The time period is from 6 October 2011 to 5 October 2017.

4.2. The Characteristics of the Optimal Portfolios

In this section, we look into the characteristics of the optimal portfolios obtained by using different portfolio selection models. In this experiment, we set $\alpha = 0.95$. According to the features of historical data, we set $G = 0.001$ and $L = 0.05$. The optimal portfolios under the Variance-CVaR MPSM are obtained by following the steps described in Section 3. We summarize the characteristics of individual stocks in Table 2, including the sample mean and risk (variance and $\text{CVaR}_{0.95}$).

One can easily observe from the fifth column of Table 2 that the optimal portfolios obtained under the MV (safety needs) model are of the most diversification while the optimal portfolios displayed in the last column that are obtained under the M-CVaR (self-actualization needs) model are of the most concentration because there is only one asset in each portfolio under this category. Nonetheless, the optimal portfolios corresponding to our proposed Variance-CVaR MPSM (13)–(21) strike a balance between the most diversification and the most concentration that the optimal portfolios are diversified but not the most diversified and not the most concentrated. In addition, with the increase of the tolerance parameter δ for the safety requirement, the optimal portfolio derived from our proposed model becomes more concentrated.

All the above empirical results are in agreement with the motivations of the portfolio selection models. The well-diversified portfolios are robust with respect to the collapse of any stock, and, thus, being safe to some extent. The concentrated portfolios (Column 8) including only stocks with the largest means gain the higher returns during the formation period. The compromising portfolios derived from our new Variance-CVaR MPSM not only contain the assets with larger means to get higher returns at one hand, but also allocate the wealth to several other stocks to enhance the diversification and robustness of the portfolios on the other hand.

Table 2. The characteristics of the stocks and the optimal portfolios under different models.

No.	Mean	Variance	CVaR$_{0.95}$	MV	Maslow (0.008)	Maslow (0.010)	M-CVaR
1	0.00404	0.00151	0.07937	0.04441	0	0	0
2	0.00292	0.00050	0.04568	0	0	0	0
3	0.00478	0.00120	0.07103	0	0	0	0
4	0.00506	0.00098	0.05577	0	0.03968	0	0
5	0.00203	0.00081	0.06450	0.01195	0	0	0
6	0.00390	0.00092	0.06374	0	0	0	0
7	0.00493	0.00186	0.08423	0	0	0	0
8	0.00178	0.00123	0.07027	0	0	0	0
9	0.00497	0.00072	0.04690	0	0.18695	0.09400	0
10	0.00146	0.00089	0.06205	0	0	0	0
11	0.00204	0.00045	0.04590	0.20159	0	0	0
12	0.00184	0.00131	0.07279	0	0	0	0
13	0.00399	0.00093	0.06051	0	0	0	0
14	0.00327	0.00076	0.05318	0	0	0	0
15	0.00585	0.00067	0.04747	0	0.59555	0.81397	1.00000
16	0.00406	0.00067	0.05145	0	0	0	0
17	0.00317	0.00095	0.06752	0	0	0	0
18	0.00319	0.00033	0.03577	0.16204	0	0	0
19	0.00414	0.00109	0.07121	0	0	0	0
20	0.00170	0.00044	0.04402	0	0	0	0
21	0.00183	0.00041	0.04445	0.17576	0	0	0
22	0.00440	0.00054	0.04456	0.03585	0	0	0
23	0.00409	0.00104	0.07210	0	0	0	0
24	0.00398	0.00087	0.05899	0	0	0	0
25	0.00211	0.00037	0.04206	0.14469	0	0	0
26	0.00205	0.00130	0.07410	0	0	0	0
27	0.00244	0.00046	0.04281	0.07682	0	0	0
28	0.00496	0.00076	0.05413	0.03914	0.17783	0.09203	0
29	0.00171	0.00055	0.05434	0.09760	0	0	0
30	0.00147	0.00057	0.05289	0.01016	0	0	0

Note: MV, M-CVaR, Maslow (0.008), and Maslow (0.010) stand for the optimal portfolios obtained under the MV model, M-CVaR model, and as our proposed model (13)–(21) with $\delta = 0.008$ and $\delta = 0.010$, respectively.

4.3. Out-of-Sample Performance

During the formation period, one could conclude that the optimal portfolios obtained under the MV (safety needs) model are of the most diversification while the optimal portfolios obtained under the M-CVaR (self-actualization needs) model are of the most concentration. However, the optimal portfolios obtained from our proposed Variance-CVaR MPSMs are neither of the most diversification nor of the most concentration. Could this be used to infer that the performance of our proposed Variance-CVaR MPSMs is not good? We say that this is not true. Whether the performance of the optimal portfolios is good or not does not depend on their performance of the portfolios in the formation period (since one cannot make any cent from the formation period) but depends on their performance on the testing period (that one can make profit in the testing period). This argument is consistent with DeMiguel et al. [24], who suggest that the out-of-sample performance is one of the most optimal ways to compare different investment strategies.

Thus, in this section, we carry out the out-of-sample analysis to evaluate the practical usefulness of different portfolio selection models. To do so, we investigate the payoffs of different optimal portfolios by using their cumulative returns and exhibit the results in Figure 1, which enable us to examine the evolution of the actual out-of-sample returns over time. Obviously, for a given time, the higher the cumulative return, the better the optimal portfolio or the corresponding portfolio selection model.

From Figure 1, one can tell that the cumulative returns of the safety strategy are always the lowest. On the other hand, the out-of-sample performance of the "most promising" strategy (M-CVaR) is barely satisfactory because this strategy suffers from the maximum losses at the very beginning,

and thus, obtains the minimum gains during the final period. Nonetheless, the two optimal portfolios determined by our proposed Variance-CVaR MPSM outperform the two strategies above significantly, which can be easily observed from the curves in Figure 1 corresponding to Maslow (0.008) and Maslow (0.010). This shows that our proposed Variance-CVaR MPSM is not only sustainable, but also obtains the best performance among the four models in our analysis.
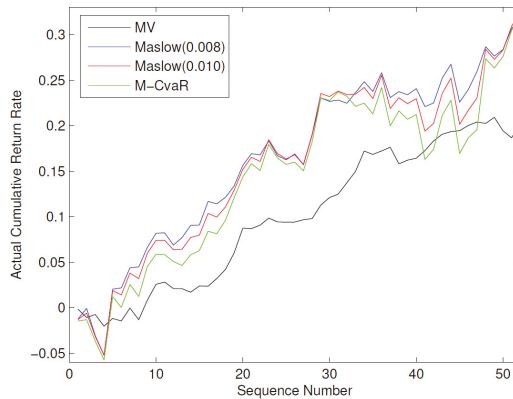
Figure 1. The cumulative return rates of the optimal portfolios got under different portfolio selection models. Note: MV, M-CVaR, Maslow (0.008) and Maslow (0.010) stand for the optimal portfolios obtained under the MV model, M-CVaR model, as well as our proposed model (13)–(21) with $\delta = 0.008$ and $\delta = 0.010$, respectively.

4.4. Further Discussion on the Results and Methods

To examine whether our proposed model outperforms the most commonly-used classical MV (safety needs) model developed in [8] and the M-CVaR (self-actualization needs) model developed in [13], we illustrate our proposed model with real American stock data from the S&P index and conduct the out-of-sample analysis to compare the performance of our proposed Variance-CVaR MPSM with both MV and M-CVaR models. Our empirical analysis shows that our proposed Variance-CVaR MPSM is not only sustainable, but also obtains the best out-of-sample performance in the sense that the optimal portfolios obtained by using our proposed Variance-CVaR MPSM obtain the highest cumulative returns in the out-of-sample period among the models used in our paper.

We note that our proposed model is mainly designed for individuals with low financial sustainability who prefer to satisfy their lower-level (safety) needs first, and, thereafter, look for higher-level (self-actualization) needs. One may believe that our proposed model is not suitable for investors with high financial sustainability. We argue that this is not true. Firstly, there are still some institutions or investors with high financial sustainability who still prefer to satisfy their lower-level (safety) needs first, and, thereafter, look for higher-level (self-actualization) needs. Our models are still suitable for these types of institutions or investors with high financial sustainability. Secondly, even for institutions or investors with high financial sustainability who prefer to look for higher-level (self-actualization) needs first, and, thereafter, satisfy their lower-level (safety) needs, our proposed models could still be suitable for them because our empirical study does find that our proposed Variance-CVaR MPSM obtains the highest cumulative returns in the out-of-sample period among the models used in our paper. Thus, we can claim that our proposed model is not only suitable for institutions or investors with high financial sustainability who still prefer to satisfy their lower-level (safety) needs first, and, thereafter, look for higher-level (self-actualization) needs, but is also suitable for institutions or investors with high financial sustainability who prefer to look for higher-level (self-actualization) needs first, and thereafter, satisfy their lower-level (safety) needs.

Nonetheless, our proposed model is designed for individuals with low financial sustainability who prefer to satisfy their lower-level (safety) needs first, and, thereafter, look for higher-level (self-actualization) needs, not for institutions or investors with high financial sustainability who prefer to look for higher-level (self-actualization) needs first, and, thereafter, satisfy their lower-level (safety) needs. One could easily modify our proposed model to obtain the model suitable for institutions or investors with high financial sustainability who prefer to look for higher-level (self-actualization) needs first, and, thereafter, satisfy their lower-level (safety) needs.

Many studies, for example, Michaud [25], have found the traditional estimation of the MV-optimized model in [8] does more harm than good for high dimensional big data. To circumvent the limitation, Bai et al. [26] develop the bootstrap-corrected estimation that could analytically circumvent the limitation and is proportionally consistent with the theoretical return parameter. Leung et al. [27] derive explicit formulas for the estimator of the optimal portfolio return. Bai et al. [28] prove that the traditional estimate for the optimal return of self-financing portfolios always over-estimates from its theoretic value. To circumvent the problem, they develop a bootstrap estimate for the optimal return of self-financing portfolios and prove that this estimate is consistent with its counterpart parameter. Bai et al. [28] further develop the spectrally-corrected estimation to improve the estimation further. Extension of our paper could incorporate both the MPSM approach developed in our paper and the bootstrap-corrected estimation or the spectrally-corrected estimation to meet the needs of individuals with low financial sustainability who prefer to satisfy their lower-level (safety) needs first, and, thereafter, look for higher-level (self-actualization) need for high-dimensional big data.

Investigating whether international diversification and home bias inertia are substitutes or complements for Americans, Abid et al. [29,30] conclude that the US investors have a 'home bias' if they prefer less risk and to be 'internationally diversified' if they prefer higher risk. Hoang et al. [31] conclude that gold is good for the diversification of stock portfolios but not for bond portfolios in the Paris gold market. On the other hand, studying the role of gold quoted on the Shanghai Gold Exchange in the diversification of Chinese portfolios, Hoang et al. [32] show that, in general, risk-averse investors prefer not to include gold while risk-seeking investors prefer to include it in their stock-bond portfolios, especially in crisis periods. They also find that risk-seekers prefer including gold in an equal-weighted portfolio while risk-averters prefer including gold in efficient portfolios. Studying the performance in the Hong Kong residential property market, Qiao and Wong [33] conclude that risk averters prefer to invest in smaller property while Tsang et al. [34] conclude that, regardless of whether the buyers eschew risk, embrace risk or are indifferent to it, they prefer to invest in smaller property. On the other hand, to study the performance of a portfolio of US/UK equities, bonds, gold, and housing, Bouri et al. [35] conclude that wine is the best investment among all individual assets under their study, and, in general, investors prefer to invest in with-wine portfolios than without-wine portfolios to gain higher expected utility. It will be interesting to readdress all the above issues and other important issues in finance and economics by using the theory developed in our paper, in order to check whether investors' preferences change when they adopt our proposed model.

5. Conclusions

In this paper, we extend the need hierarchy theory developed in [15] and the two-level optimization approach developed by Colson et al. in [17] to obtain the framework of the Maslow portfolio selection (Variance-CVaR MPSM) model by solving the two optimization problems to meet the need of individuals with low financial sustainability who prefer to satisfy their lower-level (safety) needs first, and, thereafter, look for higher-level (self-actualization) needs. Investors with low financial sustainability pay more attention to the safety of their financial investment. Investors are more concerned about the basic or lower level needing to be satisfied first, and, thereafter, care about their upper level needs to gain a higher return. In particular, investors like to use diversified strategy to obtain an efficient instrument to manage risk so that they could make the risk level affordable.

To do so, we first solve the optimization problem to minimize the risk first, and, thereafter, solve the optimization problem to maximize the optimal return.

Further study includes the extension of our new portfolio selection model to the multi-period case because investors tend to adjust their investment strategies with the emergence of new information, especially for the investors with strong financial sustainability to adopt long-term investment. In this paper, we use variance or CVaR as risk. Extensions could use other risk measures to represent risk, including the Sharpe Ratio in [36], mixed Sharpe ratio in [37], mean-variance-ratio in [38], Omega ratio in [39], Kappa Ratios in [40], and Farinelli and Tibiletti ratio in [41] to represent risk.

Author Contributions: Formal analysis: Zongxin Li, Xinge Li, Yongchang Hui, Wing-Keung Wong; Methodology: Zongxin Li, Yongchang Hui; Project administration: Yongchang Hui, Wing-Keung Wong; Software: Zongxin Li, Xinge Li; Writing-original draft: Zongxin Li, Xinge Li, Yongchang Hui, Wing-Keung Wong; Writing-review and editing: Zongxin Li, Xinge Li, Yongchang Hui, Wing-Keung Wong.

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

References

1. Canadian Institute of Chartered Accountants (CICA). *Indicators of Government Financial Condition*; CICA: Toronto, ON, Canada, 1997.
2. Bath, P. *Resource for Success Series*, 2nd ed.; USA Nature Conservancy Arlington: Arlington, VA, USA, 2001.
3. Stavins, R.; Wagner, A.; Wagner, G. Interpreting sustainability in economic terms: Dynamic efficiency plus intergenerational equity. *Econ. Lett.* **2003**, *79*, 339–343.
4. Crombrugghe, A.; Tenikue, M.; Sureda, J. Performance analysis for a sample of microfinance institutions in India. *Ann. Public Cooper. Econ.* **2008**, *79*, 269–299.
5. Cull, R.; Demirguc-Kunt, A.; Morduch, J. Does microfinance regulation curtail profitability and outreach? *World Dev.* **2011**, *39*, 949–965.
6. Hermes, N.; Lensink, R.; Meesters, A. Outreach and efficiency of microfinance institutions. *World Dev.* **2011**, *39*, 938–948.
7. Nurmakhanova, M.; Kretzschmar, G.; Fedhila, H. Trade-off between financial sustainability and outreach of microfinance institutions. *Eurasian Econ. Rev.* **2015**, *5*, 231–250.
8. Markowitz, H.M. Portfolio selection. *J. Financ.* **1952**, *7*, 77–91.
9. Szego, G. Measures of risk. *J. Bank. Financ.* **2002**, *26*, 1253–1272.
10. Artzner, P.; Delbaen, F.J.; Eber, M.; Health, D. Coherent measures of risk. *Math. Financ.* **1999**, *9*, 203–228.
11. Morgan, J.P. *Risk Metrics Technical Document*, 4th ed.; Morgan Guaranty Trust Company: New York, NY, USA, 1996.
12. Ma, C.; Wong, W.K. Stochastic dominance and risk measure: A decision-theoretic foundation for VaR and C-VaR. *Eur. J. Oper. Res.* **2010**, *207*, 927–935.
13. Rockafellar, R.T.; Uryasev, S. Conditional value-at-risk for general loss distributions. *J. Bank. Financ.* **2002**, *26*, 1443–1471.
14. Dobrovolskiene, N.; Tamosiuniene, R. Sustainability-oriented financial resource allocation in a project portfolio through multi-criteria decision-making. *Sustainability* **2016**, *8*, 485.
15. Maslow, A.H. A theory of human motivation. *Psychol. Rev.* **1943**, *50*, 370–396.
16. De Brouwer, P.J.S. Maslowian portfolio theory: An alternative formulation of the behavioural portfolio theory. *J. Asset Manag.* **2009**, *9*, 359–365.
17. Colson, B.; Marcotte, P.; Savard, G. An overview of two-level optimization. *Ann. Oper. Res.* **2007**, *59*, 617–630.
18. Acerbi, A.; Tasche, D. On the coherence of expected shortfall. *J. Bank. Financ.* **2002**, *26*, 1487–1503.

19. Chen, Z.; Hu, Q. On coherent risk measures induced by convex risk measures. *Methodol. Comput. Appl. Probab.* **2017**, *1*, 1–26.
20. Rockafellar, R.T.; Uryasev, S. Optimization of conditional value-at-risk. *J. Risk* **2000**, *2*, 21–41.
21. Chen, W.; Sim, M.; Sun, J.; Teo, C.P. From CVaR to uncertainty set: Implications in joint chance-constrained optimization. *Oper. Res.* **2010**, *58*, 470–485.
22. Frost, P.A.; Savarino, J.E. For better performance, constraint portfolio weights. *J. Portf. Manag.* **1988**, *15*, 29–34.
23. Li, D.; Yang, J.B.; Biswal, M.P. Quantitative parametric connections between methods for generating noninferior solutions in multiobjective optimization. *Eur. J. Oper. Res.* **1999**, *117*, 84–99.
24. DeMiguel, V.; Garlappi, L.; Uppal, R. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financ. Stud.* **2009**, *22*, 1915–1953.
25. Michaud, R.O. The Markowitz optimization enigma: Is 'optimized' optimal? *Financ. Anal. J.* **1989**, *45*, 31–42.
26. Bai, Z.D.; Lui, H.X.; Wong, W.K. Enhancement of the applicability of markowitz's portfolio optimization by utilizing random matrix theory. *Math. Financ.* **2009**, *19*, 639–667.
27. Leung, P.L.; Ng, H.Y.; Wong, W.K. An improved estimation to make markowitz's portfolio optimization theory users friendly and estimation accurate with application on the US stock market investment. *Eur. J. Oper. Res.* **2012**, *222*, 85–95.
28. Bai, Z.D.; Lui, H.X.; Wong, W.K. On the markowitz mean-variance analysis of self-financing portfolios. *Risk Decis. Anal.* **2009**, *1*, 35–42.
29. Abid, F.; Mroua, M.; Wong, W.K. Should americans invest internationally? The mean-variance portfolios optimization and stochastic dominance approaches. *Risk Decis. Anal.* **2013**, *4*, 89–102.
30. Abid, F.; Leung, P.L.; Mroua, M.; Wong, W.K. International Diversification versus Domestic diversification: Mean-variance portfolio optimization and stochastic dominance approaches. *J. Risk Financ. Manag.* **2014**, *7*, 45–66.
31. Hoang, T.H.V.; Lean, H.H.; Wong, W.K. Is gold good for portfolio diversification? A stochastic dominance analysis of the Paris stock exchange. *Int. Rev. Financ. Anal.* **2015**, *42*, 98–108.
32. Hoang, T.H.V.; Wong, W.K.; Zhu, Z.Z. Is gold different for risk-averse and risk-seeking investors? An empirical analysis of the Shanghai Gold Exchange. *Econ. Model.* **2015**, *50*, 200–211.
33. Qiao, Z.; Wong, W.K. Which is a better investment choice in the Hong Kong residential property market: A big or small property? *Appl. Econ.* **2015**, *47*, 1670–1685.
34. Tsang, C.K.; Wong, W.K.; Horowitz, I. Arbitrage opportunities, efficiency, and the role of risk preferences in the Hong Kong property market. *Stud. Econ. Financ.* **2016**, *33*, 735–754.
35. Bouri, E.; Gupta, R.; Wong, W.K.; Zhu, Z.Z. Is wine a good choice for investment? *Pac. Basin Finance J.* **2018**, in press.
36. Leung, P.L.; Wong, W.K. On testing the equality of the multiple Sharpe Ratios, with application on the evaluation of iShares. *J. Risk* **2008**, *10*, 1–16.
37. Wong, W.K.; Wright, J.A.; Yam, S.C.P.; Yung, S.P. A mixed sharpe ratio. *Risk Decis. Anal.* **2012**, *3*, 37–65.
38. Bai, Z.D.; Hui, Y.C.; Wong, W.K.; Zitikis, R. Evaluating prospect performance: Making a case for a non-asymptotic UMPU test. *J. Financ. Econ.* **2012**, *10*, 703–732.
39. Guo, X.; Jiang, X.J.; Wong, W.K. Stochastic dominance and omega ratio: Measures to examine market efficiency, arbitrage opportunity, and anomaly. *Economies* **2017**, *5*, 38.
40. Niu, C.Z.; Wong, W.K.; Xu, Q.F. Kappa ratios and (higher-order) stochastic dominance. *Risk Manag.* **2017**, *19*, 245–253.
41. Niu, C.Z.; Wong, W.K.; Zhu, L.X. *Farinelli and Tibiletti Ratio and Stochastic Dominance*; MPRA Paper No. 82737; University Library of Munich: Munich, Germany, 2017.

The Study of Utility Valuation of Single-Name Credit Derivatives with the Fast-Scale Stochastic Volatility Correction

Shican Liu [1], Yanli Zhou [2,*], Benchawan Wiwatanapataphee [1], Yonghong Wu [1] and Xiangyu Ge [3,*]

[1]   Department of Mathematics and Statistics, Curtin University, Perth, WA 6845, Australia;
      shican.liu@curtin.edu.au (S.L.); benchawan.wiw@curtin.edu.au (B.W.); y.wu@curtin.edu.au (Y.W.)
[2]   School of Finance, Zhongnan University of Economics and Law, Wuhan 430073, China
[3]   Department of Finance, Wuhan Technology and Business University, Wuhan 430065, China
*   Correspondence: ylzhou8507@gmail.com (Y.Z.); xiangyu_ge@hotmail.com (X.G.)

Abstract: In this paper, we study the risk aversion on valuing the single-name credit derivatives with the fast-scale stochastic volatility correction. Two specific utility forms, including the exponential utility and the power utility, are tested as examples in our work. We apply the asymptotic approximation to obtain the solution of the non-linear PDE, and make a comparison of the utility before and after the stochastic volatility modification, and we find that incorporation of fast-scale volatility will lower down the utility. By using the indifference price, we also give the yield spread impacted by the risk adverse valuation. We find that by considering the default risk, yield spread is sloping in a short period and converge in a long run.

Keywords: utility; credit derivatives; stochastic volatility; asymptotic approximation; risk aversion

1. Introduction

Credit risk, which is also known as the default risk, is the uncertainty of a firm's ability in servicing its debts and obligations. To pursue a better investment in financial contracts, it is essential but challenging to predict whether the contract linked company will default or not, which primarily explains the necessity of a risk premium. Consequently, the last few decades have witnessed the rapid development of the defaultable instruments. In recent years, due to the more and more frequent credit crisis worldwide, financial models become more complicated than ever. In particularly, the Asian financial crisis in 1997 and subprime lending crisis in 2007 have significantly raised the awareness of both regulators and academics on the evaluation of credit risk. As such, to achieve more rapid and effective management of the credit risk, more sophisticated approaches and quantitative technology are desperately required. Needless to say, conducting assessment on the credit risk is also vital for the Chinese banking system. The excellent work of Tan and Floros [1] confirmed this point of view using three various efficiency indexes and four risk indicators The reported result suggested that the credit risk played a crucial role in the entire Chinese banking industry and therefore various factors affecting the credit risk should be well identified . The efficiency and risk features of the Chinese Bank industry is also studied from an econometric point of view by them [2].

The default risk has been under investigation for a few decades. Dilip B. and M.Unal decomposed default risk into two components, i.e., timing risk and recovery risk. Subsequently, they priced the two components in future's market, and developed an estimation strategies to evaluate the recovery risks and timing risk [3]. Intensity based term structure model of the credit risk was proposed and studied by Jarrow, Lando and Turnbull. In their work both the default-free term structure and risky debt term structure were specified for a comprehensive study of the corporate debt [4].

David Lando et al. developed a model to incorporate stochastic transition intensities, and had proved that their framework could address the technical issues of modelling credit risk [5]. Hao and Zhang also contributed to the recent advance in credit risk modelling [6], in which they established a new model including the Black-sholes Merton framework, individual reduced form level, and the portfolio reduced model. Generally, the traditional Black-Scholes-Merton model (1973) [7] was based on a complete financial market, in which the payoff of the derivatives could be replicated by a certain trading strategy. However, markets in real world are never complete, and thereby market friction always exists. If unpredictable default occurred, almost all classical approaches failed, and accordingly new dynamic pricing rules were urgently needed. The work of Sircar and Zariphopoulou (2007) [8] provided insight into the risk aversion on the valuation of credit derivatives applying the utility-indifference valuation in intensity-based models where the single-name defaultable bonds and a simple representative two-name credit derivative were analysed. Later, Papageorgiou and Sircar (2008) [9] extended the work to the multiname CDOs.

In this work, we looked at the credit risk pricing problem in the framework of the structural model and utility-based portfolio selection, as the payoff of financial derivatives might be replicated by varying trading strategies of the underlying assets in a complete financial market. The issue of the portfolio optimization had a long history that dated back to 1971 [10], in which the author provided an explicit scheme to allocate investment capital between risky stocks and riskless bond. Within this framework, the underlying asset was driven by a stochastic process, which was later known as the Black-sholes model. Nonetheless, the chief disadvantage of the Black-sholes and Merton's model was the over-restrictive assumptions, especially the ones of constant interest rate and constant volatility. A great number of extensions had been made in the future research. Heston (1993) [11] took into account the stochastic volatility and derived a semi-analytic solution for the European call option by introducing a characteristic function, allowing the arbitrary correlation between the volatility and asset price. Longstaff and Schwartz (1995) incorporated stochastic short-term interest rate, which they found was negatively correlated to the asset value process [12]. Fouque et al. (2003) [13] developed an effective approximation of the option pricing problem through the incorporation of the multiscale volatility. However, the corresponding partial differential equation for option pricing was always linear, while the equation related to the optimal control problem was non-linear. For this reason, the asymptotic theory was extended to estimate the non-linear pricing problem by Fouque et al. (2015) [14].

The valuation mechanism used in our work is called indifference prices. The so called indifference price is the amount of capital that the investor pays today, so that difference between holding or not holding the derivatives was trivial. The indifference approach was first introduced by Hodges and Neuberger (1989) [15] and extended by Davis and Yoshikawa (2012) [16]. Its mechanism was based on the utility function that was a twice continuously-differentiable one strictly increasing and concave. Herein we considered the risk attitude of individuals by applying the utility based models, and specifically assessed the single-name credit default swap (CDS) that , could be treated as an insurance against the default of a reference entity. CDS is written on a single-bond issued by a reference entity. The buyer pays the seller a risk premium regularly and they in turn will get compensation if default happens. More details can be found in the work of Papageorgiou and Sircar (2008) [9].

In comparison with the aforementioned work, our work mainly features the following aspects. Firstly, we studied the credit-derivatives pricing considering the impact of both the default risk and fastscale stochastic volatility. Moreover, the problem is solved within the framework of utility-based portfolio selection, which might lead to a high dimensional non-linear partial differential equation (PDE). As the closed-form solution of high dimensional non-linear PDE was hard to be solved via existing methods, we accordingly applied asymptotic approximation to decrease the high dimensional non-linear PDE into low dimensional PDEs. Last but not least, we exhibited our results in two specific cases and numerically analyse them.

The rest of the paper is organised as follows. We established the model in Section 2. In Section 3, we applied the perturbation asymptotic method to approximate the explicit solution of our non-linear

PDE. In Section 4, we derived the full solution for the case with constant intensity process. In Section 5, we presented two special utilities and study it numerically, and also investigated the value function, maximizer and the yield spread. We concluded this work and suggested a few future works in the last section.

2. Model Setup

Generally, there are two approaches for pricing credit derivatives, including the structural model and the intensity-based model. Our work here is mainly based on the intensity-based model (or reduced form model), in which defaults happen at the jump process of poisson intensity. We start our model with simple singlename defaultable bonds with fast stochastic volatilities and then extend it to multi-name and multi-scale cases.

Unlike the traditional structural model, our model is based on the assumption that default happens at an unpredictable stopping time τ with stochastic intensity process λ, which incorporates information from the firm's stock price S and is called a hybrid model. The stock price S follows a geometric Brownian motion with the intensity process $\lambda(Z_t)$, where $\lambda(\cdot)$ is a non-negative, locally Lipschitz, smooth and bounded function. Our model takes the following form:

$$\frac{dS_t}{S_t} = \mu(Y_t)dt + \sigma(Y_t)dW_t, \tag{1}$$

$$dY_t = \frac{1}{\xi}b(Y_t)dt + \frac{1}{\sqrt{\xi}}a(Y_t)dW_t^{(1)}, \tag{2}$$

$$dZ_t = g(Z_t)dt + c(Z_t)dW_t^{(2)}, \tag{3}$$

where the Browning motion $W_t, W_t^{(1)}, W_t^{(2)}$ are correlated as follows:

$$Cov(W_t, W_t^{(1)}) = \rho_1, Cov(W_t, W_t^{(2)}) = \rho_2, Cov(W_t^{(1)}, W_t^{(2)}) = \rho_{12}. \tag{4}$$

ρ_1 measures the correlation between the Brownian motion for volatility Y and the Brownian motion for stock prices, ρ_2 measures the instantaneous correlation between the Brownian motion for the stock price S and the Brownian motion for the intensity process Z, and they satisfy $|\rho_1| < 1, |\rho_2| < 1, |\rho_{12}| < 1$, and $1 + 2\rho_1\rho_2\rho_{12} - \rho_1^2 - \rho_2^2 - \rho_{12}^2 > 0$. When the parameter ξ is small, the stochastic processes Y_t and Z_t represent the fast volatility process and the intensity process, respectively. Here we assume that $Y_{\frac{t}{\xi}}^{(1)}$ is an ergodic diffusion process and has the same unique invariant distribution as Y_t, and for more details we refer the reader to Section 4 of the reference due to Fouque et al. (2011) [17]. The drift part of dY_t is always assumed to be mean-reverted with the long term drift θ, while the volatility of volatility could be a constant σ so that the underlying distribution of dY_t is a normal distribution. However, other specific forms can also be fit in volatility, like CIR process, $\frac{3}{2}$ stochastic volatility process and $\frac{4}{2}$ stochastic volatility process. In our work, we assume the constant volatility of volatility in terms of simplicity. The default time τ of the firm is defined by the first time when the cumulated intensity reaches the random threshold ε.

$$\tau_t = inf\{s \geq t : \int_t^s \lambda(Z_s)ds = \varepsilon\}. \tag{5}$$

2.1. Maximal Expected Utility Problem

Let X_t be the wealth process and π_t denote the money we invest in the stock at time t, where $t \in [0, T], t < \tau \wedge T$, then the wealth process is as follows:

$$dX_t = \pi_t \frac{dS_t}{S_t} + r(X_t - \pi_t)dt$$
$$= (rX_t + \pi_t(\mu(Y_t) - r))dt + \pi_t\sigma(Y_t)dW_t, \tag{6}$$

where π_t is $\mathcal{F}_t$-measurable and satisfies the integrability constraint $E\int_0^T \pi_s^2 ds < \infty$. Under the utility form $\tilde{U}(X)$, the maximum expected utility payoff takes the general form of

$$sup_{\pi_t \in \mathcal{A}} E\{\tilde{U}(e^{-rT}X_T)\mathbf{1}_{\{\tau > T\}} + \tilde{U}(e^{-r\tau}X_\tau)\mathbf{1}_{\{\tau \leq T\}}\}, \tag{7}$$

where $\mathcal{A}$ is the set of π.

To simplify the formulation, we denote $e^{-rT}X_t$ by X_t and $\mu - r$ by μ, then the wealth process can be described by

$$dX_t = \pi_t \mu(Y_t, Z_t)dt + \pi_t \sigma(Y_t, Z_t)dW_t^{(1)}. \tag{8}$$

If default happens, stock of the firm cannot be traded, and investors have to liquidate holdings in the stock and deposit them in the bank account. For simplicity, we assume that the investors get full amount of the liquidated pre-default stocks and invest all of them into the bank account. Therefore, we obtain

$$X_T = X_\tau e^{r(T-\tau)}. \tag{9}$$

The problem here is to maximize the expected utility payoff at time zero, which takes the form as follows:

$$V(t, x, y, z) = sup_{\pi \in \mathcal{A}} E\{\tilde{U}(X_T)\mathbf{1}\{\tau_t > T\} + \tilde{U}(X_{\tau_t})\mathbf{1}\{\tau_t \leq T\} \mid X_t = x, Y_t = y, Z_t = z\} \tag{10}$$

Proposition 1. *The HJB equation of the value function is*

$$V_t + \frac{1}{\xi}\mathcal{L}\dagger V + V + \frac{1}{\sqrt{\xi}}\rho_{12}a(y)c(z)V_{yz} + max\left\{\pi\mu(y)V_x + \frac{1}{2}\pi^2\sigma^2(y)V_{xx}\right.$$
$$\left. + \frac{1}{\sqrt{\xi}}\pi\rho_1\sigma(y)a(y)V_{xy} + \rho_2 V_{xz}\pi\sigma(y)c(z)\right\} + \lambda(z)(\tilde{U}(x) - V) = 0 \tag{11}$$

with $V(T, x, y, z) = \tilde{U}(x)$ and the operators $\mathcal{L}\dagger$ and $\mathcal{L}\ddagger$ are defined by

$$\mathcal{L}\dagger = b(y)\frac{\partial}{\partial y} + \frac{1}{2}a^2(y)\frac{\partial^2}{\partial y^2} \tag{12}$$

$$\mathcal{L}\ddagger = g(z)\frac{\partial}{\partial z} + \frac{1}{2}c^2(y)\frac{\partial^2}{\partial z^2}. \tag{13}$$

where x represents the wealth process, y is a stochastic volatility process, and z is an intensity process.

Proof. The proof follows by the extension of the arguments used in Theorem 4.1 of Duffie and Zariphopoulou (1993) [18] and thus is omitted here. For more details and applications, we refer the reader to Sircar and Zariphopoulou (2007) [8] , Sircar and Zariphopoulou (2010) [19], and Brémand (1981) [20]. □

2.2. Bond Holder's Problem and Indifference Price

In this section we assume that the investor owns a bond of the firm, which is defaultable and pays 1 dollar at maturity. We then construct a similar problem, i.e.,

$$U(t, x, y, z) = sup_{\pi \in \mathcal{A}} E\{\tilde{U}(X_T + c)\mathbf{1}\{\tau_t > T\} + \tilde{U}(X_{\tau_t})\mathbf{1}\{\tau_t \leq T\} \mid X_t = x, Y_t = y, Z_t = z\} \tag{14}$$

where c denotes e^{-rT} .

Proposition 2. *The HJB equation of Bond Holder's value function is*

$$U_t + \frac{1}{\xi}\mathcal{L}\dagger U + \mathcal{L}\ddagger U + \frac{1}{\sqrt{\xi}}\rho_{12}a(y)c(z)U_{yz} + max\{\pi\mu(y)U_x + \frac{1}{2}\pi^2\sigma^2(y)U_{xx}$$

$$+ \frac{1}{\sqrt{\xi}}\pi\rho_1\sigma(y)a(y)U_{xy} + \rho_2 U_{xz}\pi\sigma(y)c(z)\} + \lambda(z)(\tilde{U}(x) - U) = 0, \tag{15}$$

with $U(T, x, y, z) = \tilde{U}(x + c)$.

We can then have the following definition

Definition 1. *The indifference price to an investor is defined at time zero by*

$$V(0, x, y, z) = U(0, x - p_0, y, z), \tag{16}$$

which aims to keep the utility indifference between holding or not holding the bond. The bond holder should lower his initial wealth level. And the yield spread is defined as

$$y_0(T) = -\frac{1}{T}log(p_0(T)) - \gamma, \tag{17}$$

which is non-negative for all $T > 0$ *and* $p_0(T)$ *represents the indifference price at time T.*

3. Asymptotic Approximation

For simplicity, we start our analysis under exponential utility, for the reason that the analytic form of solution is easy to obtain for an exponential affine structure. But the idea behind is the same, and the analysis of the constant-relative risk aversion (CRRA) utility is shown in the subsequent section. By necessary conditions for extreme values, we obtain the maximizer π^* for the optimization problem (11),

$$\pi^* = -\frac{\frac{1}{\sqrt{\xi}}\rho_1\sigma(y)a(y)V_{xy} + \mu(y)V_x + \rho_2\sigma(y)c(z)V_{xz}}{\sigma^2(y)V_{xx}}. \tag{18}$$

Substituting (18) into (11), we obtain the following nonlinear PDE,

$$V_t + \frac{1}{\xi}\mathcal{L}\dagger V + \mathcal{L}\ddagger V + \frac{1}{\sqrt{\xi}}\rho_{12}a(y)c(z)V_{yz}$$

$$- \frac{[\theta(y)V_x + \frac{1}{\sqrt{\xi}}\rho_1 a(y)V_{xy} + \rho_2 c(z)V_{xz}]^2}{2V_{xx}} + \lambda(z)(-e^{-\gamma x} - V) = 0, \tag{19}$$

where

$$\theta(y) = \frac{\mu(y)}{\sigma(y)}. \tag{20}$$

It is hard to get the explicit solution of the nonlinear PDE. Thus, we use the perturbation method to solve the problem.

Firstly, we expand the V as follows

$$V^\xi = V^{(0)} + \xi^{1/2}V^{(1)} + \xi V^{(2)} + \xi^{3/2}V^{(3)} + \cdots \tag{21}$$

According to the term ξ^{-1} derived from (19) and (21), we can prove that $V^{(0)}$ is independent of y, because $\mathcal{L}\dagger V^{(0)} = 0$, and $\mathcal{L}\dagger$ is an operator based on y. Similarly, by the terms $\xi^{-\frac{1}{2}}$, we can prove

that $V^{(1)}$ is independent of y, which means $V^{(0)}$ and $V^{(1)}$ are functions of t and x. The variable y is involved only in the expansion of the term $V^{(2)}$. By extracting the coefficient of the term ξ^0, we obtain

$$V_t^{(0)} + \mathcal{L}\dagger V^{(2)} + \mathcal{L}\ddagger V^{(0)} + NL(1) + \lambda(z)(-e^{-rx} - V^{(0)}) = 0. \tag{22}$$

By extracting the coefficient of the term $\xi^{\frac{1}{2}}$, we obtain

$$V_t^{(1)} + \mathcal{L}\dagger V^{(3)} + \mathcal{L}\ddagger V^{(1)} + NL(2) - \lambda(z)V^{(1)} = 0. \tag{23}$$

Now we consider the expansion about $NL(i)(i = 1, 2)$. By using the Taylor expansion and the fact that $V^{(0)}$ and $V^{(1)}$ are independent of y, we get

$$
\begin{aligned}
NL(i) &= -\frac{[\theta(y)V_x + \frac{1}{\sqrt{\xi}}\rho_1 a(y)V_{xy} + \rho_2 c(z)V_{xz}]^2}{2V_{xx}} \\
&= -[\theta(y)(V_x^{(0)} + \sqrt{\xi}V_x^{(1)}) + \frac{1}{\sqrt{\xi}}\rho_1 a(y)(V_{xy}^{(0)} + \sqrt{\xi}V_{xy}^{(1)}) + \xi V_{xy}^{(2)} \\
&\quad + \rho_2 c(z)(V_{xz}^{(0)} + \sqrt{\xi}V_{xy}^{(1)})]^2 \frac{1}{2V_{xx}^{(0)}}(1 - \sqrt{\xi}\frac{V_{xx}^{(1)}}{V_{xx}^{(0)}} - \xi\frac{V_{xx}^{(1)}}{V_{xx}^{(0)}}) \\
&= -\frac{1}{2V_{xx}^{(0)}}[\theta(y)V_x^{(0)} + \rho_2 c(z)V_{xz}^{(0)}]^2 - \sqrt{\xi}\{-\frac{V_{xx}^{(1)}}{2(V_{xx}^{(0)})^2}[\theta(y)V_x^{(0)} + \rho_2 c(z)V_{xz}^{(0)}]^2 \\
&\quad + \frac{1}{V_{xx}^{(0)}}[\theta(y)V_x^{(0)} + \rho_2 c(z)V_{xz}^{(0)}][\theta(y)V_x^{(1)} + \rho_1 a(y)V_{xy}^{(2)}]\}
\end{aligned}
\tag{24}
$$

Then we have

$$NL(1) = -\frac{1}{2V_{xx}^{(0)}}[\theta(y)V_x^{(0)} + \rho_2 c(z)V_{xz}^{(0)}]^2, \tag{25}$$

and

$$
\begin{aligned}
NL(2) &= \frac{V_{xx}^{(1)}}{2(V_{xx}^{(0)})^2}[\theta(y)V_x^{(0)} + \rho_2 c(z)V_{xz}^{(0)}]^2 \\
&\quad - \frac{1}{V_{xx}^{(0)}}[\theta(y)V_x^{(0)} + \rho_2 c(z)V_{xz}^{(0)}][\theta(y)V_x^{(1)} + \rho_1 a(y)V_{xy}^{(2)} + \rho_2 c(z)V_{xz}^{(1)}].
\end{aligned}
\tag{26}
$$

3.1. Analysis of the Zero-Strategy Leading Term

According to Fredholm's Alternative solvability condition specified in Equation (22) in Fouque et al. (2011) [17], we obtained

$$V_t^{(0)} + \mathcal{L}\ddagger V^{(0)} - \frac{(\hat{\theta}V^{(0)} + \rho_2 c(z)V_{xz}^{(0)})^2}{2V_{xx}^{(0)}} + \lambda(-e^{-\gamma x} - V^{(0)}) = 0, \tag{27}$$

where

$$V(t, x, y, z) = -e^{-\gamma x} \tag{28}$$

The Equation (27) can be simplified by a distortion scaling

$$V^{(0)}(t, x, z) = -e^{-\gamma x}M(t, z)^{\frac{1}{1-\rho_2^2}}, \tag{29}$$

to become

$$M_t + \tilde{\mathcal{L}}\ddagger M - (1 - \rho_2^2)(\frac{\theta^2}{2} + \lambda)M - \lambda(1 - \rho_2^2)M^\alpha = 0, \tag{30}$$

where

$$\alpha = \frac{\rho_2^2}{\rho_2^2 - 1}, \quad \tilde{\mathcal{L}\ddagger} = \mathcal{L}\ddagger - \rho_2 \hat{\theta} c(z) \frac{\partial}{\partial z}. \tag{31}$$

The only difference between holding or not holding the bond is the initial condition of the leading term. The differential equation follows:

$$U_t^{(0)} + \mathcal{L}\ddagger U^{(0)} - \frac{(\hat{\theta} U^{(0)} + \rho_2 c(z) U_{xz}^{(0)})^2}{2 U_{xx}^{(0)}} + \lambda(-e^{-\gamma x} - U^{(0)}) = 0, \tag{32}$$

where

$$U(t, x, y, z) = -e^{-\gamma(x+c)} \tag{33}$$

The above equation can be simplified by a distortion scaling

$$U^{(0)}(t, x, z) = -e^{-\gamma(x+c)} N(t, z)^{\frac{1}{1-\rho_2^2}}, \tag{34}$$

and becomes

$$N_t + \tilde{\mathcal{L}\ddagger} N - (1 - \rho_2^2)(\frac{\hat{\theta}^2}{2} + \lambda) N - \lambda(1 - \rho_2^2) e^{\lambda c} N^\alpha = 0, \tag{35}$$

where

$$\alpha = \frac{\rho_2^2}{\rho_2^2 - 1}, \quad \tilde{\mathcal{L}\ddagger} = \mathcal{L}\ddagger - \rho_2 \hat{\theta} c(z) \frac{\partial}{\partial z}. \tag{36}$$

3.2. Analysis of the Fast Modification Term

Firstly, we give the following notations

$$\phi_1 = -\frac{\theta(y) V_x^{(0)} + \rho_2 c(z) V_{xz}^{(0)}}{V_{xx}^{(0)}} \frac{\partial}{\partial x}, \tag{37}$$

$$\phi_2 = [\frac{\theta(y) V_x^{(0)} + \rho_2 c(z) V_{xz}^{(0)}}{V_{xx}^{(0)}}]^2 \frac{\partial^2}{\partial x^2}. \tag{38}$$

Substituting (37) and (38) into the non-linear term of (23), we get

$$\mathcal{L}\dagger V^{(3)} + V_t^{(1)} + \mathcal{L}\ddagger V^{(1)} + \frac{1}{2}\phi_2 V^{(1)} + \theta\phi_1 V^{(1)} + \rho_1 a\phi_1 V_y^{(2)} - \lambda(z) V^{(1)} + \rho_2 c\phi_1 V_z^{(0)} = 0. \tag{39}$$

Similarly, using ϕ_1 and ϕ_2, Equation (22) can be written as

$$\mathcal{L}\dagger V^{(2)} + V_t^{(0)} + \mathcal{L}\ddagger V^{(0)} - \lambda(z) V^{(0)} + \phi_2 V^{(0)} + \theta\phi_1 V^{(1)} + \rho_2 c\phi_1 V_z^{(0)} = \lambda(z) e^{-\gamma x}. \tag{40}$$

By using the Fredholm Alternative theorem as before, we obtain

$$V_t^{(1)} + \mathcal{L}\ddagger V^{(1)} + \frac{1}{2}\hat{\phi}_2 V^{(1)} + \theta\hat{\phi}_1 V^{(1)} - \lambda(z) V^{(1)} + \rho_2 c\hat{\phi}_1 V_z^{(0)} = -\rho_1 a\hat{\phi}_1 V_y^{(2)}. \tag{41}$$

$$V_t^{(0)} + \mathcal{L}\ddagger V^{(0)} + \frac{1}{2}\hat{\phi}_2 V^{(0)} + \theta\hat{\phi}_1 V^{(0)} - \lambda(z) V^{(1)} + \rho_2 c\hat{\phi}_1 V_z^{(0)} = \lambda e^{-\gamma x}. \tag{42}$$

By comparing the above two equations, it is easy to guess the solution as

$$V^{(1)} = -(T - t)\rho_1 a\hat{\phi}_1 V_y^{(2)} + c(t, x), \tag{43}$$

where $V^{(2)}$ is a function of $V^{(0)}$ and $c(t, x)$ can be determined by substituting (43) into (41).

4. Analysis of Fast-Scale Correction under the Exponential Utility Assumption

For simplification of the problem, we assume λ to be a constant. Firstly, we consider our problem under the fast mean-reverting stochastic volatility, namely the volatility of the stock process is only related to Y. We then have the following model:

$$\frac{dS_t}{S_t} = \mu(Y_t)dt + \sigma(Y_t)dW_t, \tag{44}$$

$$dY_t = \frac{1}{\xi}b(Y_t)dt + \frac{1}{\sqrt{\xi}}a(Y_t)dW_t^{(1)}. \tag{45}$$

4.1. Fast-Scale Expansion for Single Name Derivatives

The HJB equation is transformed into the following form

$$V_t^\xi + \frac{1}{\xi}\mathcal{L}\mathsf{t}_0 V^\xi + \lambda(-e^{-\gamma x} - V) + \mathcal{F}V = 0, \tag{46}$$

where

$$\mathcal{F}V = sup_{\pi_t \in \mathcal{A}}[\pi_t \mu(y)V_x + \frac{1}{2}(\pi_t)^2 \sigma(y)^2 V_{xx} + \pi_t \frac{1}{\sqrt{\xi}}\rho_1 a(y)\sigma(y)V_{xy}] \tag{47}$$

By solving the optimization problem in (47), we obtain π_t^* as follows

$$\pi_t^* = -\frac{\mu(y)}{\sigma^2(y)}\frac{V_x}{V_{xx}} - \frac{1}{\sqrt{\xi}}\rho_1 \frac{a(y)}{\sigma(y)}\frac{V_{xy}}{V_{xx}}. \tag{48}$$

Substituting (48) into (46), the non-linear equation becomes

$$V_t^\xi + \frac{1}{\xi}\mathcal{L}\mathsf{t}_0 V^\xi - \frac{(\theta(y)V_x^\xi + \frac{\rho_1 a(y)}{\sqrt{\xi}}V_{xy}^\xi)^2}{2V_{xx}^\xi} + \lambda(-e^{-rx} - V) = 0, \tag{49}$$

where

$$\theta(y) = \frac{\mu(y)}{\sigma(y)}. \tag{50}$$

Then we can look for an expansion of the value function:

$$V^\xi = V^{(0)} + \sqrt{\xi}V^{(1)} + \xi V^{(2)} + \xi^{3/2}V^{(3)} + \cdots. \tag{51}$$

By Substituting (51) into (49) and collecting the coefficients of the terms ξ^{-1} and $\xi^{-\frac{1}{2}}$, we can get the conclusion that $V^{(0)}$ and $V^{(1)}$ are independent of Y. From the the coefficients of the constant term and the term ξ^{-1}, we get the following two equations:

$$V_t^{(0)} + \mathcal{L}\mathsf{t}_0 V^{(2)} - \frac{1}{2}\theta^2(y)\frac{(V_x^{(0)})^2}{V_{xx}^{(0)}} - \lambda V^{(0)} = \lambda e^{-rx}, \tag{52}$$

$$V_t^{(1)} + \mathcal{L}\mathsf{t}_0 V^{(3)} - NL(1) - \lambda V^{(1)} = 0, \tag{53}$$

where

$$NL(1) = -\frac{\theta(y)}{V_{xx}^{(0)}}V_x^{(0)}[\lambda(y)V_x^{(1)} + \rho_1 a(y)V_{xy}^{(2)}] + \frac{V_{xx}^{(1)}}{2(V_x^{(0)})^2}\theta^2(y)(V_x^{(0)})^2. \tag{54}$$

From Fredholm's alternative solvability condition, we get

$$V_t^{(0)} - \frac{1}{2}\hat{\theta}^2 \frac{(V_x^{(0)})^2}{V_{xx}^{(0)}} - \lambda V^{(0)} \;=\; \lambda e^{-rx}, \tag{55}$$

$$V_t^{(1)} - < NL(1) > -\lambda V^{(1)} \;=\; 0. \tag{56}$$

where $< \cdot >$ denotes the average of y. From Equation (55), we get the leading term $V^{(0)}$, and from (42), we can get the relationship between $V^{(0)}$ and $V^{(1)}$, and then we can get the approximation of V^{ξ}.

Proposition 3. *The explicit solution of Equation (55) is*

$$V^{(0)}(t,x) = -\frac{\lambda}{\frac{1}{2}\hat{\theta}^2 + \lambda} e^{-\gamma x} + (1 - \frac{\lambda}{\frac{1}{2}\hat{\theta}^2 + \lambda}) e^{-(\frac{1}{2}\hat{\theta}^2 + \lambda)(T-t)} e^{-\gamma x}, \tag{57}$$

where $\hat{\theta}$ is the average value of $\theta(y)$ with the distribution of Π, namely

$$\hat{\theta} = \int \theta(y) \Pi(dy) \tag{58}$$

Proof. We firstly transform the PDE by averaging $\theta(y)$. Because $V^{(0)}$ is independent of y, we get the following PDE,

$$V_t^{(0)} - \frac{1}{2}\hat{\theta}^2 \frac{(V_x^{(0)})^2}{V_{xx}^{(0)}} - \lambda V^{(0)} = \lambda e^{-\gamma x}, \quad V_T^{(0)} = e^{-\gamma x} \tag{59}$$

By making the substitution of $V_T^{(0)} = -e^{-\gamma x} M$, we get the following ODE,

$$M_t - (\lambda + \frac{1}{2}\hat{\theta}^2)M = -\lambda, \quad M_T = 1 \tag{60}$$

Then we can obtain the solution of (55) by solving the above equation. □

We then introduce

$$R^{(0)} \;=\; -\frac{V_x^{(0)}}{V_{xx}^{(0)}} \tag{61}$$

$$D_k \;=\; (R^{(0)})^k \frac{\partial^k}{\partial x^k}, k = 1, 2, \cdots \tag{62}$$

$$\mathcal{L}\dagger_{t,x,y}^e \;=\; \frac{\partial}{\partial t} + \frac{1}{2}\theta^2(y)D_2 + \theta^2(y)D_1 - \lambda \tag{63}$$

$$\mathcal{L}\dagger_{t,x}^e \;=\; \frac{\partial}{\partial t} + \frac{1}{2}\hat{\theta}^2 D_2 + \hat{\theta}^2 D_1 - \lambda \tag{64}$$

Equations (109) and (55) become

$$\mathcal{L}\dagger_0 V^{(2)} + \mathcal{L}\dagger_{t,x,y} V^{(0)} = \lambda e^{-\gamma x}, \tag{65}$$

$$\mathcal{L}\dagger_{t,x} V^{(0)} = \lambda e^{-\gamma x}. \tag{66}$$

Subtracting (65) by (66), we get

$$V^{(2)} = -\eta(y)(\frac{1}{2}D_2 + D_1)V^{(0)}, \tag{67}$$

$$\eta(y) = \mathcal{L}\dagger_0^{-1}(\theta^2(y) - \hat{\theta}). \tag{68}$$

Substituting (67) into (54), we can get the following proposition

Proposition 4. *The value of the fast modification form is the solution of the equation below,*

$$\mathcal{L}\dagger^e_{t,x,y} V^{(1)} = \frac{1}{2}\rho_1 B D_1^2 V^{(0)}(t,x), \quad V^{(1)}(T,x) = 0, \tag{69}$$

where $B = \theta(y)a(y)\eta(y)$.

Proof. As $D_2 = -D_1$, we have

$$V^{(2)} = -\eta(y)(\frac{1}{2}D_2 + D_1)V^{(0)} = -\frac{1}{2}\eta(y)D_1 V^{(0)}. \tag{70}$$

Based on (56) and (54), we have

$$V_t^{(1)} - \langle \frac{V_x^{(0)}}{V_{xx}^{(0)}}\theta(y)[V_x^{(1)} + \rho_1 a(y)V_{xy}^{(2)}] - \frac{(V_x^{(0)})^2}{2(V_{xx}^{(0)})^2}V_{xx}^{(1)}\theta^2(y)\rangle - \lambda V^{(1)}$$

$$= V_t^{(1)} - \langle -\theta(y)D_1 V_x^{(1)} - \rho_1 a(y)\theta(y)D_1 V_y^{(2)} - \frac{1}{2}\theta^2(y)D_2 V^{(1)}\rangle - \lambda V^{(1)} \tag{71}$$

$$= V_t^{(1)} + \hat{\theta}(y)D_1 V^{(1)} - \frac{1}{2}\rho_1 B D_1^2 V^{(0)} + \frac{1}{2}\hat{\theta}^2(y)D_2 V^{(1)} - \lambda V^{(1)},$$

where $B = \langle a(y)\theta(y)\eta'(y)\rangle$. □

Lemma 1. *The operators* $\mathcal{L}\dagger^e_{t,x}$ *and* D_1 *acting on smooth functions of* (t,x) *commute:*

$$\mathcal{L}\dagger^e_{t,x} D_1 = D_1 \mathcal{L}\dagger^e_{t,x}. \tag{72}$$

Proof.

$$D_2 D_1 - D_1 D_2 = (R^{(0)})^2 \frac{\partial^2}{\partial x^2}(R^{(0)}w_x) - R^{(0)}\frac{\partial}{\partial x}((R^{(0)})^2 w_{xx}) = (R^{(0)})^2 R_{xx}^{(0)} w_x \tag{73}$$

$$\begin{aligned}
\mathcal{L}\dagger^e_{t,x} D_1 w &= (\frac{\partial}{\partial t} + \frac{1}{2}\hat{\theta}^2 D_2 + \hat{\theta}D_1 - \lambda)D_1 w \\
&= D_1 \frac{\partial}{\partial t} + \frac{1}{2}\hat{\theta}^2 D_1 D_2 + \hat{\theta}D_1^2 - \lambda D_1)w + (R_t^{(0)} + \frac{1}{2}\hat{\theta}^2(R^{(0)})^2 R_{xx}^{(0)})w_x \\
&= D_1 \mathcal{L}\dagger^e_{t,x} w.
\end{aligned} \tag{74}$$

□

From Lemma 1 we can draw the conclusion that $\mathcal{L}\dagger^e_{t,x}(D_1^k V^{(0)}) = D_1^k \mathcal{L}\dagger^e_{t,x} V^{(0)}$, which leads to the following proposition.

Proposition 5. *The solution of (69) is*

$$V^{(1)} = -(T-t)\frac{1}{2}\rho_1 B D_1^2 V^{(0)}(t,x) + c(t,x), \tag{75}$$

where $B = \theta(y)a(y)\eta'(y)$, and

$$c(t,x) = (\frac{M'}{N'}(T-t) + \frac{M'}{N'^2} - \frac{M'}{N'^2}e^{N'(T-t)})e^{-\gamma x}, \tag{76}$$

$$M' = \frac{1}{2}\rho_1 B\lambda\gamma^2, \tag{77}$$

$$N' = \frac{1}{2}\hat{\theta}^2(R^{(0)})^2 - \hat{\theta}R^{(0)}\gamma - \lambda. \tag{78}$$

Proof. We firstly assume that the solution of (69) is

$$V^{(1)} = -(T-t)\frac{1}{2}\rho_1 BD_1^2 V^{(0)}(t,x) + c(t,x). \tag{79}$$

Substituting (79) into (69), we obtain

$$\frac{1}{2}\rho_1 BD_1^2 V^{(0)} - (T-t)\frac{1}{2}\rho_1 BD_1^2 \mathcal{L}\dagger_{t,x}V^{(0)} + \mathcal{L}\dagger_{t,x}c(t,x) = \frac{1}{2}\rho_1 BD_1^2 V^{(0)}. \tag{80}$$

Then we obtain

$$\mathcal{L}\dagger_{t,x}c(t,x) = (T-t)\frac{1}{2}\rho_1 BD_1^2 \mathcal{L}\dagger_{t,x}V^{(0)}. \tag{81}$$

Because $\mathcal{L}\dagger_{t,x}V^{(0)} = \lambda e^{-\gamma x}$, we obtain the PDE as follows

$$\mathcal{L}\dagger_{t,x}c(t,x) = (T-t)\frac{1}{2}\rho_1 B\gamma^2\lambda e^{-\gamma x}, \quad c(T,x) = 0. \tag{82}$$

Assume $c(t,x) = A(t)e^{-\gamma x}$, then we obtain

$$A_t + N'A = (T-t)M', \quad A(T,x) = 0, \tag{83}$$

where

$$M' = \frac{1}{2}\rho_1 B\lambda\gamma^2, \quad N' = \frac{1}{2}\hat{\theta}^2(R^{(0)})^2 - \hat{\theta}R^{(0)}\gamma - \lambda. \tag{84}$$

The terminal condition here is arisen from the condition $V^{(1)}(T,x) = c(T,x) = 0$. By solving the ODE for A, we get

$$A = \frac{M'}{N'}(T-t) + \frac{M'}{N'^2} - \frac{M'}{N'^2}e^{N'(T-t)}, \tag{85}$$

and

$$c(t,x) = (\frac{M'}{N'}(T-t) + \frac{M'}{N'^2} - \frac{M'}{N'^2}e^{N'(T-t)})e^{-\gamma x}. \tag{86}$$

□

From the expansion (51), and the solution of $V^{(0)}$, $V^{(1)}$ and $V^{(2)}$, we obtain

$$\begin{aligned}
V^{(\xi)} &= V^{(0)} + \sqrt{\xi}V^{(1)} + \xi V^{(2)} + o(\xi^{\frac{3}{2}}) \\
&= (1 - \sqrt{\xi}\frac{1}{2}(T-t)\rho_1 BD_1^2)V^{(0)}(t,x) + \sqrt{\xi}c(t,x) + o(\xi^{\frac{3}{2}}).
\end{aligned} \tag{87}$$

Then we analyse the approximation of the maximizer π^* as given in (48).

Using Taylor expansion, we get

$$
\begin{aligned}
\frac{V_x}{V_{xx}} &= \frac{V_x^{(0)} + \sqrt{\xi}V_x^{(1)}}{V_{xx}^{(0)} + \sqrt{\xi}V_{xx}^{(1)}} \\
&= \frac{1}{V_{xx}^{(0)}}(V_x^{(0)} + \sqrt{\xi}V_x^{(1)})(1 - \sqrt{\xi}\frac{V_{xx}^{(1)}}{V_{xx}^{(0)}}) \\
&= \frac{V_x^{(0)}}{V_{xx}^{(0)}} + \sqrt{\xi}(\frac{V_x^{(1)}}{V_{xx}^{(0)}} - \frac{V_x^{(0)}V_{xx}^{(1)}}{(V_{xx}^{(0)})^2}) \\
&= \frac{V_x^{(0)}}{V_{xx}^{(0)}} + \frac{\sqrt{\xi}}{V_{xx}^{(0)}}(c_x + R^{(0)}c_{xx}) \\
&= \frac{V_x^{(0)}}{V_{xx}^{(0)}} - \sqrt{\xi}\frac{1}{V_x^{(0)}}(D_1 + D_2)c + o(\xi),
\end{aligned}
\tag{88}
$$

and

$$
\begin{aligned}
\frac{V_{xy}}{V_{xx}} &= \frac{V_{xy}^{(0)} + \sqrt{\xi}V_{xy}^{(1)} + \xi V_{xy}^{(2)}}{V_{xx}^{(0)} + \sqrt{\xi}V_{xx}^{(1)} + \xi V_{xx}^{(2)}} \\
&= \xi\frac{V_{xy}^{(2)}}{V_{xx}^{(0)}} = -\xi\frac{1}{V_x^{(0)}}\eta(y)\frac{1}{2}D_2D_1V_y^{(0)}.
\end{aligned}
\tag{89}
$$

Substituting the above into (48) yields

$$
\pi^* = -\frac{\theta(y)}{\delta(y)}\frac{V_x^{(0)}}{V_{xx}^{(0)}} + \frac{\sqrt{\xi}}{V_x^{(0)}}[\frac{\theta(y)}{\delta y}(D_1 + D_2)c + \rho_1\eta(y)\frac{1}{2}D_1D_2V_y^{(0)}]
\tag{90}
$$

Similarly, the solution of the bond holders' problem is given in the following properties,

Proposition 6. *The leading term of the bond holder's problem is*

$$
U^{(0)} = -\frac{\lambda}{\frac{1}{2}\hat{\theta}^2 + \lambda}e^{-\gamma x} + (1 - \frac{\lambda e^{\gamma c}}{\frac{1}{2}\hat{\theta}^2 + \lambda})e^{-(\frac{1}{2}\hat{\theta}^2 + \lambda)(T-t)-\gamma(x+c)}
\tag{91}
$$

where $\hat{\theta}$ is the average of $\theta(y)$ with respect to the distribution Π, namely

$$
\hat{\theta} = \int \theta(y)\Pi(dy).
\tag{92}
$$

The fast-scale modification term of the bond holder's problem is

$$
(1 - \sqrt{\xi}\frac{1}{2}(T - t)\rho_1BD_1^2)U^{(0)}(t, x) + \sqrt{\xi}C(t, x) + o(\xi^{\frac{3}{2}}),
\tag{93}
$$

where

$$
C(t, x) = (\frac{M'}{N'}(T - t) + \frac{M'}{N'^2} - \frac{M'}{N'^2}e^{D(T-t)})e^{-\gamma x}.
\tag{94}
$$

So the approximation of the bond holder's value function is

$$
\begin{aligned}
U^{(\xi)} &= U^{(0)} + \sqrt{\xi}U^{(1)} + \xi U^{(2)} + o(\xi^{\frac{3}{2}}) \\
&= (1 - \sqrt{\xi}\frac{1}{2}(T - t)\rho_1BD_1^2)U^{(0)}(t, x) + \sqrt{\xi}C(t, x) + o(\xi^{\frac{3}{2}}).
\end{aligned}
\tag{95}
$$

5. Numerical Study of Exponential Utility

5.1. Analysis of the Value Function

The utility we use from Bond seller is exponential and is given by

$$V(x) = -e^{-\gamma x}, \tag{96}$$

where $\gamma > 0$ represents the risk aversion parameter. We can prove that the utility function is concave and increasing since

$$V'(x) = \gamma e^{-\gamma x} > 0, V''(x) = -\gamma^2 e^{-\gamma x} < 0. \tag{97}$$

The concave property of the utility function implies that the bond seller is risk aversion. The risk aversion rate is calculated by the Arrow-Pratt index,

$$AP[U] := -\frac{U''(x)}{U'(x)} = \gamma, \tag{98}$$

where the larger the γ is, the higher risk averse the agent is. The risk-tolerance function at terminal time T is

$$R(T, x) = -\frac{U'}{U''} = \frac{1}{\gamma}. \tag{99}$$

5.2. The Effect of Volatility Correction

The study above is all based on the general form. In order to demonstrate the result graphically, we give the special case as follows:

$$\frac{dS_t}{S_t} = Y_t dt + \sqrt{Y_t} dW_t, \tag{100}$$

$$dY_t = \frac{1}{\xi}(m_1 - Y_t)dt + \sqrt{\frac{2}{\xi}} v dW_t^{(1)}. \tag{101}$$

If Y_t is an ergodic process, it has the distribution of $N(m_1, v^2)$. Assuming that $m_1 = 0.01$, $v^2 = 0.25$, $\xi = \frac{1}{200}$, and based on the definition of $\hat{\theta}$, we obtain

$$\hat{\theta} = \frac{1}{\sqrt{2\pi} v} \int_{-\infty}^{\infty} \sqrt{y} e^{-\frac{(y-m_1)^2}{2v^2}} dy \tag{102}$$

According to (57) and (79), we get the solution of $V^{(0)}$, and also the fast modification term of $V^{(1)}$, we then calculate the utility term as $V^{(0)} + \sqrt{\xi} V^{(1)}$. The approximations to the value functions are plotted in Figures 1 and 2.

Also, since the bond pays \$1 on maturity date T if the firm has survived till then, the bond seller's value function will be higher than the bond holder's value function. The comparison of the value function is shown in Figures 3 and 4. The Stochastic Volatility Model(SVM) in Figure 4 represents the Stochastic Volatility(SV) modification form.

The approximate indirect utilities $V^{(0)}$ or $V^{(0)} + \sqrt{\xi} V^{(1)}$ can also be represented by their certainty equivalents $U^{-1}(V^{(0)})$ and $U^{-1}(V^{(0)} + \sqrt{\xi} V^{(1)})$, which are shown in Figures 5 and 6.

In Figures 1 and 2, the original value function is denoted by blue line, while the dashed blue line is the value function with stochastic volatility correction. We can see clearly that the correction line is a little lower than the original line. In Figures 3 and 4, we make a comparison of the value function for holding and not holding the bond. Figure 3 shows the relationship before SV correction while Figure 4 shows the relationship after SV correction. Figures 5 and 6 show the certainty equivalent before or

after the correction; the solid line shows the certainty equivalent before the correction and the dashed line shows the after situation.
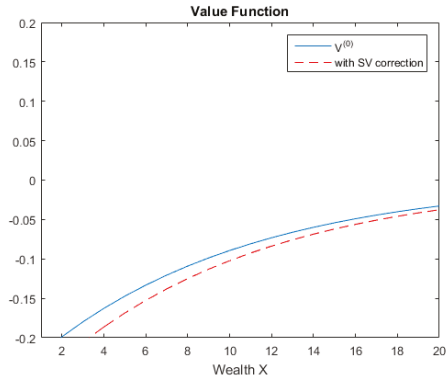
Figure 1. Value Function of Bond Seller.
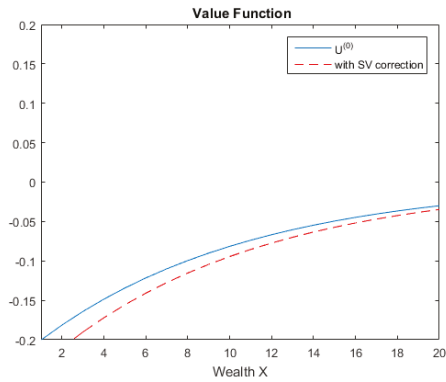
Figure 2. Value Function of Bond Holder.
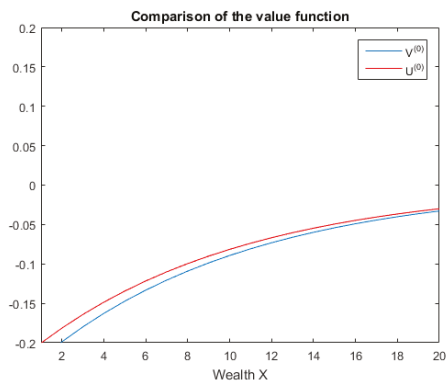
Figure 3. Leading Term Value Function.
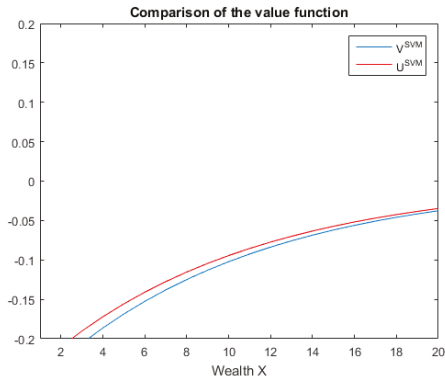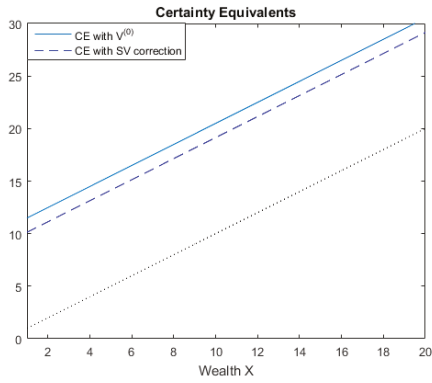
Figure 4. SV Modification Value Function.
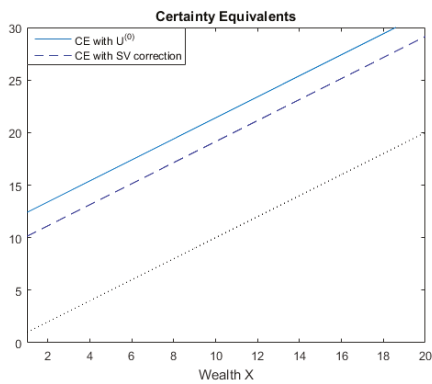

Figure 5. Certainty Equivalents of Bond Seller.

Figure 6. Certainty Equivalents of Bond Holder.

Therefore, we can draw the conclusion that by adding a stochastic volatility process into model (101), the investor becomes more risk adverse. The stochastic Volatility is lower than both the utility function and the certainty equivalent. Also, as bond holder will get a fixed pay at the maturity

date if default does not happen, the value function of the bond holder will be a little higher than that of the bond seller. That is why we give the definition of indifference price p_0. By cutting down the initial wealth of bond holder, the expected utility of bond holder should be the same as that of the bond seller. In the following subsection, we will analyse the indifference and yield spread numerically.

5.3. Analysis of Yield Spread

According to the Definition 1, it is easy to calculate p_0 and the yield spread. Without the modification term, the indifference price $p_0^{(0)}$ is given by

$$p_0^{(0)} = e^{-rT} + \frac{1}{\gamma} \ln \frac{u - (1-u)e^{-(\frac{1}{2}\hat{\theta}^2 + \lambda)T}}{ue^{\gamma c} - (1 - ue^{\gamma c})e^{-(\frac{1}{2}\hat{\theta}^2 + \lambda)T}},$$

(103)

where

$$u = \frac{\lambda}{\frac{1}{2}\hat{\theta}^2 + \lambda}.$$

(104)

If γ takes the value of 0.05, 0.1, 0.25, 0.5 and 0.75 respectively, we obtain the profile of yield spread $y_0(T) = -\frac{1}{T} log(p_0^{(0)}(T)) - r$ as shown in Figure 7.
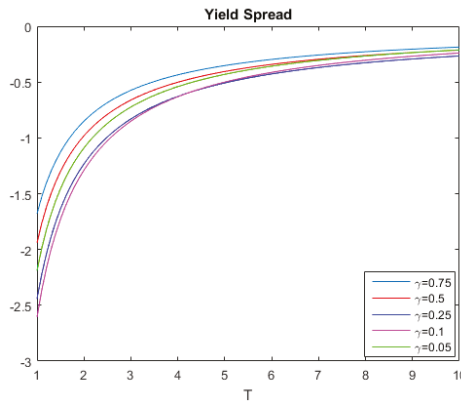
Figure 7. Yield Spread.

It is noted that the yield spread is not flat even though the intensity is a constant, and this is due to the effect of the intensity rate λ upon T. When T goes to infinity, yield spread will convergent to a long time level and become flat. As we can read from Figure 7, the yield spread for the investor is upward sloping and is approximated to a long time level due to the different maturity time.

6. Numerical Study of CRRA Utility

The utility we use from Bond seller is exponential and given by

$$V(x) = c_0 \frac{x^{1-\gamma}}{1-\gamma},$$

(105)

where $\gamma > 0$ represents the risk aversion parameter. We can prove that the utility function is concave and increasing since

$$V'(x) = x^{-\gamma} > 0, V''(x) = -\gamma x^{-\gamma - 1} < 0.$$

(106)

The concave property of the utility function implies that the bond seller is risk aversion. The risk aversion rate is calculated by the Arrow-Pratt index,

$$AP[U] := -\frac{U''(x)}{U'(x)} = \gamma/x, \tag{107}$$

where the larger the γ is, the higher risk averse the agent is. And the risk-tolerance function at terminal time T is

$$R(T,x) = -\frac{U'}{U''} = \frac{1}{\gamma}x. \tag{108}$$

Under the assumption of the CRRA utility form, the above leading term and the first-order correction term reduce to

$$V_t^{(0)} - \frac{1}{2}\hat{\theta}^2\frac{(V_x^{(0)})^2}{V_{xx}^{(0)}} - \lambda V^{(0)} = -\lambda c_0\frac{x^{1-\gamma}}{1-\gamma}, \tag{109}$$

$$V_t^{(1)} - <NL(1)> -\lambda V^{(1)} = 0, \tag{110}$$

with the terminal condition $V^{(0)}(T,x) = c_0\frac{x^{1-\gamma}}{1-\gamma}$ and $V^{(1)}(T,x) = 0$. The leading term can be solved analytically by assuming

$$V^{(0)} = x^{1-\gamma}M(t). \tag{111}$$

Substituting (111) into (109), we obtain the following ordinary differential equation(ODE)

$$M_t + (\frac{1-\gamma}{2}\hat{\theta}^2 - \lambda)M = -\frac{\lambda c_0}{1-\gamma}, \tag{112}$$

with the terminal condition $M(T) = \frac{1}{1-\gamma}$. Solving the above ODE together with the initial condition, we obtain

$$V^{(0)} = \left(-\frac{\lambda c_0}{\lambda - \frac{1-\gamma}{2\gamma}\hat{\theta}^2} + \left(\frac{1}{1-\gamma} + \frac{\lambda c_0}{\lambda - \frac{1-\gamma}{2\gamma}\hat{\theta}^2}\right)e^{-(\lambda - \frac{1-\gamma}{2\gamma}\theta^2)(T-t)}\right)x^{1-\gamma}, \tag{113}$$

Similarly, we solve U_0 with respect to the terminal condition of $U^{(0)}(T,x) = (x+c)^{1-\gamma}/(1-\gamma)$, and obtain

$$U^{(0)} = \left(-\frac{\lambda c_0}{\lambda - \frac{1-\gamma}{2\gamma}\hat{\theta}^2} + \left(\frac{1}{1-\gamma}(\frac{x+c}{x})^{1-\gamma} + \frac{\lambda c_0}{\lambda - \frac{1-\gamma}{2\gamma}\hat{\theta}^2}\right)e^{-(\lambda - \frac{1-\gamma}{2\gamma}\theta^2)(T-t)}\right)x^{1-\gamma}, \tag{114}$$

In order to obtain the first order correction term $V^{(1)}$, we assume $V^{(1)} = -(T-t)\frac{1}{2}\rho_1 BD_1^2 V^{(0)}(t,x+c(t,x))$, with $c(t,x)$ satisfying

$$\mathcal{L}\dagger_{t,x}c(t,x) = (T-t)\frac{1}{2}\rho_1 B\lambda c_0 D_1^2\frac{x^{1-\gamma}}{1-\gamma}, \quad c(T,x) = 0. \tag{115}$$

We can not derive the analytic form of the solution of the above equation because the right hand side of the equation is not necessary an affine structure. In this case, we solve it numerically by finite element discretization shown in Appendix A.

The results are shown in Figures 8 and 9, from which we know that value function is concave and increasing considering the risk attitude. However, the fast scale stochastic volatility correction drag the value function a little bit downward, and the main reason is that incorporation of uncertain volatility gives the investors more risk exposure.
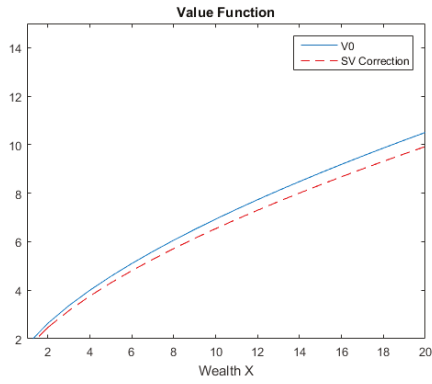
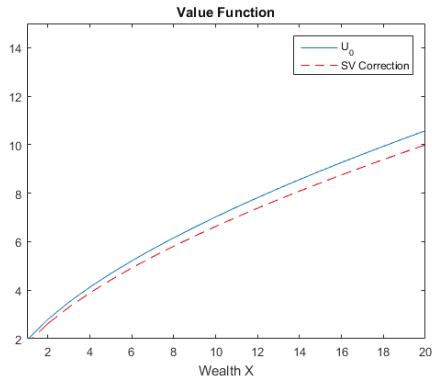**Figure 8.** Value Function of Bond Seller.

Figure 9. Value Function of Bond Holder.

Similar results are also shown in the certainty equivalents, that is incorporation of stochastic volatility lower down the value function, as shown in Figures 10 and 11.
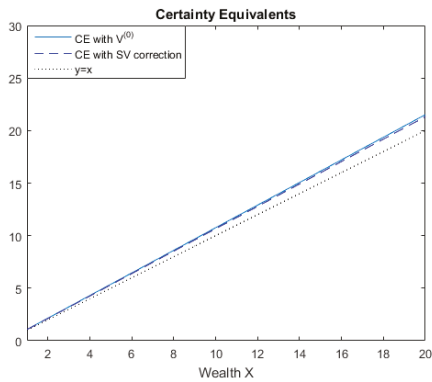

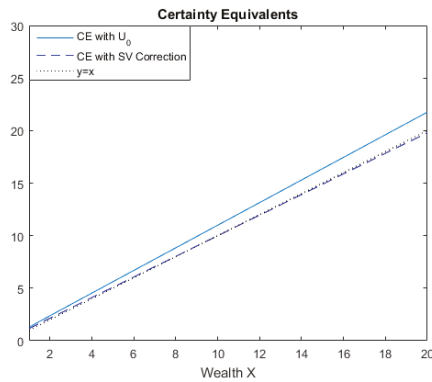**Figure 10.** Certainty Equivalents of Bond Seller.

Figure 11. Certainty Equivalents of Bond Holder.

Table 1 gives the results of the percentage change of the mean value when the model parameter is given a 1% change. Clearly, the value function is sensitive to the correlation rate, and the long term mean variance $m1$ is more sensitive compared to the volatility of volatility.

Table 1. Sensitivity study of parameters.

Name	Value	Sensitivity
$m1$	0.11	3.65%
v	0.25	1.80%
λ	0.5	1.32%
ρ	0.5	-19.61%

7. Conclusions and Future Work

In this paper, we study the single-name bond under the stochastic intensity and the stochastic volatility. In order to solve the non-linear PDE, we use the method of asymptotic approximation. We establish the expression of leading term $V^{(0)}$, and fast-scale modification term $V^{(1)}$. By comparing the leading term and the utility with fast scale modification, we can draw the conclusion that by considering the effects of the fast-scale volatility, investors become more and more risk aversive, which lowers down their utility and increases the certainty equivalents. Also, according to the analysis above, we prove that the yield spread of the investor goes up with the maturity time and converges to a long time level. The advantage of the asymptotic method is that it reduces the high dimensional problem into a lower dimensional problem, which is relatively easy to solve. However, the limitation of this approach is that it only works for a specific utility model, and for other utilities, the analytic solutions may not be obtained so that numerical method is needed. In our future research, the effect of multiscale volatility and stochastic interest rate will be taken into consideration.

Author Contributions: All the authors contributed to the development of credit derivatives and this manuscript. Shican Liu and Yanli Zhou conceived this research, analyzed the data and wrote the manuscript. Benchawan Wiwatanapataphee offered advises on the simulations. Yonghong Wu guided the research direction and proposed the framework of analysis. Xiangyu Ge contributed to the manuscript revising process and the overall quality of the manuscript. All of the authors revised and approved the final manuscript.

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

Appendix A

In order to obtain the first correction of the CRRA utility, we solve the following parabolic equation numerically,

$$V_t^{(1)} + \frac{1}{2\gamma^2}\hat{\theta}^2 x^2 V_{xx}^{(1)} + \frac{1}{\gamma}\hat{\theta}^2 x V^{(1)} - \lambda V^{(1)} = f(T - t, x), \tag{A1}$$

with $f(T - t, x) = (T - t)\frac{1}{2}\rho_1 B\lambda c_0 D_1^2 \frac{x^{1-\gamma}}{1-\gamma}$, $c(T, x) = 0$, and terminal condition $V^{(1)}(T, x) = 0$. Let $\tau = T - t$, we can obtain the weak form of (A1),

$$(V_\tau^{(1)}, V) + \frac{1}{2}\hat{\theta}^2(x^2 V_x^{(1)}, U_x) - \hat{\theta}^2\frac{1}{\gamma}(1 + \frac{1}{\gamma^2})(x V_x^{(1)}, U) + \lambda(V^{(1)}, U) = (f(\tau, x), U). \tag{A2}$$

The basis function $V^{(1)}$, the test function U and the function f can be approximated by the following form

$$V^{(1)}(x, \tau) = \sum_{i=1}^{N} u_i(\tau)\Phi_i(x)$$

$$U(x) = \sum_{j=1}^{N} v_j\Phi_j(x), \tag{A3}$$

$$f(x, \tau) = \sum_{i=1}^{N} f_i(\tau)$$

We then obtain the ODE systems,

$$M\dot{U} + RU = F,$$
$$U(0) = 0, \tau \in [0, T] \tag{A4}$$

where

$$M = (\Phi_i, \Phi_j);$$
$$R = \frac{1}{2}\hat{\theta}^2(x_i^2 \frac{\partial \Phi_i}{\partial x}, \frac{\partial \Phi_j}{\partial x}) - \hat{\theta}^2\frac{1}{\gamma}(1 + \frac{1}{\gamma^2})(x_i \frac{\partial \Phi_i}{\partial x}, \Phi_j) + \lambda(\Phi_i, \Phi_j); \tag{A5}$$
$$F = (f_i, \Phi_j)$$

We apply the backward Euler method to solve the above dynamic ODE system and yields

$$(\frac{M}{\Delta t} + R)U_{n+1} = M\frac{U_n}{\Delta t} + F. \tag{A6}$$

References

1. Tan, Y.; Floros, C. Risk, capital and efficiency in Chinese banking. *J. Int. Financ. Mark. Inst. Money* **2013**, *26*, 378–393.
2. Tan, Y. The impacts of risk and competition on bank profitability in China. *J. Int. Financ. Mark. Inst. Money* **2016**, *40*, 85–110.
3. Madan, D.B.; Unal, H. Pricing the risks of default. *Rev. Deriv. Res.* **1998**, *2*, 121–160.
4. Jarrow, R.A.; Turnbull, S.M. Pricing derivatives on financial securities subject to credit risk. *J. Financ.* **1995**, *50*, 53–85.
5. Lando, D. On Cox processes and credit risky securities. *Rev. Deriv. Res.* **1998**, *2*, 99–120.

6.   Hao, C.; Zhang, B.; Carling, K.; Alam, M.M. *Review of the Literature on Credit Risk Modeling: Development of the Recent 10 Years*; Business Perspectives: Sumy, Ukraine, 2009.

7.   Black, F.; Scholes, M. The pricing of options and corporate liabilities. *J. Political Econ.* **1973**, *81*, 637–654.

8.   Sircar, R.; Zariphopoulou, T. Utility valuation of credit derivatives: Single and two-name cases. *Adv. Math. Financ.* **2007**, 279–301, doi:10.1007/978-0-8176-4545-8_15.

9.   Papageorgiou, E.; Sircar, R. Multiscale intensity models for single name credit derivatives. *Appl. Math. Financ.* **2008**, *15*, 73–105.

10. Merton, R.C. Optimum consumption and portfolio rules in a continuous-time model. *J. Econ. Theory* **1971**, *3*, 373–413.

11. Heston, S.L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.* **1993**, *6*, 327–343.

12. Longstaff, F.A.; Schwartz, E.S. A simple approach to valuing risky fixed and floating rate debt. *J. Financ.* **1995**, *50*, 789–819.

13. Fouque, J.P.; Papanicolaou, G.; Sircar, R.; Solna, K. Multiscale stochastic volatility asymptotics. *Multiscale Model. Simul.* **2003**, *2*, 22–42.

14. Fouque, J.P.; Sircar, R.; Zariphopoulou, T. Portfolio optimization and stochastic volatility asymptotics. *Math. Financ.* **2017**, *27*, 704–745, doi:10.1111/mafi.12109.

15. Hodges, S.D.; Neuberger, A. Optimal Replication of Contingent Claims under Transaction Costs. *Rev. Futures Mark.* **1989**, *8*, 222–239.

16. Davis, M.; Yoshikawa, D. An equilibrium approach to indifference pricing. *Adv. Financ. Eng.* **2012**, 29–56, doi:10.2139/ssrn.1568856.

17. Fouque, J.P.; Papanicolaou, G.; Sircar, R.; Solna, K. *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*; Cambridge University Press: Cambridge, UK, 2011.

18. Duffie, D.; Zariphopoulou, T. Optimal investment with undiversifiable income risk. *Math. Financ.* **1993**, *3*, 135–148.

19. Sircar, R.; Zariphopoulou, T. Utility valuation of multi-name credit derivatives and application to CDOs. *Quant. Financ.* **2010**, *10*, 195–208.

20. Brémand, P. *Point Processes and Queues*; Springer: New York, NY, USA, 1981.

Article

Dependence Structures and Systemic Risk of Government Securities Markets in Central and Eastern Europe: A CoVaR-Copula Approach

Lu Yang [1], Jason Z. Ma [1] and Shigeyuki Hamori [2,*]

[1]   School of Finance, Zhongnan University of Economics and Law, 182# Nanhu Avenue, East Lake High-Tech Development Zone, Wuhan 430-073, China; kudeyang@gmail.com (L.Y.); zhejma@hotmail.com (J.Z.M.)
[2]   Faculty of Economics, Kobe University, 2-1, Rokkodai, Nada-Ku, Kobe 657-8501, Japan
*   Correspondence: hamori@econ.kobe-u.ac.jp; Tel.: +81-78-803-6832

Abstract: In this study, we proposed a new empirical method by combining generalized autoregressive score functions and a copula model with high-frequency data to model the conditional time-varying joint distribution of the government bond yields between Poland/Czech Republic/Hungary, and Germany. Capturing the conditional time-varying joint distribution of these bond yields allowed us to precisely measure the dependence of the government securities markets. In particular, we found a high dependence of these government securities markets in the long term, but a low dependence in the short term. In addition, we report that the Czech Republic showed the highest dependence with Germany, while Hungary showed the lowest. Moreover, we found that the systemic risk dynamics were consistent with the idea that the global financial crisis not only had spillover effects on countries with weak economic fundamentals (e.g., Hungary, which had the highest systemic risk), but also had contagion effects for both CEEC-3 countries and Germany. Finally, we confirm that three major market events, namely the EU accession, the global financial crisis, and the European debt crisis, caused structural changes to the dynamic correlation.

Keywords: dynamic conditional correlation; generalized autoregressive score functions; time-varying copula function; CoVaR

---

1. Introduction

Measuring the dependence structures of government securities markets is garnering considerable attention from academia as well as from financial institutions, given the continuing expansion of the European Union (EU). In 2004, 10 countries from Central and Eastern Europe and the Mediterranean region joined the EU, which served as a historic step towards unifying Europe after several decades of division that had resulted from the Cold War. In this study, rather than investigating correlations, we propose a new approach to investigate the dependence structures among these countries' financial markets including the investigation of general correlations as well as tail correlation.

Financial markets become integrated when economies strongly depend on one other. This process not only reduces transaction costs, but also improves the efficiency of information sharing. However, although financial integration increases overall market efficiency, it reduces the diversification benefits available to prospective investors. Thus, investigating the dynamic process of financial integration allows us not only to measure the interdependence of economies, but also to provide useful information for investors.

Here, we propose a new method for evaluating the degree to which the integration processes and risk spillovers in Central and Eastern Europe have evolved over time. To simplify our analysis, we chose Poland, the Czech Republic, and Hungary (termed as the CEEC (Central and Eastern European countries)-3

hereafter) to represent Central and Eastern Europe given that these countries have the largest economies and financial markets in the region as well as the best data availability. To represent the EU, we chose Germany because of its economic background and geographic factors. Therefore, we investigated the differences in the dependence structures of the government securities markets in the CEEC-3 and Germany.

Two types of approaches tend to be used to study dependence structures. The first type includes observation-based methods such as those based on the generalized autoregressive conditional heteroskedasticity (GARCH) framework [1,2]. The dynamic conditional correlation (DCC-GARCH)-based approach [3–5] and copula-GARCH-based approach [6–8] are representative examples. The second type is parameter-based methods. The classical analysis of this type focuses on time-varying parameters, which allows us to better characterize the dynamic correlations in government securities markets by using easy estimations. For example, Pozzi and Wolswijk [9] employed a linear state space approach to estimate the latent factor decomposition of the excess returns or risk premiums suggested by a standard international capital asset pricing model for government bonds. They found that the government bond markets in the Eurozone under investigation were almost fully integrated by the end of 2006, showing that an important part of the achieved convergence was reversed during 2007–2009. Bekiros [10] also provided evidence that time-varying parameter models more accurately forecast Eurozone economies than other models.

In this study, we employed a parameter-driven model, namely the generalized autoregressive score (GAS) model, to investigate the dynamic integrated process of European government securities markets. For example, Creal et al. [11] employed the GAS model to analyze the dynamic correlation between the euro and yen, and between the euro and pound. Meanwhile, Oh and Patton [12] and Creal and Tsey [13] provided evidence that the GAS model could be employed with high dimensional copula to investigate the interdependence among different assets. With regard to the topics of the present study, Boubakri and Guillaumin [14] provided evidence that financial integration was not perfect, but was increasing based on the dynamic correlation of the foreign exchange rate. Furthermore, they also showed that financial contagion occurred during the global financial crisis.

Instead of focusing on the foreign exchange rate, in this study, we investigated the integration of these countries based on interest rates (e.g., bond yields). Moreover, in contrast to the studies of Yang and Hamori [5,7,8] who focused on investigating observations, we computed time-varying parameters. The technique adopted herein was based on the score function of the predictive model density at time t by incorporating the non-linear property. In addition, in contrast to observation-driven models, the GAS model has the advantage of exploiting the complete density structure rather than only means and higher moments. Furthermore, its applications can be extended to asymmetric, long memory, and other more complicated dynamics without increasing model complexity. Therefore, by employing the GAS framework, we restructured the time-varying copula model to investigate the dynamic integrations of the government securities markets in Eastern Europe.

To understand the risk spillover effect between the CEEC-3 and Germany, we employed copulas to compute the conditional value-at-risk (CoVaR) by providing quantitative evidence on the systemic risk spillovers in government securities markets. Furthermore, we evaluated how the deteriorating financial position of a sovereign market could impair the performance of other government securities markets during a crisis. In particular, we used the CoVaR measures originally proposed by Adrian and Brunnermeier [15] and generalized by Girardi and Ergün [16], which allowed us to capture the possible risk spillovers between markets by providing information on the value-at-risk (VaR) of a market, conditional on the fact that another market is in financial distress.

By adopting a two-step procedure, we easily obtained the value of the CoVaR. In the first step, we computed the cumulative probability of the CoVaR from a copula function by assuming the cumulative probability of the VaR of the market in financial distress, and the confidence level of the CoVaR. In the second step, we obtained the value of the CoVaR by inverting the marginal distribution function for this cumulative probability. Moreover, by employing GAS specifications, we obtained more sensitive information on the risk spillover effect in the government securities markets of the CEEC-3 and Germany.

Our contributions to the body of knowledge are threefold. First, we provide more specific details on the dependence across different maturities when compared with previous studies. Second, we implemented a new approach (i.e., the GAS-based dynamic Gaussian copula) to investigate the dynamic correlations among these markets, which can provide us with more sensitive correlations to the structural changes. This approach allowed us to analyze how the degree of dependence changed according to major market events, namely the EU accession (2004), the global financial crisis (2008), and the European debt crisis (2012). Third, we compared and contrasted the risk spillover effect in the government securities markets of the CEEC-3 and Germany by employing both the Gaussian copula model and the Gaussian copula GAS model. Finally, we employed the Symmetrized Joe-Clayton copula (SJC copula; [17]) to investigate the tail dependence of these markets and compared them with the results from the GAS-based model to verify the robustness of the results.

The remainder of this article is organized as follows. Section 2 discusses the copulas and verifies the time-varying dependence structure. Section 3 describes the data and statistical issues. Section 4 provides the empirical results and Section 5 concludes.

2. Method

In this section, we first describe the margins of the return distributions based on our empirical model. Second, we introduce the specifications of the dynamic copula model. Then, we selected one particular elliptical copula (Gaussian copula) model to investigate the dependence of the government securities markets in Eastern Europe. Furthermore, we estimated the systemic risk of these countries based on both the Gaussian copula and the Gaussian copula GAS models. Finally, to justify the empirical findings, we employed the SJC copula to examine the dynamic tail dependence of the examined government securities markets.

2.1. Marginal Distribution Specifications

The marginal distribution for each return series is characterized by a Glosten-Jagannathan-Runkle GARCH (GJR-AR(k)-GARCH(1,1)-Skew-t; [18]) model that considers the effects of asymmetric information [18–20]. Assume $R_{i,t}$ and $h_{i,t}$ to be bond i's return and conditional variance for period t, respectively. Thus, the GJR-AR(k)-GARCH(1,1)-Skew-t model for the bond return is

$$R_{it} = \mu_i + \alpha_{i,1}R_{i,t-1} + \alpha_{i,2}R_{i,t-2} + \cdots + \alpha_{i,k}R_{i,t-k} + \varepsilon_{i,t} \tag{1}$$

$$h_{i,t} = \omega_i + \beta_i h_{i,t-1} + \delta_i \varepsilon_{i,t-1}^2 + \cdots + \gamma_i s_{i,t-1} \varepsilon_{i,t-1}^2 \tag{2}$$

where $s_{i,t-1} = 1$ when $\varepsilon_{i,t-1}$ is negative and $s_{i,t-1} = 0$ otherwise. We assumed that the error term $\varepsilon_{i,t}$ followed the skew-t distribution with the density function $f(v_t, \lambda_t)$, such that

$$f(y_t; v_t, \lambda_t) = \begin{cases} bc\left(1 + \frac{1}{v_t}\left(\frac{by_t+a}{1-\lambda_t}\right)^2\right)^{-\frac{v_t+1}{2}} & for \quad y_t \leq -\frac{a}{b} \\ bc\left(1 + \frac{1}{v_t}\left(\frac{by_t+a}{1+\lambda_t}\right)^2\right)^{-\frac{v_t+1}{2}} & for \quad y_t > -\frac{a}{b} \end{cases} \tag{3}$$

where $c = \frac{\Gamma(\frac{v_t+1}{2})}{\Gamma(\frac{v_t}{2}\sqrt{\pi(v_t-2)})}$, $b = \sqrt{1 + 3\lambda_t^2 - a^2}$, and $a = 4\lambda_t c(\frac{v_t-2}{v_t-1})$. This density is defined for $2 < v_t < \infty$ and $-1 < \lambda_t < 1$ [21]. For the GJR (1, 1) model, the constraints applied to Equation (3) are $\delta + \beta + 2\gamma < 2$, $\delta > -\gamma$, and $\beta \in (0, 1)$, and we chose k based on the Akaike information criterion (AIC) [22].

2.2. A Copula with GAS Dynamics

After determining the suitable marginal distribution, we proceeded to the copula function. A dynamic copula model is typically used to model the dependence of government securities markets in Eastern Europe in a dynamic process. However, an important contribution of our research

was to calculate the time-varying correlations between the CEEC-3 and Germany. Two types of specifications allow the parameters to vary over time. First, studies of copula-based analysis such as Hafner and Manner [23] and Manner and Segers [24] have proposed a stochastic copula model that allows the parameters to evolve as a latent time series. Second, ARCH (autoregressive conditional heteroskedasticity)-type models such as dynamic conditional correlation (DCC) [3] and their related models for copulas [11,17] permit the time-varying parameters to vary according to the functions of the lagged observables. One advantage of the second approach is that it avoids the need to "integrate out" the innovation terms driving the latent time series processes [25,26]. In addition, as pointed out by McAleer [27], DCC may suffer from the problem of the derivation of asymptotic properties of the Quasi-Maximum Likelihood Estimators. Therefore, based on the parameter-driven methodology, the Generalized Autoregressive Score (GAS) model provided us with another view of the conditional correction model as well as the CoVaR approach.

As our empirical model, we employed the GAS model of Creal et al. [11]. This function describes the time-varying copula parameter (δ_t) as a combination of the lagged copula parameter and a forcing variable related to the standardized score of the copula log-likelihood. Following Creal et al. [11], a copula with GAS dynamics can be expressed as

$$U_t | \mathcal{F}_{t-1} \sim C_t(\delta_t(\gamma)) \qquad (4)$$

where γ is the copula function's parameter; and $U_t = [U_{1t}, U_{2t}]'$ is the vector of the marginal conditional probability integral transform. To ensure that the correlation of the normal copula falls between the values of -1 and 1, Creal et al. [11] suggested transforming the copula parameter by using an increasing invertible function (e.g., logarithmic, logistic) to the parameter:

$$\kappa_t = h(\delta_t) \Longleftrightarrow \delta_t = h^{-1}(\kappa_t) \qquad (5)$$

For a copula with a transformed time-varying parameter δ_t, a GAS (1,1) model can be described as

$$\kappa_{t+1} = \omega + b\kappa_t + aI_t^{-\frac{1}{2}} h(\delta_t) s_t \qquad (6)$$

$$s_t \equiv \frac{\partial log(u_y; \delta_t)}{\partial \delta_t} \qquad (7)$$

$$I_t \equiv E_{t-1}\left[s_t s_t' \right] = I(\delta_t) \qquad (8)$$

Although the functions for the time-varying parameters are arbitrary, they can nest a variety of popular approaches from conditional variance models to trade duration and count models. Nonetheless, in contrast to the approach taken by Patton [17], GAS models are more sensitive to correlation shocks (for a comparison of the two models, see [11]).

Since we examined the dynamic process of the dependence of the government securities markets in the CEEC-3 and Germany, we employed the time-varying Gaussian copula. The conditional Gaussian copula function is defined as the density of the joint standard uniform variables (u_t, v_t) with a time-varying correlation ρ_t. Moreover, we assumed that $x_t = \phi^{-1}(u_t)$ and $y_t = \phi^{-1}(v_t)$, where $\phi^{-1}(\cdot)$ represents the inverse of the cumulative density function of the standard normal distribution. Then, the density of the time-varying Gaussian copula is expressed as

$$C_t^{Gau}(u_t, v_t | \rho_t) = \frac{1}{\sqrt{1-\rho_t^2}} exp\left(\frac{x_t^2 + y_t^2}{2} - \frac{x_t^2 - 2\rho_t x_t y_t + y_t^2}{2(1-\rho_t^2)} \right) \qquad (9)$$

Thus, by combining Equation (6) with Equation (9), the Gaussian correlation parameter ρ_t is modeled by the transformed parameter $\rho_t = (1 - exp(-\kappa_t))/(1 + exp(-\kappa_t))$, and the additional scaling factor $\delta_t = 2/(1-\rho_t^2)$ in Equation (6) is the consequence of modeling the transformed

correlation parameter κ_t rather than ρ_t directly. Hence, we compared and contrasted the GAS Gaussian copula estimation across maturities.

2.3. CoVaR

In this section, we quantified the VaR (Value at Risk) and CoVaR (Conditional Value at Risk) for the government securities markets in the CEEC-3 and Germany. Given the strong linkages of these markets [7], we considered the impact of financial distress in the German market (as measured by its VaR) on the VaR of the CEEC-3 market and vice versa. Following the studies of Adrian and Brunnermeier [15] and Girardi and Ergün [16], the CoVaR for asset *i* is the VaR for asset *i* conditional on the fact that asset *j* exhibits an extreme movement.

Let r_t^c be the returns for the CEEC-3 government securities market and r_t^d be the returns for the German government securities market. The downside CoVaR for stock returns for an extreme downward oil movement and a confidence level $1 - \beta$ can be formally expressed as the β-quantile of the conditional distribution of r_t^c as

$$Pr\left(r_t^c \leq CoVaR_{\beta,t}^c(q,p)\Big|r_t^d \leq VaR_{\alpha,t}^d\right) = \beta \tag{10}$$

where $VaR_{\alpha,t}^d$ is the α-quantile of the German government securities market return distribution and $Pr\left(r_t^d \leq VaR_{\alpha,t}^d\right) = \alpha$ measures the maximum loss that the German government securities market returns may experience for a confidence level $1 - \alpha$ and a specific time horizon.

Moreover, we measured the systemic impact of the CEEC-3 government securities market on the German government securities market by considering the CoVaR for the latter instead of the former as in Equation (10). The CoVaR in those equations can be represented in terms of copulas, since the conditional probabilities can be rewritten, respectively, as

$$C\left(F_{r_t^c}\left(CoVaR_{\beta,t}^c\right), F_{r_t^d}\left(VaR_{\alpha,t}^d\right)\right) = \alpha\beta \tag{11}$$

where $F_{r_t^c}$ and $F_{r_t^d}$ are the marginal distributions of the CEEC-3 government securities market and German government securities market returns, respectively. We followed Reboredo and Ugolini [25] in computing the CoVaR by following a two-step procedure. Following the studies of Adrian and Brunnermeier [15] and Girardi and Ergün [16], the systemic risk contribution of market *j* as the delta CoVaR ($\Delta CoVaR$) can be defined as the difference between the VaR of the overall German government securities market conditional on the distressed state of the CEEC-3 government securities market $\left(R_t^c \leq VaR_{\alpha,t}^c\right)$. The VaR of each of the individual CEEC-3 government securities markets can then be treated as a whole conditional on the benchmark state of the market, considering it to be the median of the return distribution of the market, or, alternatively, the VaR for $\alpha = 0.5$. The systemic risk contribution of the market for each CEEC-3 country is the government securities market thus defined as

$$\Delta CoVaR_t^{c/d} = \frac{CoVaR_{\beta,t}^{c/d} - CoVaR_{\beta,t}^{c/d,\alpha=0.5}}{CoVaR_{\beta,t}^{d,\alpha=0.5}} \tag{12}$$

The primary shortcoming of such a specification is that it estimates the contemporaneous correlation with the market to gauge the size of the potential tail spillover effects. In other words, it is useful as it captures the marginal contribution of markets to the overall systemic risk. In this study, we investigated the risk spillover effects between the CEEC-3 countries and Germany by employing both the Gaussian copula model and Gaussian copula GAS model.

2.4. Estimation Method

In the final step, we employed the multi-stage maximum likelihood (MSML) estimation method to calculate the dynamic relationships between the government securities markets in the CEEC-3 and

Germany. First, we estimated the marginal distributions separately. In the second step, we estimated the copula model conditioned on the estimated marginal distribution parameters. Therefore, the final dynamic copula with the GAS process based on the GARCH model can be specified as

$$\mathcal{L}(\theta) = \sum_{i=1}^{T} log(f_t(X_t;\theta)) = \sum_{i=1}^{T} \log(f_{1t}(X_{1,t};\theta_1)) + \sum_{i=1}^{T} \log(f_{2t}(X_{2,t};\theta_2))$$
$$+ \sum_{i=1}^{T} \log(c_t(F_{1,t}(X_{1,t};\theta), F_{2,t}(X_{2,t};\theta_c))) \tag{13}$$

where $\theta = (\phi', \gamma')'$ is the estimated vector of all the parameters including those of the marginal distributions ϕ and of the copula γ.

3. Data

To investigate the dependence of the CEEC-3 and Germany across maturities, we employed 3-month, 1-year, 3-year, 5-year, and 10-year government bond yields based on a daily frequency. In particular, we focused on 3-year, 5-year, and 10-year government bond yields and omitted 3-month and 1-year government bond yields due to the availability of data and empirical results. For instance, the short-term interest rate for 3-month and 1-year yields cannot model the stable dynamic correlation between Hungary and Germany since the estimation procedure does not converge. Thus, the data on 3-month and 1-year yields did not fit the model well as there were too many poorly fitting observations. Moreover, the marginal distribution for Poland was not well specified since the GARCH process was hardly justified.

The sample period ran from 1 January 2002 to 31 December 2016. The total dataset was comprised of 3914 valid observations. In all cases, bond returns were calculated as the first differences of the logs of yields. Table 1 reports the descriptive statistics of the return series. Particularly, we witnessed the increasing of interest rate for the CEEC-3 countries across the different term structures during our sample periods. In addition, the negative returns of the bond yields also indicated the bad credit environment in the CEEC-3 countries where investors require higher nominal interests. The reason may be due to the saving-investment imbalance with other developed countries such as Germany, whose mean return for ten-year bond yield was still positive. Compared to Germany, the CEEC-3 countries have to deal with their debt problem. For example, the government of Hungary faces a great fiscal deficit and struggles to solve its debt problem. The results of the Jarque–Bera (JB) test showed that the null hypothesis of the normal distribution was rejected in all cases.

Table 1. Summary statistics across different maturities.

	Poland	Hungary	Czech Republic	Germany
3-year				
Mean	−0.000366	−0.000562	−0.000653	−0.000225
Std. Dev.	0.016497	0.017400	0.140969	0.167249
Skewness	0.678808	1.424991	−0.416608	0.682362
Kurtosis	11.01275	22.50707	87.16472	65.51154
JB	10,771.20 ***	63,382.03 ***	1,145,605 ***	628,135.2 ***
Observations	3914	3914	3914	3914
5-year				
Mean	−0.000303	−0.000380	−0.000874	−0.000848
Std. Dev.	0.015542	0.017361	0.169736	0.165198
Skewness	0.447378	0.761848	1.162340	2.072385
Kurtosis	13.77973	15.21576	188.6507	218.0410
JB	19,081.27 ***	24,714.69 ***	5,574,346 ***	7,432,406 ***
Observations	3914	3914	3914	3914
10-year				
Mean	−0.000225	−0.000204	−0.000636	0.000268
Std. Dev.	0.013174	0.015519	0.022779	0.124821
Skewness	0.496346	0.154218	0.336106	−0.555559
Kurtosis	13.80813	10.68544	28.27690	206.1039
JB	19,211.42 ***	9648.186 ***	103,392 ***	6,627,889 ***
Observations	3914	3914	3914	3914

Notes: *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

4. Empirical Results

4.1. Marginal Distribution Estimations

In the first step, we employed univariate GJR-AR(k)-GARCH(1,1)-Skew-t models to model the marginal distributions. Based on the SBIC (Schwarz Bayesian information criterion) [28], we selected $k = 2$ for the 3-year maturity and $k = 1$ for the 5-year and 10-year maturities. Tables 2–4 report our estimation results. We found that all the coefficients of the conditional variance term (β) with values close to one were statistically significant at the 1% level. The coefficients of the asymmetric effect (γ) were also statistically significant at the 1% level for the Czech Republic and Germany for the 3-year maturity, and Poland and Germany for the 10-year maturity. Furthermore, the degrees of freedom parameters (v) were statistically significant at the 1% level with values above two, suggesting that the tails of the error terms were heavier when compared with the normal distribution. Although the skew terms (λ) were not statistically significant with positive values in most cases except Germany, we still used the skew-student-t distribution since all the countries must correlate with Germany.

Table 2. Estimation results of the marginal distribution for 3-year yields.

	Poland	Hungary	Czech Republic	Germany
Mean Equation				
$\mu_1 \times 10^{-4}$	−5.211 (2.115) ***	−4.897 (2.511) **	−0.788 (0.745)	1.388 (1.546)
α	−0.024 (0.015)	−0.018 (0.015)	−0.051 (0.015) ***	−0.029 (0.023)
Variance Equation				
$\omega \times 10^{-5}$	3.414 (1.125) ***	2.251 (2.332)	3.112 (1.052) ***	1.718 (0.344)
δ	0.108 (0.053) ***	0.149 (0.607) ***	0.249 (0.075) ***	0.206 (0.051) ***
β	0.805 (0.039) ***	0.716 (0.101) ***	0.753 (0.038) ***	0.772 (0.031) ***
γ	0.023 (0.039)	−0.129 (0.365)	0.213 (0.065) ***	0.250 (0.012) ***
v	3.138 (0.209) ***	2.492 (0.416) ***	2.646 (0.106) ***	3.384 (0.581) ***
λ	0.023 (0.018)	0.014 (0.016)	0.048 (0.015) ***	0.049 (0.025) *
Diagnostic				
$Q(20)$	23.21 [0.588]	36.54 [0.251]	81.22 [0.245]	21.18 [0.227]
$Q^2(20)$	13.23 [0.786]	21.55 [0.127]	44.87 [0.621]	17.97 [0.419]
Log-Likelihood	11,202.57	10,244.36	8596.28	8496.57

Notes: The numbers in parentheses are standard errors. The numbers in square brackets are p-values. $Q(20)$ ($Q^2(20)$) is the Ljung–Box Q statistic for the null hypothesis that there is no autocorrelation up to order 20 for the standardized residuals (standardized squared residuals). *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Table 3. Estimation results of the marginal distribution for 5-year yields.

	Poland	Hungary	Czech Republic	Germany
Mean Equation				
$\mu_1 \times 10^{-4}$	−2.618 (2.110)	5.124 (2.221) **	−5.428 (4.775)	3.781 (1.546) ***
α	−0.008 (0.019)	0.031 (0.015) **	−0.086 (0.017) ***	0.332 (0.015) ***
Variance Equation				
$\omega \times 10^{-5}$	2.414 (3.125)	4.141 (1.128)	1.787 (1.188)	1.221 (1.188)
δ	0.091 (0.037) ***	0.154 (0.013) ***	0.073 (0.001) ***	0.012 (0.002) ***
β	0.911 (0.028) ***	0.842 (0.039) ***	0.891 (0.028) ***	0.944 (0.015) ***
γ	0.008 (0.028)	0.139 (0.005)	0.145 (0.016) ***	0.099 (0.005) ***
v	3.806 (0.281) ***	2.445 (0.055) ***	2.836 (0.588) ***	7.367 (0.568) ***
λ	0.013 (0.053)	0.015 (0.016)	0.027 (0.021)	0.071 (0.101)
Diagnostic				
$Q(20)$	15.21 [0.448]	41.27 [0.651]	82.12 [0.245]	22.54 [0.347]
$Q^2(20)$	3.286 [1.000]	14.22 [0.234]	44.11 [0.621]	14.27 [0.721]
Log-Likelihood	11,235.812	10,113.699	10,244.87	9853.126

Notes: The numbers in parentheses are standard errors. The numbers in square brackets are p-values. $Q(20)$ ($Q^2(20)$) is the Ljung–Box Q statistic for the null hypothesis that there is no autocorrelation up to order 20 for the standardized residuals (standardized squared residuals). *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Table 4. Estimation results of the marginal distribution for 10-year yields.

	Poland	Hungary	Czech Republic	Germany
Mean Equation				
$\mu_1 \times 10^{-4}$	2.568 (0.221) ***	6.351 (1.121) ***	−4.298 (1.125) ***	2.121 (0.285) ***
α	−0.042 (0.011)	0.046 (0.031)	−0.009 (0.018)	0.089 (0.119)
Variance Equation				
$\omega \times 10^{-5}$	3.122 (1.155) ***	2.886 (1.085) ***	1.987 (0.788) ***	4.221 (1.688) ***
δ	0.116 (0.032) ***	0.111 (0.003) ***	0.166 (0.078) ***	0.017 (0.007) ***
β	0.891 (0.025) ***	0.832 (0.054) ***	0.835 (0.053) ***	0.961 (0.006) ***
γ	−0.023 (0.021)	0.024 (0.469)	0.046 (0.041)	0.041 (0.009) ***
υ	4.012 (0.324) ***	2.319 (0.607) ***	2.996 (0.256) ***	11.621 (2.152) ***
λ	0.009 (0.012)	−0.055 (0.019)	0.015 (0.017)	0.026 (0.022)
Diagnostic				
$Q(20)$	15.26 [0.541]	69.17 [0.265]	55.32 [0.185]	12.96 [0.899]
$Q^2(20)$	1.565 [1.000]	24.75 [0.631]	42.25 [0.331]	9.54 [0.841]
Log-Likelihood	11,116.610	10,004.310	10,522.495	9826.073

Notes: The numbers in parentheses are standard errors. The numbers in square brackets are *p*-values. $Q(20)$ ($Q^2(20)$) is the Ljung–Box Q statistic for the null hypothesis that there is no autocorrelation up to order 20 for standardized residuals (standardized squared residuals). *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Table 2 shows the $Q(s)$ and $Q^2(s)$ statistics to justify the empirical results of the GJR-AR(k)-GARCH(1,1)-Skew-t models. The $Q(s)$ statistic at lag s is a test statistic following an asymptotical distribution with degrees of freedom equal to the number of autocorrelations less the number of parameters. Its null hypothesis assumes that there is no autocorrelation up to lag s for the standardized residuals. The $Q^2(s)$ statistic at lag s proposes a null hypothesis of no autocorrelation up to order s for the standardized squared residuals. As shown in Tables 2–4, the null hypothesis of no autocorrelation up to order 20 for the standardized residuals and standardized squared residuals was accepted for all currencies, supporting our model specifications.

4.2. Dynamic Copula Estimations

In the second step, we transformed the standardized residuals obtained from the GARCH model into uniform variates based on the cumulative distribution function. By applying this step, we obtained the vector of filtered returns to estimate the copula functions in the CEEC-3 government securities markets. Therefore, we estimated both the dynamic Gaussian copula and the dynamic Gaussian copula based on the GAS framework by using the filtered return in the first step. Table 5 reports the estimation results. According to Creal et al. [11] and Creal and Tsay [13], the GAS specification can provide a more persistently time-varying correlation process. Since the log-likelihood was the largest for the 10-year yields when compared with the other two, the long-term yields also provided the most persistently time-varying correlation process. In addition, the terms (a, b) for the GAS framework estimations were significant in most cases, which indicated that the GAS framework models the Gaussian copula well.

To illustrate the integration process between the CEEC-3 and Germany, Figures 1–3 plot their estimated dynamic correlations from the Gaussian copula GAS model for the 3-year, 5-year, and 10-year yields. These figures illustrate the high (low) dependence of the government securities markets in the long term (short term). In addition, the Czech Republic showed the highest dependence with Germany, while Hungary showed the lowest. In particular, the structures of dynamic correlations for Hungary were different from that of Poland and the Czech Republic, which may due to the fact that Hungary has been experiencing a fiscal crisis since 2012.

Meanwhile, to see how EU accession, the global financial crisis, and the European debt crisis affected dependence, we employed the multiple breakpoint test to examine the influence of dependence based on global information citations (Table 6). In general, we found that these three events affected

dependence significantly. As shown in Figures 1–3, the correlation significantly increased before the examined CEEC-3 countries became EU members, in the global financial crisis period, and in the European debt crisis period. Combining the results presented in Table 6 confirmed that financial contagion occurred during these two crises. Meanwhile, the significant increase in correlation before EU accession may have been caused by the expectations of market participants and requirements of being EU members. After the global financial crisis, there was a significant decrease in dependence, perhaps because of capital regulations and market segmentation [14].

Table 5. Estimation results of the Gaussian copula and Gaussian copula GAS (1,1) models.

	3-Year	5-Year	10-Year
Gaussian Copula Model			
Poland–Germany			
ω	0.001 (0.007)	0.004 (0.011)	0.179 (0.059) ***
a	0.026 (0.004) ***	0.079 (0.006) ***	0.481 (0.014) ***
b	1.981 (0.018) ***	1.923 (0.012) ***	0.556 (0.041) ***
Log-Likelihood	30.528	76.522	90.791
Hungary–Germany			
ω	−0.021 (0.048)	−0.043 (0.063) ***	−0.057 (0.022) **
a	0.075 (0.004) ***	0.022 (0.007) ***	0.490 (0.004) ***
b	0. 821 (0.257) ***	0.975 (0.016) ***	−0.477 (0.086) ***
Log-Likelihood	11.344	18.252	22.671
Czech–Germany			
ω	0.001 (0.118)	0.006 (0.361)	0.352 (0.102) ***
a	0.025 (0.001) ***	0.071 (0.003) ***	0.395 (0.012) ***
b	1.994 (0.147) ***	1.984 (0.004) ***	0.749 (0.221) ***
Log-Likelihood	61.300	173.693	254.458
Gaussian Copula GAS (1,1) Model			
Poland–Germany			
ω	0.096 (0.085)	0.187 (0.091) **	0.328 (0.094) ***
a	0.013 (0.003) ***	0.032 (0.007) ***	0.021 (0.004) ***
b	0.992 (0.004) ***	0.983 (0.008) ***	0.990 (0.004) ***
Log-Likelihood	30.551	84.951	110.213
Hungary–Germany			
ω	−0.029 (0.048)	−0.062 (0.061)	−0.052 (0.067)
a	0.018 (0.007) **	0.019 (0.006) ***	0.014 (0.006) **
b	0.963 (0.030) ***	0.978 (0.014) ***	0.987 (0.012) ***
Log-Likelihood	11.136	19.819	18.023
Czech–Germany			
ω	0.796 (0.198) ***	2.004 (0.358) ***	0.872 (0.187) ***
A	0.005 (0.001) ***	0.0163 (0.028) ***	0.001 (0.000) ***
B	0.988 (0.023) ***	0.998 (0.004) ***	0.998 (0.001) ***
Log-Likelihood	79.386	225.322	346.200

Notes: *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.
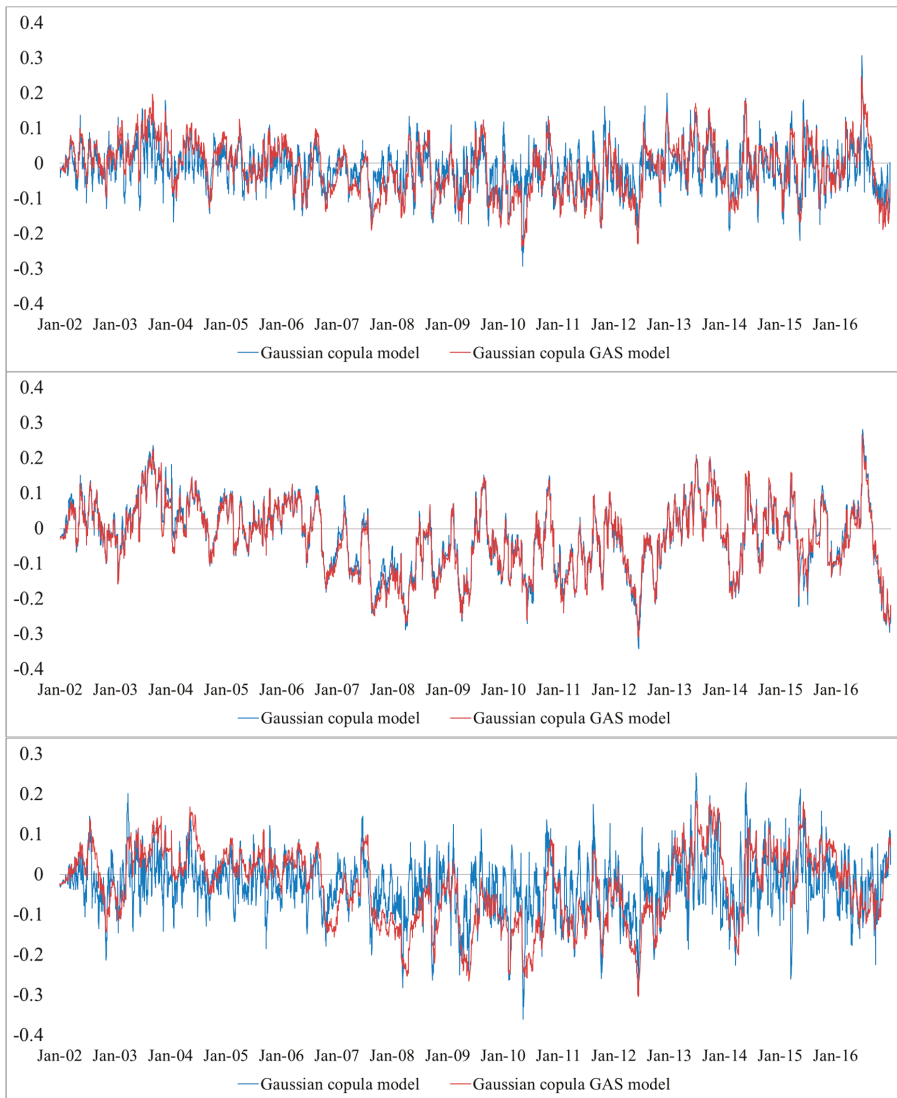
Figure 1. Dynamic correlations between Hungary and Germany. Notes: This figure plots the estimated dynamic correlations between Hungary and Germany from the Gaussian copula and Gaussian copula GAS models for 3-year yields (**top**), 5-year yields (**middle**), and 10-year yields (**bottom**).
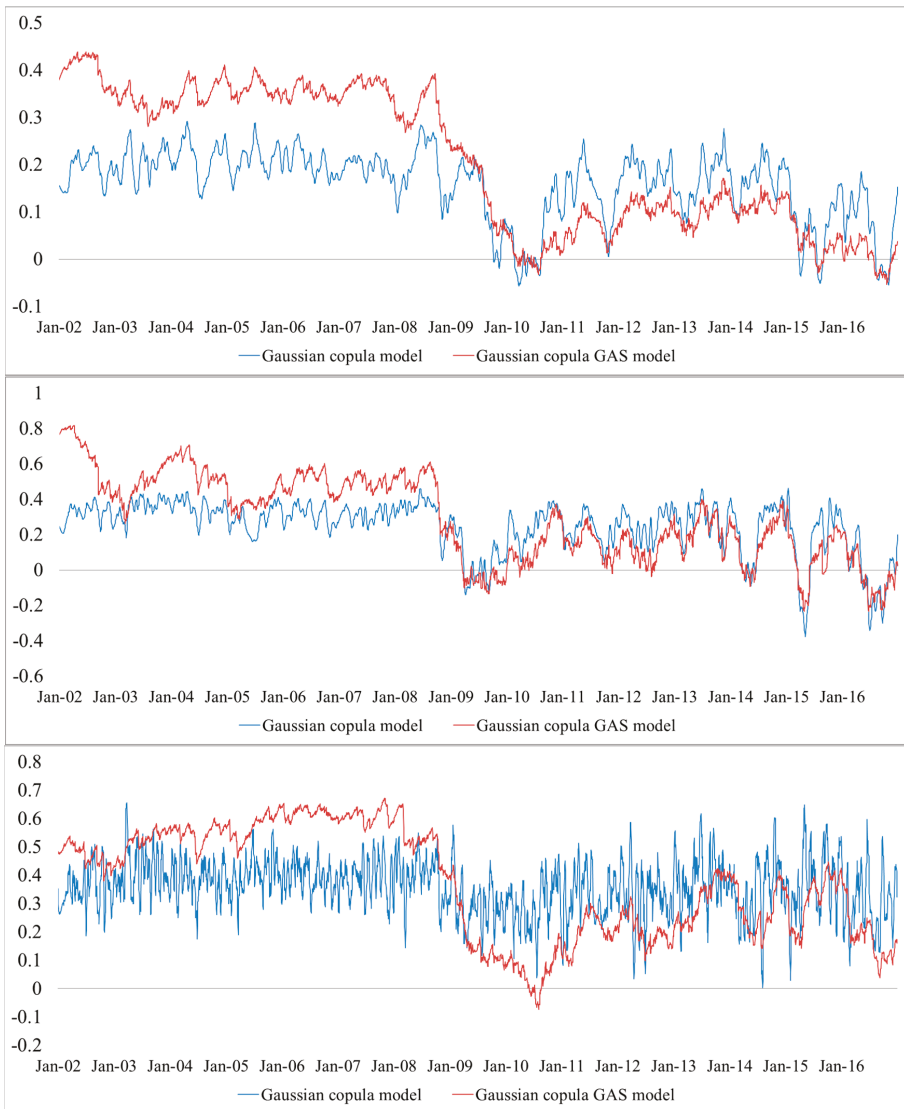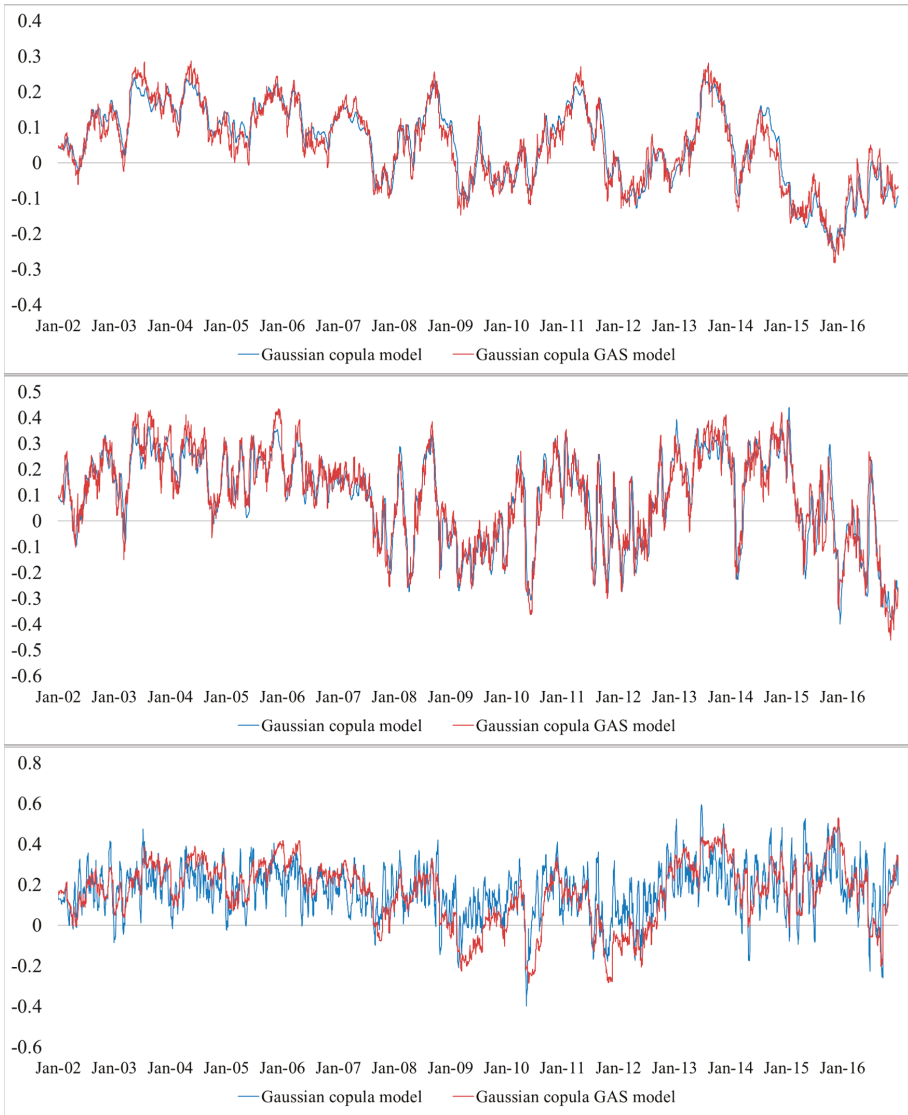
Figure 2. Dynamic correlations between Czech Republic and Germany. Notes: This figure plots the estimated dynamic correlations between Czech Republic and Germany from the Gaussian copula and Gaussian copula GAS models for 3-year yields (**top**), 5-year yields (**middle**), and 10-year yields (**bottom**).

Figure 3. Dynamic correlations between Poland and Germany. Notes: This figure plots the estimated dynamic correlations between Poland and Germany from the Gaussian copula and Gaussian copula GAS models for 3-year yields (**top**), 5-year yields (**middle**), and 10-year yields (**bottom**).

Table 6. Breakpoint test based on global information citations.

	3-Year	5-Year	10-Year
Czech–Germany			
Breakpoint 1	9/28/2004	12/24/2004	12/12/2003
Breakpoint 2	8/23/2007	10/17/2008	3/05/2007
Breakpoint 3	8/04/2009	9/29/2010	2/12/2009
Breakpoint 4	12/28/2011	10/22/2012	1/25/2011
Breakpoint 5			1/11/2013
Poland–Germany			
Breakpoint 1	5/01/2006	4/01/2004	4/01/2004
Breakpoint 2	11/05/2008	5/15/2006	5/15/2006
Breakpoint 3	10/18/2010	10/22/2008	10/22/2008
Breakpoint 4	1/21/2013	10/04/2010	10/04/2010
Breakpoint 5		11/07/2012	11/07/2012
Hungary–Germany			
Breakpoint 1	3/20/2007	9/25/2006	8/07/2007
Breakpoint 2	12/04/2009	6/05/2009	9/27/2010
Breakpoint 3	12/10/2012	12/07/2012	1/04/2013
Breakpoint 4			
Breakpoint 5			

Notes: The date is given by Month/Day/Year. We chose the numbers of the breakpoint date according to the SIC.

4.3. Risk Spillovers

Figures 4–6 plot the estimations of ΔCoVaR. Specifically, the blue line reflects the spillover effect from Germany to the CEEC-3 and the red line reflects the spillover effect from the CEEC-3 to Germany (The CoVaR estimations are available from the authors upon request). As shown in these figures, the GAS-based Gaussian copula model was more sensitive than the Gaussian copula model as expected. Moreover, the empirical evidence indicated that the German systemic risk was low and relatively stable, while the CEEC-3 systemic risk was high and variant. Specifically, Poland showed the lowest systemic risk, whereas Hungary showed the highest. Since the impact of the global financial crisis was reflected in the abrupt increase in the ΔCoVaR value, we observed that the European debt crisis increased the ΔCoVaR value for both the German systemic risk and the CEEC-3 systemic risk. Finally, the ΔCoVaR of long-term government securities fluctuated more widely than that for short-term government securities in these countries. These results suggest that the systemic risk is higher for both the CEEC-3 countries and for longer-term bonds.
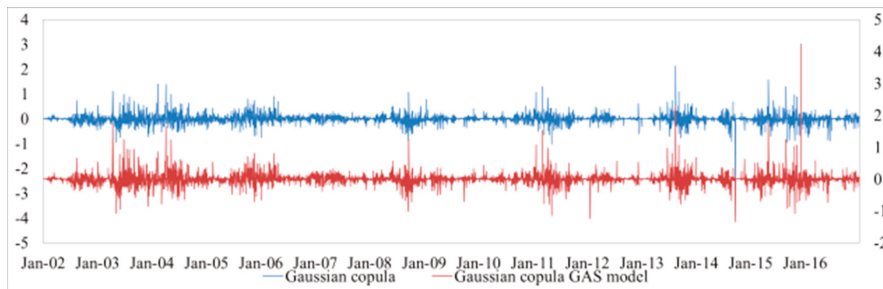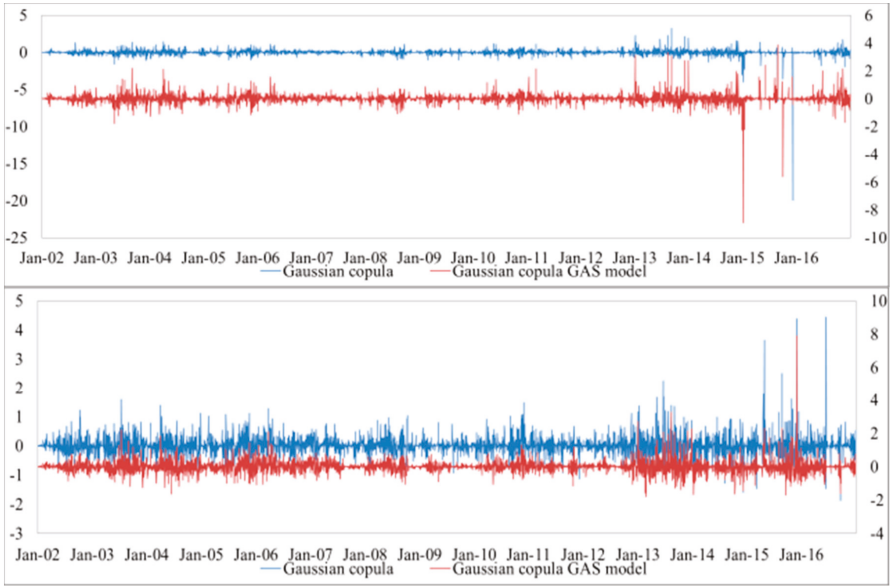
Panel A: The risk spillover from Germany to Poland

Figure 4. *Cont.*

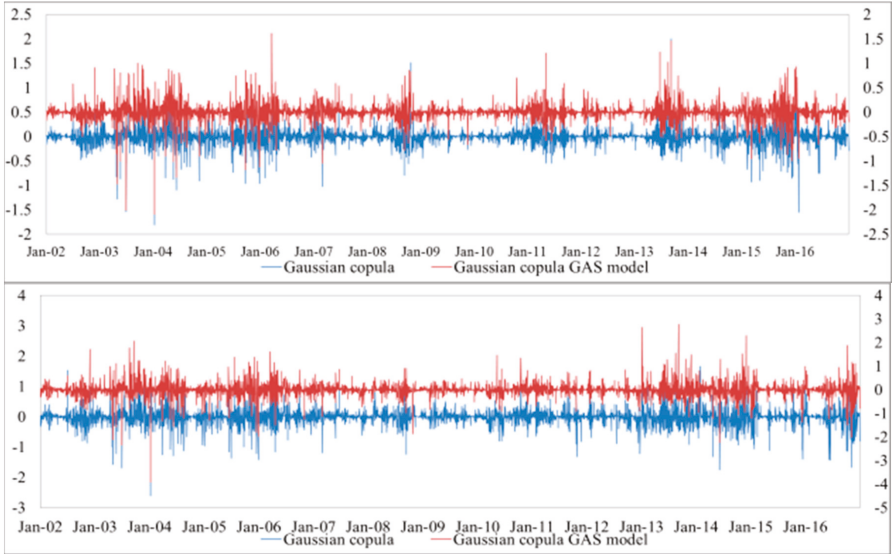Panel B: The risk spillover from Poland to Germany.
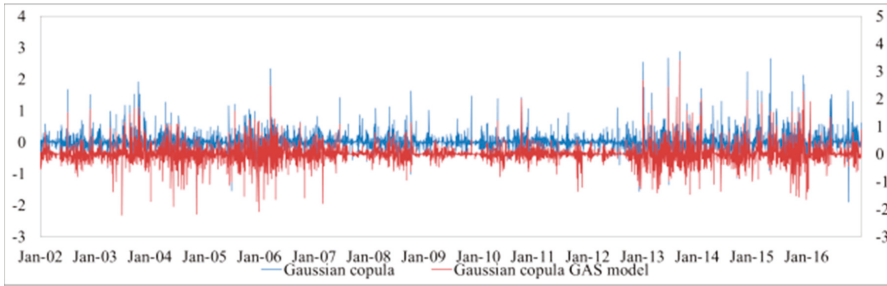

Figure 4. *Cont.*

Figure 4. ΔCoVaR between Poland and Germany. Notes: This figure plots the estimated ΔCoVaR between Poland and Germany from the Gaussian copula and Gaussian copula GAS models for 3-year yields (**top**), 5-year yields (**middle**), and 10-year yields (**bottom**) in the panel (**A**,**B**).
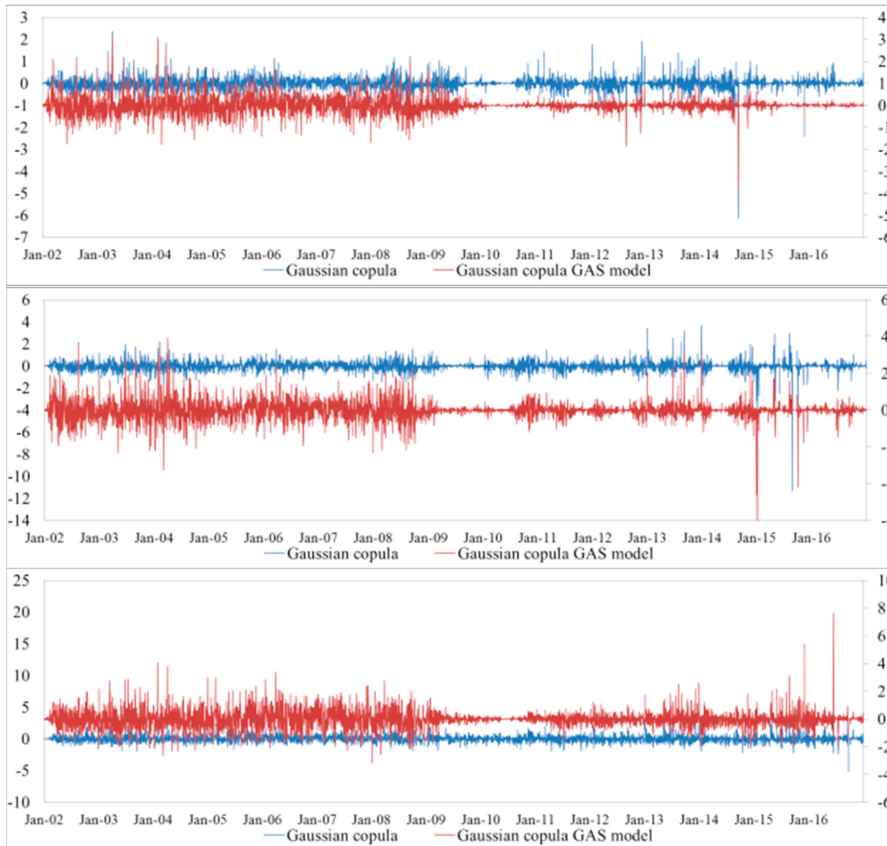
Panel A: The risk spillover from Germany to the Czech Republic.

Figure 5. *Cont.*

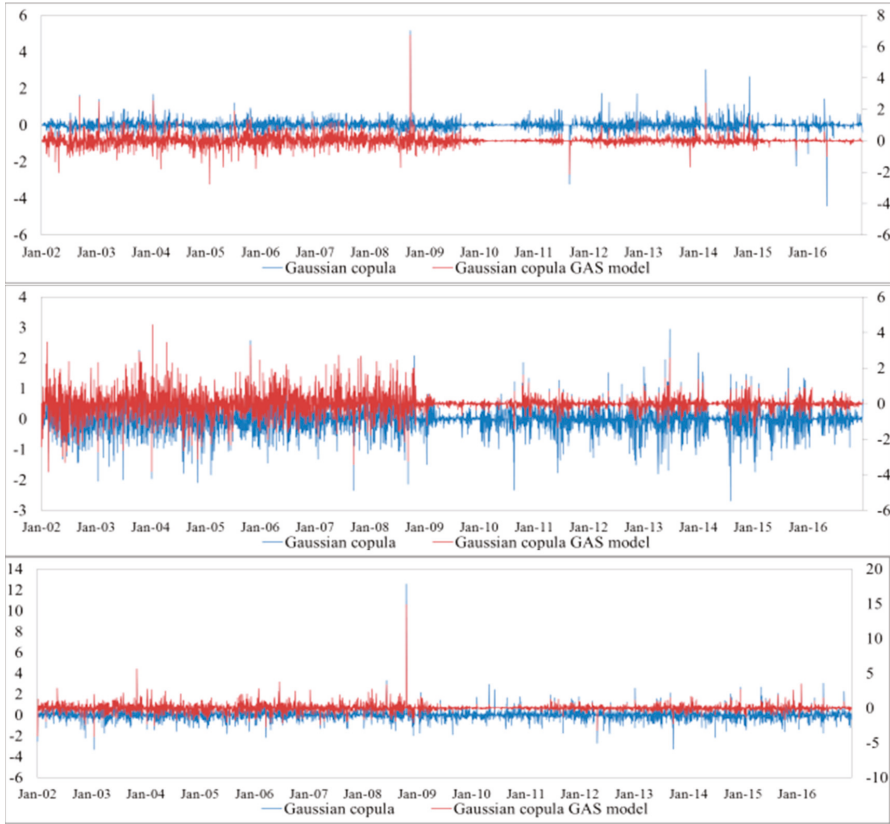Panel B: The risk spillover from the Czech Republic to Germany.

Figure 5. ΔCoVaR between the Czech Republic and Germany. Note: This figure plots the estimated ΔCoVaR between the Czech Republic and Germany from the Gaussian copula and Gaussian copula GAS models for 3-year yields (**top**), 5-year yields (**middle**), and 10-year yields (**bottom**) in the panel (**A**,**B**).

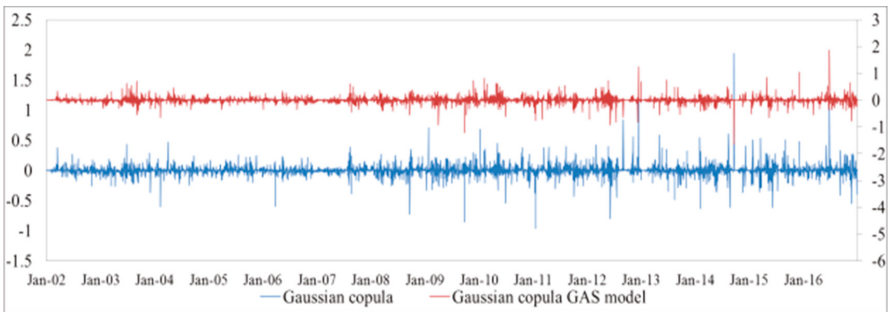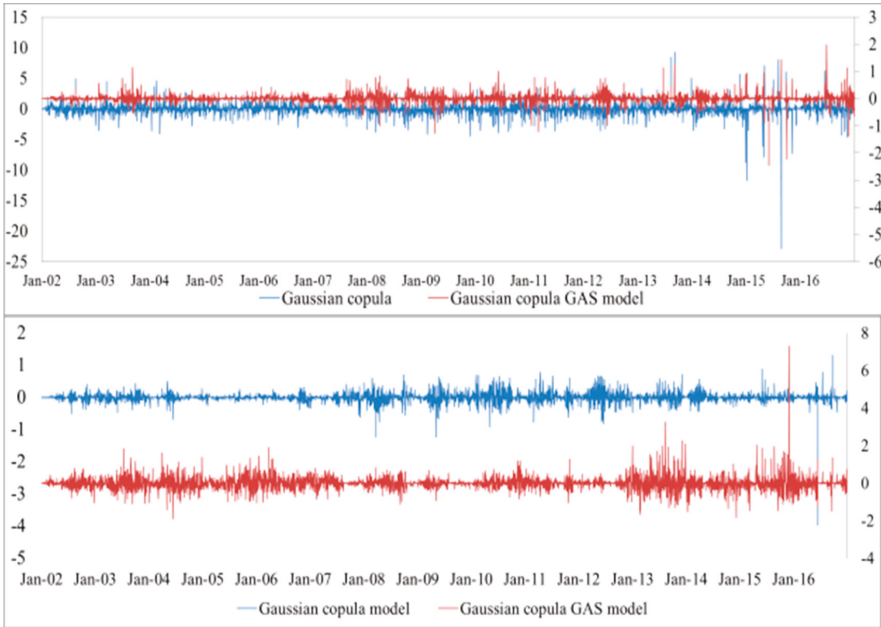Panel A: The risk spillover from Germany to Hungary.

Figure 6. *Cont.*

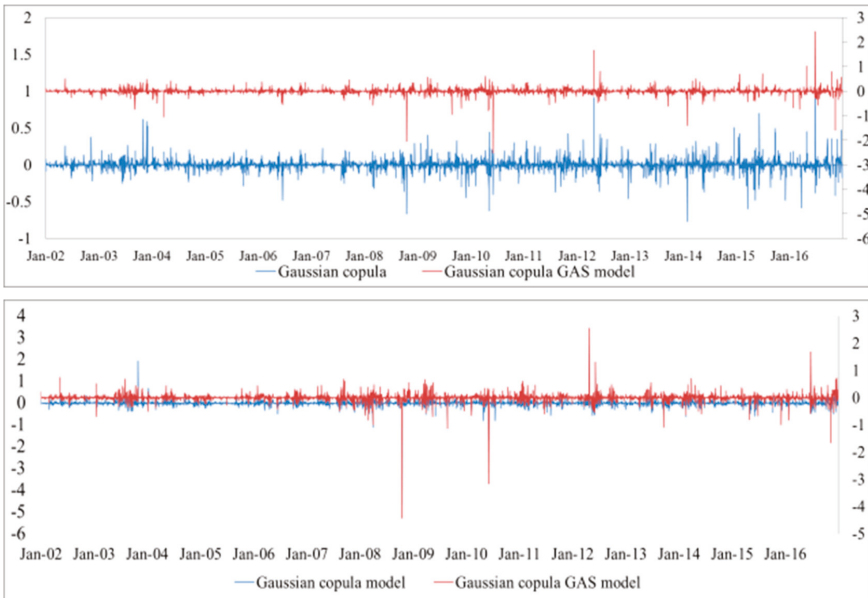Panel B: The risk spillover from Hungary to Germany.
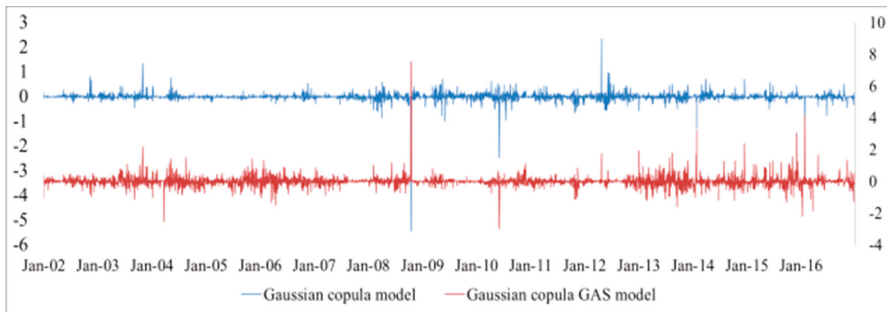

Figure 6. *Cont.*

Figure 6. ΔCoVaR between Hungary and Germany. Note: This figure plots the estimated ΔCoVaR between Hungary and Germany from the Gaussian copula and Gaussian copula GAS models for 3-year yields (**top**), 5-year yields (**middle**), and 10-year yields (**bottom**) in the panel (**A,B**).

Furthermore, our empirical evidence also showed that ΔCoVaR volatility increased substantially for the countries in crisis. The reason may be the uncertainty of the government securities markets and implementation of stabilization policies by the European Central Bank and International Monetary Fund. These actions also provoke sudden changes in investor expectations. All the evidence on the systemic risk dynamics was consistent with the idea that the crisis not only had spillover effects on countries with weak economic fundamentals (e.g., Hungary, which had the highest systemic risk), but also had contagion effects for both the CEEC-3 and Germany.

4.4. Dynamic SJC Copula

To ascertain how these events affected the dependence of the government securities markets in CEEC-3 and Germany, we employed the dynamic SJC (symmetrized Joe-Clayton) copula proposed by Patton [17] to investigate positive and negative events. In particular, we examined the dynamic tail correlations in these markets to find the possibility of contagion or fight to quality. Generally, correlations exist across the markets, but tail correlations do not. If the tail correlations exist across the markets, the contagion or fight to quality will more likely occur as the contagion is more likely to be related to the lower tail dependence, while the fight to quality is more likely to be connected to the upper dependence. Following Patton [17], the density of the SJC copula is

$$c_{SJC}\left(u, v | \tau^U, \tau^L\right) = 0.5[c_{JC}\left(u, v | \tau^U, \tau^L\right) + c_{JC}\left(1 - u, 1 - v | \tau^U, \tau^L\right) + u + v - 1] \tag{14}$$

The SJC copula is symmetric when $\tau^U = \tau^L$ and asymmetric otherwise. To estimate the time-varying dependence structure for the conditional copula, we assumed that the dependence parameter was determined by past information and that it followed an autoregressive moving average, or ARMA (1,10)-type process. Therefore, the dynamics of upper and lower tail dependence can be expressed as Equations (15) and (16), respectively:

$$\tau_t^U = \prod \left(\beta_U^{SJC} \tau_{t-1}^U + \omega_U^{SJC} + \gamma_U^{SJC} \frac{1}{10} \sum_{i=1}^{10} |u_{t-i} - v_{t-i}|\right) \tag{15}$$

$$\tau_t^L = \prod \left(\beta_L^{SJC} \tau_{t-1}^L + \omega_L^{SJC} + \gamma_L^{SJC} \frac{1}{10} \sum_{i=1}^{10} |u_{t-i} - v_{t-i}|\right) \tag{16}$$

where $\prod$ is the logistic transformation to keep τ^U and τ^L within the (0, 1) interval. We also estimated the parameters based on the MSML estimation method.

Table 6 reports the estimation results. For the copula function, β denotes the degree of persistence and γ represents the adjustment in the dependence process. As shown in Table 7, the parameters γ and β are significant only for the Czech Republic and Germany for all maturities, and for Poland and Germany for the 10-year maturity, suggesting that significant variance and strong dependency existed over time in these pairs.

Table 7. Estimation results of the SJC copula model.

	3-Year	5-Year	10-Year
Czech–Germany			
ω^U	0.439 (0.302)	0.072 (0.078)	0.041 (0.015) ***
γ^U	−2.671 (0.431) ***	−0.368 (0.027) ***	−0.196 (0.072) ***
β^U	0.914 (0.051) ***	0.988 (0.017) ***	0.991 (0.004) ***
ω^L	0.044 (0.059)	0.094 (0.092)	0.079 (0.023) ***
γ^L	−0.203 (0.078) ***	−0.453 (0.466)	−0.399 (0.176) **
β^L	0.995 (0.008) ***	0.986 (0.017) ***	0.988 (0.007) ***
Log-Likelihood	78.469	230.020	352.096
Poland–Germany			
ω^U	−1.035 (0.131) ***	−0.180 (0.696)	0.162 (0.065) ***
γ^U	−0.484 (0.465)	−0.538 (1.975)	−0.961 (0.358) ***
β^U	2.106 (2.253)	1.256 (1.353)	0.963 (0.014) ***
ω^L	−3.912 (3.450)	−4.751 (2.950)	−1.115 (1.018)
γ^L	−0.917 (1.507)	−0.658 (0.900)	−9.421 (1.431) ***
β^L	0.903 (1.659)	0.356 (0.371)	−0.112 (0.237)
Log-Likelihood	2.261	22.933	83.667
Hungary–Germany			
ω^U	−4.377 (1.259) ***	−4.849 (3.180)	−4.556 (2.918)
γ^U	−1.021 (18.862)	−1.132 (26.387)	−1.079 (37.612)
β^U	0.801 (1.039)	0.707 (0.415) *	0.795 (1.840)
ω^L	−9.462 (1.007) ***	−9.534 (0.980) ***	−9.473 (1.071) ***
γ^L	−1.185 (2.827)	−1.406 (2.367)	−1.351 (5.567)
β^L	0.303 (0.111) **	0.301 (0.106) ***	0.317 (0.154) **
Log-Likelihood	−11.042	−17.312	−17.291

Notes: *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Figures 7 and 8 compare the time paths of the conditional lower and upper tail dependence based on the SJC copula for Poland and the Czech Republic, respectively. In general, we found that the conditional upper tail dependence was greater and fluctuated more than the conditional lower tail dependence for Poland and for the 3- and 5-year government securities markets because the value of ω^L was less than that of ω^U. Moreover, the variation degree increased as maturities increased. However, the conditional upper tail dependence fluctuated less than the conditional lower tail dependence for the Czech Republic in the 10-year government securities market. In addition, the dynamic process of tail dependence was not well specified for the Poland–Germany and Hungary–Germany pairs since the parameters γ and β were insignificant. Thus, they were omitted.

Meanwhile, these results also indicated that the Czech Republic showed the highest dependence with Germany. In addition, both positive and negative news from Germany significantly affected dependence, with the former having a larger influence than the latter, which was consistent with the findings of Yang and Hamori [7]. In contrast to Büttner and Hayo [4] as well as Yang and Hamori [7,8], however, we provided the dynamic process of dependence between the CEEC-3 and Germany and showed that the positive and negative news affected dependence dynamically. Figures 7 and 8 also confirmed that financial contagion occurred during the global financial and European debt crises, consistent with the evidence provided by Boubakri and Guillaumin [14].
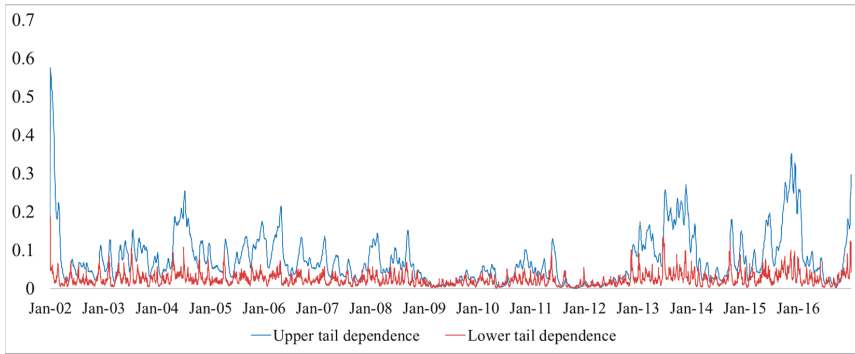
Figure 7. Dynamic tail correlations between Poland and Germany. Note: This figure plots the estimated dynamic tail correlations between Poland and Germany from the SJC copula model for 10-year yields.
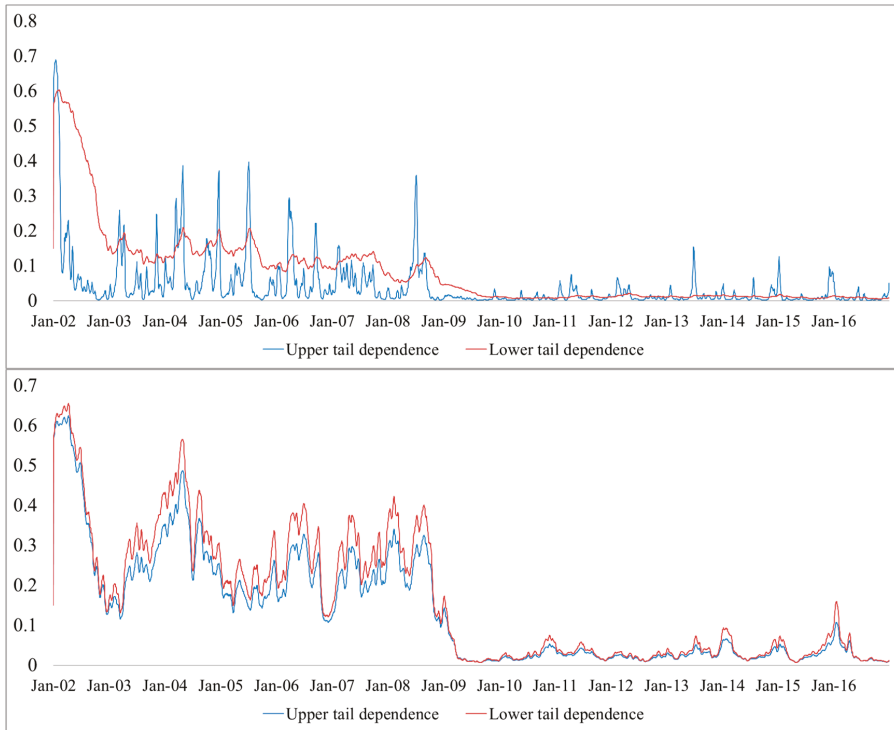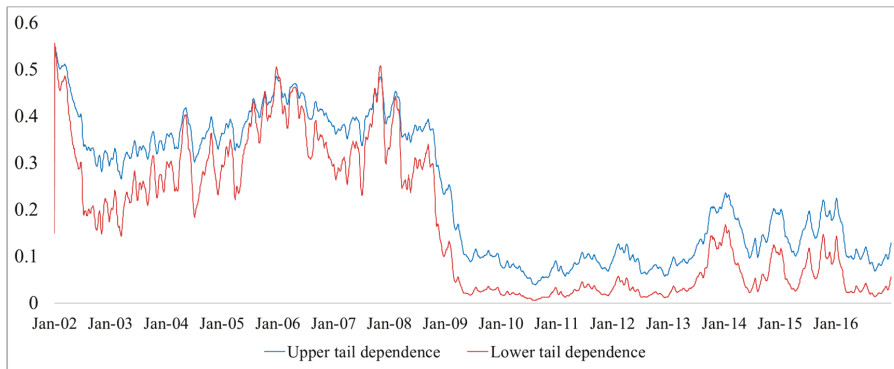

Figure 8. *Cont.*

Figure 8. Dynamic tail correlations between the Czech Republic and Germany. Note: This figure plots the estimated dynamic tail correlations between the Czech Republic and Germany from the SJC copula model for 3-year yields (**top**), 5-year yields (**middle**), and 10-year yields (**bottom**).

5. Conclusions

In this study, we investigated the dependence of the government securities markets in the CEEC-3 and Germany across maturities by employing the GAS-based dynamic Gaussian copula model. We found a high dependence of these government securities markets in the long maturity, but low dependence in the short maturity. In addition, the Czech Republic showed the highest dependence with Germany, while Hungary showed the lowest. Consistent with the findings of Pozzi and Wolswijk [9], by employing the breakpoint test, we also confirmed that EU accession, the global financial crisis, and the European debt crisis caused structural changes in the dynamic correlation.

Furthermore, by employing the ΔCoVaR risk measure, we observed that the German systemic risk was low and relatively stable, while the CEEC-3 systemic risk was high and variant. By considering different time horizons, we showed that the long-run bond ΔCoVaR was higher than the short-run bond ΔCoVaR. This evidence on the systemic risk dynamics shows that the crisis not only had spillover effects on countries with weak economic fundamentals (e.g., Hungary, which has the highest systemic risk), but also had contagion effects for both the CEEC-3 and Germany.

We also employed the SJC copula to examine the dynamic tail dependence among these countries. By comparing and contrasting the results from the dynamic Gaussian copula, we found that both positive and negative news from Germany significantly affected dependence with the Czech Republic, with the former having a larger influence than the latter. These results also showed that the dependence structure between the CEEC-3 and Germany was asymmetric. In addition, we confirmed that the Czech Republic showed the highest dependence with Germany and that financial contagion occurred during the global financial crisis and European debt crisis.

Our results have at least one implication for policymakers and two implications for investors. For policymakers, although the integration of the financial markets in the CEEC-3 has decreased since 2008 owing to market segmentation, becoming an EU member has increased the degree of dependence with European financial markets. For investors, diversification benefits still exist, especially since the global financial crisis. In addition, the dynamic correlations for these countries are more sensitive to positive shocks, indicating that government securities markets remain a good investment, even during a crisis period. Additionally, the risk spillovers from the German government securities market may not be a large concern when compared with those from the CEEC-3 countries.

References

1. Engle, R. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **1982**, *50*, 987–1007. [CrossRef]
2. Bollerslev, T. A conditionally heteroskedastic time series model for speculative prices and rates or return. *Rev. Econ. Stat.* **1987**, *69*, 542–547. [CrossRef]
3. Engle, R. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econ. Stat.* **2002**, *20*, 339–350. [CrossRef]
4. Büttner, D.; Hayo, B. News and correlations of CEEC-3 financial markets. *Econ. Model.* **2010**, *27*, 915–922. [CrossRef]
5. Yang, L.; Hamori, S. EU accession, contagion effects, and financial integration: Evidence from CEEC-3 bond markets. *Transit. Stud. Rev.* **2013**, *20*, 179–189. [CrossRef]
6. Kumar, M.S.; Okimoto, T. Dynamics of international integration of government securities' markets. *J. Bank. Financ.* **2011**, *35*, 142–154. [CrossRef]
7. Yang, L.; Hamori, S. Dependence structure between CEEC-3 and German government securities markets. *J. Int. Financ. Mark. Inst. Money* **2014**, *2*, 109–125. [CrossRef]
8. Yang, L.; Hamori, S. Interdependence between the bond markets of CEEC-3 and Germany: A wavelet coherence analysis. *N. Am. J. Econ. Financ.* **2015**, *32*, 124–138. [CrossRef]
9. Pozzi, L.; Wolswijk, G. The time-varying integration of euro area government bond markets. *Eur. Econ. Rev.* **2012**, *56*, 36–53. [CrossRef]
10. Bekiros, S. Forecasting with a state space time-varying parameter VAR model: Evidence from the Euro area. *Econ. Model.* **2014**, *38*, 619–626. [CrossRef]
11. Creal, D.D.; Koopman, S.J.; Lucas, A. Generalized autoregressive score models with applications. *J. Appl. Econ.* **2013**, *28*, 777–795. [CrossRef]
12. Oh, D.H.; Patton, A.J. Modelling dependence in high dimensions with factor copulas. *J. Bus. Econ. Stat.* **2017**, *35*, 139–154. [CrossRef]
13. Creal, D.D.; Tsay, R.S. High dimensional dynamic stochastic copula models. *J. Econ.* **2015**, *189*, 335–345. [CrossRef]
14. Boubakri, B.; Guillaumin, C. Financial integration and currency risk premium in CEECs: Evidence from the ICAPM. *Emerg. Mark. Rev.* **2011**, *12*, 460–484. [CrossRef]
15. Tobias, A.; Brunnermeier, M.K. CoVaR. *Am. Econ. Rev.* **2016**, *106*, 1705–1741.
16. Girardi, G.; Ergün, A.T. Systemic risk measurement: Multivariate GARCH estimation of CoVaR. *J. Bank. Financ.* **2013**, *37*, 3169–3180. [CrossRef]
17. Patton, A.J. Modelling asymmetric exchange rate dependence. *Int. Econ. Rev.* **2006**, *47*, 527–556. [CrossRef]
18. Glosten, L.R.; Jagannathan, R.; Runkle, D. On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Financ.* **1993**, *48*, 1779–1801. [CrossRef]
19. Engle, R.; Ng, V.K. Measuring and testing the impact of news on volatility. *J. Financ.* **1993**, *48*, 1749–1778. [CrossRef]
20. Nelson, D.B. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* **1991**, *59*, 347–370. [CrossRef]
21. Hansen, B.E. Autoregressive conditional density estimation. *Int. Econ. Rev.* **1994**, *35*, 705–730. [CrossRef]
22. Akaike, H. Information theory and an extension of the maximum likelihood principle. In *2nd International Symposium on Information Theory*; Petrov, B., Csaki, F., Eds.; Akademia Kadio: Budapest, Hungary, 1973; pp. 267–281.
23. Hafner, C.M.; Manner, H. Dynamic stochastic copula models: Estimation, inference and applications. *J. Appl. Econ.* **2012**, *27*, 269–295. [CrossRef]

24. Manner, H.; Segers, J. Tails of correlation mixtures of elliptical copulas. *Insur. Math. Econ.* **2011**, *48*, 153–160. [CrossRef]
25. Avdulaj, K.; Barunik, J. Are benefits from oil–stocks diversification gone? New evidence from a dynamic copula and high frequency data. *Energy Econ.* **2015**, *51*, 31–44. [CrossRef]
26. Reboredo, J.C.; Ugolini, A. Systemic risk in European sovereign debt markets: A CoVaR-copula approach. *J. Int. Money Financ.* **2015**, *51*, 214–244. [CrossRef]
27. McAleer, M. Stationarity and invertibility of a dynamic correlation matrix. *Kybernetika* **2018**, forthcoming. [CrossRef]
28. Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]

Article

e-Purchase Intention of Taiwanese Consumers: Sustainable Mediation of Perceived Usefulness and Perceived Ease of Use

Massoud Moslehpour [1], Van Kien Pham [2], Wing-Keung Wong [3,4,5,*] and İsmail Bilgiçli [6]

[1] Department of Business Administration, Asia University, Taichung 41354, Taiwan; writetodrm@gmail.com
[2] International Economic Faculty, Banking University of Ho Chi Minh City, Ho Chi Minh 84, Vietnam; kienpv@buh.edu.vn
[3] Department of Finance, , Fintech Center, and Big Data Research Center, Asia University, Taichung 41354, Taiwan
[4] Department of Economics and Finance, Hang Seng Management College, Shatin 999077, Hong Kong, China
[5] Department of Economics, Lingnan University, Tuen Mun 999077, Hong Kong, China
[6] Tourism and Hotel Management, Sakarya Üniversitesi Esentepe Kampüsü, Serdivan/Sakarya 54050, Turkey; bilgicli@sakarya.edu.tr
* Correspondence: wong@asia.edu.tw; Tel.: +886-4-2332-3456

Abstract: This study proposes a new model by partially combining personality traits (PT) and Technology Acceptance Model (TAM) attributes to examine the influences of personality characteristics (conscientiousness, openness) and perception of technology (perceived usefulness, perceives ease of use) on e-purchase intention. We use truncate sampling technique and survey questionnaire to target the sample of Taiwanese online consumers and collect data. We find that consciousness (CON) (personality attribute) significantly influences perceived usefulness (PU) (technology perception attributes), perceived ease of use (PEOU) (technology perception attributes) and openness to experience (OPE) (personality attribute). PU, PEOU and OPE have significant impacts on e-purchase intention (INT). PEOU has the strongest positive impact on (INT). In addition, PU, PEOU and OPE combined together mediate the relationship between CON and INT. Further post hoc analysis of the mediation shows that both PU and PEOU are sustainable mediators. However, OPE is not a significant mediator.

---

1. Introduction

Worldwide electronic commerce has significantly advanced due to the prominent development of the internet technology in the past decade. Internet technology has become a growing tool in providing variety website services [1]. Customers seek for products' features, price and functionality online. Consumers no longer need to go to physical stores to browse and compare prices to shop for goods [2]. Accordingly, internet technology has reshaped the way customers buy merchandises. The top categories of online sales are related to technological and electronic products, crafts, handmade products, accessories and clothes [3]. The direct influence of the components of the web marketing mix and purchase intention in China [4] indicates e-commerce in the Asia-Pacific region is becoming more and more popular.

E-commerce in the Asia-Pacific region has increased 30% in 2013, accounting for over one-third of total e-commerce sales all over the world. In 2017, Taiwan's strong online infrastructure growth in terms of internet penetration and smartphone adoption surpasses some of Asia Pacific's key

e-commerce markets including South Korea, Japan, Singapore and Hong Kong. As of 31 March 2017, Taiwan's internet penetration rate reaches 88% with more than 20,000,000 internet users [5]. Taiwan mobile-cellular telephone subscriptions get nearly 30,000,000 subscribers with 78.04% of the population using the Internet by the end of 2015 [6]. An average Taiwanese spent NT$16,586 (US$568) annually in 2012, rising from NT$13,864 (US$475) in 2010 on e-purchasing which accounts over 60% of online people for the whole country. Taiwan's online shopping market is growing rapidly. Therefore, it is important to determine which factors drive e-purchase intention of Taiwanese consumers.

According to Teresa, Bonnie, and Yingjiao [7], personality is one of the roots of understanding consumers' purchase intention. Personality is the internal force that motivates customers to affect a particular behavior and to unconsciously motivate customers. Therefore, marketers have to understand the effect of personality as a direct link to consumers' mind. Previous studies indicate that shoppers' character is a substantial factor for the success of e-vendors [8]. Online purchase intention has been described by different views, such as consuming in relative to demographic characteristics [9–12], emotional and psychological characteristics [13,14], realizing positive and negative aspects of virtual transactions [15–18], shopping motivation [19,20] and orientation of purchasing [21]. Conscientious people are characterized as habitually focused, vigilant and well organized. Nevertheless, a few of the previous research has investigated the influence of conscientiousness on customers' e-purchase intention [22]. We believe the humanistic characteristics of a consciousness person are well matched with computer logarithm physiognomies that lead to the acceptance of technology.

Consuming behavior is sustainably influenced by individual characteristics of users [23]. Most authors conduct their studies under separate models of the Big Five Personality Trait Model [24] and Technology Acceptance Model [25] (Figure 1). Previous research has also studied purchase Intention [26]. However, as far as we know, there is no study in the literature to separate personality traits (PT) and then combine them with Technology Acceptance Model (TAM) to study the e-purchase intention (INT) of the consumers. In this paper, we bridge the gap in the literature to explore the most persuasive part of TAM and PT. We investigate e-purchase intention (INT) of the consumers; displayed as a new model in Figure 2 (we call it Model A). This study, therefore, examines the influence of the conscientiousness (CON) on e-purchase intention (INT) through perceived usefulness (PU), perceived ease of use (PEOU) and openness to experience (OPE). This paper attempts to allocate exclusive perceptions into the online shopping by discussing the determinative factors affecting buying intention of participants.

This study takes partially from the Technology Acceptance Model and partially from the Big Five Personality Trait Model. Conscientiousness is the propensity to show self-discipline and strive for aptitude and attainment [27]. It is one of the most sustainable personality predictors of attitude. Individuals perceive technology differently. Furthermore, the attributes of personality trait influence the way consumers shop online. This study aims at discovering how consumer's characteristics and perceptions about virtual vendors can be adjusted by personalizing e-vendors' environment. The aim is to find ways which encourage consumers and makes it easier for them to buy products online [8].

Taking into consideration the above-mentioned facts, this study provides the key for online sellers to comprehend e-purchase intention of customers. Previous research of individual's characteristics and perception of technology provides an essential ground tool for considering human behavior. As the most common model of personality traits, the Big Five model evaluates the most noticeable sides of personality [28]. Human's traits are basic constructing slabs of personality, which is more likely to lead constant shapes of person's thinking, feeling and manner [29]. In addition, the prototype of Big Five can confirm impacts of time's flow [30] as well as cultural exchanges [31]. The Big Five model elements include extraversion, conscientiousness, agreeableness, neuroticism (emotional stability) and openness to experience. In this paper, authors will limit the variables of personalities by using only conscientiousness as exogenous variable and openness to experiences as one of the mediating variables. Both variables, related to the area of personal behavior, interact and influence perception and use of technology. In traditional studies of The Big Five and consumer behavior the interaction

of the variables might be different. This study explores and hints to the fact that in the new age of e-commerce and online shopping, unlike the traditional models, personality traits are NOT the sole influential elements. Technological perceptions for different age groups play an important role in consumer's e-purchase intention.

In this paper, we initiate the new concept of personality traits in conjunction with technology perceptions. We provide a better understanding of e-purchase intention behavior by looking at the sustainability and consistency of the mediating variables. Conscientiousness describes the reliability, responsibility, diligence and determination of an individual. The results also suggest that although consciousness people are attracted to shopping online, with the mediation of other factors such as openness to experience, perceived ease of use and perceived usefulness the influence of this personality trait becomes insignificant. This study is not product or service specific. Further product detailed analysis is necessary to make the practical implications of this study for specific products sold online.

The rest of this paper is structured as follows. Section 2 discusses the conceptual background of our study and sets all hypotheses that will be tested in our paper. Section 3 discusses the methodology being used in our study. Section 4 discusses our findings. Section 5 concludes the paper, discusses the inferences drawn from our findings and make implications from the findings of our paper.

2. Conceptual Background and Hypotheses

This study combines attributes of Technology Acceptance Model (TAM) and Personality Traits into a new model (Model A as shown in Figure 2). The aim is to have a better understanding of online shoppers' attitudes and e-purchase intention of consumers. The model proposed in this study (Model A) contains several factors as discussed in the following.

2.1. Online Purchase Intention (INT)

Online purchase intention reflects the desire of clients to buy through the internet. It is believed that a shopper is more likely to buy from virtual stores when e-commerce sites provide satisfactory tools including: products/services catalogues, searching functions, trust in websites, pricing comparison sheets, buying carts, online payment systems and outlining devices [32–35]. Thus, considering the importance of each corresponding factor plays an important role for online vendors to draw consumers and sustain sales. Purchase and repurchase intention refer to the outcome of consumers' intention to decide a particular action [36,37] regarding the purchase of a product or service. Two main determinants of behavioral intention toward the aspects of personality are: shoppers' attitude and the subjective criteria [38]. Online shopping intention is the tendency of buyers to represent fixed behavior (i.e., buying) in fixed context (i.e., virtual stores). Pagani, Goldsmith, and Hofacker [39] (p. 252) suggest that combining internet users' personality traits and their perception of technology toward online shopping into one large category may lead to better understanding of human behavior. They suggest that understanding the elements of online users' interaction and intentions is necessary if "one wishes to study web-based phenomena, including social commerce".

2.2. Perceived Usefulness (PU)

The perceived benefits and sustainable advantages of shopping online are summarized as perceived usefulness. Finding bargain price in addition to the low cost of searching online can contribute to the notion of perceived usefulness. The advantages of purchasing from a virtual store compared to purchasing from a traditional store can also contribute to perceived usefulness. The effortlessness of comparing one virtual store to another virtual store is also another factor supporting the perceived usefulness [8]. E-commerce websites supply application software and helpful services that can facilitate shoppers' decisions of buying a product/service. Some services, however, may not be as convenient as in traditional market (e.g., immediate comparison among a variety of products that require tasting, smelling or sensing). Online shopping perceived usefulness could change

consumers' attitudes toward purchasing [40]. Son et al. [41] assert that PU has a positive effect on consumer intention to use an internet application.

2.3. Perceived Ease of Use (PEOU)

PEOU refers to the extent to which a user perceives a particular technology, accessing websites, Internet functions and web-interface is easy to use [42]. It is about perceived necessary technological elements [43]. More specifically, a technology is more favorable for using than another if it is most likely to be approved by online shoppers. In other words, the more complicated a technological application is perceived to be, the more likely the website will be used. PEOU and PU are attributes of the acceptance of technology [41]. The PEOU construct has been used in various contexts such as electronic mail [44], e-commerce [16], m-commerce [45] and intention to use internet applications [46].

2.4. Conscientiousness (CON)

CON is identified by words of "precise", "efficient", "orderly" and "persistent". People with highly conscientious characteristic are often concentrated, careful, trustworthy and well organized, whereas unconscientious persons are most likely to express their distraction, disorganization and having flexibility [47]. It is believed that conscientious persons normally concern about factor of effectiveness. Conscientiousness describes the reliability, responsibility, diligence and determination of an individual [48]. Therefore, we believe that although CON is the major contributing factor of e-purchase intention; when mediated by PU, PEOU and OPE as indicated in Figure 2; the contribution is insignificant.

2.5. Openness to Experience (OPE)

According to the study of Migliore [47], open-minded persons to experience something new are individuals of intelligence, curiosity, free thought and flexible action. Researchers advocate that consumers' "openness to experience" is more inclined (lead by their willingness and flexibility to pursue multiple stimuli) towards Facebook, blog more and engage in various online activities [49–51]. Therefore, open-mindedness is a noteworthy forecaster of the overall virtual use. Also, open-minded persons are more likely to use an online shopping via the internet [52]. The result means that open-minded individuals are more likely to use e-purchase to confirm their inquisitiveness and find out freshly adventured practices [53].

2.6. Development of Hypotheses

Devaraj, Easley, and Crant [54] have successfully observed the support of conscientiousness in the relationship between perceived usefulness of technology and intentions to use the technology such that the relationship is stronger for individuals with higher conscientiousness. Another study done by Punnoose [55] also found that conscientiousness has a significant effect on PU. Based on previous studies, we propose the Hypothesis 1 (H1) as follows:

Hypothesis 1. *CON has a positive association with PU.*

Closed people refer to those who have a simple, plain and straightforward mind. Using new technology is not easy for closed people right from the start. With lacking interest in the variety of new software a certain level of PEOU will not be enough to build up an intention to use the technology. On the other hand, open people like to spend time reflecting on new technology and new products. They enjoy unusual interaction and are willing to overcome any possible challenge when using new technology. Svendsen et al. [56] suggest that there is a positive relationship between OPE and PEOU. This idea is also supported by other researchers. For example, Shambare [57] finds a significant positive effect of OPE on PEOU. Rosen and Kluemper [58] conclude that CON significantly influences PEOU in acceptance of social networking websites technology. In previous studies the internal relationship

between CON and OPE has been ignored. Thus, we set to test the following two hypotheses for
e-purchase intention:

Hypothesis 2. *CON has positive association with PEOU.*

Hypothesis 3. *CON has positive association with OPE.*

Research on the different variables of the TAM reveals significant relationships between PEOU,
PU and INT. Davis [42] argues that PEOU and PU jointly influence behavioral intentions. Kim and
Song [59] propose that perceived usefulness is related to purchase attitude among online shoppers.
If consumers find an e-commerce website useful for shopping, they will advantageously have a better
e-purchasing attitude. Other studies by Yoon and Steege [60], Punnoose [55], Aldás-Manzano et al. [61],
Özbeka et al. [62] and Devaraj [54] also confirm the relationship between PU and INT. Several
researchers have found a positive relation between the PEOU and INT. Childers et al. [63] suggest that
clear and understandable online shopping sites, which require less mental efforts of their users to make
a purchase, are more attractive for potential customers than more complicated ones. Marjan et al. [64]
suggest that clients with "perceived ease of use" tend to have a higher intention of buying goods from
virtual stores. Thus, the ease of use relating to e-commerce website's functions and interfaces is useful
in the forecasting of users' intention towards e-purchase. Perceived ease of use is highly relevant to
clients' manner in experience online shopping accordingly to applications of the internet, which is
strongly associated with intention to purchase. Thus, we propose the following hypothesis:

Hypothesis 4. *PU has positive association with INT.*

Hypothesis 5. *PEOU has positive association with INT.*

Thorbjørnsen et al. [65] indicate that social identity expressiveness is positively related to intention
to use multimedia messaging. Based on Mowen's hierarchical model of personality, Bosnjak et al. [66]
explain and predict people's disposition to use e-purchases [67]. Almost all elements have positive
relationships, including the relationship between conscientiousness and openness to experiences.
This paper uses a hierarchy model with the following hierarchy: surface, situational, compound and
elemental. CON and OPE are among the basic hierarchical and elemental characteristics included in
the Big Five dimensions of characters [67]. Figure 1 shows the interrelations among the major three
factors. Based on this model and recent studies [68], we propose the following two hypotheses.

Hypothesis 6. *OPE has positive association with INT.*

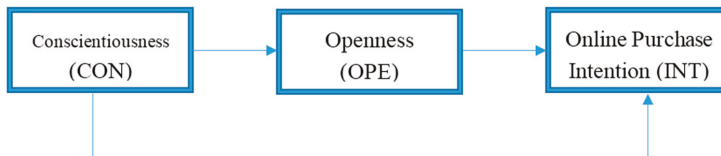**Hypothesis 7.** *CON has positive association with INT.*

Figure 1. The relationships among CON, OPE and INT [67].

Technology Acceptance Model posits that perceived PU and PEOU determine an individual's
intention to use a system, with the intention to use serving as a mediator of actual system use [42].
Previous studies test the mediating effect of PEOU and PU of behavioral intention [69]. However, to the

best of our knowledge, there is no empirical study that tests the mediating effect in the relationship between personality traits and e-purchase intention. Thus, we set the following hypothesis:

Hypothesis 8. *PU, PEOU and OPE mediate the relationship between CON and INT.*

2.7. Theoretical Framework and Hypotheses

The theoretical framework paves the way for the study hypotheses that examine the interrelations among the variables being studied. The framework hypothesizes that personality traits can affect online purchase intention. The model (Model A), depicted in Figure 2, suggests that e-purchase intention (INT) is the endogenous variable which is influenced by the CON as an exogenous variable. Moreover, it is assumed that the effect of CON on INT is fully mediated by PU, PEOU and OPE:

$$\text{Model A: } \eta = f\ (\xi_1,\ \xi_2,\ \xi_3,\ \xi_4) \tag{1}$$

where η is e-purchase intention (INT), ξ_1 be perceived usefulness (PU), ξ_2 be perceived ease of use (PEOU), ξ_3 be openness to experience (OPE) and ξ_4 be conscientiousness (CON). PU, PEOU and OPE are not only the endogenous variables but also the mediators for η. The equation can be exhibited in Figure 2.
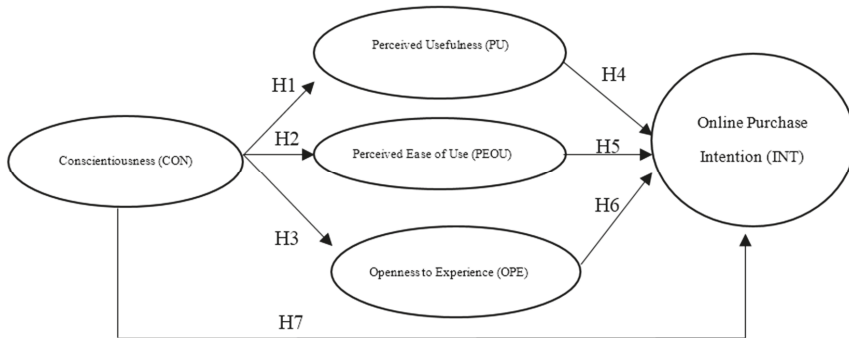
Figure 2. Model A.

3. Research Methodology

In the current study, we researched questionnaires from the previous literature review. Then, we modified the questionnaires to fit into our study in Taiwan. The purpose of the questionnaire is to evaluate the present situation about e-shopping purchase of Taiwanese. Survey instrument is then used in order to study influential attributes of personality and technology on e-purchase intention. In this section, we describe quantitative data analysis process. The results of statistical analysis will help to make the conclusions to obtain a general overview of the entire research model.

3.1. Sample Selection

This project involves the collection of empirical data regarding e-purchase. The population of this research is Taiwanese online shoppers. It is common for Taiwanese, under the age of 50 and above the legal age to shop online and to use technological applications to purchase products or services on virtual stores. Consumers who are more comfortable to access technological applications possess certain characteristics. These characteristics will lead to online shopping decisions and creates a potential e-commerce market [70]. We use truncated sampling technique to collect 380 data. The calculation of sample size is based on confidence level of 95%, Taiwanese population (15–45) of 10,402,409 [71]

and confidence internal (margin of error) of 5%. A total of 316 (out of 380) usable questionnaires from Taiwanese consumers aged from 16 to 45-year-old is collected. We call this sampling technique "truncated sampling" [72] because the data is truncated to observations from 16 to 45-year-old.

The survey questionnaires of this paper are built fundamentally by the features which are chosen based on considerations for the research framework, definition of the variables and literature reviews. Most of the items on the instrument are based on questions used in previous research. Some questions are used in their original form while others are modified slightly to address the specific nature of this study. In the design of a questionnaire for our survey, a complex construct is used so that it is rich in both meaning and multi-dimensional. As shown in Table 1, multiple measures are employed to evaluate the attributes suggested in the previous section, CON (three variables, CON1~CON3), PU (three variables, PU1~PU3), PEOU (three variables, PEOU1~PEOU3), OPE (four variables, OPE1~OPE4) and IN (two variables, IN1~IN2). Table 1 shows the brief descriptions of these variables used in our questionnaire. Our questionnaire containing research questions related to those hypotheses proposed in Section 2 were first circulated to consumers using simple random sampling technique [73]. Respondents were requested to access various attributes with respect to e-purchase and e-service. In order to assess consumers e-purchases and use of e-services, the respondents were also asked to divulge the extent to which they actually use Internet services within a particular period of time. Because Internet users are likely to appreciate the idea of e-purchase and e-service, they should be able to give reasonable expectations in response to our research questions. We "truncate" Taiwanese consumers aged between 16 and 45 years old who possess experiences with e-purchases and e-services. We call this sampling technique "truncated sampling" [72] because the data is truncated to observations from 16 to 45 year-olds who possess experiences with e-purchases. We use the Likert's scale for all of the research questions including five scores ranking from 1 to 5 in which scores 1 and 5 are corresponding to "strongly disagree" and "strongly agree," respectively. Using the truncated sampling technique, we collect 380 data and eventually choose a total of 316 (out of 380) usable questionnaires.

Table 1. Summary of Reliability and Validity.

Variables	Items *	M	SD	Factor Loadings		Cronbach's Alpha
				Model	CFA	
CON	CON1	3.728	0.778	0.695	0.694	
	CON2	3.661	0.741	0.786	0.788	0.781
	CON3	3.655	0.819	0.731	0.734	
PU	PU1	3.769	0.936	0.772	0.773	
	PU2	3.687	0.902	0.732	0.740	0.823
	PU3	3.873	0.875	0.84	0.831	
PEOU	PEOU1	3.718	0.866	0.898	0.899	
	PEOU2	3.750	0.849	0.917	0.915	0.906
	PEOU3	3.766	0.794	0.809	0.811	
OPE	OPE1	3.392	0.879	0.787	0.787	
	OPE2	3.475	0.818	0.857	0.858	0.885
	OPE3	3.415	0.833	0.940	0.939	
	OPE4	3.076	0.960	0.689	0.689	
INT	INT1	3.994	0.866	0.547	0.551	0.640
	INT2	3.915	0.924	0.853	0.856	
Instrument Total	KMO		0.796			0.815
	p-value		0.000			

* Item details are as follows: CON: I see myself as someone who . . . ; CON1: does a thorough job; CON2: can be somewhat careless; CON3: is a reliable worker. PU [42]: Online shopping . . . ; PU1: helps me to make purchases faster; PU2: helps me to make cheaper purchases; PU3: makes it easier for me to make purchases. PEOU [42]: Online shopping . . . ; PEOU1: instructions are easy to follow; PEOU2: is easy to learn how to use; PEOU3: websites are easy to operate. OPE [31]: I see myself as someone who . . . ; OPE1: is original and comes up with new ideas; OPE2: is curious about many different things; OPE3: is ingenious and a deep thinker; OPE4: has an active imagination. INT [38]: Online shopping . . . ; INT1: helps me to make purchases faster; INT2: helps me to make cheaper purchases.

3.2. Research Model

Equation (1) is the mathematical model. To estimate the parameter in Equation (1), we need to obtain the data while η is a latent endogenous variable of the observable variable y and thus, there is an error to measure the latent variable. In addition, all the exogenous variables ξ_1, ξ_2, ξ_3 and ξ_4 are latent and thus, Equation (1) become

$$y = f\ (\xi_1, \xi_2, \xi_3, \xi_4) + \zeta = \eta + \zeta \tag{2}$$

where ζ is an error term with $\Sigma = \text{Cov}\ (\zeta)$. Since all the exogenous variables ξ_1, ξ_2, ξ_3 and ξ_4 are hypothesized to lead to the latent η of endogenous variable y positively, we assume:

$$\partial\eta/\partial\xi_i > 0\ i = 1, 2, 3, 4. \tag{}$$

We use the following linear structural econometric model to approximate Equation (2):

$$y = \Gamma\ \xi + \zeta \tag{3}$$

where $\xi = [\xi_1, \xi_2, \xi_3, \xi_4]'$ the endogenous variable, y, is observable but all the exogenous variables ξ_1, ξ_2, ξ_3 and ξ_4 are unobservable. As such, several said n_i, of observed items of x_i, $(x_{ij}, j = 1, \ldots, n_i)$ are used to measure ξ_i for each $i = 1, 2, 3$ and 4. The measurement model for the exogenous latent variables is:

$$x = \Lambda_x\ \xi + \delta \tag{4}$$

where, $x = [x_{11}, x_{12}, \ldots, x_{1n1}, x_{21}, x_{22}, \ldots, x_{2n2}, x_{31}, x_{32}, \ldots, x_{3n3}, x_{41}, x_{42}, \ldots, x_{4n4}]'$ and $\xi = [\xi_1, \xi_2, \xi_3, \xi_4]'$.

We estimate the model in (2), we employ the exploratory factor analysis to examine all the items within the four exogenous variables and the endogenous variable to obtain the factor loadings. The reliability analysis is then applied to examine the reliability of the items. Thereafter, the structural equation model of the path analysis is used to perform confirmatory factor analysis (CFA) on the hypothesized model in Figure 2.

3.3. Data Analysis Procedure

Descriptive statistics show the value of the personality measure for every personality dimensions. To measure the factor loading, EFA (Explored Factor Analysis) and CFA (Confirmatory Factor Analysis) are employed. Structural equation modeling (SEM) allows the concurrent estimation of multiple multivariate equations. It consists of factor analysis, multiple regression analysis and path model analysis [74]. Path analysis of the hypothesized structural model, EFA, CFA, mediation analysis, validity and reliability estimates are performed by using SPSS 22 and AMOS 22.

4. Result and Findings

4.1. Demographic Characteristic

In this study, 316 usable survey questionnaires are gathered from Taiwanese consumers. The demographic information of the respondents includes: gender (Female, 59.5%) and age (16–25, 54.1%; 26–45, 45.9%). Regarding the age of online shoppers, up to 94.9% are youngsters from 16 to 35. With regard to the educational level, the study indicates that most of the respondents are educated beyond high school (90.2%).

4.2. Test of Reliability and Validity

Reliability and Validity tests are used to test whether the data are reliable and valid. The tests are the essential part of all research designs and although distinct, they are often discussed together [75]. The reliability coefficient of Cronbach's alpha conducted is obtained to evaluate internal consistency

among the survey instruments. If the value of coefficient alpha is between 0.6 and 0.8 and Factor loadings should be 0.5 or higher, the instrument is considered reliable and an α value over 0.8 is considered highly reliable [76]. The results of the reliability test indicate that all constructs and the instrument are reliable or highly reliable (Table 1).

Validity test is completed by exploratory factor analysis which is known as an important tool for researchers since it can be useful for refining measures, evaluating constructing validity and in some cases testing hypotheses [77]. In so doing, this study uses KMO and Bartlett's Test of Sphericity to test the validity of the instrument.

As a result, and shown in Table 1, all factor loadings are higher than 0.5. For the KMO measure of sampling adequacy, in this analysis the KMO is 0.84 which is higher than 0.6. The chi-square $\chi^2 = 3133.25$, significant with $\rho < 001$. All factor loadings are above 0.5. In conclusion, the constructs used in this study are valid and the results from factor analysis can be used for further analysis.

4.3. Confirmatory Factor Analysis (CFA)

In this process, all variables are included as exogenous variables in the proposed model using AMOS 22. In EFA, we explore factor structure while in CFA, we confirm the factor structure we extracted from EFA to improve the overall model fit of the model. The CFA results indicate that Chi-square/Degrees of freedom $(\chi^2/\mathrm{df}) = 1.54$. The GFI (goodness of fit index) = 0.95. The results meet the criteria of a good model fitting suggested by Bentler and Bonnet [78] and others. It is commonly agreed [79] that the comparative fit index (CFI) should be greater than 0.90 to indicate well fit. In this study, the CFI = 0.98, implying our model fits the data very well. The Tucker-Lewis coefficient (TLI) = 0.98. TLI value is close to 1, indicating a very good fit [78]. In addition, MacCallum, Browne, and Sugawara [80] suggest that it is an adequate fit if the root mean square error of approximation (RMSEA) is less than 0.08. In this paper, the RMSEA = 0.04. Thus, all the model fit indices of this study indicate a good fit for the structural model.

4.4. Composite Reliability, Convergent and Discriminant Validity

Convergent Validity is the extent to which indicators of a specific variable 'converge' or share a high proportion of variance in common. Convergent Validity consists of two items: composite reliability (CR) and average variance extracted (AVE). CR is a measure of reliability and internal consistency based on the square of the total of factor loadings for a construct [81]. On the other hand, AVE is a summary measure of convergence among a set of items representing a variable [82]. It is the average percent of variation explained among the items [74,81]. As suggested by Anderson and Gerbing [81], the CR should be greater than 0.7 while Fornell and Larcker [82] suggest AVE to be at least 0.5.

Besides CR and AVE as discussed before, there are a few measures that are useful for establishing validity and reliability: maximum shared variance (MSV), maximum reliability (Max R) and maximum inter-construct correlation (Max r). As indicated in Table 2 for all variables: CR > 0.7, AVE > 0.5, MSV < AVE and $\sqrt{\text{AVE}}$ > Max correlation [74,82]. The results of Table 2 indicate the validity and reliability of the instrument and the CFA. Fornell and Larcker [82] suggest comparing the square root of AVE with the maximum correlation coefficients of latent variables. The result indicates that $\sqrt{\text{AVE}}$ > Max r (Table 2). The findings of all indicators provide the evidence supporting the reliability and validity of the indicators of the proposed model.

Table 2. Test of Composite Reliability, Convergent Validity, and Discriminant Validity.

	CR	AVE	MSV	Max R	Max r	PEOU	CON	OPE	PU	INT
PEOU	0.908	0.768	0.444	0.919	0.666	0.876				
CON	0.783	0.547	0.099	0.938	0.315	0.315	0.740			
OPE	0.893	0.678	0.068	0.965	0.260	0.070	0.260	0.823		
PU	0.825	0.612	0.088	0.970	0.296	0.251	0.280	0.164	0.782	
INT	0.701	0.518	0.444	0.973	0.666	0.666	0.211	0.038	0.296	0.720

CR > 0.7; AVE > 0.5; MSV < AVE; $\sqrt{\text{AVE}}$ > Max r; $\sqrt{\text{AVE}}$ is bold face diagonal.

4.5. Test of Hypotheses

We employ Structural Equation Modeling (SEM) technique to analyze the measurement model, estimate the structural model and test the proposed research hypotheses. The fitted structural model indicates the following values: CMIN/DF = 1.54, which complies with the suggested criteria of <3, Goodness of Fit Index (GFI) = 0.95, Comparative Fit Index (CFI) = 0.98 and Root Mean Square Error of Approximation (RMSEA) = 0.04. The results shown in the structural model meet the minimum requirement of acceptable values, indicating a good fit of the conceptual model on the empirical data in this study.

Table 3 presents the result of the proposed hypotheses for the proposed research model. From the table, we find that H1, H2, H3, H4, and H5 are supported but H6 and H7 are rejected in the research model. The H1 is supported because CON shows positive (β = 0.41) and significant ($p < 0.001$) effect on PU. From this result, we can conclude that CON has a significantly positive effect on PU. Similarly, CON also shows significantly positive effects on both PEOU (β = 0.39) and OPE (β = 0.32) with significance level $p < 0.001$, implying both H2 and H3 are supported. Both PU and PEOU indicate significant impacts on INT with β = 0.17 ($p < 0.05$) and 0.79 ($p < 0.001$), respectively. OPE, however, does not influence INT significantly, resulting in the rejection of H6. The indirect influence of CON on INT (H7) with the presence of the mediating variables (PU, PEOU, OPE) is rejected with β = 0.01 ($p > 0.05$). The influence of CON on INT is supported without the presence of the mediators is significant (β = 0.28, $p < 0.01$) (Figure 3). This suggests the existence of significant mediating effects as stated in the H8.

Table 3. Results of the hypotheses.

Hypothesis	Description	Estimate	Results
H1	CON→PU	0.41 ***	Supported
H2	CON→PEOU	0.39 ***	Supported
H3	CON→OPE	0.32 ***	Supported
H4	PU→INT	0.17 *	Supported
H5	PEOU→INT	0.79 ***	Supported
H6	OPE→INT	−0.03	Rejected
H7	CON→INT (with mediation of PU, PEOU, INT)	0.01	Rejected

* $p < 0.05$, *** $p < 0.001$.

The mediating test of the PU, PEOU and OPE indicates significant full mediation with the presence of the three mediating variables. Therefore, H8 is supported as shown in Table 4. This result prompts us to look at each mediator separately to have a better understanding of all the mediation effects. The post hoc analysis of the mediation indicates that both PU and PEOU act as significant mediators between CON and INT. However, OPE is not a significant mediator. Figure 3 provides a visual representation of the proposed eight hypotheses of the study.

Table 4. Results of the mediation hypotheses.

		Indirect (ab)	Direct (c′)	Total (c)	Mediation
H8	CON→(PU, PEOU, OPE)→INT	0.31 ***	0.01 (NS)	0.32 ***	Supported
H8a	CON→PU→INT	0.09 **	0.2 (NS)	0.29 **	Supported
H8b	CON→PEOU→INT	0.30 ***	0.01 (NS)	0.31 *	Supported
H8c	CON→OPE→INT	0.01 (NS)	0.27 **	0.28 **	Rejected

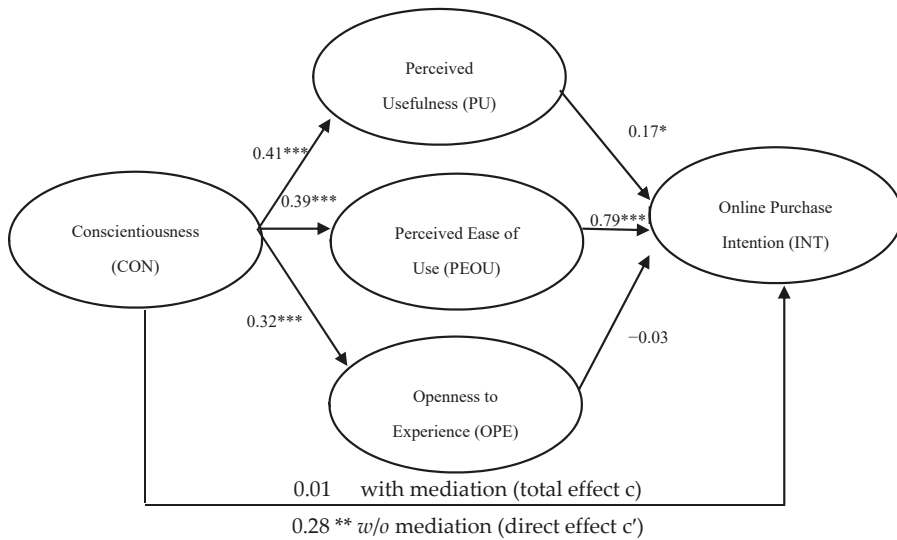* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 3. Result of Hypothesis Test. * *p* < 0.05, ** *p* < 0.01, *** *p* < 0.001.

5. Discussion, Conclusions and Managerial Implications

5.1. Discussion and Conclusions

There are some papers that study about personality traits and technology perception as they relate to online purchase intention. However, most of those studies ignore the mediating effects of some factors such as perceived usefulness, perceived ease of use and openness to experience related to personality traits. For these reasons, this study extends previous works by first combining personality traits and technology perception to offers a total of eight hypotheses. These hypotheses are created and tested by using the multiple regression analysis. The purpose of this research is to test the direct relationship between the exogenous variables (CON, PU, PEOU and OPE) and endogenous variable (INT). Furthermore, we investigate the indirect relationship between CON and INT through the mediation roles of PU, PEOU and OPE. The final results indicate that all hypotheses are supported except the mediating role of OPE. We will discuss every single hypothesis as follows.

H1 CON has a significant influence on PU. This finding is consistent with previous works like Devaraj et al. [54], Punnoose [55] and others that support the significant connection between CON and PU. Intentions to experience the technology is stronger for individuals with higher conscientiousness. This study extends the previous findings to show the influence of CON on INT through PU. Costa and McCrae [31] characterize CON personality trait as being about people who are naturally motived and those who strive to aim at achievements at a high level and perform positive actions. The signs of CON person are self-control reflected in a need for performance, order and persistence. Therefore, in the proposed model of this study, CON represents an intrinsically driving force for toward the e-purchase. Conscientious people are cautious when reviewing whether technological applications will allow them to be more effective. On the other hand, if a conscientiousness person concludes that a technology is not beneficial, then he or she will not use it [54].

H2 CON has a significant impact on PEOU. CON is the characteristic of people who think everything inside out with persistence and caution. Is the characteristic of a person expecting that actions will end with the highest possible results? Person of CON is often efficient and ordered as opposed to free and easy. It means that this person who is generally organized and dependable will perform actions based on a clear and particular schedule instead of immediately thought. PEOU refers

to technologies and interfaces on e-commerce websites which is more favorable to access than another is more likely to be accepted by the participants. Since they are very precise, perfectionistic and effective, they prefer something that is very efficient to use. When they look for a website, they will eagerly compare some features that make the website easier to understand and thus, more efficient to use. They are very aware of the usefulness of the feature of the website. For these reasons, online shopping websites have to create improvement on the usefulness to attract conscientiousness online shoppers by performing time and money saving, quick service, low risks, etc.

H3 states that CON significantly influences OPE. Conscientious personalities usually are very concerned about effectiveness, while OPE is implied as being intellectually curious, open to new ideas, involves imaginative and creative cognition styles [47]. Thus, CON personalities have to be more open to new ideas which make their life become more efficient when they use the approach of online purchasing. This study extends the previous research to show the internal relationship of personality characters toward e-purchase intention.

H4 shows that PU has a significant impact on INT. Previous studies indicate that shoppers' PU of an online store positively influences their buying intention and repurchase intention in the future. Online purchase intention is determined by perceived usefulness of technological innovation [43,83]. The results of this study imply that Taiwanese customers who perceive technology as a useful tool for shopping online tend to use more online for their purchase. The result is consistent with studies of Aldás-Manzano et al. [61] and Yoon and Steege [60]. They find that online shoppers in different countries have the same expectation of looking for advantages of purchase via the internet by themselves. As a result, PU is one of the major key drivers of usage behavior and intention to purchase.

H5 states that PEOU has a significant influence on INT. Therefore, technologies which are perceived as easy to use will stimulate customers to purchase online. In addition, Childers et al. [63] argue that clear and understandable online shopping sites, which require less mental efforts to make a purchase, are more attractive for potential customers than more complicated ones [63]. This finding is very important since online shoppers reveal that the conveniences of websites, which are free from effort in order to make an order, are part of the main factors to make virtual store to be successful.

H6 states that OPE is a sustainably determining factor related to INT. People with high OPE are more likely to explore something new and thus, like to participate in e-activities. As proposed by Tuten and Bosnjak [53], it is reasonable to meet the need of this type of people and find out new adventurous way for them. OPE is one of the important forecasting elements influencing the growth of internet's applications [52]. It means that people who exhibit this type of personality trait (openness) would prefer to purchase online than go to the traditional stores. The results are in accordance with Arnold and Reynolds [84] who argue that a person with higher level of OPE is more likely to buy goods and get services in virtual stores.

H7 formulates that CON has a significant impact on INT. The awareness of high conscientious people is very high. They prefer to make themselves to do something that is more efficient and well organized. They like to do something orderly but not something spontaneously. There is a possibility that conscientious people prefer online shopping to traditional practices since online shopping is a method that is more organized, more efficient and more specific. Moreover, going to the traditional stores that requires staying in traffic, wandering around and could end up wasting time with the possibility of buying nothing is not attractive to conscientious people. While online store provides more detailed information about prices, products and how to buy something without going to the shops but delivering to the destination place. Thus, it helps online buyers to compare the prices and products from one website to another and help them to consume products or get services that match with their need. Hence, it will be more comfortable and more effective for those who seek for efficiency and precisely.

H8 states that PU, PEOU and OPE mediate the relationship between CON and INT. As a result, both PU and PEOU play an important role of full mediators in the relationship between CON and INT while OPE does not. As discussed in the theoretical section, PU, perceived usefulness of purchasing

from an online store, is known as the perceived benefits and the overall perceived advantages of online shopping sites. PEOU, perceived ease of use, is about perceived necessary technological elements approved by online shoppers. Obviously, the influence of conscientiousness on online purchase intention is strongly meditated by the benefits and technological elements that an online store provides to customers. Therefore, online shopping service providers should highly focus on the improvement of the latest and advanced technology that enables online shoppers to purchase goods or get services easily. Furthermore, in order to increase the benefits perceived from buyers, online shopping service providers need to provide easy ordering, convenient delivery, customer training, customer consulting, maintenance and repair, returns, etc. The findings also indicate existence of a non-mediating effect of OPE. In other words, openness to experience, in the presence of PU and PEOU, is not a good mediator in the relationship between CON and INT as found out in the proposed model of this study. Thus, it is important for virtual stores to put more emphasis on PU and PEOU when they observe the influences of conscientious factor on online purchase intention.

In conclusion, our study successfully explores the direct and indirect influence of both CON on INT under the sample of Taiwanese online shoppers. Our findings provide many significant applications that advice online shopping service providers need to study clearly about conscientiousness personality of online shoppers to draw purchase intention from them. In doing so, online shopping sites should focus on utility-motivated factors like technology throughout the process of online shopping. Since consumers that have conscientiousness personality often use their knowledge to carefully access if products or services bring them the desired values.

5.2. Managerial Applications

This research shows evidence that personal traits are important determinants affecting e-shoppers' buying intention. As a result, it is a noteworthy background toward studying the field of online consuming behavior. Actually, online vendors can use the result as fundamental reference material to plan an effective consumer-oriented strategy and attract more consumers to buy goods and get services online. The results also suggest that although consciousness people are attracted to shopping online, with the mediation of other factors such as openness to experience, perceived ease of use and perceived usefulness, the influence of this personality trait becomes insignificant. Online shoppers' intention to purchase online will increase through perceived usefulness, perceived ease of use and their openness to experience of new technology. Therefore, online store owners and web designers can attract consciousness people by increasing easy and useful functions during the process of purchasing such as ease of order, ease of payment, time and money savings, while considering the other personality trait like openness to experience. Therefore, this research is a good reference for web designers, online shopping vendors and technological applications and software providers. Specifically, the web-designers need to create eye-catching sales interface of websites. Additionally, website owners have to come up with smart tools and functionalities of the technological applications to help users go online shopping easily even unprecedented access to technology previously. Duties of online providers is to find ways to encourage customers' interest in the products and services in their virtual stores by providing online shoppers easy and useful purchasing service with obvious and detailed information including products, prices, promotions, delivery, installation, warranty, guarantee, etc. It is the aim of this study to aid online vendors in using the findings in our study effectively to draw online purchase intention from conscientiousness consumers.

Last, we note that is worthwhile investigating whether the findings in our paper are consistent in some sub-samples, like dividing the sample into two subsamples and analyzing different groups based on gender, age group and income, to name a few. In addition, it is common that after one interesting case study paper published, there are many papers extend the work to use different samples to show that the results are not, by chance, confirming the theory. Hence, extensions may include comparison with any other samples. Readers may refer to Wong et al. [85] and the references therein for more information on the discussion.

Author Contributions: Massoud Moslehpour initiate the ideas, formulate the problems, do detailed planning, develop the theory, developed the initial research model and conduct the initial analysis. Van Kien Pham improved the research model, conduct the analysis and write up the first draft. Wing-Keung Wong assist Massoud Moslehpour to further formulate the problems, do detailed planning, and improve the theory, and make the theory more advanced, and check the final research model and statistical analyses İsmail Bilgiçli assists us to write up and help proof read our paper.

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

References

1. Rodríguez Bolívar, M.P. Policy makers' perceptions on the transformational effect of Web 2.0 technologies on public services delivery. *Electron. Commer. Res.* **2017**, *17*, 227–254. [CrossRef]
2. Wind, J.; Mahajan, V. *Digital Marketing: The Challenge of Digital Marketing*; John Wiley & Son: New York, NY, USA, 2001.
3. Cummins, S.; Peltier, J.W.; Schibrowsky, J.A.; Nill, A. Consumer behavior in the online context. *J. Res. Interact. Mark.* **2014**, *8*, 169–202. [CrossRef]
4. Choi, Y.; Jin, J. Is the Web Marketing Mix Sustainable in China? The Mediation Effect of Dynamic Trust. *Sustainability* **2015**, *7*, 13610–13630. [CrossRef]
5. Internet World Stats. Internet Usage in Asia. Available online: http://www.internetworldstats.com/stats3.htm (accessed on 25 July 2017).
6. International Telecommunication Union Statistics. Available online: http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx (accessed on 25 July 2017).
7. Teresa, A.S.; Bonnie, D.B.; Yingjiao, X. Predicting purchase intention of a controversial luxury apparel product. *J. Fash. Mark. Manag. Int. J.* **2006**, *10*, 405–419. [CrossRef]
8. Barkhi, R.; Wallace, L. The impact of personality type on purchasing decisions in virtual stores. *Inf. Technol. Manag.* **2007**, *8*, 313–330. [CrossRef]
9. Brown, M.; Pope, N.; Voges, K. Buying or browsing? An exploration of shopping orientations and online purchase intention. *Eur. J. Mark.* **2003**, *37*, 1666–1684. [CrossRef]
10. Korgaonkar, P.; Silverblatt, R.; Becerra, E. Hispanics and Patronage Preferences for Shopping from the Internet. *J. Comput. Med. Commun.* **2004**, *9*, 31–44. [CrossRef]
11. Park, J.; Lee, D.; Ahn, J. Risk-Focused E-Commerce Adoption Model: A Cross-Country Study. *J. Glob. Inf. Manag.* **2014**, *7*, 6–30. [CrossRef]
12. Stafford, T.F.; Turan, A.; Raisinghani, M.S. International and Cross-Cultural Influences on Online Shopping Behavior. *J. Glob. Inf. Manag.* **2004**, *7*, 70–87. [CrossRef]
13. Huang, M.H. Modeling Virtual Exploratory and Shopping Dynamics: An Environmental Psychology Approach. *Inf. Manag.* **2003**, *41*, 39–47. [CrossRef]
14. Xia, L. Affect as Information: The Role of Affect in Consumer Online Behaviors. *Adv. Consum. Res.* **2002**, *29*, 93–100.
15. Bhatnagar, A.; Ghose, S. A Latent Class Segmentation Analysis of E-Shoppers. *J. Bus. Res.* **2004**, *57*, 758–767. [CrossRef]
16. Featherman, M.S.; Pavlou, P.A. Predicting E-Services Adoption: A Perceived Risk Facets Perspective. *Int. J. Hum. Comput. Stud.* **2003**, *59*, 451–474. [CrossRef]
17. Garbarino, E.; Strahilevitz, M. Gender Differences in the Perceived Risk of Buying Online and the Effects of Receiving a Site Recommendation. *J. Bus. Res.* **2004**, *57*, 768–775. [CrossRef]
18. Huang, W.Y.; Schrank, H.; Dubinsky, A.J. Effect of Brand Name on Consumers' Risk Perceptions of Online Shopping. *J. Consum. Behav.* **2004**, *4*, 40–50. [CrossRef]
19. Novak, T.P.; Hoffman, D.L.; Yung, Y.F. Measuring the Customer Experience in Online Environments: A Structural Modeling Approach. *Mark. Sci.* **2000**, *19*, 22–42. [CrossRef]

20. Wolfinbarger, M.; Gilly, M. Shopping Online for Freedom, Control and Fun. *Calif. Manag. Rev.* **2001**, *43*, 34–56. [CrossRef]
21. Schiffman, L.G.; Kanuk, L.L. *Consumer Behavior*; Prentice-Hall: Upper Saddle River, NJ, USA, 2000.
22. Zhou, T.; Lu, Y. The Effects of Personality Traits on User Acceptance of Mobile Commerce. *Int. J. Hum. Comput. Interact.* **2011**, *27*, 545–561. [CrossRef]
23. Mount, M.K.; Barrick, M.R.; Scullen, S.M.; Rounds, J. Higher-Order Dimensions of the Big Five Personality Traits and the Big Six Vocational Interest Types. *Pers. Psychol.* **2005**, *58*, 447–478. [CrossRef]
24. Myung, J.K.; Choi, Y.H.; Kim, J.D. Effects of CEOs' Negative Traits on Corporate Social Responsibility. *Sustainability* **2017**, *9*, 543. [CrossRef]
25. Chin, J.; Lin, S.C. A Behavioral Model of Managerial Perspectives Regarding Technology Acceptance in Building Energy Management Systems. *Sustainability* **2016**, *8*, 641. [CrossRef]
26. Wei, X.; Jung, S. Understanding Chinese Consumers' Intention to Purchase Sustainable Fashion Products: The Moderating Role of Face-Saving Orientation. *Sustainability* **2017**, *9*, 1570. [CrossRef]
27. Greenberg, J.; Baron, R.A. *Behavior in Organizations*, 9th ed.; Pearson Education: Upper Saddle River, NJ, USA, 2007.
28. Huang, J.H.; Yang, Y.C. The relationship between personality traits and online shopping motivations. *Soc. Behav. Personal.* **2010**, *38*, 673–680. [CrossRef]
29. Pervin, L.A. *The Science of Personality*; Wiley: New York, NY, USA, 1996.
30. Hampson, S.E.; Goldberg, L.R. A first large-cohort study of personality-trait stability over the 40 years between elementary school and midlife. *J. Personal. Soc. Psychol.* **2006**, *91*, 763–779. [CrossRef] [PubMed]
31. Costa, P.T., Jr.; McCrae, R.R. Domains and Facets: Hierarchical Personality Assessment Using the Revised NEO Personality Inventory. *J. Personal. Assess.* **1995**, *64*, 21–50. [CrossRef] [PubMed]
32. Liao, Z.; Wong, W.K. The determinants of customer interactions with internet-enabled e-banking service. *J. Oper. Res. Soc.* **2008**, *59*, 1201–1210. [CrossRef]
33. Chen, Y.H.; Hsu, I.C.; Lin, C.C. Website attributes that increase consumer purchase intention: A conjoint analysis. *J. Bus. Res.* **2010**, *63*, 1007–1014. [CrossRef]
34. Liao, Z.; Shi, X.; Wong, W.K. Consumer Perceptions of the Smartcard in Retailing: An Empirical Study. *J. Int. Consum. Mark.* **2012**, *24*, 252–262. [CrossRef]
35. Liao, Z.; Shi, X.; Wong, W.K. Key determinants of sustainable smartcard payment. *J. Retail. Consum. Serv.* **2014**, *21*, 306–313. [CrossRef]
36. Moslehpour, M.; Wong, W.K.; Aulia, C.K.; Pham, V.K. Repurchase intention of Korean beauty products among Taiwanese consumers. *Asia Pac. J. Mark. Logist.* **2017**, *29*, 569–588. [CrossRef]
37. Moslehpour, M.; Wong, W.K.; Lin, Y.H.; Nguyen, T.L.H. Top purchase intention priorities of Vietnamese low cost carrier passengers: Expectations and satisfaction. *Eurasian Bus. Rev.* **2017**, 1–19. [CrossRef]
38. Ajzen, I. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* **1991**, *50*, 179–211. [CrossRef]
39. Pagani, M.; Goldsmith, R.E.; Hofacker, C.F. Extraversion as a stimulus for usergenerated content. *J. Res. Interact. Mark.* **2013**, *7*, 242–256. [CrossRef]
40. Mohamed, N.; Hussein, R.; Zamzuri, N.H.A.; Haghshenas, H. Insights into individual's online shopping continuance intention. *Ind. Manag. Data Syst.* **2014**, *114*, 1453–1476. [CrossRef]
41. Son, J.; Sadachar, A.; Manchiraju, S.; Fiore, A.M.; Niehm, L.S. Consumer adoption of online collaborative customer co-design. *J. Res. Interact. Mark.* **2012**, *6*, 180–197. [CrossRef]
42. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **1989**, *13*, 318–340. [CrossRef]
43. Burton-Jones, A.; Hubona, G.S. Individual differences and usage behaviour: Revisiting a technology acceptance model assumption. *Date Base Adv. Inf. Syst.* **2005**, *36*, 58–77. [CrossRef]
44. Szajna, B. Empirical Evaluation of the Revised Technology Acceptance Model. *Manag. Sci.* **1996**, *42*, 85–92. [CrossRef]
45. Bruner, G.C.; Kumar, A. Explaining consumer acceptance of handheld Internet devices. *J. Bus. Res.* **2005**, *58*, 553–558. [CrossRef]
46. King, W.R.; He, J. A meta-analysis of the technology acceptance model. *Inf. Manag.* **2006**, *43*, 740–755. [CrossRef]
47. Migliore, L.A. Relation between big five personality traits and Hofstede's cultural dimensions: Samples from the USA and India. *Cross Cult. Manag. Int. J.* **2011**, *18*, 38–54. [CrossRef]

48. Banerjee, S. Influence of consumer personality, brand personality, and corporate personality on brand preference: An empirical investigation of interaction effect. *Asia Pac. J. Mark. Logist.* **2016**, *28*, 198–216. [CrossRef]

49. Correa, T.; Hinsley, A.W.; de Zúñiga, H.G. Who interacts on the web?: The intersection of users' personality and social media use. *Comput. Hum. Behav.* **2010**, *26*, 247–253. [CrossRef]

50. Jadin, T.; Gnambs, T.; Batinic, B. Personality traits and knowledge sharing in online communities. *Comput. Hum. Behav.* **2013**, *29*, 210–216. [CrossRef]

51. Tsao, W.C. Big Five personality traits as predictors of internet usage categories. *Int. J. Manag.* **2013**, *30*, 374–386.

52. McElroy, J.C.; Hendrickson, A.R.; Townsend, A.M.; DeMarie, S.M. Dispositional factors in Internet use: Personality versus cognitive style. *MIS Q.* **2007**, *31*, 809–820. [CrossRef]

53. Tuten, T.L.; Bosnjak, M. Understanding differences in web usage: The role of need for cognition and the Five Factor Model of personality. *Soc. Behav. Personal.* **2001**, *29*, 391–398. [CrossRef]

54. Devaraj, S.; Easley, R.F.; Crant, J.M. How Does Personality Matter? Relating the Five-Factor Model to Technology Acceptance and Use. *Inf. Syst. Res.* **2008**, *19*, 93–105. [CrossRef]

55. Punnoose, A.C. Determinants of Intention to Use e-Learning Based on the Technology Acceptance Model. *J. Inf. Technol. Educ. Res.* **2012**, *11*, 301–337.

56. Svendsen, G.B.; Johnsen, J.A.K.; Almas-Sørensen, L.; Vittersø, J. Personality and technology acceptance: The influence of personality factors on the core constructs of the Technology Acceptance Model. *Behav. Inf. Technol.* **2013**, *32*, 323–334. [CrossRef]

57. Shambare, R. Factors influencing the adoption of cell phone banking by South African students. *Afr. J. Bus. Manag.* **2013**, *7*, 30–38.

58. Rosen, P.A.; Kluemper, D.H. The Impact of the Big Five Personality Traits on the Acceptance of Social Networking Website. In Proceedings of the Fourteenth Americas Conference on Information Systems (AMCIS), Toronto, ON, Canada, 14–17 August 2008; Volume 274.

59. Kim, H.; Song, J. The quality of word-of-mouth in the online shopping mall. *J. Res. Interact. Mark.* **2010**, *4*, 376–390. [CrossRef]

60. Yoon, H.S.; Barker Steege, L.M. Development of a quantitative model of the impact of customers' personality and perceptions on Internet banking use. *J. Comput. Hum. Behav.* **2013**, *29*, 1113–1141. [CrossRef]

61. Aldás-Manzano, J.; Ruiz-Mafé, C.; Sanz-Blas, S. Exploring individual personality factors as drivers of M-shopping acceptance. *Ind. Manag. Data Syst.* **2009**, *109*, 739–757. [CrossRef]

62. Özbeka, V.; Alnıaçık, Ü.; Kocc, F.; Akkılıç, E.; Kaş, E. The Impact of Personality on Technology Acceptance: A Study on Smart Phone Users. *Procedia Soc. Behav. Sci.* **2014**, *150*, 541–551. [CrossRef]

63. Childers, T.L.; Carr, C.L.; Peck, J.; Carson, S. Hedonic and utilitarian motivations for online retail shopping behavior. *J. Retail.* **2001**, *77*, 511–535. [CrossRef]

64. Marjan, M.; Mohammad, R.E.; Ali, S.B. Influencing VSN users' purchase intentions: The roles of flow, trust and eWOM. *J. Res. Interact. Mark.* **2014**, *8*, 102–123. [CrossRef]

65. Thorbjørnsen, H.; Pedersen, P.E.; Nysveen, H. This is who I am: Identity expressiveness and the theory of planned behavior. *Psychol. Mark.* **2007**, *24*, 763–785. [CrossRef]

66. Bosnjak, M.; Galesic, M.; Tuten, T. Personality determinants of online shopping: Explaining online purchase intentions using a hierarchical approach. *J. Bus. Res.* **2007**, *60*, 597–605. [CrossRef]

67. Mowen, J.C. *The 3M Model of Motivation and Personality*; Kluwer Academic Press: Norwell, MA, USA, 2000.

68. Islam, J.U.; Rahman, Z.; Hollebeek, L.D. Personality factors as predictors of online consumer engagement: An empirical investigation. *Mark. Intell. Plan.* **2017**, *35*, 510–528. [CrossRef]

69. Ramayah, T.; Lo, M.C. Impact of shared beliefs on "perceived usefulness" and "ease of use" in the implementation of an enterprise resource planning system. *Manag. Res. News* **2007**, *30*, 420–431. [CrossRef]

70. Hayhoe, C.R.; Leach, L.J.; Turner, P.R.; Bruin, M.J.; Lawrence, F.C. Differences in Spending Habits and Credit Use of College Students. *J. Consum. Aff.* **2000**, *34*, 113–133. [CrossRef]

71. National Statistics Republic of China. Available online: https://eng.stat.gov.tw (accessed on 15 November 2017).

72. Blumenthal, S. Estimating population size with truncated sampling. *Commun. Stat. Theory Methods* **1977**, *6*, 297–308. [CrossRef]

73. Cochran, W.G. *Sampling Techniques*, 3rd ed.; Wiley: New York, NY, USA, 1977.

74. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2009.
75. Vogt, W.P.; Johnson, R.B. *The SAGE Dictionary of Statistics and Methodology a Nontechnical Guide for the Social Sciences*, 5th ed.; Sage Publication: London, UK, 2015.
76. Coakes, S.J.; Steed, L.; Ong, C. *SPSS 16.0 for Windows: Analysis without Anguish*; John Wiley and Sons: Milton, QLD, Australia, 2009.
77. Chen, J.K.C.; Batchuluun, A.; Batnasan, J. Services innovation impact to customer satisfaction and customer value enhancement in airport. *Technol. Soc.* **2015**, *43*, 219–230. [CrossRef]
78. Bentler, P.M.; Bonnet, D.G. Significance Tests and Goodness of Fit in the Analysis of Covariance Structures. *Psychol. Bull.* **1980**, *88*, 588–606. [CrossRef]
79. Bentler, P.M. Comparative Fit Indexes in Structural Models. *Psychol. Bull.* **1990**, *107*, 238–246. [CrossRef] [PubMed]
80. MacCallum, R.C.; Browne, M.W.; Sugawara, H.M. Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychol. Methods* **1996**, *1*, 130–149. [CrossRef]
81. Anderson, J.C.; Gerbing, D.W. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychol. Bull.* **1988**, *103*, 411–423. [CrossRef]
82. Fornell, C.; Larcker, D.F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *J. Mark. Res.* **1981**, *18*, 39–50. [CrossRef]
83. Tong, X. A cross-national investigation of an extended technology acceptance model in the online shopping context. *Int. J. Retail Distrib. Manag.* **2010**, *38*, 742–759. [CrossRef]
84. Arnold, M.J.; Reynolds, K.E. Hedonic shopping motivation. *J. Retail.* **2003**, *79*, 77–95. [CrossRef]
85. Wong, W.K.; Chow, S.C.; Hon, T.Y.; Woo, K.Y. Empirical Study on Conservative and Representative Heuristics of Hong Kong Small Investors Adopting Momentum and Contrarian Trading Strategies. *Int. J. Rev. Manag.* **2018**, forthcoming.

Do Sustainable Stocks Offer Diversification Benefits for Conventional Portfolios? An Empirical Analysis of Risk Spillovers and Dynamic Correlations

Mehmet Balcilar [1,2,3], **Riza Demirer** [4] **and Rangan Gupta** [5,6,*]

1   Department of Economics, Eastern Mediterranean University, Turkish Republic of Northern Cyprus, via Mersin 10, Famagusta, Turkey; mehmet@mbalcilar.net
2   Department of Economics, University of Pretoria, Pretoria 0002, South Africa
3   2300 Avenue des Moulins, 34080 Montpellier, France
4   Department of Economics & Finance, Southern Illinois University Edwardsville, Edwardsville, IL 62026-1102, USA; rdemire@siue.edu
5   Department of Economics, University of Pretoria, Pretoria 0002, South Africa
6   IPAG Business School, 184 Boulevard Saint-Germain, 75006 Paris, France
*   Correspondence: rangan.gupta@up.ac.za

Abstract: This paper explores the potential diversification benefits of socially responsible investments for conventional stock portfolios by examining the risk spillovers and dynamic correlations between conventional and sustainability stock indexes from a number of regions. We observe significant unidirectional volatility transmissions from conventional to sustainable equities, suggesting that the criteria applied for socially responsible investments do not necessarily shield these securities from common market shocks. While significant dynamic correlations are observed between sustainable and conventional stocks, particularly in Europe, the analysis of both in- and out-of-sample dynamic portfolios suggests that supplementing conventional stock portfolios with sustainable counterparts improves the risk/return profile of stock portfolios in all regions. The findings overall suggest that sustainable investments can indeed provide diversification gains for conventional stock portfolios globally.

Keywords: socially responsible investment; multivariate regime-switching; time-varying correlations; volatility transmission

JEL Classification: C32; G11; G12

1. Introduction

In the wake of the recent global financial crisis, enormous negative impacts have been felt by conventional institutions and markets. Understandably, a need has been felt for exploring alternatives to conventional financial practices in order to reduce investment risks, increase returns, enhance financial stability, and reassure investors and financial markets. In this regard, academic research on socially responsible investing (SRI), though originally initiated by religious groups like Quakers and Methodists around the eighteenth century [1], has intensified, as has received attention in popular media (http://www.ussif.org/). One reason for the increased interest in SRI investments is that they combine the pursuit of financial returns with non-financial considerations relating to the environment, social issues, and governance (ESG), and hence, are perceived to be less risky compared to conventional alternatives.

As will be seen from the literature review segment below, research on SRI has primarily focused on the risk–return characteristics of these securities in relation to conventional investments. A missing area

of research in this regard is whether these securities offer diversification opportunities for conventional investments, based on a formal portfolio allocation exercise. Against this backdrop, our study is the first to address the issue of diversification (or risk hedging) between SRI and conventional investments by considering the regime-switching and volatility interactions between these two types of assets for the entire world economy and a number of regions including North America, Europe and Asia-Pacific. It must be noted that North America and Europe are the largest regions in terms of SRI assets, accounting for 99 percent of the global share for sustainable investing assets [2]. To that end, by examining the risk spillovers and dynamic correlations across SRI investments and conventional assets from different regions, this study provides a comparative analysis of the interaction of these assets with conventional markets, thus enlarging our understanding of whether or not socially responsible investing can indeed benefit investors financially.

In addition to the analysis of dynamic interactions across conventional and SRI assets, we also derive dynamic hedging strategies by adopting a Markov regime-switching Generalized Autoregressive Conditional Heteroskedasticity GARCH model with dynamic conditional correlations (MS-DCC-GARCH). This model allows us to capture both the time-variation in conditional volatility of the markets under consideration according to different regimes and their dynamic links (correlations). By utilizing a time-varying regime-switching specification, we not only account for the well-established nonlinearity that exists in financial markets, but also examine the possibility that SRI significantly reduces the downside risk [3]. Our spillover tests yield significant unidirectional volatility transmissions from conventional to sustainable equities, suggesting that the criteria applied for socially responsible investments do not necessarily shield these securities from common market shocks. While the results from the MS-DCC-GARCH model indicates significant time variation in the dynamic correlations between conventional and sustainable equities, particularly in Europe, the analysis of both in- and out-of-sample portfolios suggests that supplementing conventional stock portfolios with sustainable counterparts improves the risk/return profile of stock portfolios in all regions. Improvement in risk adjusted returns is particularly striking for the broader world index and the Asia-Pacific region when the negative risk adjusted returns for undiversified, conventional portfolios turn around to positive values when the conventional index is supplemented by the sustainable counterpart. However, our portfolio analysis also suggests that these diversification gains can only be achieved by implementing an investment strategy that aims to minimize portfolio risk and utilize sustainable assets in the short leg of the portfolio. The findings overall provide useful guidance for the implementation of effective SRI risk management and for policy regulations. A significant finding of this study is that socially responsible investment does not result in lower risk-adjusted portfolio returns when information on market regimes and dynamic investing strategies are used. This finding is important since it implies that individual investors and fund managers can pursue socially responsible investments without sacrificing returns.

The remainder of the study is organized as follows: Section 2 summarizes the relevant literature and Section 3 presents the MS-DCC-GARCH model used in the analysis. Section 4 describes the data and presents the estimation results, volatility spillover tests and dynamic correlation analysis. Section 5 provides the in- and out-of-sample portfolio performance comparisons and Section 6 concludes the paper.

2. Literature Review

In his pioneering works [4,5], Markowitz lay the foundation for the efficient diversification of investment portfolios and how spreading out a portfolio's holdings across various assets can improve the risk/return profile for investors. In applications of this concept to socially responsible investments, a number of studies including [1,3,6–9] claim that non-financial elements provide SRI investors with extra utility or satisfaction. In addition, as pointed out by [1,9–12], SRI investors tend to believe that ESG factors materially affect the returns in a positive way, which, in turn, can lead to lower costs involved in the avoidance or minimization of environmental and reputational risks, and better

management and better customer satisfaction that eventually impacts revenues in a positive way. Possibly, these are the reasons that have led the global SRI (sustainable investment) market to grow steadily both in absolute and relative terms. According to the Global Sustainable Investment Review of 2014 [2], released by the Global Sustainable Investment Association (GSIA), SRI has risen from $13.3 trillion at the outset of 2012 to $21.4 trillion at the start of 2014, which corresponds to an increase from 21.5 percent to 30.2 percent of the professionally managed assets in Europe, the United States, Canada, Asia, Japan, Australasia and Africa.

With support for SRI expanding since the 1960s due to the rise of the civil rights movement, environmentalism and concerns about globalization [1], formal research in this area is not new, and can be associated first with [13]. There are now a number of studies on SRI which have investigated the following aspects, primarily through the lens of mutual funds, but also through regional SRI indexes for not only the US, but also Europe and other major developed economies. (a) Performance (i.e., risk–return characteristics relative to conventional indexes), using mutual funds and broad market indexes [11,12,14–38] and at firm-level [3,34,39–45]. These studies, however, fail to provide clear-cut empirical evidence on whether SRI does yield higher returns after adjusting for risks. Similarly, studies on (b) ratings [46–48], and (c) screenings [49] in terms of sustainability, do not seem to provide clear cut evidence in terms of higher returns either. Studies of (d) predictability and determinants of returns and volatility [50,51], highlight the role of various forms of uncertainties related to economic policies; and (e) co-movements of SRI indexes and with conventional indexes across various regions [1,52] have been shown to exist, especially when nonlinearity is taken into account.

As can be seen from the above discussion, research on SRI has primarily focused on the risk–return characteristics of these securities in relation to conventional investments. A missing area of research in this regard is whether these securities offer diversification opportunities for conventional investments, based on a formal portfolio allocation exercise. Some tangential discussion regarding diversification is available in [52], where cointegration analysis is performed for the US between the Dow Jones Sustainability Index and the Dow Jones Industrial Average Index. The authors show that while there is no evidence of linear cointegration due to nonlinearity and regime changes, cointegration can be detected using a quantile-regression based approach. This paper then goes on to suggest that this result implies that there are no long-run diversification opportunities in the US between SRI and conventional investments. However, no formal portfolio allocation exercise is performed by [52], which is what we aim to address in this paper based on a MS-DCC-GARCH model, i.e., a variant of the original DCC-GARCH model of [53], with Markov-switching (as detailed in [54]). Note that these types of models have also been widely used in analyzing hedges and safe-haven properties of various assets (see [55] for a detailed discussion in this regard) and also comparing Islamic and conventional equities (see for example, [56] for further details), with the latter being somewhat related to our analysis, given the importance of Sharia rules imposed on screening the equities included in Islamic indices.

3. Methodology

The dynamic conditional correlation (DCC) model used in the study follows [57–59] and more recently [60]. Let $R_t = [R_{s,t}, R_{c,t}]'$ be the (2×1) vector of returns where $R_{s,t}$ and $R_{c,t}$ are the return on SRI represented by a sustainability index and the return on conventional investment represented by a conventional market index, respectively. The model is constructed in a bivariate fashion with pairs of SRI and conventional investment returns for the entire world economy and a number of regional indexes representing North America, Europe, and Asia-Pacific. The GARCH specification for the volatility spillover model follows [61] and is specified as

$$R_t = \Phi_0 + \sum_{i=1}^{p} \Phi_i R_{t-i} + \varepsilon_t$$
$$\varepsilon_t = D_t z_t$$

(1)

where $D_t = \text{diag}(h_{s,t}^{1/2}, h_{c,t}^{1/2})$ is the vector of the conditional volatility terms. The conditional mean of the return vector R_t is specified as a vector autoregressive (VAR) process of order p with (2×2) parameter matrices Φ_i, $i = 1, 2, ..., p$. The unexplained component ε_t follows a GARCH specification described as $\varepsilon_t|\psi_{t-1} \sim ID(0, P_t)$ where P_t is the time-varying variance–covariance matrix. Denoting the conditional variance matrix as $H_t = [h_{s,t}, h_{c,t}]'$, we impose the following specification which allows for volatility spillover in the model

$$H_t = c + A\varepsilon_{t-1}^{(2)} + BH_{t-1} \qquad (2)$$

where c is a (2×1) vector of constants, A and B are (2×2) matrices for the ARCH and GARCH effects and $\varepsilon_t^{(2)} = [\varepsilon_{s,t}^2, \varepsilon_{c,t}^2]'$. Note that the non-diagonal forms of the matrices A and B allow volatility spillovers across the series. Following [52], we allow conditional correlations to vary over time by specifying the variance–covariance matrix as $P_t = D_t\Gamma_t D_t$ where Γ_t is the conditional correlation matrix.

A distinct feature of the model is that the conditional correlation matrix, Γ_t, is characterized by regime-switching as governed by a discrete Markov process and is defined as $\Gamma_t = \text{diag}\{Q_t\}^{-1/2}Q_t\text{diag}\{Q_t\}^{-1/2}$. In order to incorporate regime shifts into the DCC model shown in Equations (1) and (2), we follow [57] and introduce a Markov regime-switching dynamic correlation model by specifying Q_t as

$$Q_t = [1 - \alpha(s_t) - \beta(s_t)]\overline{Q} + \alpha(s_t)\varepsilon_{t-1}^{(2)} + \beta(s_t)Q_{t-1} \qquad (3)$$

where $\overline{Q}$ is the unconditional covariance matrix of the standardized residuals. In Equation (3), $\alpha(s_t)$ and $\beta(s_t)$ are the regime-dependent parameters that control the regime-switching system dynamics where $s_t \in \{1, 2\}$ is the state or regime variable following a first-order, two-state discrete Markov process. Note that the variances in this specification are regime-independent whereas the covariances (or correlations) are both time-varying and regime-switching (We estimate the MS-DCC-GARCH model using the two-step approach of [53,62]. In the second step, we use the modified Hamilton filter proposed by [57] to solve the path-dependence problem [63–65] and estimate the regime-switching conditional covariances accordingly). As [57] note, the specification in which all parameters are regime dependent is highly unstable due to the large number of switching parameters. Therefore, we restrict the regime dependent structure to the time-varying correlations only. Thus, the model allows both volatility spillovers and regime-switching dynamic correlations. The specification is then completed by defining the transition probabilities of the Markov process as $p_{ij} = P(s_{t+1} = i|s_t = j)$ where p_{ij} is the probability of being in regime i at time $t + 1$ given that the market was in regime j at time t with regimes i and j taking values in {1, 2}. Finally, the transition probabilities satisfy $\sum_{i=1}^{2} p_{ij} = 1$.

The MS-DCC-GARCH model we specified above has several advantages over the standard DCC-GARCH model. Caporin and McAleer [66] lists and explains ten limitations of the standard DCC-GARCH model. Most of these are technical and the extent of their significance are not well known. How important the technical issues are, usually depends on the complexity of the specification and how far the data is from the assumptions. Two of the limitations, however, might have series consequences for the portfolio analysis. First, as pointed out by [66], the dynamic conditional correlations of the standard DCC-GARCH model are specified for the standardized residuals and, indeed, the standard DCC-GARCH model does not yield dynamic conditional correlations. Second, the standard DCC-GARCH model is not dynamic empirically, because the effect of news in this model is inherently extremely small. Additionally, the standard DCC-GARCH model is a single regime model and completely ignores the typical regime-switching behavior of the financial markets. The MS-DCC-GARCH model used in this study does not have these three limitations of the standard DCC-GARCH models. The MS-DCC-GARCH model has dynamic conditional time varying correlations, is asymmetric in its treatment of the conditional variance matrix and, therefore, is inherently dynamic.

4. Empirical Findings

4.1. Data

In our empirical analysis, we use daily data for Dow Jones sustainability and conventional indices obtained from Datastream. The conventional indices include the Dow Jones global indices for the World (GLOBAL), North America (AMRCS), Europe (EUROPE) and Asia-Pacific (ASPCF). Similarly, the corresponding Dow Jones sustainability indices for the above-mentioned regions are denoted by SIWORLD, SINAMR, SIEUROPE, and SIASPCF, respectively. The sample period is from 1 January 2004 to 2 September 2015, including 3044 observations. Table 1 presents the descriptive statistics for logarithmic returns.

Table 1. Descriptive statistics for returns (%).

	SIWOLRD	SINAMRC	SIEUROPE	SIASPCF	GLOBAL	AMRCS	EUROPE	ASPCF
Mean	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01
S.D.	1.14	1.14	1.45	1.32	1.10	1.22	1.42	1.16
Min	−7.77	−8.99	−9.93	−10.33	−7.89	−9.74	−10.13	−9.11
Max	8.84	9.45	10.46	10.84	9.88	10.51	10.51	9.01
Skewness	−0.30	−0.42	−0.09	−0.34	−0.44	−0.48	−0.12	−0.47
Kurtosis	8.95	11.30	7.59	6.35	9.89	11.40	7.58	6.72
JB	10,227.42 ***	16,313.24 ***	7318.62 ***	5190.91 ***	12,516.18 ***	16,620.02 ***	7301.17 ***	5855.89 ***
$Q(1)$	34.07 ***	16.53 ***	1.14	0.96	85.75 ***	9.92 ***	0.10	1.32
$Q(5)$	57.59 ***	31.22 ***	29.83 ***	7.18	98.61 ***	19.57 ***	26.79 ***	4.34
ARCH(1)	120.66 ***	202.02 ***	113.33 ***	94.02 ***	152.65 ***	140.80 ***	100.00 ***	82.70 ***
ARCH(5)	798.47 ***	797.38 ***	593.67 ***	899.63 ***	915.00 ***	789.89 ***	611.35 ***	837.62 ***
n	3044	3044	3044	3044	3044	3044	3044	3044

	Pearson Correlation Coefficient Estimates			
	World	**Americas**	**Europe**	**Asia-Pacific**
Full sample	0.966	0.987	0.995	0.976
Subprime Crises Period	0.967	0.992	0.996	0.985

Note: This table gives the descriptive statistics for logarithmic returns. SIWORLD, SINAMR, SIEUROPE, and SIASPCF denote Dow Jones Sustainability Indices (DJSI) for the World, North America, Europe, and Asia-Pacific, respectively, while GLOBAL, AMRCS, EUROPE, and ASPCF denote Dow Jones conventional Global Indices (DJGI) for the World, Americas, Europe and Asia-Pacific. The daily data covers the period 1 January 2004 to 2 September 2015 with $n = 3044$ observations. In addition to the mean, the standard deviation (S.D.), minimum (min), maximum (max), skewness, and kurtosis statistics, the table reports the Jarque–Bera normality test (JB), the Ljung–Box first ($Q(1)$), the fourth ($Q(5)$) autocorrelation tests, and the first (ARCH(1)) and the fourth (ARCH(5)) order Lagrange multiplier (LM) tests for the autoregressive conditional heteroscedasticity (ARCH), and Pearson correlations coefficient estimates. Full sample and subprime mortgage crises period (December 2007–June 2009) Pearson correlation coefficients are reported for World, Americas, Europe, and ASIA-Pacific, which represented the sustainability and conventional index pairs, (SIWORLD GLOBAL), (SINAMRC AMRCS), (SIEUROPE UROPE), and (SIASPCF ASPCF), respectively. The asterisks ***, ** and * represent significance at the 1%, 5%, and 10% levels, respectively.

Despite similar values for mean returns, we generally observe higher return volatility for the sustainability indices compared to their conventional counterparts. It can be argued that the economic, environmental and social criteria applied in the selection of firms to be included in these indices limit the potential to mitigate idiosyncratic risks in these portfolios, thus leading to higher return volatility compared to broader based conventional indices. On the other hand, all return series exhibit negative skewness, implying greater likelihood of experiencing losses. Similarly, all return series have kurtosis values higher than the normal distribution, implying the presence of extreme movements. It is possible that the inclusion of the global financial crisis (GFC) in the sample period drives the patterns observed in higher order moments. The impact of the GFC is evident in the time series plots presented in Figure 1. Both conventional and sustainable stock indices sustained significant losses during the 2007/2008 crisis period and then again during early 2012 at the height of the Eurozone crisis.
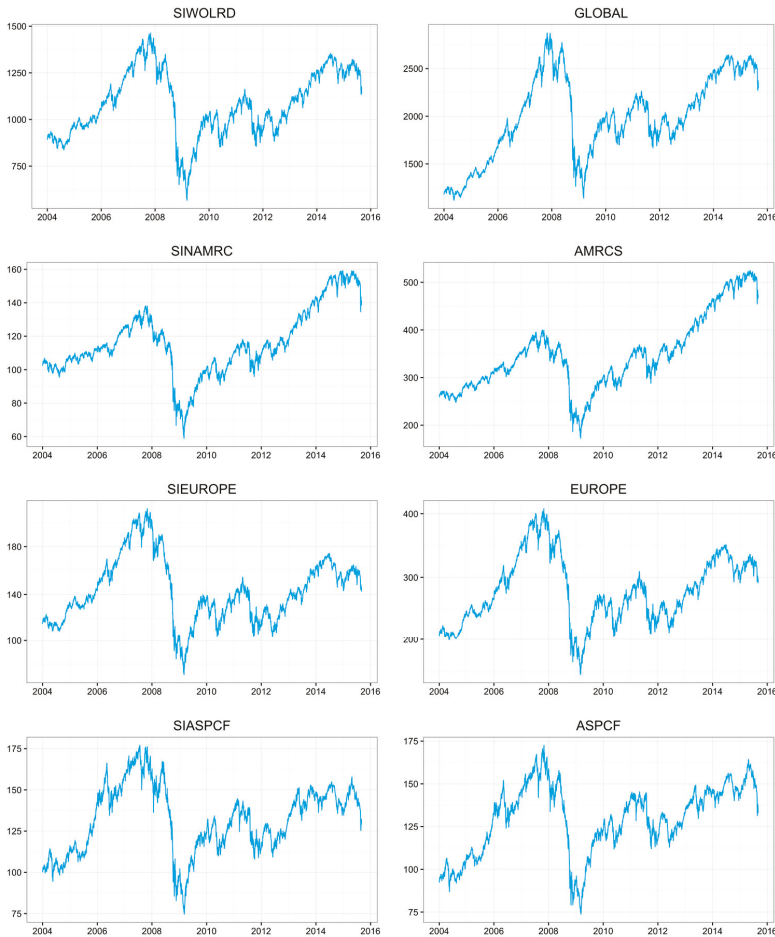
Figure 1. Time-series plots of conventional and sustainability indexes. **Note:** This figure provides the plots of the daily levels of the conventional and sustainability indices for the period 1 January 2004 to 2 September 2015. SIWORLD (GLOBAL), SINAMR (AMRCS), SIEUROPE (EUROPE), and SIASPCF (ASPCF) denote Dow Jones Sustainability (Conventional Global) Indices for the World, North America, Europe, and Asia-Pacific, respectively.

Table 1 also reports the Pearson correlation coefficient estimates for the pairs of sustainability and conventional indices for each of the four regions, i.e., World, North America, Europe, and Asia-Pacific. The correlations coefficients are reported both for the full sample and the subprime mortgage crises period (December 2007–June 2009) for comparison purposes. Estimates of the correlation coefficients for all regions, both in the full sample and subprime mortgage crises period, are found to be above 96%, suggesting a high degree of co-movement across sustainable and conventional investment returns. While we observe the highest correlation estimates in the case of Europe, we see that correlations do not exhibit a significantly different pattern during the subprime mortgage crises period.

4.2. Model Identification

The MS-DCC-GARCH model requires prior identification of the VAR order p in Equation (1) and univariate GARCH models that are used to obtain conditional volatility estimates in Equations (2) and (3). We first identified the univariate GARCH models using the Akaike information criterion (AIC) to fit the GARCH(1,1) models with a conditional mean that is specified as an autoregressive process of order p, AR(p), leading to a AR(p)-GARCH(1,1) model. We selected the AR order p using the AIC. In order check for possible misspecifications, we performed conditional heteroskedasticity and serial correlation diagnostics. The Lagrange multiplier (LM) test was used for conditional heteroskedasticity diagnosis, while the Ljung–Box portmanteau test (Q) was used for the serial correlation diagnostic.

Table 2 reports the diagnostics for the univariate AR(p)-GARCH(1,1) model and also presents the selected AR orders p where the maximum p was set equal to 10. The selected AR orders vary from 0 to 5 and Ljung–Box tests with the orders 10 and 20 show that the selected orders were sufficient to capture serial correlations in the series. The LM tests do not reject the null of no first order ARCH effects even at the 10% level, except SINAMRC, for which non-rejection occurred only at the 1% level. Given the results in Table 2, we decided that a GARCH(1,1) specification with the AR orders selected by the AIC sufficiently models the conditional heteroskedasticity in all series. In order to select the VAR orders in Equation (1), we used the Bayesian information criterion (BIC) with a maximum order equal to 10. The BIC selected an order of one for all four VAR specifications for the four regions. Finally, the MS-DCC-GARCH models were estimated using the maximum likelihood (ML) method based on these specifications.

Table 2. Univariate AR(p)-GARCH(1,1) fit diagnostics.

	ARCH-LM(1)	JB	Q(10)	Q(20)	p
SIWOLRD	2.724 (0.010)	197.383 *** (<0.001)	5.454 (0.793)	19.214 (0.443)	4
SINAMRC	5.277 ** (0.022)	427.863*** (<0.001)	6.409 (0.698)	17.684 (0.544)	2
SIEUROPE	0.122 (0.727)	230.330 *** (<0.001)	4.802 (0.851)	16.135 (0.648)	0
SIASPCF	0.001 (0.980)	92.166 *** (<0.001)	4.638 (0.865)	11.789 (0.895)	4
GLOBAL	2.160 (0.142)	244.572 *** (<0.001)	4.439 (0.880)	19.661 (0.415)	4
AMRCS	5.790 (0.016)	436.193 *** (<0.001)	7.645 (0.570)	18.204 (0.509)	2
EUROPE	0.294 (0.588)	220.475 *** (<0.001)	4.543 (0.872)	16.985 (0.591)	0
ASPCF	0.311 (0.577)	160.252 *** (<0.001)	7.192 (0.617)	16.732 (0.608)	4

Note: The table reports diagnostic tests for univariate autoregressive GARCH model fits. An AR(p)-GARCH(1,1) model was fitted to each series. The AR order p was selected by the Akaike information criterion (AIC). Table reports the Jarque–Bera normality test (JB), the Ljung-Box 10th ($Q(10)$) and the 20th ($Q(20)$) autocorrelation tests, and the first (ARCH(1)) order Lagrange multiplier (LM) tests for the autoregressive conditional heteroscedasticity (ARCH). The p-values of the tests are given in parentheses. The asterisks ***, ** and * represent significance at the 1%, 5%, and 10% levels, respectively. The symbol ">" signifies "less than" the number it precedes.

4.3. Volatility Spillover Tests

Table 3 presents the parameter estimates for the MS-DCC-GARCH model described in Equations (1)–(3). As explained earlier, the model is structured to allow for possible bidirectional volatility spillovers across the sustainable and conventional market segments for each global and regional index examined. We observe in Panel A generally insignificant shock spillovers across

the sustainable and conventional markets, indicated by insignificant a_{ij} $(i \neq j)$ estimates for all regional indexes. On the other hand, significant and positive volatility spillovers are observed from conventional to sustainable indices, implied by highly significant b_{12} estimates consistently for each region. This finding suggests that uncertainty regarding global equity markets spills over to the market for sustainable stocks, driving return volatility in this market segment. Risk transmissions, however, are found to be unidirectional, implied by insignificant spillover effects from sustainable to conventional indexes. It can thus be argued that sustainable stocks do not necessarily exhibit segmentation from their conventional counterparts and are driven by the common fundamental uncertainties affecting equity markets globally. The findings also suggest that the criteria applied in the identification of socially responsible investments do not necessarily shield these stocks from equity market shocks.

Table 3. Estimates of the MS-DCC-GARCH model.

Parameters	Models			
	World	Americas	Europe	Asia-Pasiific
Panel A: Spillover parameters				
c_s	0.0033 (0.0139)	0.0159 (0.0099)	0.0052 (0.0371)	0.0319 (0.0289)
c_c	0.0159 (0.0305)	0.0176 (0.0378)	0.0202 (0.0672)	0.0168 (0.0317)
$a_{s,s}$	0.0358 (0.0304)	0.0162 (0.0303)	0.0189 (0.0821)	0.0173 (0.0562)
$a_{s,c}$	0.0633 (0.7564)	0.0249 (1.9827)	0.0082 (2.8751)	0.0110 (2.8601)
$a_{c,s}$	0.0252 (0.8179)	0.0582 (1.7571)	0.0593 (2.9988)	0.1014 (3.7029)
$a_{c,c}$	0.0337 *** (0.0106)	0.0786 *** (0.0092)	0.0772 ** (0.0365)	0.0945 *** (0.0192)
$b_{s,s}$	0.1266 *** (0.0265)	0.1496 *** (0.0388)	0.2183 *** (0.0590)	0.1549 *** (0.0229)
$b_{s,c}$	0.4627 *** (0.0253)	0.6228 *** (0.0306)	0.4593 *** (0.0721)	0.3965 *** (0.0526)
$b_{c,s}$	0.8503 (0.6626)	0.6612 (2.3041)	0.7239 (2.9214)	0.9425 (2.1682)
$b_{c,c}$	0.4003 (0.7189)	0.3353 (2.0511)	0.4324 (3.0450)	0.3686 (2.8084)
Panel B: DCC parameters				
$\alpha(s_t = 1)$	0.0181 *** (0.0036)	0.0427 *** (0.0040)	0.0880 *** (0.0054)	0.0361 *** (0.0060)
$\beta(s_t = 1)$	0.9750 *** (0.0063)	0.9430 *** (0.0058)	0.8528 *** (0.0102)	0.9553 *** (0.0147)
$\alpha(s_t = 2)$	0.0677 *** (0.0250)	0.0839 *** (0.0108)	0.1073 *** (0.0301)	0.0778 * (0.0444)
$\beta(s_t = 2)$	0.7769 *** (0.0999)	0.8730 *** (0.0172)	0.8095 *** (0.0602)	0.8314 *** (0.1668)
Panel C: Regime Inference				
log *L* of MS-DCC	−4029.247	−3084.625	−2785.198	−4495.000
log *L* of DCC	−5103.762	−4034.694	−3898.901	−5275.907
AIC of MS-DCC	2.661	2.041	1.844	2.968
AIC of DCC	3.360	2.658	2.569	3.474
LR linearity Test	2149.030 ***	1900.138 ***	2227.405 ***	1561.813 ***
p_{11}	0.982	0.984	0.969	0.979
p_{22}	0.935	0.930	0.889	0.930
n_1	2387.100	2478.600	2377.700	2352.100
n_2	655.900	564.400	665.300	690.900
Prob(Regime 1)	0.781	0.812	0.781	0.770
Prob(Regime 2)	0.219	0.188	0.219	0.230
Duration of Regime 1	55.140	61.670	32.080	47.740
Duration of Regime 2	15.420	14.320	8.980	14.220

Note: This table reports the estimates of the MS-DCC-GARCH model given in Equations (1)–(3). The matrix *R* for the World, Americas, Europe, and Asia-Pacific models are formed as R = (SIWORLD GLOBAL), R = (SINAMRC AMRCS), R = (SIEUROPE UROPE), and R = (SIASPCF ASPCF), respectively. The GARCH part of the model is specified as a GARCH(1,1). The subscript *s* denotes the SRI return series while subscript *c* denotes conventional return series. The models are estimates over the full sample period 1 January 2004–2 September 2015 with n = 3044 observations. The lag order for the Vector Autoregressive VAR part of the model was selected by the AIC and is one for all four models. The MS-DCC-GARCH model was estimated using the maximum likelihood (ML) method. The likelihood ratio (LR) linearity test is reported with *p*-value of the [67]. Standard errors of the estimates are given in parentheses. log *L* stands for the log likelihood, p_{ii} for the regime transition probabilities, *Prob* (Regime *i*) for the ergodic (limit) probability of regime *i*, and n_i for the number of observations falling in regime *i* according to the ergodic probability. ***, ** and * represent significance at the 1%, 5%, and 10% levels, respectively.

Examining the volatility persistence coefficients measured by $(a_{ii} + b_{ii})$, we generally observe moderate to weak volatility persistence, relatively weaker in the case of sustainable indexes. The volatility persistence coefficients for the conventional (sustainable) indices are estimated as 0.433 (0.162), 0.413 (0.165), 0.509 (0.237), and 0.463 (0.172) for the World, Americas, Europe, and Asia-Pacific regions, respectively. Considering positive own volatility shocks observed in the case of sustainable indexes, implied by highly significant b_{11} estimates, it can be argued that historical information on return and volatility in sustainable equity markets could be utilized in forecasting future volatility despite the evidence of weak volatility persistence in these markets.

Formal tests of causality in volatility between the conventional and sustainable stock markets are presented in Table 4. Four alternative spillover tests are utilized to test the null hypothesis of no unidirectional volatility spillover from market X to market Y ($X \Rightarrow Y$) and no bidirectional spillover between markets X and Y ($X \Leftrightarrow Y$). The first test is a Wald test involving two zero restrictions on the relevant parameters in matrices A and B in Equation (2). The next two tests are the LM-based robust (NT-R) and non-robust (NT-NR) tests of causality in conditional variance proposed by [68]. Finally, the fourth test (HH) is the Hafner-Herwartz [69] LM test of causality on conditional variance.

Table 4. Volatility spillover tests.

Test Type	Wald	NT-R	NT-NR	HH
Panel A: Unidirectional volatility spillovers from conventional to sustainable				
H_0 : GLOBAL $\Rightarrow$ SIWORLD	26.0335 ***	33.5088 ***	9.9801 ***	7.9773 **
H_0 : AMRCS $\Rightarrow$ SINMARC	3.9563	1.7234	5.2534	2.3548
H_0 : EUROPE $\Rightarrow$ SIEUROPE	6.9236	3.9381	5.1069	9.6125 **
H_0 : ASPCF $\Rightarrow$ SIASPCF	7.2269	5.2085	4.9439	7.6233
Panel B: Unidirectional volatility spillovers from sustainable to conventional				
H_0 : SIWORLD $\Rightarrow$ GLOBAL	5.7180 *	9.9768 ***	13.6846 ***	3.0029
H_0 : SINMARC $\Rightarrow$ AMRCS	1.8908	1.1614	5.285	2.1588
H_0 : SIEUROPE $\Rightarrow$ EUROPE	4.4583	2.2569	3.0597	0.338
H_0 : SIASPCF $\Rightarrow$ ASPCF	4.7144	3.1005	2.1225	0.3585
Panel C: Bi-directional volatility spillovers between sustainable and conventional				
H_0 : GLOBAL $\Leftrightarrow$ SIWORLD	19.4387 ***	0.2948	32.5822 ***	42.6304 ***
H_0 : AMRCS $\Leftrightarrow$ SINMARC	5.847	2.8848	10.5384 **	4.5136
H_0 : EUROPE $\Leftrightarrow$ SIEUROPE	11.3819 **	6.195	8.1666 *	9.9505 **
H_0 : ASPCF $\Leftrightarrow$ SIASPCF	11.9413 **	8.3090 *	7.0664	7.9818 *

Note: The table reports causality tests for testing the null hypothesis of no one unidirectional volatility spillover from variable X to variable Y, demoted, $X \Rightarrow Y$ as well as the bidirectional volatility spillover, denoted $X \Leftrightarrow Y$. The Wald tests for testing the no volatility spillover restrictions were imposed on Equation (1). The tests report that the tests are distributed as Chi-square with 2 and 4 degrees of freedom, respectively, for unidirectional and bidirectional tests. The HH test is the [69] LM test of causality on conditional variance. NT-R is the [68] robust test of the causality in conditional variance, while the NT-NR is the non-robust version of the [68] test. HH, NT-R, and NT-NR tests are LM tests and the univariate specification for conditional variances is a GARCH(1,1) model. We compute HH, NT-R, and NT-NR tests to tests only causality in conditional variance from X variable (Japan or US) to Y variable. ***, ** and * represent significance at the 1%, 5%, and 10% levels, respectively.

Examining the unidirectional spillover tests from the conventional to sustainable indices reported in Panel A, we find that all four tests consistently reject no causality in variance in the case of the broader world index, further supporting prior evidence of significant volatility spillovers from conventional to sustainable stocks. Although not as consistently significant as in the conventional-to-sustainable case, some evidence of volatility spillover in the opposite direction is also found for the world index in Panel B, supported particularly by the causality tests of [68]. On the other hand, the formal unidirectional tests for the other regions reported in Panels A and B did not generally yield evidence of risk transmissions in either direction for regional indices. The tests for bidirectional spillover effects reported in Panel C further support prior findings for the world index, indicating bi-directional risk transmissions across the sustainable and conventional stock indices. On the other hand, we observe largely inconsistent test

results for regional indices, consistent with the findings in Panels A and B. Overall, the format tests clearly indicate significant risk transmissions from conventional to sustainable stocks in the case of the world index while somewhat weaker evidence of volatility spillover in the opposite direction is also observed.

4.4. Dynamic Correlations

The regime-switching specification that governs the data is tested against the static alternative using a battery of specification tests including the likelihood ratio (LR) linearity test with a *p*-value of [64], further supported the Akaike (AIC) information criteria. Both formal tests and the information criteria reported in Panel C of Table 3 consistently favor a two-regime MS-DCC-GARCH specification over the static DCC-GARCH alternative, indicating strong support for the presence of two distinct market regimes. The smoothed probability plots for the first regime reported in Figure 2 indicate that the first regime largely corresponds to normal market periods with the smoothed probabilities for this regime dropping to near zero values during the GFC period, as well as the late-2011 and early-2012 periods when the Eurozone uncertainty hit its peak. Therefore, we conclude that the first regime characterizes normal (or low) volatility periods while the second regime is the high volatility regime.
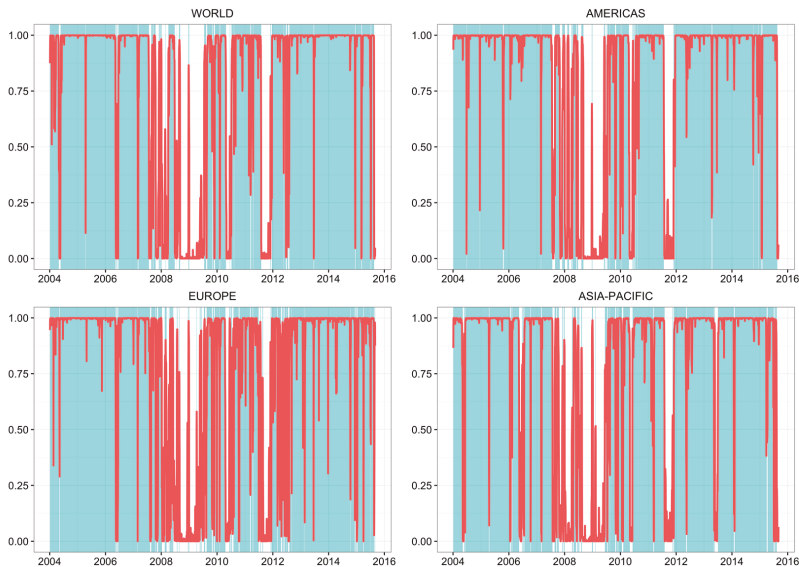
Figure 2. Smoothed probability estimates of regime 1. **Note:** The figure plots the smoothed probability estimates of the low volatility regime (regime 1). The shaded regions in the figures correspond to the periods where the smoothed probability of regime 1 is the maximum.

Panel B in Table 3 presents the parameter estimates for the MS-DCC-GARCH model that generates the regime-specific conditional correlations. We observe highly significant $\alpha(s_t)$ and $\beta(s_t)$ estimates in both regime 1 (low volatility) and regime 2 (high volatility), implying significant correlations between the conventional and sustainable market indices in both regimes. The sums $\alpha(s_t) + \beta(s_t)$ are estimated as 0.99 (0.83), 0.98 (0.95), 0.94 (0.90) and 0.99 (0.90) for the low (high) volatility regime for the World, North America, Europe, and Asia-Pacific regions, respectively, suggesting that correlations are highly persistent in both regimes consistently across all regions. Relatively higher values of $\alpha(s_t) + \beta(s_t)$ for the regional indices in both regimes imply that the correlation persistence is more pronounced at the regional level, possibly driven by regional fundamentals driving return dynamics in equity markets.

The inferences from the MS-DCC parameter estimates reported in Panel B are further supported by the probability weighted dynamic conditional correlations reported in Figure 3 (The probability weighted time-varying conditional correlations $\rho_{ij,t}$ are calculated as $\rho_{ij,t} = p_{1,t}\rho_{ij,1,t} + (1 - p_{1,t})\rho_{ij,2,t}$, where $\rho_{ij,k,t}$, $k = 1, 2$, are the time-varying conditional correlations in regime k and $p_{1,t} = P(s_t = 1|\psi_{t-1})$ is the predictive probability of being in regime 1 at time t given the information set ψ_{t-1} available through time $t - 1$). The dynamic correlations are highly time-varying for most regions, with the exception of European markets where correlations consistently range in the upper 90%. The significant time variation in the case of the other regional indices, however, further confirms the use of the DCC specification against the constant correlation alternative. Examining the plots in Figure 3, we see that both the global and regional indices exhibit a high degree of association between conventional and sustainable stocks, more consistently in the case of European stocks. Despite the high level of correlations found across all regional indices, however, a somewhat decreasing pattern in conditional correlations is observed for the Asia-Pacific region, suggesting that sustainable securities might have relatively better diversification potential for equity investors in this region. Nevertheless, the dynamic correlations clearly indicate a high degree of association between sustainable and conventional market indices, suggesting that sustainable stocks may have limited diversification benefits for conventional equity portfolios globally.
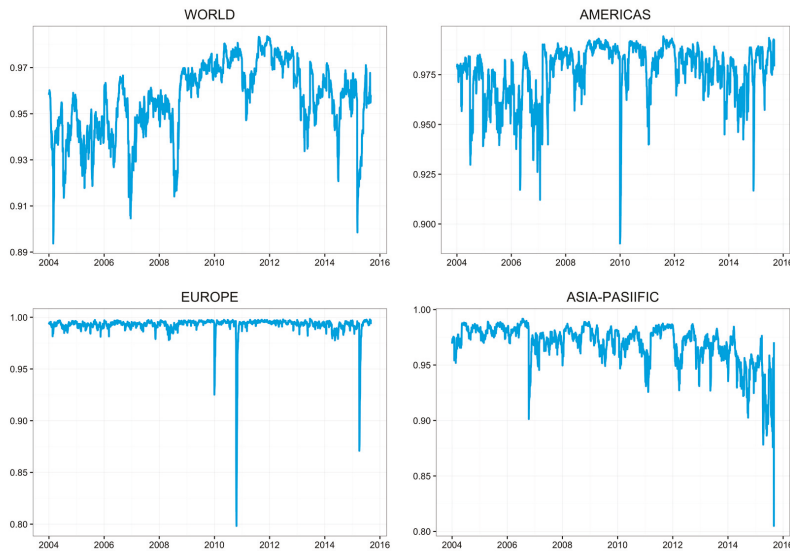
Figure 3. Dynamic correlation estimates from the MS-DCC-GARCH. **Note:** Figure plots the dynamic correlation estimates from the MS-DCC-GARCH model given in Equations (1)–(3). The correlations are obtained as the correlation coefficients are regime dependent and directly obtained from Equations (1)–(4) using the ML estimation. Since the correlations are regime-dependent and the two sets of correlations $\rho_{ij,1,t}$ and $\rho_{ij,2,t}$ are estimated for regimes 1 and 2, we obtain $\rho_{ij,t} = p_{1,t}\rho_{ij,1,t} + (1 - p_{1,t})\rho_{ij,2,t}$, where $p_{1,t} = P(s_t = 1|\psi_{t-1})$ is the predictive probability of being in regime 1 at time t given the information set ψ_{t-1} available through time $t - 1$. See Note to Table 3 for model details.

5. Portfolio Analysis

Having examined the dynamic conditional correlations between sustainable and conventional stocks, we next focus our attention on the risk and return tradeoffs offered by sustainable stocks for conventional equity investors. For this purpose, we consider a currently 'undiversified' investor,

i.e., an investor who is fully invested in a conventional stock index, and form bivariate portfolios by supplementing the undiversified portfolios with sustainable counterparts one at a time. Two alternative bivariate portfolios are examined, one based on the risk-minimizing portfolio strategy of [70]. (This model follows the dynamic risk-minimizing hedge ratio of [70] computed as $\theta_t^* = -h_{12,t}/h_{2,t}$ where $h_{i,t} = var(R_{i,t})$ and $h_{12,t} = cov(R_{1,t}, R_{2,t})$ with the subscripts 1 and 2 representing the assets in the bivariate portfolio. In our application, this is based on a \$1 long position in the conventional portfolio.) The other is based on the optimal portfolio weight of [71]. (This model follows the minimum-variance portfolio formula of [71], where the regime-independent covariances used in the computation of portfolio weights are obtained as the probability weighted average of regime-dependent covariances with the corresponding predictive regime probabilities as the weights.) A similar procedure is applied in a similar context in [58–60,72].

Table 5 presents the summary statistics for the in-sample period covering 2 January 2004–19 February 2014, with 2644 observations. We report in the table the summary statistics for portfolio returns as well as the optimal portfolio weights based on the portfolio strategies of [70,71]. Hedge effectiveness (HE), measured as the percentage of portfolio return volatility that is reduced by supplementing the undiversified portfolio with the sustainable index, along with the corresponding Sharpe ratios, are also reported in the table. Panels A, B, C and D in Table 5 present the findings for the 'undiversified' stock portfolios representing an investor who is currently fully invested in the conventional Dow Jones World, Americas, Europe, and Asia-Pacific indices, respectively. In each panel, the row labeled 'undiversified' provides the summary statistics for an undiversified investor who is currently fully invested in the corresponding conventional market.

Table 5. Summary statistics for in-sample portfolios.

	Mean	S.D.	Min	Max	HE	Sharpe Ratio
Panel A: World Market						
Undiversified Portfolio Return	0.028	1.154	−7.886	9.883	–	0.024
MS-DCC-GARCH Hedged Portfolio Return	0.018	0.295	−1.928	1.961	93.567	0.061
DCC-GARCH Hedged Portfolio Return	0.017	0.293	−1.931	2.026	93.478	0.058
MS-DCC-GARCH Optimal Portfolio Return	0.034	1.012	−8.413	7.775	23.082	0.034
DCC-GARCH Optimal Portfolio Return	0.024	0.940	−6.590	9.883	33.664	0.026
MS-DCC-GARCH Optimal Hedge Ratio	0.929	0.070	0.780	1.209	–	–
DCC-GARCH Optimal Hedge Ratio	0.935	0.059	0.814	1.122	–	–
MS-DCC-GARCH Optimal Portfolio Weight	0.618	0.417	0.000	1.000	–	–
DCC-GARCH Optimal Portfolio Weight	0.635	0.407	0.000	1.000	–	–
Panel B: Americas Market						
Undiversified Portfolio Return	0.022	1.268	−9.736	10.515	–	0.017
MS-DCC-GARCH Hedged Portfolio Return	0.010	0.199	−1.177	0.959	97.533	0.050
DCC-GARCH Hedged Portfolio Return	0.009	0.201	−1.177	1.068	97.478	0.045
MS-DCC-GARCH Optimal Portfolio Return	0.009	0.458	−9.453	8.993	16.666	0.020
DCC-GARCH Optimal Portfolio Return	0.009	0.472	−9.453	8.993	14.638	0.019
MS-DCC-GARCH Optimal Hedge Ratio	1.047	0.053	0.828	1.221	–	–
DCC-GARCH Optimal Hedge Ratio	1.042	0.015	0.960	1.080	–	–
MS-DCC-GARCH Optimal Portfolio Weight	0.078	0.216	0.000	1.000	–	–
DCC-GARCH Optimal Portfolio Weight	0.002	0.027	0.000	1.000	–	–
Panel C: European Market						
Undiversified Portfolio Return	0.019	1.476	−10.130	10.512	–	0.013
MS-DCC-GARCH Hedged Portfolio Return	0.006	0.149	−1.804	1.601	98.987	0.040
DCC-GARCH Hedged Portfolio Return	0.005	0.149	−1.804	1.601	98.979	0.034
MS-DCC-GARCH Optimal Portfolio Return	0.038	1.429	−10.130	10.512	6.342	0.027
DCC-GARCH Optimal Portfolio Return	0.015	1.466	−10.130	10.512	1.354	0.010
MS-DCC-GARCH Optimal Hedge Ratio	0.977	0.028	0.865	1.088	–	–
DCC-GARCH Optimal Hedge Ratio	0.977	0.006	0.956	0.996	–	–
MS-DCC-GARCH Optimal Portfolio Weight	0.744	0.409	0.000	1.000	–	–
DCC-GARCH Optimal Portfolio Weight	0.980	0.073	0.295	1.000	–	–
Panel D: Asia-Pacific Market						
Undiversified Portfolio Return	0.017	1.218	−9.114	9.008	–	0.014
MS-DCC-GARCH Hedged Portfolio Return	0.005	0.239	−1.431	1.434	96.135	0.021
DCC-GARCH Hedged Portfolio Return	0.003	0.240	−1.340	1.295	96.104	0.013
MS-DCC-GARCH Optimal Portfolio Return	0.016	1.094	−9.114	9.008	3.775	0.015
DCC-GARCH Optimal Portfolio Return	0.017	1.192	−9.114	9.008	0.000	0.014
MS-DCC-GARCH Optimal Hedge Ratio	0.854	0.052	0.700	1.052	–	–
DCC-GARCH Optimal Hedge Ratio	0.851	0.005	0.837	0.868	–	–
MS-DCC-GARCH Optimal Portfolio Weight	0.977	0.130	0.000	1.000	–	–
DCC-GARCH Optimal Portfolio Weight	1.000	0.000	1.000	1.000	–	–

Note: The in-sample period covers 2 January 2004–19 February 2014 with 2644 observations. HE stands for the hedge effectiveness index.

As expected, the risk-minimizing portfolio strategy of [70] yields the largest reduction in return volatility, consistently in all panels. For example, focusing on Panel A, while the undiversified portfolio that is fully invested in the conventional world index has return volatility of 1.154%, supplementing the portfolio with the sustainable counterpart helps reduce portfolio risk down to 0.295% (0.293%), leading to a 93.5% (93.4%) reduction in portfolio volatility based on the MS-DCC (DCC) specification, respectively. Clearly the high conditional correlations between the conventional and sustainable stock indices reported earlier help reduce return volatility in the hedged portfolio as the strategy by [70] takes a short position in the corresponding sustainable index. On the other hand, the optimal portfolio weight strategy of [71] does not work as effectively in mitigating portfolio risk, yielding about 33% risk reduction at best in the case of the world index in Panel A.

Examining the Sharpe ratios reported in the last column in each panel, we observe that supplementing the conventional portfolio with a position in the sustainable counterpart leads to a significant improvement in risk-adjusted returns in all panels. The improvement in Sharpe ratios is especially evident in the case of the risk-minimizing portfolio strategy of [70], where risk-adjusted returns are more than double in most regions, with the exception of Asia-Pacific in Panel D. Furthermore, comparing the risk adjusted returns and hedge effectiveness values for the MS-DCC-GARCH- and DCC-GARCH-based portfolios, we observe that the MS-DCC-GARCH model yields more favorable outcomes across all panels, underscoring the superiority of dynamic specification over the static counterpart. Overall, the in-sample portfolio findings reported in Table 5 suggest that supplementing conventional stock portfolios with their sustainable counterparts could both help reduce portfolio volatility and yield much improved risk-adjusted returns. However, this can only be achieved following the risk-minimizing portfolio strategy of [70], which takes advantage of the high correlations between the conventional and sustainable stocks by taking a short position in the sustainable index.

The in-sample portfolio results reported in Table 5 are further supported by the out-of-sample results reported in Table 6. The out-of-sample period covers 20 February 2014–2 September 2014, including 400 observations, with the estimates obtained as one-step forecasts recursively during the out-of-sample period. Consistent with the findings in Table 5, we observe that the risk-minimizing portfolio strategy yields a significant reduction in portfolio risk when the conventional index is supplemented by a position in the sustainable counterpart. The largest risk reduction is observed for the Americas (Panel B) and Europe (Panel C), with more than 96% of return volatility eliminated in the hedged portfolio. Interestingly, hedging the conventional portfolio risk with a short position in the sustainable counterpart also helps improve the risk/return profile of the portfolio in all regions. More strikingly, the negative Sharpe ratios observed for the World and Asia-Pacific indexes turn around to positive risk adjusted returns when the conventional index is supplemented by the sustainable counterpart. A similar improvement in risk-adjusted returns is also observed in other panels, indicating significant diversification benefits from sustainable stocks. In sum, despite the high conditional correlations observed between conventional and sustainable market indices, the analysis of both in- and out-of-sample portfolios clearly suggest significant diversification gains from supplementing conventional portfolios by positions in sustainable stocks. However, these diversification gains can only be achieved by implementing the risk-minimizing portfolio strategy of [67], which takes advantage of the high correlations by taking opposite positions in the conventional and sustainable portfolios.

Table 6. Summary statistics for out-of-sample portfolios.

	Mean	S.D.	Min	Max	HE	Sharpe Ratio
Panel A: World Market						
Undiversified Portfolio Return	−0.015	0.686	−3.986	2.119	–	-0.022
MS-DCC-GARCH Hedged Portfolio Return	0.008	0.219	−1.154	1.369	89.818	0.037
DCC-GARCH Hedged Portfolio Return	0.008	0.220	−1.206	1.375	89.712	0.036
MS-DCC-GARCH Optimal Portfolio Return	0.018	0.525	−2.284	3.038	41.380	0.034
DCC-GARCH Optimal Portfolio Return	0.010	0.566	−2.284	2.488	32.011	0.018
MS-DCC-GARCH Optimal Hedge Ratio	0.916	0.069	0.760	1.071	–	–
DCC-GARCH Optimal Hedge Ratio	0.927	0.041	0.844	1.025	–	–
MS-DCC-GARCH Optimal Portfolio Weight	0.648	0.390	0.000	1.000	–	–
DCC-GARCH Optimal Portfolio Weight	0.734	0.299	0.000	1.000	–	–
Panel B: Americas Market						
Undiversified Portfolio Return	0.002	0.792	−3.988	3.438	–	0.003
MS-DCC-GARCH Hedged Portfolio Return	0.006	0.141	−0.661	0.458	96.830	0.043
DCC-GARCH Hedged Portfolio Return	0.006	0.152	−0.649	0.500	96.287	0.039
MS-DCC-GARCH Optimal Portfolio Return	0.005	0.700	−2.903	3.438	21.883	0.007
DCC-GARCH Optimal Portfolio Return	0.003	0.800	−3.552	3.853	−2.088	0.004
MS-DCC-GARCH Optimal Hedge Ratio	0.995	0.050	0.864	1.161	–	–
DCC-GARCH Optimal Hedge Ratio	1.037	0.011	0.994	1.073	–	–
MS-DCC-GARCH Optimal Portfolio Weight	0.370	0.421	0.000	1.000	–	–
DCC-GARCH Optimal Portfolio Weight	0.001	0.011	0.000	0.153	–	–
Panel C: European Market						
Undiversified Portfolio Return	−0.040	0.917	−3.182	3.122	–	−0.044
MS-DCC-GARCH Hedged Portfolio Return	−0.001	0.132	−1.067	1.181	97.930	−0.008
DCC-GARCH Hedged Portfolio Return	−0.002	0.132	−1.095	1.181	97.928	−0.015
MS-DCC-GARCH Optimal Portfolio Return	−0.032	0.827	−2.911	3.122	18.625	−0.039
DCC-GARCH Optimal Portfolio Return	−0.039	0.891	−3.182	3.122	5.500	−0.044
MS-DCC-GARCH Optimal Hedge Ratio	0.975	0.029	0.840	1.060	–	–
DCC-GARCH Optimal Hedge Ratio	0.981	0.006	0.967	0.997	–	–
MS-DCC-GARCH Optimal Portfolio Weight	0.734	0.393	0.000	1.000	–	–
DCC-GARCH Optimal Portfolio Weight	0.933	0.165	0.200	1.000	–	–
Panel D: Asia-Pacific Market						
Undiversified Portfolio Return	−0.019	0.708	−4.425	2.146	–	−0.027
MS-DCC-GARCH Hedged Portfolio Return	0.003	0.328	−4.267	0.851	78.579	0.009
DCC-GARCH Hedged Portfolio Return	0.002	0.332	−4.370	0.886	78.034	0.006
MS-DCC-GARCH Optimal Portfolio Return	−0.007	0.639	−2.305	2.146	18.430	−0.011
DCC-GARCH Optimal Portfolio Return	−0.019	0.708	−4.425	2.146	0.000	−0.027
MS-DCC-GARCH Optimal Hedge Ratio	0.824	0.063	0.686	1.080	–	–
DCC-GARCH Optimal Hedge Ratio	0.845	0.005	0.835	0.861	–	–
MS-DCC-GARCH Optimal Portfolio Weight	0.928	0.222	0.000	1.000	–	–
DCC-GARCH Optimal Portfolio Weight	1.000	0.000	1.000	1.000	–	–

Note: The out-of-sample period covers 20 February 2014–2 September 2014 with 400 observations. HE stands for the hedge effectiveness index.

6. Conclusions

This paper explores the potential diversification benefits of socially responsible investments for conventional stock portfolios by examining the risk transmissions and dynamic correlations between conventional and sustainable stock indices from a number of regions. Utilizing a Markov regime-switching GARCH model with dynamic conditional correlations (MS-DCC-GARCH), we find evidence of significant and positive volatility spillovers from conventional to sustainable equities, suggesting that uncertainty regarding global equity markets spills over to the market for sustainable stocks, driving return volatility in this market segment. Risk transmissions, however, are found to be unidirectional, implied by largely insignificant spillover effects from sustainable to conventional indexes. We argue that the economic, environmental and social criteria applied in the selection of firms to be included in socially responsible indices do not necessarily shield these stocks from common equity market shocks. Despite the presence of risk transmissions from conventional markets, however, our findings also suggest that historical information on return and volatility in sustainable equity markets could be utilized in forecasting future volatility in these markets. Thus, investors and trustees of institutional funds who are concerned about stability in the market for sustainable investments should not only monitor volatility in global conventional markets, but also supplement their volatility forecasting models by measures of historical risk and return dynamics in these markets.

Similarly, the analysis of dynamic conditional correlations suggests that both the global and regional indices exhibit a high degree of association between conventional and sustainable stocks, more consistently in the case of European stocks. Although significant time-variations in the dynamic correlations are observed between conventional and sustainable stock returns, we estimate particularly high correlations that consistently range in the upper 90% in the case of Europe. Interestingly, however, despite the high correlations observed, the analysis of both in- and out-of-sample portfolios suggests that significant diversification gains can be obtained from supplementing conventional portfolios by positions in sustainable stocks. Improvement in risk adjusted returns is particularly striking for the broader world index and the Asia-Pacific region when the negative Sharpe ratios for undiversified, conventional portfolios turn around to positive values when the conventional index is supplemented by the sustainable counterpart. However, our portfolio analysis also suggests that these diversification gains can only be achieved by implementing an investment strategy that aims to minimize portfolio risk and utilize sustainable assets in the short leg of the portfolio.

Given the availability of various exchange-traded funds that allow investors to choose investments based on social and personal criteria, our findings have significant implications for both retail and institutional investors. Thanks to the rapid growth experienced in the SRI market segment, investors have their choices when it comes to allocating parts of their portfolios in various exchange traded funds that reflect this growing segment. Furthermore, the fact that these funds are offered to investors at low cost makes transaction costs less of a concern from a retail investor perspective. More importantly, unlike the case for individual stocks, for which uptick rules apply, diversifying into SRIs via short positions in exchange traded funds that do not have the uptick rules means that investors will have greater flexibility in the creation of diversified portfolios as we recommend in our empirical results. Overall, the findings suggest that sustainable investments can indeed provide significant diversification gains for conventional stock portfolios globally and the fact that these investments are easily accessible at low cost via a myriad of exchange traded funds makes them an appealing investment tool both for retail and institutional investors.

References

1. Roca, E.; Wong, V.S.; Anand, T.G. Are socially responsible investment markets worldwide integrated? *Account. Res. J.* **2010**, *23*, 281–301. [CrossRef]
2. Global Sustainable Investment Association. *Global Sustainable Investment Review 2014*; Global Sustainable Investment Association (GSIA): Washington, DC, USA, 2014.
3. Nofsinger, J.; Varma, A. Socially responsible funds and market crises. *J. Bank. Financ.* **2014**, *48*, 180–193. [CrossRef]
4. Markowitz, H. Portfolio selection. *J. Financ.* **1952**, *7*, 77–91.
5. Markowitz, H. *Portfolio Selection: Efficient Diversification of Investments*; Wiley: New York, NY, USA, 1959.
6. Statman, M. What do investors want? *J. Portf. Manag.* **2004**, *30*, 153–161. [CrossRef]
7. Bollen, N.P. Mutual fund attributes and investor behavior. *J. Financ. Quant. Anal.* **2007**, *42*, 683–708. [CrossRef]
8. Benson, K.L.; Humphrey, J.E. Socially responsible investment funds: Investor reaction to current and past returns. *J. Bank. Financ.* **2008**, *32*, 1850–1859. [CrossRef]
9. Renneboog, L.; Horst, T.J.; Zhang, C. Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *J. Financ. Intermed.* **2011**, *20*, 562–588. [CrossRef]
10. Renneboog, L.; Horst, T.J.; Zhang, C. *Is Ethical Money Financially Smart?* ECGI-Finance Working Paper; European Corporate Governance Institute: Brussels, Belgium, 2006; Volume 117.

11. Renneboog, L.; Ter Horst, J.; Zhang, C. The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *J. Corp. Financ.* **2008**, *14*, 302–322. [CrossRef]

12. Renneboog, L.; Horst, T.J.; Zhang, C. Socially responsible investments: Institutional aspects, performance, and investor behavior. *J. Bank. Financ.* **2008**, *32*, 1723–1742. [CrossRef]

13. Moskowitz, M. Choosing socially responsible stocks. *Bus. Soc. Rev.* **1972**, *1*, 71–75.

14. Luther, R.G.; Matatko, J.; Corner, D.C. The Investment Performance of UK Ethical Unit Trusts. *Account. Audit. Account. J.* **1992**, *5*, 57–70. [CrossRef]

15. Hamilton, S.; Jo, H.; Statman, M. Doing well while doing good? The investment performance of socially responsible mutual funds. *Financ. Anal. J.* **1993**, *49*, 62–66. [CrossRef]

16. Luther, R.G.; Matatko, J. The performance of ethical unit trusts: Choosing an appropriate benchmark. *Br. Account. Rev.* **1994**, *26*, 77–89. [CrossRef]

17. Mallin, C.A.; Saadouni, B.; Briston, R.J. The financial performance of ethical investment funds. *J. Bus. Financ. Account.* **1995**, *22*, 483–496. [CrossRef]

18. White, M.A. The performance of environmental mutual funds in the United States and Germany: Is there economic hope for green investors. *Res. Corp. Soc. Perform. Policy* **1995**, *1*, 323–344.

19. Kurtz, L.; DiBartolomeo, D. Socially screened portfolios: An attribution analysis of relative performance. *J. Invest.* **1996**, *5*, 35–41. [CrossRef]

20. Gregory, A.; Matatko, J.; Luther, R. Ethical unit trust financial performance: Small company effects and fund size effects. *J. Bus. Financ. Account.* **1997**, *24*, 705–725. [CrossRef]

21. Russo, M.V.; Fouts, P.A. A resource-based perspective on corporate environmental performance and profitability. *Acad. Manag. J.* **1997**, *40*, 534–559. [CrossRef]

22. Sauer, D.A. The impact of social-responsibility screens on investment performance: Evidence from the Domini 400 Social Index and Domini Equity Mutual Fund. *Rev. Financ. Econ.* **1997**, *6*, 137–149. [CrossRef]

23. DiBartolomeo, D.; Kurtz, L. *Managing Risk Exposures of Socially Screened Portfolios*; Northfield Information Services: Boston, MA, USA, 1999; pp. 1–17.

24. Goldreyer, E.F.; Diltz, J.D. The performance of socially responsible mutual funds: Incorporating sociopolitical information in portfolio selection. *Manag. Financ.* **1999**, *25*, 23–36. [CrossRef]

25. Statman, M. Socially responsible mutual funds. *Financ. Anal. J.* **2000**, *56*, 30–39. [CrossRef]

26. Stone, B.K.; Guerard, J.B., Jr.; Gultekin, M.N.; Adams, G. *Socially Responsible Investment Screening: Strong Evidence of No Significant Cost for Actively Managed Portfolios*; Moskowitz Prize Honorable Mention: San Francisco, CA, USA, 2001.

27. Garz, H.; Volk, C.; Gilles, M. More Gain than Pain, SRI: Sustainability Pays Off. West LB Panmure, 2002. Available online: http://www.westlbpanmure.com/sri/pdf/sri_nov2002.pdf (accessed on 27 September 2017).

28. Kreander, N.; Gray, R.; Power, D.; Sinclair, C. The Financial Performance of European Ethical Funds 1996–1998. *J. Account. Financ.* **2002**, *1*, 3–22.

29. Geczy, C.; Stambaugh, R.F.; Levin, D. Investing in Socially Responsible Mutual Funds, Wharton School Working Paper. 2005. Available online: http://finance.wharton.upenn.edu/~rlwctr/papers/0402.pdf (accessed on 27 September 2017).

30. Orlitzky, M.; Schmidt, F.L.; Rynes, S.L. Corporate social and financial performance: A meta-analysis. *Organ. Stud.* **2003**, *24*, 403–441. [CrossRef]

31. Bauer, R.; Koedijk, K.; Otten, R. International evidence on ethical mutual fund performance and investment style. *J. Bank. Financ.* **2005**, *29*, 1751–1767. [CrossRef]

32. Shank, T.; Manullang, D.; Hill, R. Doing well while doing good revisited: A study of socially responsible firms' short-term versus long-term performance. *Manag. Financ.* **2005**, *31*, 33–46. [CrossRef]

33. Bauer, R.; Derwall, J.; Otten, R. The ethical mutual fund performance debate: New evidence from Canada. *J. Bus. Eth.* **2007**, *70*, 111–124. [CrossRef]

34. Edmans, A. Does the stock market fully value intangibles? Employee satisfaction and equity prices. *J. Financ. Econ.* **2011**, *101*, 621–640. [CrossRef]

35. Girard, E.C.; Rahman, H.; Stone, B.A. Socially responsible investments: Goody-two-shoes or bad to the bone? *J. Invest.* **2007**, *16*, 96–110. [CrossRef]

36. Schröder, M. Is there a difference? The performance characteristics of SRI equity indices. *J. Bus. Financ. Account.* **2007**, *34*, 331–348. [CrossRef]

37. Adler, T.; Kritzman, M. The cost of socially responsible investing. *J. Portf. Manag.* **2008**, *35*, 52–56. [CrossRef]
38. Galema, R.; Plantinga, A.; Scholtens, B. The stocks at stake: Return and risk in socially responsible investment. *J. Bank. Financ.* **2008**, *32*, 2646–2654. [CrossRef]
39. Derwall, J.; Guenster, N.; Bauer, R.; Koedijk, K. The eco-efficiency premium puzzle. *Financ. Anal. J.* **2005**, *61*, 51–63. [CrossRef]
40. Kempf, A.; Osthoff, P. The effect of socially responsible investing on portfolio performance. *Eur. Financ. Manag.* **2007**, *13*, 908–922. [CrossRef]
41. Hong, H.; Kacperczyk, M. The price of sin: The effects of social norms on markets. *J. Financ. Econ.* **2009**, *93*, 15–36. [CrossRef]
42. Statman, M.; Glushkov, D. The wages of social responsibility. *Financ. Anal. J.* **2009**, *65*, 33–46. [CrossRef]
43. Kim, I.; Venkatachalam, M. Are sin stocks paying the price for accounting sins? *J. Account. Audit. Financ.* **2011**, *26*, 415–442. [CrossRef]
44. Guenster, N. Performance implications of SR investing: Past versus future. *Soc. Responsib. Financ. Invest.* **2012**, 443–454. [CrossRef]
45. Borgers, A.; Derwall, J.; Koedijk, K.; Ter Horst, J. Stakeholder relations and stock returns: On errors in investors' expectations and learning. *J. Empir. Financ.* **2013**, *22*, 159–175. [CrossRef]
46. Angel, J.J.; Rivoli, P. Does ethical investing impose a cost upon the firm? A theoretical perspective. *J. Invest.* **1997**, *6*, 57–61. [CrossRef]
47. Lee, C.; Ng, D.T. *Corruption and International Valuation: Does Virtue Pay? Working Paper*; Cornell University: Ithaca, NY, USA, 2006.
48. Guenster, N.; Bauer, R.; Derwall, J.; Koedijk, K. The economic value of corporate eco-efficiency. *Eur. Financ. Manag.* **2011**, *17*, 679–704. [CrossRef]
49. Guerard, J.B., Jr. Is there a cost to being socially responsible in investing? *J. Invest.* **1997**, *6*, 11–18. [CrossRef]
50. Lean, H.H.; Nguyen, D.K. Policy uncertainty and performance characteristics of sustainable investments across regions around the global financial crisis. *Appl. Financ. Econ.* **2014**, *24*, 1367–1373. [CrossRef]
51. Antonakakis, N.; Babalos, V.; Kyei, C. Predictability of sustainable investments and the role of uncertainty: Evidence from a non-parametric causality-in-quantiles test. *Appl. Econ.* **2017**, in press. [CrossRef]
52. Apergis, N.; Babalos, V.; Christou, C.; Gupta, R. *Identifying Asymmetries between Socially Responsible and Conventional Investments*; Working Paper No. 201537; Department of Economics, University of Pretoria: Pretoria, South Africa, 2015.
53. Engle, R.F. Dynamic conditional correlation: A new simple class of multivariate GARCH models. *J. Bus. Econ. Stat.* **2002**, *20*, 339–350. [CrossRef]
54. Kim, C. Dynamic linear models with Markov-switching. *J. Econ.* **1994**, *60*, 1–22. [CrossRef]
55. Ciner, C.; Gurdgiev, C.; Lucey, B. Hedges and safe havens: An examination of stocks, bonds, gold, oil and exchange rates. *Int. Rev. Financ. Anal.* **2013**, *29*, 202–211. [CrossRef]
56. Jawadi, F.; Jawadi, N.; Louhichi, W. Conventional and Islamic stock price performance: An empirical investigation. *Int. Econ.* **2014**, *137*, 73–87. [CrossRef]
57. Billio, M.; Caporin, M. Multivariate Markov switching dynamic conditional correlation GARCH representations for contagion analysis. *Stat. Methods Appl.* **2005**, *14*, 145–161. [CrossRef]
58. Lee, H.-T. Regime switching correlation hedging. *J. Bank. Financ.* **2010**, *34*, 2728–2741. [CrossRef]
59. Chang, C.-L.; McAleer, M.; Tansuchat, R. Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Econ.* **2011**, *33*, 912–923. [CrossRef]
60. Balcilar, M.; Demirer, R.; Hammoudeh, S.; Nguyen, D.K. Risk spillovers across the energy and carbon markets and hedging strategies for carbon risk. *Energy Econ.* **2016**, *54*, 159–172. [CrossRef]
61. Ling, S.; McAleer, M. Asymptotic theory for a vector ARMA-GARCH model. *Econ. Theory* **2003**, *19*, 280–310. [CrossRef]
62. Engle, R.F.; Sheppard, K. *Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH*; Discussion Paper No. 2001-15; NBER Working Paper No. 8554; University of California: San Diego, CA, USA, 2001.
63. Cai, J. A Markov model of switching-regime ARCH. *J. Bus. Econ. Stat.* **1994**, *12*, 309–316.
64. Hamilton, J.D.; Susmel, R. Autoregressive conditional heteroskedasticity and changes in regime. *J. Econ.* **1994**, *64*, 307–333. [CrossRef]

65. Gray, S.F. Modeling the conditional distribution of interest rates as a regime-switching process. *J. Financ. Econ.* **1996**, *42*, 27–62. [CrossRef]

66. Caporin, M.; McAleer, M. Ten things you should know about the dynamic conditional correlation representation. *Econometrics* **2013**, *1*, 115–126. [CrossRef]

67. Davies, R.B. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* **1987**, *74*, 33–43.

68. Nakatani, T.; Teräsvirta, T. *An Alternative Test for Causality in Variance in the Conditional Correlation GARCH Models*; Stockholm School of Economics: Stockholm, Sweden, 2010.

69. Hafner, C.M.; Herwartz, H. A Lagrange multiplier test for causality in variance. *Econ. Lett.* **2006**, *93*, 137–141. [CrossRef]

70. Kroner, K.F.; Sultan, J. Time dynamic varying distributions and dynamic hedging with foreign currency futures. *J. Financ. Quant. Anal.* **1993**, *28*, 535–551. [CrossRef]

71. Kroner, K.; Ng, V. Modeling asymmetric comovements of asset returns. *Rev. Financ. Stud.* **1998**, *11*, 817–844. [CrossRef]

72. Hammoudeh, S.; Yuan, Y.; McAleer, M.; Thompson, M. Precious metals-exchange rate volatility transmissions and hedging strategies. *Int. Rev. Econ. Financ.* **2010**, *20*, 633–647. [CrossRef]

Article

Volatility Spillovers and Causality of Carbon Emissions, Oil and Coal Spot and Futures for the EU and USA

Chia-Lin Chang [1], Michael McAleer [2,3,4,5,6,*] and Guangdong Zuo [2]

[1] Department of Applied Economics, Department of Finance, National Chung Hsing University, Taichung 402, Taiwan; changchialin@email.nchu.edu.tw
[2] Department of Quantitative Finance, National Tsing Hua University, Hsinchu 300, Taiwan; guangdong.zuo@foxmail.com
[3] Discipline of Business Analytics, University of Sydney Business School, Sydney, NSW 2006, Australia
[4] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, 3062 Rotterdam, The Netherlands
[5] Department of Quantitative Economics, Complutense University of Madrid, 28040 Madrid, Spain
[6] Institute of Advanced Sciences, Yokohama National University, Yokohama 240-0067, Japan
* Correspondence: michael.mcaleer@gmail.com

Abstract: Recent research shows that the efforts to limit climate change should focus on reducing the emissions of carbon dioxide over other greenhouse gases or air pollutants. Many countries are paying substantial attention to carbon emissions to improve air quality and public health. The largest source of carbon emissions from human activities in some countries in Europe and elsewhere is from burning fossil fuels for electricity, heat, and transportation. The prices of fuel and carbon emissions can influence each other. Owing to the importance of carbon emissions and their connection to fossil fuels, and the possibility of Granger causality in spot and futures prices, returns, and volatility of carbon emissions, crude oil and coal have recently become very important research topics. For the USA, daily spot and futures prices are available for crude oil and coal, but there are no daily futures prices for carbon emissions. For the European Union (EU), there are no daily spot prices for coal or carbon emissions, but there are daily futures prices for crude oil, coal and carbon emissions. For this reason, daily prices will be used to analyse Granger causality and volatility spillovers in spot and futures prices of carbon emissions, crude oil, and coal. As the estimators are based on quasi-maximum likelihood estimators (QMLE) under the incorrect assumption of a normal distribution, we modify the likelihood ratio (LR) test to a quasi-likelihood ratio test (QLR) to test the multivariate conditional volatility Diagonal BEKK model, which estimates and tests volatility spillovers, and has valid regularity conditions and asymptotic properties, against the alternative Full BEKK model, which also estimates volatility spillovers, but has valid regularity conditions and asymptotic properties only under the null hypothesis of zero off-diagonal elements. Dynamic hedging strategies by using optimal hedge ratios are suggested to analyse market fluctuations in the spot and futures returns and volatility of carbon emissions, crude oil, and coal prices.

Keywords: carbon emissions; fossil fuels; crude oil; coal; low carbon targets; green energy; spot and futures prices; Granger causality; volatility spillovers; quasi likelihood ratio (QLR) test; diagonal BEKK; full BEKK; dynamic hedging

JEL Classification: C58; L71; O13; P28; Q42

1. Introduction

Recent research shows that efforts to limit climate change should focus on reducing emissions of carbon dioxide over other greenhouse gases or air pollutants. Many countries are paying substantially greater attention to carbon emissions to improve air quality and public health. Carbon emissions trading programs have been established at the international, regional, national, and sub-national levels (see Figure 1).
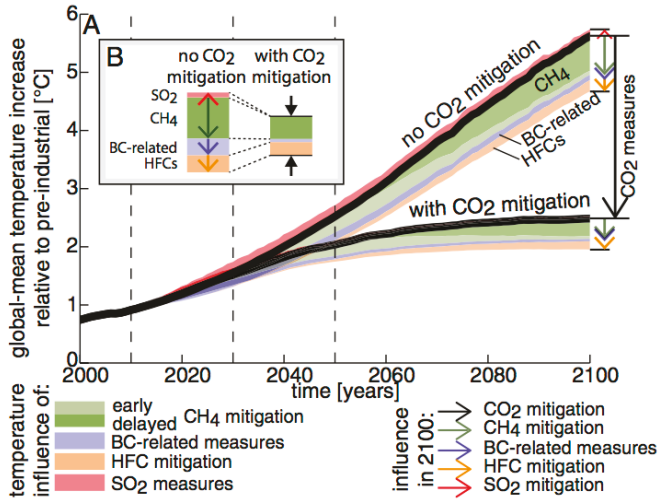
Figure 1. Global mean temperatures. With and without carbon dioxide mitigation. Source: [1] Rogelj et al. (2014).

As can be seen from Figure 1, in a scenario of 'no carbon dioxide mitigation', global temperatures would be predicted to rise by over five degrees Celsius by 2100, but cutting emissions of methane, HFCs, and black carbon would reduce this rise to around one degree Celsius. The results suggest that carbon dioxide should certainly remain central to greenhouse gas emission cuts.

Figure 2 shows that projects and regions such as the CDM (Clean Development Mechanism), RGGI (Regional Greenhouse Gas Initiative), and European Union (EU), countries like New Zealand, Australia, and South Korea, the State of California in the USA, and the Province of Quebec inn Canada, have passed and implemented programs to mitigate carbon emissions.
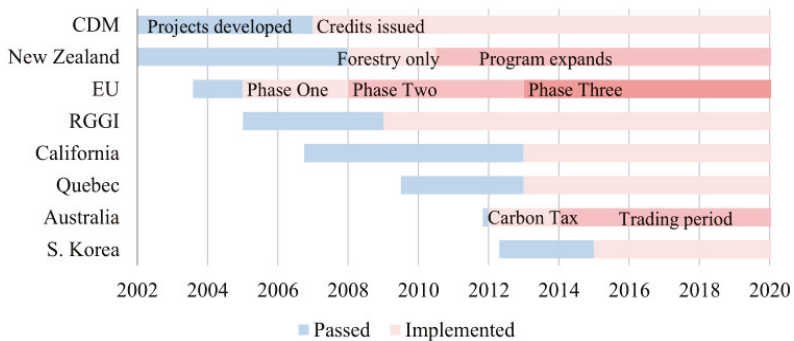
Figure 2. Implementation of programs to mitigate carbon emissions.

The programs have operated in phases, with a pilot phase from 2005 to 2007 covering the power sector and certain heavy industries, a second phase from 2008 to 2012 expanding coverage slightly, and a third phase for 2013–2020 that adds a significant range of industrial activities.

The largest source of carbon emissions from human activities in some countries in Europe and elsewhere is from burning fossil fuels for electricity, heat, and transportation. The price of fuel influences carbon emissions, but the price of carbon emissions can also influence the price of fuel.

Owing to the importance of carbon emissions and their connection to fossil fuels, and the possibility of [2] Granger (1980) causality in spot and futures prices, returns and volatility of carbon emissions, it is not surprising that crude oil and coal have recently become a very important public policy issue, and hence also a significant research topic.

Energy markets have recently expanded considerably due in large part to the rapidly accelerating behaviour of investors in financial markets. The synergy between financial and energy markets is that the financial aspect of fossil fuels and carbon emissions need to be analysed more carefully by using advanced financial econometric methods. An important reference in the field of energy prices and its consequences on financial markets are the empirical studies presented in [3] Ramos and Veiga (2014). These macroeconomic variables include risk factors in the oil industry, risk taking in the airline industry, prices, volatility, and shocks in the oil industry, oil shock spillovers to stock market returns, equity returns, bond returns, and volatility market risks.

In a more microeconomic context, [4] Sawik, Faulin and Pérez-Bernabeu (2017a) examine energy and environment issues with respect to multi-criteria analysis and multi-objective green logistics optimization. The optimality criteria are presented in terms of environmental costs, that is, the minimization of externality costs for noise, pollution, and fuel costs as compared with their minimization. In a separate contribution, [5] Sawik, Faulin and Pérez-Bernabeu (2017b) solve a multi-objective formulation problem by minimizing the total distance, and hence the costs to a delivery company, and the amount of CO_2 emissions. [6] Sawik, Faulin and Pérez-Bernabeu (2017c) optimize a multi-criteria formulation for green vehicle routing problems by mixed integer programming, specifically to decide the best delivery route to minimize the travel costs and optimize the transportation route of a delivery company.

The plan of the remainder of the paper is as follows. Section 2 discusses the spot and futures data for carbon emissions, coal, and oil that will be used in the empirical analysis for the EU and USA. Section 3 discusses methodological issues, including univariate and multivariate conditional volatility models, Granger causality, volatility spillovers, optimal hedge ratios, causality in returns and volatility, as well as an interesting and novel adaptation of the likelihood ratio (LR) test to a quasi likelihood ratio (QLR) test of the Diagonal BEKK model against the alternative of a Full BEKK model. Section 4 examines the alternative unit root tests that are used to test for stationarity in the data. Granger Causality and Spillovers in Returns and Volatilities are analysed in Section 5. Section 6 provides some concluding remarks.

2. Data

The length of the sample period for the empirical analysis was dictated by the availability of data on carbon, coal, and crude oil spot and futures prices in the EU and the United States of America (USA). The carbon emission trading market of the EU has the longest trading period for futures prices, but not for spot prices. The USA is the leader in developing a wide range of financial derivatives, such as futures prices, for financial, energy, and commodities, but not for carbon emissions, where only spot prices are available.

Data for EU carbon emission, crude oil, and coal futures are available from 1 April 2008 to 20 May 2017, and these will be analyzed in the paper. Coal spot price in the EU is available on a weekly basis. The spot prices of carbon emission and crude oil have a high correlation with the corresponding futures prices. The volume of trades in the spot market of carbon emissions is much smaller than in the futures market, as shown in Figure 3.
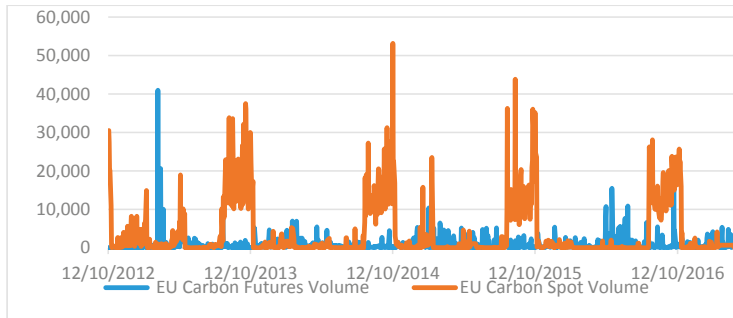
Figure 3. Carbon futures and spot volumes for European Union (EU) 10 December 2012–19 May 2017.

Data for crude oil are available prior to 2000. However, the data for the spot prices of coal and carbon emissions start from 17 July 2006 and 1 April 2008, respectively. Therefore, the data in the empirical analysis for the European Union starts from the latest date for crude oil, coal, and carbon emissions, namely 1 April 2008.

Data for carbon, coal, and oil spot prices from 5 January 2016 to 20 May 2017 for the USA will also be analyzed in the paper, but data for futures prices of carbon emissions are not available for the USA. Spot prices for coal and crude oil start prior to 2000. However, data for carbon emissions start from 1 May 2016. Consequently, the spot price data in the empirical analysis for the USA starts from the latest date for oil, coal, and carbon emissions, namely 5 January 2016.

The transaction markets and units for the variables are different. EU carbon futures is the Intercontinental Exchange EU allowance, which is traded in the ICE-ICE Futures Europe Commodities market and is expressed in Euros per metric ton. EU coal futures is ICE Rotterdam Monthly Coal Futures Contract, and is traded in the ICE-ICE Futures Europe Commodities market. EU oil futures is the current pipeline export quality Brent blend, as supplied at Sullom Voe, is traded in the ICE-ICE Futures Europe Commodities market, and is expressed in USDs per bbl.

Carbon spot prices in the USA are given as the United States Carbon Dioxide RGGI Allowance, and are expressed in USDs per allowance. Coal spot prices are given as the Dow Jones US Total Market Coal Index, which is expressed in USD. Oil spot returns are given as the West Texas Intermediate Cushing Crude Oil, which is expressed in USDs per bbl. All of the currency units are transformed to USD in the empirical analysis.

The endogenous variables used in the empirical analysis are daily returns, where the rate of return is obtained as the first difference in the natural logarithm of the relevant daily price data. The mnemonics EUcarbon$_{fr}$, EUcoal$_{fr}$, EUoil$_{fr}$ denote, respectively, the future returns of carbon emission, coal, and oil in the European Union. Similarly, the mnemonics UScarbon$_{sr}$, UScoal$_{sr}$, USoil$_{sr}$ denote, respectively, the spot returns for carbon emission, coal, and oil in the USA.

The variable sources and definitions are given in Table 1, with respect to the futures returns for the EU and spot returns for the USA, as well as their transactions markets, and the descriptions of the data.

For the USA, daily spot and futures prices are available for crude oil and coal, but there are no daily spot or futures prices for carbon emissions. For the EU, there are no daily spot prices for coal or carbon emissions, but there are daily futures prices for crude oil, coal, and carbon emissions.

For this reason, daily futures prices will be used to analyse Granger causality and volatility spillovers in spot and futures prices of carbon emissions, crude oil, and coal. This will be based on the Lagrange multiplier test of univariate causality in variance (strictly, causality in conditional volatility) of [7] Hafner and Herwartz (2006), and more recently, [8] Chang and McAleer (2017). An extension to multivariate tests of causality in conditional volatility will be a focus of the paper.

Table 1. Data Sources and Definitions.

Variable Name	Definitions	Transaction Market	Description
EUcarbon$_{fr}$	EU carbon futures return	ICE-ICE Futures Europe Commodities	ICE EUA Futures Contract EUR/MT
EUcoal$_{fr}$	EU coal futures return	ICE-ICE Futures Europe Commodities	ICE Rotterdam Monthly Coal Futures Contract USD/MT
EUoil$_{fr}$	EU oil futures return	ICE-ICE Futures Europe Commodities	Current pipeline export quality Brent blend as supplied at Sullom Voe USD/bbl
UScarbon$_{sr}$	US carbon spot return	over the counter	United States Carbon Dioxide RGGI Allowance USD/Allowance
UScoal$_{sr}$	US coal spot return	over the counter	Dow Jones US Total Market Coal Index USD
USoil$_{sr}$	US oil spot return	over the counter	West Texas Intermediate Cushing Crude Oil USD/bbl

ICE is the Intercontinental Exchange; EUA is the EU allowance; MT is metric ton; RGGI (Regional Greenhouse Gas Initiative) is a CO_2 cap-and-trade emissions trading program that is comprised of ten New England and Mid-Atlantic States that will commence in 2009 and aims to reduce emissions from the power sector. RGGI will be the first government mandated CO_2 emissions trading program in USA.

As the estimators are based on Quasi-Maximum Likelihood Estimators (QMLE) under the incorrect assumption of a normal likelihood function, we will modify the likelihood ratio (LR) test to a novel quasi-likelihood ratio test (QLR).

Definition of QLR test statistic: QLR = 2 (quasi maximized log likelihood value under the alternative hypothesis − quasi maximized log likelihood value under the null hypothesis).

The QLR test statistic tests the multivariate conditional volatility Diagonal BEKK model, which is used to estimate and test spillovers, and which has valid regularity conditions and asymptotic properties, against the alternative Full BEKK model, which is used to estimate spillovers, but has valid regularity conditions and asymptotic properties only under the null hypothesis of zero off-diagonal elements. Dynamic hedging strategies using optimal hedge ratios will be suggested to analyse market fluctuations in the spot and futures returns and volatility of carbon emissions, crude oil, and coal prices.

The QLR statistic has an asymptotic chi-squared distribution under the null hypothesis, with degrees of freedom (*df*) equivalent to the number of off-diagonal terms in the two $m \times m$ matrices, the weighting matrix, A, and the stability matrix, B, of the Full BEKK model, namely $2m(m - 1)$.

The descriptive statistics for the endogenous returns of the variables are given in Table 2. The highest standard deviation for the EU over the sample period is for carbon futures, followed by oil and coal futures. Similarly, the highest standard deviation for the US market is for coal spot returns, followed by carbon emission spot returns.

Table 2. Descriptive Statistics 2 April 2008–19 May 2017 for EU 6 January 2016–19 May 2017 for United States of America (USA).

Variable	Mean	Median	Max	Min	SD	Skewness	Kurtosis	Jarque-Bera
EUcarbon$_{fr}$	−0.078	−0.038	24.561	−42.457	3.349	−0.708	17.624	21,434.2
EUcoal$_{fr}$	−0.022	0	17.419	−22.859	1.599	−1.268	44.924	175,155.8
EUoil$_{fr}$	−0.026	−0.015	12.707	−10.946	2.246	0.054	6.522	1232.8
UScarbon$_{sr}$	−0.248	0	13.937	−36.446	2.986	−5.236	66.269	61,346.8
UScoal$_{sr}$	0.177	0.104	17.458	−14.183	4.041	0.047	5.343	81.99
USoil$_{sr}$	0.094	0.037	11.621	−8.763	2.712	0.431	4.690	53.69

The Jarque-Bera Lagrange multiplier statistic for normality is based on testing the empirical skewness and kurtosis against their normal counterparts.

The returns have different degrees of skewness. The futures and spot returns of oil in the EU and US markets, and coal spot returns in the USA are skewed to the left, indicating that these series have longer left tails (extreme losses) than right tails (extreme gains). However, other returns are all skewed to the right, especially carbon emission spot return in the USA, for which the value of the skewness is high, indicating that these series have more extreme gains than extreme losses.

These stylized facts should be of interest to participants in commodity markets. All of the price distributions have kurtosis that is significantly higher than three, implying that higher probabilities of extreme market movements in either direction (gains or losses) occur in these futures markets, with greater frequency in practice than would be expected under the normal distribution.

In the EU market, the highest kurtosis is for coal futures, followed by carbon futures and oil futures. For the US market, the highest kurtosis is for carbon spot, followed by coal spot. The Jarque-Bera Lagrange multiplier statistic is based on testing the empirical skewness and kurtosis against their normal counterparts, and confirms the non-normal distributions for all of the returns series.

3. Methodology

Although financial and energy returns are almost certainly stationary, the empirical analysis will commence with tests of unit roots based on ADF, DF-GLS, and KPSS. This will be followed by an analysis and estimation of univariate GARCH and multivariate diagonal BEKK models (see [9] Baba et al. (1985) [10] Engle and Kroner (1995)), from which the conditional covariances will be used for testing co-volatility spillovers, that is, Granger causality in conditional volatility.

Despite the empirical applications of a wide range of conditional volatility models in numerous papers in empirical finance, there are theoretical problems associated with virtually all of them. The CCC ([11] Bollerslev (1990)), VARMA-GARCH ([12] Ling and McAleer (2003), and its asymmetric counterpart, VARMA-AGARCH [13] McAleer et al. (2009)), models have static conditional covariances and correlations, which means that accommodating volatility spillovers is not possible.

Apart from the diagonal version, the multivariate Full BEKK model of conditional covariances has been shown to have no regularity conditions, and hence no statistical properties (see [14] McAleer et al. (2008) and [15] Chang and McAleer (2017b), and the discussion below, for further details). Therefore, spillovers can be considered only for the special case of Diagonal BEKK. The multivariate DCC model of (purported) conditional correlations has been shown to have no regularity conditions, and hence no statistical properties (see [16] Hafner and McAleer (2014) and [17] McAleer (2017) for further details).

The analysis of univariate and multivariate conditional volatility models below is a summary of what has been presented in the literature (see, for example [18] Caporin and McAleer (2012) [19] Chang et al. (2015), and especially [20] Chang et al. (2017)), although a comprehensive discussion of the Full and Diagonal BEKK models is not available in any published source. In particular, the application of the quasi likelihood ratio (QLR) test of the Diagonal BEKK model as the null hypothesis against the alternative hypothesis of a Full BEKK model does not seem to have been considered in the literature.

The first step in estimating multivariate models is to obtain the standardized residuals from the conditional mean returns shocks. For this reason, the most widely used univariate conditional volatility model, namely GARCH, will be presented briefly, followed by the two most widely estimated multivariate conditional covariance models, namely the Diagonal and Full BEKK models.

3.1. Univariate Conditional Volatility

Consider the conditional mean of financial returns, as follows:

$$y_t = E(y_t|I_{t-1}) + \varepsilon_t \tag{1}$$

where the financial returns, $y_t = \Delta log P_t$, represent the log-difference in the financial commodity or agricultural prices, P_t, I_{t-1} is the information set at time $t-1$, and ε_t is a conditionally heteroskedastic

error term, or returns shock. In order to derive conditional volatility specifications, it is necessary to specify the stochastic processes underlying the returns shocks, ε_t. The most popular univariate conditional volatility model, GARCH model, is discussed below.

Now consider the random coefficient AR (1) process underlying the return shocks, ε_t:

$$\varepsilon_t = \phi_t \varepsilon_{t-1} + \eta_t \tag{2}$$

where $\phi_t \sim iid(0, \alpha)$, $\alpha \geq 0$, $\eta_t \sim iid(0, \omega)$, $\omega \geq 0$, $\eta_t = \varepsilon_t / \sqrt{h_t}$ is the standardized residual, with h_t defined below. [21] Tsay (1987) derived the ARCH (1) model of [22] Engle (1982) and [23] Bollerslev (1986) from Equation (2) as:

$$h_t \equiv E(\varepsilon_t^2 | I_{t-1}) = \omega + \alpha \varepsilon_{t-1}^2 \tag{3}$$

where h_t represents conditional volatility, and I_{t-1} is the information set available at time $t-1$. A lagged dependent variable, h_{t-1}, is typically added to Equation (3) to improve the sample fit:

$$h_t \equiv E(\varepsilon_t^2 | I_{t-1}) = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{4}$$

From the specification of Equation (2), it is clear that both ω and α should be positive, as they are the unconditional variances of two different stochastic processes.

Given the non-normality of the returns shocks, the Quasi-Maximum Likelihood Estimators (QMLE) of the parameters have been shown to be consistent and asymptotically normal in several papers. For example [12] Ling and McAleer (2003) showed that the QMLE for a generalized ARCH(*p,q*) (or GARCH(*p,q*)) is consistent if the second moment is finite. A sufficient condition for the QMLE of GARCH(1,1) in Equation (4) to be consistent and asymptotically normal is $\alpha + \beta < 1$.

In general, the proofs of the asymptotic properties follow from the fact that GARCH can be derived from a random coefficient autoregressive process. Ref. [13] McAleer et al. (2008) give a general proof of asymptotic normality for multivariate models that are based on proving that the regularity conditions satisfy the conditions given in [24] Jeantheau (1998) for consistency, and the conditions given in Theorem 4.1.3 in [25] Amemiya (1985) for asymptotic normality.

3.2. Multivariate Conditional Volatility

The multivariate extension of the univariate ARCH and GARCH models is given in [9] Baba et al. (1985) and [10] Engle and Kroner (1995) (for caveats regarding Full BEKK, see [15] Chang and McAleer (2017b)). In order to establish volatility spillovers in a multivariate framework, it is useful to define the multivariate extension of the relationship between the returns shocks and the standardized residuals, that is, $\eta_t = \varepsilon_t / \sqrt{h_t}$.

The multivariate extension of Equation (1), namely $y_t = E(y_t | I_{t-1}) + \varepsilon_t$, can remain unchanged by assuming that the three components are now $m \times 1$ vectors, where m is the number of financial assets. The multivariate definition of the relationship between ε_t and η_t is given as:

$$\varepsilon_t = D_t^{1/2} \eta_t \tag{5}$$

where $D_t = diag(h_{1t}, h_{2t}, \ldots, h_{mt})$ is a diagonal matrix comprising the univariate conditional volatilities.

Define the conditional covariance matrix of ε_t as Q_t. As the $m \times 1$ vector, η_t, is assumed to be *iid* for all m elements, the conditional correlation matrix of ε_t, which is equivalent to the conditional correlation matrix of η_t, is given by Γ_t. Therefore, the conditional expectation of (5) is defined as:

$$Q_t = D_t^{1/2} \Gamma_t D_t^{1/2} \tag{6}$$

Equivalently, the conditional correlation matrix, Γ_t, can be defined as:

$$\Gamma_t = D_t^{-1/2} Q_t D_t^{-1/2}. \tag{7}$$

Equation (6) is useful if a model of Γ_t is available for purposes of estimating Q_t, whereas (7) is useful if a model of Q_t is available for the purposes of estimating Γ_t.

Equation (6) is convenient for a discussion of volatility spillover effects, while both Equations (6) and (7) are instructive for a discussion of asymptotic properties. As the elements of D_t are consistent and asymptotically normal, the consistency of Q_t in (6) depends on the consistent estimation of Γ_t, whereas the consistency of Γ_t in (7) depends on the consistent estimation of Q_t. As both Q_t and Γ_t are products of matrices, with inverses in (7), neither the QMLE of Q_t nor Γ_t will be asymptotically normal based on the definitions given in Equations (6) and (7).

3.3. Diagonal BEKK

The Diagonal BEKK model can be derived from a vector random coefficient autoregressive process of order one, which is the multivariate extension of the univariate process given in Equation (2):

$$\varepsilon_t = \Phi_t \varepsilon_{t-1} + \eta_t \tag{8}$$

where ε_t and η_t are $m \times 1$ vectors, Φ_t is an $m \times m$ matrix of random coefficients, $\Phi_t \sim iid(0, A)$, A is positive definite, $\eta_t \sim iid(0, C)$, C is an $m \times m$ matrix.

Vectorization of a full matrix A to $vec\, A$ can have dimension as high as $m^2 \times m^2$, whereas vectorization of a symmetric matrix A to $vech\, A$ can have a smaller dimension of $m(m+1)/2 \times m(m+1)/2$.

In a case where A is a diagonal matrix, with $a_{ii} > 0$ for all $i = 1, \ldots, m$ and $|b_{jj}| < 1$ for all $j = 1, \ldots, m$, so that A has dimension $m \times m$, [13] McAleer et al. (2008) showed that the multivariate extension of GARCH(1,1) from Equation (8) is given as the Diagonal BEKK model, namely:

$$Q_t = CC' + A\varepsilon_{t-1}\varepsilon'_{t-1}A' + BQ_{t-1}B' \tag{9}$$

where A and B are both diagonal matrices, though the last term in Equation (9) need not come from an underlying stochastic process. The diagonality of the positive definite matrix A is essential for matrix multiplication as $\varepsilon_{t-1}\varepsilon'_{t-1}$ is an $m \times m$ matrix; otherwise, Equation (9) could not be derived from the vector random coefficient autoregressive process in Equation (8).

3.4. Full, Triangular and Hadamard BEKK

The full BEKK model in [9] Baba et al. (1985) and [10] Engle and Kroner (1995), who do not derive the model from an underlying stochastic process, is presented as:

$$Q_t = CC' + A\varepsilon_{t-1}\varepsilon'_{t-1}A' + BQ_{t-1}B' \tag{10}$$

except that A and (possibly) B in Equation (10) are now both full matrices, rather than the diagonal matrices that were derived in Equation (9) by using the stochastic process in Equation (8). The full BEKK model can be replaced by the triangular or Hadamard (element-by-element multiplication) BEKK models, with similar problems of identification and (lack of) existence.

A fundamental technical problem is that the full, triangular, and Hadamard BEKK models cannot be derived from any known underlying stochastic processes, which means that there are no regularity conditions (except by assumption) for checking the internal consistency of the alternative models, and consequently no valid asymptotic properties of the QMLE of the associated parameters (except by assumption).

Moreover, as the number of parameters in a full BEKK model can be as much as $3m(m + 1)/2$, the "curse of dimensionality" will be likely to arise, which means that the convergence of the estimation algorithm can become problematic and less reliable when there is a large number of parameters to be estimated.

As a matter of empirical fact, the estimation of the full BEKK can be problematic even when m is as low as five financial assets. Such computational difficulties do not arise for the Diagonal BEKK model. Convergence of the estimation algorithm is more likely when the number of commodities is less than four, though this is nevertheless problematic in terms of interpretation.

Therefore, in the empirical analysis, in order to investigate volatility spillover effects, the solution is to use the Diagonal BEKK model for estimation. A quasi likelihood ratio (QLR) test is developed to test the multivariate conditional volatility Diagonal BEKK model in Equation (9) (where A and B are both diagonal matrices), which has valid regularity conditions and asymptotic properties, against the alternative Full BEKK model in Equation (10) (where A and B in are now both full matrices), which has valid regularity conditions and asymptotic properties only under the null hypothesis of zero off-diagonal elements. The quasi likelihood ratio (QLR) test of the null Diagonal BEKK model against the alternative of the Full BEKK model does not yet seem to have been presented in the literature.

3.5. Granger Causality, Volatility Spillovers, and Optimal Hedge Ratios

[13] McAleer et al. (2008) showed that the QMLE of the parameters of the Diagonal BEKK model were consistent and asymptotically normal, so that standard statistical inference on testing hypotheses is valid. Moreover, as Q_t in (9) can be estimated consistently, Γ_t in Equation (7) can also be estimated consistently.

The Diagonal BEKK model is given as Equation (9), where the matrices A and B are given as:

$$
A = \begin{bmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{mm} \end{bmatrix}, B = \begin{bmatrix} b_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & b_{mm} \end{bmatrix} \tag{11}
$$

The Diagonal BEKK model permits a test of Co-volatility Spillover effects, which is the effect of a shock in commodity j at $t-1$ on the subsequent co-volatility between j and another commodity at t. Given the Diagonal BEKK model, as expressed in Equations (9) and (10), the subsequent co-volatility must only be between commodities j and i at time t. [19] Chang et al. (2015) define Full and Partial Volatility and Covolatility Spillovers in the context of Diagonal and Full BEKK models. Volatility spillovers are defined as the delayed effect of a returns shock in one asset on the subsequent volatility or covolatility in another asset. Therefore, a model relating Q_t to returns shocks is essential, and this will be addressed in the following sub-section. Spillovers can be defined in terms of full volatility spillovers and full covolatility spillovers, as well as partial covolatility spillovers, as follows, for $i, j, k = 1, \ldots, m$:

(1) **Full volatility spillovers:**

$$
\partial Q_{iit} / \partial \varepsilon_{kt-1}, k \neq i; \tag{12}
$$

(2) **Full covolatility spillovers:**

$$
\partial Q_{ijt} / \partial \varepsilon_{kt-1}, i \neq j, \ k \neq i, j; \tag{13}
$$

(3) **Partial covolatility spillovers:**

$$
\partial Q_{ijt} / \partial \varepsilon_{kt-1}, i \neq j, \ k = either \ i \ or \ j. \tag{14}
$$

Full volatility spillovers occur when the returns shock from financial asset k affects the volatility of a different financial asset i.

Full covolatility spillovers occur when the returns shock from financial asset k affects the covolatility between two different financial assets, i and j.

Partial covolatility spillovers occur when the returns shock from financial asset k affects the covolatility between two financial assets, i and j, one of which can be asset k.

When $m = 2$, only spillovers (1) and (3) are possible as full covolatility spillovers depend on the existence of a third financial asset.

This leads to the definition of a **Co-volatility Spillover Effect** as:

$$\frac{\partial H_{ij,t}}{\partial \varepsilon_{j,t-1}} = a_{ii} \times a_{jj} \times \varepsilon_{i,t-1}, \ i \neq j$$

As $a_{ii} > 0$ for all i, a test of the co-volatility spillover effect is given as a test of the null hypothesis:

$$H_0 : a_{ii}a_{jj} = 0$$

which is a test of the significance of the estimate of $a_{ii}a_{jj}$ in the following co-volatility spillover effect, as $\varepsilon_{i,t-1} \neq 0$:

$$\frac{\partial H_{ij,t}}{\partial \varepsilon_{j,t-1}} = a_{ii}a_{jj}\varepsilon_{i,t-1}, \ i \neq j.$$

If H_0 is rejected against the alternative hypothesis, $H_1 : a_{ii}a_{jj} \neq 0$, there is a spillover from the returns shock of commodity j at $t - 1$ to the co-volatility between commodities i and j at t that depends only on the returns shock of commodity i at $t - 1$. It should be emphasized that the returns shock of commodity j at $t - 1$ does not affect the co-volatility spillover of commodity j on the co-volatility between the commodities i and j at t. Moreover, spillovers can and do vary for each observation $t - 1$, so that the empirical results average co-volatility spillovers will be presented, based on the average return shocks over the sample period.

Granger (1980) [2] causality is based on the following vector AR (VAR(m,n)) models:

$$x(t) = a_0 + a_1 x(t - 1) + \cdots + a_m x(t - m) + b_1 y(t - 1) + \cdots + b_n y(t - n) + u(t), \tag{15}$$

$$y(t) = c_0 + c_1 y(t - 1) + \cdots + c_n y(t - n) + d_1 x(t - 1) + \cdots + d_m x(t - m) + v(t) \tag{16}$$

The null hypothesis of Granger non-causality of $y(t - 1)$ on $x(t)$ is based on testing:

$$H_0: b_i = 0 \text{ for all } i = 1, \cdots n$$

in Equation (12), while the null hypothesis of Granger non-causality of $x(t)$ on $y(t - 1)$ is based on testing:

$$H_0: d_i = 0 \text{ for all } i = 1, \cdots m$$

in Equation (13). In the empirical analysis, $m = n = 1$ as daily data are used.

For the multivariate conditional mean returns equation:

$$y_{it} = E(y_{it}|I_{t-1}) + \varepsilon_{it}, i = 1, 2, \cdots, m \tag{17}$$

the bivariate random coefficient autoregressive process for ε_{it} is given as:

$$\varepsilon_{it} = \phi_{it}\varepsilon_{it-1} + \phi_{jt}\varepsilon_{jt-1} + \eta_{it}, \ i \neq j \tag{18}$$

where $\phi_{it} \sim iid(0, \alpha_i)$, $\alpha_i \geq 0$, $\phi_{jt} \sim iid(0, \alpha_j)$, $\alpha_j \geq 0$, $\eta_{it} \sim iid(0, \omega_i)$, $\omega_i \geq 0$, $\eta_{it} = \varepsilon_{it}/\sqrt{h_{it}}$ is the standardized residual, h_{it} is the conditional volatility obtained by setting $\phi_{jt} = 0$ in bivariate Equation (15):

$$\varepsilon_{it} = \phi_{it}\varepsilon_{it-1} + \eta_{it}$$

$$E\left(\varepsilon_{it}^2 \middle| I_{t-1}\right) \equiv h_{it} = \omega_i + \alpha_i \varepsilon_{it-1}^2$$

Adding another commodity, as in the bivariate Equation (15), gives:

$$\varepsilon_{it} = \phi_{it}\varepsilon_{it-1} + \phi_{jt}\varepsilon_{jt-1} + \eta_{it}, \ i \neq j$$

$$E\left(\varepsilon_{it}^2 \middle| I_{t-1}\right) \equiv h_{it} = \omega_i + \alpha_i\varepsilon_{it-1}^2 + \alpha_j\varepsilon_{jt-1}^2$$

while adding first-order lags of h_{it} and h_{jt} gives:

$$h_{it} = \omega_i + \alpha_i\varepsilon_{it-1}^2 + \alpha_j\varepsilon_{jt-1}^2 + \beta_i h_{it-1} + \beta_j h_{jt-1}$$

where

$$\alpha_i \geq 0, \alpha_j \geq 0, \beta_i \in (-1, 1), \beta_j \in (-1, 1)$$

The null hypothesis of non-causality in volatility is given as a test of:

$$H_0: \alpha_j = \beta_j = 0$$

Based on the empirical results, dynamic hedging strategies using optimal hedge ratios will be suggested to analyse market fluctuations in the spot and futures returns and volatility of carbon emissions, crude oil, and coal prices.

Using the hedge ratio: $R_{H,t} = R_{S,t} - \gamma_t R_{F,t}$ and its variance, namely:

$$var(R_{H,t}|\Omega_{t-1}) = var(R_{S,t}|\Omega_{t-1}) - 2\gamma_t cov(R_{S,t}, R_{F,t}|\Omega_{t-1}) + \gamma_t^2 \ var(R_{F,t}|\Omega_{t-1})$$

the optimal hedge ratio is given as:

$$\gamma_t|\Omega_{t-1} = cov(R_{S,t}, R_{F,t}|\Omega_{t-1})/var(R_{F,t}|\Omega_{t-1})$$

An extension of the recent research on the realized matrix-exponential stochastic volatility with asymmetry, long memory, and spillovers, in [26] Asai, Chang and McAleer (2017), to multivariate conditional volatility models, especially the use of the matrix-exponential transformation to ensure a positive definite covariance matrix, will enable a significant extension of the univariate Granger causality tests to be extended to multivariate Granger causality tests. This would be a novel extension of the paper.

4. Unit Root Tests

In order to evaluate the characteristics of the data, we investigate whether shocks to a series are temporary or permanent in nature. We will use the ADF test ([27] Dickey and Fuller, 1979; [28] Dickey and Fuller, 1982; [29] Said and Dickey, 1984), DF-GLS test ([30] Elliott et al., 1996), and the KPSS test ([31] Kwiatkowski et al., 1992) to test for unit roots in the individual returns series. The ADF and DF-GLS tests are designed to test for the null hypothesis of a unit root, while the KPSS test is used for the null hypothesis of stationarity.

In Table 3, based on the ADF test results, the large negative values in all of the cases indicate a rejection of the null hypothesis of unit roots at the 1% level. Based on the KPSS test, the small positive values in all of the cases do not reject the null hypothesis of stationary at the 1% level. For the DF-GLS test, the futures returns of carbon emissions and of coal in the EU, and the spot returns of carbon emissions in the USA, reject the null hypothesis of unit roots at the 1% level. However, the results of the coal and oil spot returns do not reject the null hypothesis. It should be noted that, for the USA, a relatively small sample size of 357 observations is used.

Table 3. Unit Root Tests 2 April 2008–19 May 2017 for EU 6 January 2016–19 May 2017 for USA.

Variables	ADF	DF-GLS	KPSS
$EUcarbon_{fr}$	−37.79 *	−3.09 *	0.05 *
$EUcoal_{fr}$	−35.48 *	−10.34 *	0.12 *
$EUoil_{fr}$	−51.97 *	−1.53	0.10 *
$UScarbon_{sr}$	−10.64 *	−1.46	0.06 *
$UScoal_{sr}$	−19.30 *	−0.43	0.18 *
$USoil_{sr}$	−20.96 *	−0.78	0.07 *

* Denotes the null hypothesis of a unit root is rejected at 1%.

5. Granger Causality and Spillovers in Returns and Volatilities

Table 4 reports the results for the [2] Granger (1980) causality and spillover tests in returns, with one lag being used throughout the empirical analysis. There is no evidence of bidirectional Granger causality between carbon and coal futures for the EU. However, oil futures in the EU has a causal effect on carbon emissions futures in the EU. For the USA, the carbon emissions spot has a causal effect on the coal spot, as well as on the oil spot.

Table 4. Granger Causality Test for Returns 2 April 2008–19 May 2017 for EU 6 January 2016–19 May 2017 for USA.

Variables		Lags	Outcome			Null Hypothesis			
						A Does Not Cause B		B Does Not Cause A	
A	B					F-Test	p-Value	F-Test	p-Value
$EUcarbon_{fr}$	$EUcoal_{fr}$	1	$EUcarbon_{fr}$	←	$EUcoal_{fr}$	0.6190	0.4315	5.7112	0.0169
$EUcarbon_{fr}$	$EUoil_{fr}$	1	$EUcarbon_{fr}$	←	$EUoil_{fr}$	0.2337	0.6289	4.1837	0.0409
$UScarbon_{sr}$	$UScoal_{sr}$	1	$UScarbon_{sr}$	→	$UScoal_{sr}$	4.6809	0.0312	0.9142	0.3397
$UScarbon_{sr}$	$USoil_{sr}$	1	$UScarbon_{sr}$	→	$USoil_{sr}$	5.1310	0.0241	0.0075	0.9313

Estimates of the DBEKK and Full BEKK models for EU Carbon, Coal, and Oil Futures returns are given in Table 5. The estimates of the weighting coefficients, A(1,1), are similar for the two models, but the estimates of the weighting coefficients A(2,2) and A(3,3) are different for the two models. Similar comments apply to the estimates of the matrix stability coefficients, B(1,1), B(2,2), and B(3,43), respectively.

Table 5. DBEKK and Full BEKK for EU Carbon, Coal, and Oil Futures 2 April 2008–19 May 2017.

DBEKK	C				A			B		
$CARBON_{fr}$	0.379 *** (0.055)	0.024 ** (0.010)	0.128 *** (0.024)	0.311 *** (0.025)				0.947 *** (0.009)		
$COAL_{fr}$		0.088 *** (0.010)	0.022 (0.075)		0.118 *** (0.007)			0.991 *** (0.001)		
OIL_{fr}			0.000 (0.077)			−0.205 *** (0.013)			−0.977 *** (0.003)	
Full BEKK	**C**			**A**			**B**			
$CARBON_{fr}$	0.435 *** (0.055)	−0.067 * (0.038)	0.077 (0.072)	0.331 *** (0.023)	−0.014*** (0.004)	0.007 (0.006)	0.936 *** (0.009)	0.009 (0.007)	−0.005 (0.010)
$COAL_{fr}$		0.000 (0.068)	0.000 (0.103)	0.037 (0.029)	−0.086 *** (0.011)	0.120 *** (0.017)	0.274 *** (0.036))	0.737 *** (0.015)	1.110 *** (0.023)
OIL_{fr}		−0.000 (0.101)	−0.104 *** (0.026)	−0.032 ** (0.013)	−0.168 *** (0.010)	−0189 *** (0.024)	−0.052 *** (0.011)	0.054 *** (0.015)	

1. $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$, $C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$. 2. Standard errors are in parentheses, *** denotes significant at 1%, ** denotes significant at 5%, * denotes significant at 10%.

Given the differences in two of the three weighting coefficients in A in Table 5, it is not particularly surprising that the quasi likelihood ratio (QLR) test in Table 6 of the null hypothesis, DBEKK, against the alternative hypothesis, Full BEKK, leads to the rejection of the null hypothesis that the off-diagonal elements of A and B are zero. The calculated chi-squared statistic with 12 degrees of freedom, at 34.32, is greater than the critical value of 26.22 at the 1% level. Therefore, DBEKK is rejected, but Full BEKK is not appropriate as it is valid only under the null hypothesis of zero off-diagonal coefficients for the weighting matrix A and for the stability matrix B. In short, the Diagonal BEKK model is rejected, but the full BEKK model is not an appropriate replacement.

Table 6. Quasi Likelihood Ratio (QLR) Test of DBEKK and Full BEKK for EU Futures 2 April 2008–19 May 2017.

Quasi Log-likelihood value for DBEKK $-14{,}815.88$
Quasi Log-likelihood value for Full BEKK $-14{,}798.72$
QLR test statistic 34.32
Critical value at 1% with 12 df 26.22

Estimates of the DBEKK and Full BEKK models for US Carbon, Coal, and Oil Spot returns are given in Table 7. The estimates of the three weighting coefficients, A(1,1), A(2,2), and A(3,3), are reasonably similar for the two models, as are the estimates of the stability coefficients B(1,1) and B(2,2), though the estimates of B(3,3) are different for the two models.

In view of the similarities in the estimates of the three weighting coefficients in A in Table 7, the quasi likelihood ratio (QLR) test in Table 8 of the null hypothesis, DBEKK, against the alternative hypothesis, Full BEKK, leads to the non-rejection of the null hypothesis that the off-diagonal elements of A and B are zero, as compared with the outcome in Table 6. The calculated chi-squared statistic with 12 degrees of freedom, at 22.18, is less than the critical value of 26.22 at the 1% level. Therefore, DBEKK is not rejected against Full BEKK, which is valid only under the null hypothesis of zero off-diagonal coefficients for the weighting matrix A and stability matrix B. In short, the Diagonal BEKK model is empirically supported by the data.

Table 7. DBEKK and Full BEKK for US Carbon, Coal, and Oil Spot 6 January 2016–19 May 2017.

DBEKK	C			A			B		
CARBON$_{sr}$	0.854 ***	−0.276	0.129	0.707 ***			0.757 ***		
	(0.105)	(0.294)	(0.332)	(0.073)			(0.038)		
COAL$_{sr}$		0.256	0.299 **		−0.199 ***			0.972 ***	
		(0.314)	(0.154)		(0.034)			(0.008)	
OIL$_{sr}$			0.000			−0.222 ***			−0.964 ***
			(1.029)			(0.0035)			(0.010)
Full BEKK	**C**			**A**			**B**		
CARBON$_{sr}$	0.772 ***	0.119	0.685 ***	0.632 ***	−0.023	−0.077	0.791 ***	0.004	−0.034
	(0.092)	(0.606)	(0.178)	(0.054)	(0.089)	(0.064)	(0.025)	(0.112)	(0.063)
COAL$_{sr}$		0.000	0.000	0.002	−0.320 ***	0.036	−0.042	0.900 ***	0.578 ***
		(0.528)	(0.715)	(0.033)	(0.058)	(0.041)	(0.046)	(0.056)	(0.044)
OIL$_{sr}$			0.000	−0.028	−0.072	−0.252 ***	0.010	−1.267 ***	0.140 **
			(0.721)	(0.049)	(0.092)	(0.060)	(0.080)	(0.074)	(0.082)

1. $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$, $C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$. 2. Standard errors are in parentheses, *** denotes significant at 1%, ** denotes significant at 5%.

Table 8. QLR Test of DBEKK and Full BEKK for US Spot 6 January 2016–19 May 2017.

Quasi Log-likelihood value for DBEKK -2499.27
Quasi Log-likelihood value for Full BEKK -2488.18
QLR test statistic 22.18
Critical value at 1% with 12 df 26.22

In light of the discussion based on Equations (14), partial co-volatility spillovers with DBEKK are presented in Table 9. Based on the estimates of the weighting matrix A, six of the eight partial co-volatility spillovers are negative, which means that a shock in one of carbon emission, coal, or oil will have a one-period delayed negative impact on the conditional correlation between itself and one of the other two commodities. Two of the eight partial co-volatility spillovers are positive, so an opposite effect will be observed.

Table 9. Partial Co-volatility Spillovers with DBEKK for EU and USA 2 April 2008–19 May 2017 for EU 6 January 2016–19 May 2017 for USA.

Market	$\left(\frac{\partial H_{ij,t}}{\partial \varepsilon_{k,t-1}}\right)$		Average Co-Volatility Spillovers
EU	$j = k = \text{coal}_{fr},$	$i = \text{carbon}_{fr}$	$-0.001 = -0.030 \times 0.311 \times 0.118$
	$j = k = \text{carbon}_{fr},$	$i = \text{coal}_{fr}.$	$0.001 = 0.026 \times 0.311 \times 0.118$
	$j = k = \text{oil}_{fr},$	$i = \text{carbon}_{fr}$	$0.002 = -0.030 \times 0.311 \times -0.205$
	$j = k = \text{carbon}_{fr},$	$i = \text{oil}_{fr}.$	$0.001 = -0.023 \times 0.311 \times -0.205$
USA	$j = k = \text{coal}_{sr},$	$i = \text{carbon}_{sr}$	$0.020 = -0.140 \times 0.707 \times -0.199$
	$j = k = \text{carbon}_{sr},$	$i = \text{coal}_{sr}$	$-0.002 = 0.012 \times 0.707 \times -0.199$
	$j = k = \text{oil}_{sr},$	$i = \text{carbon}_{sr}$	$0.022 = -0.140 \times 0.707 \times -0.222$
	$j = k = \text{carbon}_{sr},$	$i = \text{oil}_{sr}$	$0.003 = -0.022 \times 0.707 \times -0.222$

Co-volatility Spillovers: $\frac{\partial H_{ij,t}}{\partial \varepsilon_{k,t-1}} = a_{ii}a_{jj}\varepsilon_{i,t-1}$.

Given the discussion based on Equations (12) and (13), full co-volatility spillovers with DBEKK are presented in Table 10. Based on the estimates of the weighting matrix A, two of the six full co-volatility spillovers are negative, which means that a shock in one of carbon emission, coal, or oil will have a one-period delayed negative impact on the conditional correlation between two of the other commodities. Two of the six full co-volatility spillovers are positive, so an opposite effect will be observed, while two of the six full co-volatility spillovers are zero, in which case there will be no spillovers.

Table 10. Full Co-volatility Spillovers with Full BEKK for EU and USA 2 April 2008–19 May 2017 for EU 6 January 2016–19 May 2017 for USA.

Market	$\left(\frac{\partial H_{ij,t}}{\partial \varepsilon_{k,t-1}}\right)$			Co-Volatility Spillovers
EU	$j = \text{coal}_{fr},$	$i = \text{carbon}_{fr}$	$k = \text{oil}_{fr}$	-0.001
	$j = \text{oil}_{fr},$	$i = \text{carbon}_{fr}$	$k = \text{coal}_{fr},$	0
	$j = \text{coal}_{fr},$	$i = \text{oil}_{fr}$	$k = \text{carbon}_{fr}$	0.001
USA	$j = \text{coal}_{sr},$	$i = \text{carbon}_{sr}$	$k = \text{oil}_{sr}$	-0.002
	$j = \text{oil}_{sr},$	$i = \text{carbon}_{sr}$	$k = \text{coal}_{sr}$	0.004
	$j = \text{coal}_{sr},$	$i = \text{oil}_{sr}$	$k = \text{carbon}_{sr}$	0

Co-volatility Spillovers: $\frac{\partial H_{ij,t}}{\partial \varepsilon_{k,t-1}} = a_{ii}a_{jk}\varepsilon_{i,t-1} + a_{ij}a_{jk}\varepsilon_{j,t-1} + a_{ik}a_{ji}\varepsilon_{i,t-1} + a_{ik}a_{jj}\varepsilon_{j,t-1} + 2a_{ik}a_{jk}\varepsilon_{k,t-1}$. A co-volatility spillover of 0 is to three decimal places.

The results for full co-volatility spillovers in Table 10 are not as clear or as helpful as in the case of the partial co-volatility spillovers in Table 9, as the estimates of the off-diagonal elements in the weighting matrix A are not especially large.

The unconditional and conditional volatility of carbon, coal, and oil futures returns for the EU are shown in Figure 4a–f, while the unconditional and conditional volatility of carbon, coal, and oil spot returns for the USA are shown in Figure 5a–f. The conditional volatility estimates are forecasts of the unconditional volatilities. Both figures show that there is a significant difference between the conditional and unconditional volatilities. As one of the purposes of the paper is to use conditional volatilities to forecast optimal hedge ratios for the various spot and futures returns, any differences between the unconditional and conditional volatilities is based on the unconditional volatilities being unpredictable as compared to the conditional volatilities.
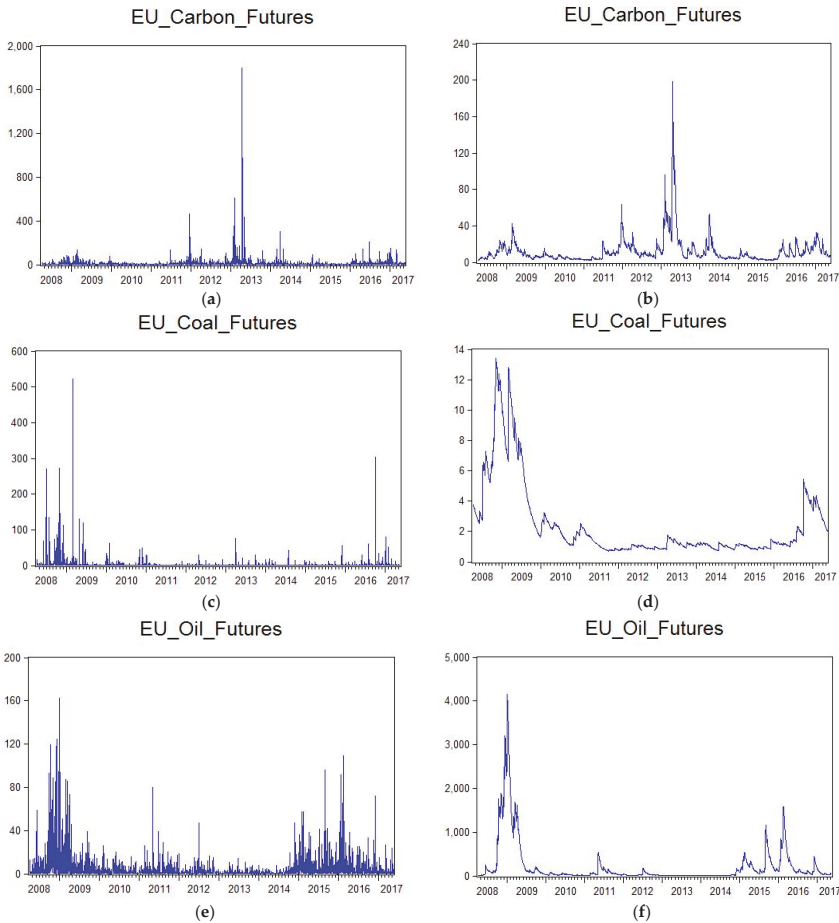
Figure 4. Unconditional (**a**,**c**,**e**) and Conditional (**b**,**d**,**f**) Volatility of Carbon, Coal, and Oil Futures for EU 2 April 2008–19 May 2017.
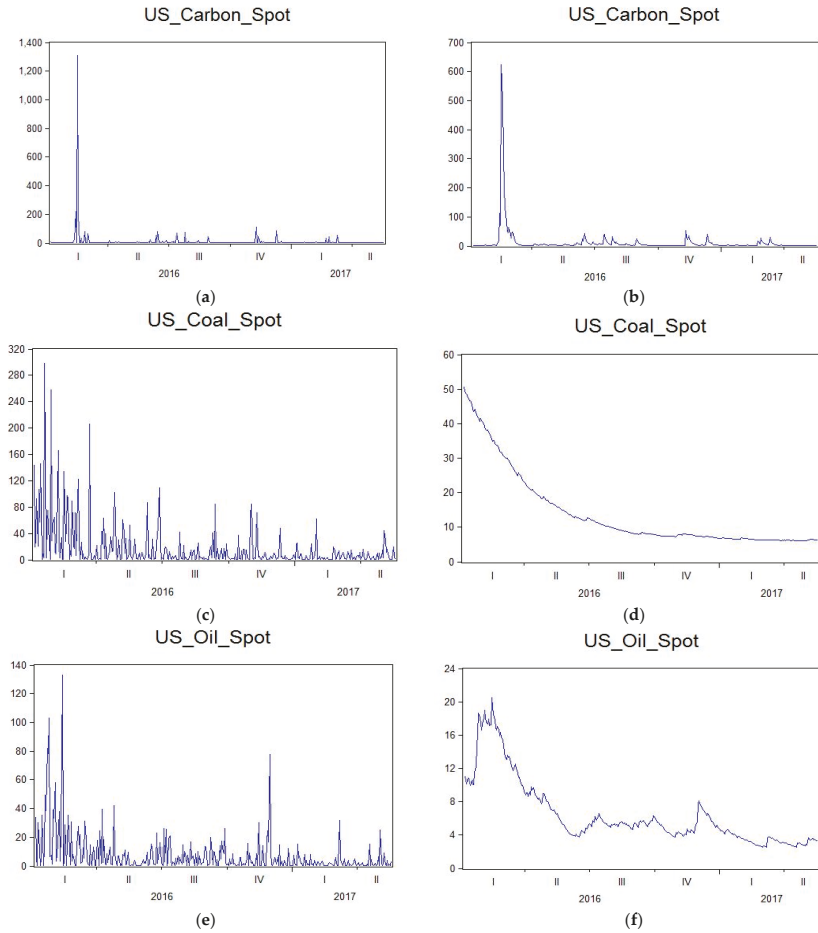
Figure 5. Unconditional (**a**,**c**,**e**) and Conditional (**b**,**d**,**f**) Volatility of Carbon, Coal, Oil Spot for USA 6 January 2016–19 May 2017.

The conditional co-volatility correlations for carbon, coal, and oil futures returns for the EU are shown in Figure 6a–f, while the conditional co-volatility correlations for carbon, coal, and oil spot returns for the USA are shown in Figure 7a–f. Both of the figures show that there are substantial differences in the correlations of conditional co-volatility across the two markets and time periods for carbon, coal, and oil futures returns.
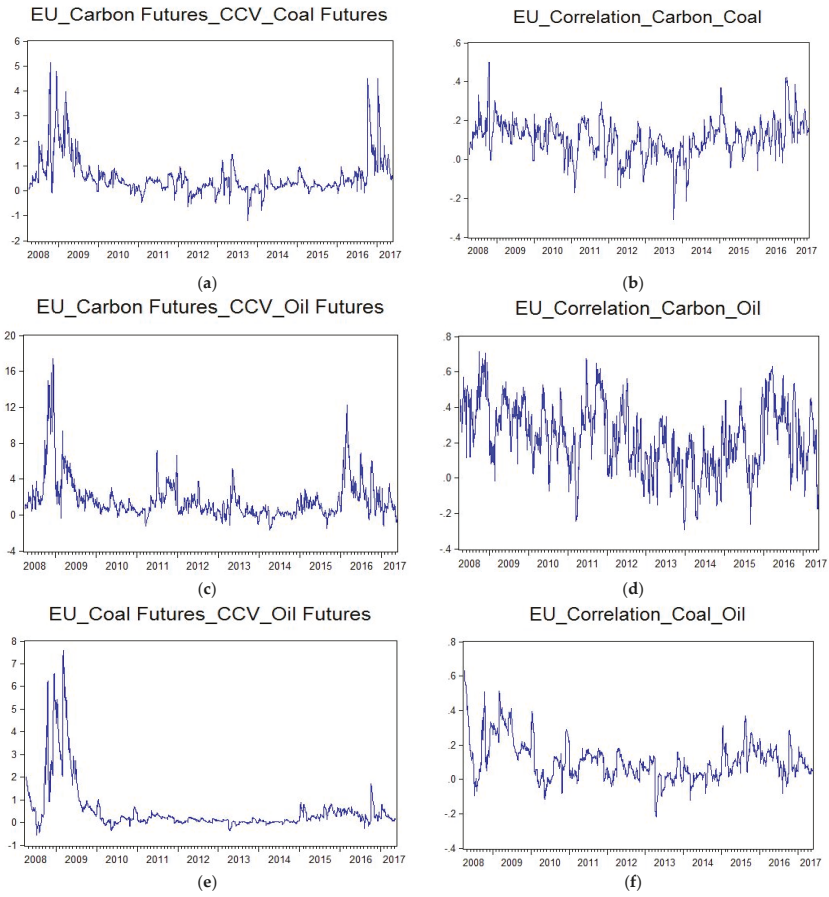
Figure 6. Conditional Co-volatility (**a**,**c**,**e**) and Correlations (**b**,**d**,**f**) for Carbon, Coal, and Oil Futures for EU 2 April 2008–18 May 2017.

Figure 7. *Cont.*

US_Carbon Spot_CCV_Oil Spot

US_Correlation_Carbon_Oil

(c)

(d)

US_Coal Spot_CCV_Oil Spot

US_Correlation_Coal_Oil

(e)

(f)
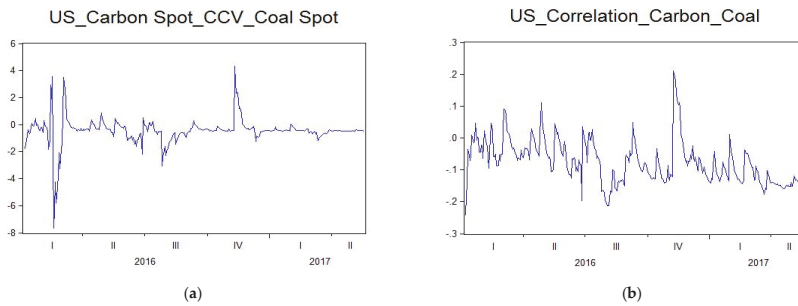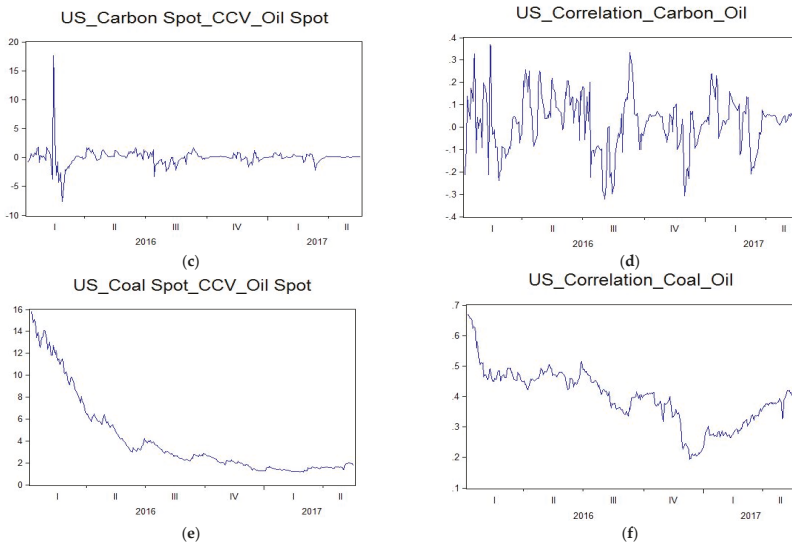
Figure 7. Conditional Co-volatility (**a**,**c**,**e**) and Correlations (**b**,**d**,**f**) for Carbon, Coal, Oil Spot for USA 6 January 2016–18 May 2017.

The optimal hedge ratios for carbon, coal, and oil futures returns for the EU, and optimal hedge ratios for carbon, coal, and oil spot returns for the USA, are given in Figures 8a–f and 9a–f, respectively. The hedge ratios show how the covariances in returns between two assets changes relative to the variance of the hedging instrument. Both figures show that there is substantial variation in the optimal hedge ratios, so that the futures and spot prices of carbon emissions, coal, and oil should be considered contemporaneously and simultaneously in a portfolio that links the prices, returns, and volatilities of carbon emissions to the use of fossil fuels.

EU_Hedge Carbon with Coal

EU_Hedge Carbon with Oil

(a)

(b)

EU_Hedge Coal with Carbon

EU_Hedge Oil with Carbon

(c)

(d)

Figure 8. *Cont.*

Figure 8. Optimal Hedge Ratios for Carbon (**a**,**b**), Coal (**c**,**e**), and Oil (**d**,**f**) Futures for EU 2 April 2008–19 May 2017.
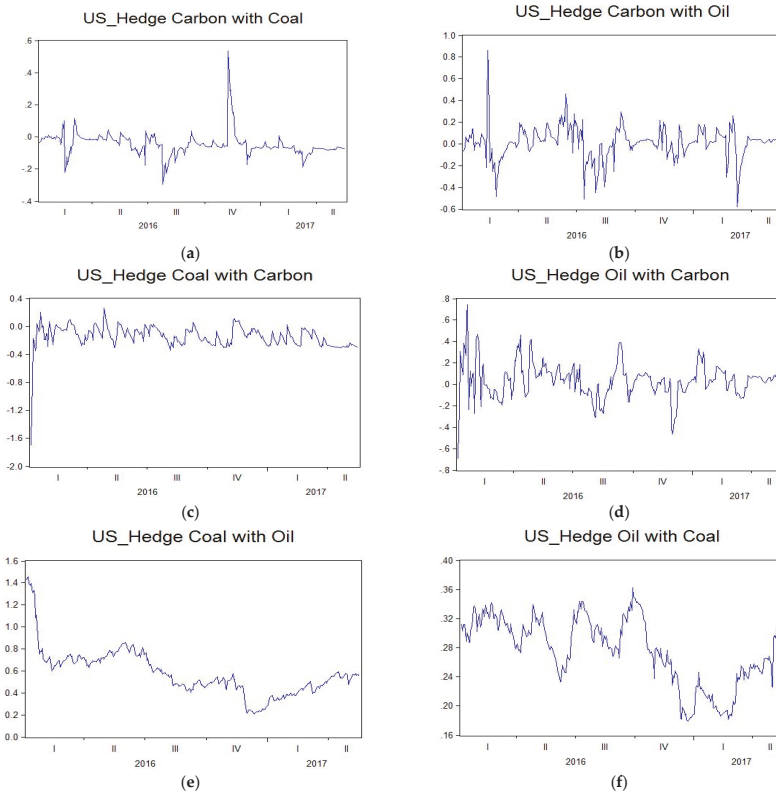
Figure 9. Optimal Hedge Ratios for Carbon (**a**,**b**), Coal (**c**,**e**), and Oil (**d**,**f**) Spot for USA 6 January 2016–18 May 2017.

Finally, Figure 10a–d show the optimal hedge ratios for carbon futures returns for the EU and both coal and oil spot returns for the USA. In all cases, the optimal hedge ratios vary substantially, which suggests that it would be sensible to use both markets to hedge carbon emission futures returns in the EU against both coal and oil spot price returns in the USA.
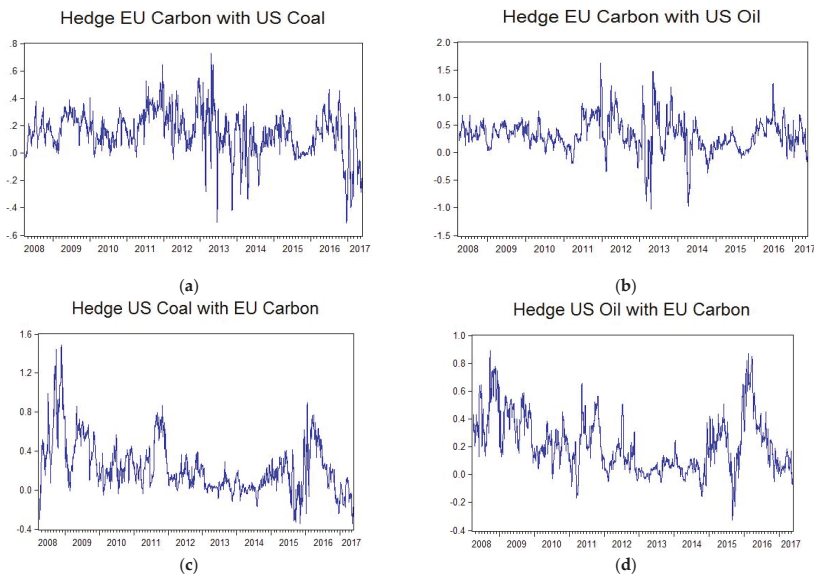
Figure 10. Optimal Hedge Ratios for Carbon (**a**,**b**) Futures of EU, and Coal (**c**) and Oil (**d**) Spot of USA 2 April 2008–18 May 2017.

6. Concluding Remarks

The paper discussed recent research that showed the efforts to limit climate change have been focusing on the reduction of carbon dioxide emissions over other greenhouse gases or air pollutants. Many countries have paid great attention to carbon emissions in order to improve air quality and public health. The largest source of carbon emissions from human activities in many countries in Europe and around the world has been from burning fossil fuels. The prices of both fuel and carbon emissions can and do have simultaneous and contemporaneous effects on each other.

Owing to the importance of carbon emissions and their interconnection to the prices, financial returns, and associated volatilities of fossil fuels, and the possibility of Granger causality in spot and futures prices, returns, and volatility of carbon emissions, it is not surprising that crude oil and coal, and their interactions with carbon emission prices, returns and volatility, have recently become very important for public policy and an associated research topic.

For the USA, daily spot and futures prices are available for crude oil and coal, but there are no daily spot or futures prices for carbon emissions. For the EU, there are no daily spot prices for coal or carbon emissions, but there are daily futures prices for crude oil, coal, and carbon emissions. For this reason, daily prices were used to analyse Granger causality and volatility spillovers in spot and futures prices of carbon emissions, crude oil, and coal.

A quasi likelihood ratio (QLR) test was developed to test the multivariate conditional volatility Diagonal BEKK model, which has valid regularity conditions and asymptotic properties, against the alternative Full BEKK model, which has valid regularity conditions and asymptotic properties only under the null hypothesis of zero off-diagonal elements. In short, Full BEKK has no desirable mathematical or statistical properties, except either under the null hypothesis of zero off-diagonal elements of the weighting matrix, or simply by assumption.

In the empirical analysis, DBEKK was rejected against the Full BEKK model for EU futures returns, but DBEKK was not rejected against Full BEKK for US spot returns. Therefore, further work would seem to be required for DBEKK in the case of EU futures returns, whereas DBEKK is empirically supported by the data for US spot returns.

Dynamic hedging strategies using optimal hedge ratios were suggested to analyse market fluctuations in the spot and futures returns and volatility of carbon emissions, crude oil, and coal prices. It was suggested that the futures and spot prices of carbon emissions, coal, and oil should be considered contemporaneously and simultaneously in a portfolio that links the prices, returns, and volat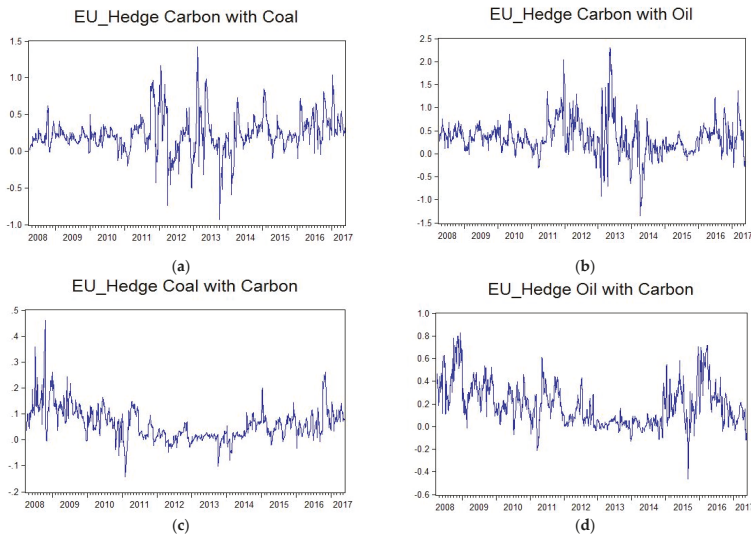ilities of carbon emissions to the use of fossil fuels. It would also b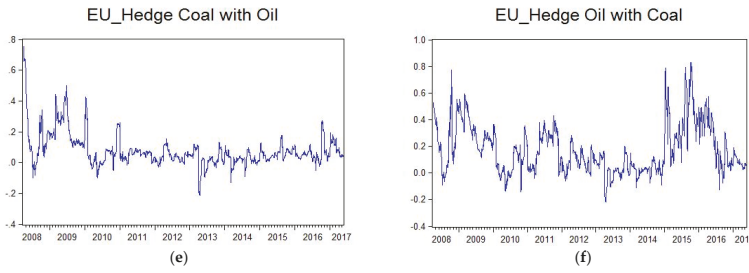e sensible to use the prices in both markets to hedge carbon emission price returns in the EU against both coal and oil spot price returns in the USA.

Author Contributions: Chia-Lin Chang and Michael McAleer conceived and designed the ideas. McAleer wrote the first draft. Guangdong Zuo collected the data and estimated the models. Chang and McAleer analysed the data and discussed the empirical results. McAleer revised the paper.

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

References

1. Rogelj, J.; Meinshausen, M.; Sedláček, J.; Knutti, R. Implications of potentially lower climate sensitivity on climate projections and policy. *Environ. Res. Lett.* **2014**, *9*, 3–10. [CrossRef]
2. Granger, C.W.J. Testing for causality: A personal viewpoint. *J. Econ. Dyn. Control* **1980**, *2*, 329–352. [CrossRef]
3. Ramos, S.; Veiga, H. (Eds.) The Interrelationship between Financial and Energy Markets. In *Lecture Notes in Energy*; Springer: Berlin, Germany, 2014; Volume 54.
4. Sawik, B.; Faulin, J.; Pérez-Bernabeu, E. A multicriteria analysis for the green VRP: A case discussion for the distribution problem of a Spanish retailer. *Transp. Res. Proced.* **2017**, *22*, 305–313. [CrossRef]
5. Sawik, B.; Faulin, J.; Pérez-Bernabeu, E. Multi-objective traveling salesman and transportation problem with environmental aspects. In *Applications of Management Science*; Lawrence, K.D., Kleinman, G., Eds.; Emerald: Bingley, UK, 2017; Volume 18, pp. 21–56.
6. Sawik, B.; Faulin, J.; Pérez-Bernabeu, E. Selected multi-criteria green vehicle routing problems. In *Applications of Management Science*; Lawrence, K.D., Kleinman, G., Eds.; Emerald: Bingley, UK, 2017; Volume 18, pp. 57–84.
7. Hafner, C.M.; Herwartz, H. A Lagrange multiplier test for causality in variance. *Econ. Lett.* **2006**, *93*, 137–141. [CrossRef]
8. Chang, C.-L.; McAleer, M. A simple test for causality in volatility. *Econometrics* **2017**, *5*, 15. [CrossRef]
9. Baba, Y.; Engle, R.F.; Kraft, D.; Kroner, K.F. *Multivariate Simultaneous Generalized ARCH*; Unpublished manuscript; Department of Economics, University of California: San Diego, CA, USA, 1985.
10. Engle, R.F.; Kroner, K.F. Multivariate simultaneous generalized ARCH. *Econom. Theory* **1995**, *11*, 122–150. [CrossRef]
11. Bollerslev, T. Modelling the coherence in short-run nominal exchange rate: A multivariate generalized ARCH approach. *Rev. Econ. Stat.* **1990**, *72*, 498–505. [CrossRef]
12. Ling, S.; McAleer, M. Asymptotic theory for a vector ARMA-GARCH model. *Econom. Theory* **2003**, *19*, 280–310. [CrossRef]
13. McAleer, M.S.; Hoti, F. Chan Structure and asymptotic theory for multivariate asymmetric conditional volatility. *Econom. Rev.* **2009**, *28*, 422–440. [CrossRef]
14. McAleer, M.F.; Chan, S.; Hoti, O. Lieberman Generalized autoregressive conditional correlation. *Econom. Theory* **2008**, *24*, 1554–1583. [CrossRef]
15. Chang, C.-L.; McAleer, M. *The Fiction of Full BEKK*; Tinbergen Institute Discussion Paper 17–015; Tinbergen Institute: Amsterdam, The Netherlands, 2017.
16. Hafner, C.; McAleer, M. *A One Line Derivation of DCC: Application of a Vector Random Coefficient Moving Average Process*; Tinbergen Institute Discussion Paper 14–087; Tinbergen Institute: Amsterdam, The Netherland, 2014.
17. McAleer, M. *Stationarity and Invertibility of a Dynamic Correlation Matrix*; Tinbergen Institute Discussion Paper 17–082; Tinbergen Institute: Amsterdam, The Netherland, 2017.
18. Caporin, M.; McAleer, M. Do we really need both BEKK and DCC? A tale of two multivariate GARCH models. *J. Econ. Surv.* **2012**, *26*, 736–751. [CrossRef]

19. Chang, C.-L.; Li, Y.-Y.; McAleer, M. *Volatility Spillovers between Energy and Agricultural Markets: A Critical Appraisal of Theory and Practice*; Tinbergen Institute Discussion Paper 15–077/III; Tinbergen Institute: Amsterdam, The Netherland, 2015.

20. Chang, C.-L.; McAleer, M.; Wang, Y.-A. Modelling volatility spillovers for bio-ethanol, sugarcane and corn spot and futures prices. *Renew. Sustain. Energy Rev.* **2017**, *81*, 1002–1018. [CrossRef]

21. Tsay, R.S. Conditional heteroscedastic time series models. *J. Am. Stat. Assoc.* **1987**, *82*, 590–604. [CrossRef]

22. Engle, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **1982**, *50*, 987–1007. [CrossRef]

23. Bollerslev, T. Generalized autoregressive conditional heteroscedasticity. *J. Econom.* **1986**, *31*, 307–327. [CrossRef]

24. Jeantheau, T. Strong consistency of estimators for multivariate ARCH models. *Econom. Theory* **1998**, *14*, 70–86. [CrossRef]

25. Amemiya, T. *Advanced Econometrics*; Harvard University Press: Cambridge, MA, USA, 1985.

26. Asai, M.; Chang, C.-L.; McAleer, M. Realized stochastic volatility with general asymmetry and long memory. *J. Econom.* **2017**, *199*, 202–212. [CrossRef]

27. Dickey, D.A.; Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **1979**, *74*, 427–431.

28. Dickey, D.A.; Fuller, W.A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **1981**, *49*, 1057–1072. [CrossRef]

29. Said, S.E.; Dickey, D.A. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* **1984**, *71*, 599–607. [CrossRef]

30. Elliott, G.; Rothenberg, T.J.; Stock, J.H. Efficient tests for an autoregressive unit root. *Econometrica* **1996**, *64*, 813–836. [CrossRef]

31. Kwiatkowski, D.; Phillips, P.C.B.; Schmidt, P.; Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econom.* **1992**, *54*, 159–178. [CrossRef]

Risk Measurement and Risk Modelling Using Applications of Vine Copulas

David E. Allen [1,2,*], Michael McAleer [3,4,5,6,7] and Abhay K. Singh [2]

[1] School of Mathenatics and Statistics, Sydney University, Sydney, NSW 2006, Australia
[2] School of Business and Law, Edith Cowan University, Joondalup, WA 6027, Australia;
     abhaykumarsingh1985@gmail.com
[3] Department of Quantitative Finance, National Tsing Hua University, 30013 Hsinchu City, Taiwan;
     michael.mcaleer@gmail.com
[4] Discipline of Business Analytics, University of Sydney Business School, Sydney, NSW 2006, Australia
[5] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam,
     3062 PA Rotterdam, The Netherlands
[6] Department of Quantitative Economics, Complutense University of Madrid, 28040 Madrid, Spain
[7] Institute of Advanced Sciences, Yokohama National University, 240-8501 Yokohama, Japan
* Correspondence: profallen2007@gmail.com

Abstract: This paper features an application of Regular Vine copulas which are a novel and recently developed statistical and mathematical tool which can be applied in the assessment of composite financial risk. Copula-based dependence modelling is a popular tool in financial applications, but is usually applied to pairs of securities. By contrast, Vine copulas provide greater flexibility and permit the modelling of complex dependency patterns using the rich variety of bivariate copulas which may be arranged and analysed in a tree structure to explore multiple dependencies. The paper features the use of Regular Vine copulas in an analysis of the co-dependencies of 10 major European Stock Markets, as represented by individual market indices and the composite STOXX 50 index. The sample runs from 2005 to the end of 2013 to permit an exploration of how correlations change indifferent economic circumstances using three different sample periods: pre-GFC (January 2005–July 2007), GFC (July 2007– September 2009), and post-GFC periods (September 2009–December 2013). The empirical results suggest that the dependencies change in a complex manner, and are subject to change in different economic circumstances. One of the attractions of this approach to risk modelling is the flexibility in the choice of distributions used to model co-dependencies. The practical application of Regular Vine metrics is demonstrated via an example of the calculation of the VaR of a portfolio made up of the indices.

Keywords: regular vine copulas; tree structures; co-dependence modelling; European stock markets

1. Introduction

In the last decade copula modelling has become a frequently used tool in financial economics. Accounts of copula theory are available in [1,2]. Hierarchical, copula-based structures have recently been used in some new developments in multivariate modelling; notable among these structures is the pair-copula construction (PCC). Joe (1996) [3] originally proposed the PCC and further exploration of its properties has been undertaken by Bedford and Cooke [4,5] and Kurowicka and Cooke (2006) [6]. Aas et al., (2009) [7] provided key inferential insights which have stimulated the use of the PCC in various applications, (see, for example, Schirmacher and Schirmacher (2008) [8], Chollete et al. [9], Heinen and Valdesogo [10], Berg and Aas [11], Min and Czado [12] and Smith et al. [13]. Allen et al., (2013) [14] provide an illustration of the use of R-Vine copulas in the modelling of the dependences amongst Dow Jones Industrial Average component stocks, and this study is a companion piece.

There have also been some recent applications of copulas in the context of time series models (see the survey by Patton (2009) [15], and the recently developed COPAR model of Breckmann and Czado [16], which provides a vector autoregressive VAR model for analysing the non-linear and asymmetric co-dependencies between two series). Nevertheless, in this paper we focus on static modelling of dependencies based on R Vines in the context of modelling the co-dependencies of ten major European markets as captured by ten major indices and one composite European index. We use the British market represented by the FTSE100, the German market as captured by the DAX, the French market via the CAC40, the Netherlands, via the AEX index, the Spanish market represented by the IBEX35, the Danish market by means of the OMX Copenhagen 20, the Swedish market represented by the OMX Stockholm PI Index, the Finnish market using the OMXHPI, the Portuguese market using the PSI General Index (BVLG) and the Belgian market via the Belgian market via the Bell 20 Index (BFX). We also use the EURO STOXX 50 Index, Europe's leading Blue-chip index for the Eurozone, which consists of 50 major stocks from 12 Eurozone countries: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, The Netherlands, Portugal and Spain. We undertake our analysis in three different sample periods which include the GFC; pre-GFC (Jan 2005–July 2007), GFC (July 2007–September 2009), and post-GFC periods (September 2009–December 2013). To further show the capabilities of this flexible modelling technique, we also use R-Vine Copulas to quantify Value at Risk for an equally weighted portfolio of our eleven European indices, as an empirical example. The main aim of the paper is to demonstrate the useful application of both C-Vine and R-Vine measures of co-dependency at at time of extreme financial stress and its effectiveness in teasing out changes in co-dependency.

The paper is divided into five sections: the next section provides a review of the background theory and models applied, Section 3 introduces the sample, Sections 4 and 5 present the results for our analyses featuring C-Vine and R-Vines, Section 6 provides an example of the use of R-Vines to forecast the Value-at-Risk (VaR) and a brief conclusion follows in Section 7.

2. Background and Models

Sklar (1959) [17] provides the basic theorem describing the role of copulas for describing dependence in statistics, providing the link between multivariate distribution functions and their univariate margins. We can speak generally of the copula of continuous random variables $X = (X_1, ...X_d) \sim F$. The problem in practical applications is the identification of the appropriate copula.

Standard multivariate copulas, such as the multivariate Gaussian or Student-t, as well as exchangeable Archimedean copulas, lack the exibility of accurately modelling the dependence among larger numbers of variables. Generalizations of these offer some improvement, but typically become rather intricate in their structure, and hence exhibit other limitations such as parameter restrictions. Vine copulas do not suffer from any of these problems.

Initially proposed by Joe [3] and developed in greater detail in Bedford and Cooke [4,5] and in Kurowicka and Cooke [6], vines are a flexible graphical model for describing multivariate copulas built up using a cascade of bivariate copulas, so-called pair-copulas. Their statistical breakthrough was due to Aas, Czado, Frigessi, and Bakken [7] who described statistical inference techniques for the two classes of canonical C-vines and D-vines. These belong to a general class of Regular Vines, or R-vines which can be depicted in a graphical theoretic model to determine which pairs are included in a pair-copula decomposition. Therefore a vine is a graphical tool for labelling constraints in high-dimensional distributions.

This area of the literature has expanded rapidly. Joe et al., (2010) [18] explore the tail dependence and conditional tail dependence functions of vine copulas of lower-dimensional margins. In addition, the effect of tail dependence of bivariate linking copulas on that of a vine copula is investigated. Geidosch and Fisher (2016) [19] show the superiority of vine copulas over conventional copulas when modeling the dependence structure of a credit portfolio. Fischer et al. [20] use vine copula based quantile regression to stress testing German industry sectors.

One drawback in the application of vine copulas is that even for a moderate number of variables, the number of alternative vine decompositions is very large and there is also a large set of

candidate bivariate copula families that can be used as building blocks in any given decomposition. Pangiotelis et al. [21] address this issue via the consideration of two greedy algorithms which automatically select vine structures and component pair-copula building blocks, so as to reduce computional demands, and report positive results from simulations and applications to data drawn from the retail sector. In a similar vein, Bedford et al. [22] demonstrate how the application of vines can approximate any density as closely as required. They operationalize their result by showing that minimum information copulas can be used to provide parametric classes of copulas that have required levels of approximation. Scheffer and Weiÿ [23] use nonparametric Bernstein vine copulas as bivariate pair-copulas to model VaR in a GARCH context. Aas (2016) [24] provides a review of both inference methods and goodness-of-fit tests for pair-copula constructions for financial applications, plus empirical applications of these models in finance and economics, whilst Fermanian [25] similarly reviews recent developments in copula models.

A regular vine is a special case for which all constraints are two-dimensional or conditional two-dimensional. Regular vines generalize trees, and are themselves specializations of Cantor trees. Combined with copulas, regular vines have proven to be a flexible tool in high-dimensional dependence modelling. Copulas are multivariate distributions with uniform univariate margins. Representing a joint distribution as univariate margins plus copulas allows the separation of the problems of estimating univariate distributions from problems of estimating dependence.

Figure 1 provides an example of two different vine structures, with a regular vine on the left and a non-regular vine on the right, both for four variables.
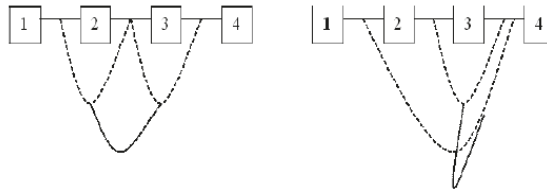
Figure 1. Vines.

A vine V on n variables is a nested set of connected trees $V = \{T_1, ..., T_{n-1}\}$, where the edges of tree j are the nodes of tree $j + 1$, $j = 1, ..., n - 2$. A regular vine on n variables is a vine in which two edges in tree j are joined by an edge in tree $j = 1$ only if these edges share a common node, $j = 1, ..., n - 2$. Kurowicka and Cook [26] provide the following definition of a Regular vine.

Definition 1. *(Regular vine)*

   V is a regular vine on n elements with $E(V) = E_1 \cup ... \cup E_{n-1}$ denoting the set of edges of V if

1. $V = \{T_1, ..., T_{n-1}\}$,
2. T_1 *is a connected tree with nodes $N_1 = \{1, ..., n\}$, plus edges E_1; for $i = 2, ..., n - 1$, T_i is a tree with nodes $N_i = E_{i-1}$,*
3. *(proximity) for $i = 2, ..., n - 1$, and $\{a, b\} \in E_i$, with $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$ it must hold that $\{\#(a \cap b) = 1$, where $\cap$ denotes the symmetric difference operator and # denotes the cardinality of a set.*

An edge in a tree T_j is an unordered pair of nodes of T_j or equivalently, an unordered pair of edges of T_{j-1}. By definition, the order of an edge in tree T_j is $j - 1$, $j = 1, ..., n - 1$. The degree of a node is determined by the number of edges attached to that node. A regular vine is called a *canonical vine*, or *C-vine*, if each tree T_i has a unique node of degree $n - 1$ and therefore, has the maximum degree. A regular vine is termed a *D-vine* if all the nodes in T_1 have degrees no higher than 2.

Definition 2. *(The following definition is taken from Cook et al., (2011) [27]). For $e \in E_i$, $i \leq n - 1$, the constraint set associated with e is the complete union of U_e^* of e, which is the subset of $\{1, ..., n\}$ reachable from e by the membership relation.*

For $i = 1, ..., n - 1$, $e \in E_i$, if $e = \{j, k\}$, then the conditioning set associated with e is

$$D_e = U_j^* \cap U_k^*$$

and the conditioned set associated with e is

$$\{C_{e,j}, C_{e,k}\} = \left\{ U_j^* \setminus D_e, U_k^* \setminus D_e \right\}.$$

Figure 2 below shows a D-Vine with 5 dimensions.

Figure 2. D-Vine 5 Dimensions.

Figure 3 shows an R-Vine on 4 variables, and is sourced from Dissman (2010) [28]. The node names appear in the circles in the trees and the edge names appear below the edges in the trees. Given that an edge is a set of two nodes, an edge in the third tree is a set of a set. The proximity condition can be seen in tree T_2, where the first edge connects the nodes $\{1, 2\}$ and $\{2, 3\}$, and both share node 2 in tree T_1.

No.	Elliptical distribution	Parameter range	Kendall's τ	Tail dependence
1	Gaussian	$\rho \in (-1, 1)$	$\frac{2}{\pi} \arcsin(\rho)$	0
2	Student-t	$\rho \in (-1, 1)$, $\nu > 2$	$\frac{2}{\pi} \arcsin(\rho)$	$2t_{\nu+1}\left(-\sqrt{\nu + 1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$

No.	Name	Generator function	Parameter range	Kendall's τ	Tail dependence (lower, upper)
3	Clayton	$\frac{1}{\theta}(t^{-\theta} - 1)$	$\theta > 0$	$\frac{\theta}{\theta+2}$	$(2^{-1/\theta}, 0)$
4	Gumbel	$(-\log t)^\theta$	$\theta \geq 1$	$1 - \frac{1}{\theta}$	$(0, 2 - 2^{1/\theta})$
5	Frank[a]	$-\log[\frac{e^{-\theta t}-1}{e^{-\theta}-1}]$	$\theta \in \mathbb{R} \setminus \{0\}$	$1 - \frac{4}{\theta} + 4\frac{D_1(\theta)}{\theta}$	$(0, 0)$
6	Joe	$-\log[1 - (1 - t)^\theta]$	$\theta > 1$	$1 + \frac{4}{\theta^2}\int_0^1 t \log(t)(1 - t)^{\frac{2(1-\theta)}{\theta}} dt$	$(0, 2 - 2^{1/\theta})$
7	BB1	$(t^{-\theta} - 1)^\delta$	$\theta > 0$, $\delta \geq 1$	$1 - \frac{2}{\delta(\theta+2)}$	$(2^{-1/(\theta\delta)}, 2 - 2^{1/\delta})$
8	BB6	$(-\log[1 - (1 - t)^\theta])^\delta$	$\theta \geq 1, \delta \geq 1$	$1 + 4\int_0^1 \left(-\log(-(1-t)^\theta + 1)\right) \times \frac{(1-t-(1-t)^{-\theta}+t(1-t)^{-\theta})}{\delta\theta}\big) dt$	$(0, 2 - 2^{1/(\theta\delta)})$
9	BB7[b]	$[1 - (1 - t)^\theta)]^{-\delta} - 1$	$\theta \geq 1$, $\delta > 0$	$1 - \frac{2}{\delta(2-\theta)} + \frac{4}{\theta^2\delta}B(\frac{2-\theta}{\theta}, \delta + 2)$	$(2^{-1/\delta}, 2 - 2^{1/\theta})$
10	BB8	$-\log\left[\frac{1-(1-\delta t)^\theta}{1-(1-\delta)^\theta}\right]$	$\theta \geq 1, 0 < \delta \leq 1$	$1 + 4\int_0^1 \left(-\log\left(\frac{(1-t\delta)^\theta-1}{(1-\delta)^\theta-1}\right)\right) \times \frac{1-t\delta-(1-t\delta)^{-\theta}+t\delta(1-t\delta)^{-\theta}}{\theta\delta}\big) dt$	$(0, 0^c)$

Figure 3. Notation and Properties of Bivariate Elliptical and Archimedean Copula Families included in CDVine.

2.1. Modelling Vines

Vine structures are developed from pair-copula constructions, in which $d(d-1)/2$ pair-copulas are arranged in $d-1$ trees (in the form of connected acyclic graphs with nodes and edges). At the start of the first C-vine tree, the first root node models the dependence with respect to one particular variable, using bivariate copulas for each pair. Conditioned on this variable, pairwise dependencies with respect to a second variable are modelled, the second root node. The tree is thus expanded in this manner; a root node is chosen for each tree and all pairwise dependencies with respect to this node are modelled conditioned on all previous root nodes. It follows that C-vine trees have a star structure. Brechmann and Schepsmeier (2012) [29] use the following decomposition in their account of the routines incorporated in the R Library CDVine, which was used for the empirical work in this paper. The multivariate density, the *C Vine density w.l.o.g. root nodes* $1, ..., d$,

$$f(x) = \prod_{k=1}^{d} f_k(x_k) \times \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{i,i+j|1:(i-1)}\left(F(x_i \mid x_1, ..., x_{i-1}), F(x_{i+j} \mid x_1, ..., x_{i-1}) \mid \theta_{i,+j|1:(i-1)}\right) \quad (1)$$

where f_k, $k = 1, ..., d$, denote the marginal densities and $c_{i,i+j|1:(i-1)}$ bivariate copula densities with parameter(s) $\theta_{i,i+j|1:(i-1)}$ (in general, $i_k : i_m$ means $i_k, ..., i_m$). The outer product runs over the $d-1$ trees and root nodes i, while the inner product refers to the $d-i$ pair copulas in each tree $i = 1, .., d-1$.

D-Vines follow a similar process of construction by choosing a specific order for the variables. The first tree models the dependence of the first and second variables, of the second and third, and so on, ... using pair copulas. If we assume the order is $1, ..., d$, then first the pairs $(1, 2), (2, 3), (3, 4)$ are modelled. In the second tree, the co-dependence analysis can proceed by modelling the conditional dependence of the first and the third variables, given the second variable; the pair $(2, 4 \mid 3)$, and so forth. This process can then be continued in the next tree, in which variables can be conditioned on those lying between entries *a and b* in the first tree, for example, the pair $(1, 5 \mid 2, 3, 4)$. The D-Vine tree has a path structure which leads to the construction of the *D*-vine density, which can be constructed as follows:

$$f(x) = \prod_{k=1}^{d} f_k(x_k) \times \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{j,j+i|(j+1):(j+i-1)}\left(F(x_j \mid x_{j+1}, ..., x_{j+i-1}), F(x_{j+i} \mid x_{j+1}, ..., x_{j+i-1}) \mid \theta_{j,j+i|(j+1):(j+i-1)}\right) \quad (2)$$

The outer product runs over $d-1$ trees, while the pairs in each tree are determined according to the inner product. The conditional distribution functions $F(x \mid v)$ can be obtained for an $m-dimensional$ vector v. This can be done in a pair copula term in tree $m-1$, by using the pair-copulas of the previous trees $1, ..., m$, and by sequentially applying the following relationship:

$$h(x \mid v, \theta) := F(x \mid v) = \frac{\partial C_{xv_j|v_{-j}}\left(F(x \mid v_{-j}), F(v_j \mid v_{-j}) \mid \theta\right)}{\partial F(v_j \mid v_{-j})} \quad (3)$$

where v_j is an arbitrary component of v, and v_{-j} denotes the $(m-1)$- dimensional vector v excluding v_j. The bivariate copula function is specified by $C_{xv_j|v_{-j}}$ with parameters θ specified in tree m.

The model of dependency can be constructed in a very flexible way because a variety of pair copula terms can be fitted between the various pairs of variables. In this manner, asymmetric dependence or strong tail behaviour can be accommodated. Figure 3 shows the various copulae available in the CDVine library in R.

2.2. Regular Vines

Until recently, the focus had been on modelling using C and D vines. However, Dissmann [28] has pointed the direction for constructing regular vines using graph theoretical algorithms. This interest in pair-copula constructions/regular vines is doubtlessly linked to their high flexibility as they can model a wide range of complex dependencies.
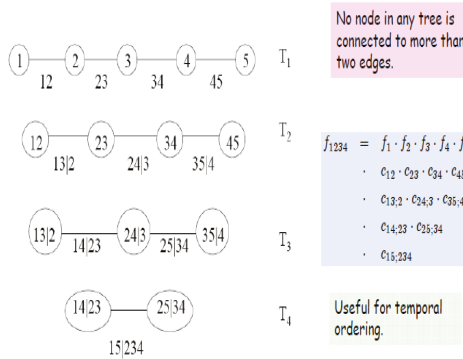
Figure 4 shows an R-Vine on 4 variables, and is sourced from Dissman [28]. The node names appear in the circles in the trees and the edge names appear below the edges in the trees. Given that an edge is a set of two nodes, an edge in the third tree is a set of a set. The proximity condition can be seen in tree T_2, where the first edge connects the nodes $\{1, 2\}$ and $\{2, 3\}$, plus both share the node 2 in tree T_1.
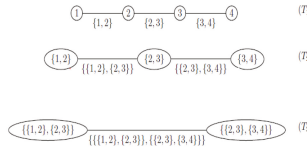
Figure 4. Example of R-Vine on 4 Variables. (Source Dissman (2010)).

The drawback is the curse of dimensionality: the computational effort required to estimate all parameters grows exponentially with the dimension. Morales-Nápoles et al. [30] demonstrate that there are $(\frac{n}{2}) \times (n-2)! \times 2^{\binom{n-2}{2}}$ possible R-Vines on n nodes. The key to the problem is whether the regular vine can be either truncated or simplified. Brechmann et al. [28] (p. 2) discuss such simplification methods. They explain that: "by a pairwisely truncated regular vine at level K, we mean a regular vine where all pair-copulas with conditioning set equal to or larger than K are replaced by independence copulas". They pairwise simplify a regular vine at level K by replacing the same pair-copulas with Gaussian copulas. Gaussian copulas mean a simplification since they are easier to specify than other copulas, easy to interpret in terms of the correlation parameter, and quicker to estimate.

They identify the most appropriate truncation/simplification level by means of statistical model selection methods; specifically, the AIC, BIC and the likelihood-ratio based test proposed by Vuong (1989) [31]. For R-vines, in general, there are no expressions like Equations (2) and (3). This means that an efficient method for storing the indices of the pair copulas required in the joint density function, as depicted in Equation (4), is required; (4) is a more general case of (2) and (3).

$$f(x_1, ..., x_d) = \left[\prod_{k=1}^{d} f_x(x_k) \right] \times \left[\prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)}(F(x_{j(e)} \mid \mathbf{x}_{D(e)}), F(x_{k(e)} \mid \mathbf{x}_{D(e)})) \right] \quad (4)$$

Kurowicka [32] and Dissman [28] have suggested a method of proceeding which involves specifying a lower triangular matrix $M = (m_{i,j} \mid i, j = 1, ..., d) \in \{0, ..., d\}^{d \times d}$, with $m_{i,i} = d - i + 1$. This means that the diagonal entries of M are the numbers $1, ..., d$ in descending order. In this matrix, each row proceeding from the bottom represents a tree, the diagonal entry represents the conditioned set and by the corresponding column entry of the row under consideration. The conditioning set is given by the column entries below this row. The corresponding parameters and types of copula can be stored in matrices relating to M. The following example in Figure 5 is taken from Dissman [28].

The first section of Figure 5 provides a key to indicate the 5 different types of copulas used in this example, ranging from Gaussian (1) to Frank (5). The second lower triangular matrix T_1 shows the application of particular types of copulas in the trees, P_1^1 shows the parameters estimated, and P_1^2 provides the extra parameters needed when we apply the t copula.
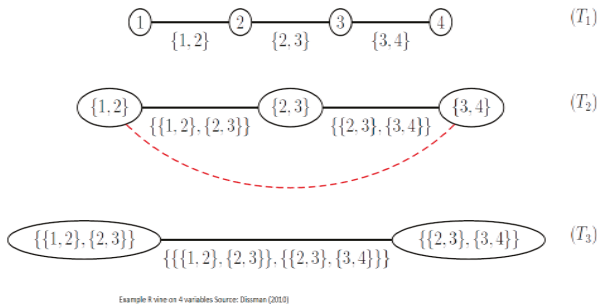
Figure 5. Matrix Mapping of vine copulas (source Dissman (2010)).

In Figure 6 the bottom row of M_1 corresponds to T_1, the second row to T_2, and so on. In order to determine the edges in T_1, we combine the numbers in the bottom row with the diagonal elements in the corresponding columns, for example the edges are (4, 3), (5, 2), (1, 2) and so on. In order to determine the edges in T_2, we combine the numbers in the second row from the bottom with the diagonal elements in the corresponding columns and condition on the elements in the bottom row. This would give edges $(4, 2 \mid 3)$, $(5, 3 \mid 2)$, $(1, 3 \mid 2)$, and so on The final entry is given by the upper entries to the left of the matrix $(4, 7 \mid 65123)$.

Figure 6. Use of Matrices to Store R-vine Information (source: Dissman (2010)).

2.3. Prior Work with R-Vines

The literature was initially mainly concerned with illustrative examples, (see, for example, Aas et al. [7], Berg and Aas [11], Min and Czado [12] and Czado et al. [33]). Mendes et al., (2010) [34] use a D-Vine copula model to a six-dimensional data set and consider its use for portfolio management. Dissman [28] uses R-Vines to analyse dependencies between 16 financial indices covering different European regions and different asset classes, including five equity, nine fixed income (bonds), and two commodity indices. He assesses the relative effectiveness of the use of copulas, based on mixed distributions, t distributions and Gaussian distributions, and explores the loss of information from truncating the R-Vine at earlier stages of the analysis and the substitution of independence copula. He also analyses exchange rates and windspeed data sets with fewer variables.

The research in this paper extends the work of Dissman [28] applying R-Vines to a European financial data set using a set of eleven European stock indices and features an exploration of how their dependency structures change through periods of extreme stress as represented by the GFC. The paper also features an example of how the dependencies captured by the R-Vine analysis can be used to assess portfolio Value at Risk (VaR) in a manner that closely parallels Breckmann and Czado [35] who adopted a factor model approach discussed below.

There have been other studies on European stock return series: Heinen and Valdesogo [10] constructed a CAPM extension using their *Canonical Vine Autoregressive* (CAVA) model using marginal GARCH models and a canonical vine copula structure. Breckmann and Czado [35] develop a regular vine market sector factor model for asset returns that uses GARCH models for margins, and which is similarly developed in a CAPM framework. They explore systematic and unsystematic risk for individual stocks, and consider how vine copula models can be used for active and passive portfolio management and VaR forecasting.

3. Sample

We use a data set of daily returns, which runs from 1 January 2005 to 31 January 2013 for ten European indices and the composite blue chip STOXX50 European index. We use the British FTSE 100 Index, the German DAX Index, the French CAC 40 Index, the Netherlands AEX Amsterdam Index, the Spanish Ibex 35 Index, the Danish OMX Copenhagen 20 Index, the Swedish OMX Stockholm All Share Index, the Finnish OMX Helsinki All Share Index, the Portuguese PSI General Index, and the Belgian Bell 20 Index. As a composite European market index we use the STOXX 50. This index covers 50 stocks from 12 Eurozone countries: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain. We divide our sample into returns for the pre-GFC (January 2005–July 2007), GFC (July 2007–September 2009) and post-GFC (September 2009–December 2013) periods. The sample is shown in Table 1.

Table 1. Index Data.

Reuters RIC Code	Index
.FTSE	British FTSE Index
.GDAXI	German DAX Index
.FCHI	French CAC 40 Index
.AEX	AEX Amsterdam Index
.IBEX	Spanish Ibex 35 Index
.OMXC20	OMX Copenhagen 20 Index
.OMXSPI	OMX Stockholm All Share Index
.OMXHPI	OMX Helsinki All Share Index
.BVLG	Portuguese PSI General
.BFX	Belgian Bell 20 Index
.STOXX50	European STOXX 50

Tables 2 and 3 provide descriptive statistics for the ten European market indices and the composite European STOXX50 index broken down into our three periods; pre-GFC (January 2005–July 2007), GFC (July 2007–September 2009) and post-GFC (September 2009–December 2013). It is apparent that the mean and median returns are uniformly positive in the pre-GFC period, and uniformly negative in the GFC period, whilst the median return is either zero or positive for all but two of the indices during this period. In the post-GFC the mean and median returns for most markets are positive or zero except in the cases of the Spanish and Portuguese markets where there are negative mean returns. The standard deviation is higher in all markets in the GFC period. The Bera-Jarque test significantly rejects normality of the daily return distributions for all indices in all periods. The returns are skewed but in many cases change the direction of the skew from positive to negative in different periods. Only three markets display negative skewness in the GFC period; the Danish, the Portuguese and the Belgian markets.

All markets except the Swedish one show greater excess-kurtosis during the GFC period. The GFC period is also characterised by a higher value of the Hurst exponent in all markets, with a value greater than 0.57 in all markets, suggesting the markets display long memory in times of crisis.

The descriptive statistics provided in Tables 2 and 3 suggest that the European index return series in our sample are non-Gaussian and are subject to changes in skewness and kurtosis in the different sample sub-periods. This suggests they should be amenable to analysis by copulas which may capture the effects of fat tails and changes in distributional characteristics.

Table 2. Descriptive statistics for indices by sub-period: Pre-GFC, GFC, and Post-GFC.

Index	PRE-GFC January 2005–June 2007	GFC July 2007–August 2009	Post GFC September 2009–December 2013
		.FTSE	
Mean	0.000555640	−0.000895714	0.000295718
Median	0.000738054	0.000203498	0.000584525
St. Deviation	0.00813613	0.0238888	0.0129783
Skewness	−0.149519	0.0528700	−0.224389
Ex-Kurtosis	1.19971	4.78545	2.23948
Bera-Jarque test	41.467 (0.00)	539.38 0.000	245.837 (0.00)
Hurst Exponent	41.467	0.57554	0.453993
		.GDAXI	
Mean	0.000965496	−0.000573968	0.000457927
Median	0.00117051	0.000537208	0.000326914
St. Deviation	0.00984275	0.0235477	0.0166782
Skewness	−0.198918	0.188195	−0.189233
Ex-Kurtosis	1.06819	4.83638	2.14630
Bera-Jarque test	35.2434 (0.00)	553.989 (0.00)	223.835
Hurst exponent	0.513785	0.577037	0.503562
		.FCHI	
Mean	0.000704096	−0.000794553	0.000107354
Median	0.000603012	0.000142254	0.000167647
St. Deviation	0.00919530	0.0244639	0.0178098
Skewness	−0.225337	0.187499	−0.0293812
Ex-Kurtosis	1.29375	4.94817	2.75582
Bera-Jarque test	50.9105 (0.00)	579.714 (0.00)	2.75582
Hurst exponent	0.480051	0.570031	0.4944
		.AEX	
Mean	0.000706398	−0.00100306	0.000233513
Median	0.000689552	0.000104563	0.000746154
St. Deviation	0.00866571	0.0252091	0.0151689
Skewness	−0.168285	0.0194178	−0.106574
Ex-Kurtosis	1.79428	4.93931	2.18174
Bera-Jarque test	90.3996 (0.00)	574.377 (0.00)	226.455 (0.00)
Hurst exponent	0.510652	0.606603	0.503613
		.IBEX	
Mean	0.000754932	−0.000376594	−0.000156397
Median	0.000457241	0.00000	0.00000
St. Deviation	0.00896277	0.0237746	0.0199591
Skewness	−0.179712	0.0160456	0.176078
Ex-Kurtosis	1.19903	4.57761	3.89965
Bera-Jarque test	42.5008	493.327 (0.00)	722.488 (0.00)
Hurst exponent	0.51974	0.602261	0.552881
		.OMXC20	
Mean	0.000812167	−0.000564910	0.000499859
Median	0.00132764	0.00000	0.000802048
St. Deviation	0.00993412	0.0239638	0.0144641
Skewness	−0.765524	−0.187144	−0.116423
Ex-Kurtosis	2.80674	4.22380	2.12358
Bera-Jarque test	277.268 (0.00)	423.095 (0.00)	215.069
Hurst exponent	0.498681	0.59553	0.520375

Table 3. Descriptive statistics for indices by sub-period: Pre-GFC, GFC, and Post-GFC (Contd).

Index	PRE-GFC January 2005–June 2007	GFC July 2007–August 2009	Post GFC September 2009–December 2013
.OMXSPI			
Mean	0.000865551	−0.000760510	0.000458071
Median	0.000980873	0.00000	0.000724200
St. Deviation	0.0111512	0.0276237	0.0177400
Skewness	−0.295747	0.230350	−0.208441
Ex-Kurtosis	3.58218	2.34544	2.67172
Bera-Jarque test	357.559	134.501 (0.00)	344.573 (0.00)
Hurst exponent	0.522449	0.574409	0.497336
.OMXHPI			
Mean	0.000938877	−0.000984965	0.000108658
Median	0.000484177	−0.000943853	0.000373821
St. Deviation	0.0103049	0.0242644	0.0167601
Skewness	−0.115052	0.139482	−0.115636
Ex-Kurtosis	2.36500	2.40933	2.21433
Bera-Jarque test	153.152	138.489	233.587 (0.00)
Hurst exponent	0.49008	0.61372	0.538587
.BVLG			
Mean	0.000891032	−0.000884323	−0.000169933
Median	0.000892271	−8.71710e-005	5.77024e-005
St. Deviation	0.00703327	0.0199733	0.0159551
Skewness	0.0898224	−0.160906	−0.00138346
Ex-Kurtosis	1.05671	6.43503	3.74001
Bera-Jarque test	31.64 (0.01)	977.29 (0.00)	659.168 (0.00)
Hurst exponent	0.6217	0.614598	0.556314
.BFX			
Mean	0.000692830	−0.00108296	0.000151703
Median	0.000817080	0.00000	0.000179149
St. Deviation	0.00877648	0.0222434	0.0158848
Skewness	−0.225903	−0.131889	0.00987544
Ex-Kurtosis	1.59909	3.23142	2.90360
Bera-Jarque test	74.898	247.462 (0.00)	397.323 (0.00)
Hurst exponent	0.54556	0.641453	0.497547
.STOXX50			
Mean	0.000642054	−0.000752410	6.45717e-005
Median	0.000635064	0.000153018	0.00000
St. Deviation	0.00922232	0.0241936	0.0179815
Skewness	−0.154465	0.0851528	0.00350895
Ex-Kurtosis	1.33626	4.29802	2.88743
Bera-Jarque test	51.0231 (0.00)	435.568 (0.00)	392.895 (0.00)
Hurst exponent	0.474577	0.587196	0.504895

4. Results

The results are presented here in two parts. In the first subsection below we model the dependence structure of the European set of indices, in three subperiods covering GFC. The second subsection gives results from an empirical exercise modelling VaR using R-Vine Copulas for a 10 asset portfolio and contrasts it with the results of a more traditional Gaussian approach undertaken in a GARCH framework.

4.1. Dependence Modelling Using Vine Copula

We divide the data into three time periods covering the pre-GFC (January 2005–July 2007), GFC (July 2007–September 2009), and post-GFC periods (September 2009–December 2011) to run the C-Vine and R-Vine dependence analysis for the stocks comprising Dow Jones Index. Before we can do this we require appropriately standardised marginal distributions for the basic company return series. These appropriate marginal time series models for the Dow Jones data have to be found in the first step of our two step estimation approach. The following time series models are selected in a stepwise procedure: GARCH (1, 1), ARMA (1, 1), AR(1), GARCH(1, 1), MA(1)-GARCH(1, 1). These are applied to the return data series and we select the model with the highest p-value, so that the residuals

can be taken to be i.i.d. The residuals are standardized and the marginals are obtained from the standardized residuals using the Ranks method. These marginals are then used as inputs to the Copula selection routine. The copula are selected using the AIC criterion. We first discuss the results obtained from the pre-GFC period data followed by the GFC and post-GFC periods.

4.2. Pre-GFC

The following Figure 7 presents the structure of the C-Vines.
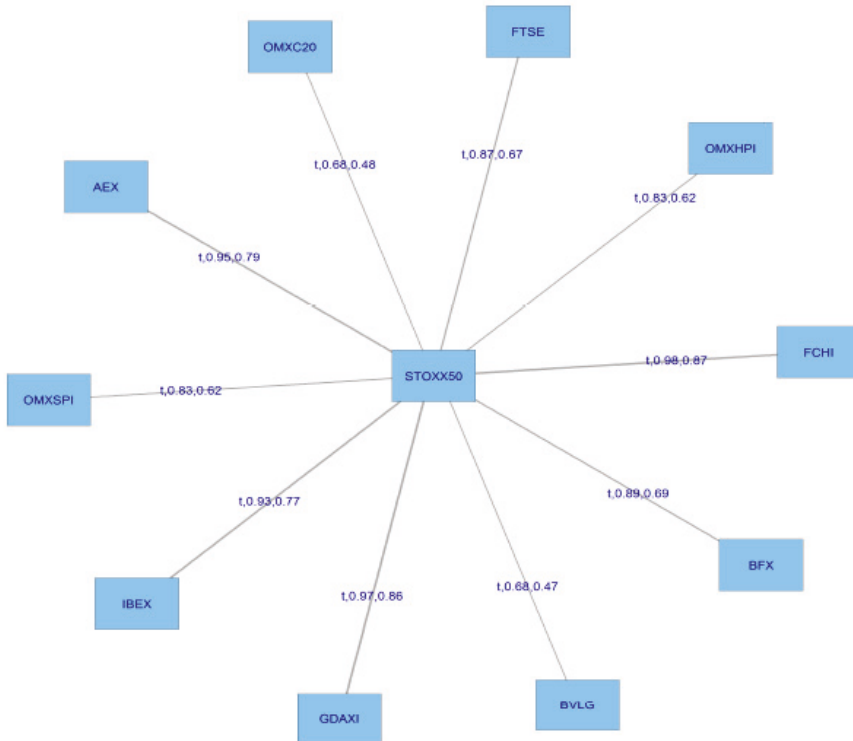
Figure 7. Results-C-Vine Tree-1 Pre-GFC.

For this C Vine selection, we choose as root node the node that maximizes the sum of pairwise dependencies to this node. We commence by linking all the stocks to the STOXX50 index which is at the centre of this diagram. We use a range of Copulas from for selection purposes; the range being (1:6). We apply AIC as the selection criterion to select from the following menu of copulae: 1 = Gaussian copula, 2 = Student t copula (t-copula), 3 = Clayton copula, 4 = Gumbel copula, 5 = Frank copula, 6 = Joe copula.

We then compute transformed observations from the estimated pair copulas and these are used as input parameters for the next trees, which are obtained similarly by constructing a graph according to the above C-Vine construction principles (proximity conditions), and finding a maximum dependence tree. The C-Vine tree for period 2 is shown Figure 8.

The pre-GFC C-Vine copula specification matrix is displayed in Table 4 below. It can be seen from the top and bottom of the first column in Table 4 that in the pre-GFC period the strongest correlations

are between the FTSE and Belgian Index BFX. The BFX remains at the bottom across all columns in the last row of Table 4.

Figure 8. C-Vine Tree 2 Pre-GFC.

From Table 4, it can be seen that the strongest individual correlations in the pre-GFC period, are between the FTSE at the top of the first column, BFX in the final row, and the individual diagonal entries starting with the FTSE at the top of the first column, which define the edges. The FTSE is correlated with BVLG (security 11), then conditioned by its relationship with OMXHPI (security 8), the Helsinki exchange index, then OMXSPI (security 5), the Stockholm index, then OMXC20 (security 3), the Copenhagen index, and so on. It can also be seen in Table 2 that C Vines are less flexible in that the same security number can usually be seen to appear across the rows. This means that it is always appearing in the nodes at that level in the tree. R Vines are more flexible and do not have this requirement. Later in the paper, we will concentrate on the results of the R Vine analysis.

Table 5 shows which copula are fitted to capture dependencies between the various pairs of indices. At the bottom of column 1 in Table 5 we can see that number 2 copula, the Student t copula is applied, to capture the dependency between FTSE and BFX, and then it is conditioned by the relationship with BVLG but this relationship uses a Frank copula (5), and so forth. All 6 categories of copula are used in Table 5 but the Student t copula appears most frequently in the table, followed by the Frank copula, the Gaussian copula, the Clayton copula and finally the Joe copula and the Gumbel copula appear once each.

It can be seen in Table 6 in the entries in the bottom row that there are strong positive dependencies between subsets of the markets concerned. The entry in the bottom of the first column shows the strong positive dependency between the FTSE and BFX. All the entries in the bottom row of Table 6 are strongly positive. We can see in the first column, that once we have conditioned the FTSE on its relationships with the markets in the bottom half of the column it is strongly positively related to the STOXX50. Not all the dependencies indicated in Table 6 are positive though, and there are 11 cases of negative co-dependency, once the relationship across other nodes has been taken into account.

Table 4. Pre-GFC C-Vine Copula Structure.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	1	0	0	0	0	0	0	0	0	0	0
GDAXI	9	3	0	0	0	0	0	0	0	0	0
FCHI	3	9	9	0	0	0	0	0	0	0	0
AEX	10	10	10	2	0	0	0	0	0	0	0
IBEX	2	2	2	10	4	0	0	0	0	0	0
STOXX50	4	4	4	4	10	5	0	0	0	0	0
OMXC20	5	5	5	5	5	10	7	0	0	0	0
OMXSPI	7	7	7	7	7	7	10	8	0	0	0
OMXHPI	8	8	8	8	8	8	8	10	10	0	0
BVLG	11	11	11	11	11	11	11	11	11	11	0
BFX	6	6	6	6	6	6	6	6	6	6	6

Table 5. Pre-GFC C-Vine Copula Specification Matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0	0	0	0	0	0	0	0	0	0	0
GDAXI	2	0	0	0	0	0	0	0	0	0	0
FCHI	2	2	0	0	0	0	0	0	0	0	0
AEX	2	3	5	0	0	0	0	0	0	0	0
IBEX	5	2	2	1	0	0	0	0	0	0	0
STOXX50	5	2	1	5	5	0	0	0	0	0	0
OMXC20	5	2	3	1	1	1	0	0	0	0	0
OMXSPI	5	1	5	5	3	4	5	0	0	0	0
OMXHPI	1	3	2	3	6	5	1	5	0	0	0
BVLG	5	1	5	5	5	2	1	3	1	0	0
BFX	2	2	2	2	2	2	2	2	2	2	0

Table 6. Pre-GFC C-Vine Copula Parameter Estimates.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0	0	0	0	0	0	0	0	0	0	0
GDAXI	0.152236	0	0	0	0	0	0	0	0	0	0
FCHI	0.786802	0.437083	0	0	0	0	0	0	0	0	0
AEX	1.418514	0.053894	1.036334	0	0	0	0	0	0	0	0
IBEX	1.033914	0.672986	0.162141	0.137522	0	0	0	0	0	0	0
STOXX50	−0.18660	1.068144	0.155770	0.076751	0.205075	0	0	0	0	0	0
OMXC20	−0.06799	−0.011583	−0.092367	0.201391	0.973316	0.083272	0	0	0	0	0
OMXSPI	−0.16274	0.016648	−0.261361	0.182827	0.714802	0.047323	−0.025852	0	0	0	0
OMXHPI	1.05367	0.087220	0.0779183	0.126660	0.071290	−0.004308	0.031866	1.223109	0	0	0
BVLG	1.56654	0.164849	1.2442506	0.408728	−0.130783	−0.129392	0.148569	0.226344	0.196334	0	0
BFX	0.94972	0.871079	0.9334080	0.8309762	0.827286	0.973728	0.978593	0.693275	0.890633	0.718538	0

Table 7 shows the second set of parameters, in cases where one is needed, for example the Student t copula.

Table 7. Pre-GFC C-Vine Copula Second Parameter Estimates.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0	0	0	0	0	0	0	0	0	0	0
GDAXI	15.375671	0	0	0	0	0	0	0	0	0	0
FCHI	9.544367	12.310803	0	0	0	0	0	0	0	0	0
AEX	9.401744	0	0	0	0	0	0	0	0	0	0
IBEX	0	10.267206	10.233424	0	0	0	0	0	0	0	0
STOXX50	0	10.124756	0	0	0	0	0	0	0	0	0
OMXC20	0	8.646548	0	0	0	0	0	0	0	0	0
OMXSPI	0	0	0	0	0	0	0	0	0	0	0
OMXHPI	0	0	8.390870	0	0	0	0	0	0	0	0
BVLG	0	0	0	0	0	12.429237	0	0	0	0	0
BFX	8.686229	5.378347	3.377834	8.575454	11.885624	7.882211	6.454538	9.626281	13.482133	8.332783	0

Table 8 shows the tau matrix for the C Vine copulas in the pre-GFC period.

The bottom row of Table 8 captures the strongest dependencies between the pairs of markets, as represented by their respective indices.

Table 8. Pre-GFC C-Vine Copula Tau matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
GDAXI	0.004067	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
FCHI	0.036096	−0.029548	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
AEX	0.014419	0.030447	0.068843	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
IBEX	0.005771	−0.091656	−0.099620	0.025585	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
STOXX50	0.155059	−0.031242	−0.053529	−0.007041	0.075579	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
OMXC20	0.036919	−0.085254	0.030712	−0.029132	−0.109944	0.101804	0.000000	0.000000	0.000000	0.000000	0.000000
OMXSPI	0.062762	0.067815	0.106788	−0.007854	0.069387	0.037842	0.118612	0.000000	0.000000	0.000000	0.000000
OMXHPI	0.084060	0.056764	0.258504	0.035254	0.039828	0.116247	0.152900	0.096636	0.000000	0.000000	0.000000
BVLG	0.121015	0.168311	0.082120	−0.024858	0.162175	0.101571	0.173707	0.120451	0.224388	0.000000	0.000000
BFX	0.673160	0.868040	0.620234	0.853751	0.797280	0.766361	0.478690	0.624435	0.472402	0.692177	0.000000

A key concern in this paper is the issue of how dependencies have changed as a result of the GFC?

4.3. GFC Period

Figure 9 shows tree 1 for C-Vine copula estimates in the GFC period, and Figure 10 shows tree 2 for the same period.
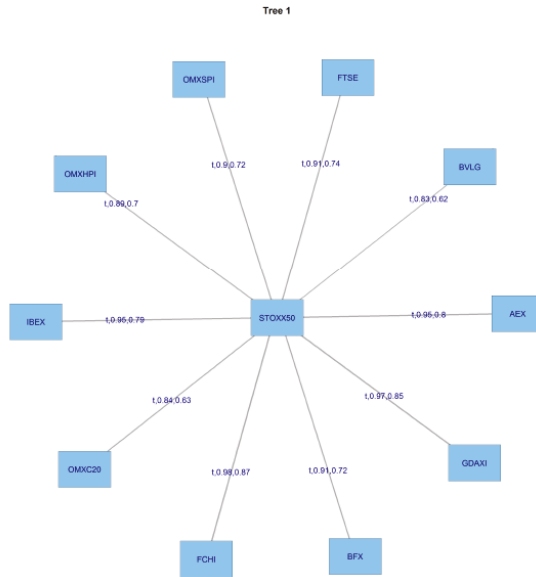

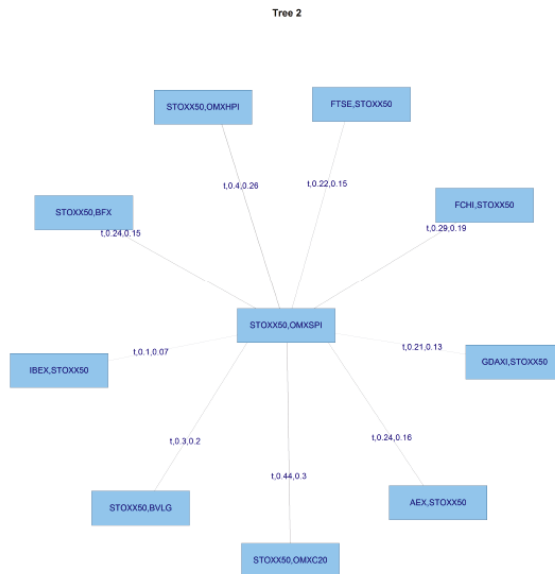**Figure 9.** Results-C-Vine Tree-1 GFC.

Figure 10. C-Vine Tree 2 GFC.

We are interesting in examing whether the major financial shock which constituted the GFC caused a noticeable change in dependencies?

Table 9 and 10 depicts the copulas chosen to capture dependency relationships during the GFC period.

Table 9. GFC C-Vine Copula Structure.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	1	0	0	0	0	0	0	0	0	0	0
GDAXI	9	9	0	0	0	0	0	0	0	0	0
FCHI	11	11	5	0	0	0	0	0	0	0	0
AEX	5	5	11	2	0	0	0	0	0	0	0
IBEX	2	2	2	11	10	0	0	0	0	0	0
STOXX50	10	10	10	10	11	3	0	0	0	0	0
OMXC20	3	3	3	3	3	11	4	0	0	0	0
OMXSPI	4	4	4	4	4	4	11	7	0	0	0
OMXHPI	7	7	7	7	7	7	7	11	8	0	0
BVLG	8	8	8	8	8	8	8	8	11	11	0
BFX	6	6	6	6	6	6	6	6	6	6	6

Table 10. GFC C-Vine Copula Specification Matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0	0	0	0	0	0	0	0	0	0	0
GDAXI	6	0	0	0	0	0	0	0	0	0	0
FCHI	2	2	0	0	0	0	0	0	0	0	0
AEX	3	1	5	0	0	0	0	0	0	0	0
IBEX	1	2	2	2	0	0	0	0	0	0	0
STOXX50	5	3	4	5	2	0	0	0	0	0	0
OMXC20	5	2	6	4	2	4	0	0	0	0	0
OMXSPI	2	2	2	2	2	2	2	0	0	0	0
OMXHPI	1	4	2	2	5	2	2	5	0	0	0
BVLG	2	2	2	2	2	2	2	2	2	0	0
BFX	2	2	2	2	2	2	2	2	2	2	0

A comparison of the entries in Table 10, the copula specification matrix for the GFC, with those in Table 5, the pre-GFC copula specification matrix, reveals that there is much less us of Gaussian copulas, 3 in Table 10, compared with 11 in Table 5. There is now a much greater use made of the Student T copula, on 36 occasions in Table 10, compared with 18 in Table 5. The use of the Gumbel copulas has increased from 1 to 4 occasions and the Clayton copula is only used on 2 occasions compared with 5 pre-GFC. The use of the Frank copula has declined from 15 to 6, whilst the Joe copula, now makes 2 appearances compared to 1 pre-GFC. The massive expansion of the use of the Student t copula, together with the other changes mentioned, is consistent with greater weight being placed on the tails of the distribution durng the GFC period.

The dependencies are captured in the Tau matrix shown in Table 11. A comparison of the values in Table 11, the tau matrix for the GFC, with those in Table 8, the tau matrix for the GFC period, reveals that the relationships have become more pronounced. If we look at the dependencies in the bottom row of Table 11, in 7 from the total of 10 cases the dependencies have increased. It is also true that there has also been a marginal increase in negative dependencies, from 10 pre-GFC to 12 during the GFC, but the values of these are of a low order. The picture that emerges from Table 11 is one of an increase in dependencies between these major European stock markets during an economic down-turn.

Table 11. GFC C-Vine Copula Tau matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
GDAXI	0.015375	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
FCHI	−0.014081	−0.039582	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
AEX	0.056810	−0.066848	−0.006416	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
IBEX	−0.025651	−0.042298	−0.058784	−0.083306	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
STOXX50	0.036874	0.051407	0.139010	−0.030941	0.108496	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
OMXC20	0.153082	0.015462	0.035938	0.124012	−0.006829	0.041608	0.000000	0.000000	0.000000	0.000000	0.000000
OMXSPI	0.131300	0.032328	−0.067132	−0.064857	0.029947	0.132041	0.162546	0.000000	0.000000	0.000000	0.000000
OMXHPI	0.124725	0.124603	0.080170	0.011683	0.194322	0.106381	0.103593	0.126699	0.000000	0.000000	0.000000
BVLG	0.143013	0.264481	0.064391	0.133197	0.195961	0.189751	0.154686	0.292464	0.153108	0.000000	0.000000
BFX	0.735492	0.702116	0.790637	0.844666	0.625231	0.867456	0.801941	0.630802	0.718425	0.720695	0.000000

4.4. Post-GFC Period

We will now turn our attention to the post-GFC period. In the case of the European markets, this is likely to be less-clear cut, given that it was characterised by economic turmoil related to the subsequent post-GFC European Sovereign debt crisis. Figure 11 displays the first tree post-GFC, and Figure 12 the second tree.
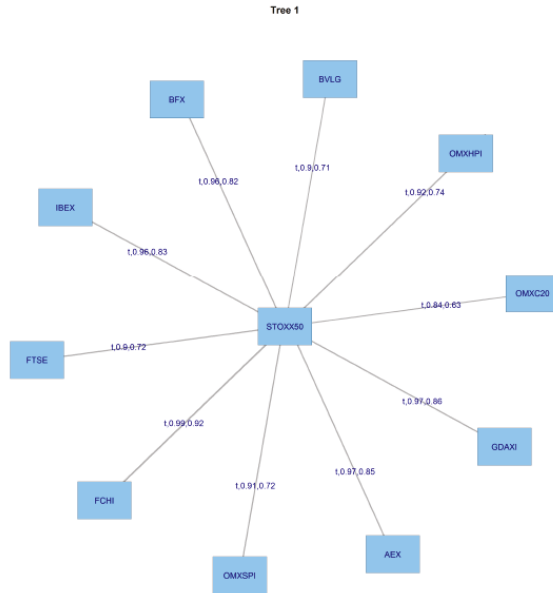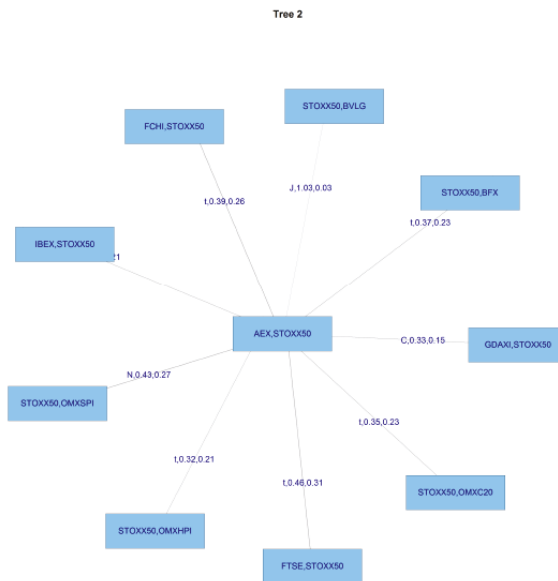
Figure 11. Results-C-Vine Tree-1 post-GFC.

Figure 12. Results-C-Vine Tree-2 post-GFC.

Table 12 shows the post-GFC C-Vine copula structure, and Table 13 the post-GFC C-Vine Copula Specification Matrix.

Table 12. Post-GFC C-Vine Copula Structure.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	2	0	0	0	0	0	0	0	0	0	0
GDAXI	11	7	0	0	0	0	0	0	0	0	0
FCHI	7	11	1	0	0	0	0	0	0	0	0
AEX	1	1	11	3	0	0	0	0	0	0	0
IBEX	3	3	3	11	9	0	0	0	0	0	0
STOXX50	9	9	9	9	11	10	0	0	0	0	0
OMXC20	10	10	10	10	10	11	5	0	0	0	0
OMXSPI	5	5	5	5	5	5	11	8	0	0	0
OMXHPI	8	8	8	8	8	8	8	11	4	0	0
BVLG	4	4	4	4	4	4	4	4	11	11	0
BFX	6	6	6	6	6	6	6	6	6	6	6

Table 13. Post-GFC C-Vine Copula Specification Matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0	0	0	0	0	0	0	0	0	0	0
GDAXI	3	0	0	0	0	0	0	0	0	0	0
FCHI	1	3	0	0	0	0	0	0	0	0	0
AEX	3	1	4	0	0	0	0	0	0	0	0
IBEX	2	4	2	3	0	0	0	0	0	0	0
STOXX50	3	5	1	5	4	0	0	0	0	0	0
OMXC20	3	4	1	1	4	2	0	0	0	0	0
OMXSPI	2	2	5	2	1	1	3	0	0	0	0
OMXHPI	4	2	2	1	2	5	5	1	0	0	0
BVLG	3	2	2	2	2	6	1	1	2	0	0
BFX	2	2	2	2	2	2	2	2	2	2	0

It can be seen in Table 13 that there is a marked change in the type of copula used to capture dependencies in the post-GFC period. The use of the Gaussian copula has risen from 3 during the GFC period to 10 in the post-GFC period, and the application of the Student t copula has dropped from 36 during the GFC to 24 in the post GFC period, whilst the use of the Clayton copula in the post-GFC period rises to 8 from 2 in the GFC period. The Gumbel copula is used on 6 occasions, whilst the Frank copula appears only 5 times, compared with 15 in the pre-GFC period. Finally, the Joe copula, is made use of on 1 occasion. The increase in the use of the Gaussian copula and the reduction in the use of the Student t copula suggests there is much less emphasis on the tails of the distributions in the post-GFC period.

The post-GFC tau matrix is shown in Table 14. The structure of dependencies that emerges in Table 14 is quite complex when compared to those of the GFC period. In the bottom row the positive dependencies captured in the tau statistics have increased in 7 of the total of 10 cases. In the GFC period there were 12 negative tau coefficients in the matrix, where as in the post-GFC period this number has reduced to 10. Thus, the broad picture that emerges in the post-GFC period, based on the use of C-Vine copulas, is that overall dependencies increased in the post-GFC period across the major European markets, in association with their experience of the European Sovereign debt crisis. The greater use of Gaussian copulas and the reduction in the use of Student t copulas in this period, suggests that tail behaviour was less important.

We now switch to the more flexible R-Vine framework to compare the two approaches.

Table 14. Post-GFC C-Vine Copula Tau matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
GDAXI	0.016107	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
FCHI	0.036144	0.035753	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
AEX	0.052784	0.024838	0.031775	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
IBEX	−0.130688	0.044635	0.051311	0.046822	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
STOXX50	0.047183	0.135747	−0.023586	−0.027756	0.086703	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
OMXC20	0.017433	0.131760	−0.035346	0.042679	0.088264	0.198073	0.000000	0.000000	0.000000	0.000000	0.000000
OMXSPI	−0.264456	−0.021457	−0.078220	−0.228646	0.003425	0.202873	0.059906	0.000000	0.000000	0.000000	0.000000
OMXHPI	0.118687	0.228880	0.072495	0.108799	0.324425	0.072847	−0.019366	0.131302	0.000000	0.000000	0.000000
BVLG	0.142026	0.226714	0.302870	0.255696	0.209531	0.019685	−0.207497	0.284246	0.239010	0.000000	0.000000
BFX	0.856993	0.629740	0.718863	0.912498	0.736706	0.707156	0.819042	0.721526	0.847664	0.818948	0.000000

5. R Vine Copulas

5.1. The Pre-GFC Period

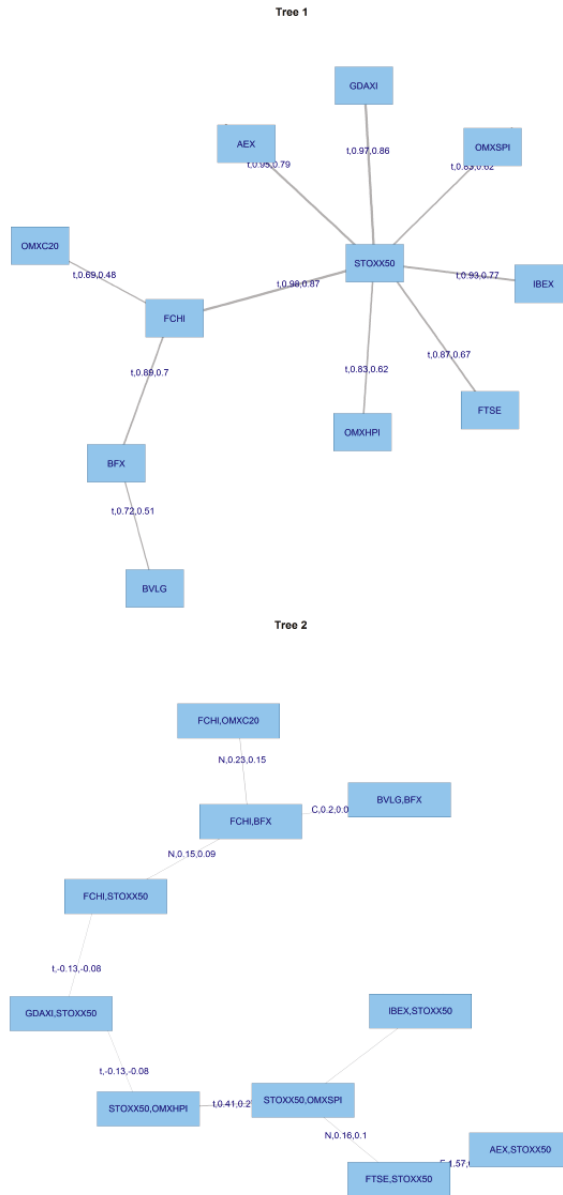The trees for the pre-GFC period are shown in Figures 13 and 14.

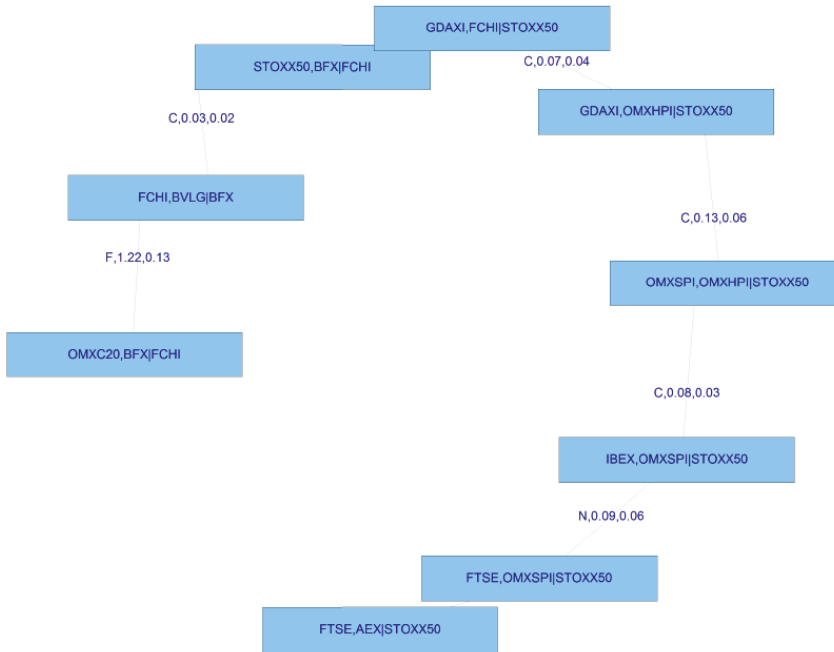**Figure 13.** Results-R-Vine Trees-1 and 2 pre-GFC.

Tree 3

Figure 14. Results-R-Vine Tree-3 pre-GFC.

It can be seen in Figures 13 and 14 above that the R Vine structure is more flexible. Tree 1 shows that a sub-group of the European markets are linked together; namely the Portuguese (BVLG), Brussels (BFY), the French (FCHI) and the Danish (OMXC20), they are then linked to the European Index (STOXX50). The other markets; Amsterdam (AEX), Germany (GDAXI), Stockholm (OMXSPI), Spain (IBEX), the UK (FTSE), and Helsinki (OMXHPI), have the strongest co-dependency with the European Index (STOXX50). This is also apparent in Tables 15 and 16 which show.

Table 15. Pre-GFC R-Vine Copula Structure.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	4	0	0	0	0	0	0	0	0	0	0
GDAXI	7	1	0	0	0	0	0	0	0	0	0
FCHI	10	7	5	0	0	0	0	0	0	0	0
AEX	11	10	7	8	0	0	0	0	0	0	0
IBEX	3	11	10	7	9	0	0	0	0	0	0
STOXX50	2	3	11	10	7	2	0	0	0	0	0
OMXC20	9	2	3	11	10	7	6	0	0	0	0
OMXSPI	5	9	2	3	11	10	7	7	0	0	0
OMXHPI	8	5	9	2	3	11	10	10	3	0	0
BVLG	1	8	8	9	2	3	11	11	10	11	0
BFX	6	6	6	6	6	6	3	3	11	10	10

Table 16. Pre-GFC R-Vine Copula Tau matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
GDAXI	0.070734	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
FCHI	0.086887	0.048472	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
AEX	0.154546	0.034327	0.035060	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
IBEX	0.019231	0.074440	0.103680	0.087828	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
STOXX50	−0.020726	0.063797	0.099572	0.036958	0.131488	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
OMXC20	−0.043319	−0.007375	−0.058887	0.091484	0.107138	0.009252	0.000000	0.000000	0.000000	0.000000	0.000000
OMXSPI	−0.104071	0.010599	−0.029020	0.083757	0.079020	0.030138	−0.016460	0.000000	0.000000	0.000000	0.000000
OMXHPI	0.050943	0.055597	0.037498	0.059558	0.034418	−0.002743	0.015683	0.133918	0.000000	0.000000	0.000000
BVLG	0.169959	0.105428	0.136164	0.268055	−0.083498	−0.082605	0.094934	0.145355	0.089392	0.000000	0.000000
BFX	0.797280	0.673160	0.766361	0.624435	0.620234	0.853751	0.868040	0.487667	0.699477	0.510377	0.000000

Table 15 shows the types of copulas fitted in the empirical analysis.

The advantage of the use of R Vines is apparent in Table 15. Complex patterns of dependency can be readily captured. It can be seen that at different dependencies conditioned across the same node six different copulas are used. For example, in column 1 the first copula used is the Clayton copula (no 3), followed by the Frank copula (no 5) for a couple of levels, then the Joe Copula (no 6), the Frank copula (no 5), two cases of the Gaussian (no 1), then the Gumbel (no 4), then the Frank copula again, and finally, the Student t (no 2). This variety of usage is apparent across Table 15 at various levels in the tree structures used to capture dependencies. The bottom row consists entirely of Student t copulas.

The copulas used to capture co-dependencies are different from the pre-GFC period C-Vine analysis. In that case, illustrated in Table 4; 11 Gaussian, 18 Student t copulas, 5 Clayton copulas, I Gumbel, 15 Frank copulas, and 1 Joe Copula were used. By contrast, in Table 15, 11 Gaussian, 18 Student t, 9 Clayton, 3 Gumbel, 12 Frank and 1 Joe copula are used. This follows, given that different co-dependencies are captured in the tree because there are not constraints on the pairings in R Vine copulas.

In the interests of brevity the details of the parameters estimated are not tabulated but the tau matrix, is shown in Table 16. The entries in Table 16 for R-Vines can be contrasted with those in Table 8 for C-Vines. Once again, given the nature of the analysis, the strongest dependencies between the various indices are captured by the entries in the bottom row of the table. Overall, the picture of dependencies is similar to those captured by the C-Vine analysis. The biggest change is in the first column of Table 16 in that the relationships between the FTSE and STOXX50, OMXC20 and OMXSPI have now become negative, but it has to be born in mind that the relationship is now conditioned on the much stronger relationship between the FTSE and BFX.

5.2. R-Vines GFC

Figure 15 provides the trees for the R-Vine analysis in the GFC period.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0	0	0	0	0	0	0	0	0	0	0
GDAXI	3	0	0	0	0	0	0	0	0	0	0
FCHI	5	5	0	0	0	0	0	0	0	0	0
AEX	5	2	4	0	0	0	0	0	0	0	0
IBEX	6	5	1	2	0	0	0	0	0	0	0
STOXX50	5	4	2	3	1	0	0	0	0	0	0
OMXC20	1	1	2	3	5	5	0	0	0	0	0
OMXSPI	1	2	5	3	5	1	1	0	0	0	0
OMXHPI	4	1	3	3	3	1	3	5	0	0	0
BVLG	5	1	5	2	2	2	1	1	3	0	0
BFX	2	2	2	2	2	2	2	2	2	2	0

Figure 15. Pre-GFC R-Vine Copula Specification Matrix.

The trees shown in Figure 16 indicate that dependencies have changed because of the influence of the GFC and the FTSE is now linked via the French FCHI to the STOXX50, whilst the OMXC20 and OMXSPI are now linked via the FCHI to the STOXX50. Previously, in the pre-GFC period the BVLG and the BFX were linked by the FCHI, but this is no longer the case.

Table 17 once again suggests the importance of capturing tail risk in financial and economic downturns plus the importance of fat-tailed distributions. Only 4 Gaussian copulas are applied in Table 17, where as the Student t copula dominates, being used on 38 occasions. There are 2 applications of the Clayton copula, 5 of the Gumbel and 4 of the Frank, whilst the Joe copula is used on 1 occasion.
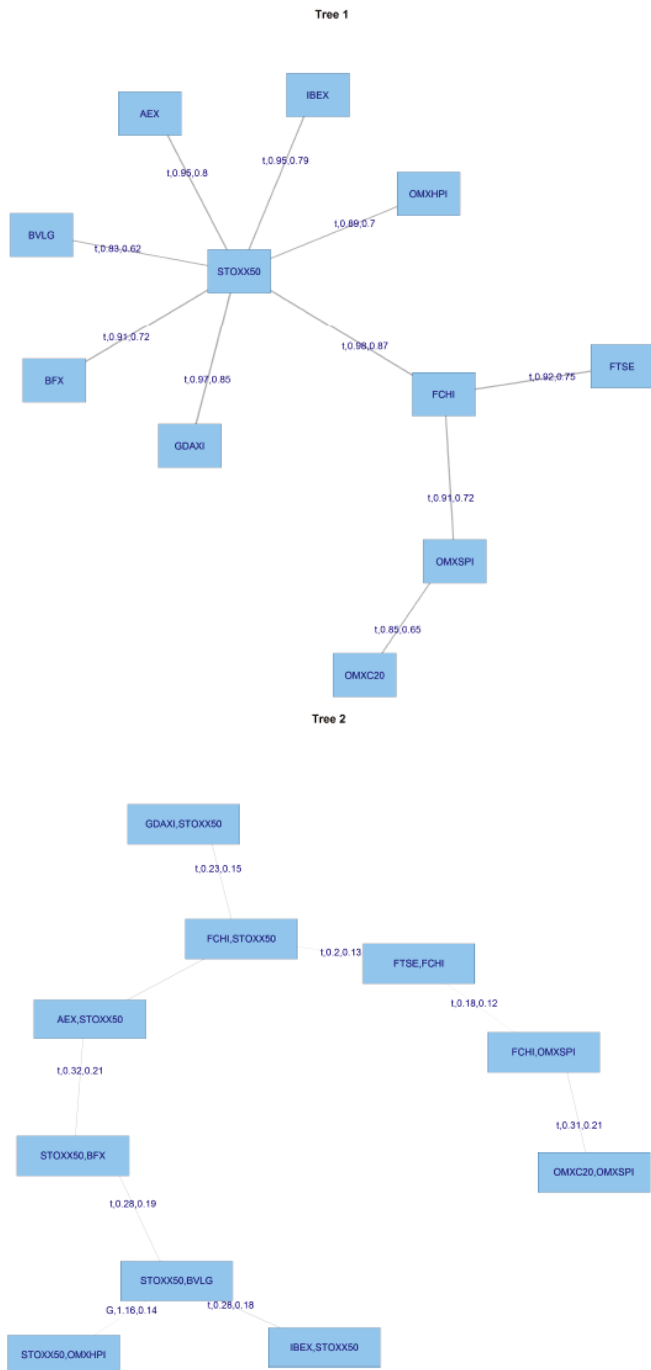
Figure 16. Results-R-Vine Trees-1 and 2 GFC.

Table 17. GFC R-Vine Copula Specification Matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0	0	0	0	0	0	0	0	0	0	0
GDAXI	3	0	0	0	0	0	0	0	0	0	0
FCHI	2	2	0	0	0	0	0	0	0	0	0
AEX	4	1	2	0	0	0	0	0	0	0	0
IBEX	5	5	2	2	0	0	0	0	0	0	0
STOXX50	3	2	1	6	4	0	0	0	0	0	0
OMXC20	2	1	2	5	2	2	0	0	0	0	0
OMXSPI	1	2	4	2	4	2	2	0	0	0	0
OMXHPI	5	2	2	2	2	2	2	2	0	0	0
BVLG	2	2	2	2	2	2	2	4	2	0	0
BFX	2	2	2	2	2	2	2	2	2	2	0

Table 18 provides details of the tau matrix for the GFC period. The change in dependencies in the R-Vine analysis following the GFC is complex and difficult to interpret in a clear-cut fashion. In terms of the dependencies captured in the bottom row of Table 18, 5 show and increase in their values, compared with the pre-GFC entries in Table 1, 6 but 5 also show a decrease. In terms of the whole matrix, the number of negative entries in Table 18 is 10, the same as the number in Table 16, but because of complex changes in patterns of dependencies, they now occur at predominantly different positions in the matrix.

We will therefore move on to the post-GFC R-Vine analysis.

5.3. Post-GFC R-Vines

Figure 17 shows the R-Vine trees in the post-GFC period. Figure 17 reveals that the relationships between the markets have changed in a complex manner in the post GFC period. It can be seen in tree 1 that the FTSE is now linked to the STOXX50 via the Dutch and French Indices. The Finnish, Danish and Swedish markets are also linked via the Durch and French markets to the STOXX50. The German and Spanish markets have individual links to the STOXX50, whilst the Portuguese market is linked via the Belgian index to the STOXX50. Table 19 shows the types of copulas used to map dependencies in the post-GFC period.

The Gaussian copula is used on 9 occasions whilst the Student t copula again dominates with 25 entries in Table 19, a considerable reduction on the 38 times it was applied during the GFC period. The Clayton copula appears 4 times, the Gumbel on 2 occasions. Greater use is made of the Frank copula, which appears 8 times and finally the Joe copula is used on 4 occasions.

The tau dependency matrix is shown in Table 20.

The tau matrix in Table 20 shows that dependencies have again changed in a complex manner in the post-GFC period which coincides with the European Sovereign debt crisis. The large dependencies in the bottom row have increased in 6 of the 10 cases in the post-GFC period. However, there are 12 cases of negative relationships in Table 20 as opposed to 8 in Table 18 representing the GFC period. These changes are interesting but do not give a direct indication of the usefulness of R-Vine modelling. We undertake an empirical application in the next section, which features a Value at Risk, (VaR) analysis, and this provides an illustration of its use in risk-assessment.

Fink et al., (2017) [36] use a Markov-switching R-vine model to explore the existence of different global dependence regimes. They explore the relationships between stock and volatility indices in Asia, Europe and the USA. They confirm the presence of normal and abnormal regimes. Our analysis is different in that we choose a particular time period to represent the GFC, whereas they use smoothed rolling windows, in an attempt to tease out changes in parameters in a Markov-switching analysis. They mention greater reliance on the Gumbel copula in the GFC period which we similarly note becomes more prominent in this period.
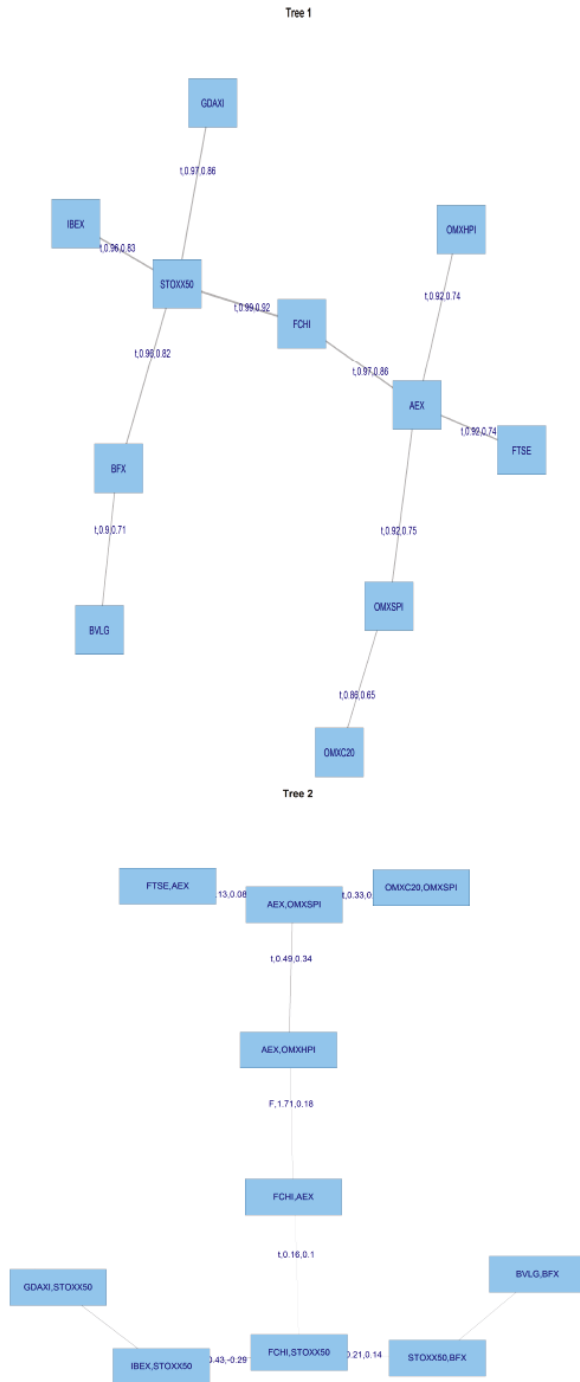
Figure 17. Results-R-Vine Trees-1 and 2 post-GFC.

Table 18. GFC R-Vine Copula Tau matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
GDAXI	0.020428	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
FCHI	0.040243	-0.050241	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
AEX	0.090111	-0.051601	0.034037	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
IBEX	0.174882	-0.023305	0.201117	0.060031	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
STOXX50	0.085350	0.133903	0.134654	0.037200	0.047610	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
OMXC20	0.074012	-0.026941	0.088942	0.067369	0.074824	-0.048008	0.000000	0.000000	0.000000	0.000000	0.000000
OMXSPI	-0.020369	-0.069500	0.087310	0.014613	0.038448	0.063628	-0.004543	0.000000	0.000000	0.000000	0.000000
OMXHPI	0.106049	-0.070206	0.093532	0.133080	0.076088	0.052608	-0.033877	0.031725	0.000000	0.000000	0.000000
BVLG	0.203436	0.147673	0.112357	0.125506	0.176601	0.208258	0.177847	0.135507	0.183817	0.000000	0.000000
BFX	0.651592	0.844666	0.725372	0.749098	0.867456	0.801941	0.790637	0.702116	0.625231	0.720695	0.000000

Table 19. Post-GFC R-Vine Copula Specification Matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0	0	0	0	0	0	0	0	0	0	0
GDAXI	6	0	0	0	0	0	0	0	0	0	0
FCHI	2	3	0	0	0	0	0	0	0	0	0
AEX	5	5	4	0	0	0	0	0	0	0	0
IBEX	1	5	5	2	0	0	0	0	0	0	0
STOXX50	1	6	5	1	1	0	0	0	0	0	0
OMXC20	3	5	3	6	3	1	0	0	0	0	0
OMXSPI	6	2	1	4	5	3	1	0	0	0	0
OMXHPI	5	1	5	2	2	1	2	2	0	0	0
BVLG	2	2	2	5	2	2	2	2	2	0	0
BFX	2	2	2	2	2	2	2	2	2	2	0

Table 20. Post-GFC R-Vine Copula Tau matrix.

	FTSE	GDAXI	FCHI	AEX	IBEX	STOXX50	OMXC20	OMXSPI	OMXHPI	BVLG	BFX
FTSE	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
GDAXI	0.021969	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
FCHI	-0.025819	0.051738	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
AEX	0.109136	-0.042114	0.091242	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
IBEX	0.068208	-0.055661	0.014128	0.092498	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
STOXX50	-0.073813	0.023514	0.001263	-0.036656	0.017084	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
OMXC20	0.015055	-0.035009	0.063244	0.067810	0.060644	0.183691	0.000000	0.000000	0.000000	0.000000	0.000000
OMXSPI	0.013031	0.069506	-0.091398	0.144184	0.117396	0.071469	-0.043601	0.000000	0.000000	0.000000	0.000000
OMXHPI	0.140392	-0.019256	0.053089	0.065557	-0.070606	-0.119362	0.192945	-0.047774	0.000000	0.000000	0.000000
BVLG	0.211872	0.082607	0.329059	0.184564	-0.294979	-0.283996	0.101914	0.135948	0.179939	0.000000	0.000000
BFX	0.657377	0.748438	0.750231	0.744893	0.856993	0.819042	0.855328	0.912498	0.818948	0.711242	0.000000

There are some further limited parallels between our work and that of Beil (2013) [37], who applies vine copula analysis to global indices and includes the DAX and STOXX in a broader global analysis which includes a separate period for 2007–2008, which is designated as the GFC period. However, Beil [37] makes relatively less use of the Gumbel copula in the truncated crisis period of this analysis. Furthermore, the overall sample of indices used is very different, and includes both US and Asian stock indices, and their corresponding volatility indices.

6. An Empirical Application

Empirical Example

We have used C-Vine and R-Vine Copulas to map dependence structures between some of the major European markets. These, in turn, can be used for portfolio evaluation and risk modelling. The R-Vine approach potentially gives better results than usual bivariate copula approach given that the copulas selected via Vine copulas are more sensitive to the asset's return distributions.

The co-dependencies calculated by R-Vine copulas can be used for portfolio Value at Risk quantification. We construct an equally weighted portfolio of the eleven market indices to explore the use of Vine copulas in modelling VaR using a portfolio example. The data used for this part of the analysis is from 3 January 2010 to 31 December 2011 with total 504 returns per asset, the eleven selected assets in the portfolio are our eleven European market indices. We use a 250 days moving window dynamic approach to forecast the VaR for this equally weighted portfolio which results in 254 forecasts. The main steps of the approach are as outlined below:

1. Convert the data sample to log returns.
2. Select a moving window of 250 returns.
3. Fit GARCH(1,1) with Student-t innovations to convert the log returns into an i.i.d. series. We fit the same GARCH(1,1) with student-t in all the iterations to maintain uniformity in the method, and this approach also makes the method a little less computationally intensive.
4. Extract the residuals from Step-3 and standardize them with the Standard deviations obtained from Step-3.
5. Convert the standardized residuals to student-t marginals for Copula estimation. The steps above are repeated for all the 10 stocks to obtain a multivariate matrix of uniform marginals.
6. Fit an R-Vine to the multivariate data with the same copulas as used in Section 1.
7. Generate simulations using the fitted R-Vine model. We generate 1000 simulations per stock for forecasting a day ahead VaR.
8. Convert the simulated uniform marginals to standardized residuals.
9. Simulate returns from the simulated standardized residuals using GARCH simulations.
10. Generate a series of simulated daily portfolio returns to forecast 1% and 5% VaR.
11. Repeat step 1 to 10 for a moving window.

The approach above results in VaR forecasts which whilst not dependent in time have the advantage of being co-dependent on the stocks in the portfolio. We use this approach as a demonstration of a practical application of the information about co-dependencies captured by the flexible Vine Copula approach applied to construct VaR forecasts. Figure 18 and Table 21 plots the 1% and 5% VaR forecasts along with original portfolio return series obtained from the method. The plot shows that the VaR forecasts closely follow the daily returns with few violations.

Table 13 below gives the results from Unconditional Coverage (Kupiec) and Conditional Coverage (Christoffersen) (Christoffersen [38] and Christoffersen, Hahn & Inoue, 2001 [39]) which are based on the number of VaR violations compared to the actual portfolio returns. According to the results in the table both the tests accept both 1% and 5% VaR models for the forecasting period.

As a direct contrast we also use our series of index returns combined into an equally weighted portfolio to construct a simulation of a VaR analysis based on the use of a GARCH(1, 1) model. The relative number of violations of the VaR set at 1% and 5% should indicate whether our vine copula approach better captures the complex structure of dependencies and is better suited to VaR analysis.

We proceed as follows:

1. Convert the data sample to log returns.
2. Select a moving window of 250 returns.
3. Fit GARCH(1,1) with Normal innovations to convert the log returns into an i.i.d. series.
4. Extract the fit from step-3 and simulate 1000 returns per asset.
5. Repeat step-3 and 4 for all the stocks and then calculate the portfolio return from the simulated series.
6. Generate a series of simulated daily portfolio returns to forecast 1% and 5% VaR.
7. Repeat step 1 to 10 for a moving window.
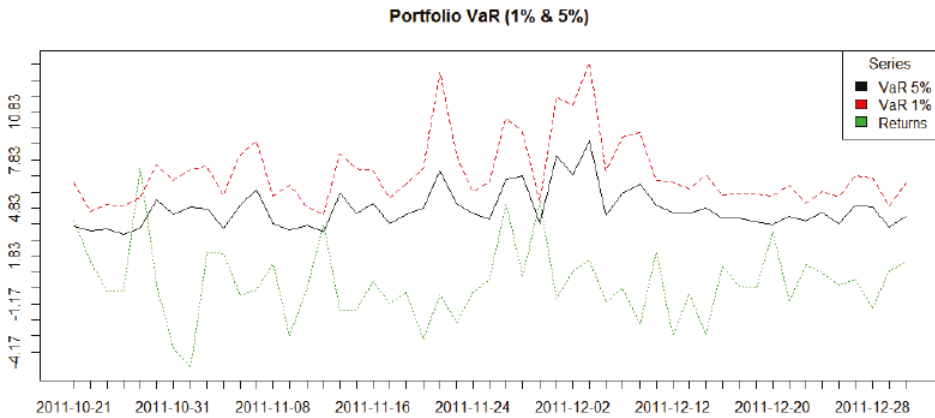
Figure 18. Portfolio Value-at-Risk analysis based on application of vine copulas.

A plot of the results of this exercise is shown in Figure 19 and Table 22. A brief glance at this shows that the application of a GARCH(1,1) model and the Gaussian distribution leads to multiple violations of the VaR 5 per cent (black line) and the VaR 1 percent (red line), whereas the VaR calculated on the basis of vine copulas, as shown in Figure 18, lead at most, to two breaches of the VaR.
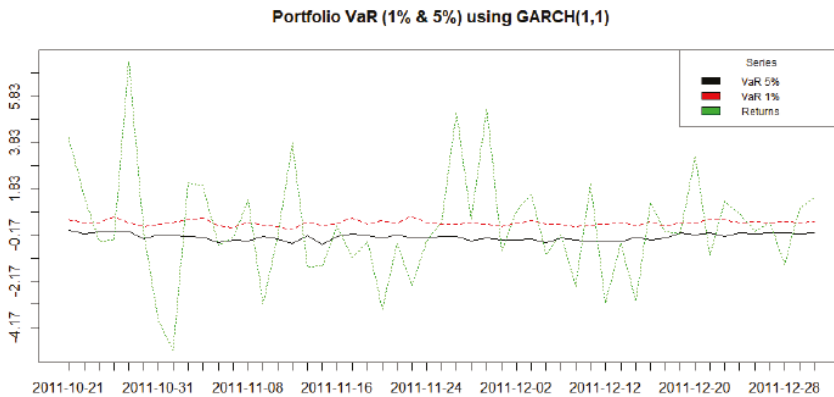
Figure 19. Portfolio Value-at-Risk analysis based on application of a GARCH(1,1) model.

Table 21. VaR Back-Test Results Vines.

	Expected Exceed	Actual Exceed	uc. H0	uc LRstat	uc. Critical	u LRp	uc Decision	cc H0	cc LRstat	cc Critical	cc LRp	cc Decision
1	2	1	Correct Exceedances	1.276879991	3.841458821	0.258479955	Fail to Reject H0	Correct Exceedances & Independent	1.317699151	5.991464547	0.517446275	Fail to Reject H0
2	0	0	Correct Exceedances	1.025134257	3.841458821	0.311304237	Fail to Reject H0	Correct Exceedances & Independent	1.025134257	5.991464547	0.598956006	Fail to Reject H0

p-Value > 0.05 (95% confidence) results in the acceptance of the null hypothesis.

Table 22. VaR Back-Test Results GARCH.

	Expected Exceed	Actual Exceed	uc. H0	uc LRstat	uc. Critical	uc LRp	uc Decision	cc H0	cc LRstar	cc. Critical	cc. LRp	cc. Decision
1	2	29	Correct Exceedances	106.27218836828 4	3.841458882069412	0	Reject H0	Correct Exceedances & Independent	106.28310314368 7	5.991464547107 98	0	Reject H0
2	0	21	Correct Exceedances	124.915738708942	3.841458882069412	0	Reject H0	Correct Exceedances & Independent	124.92665348434 5	5.991464547107 98	0	Reject H0

p-Value < 0.05 (95% confidence) results in the rejection of the null hypothesis.

7. Conclusions

In this paper we used the recently developed R Vine copula methods (see Aas et al. [7], Berg and Aas [11], Min and Czado [12] and Czado et al. [33]) to analyse the changes in the co-dependencies of ten European stock market indices and the composite STOXX50 index for three periods spanning the GFC: pre-GFC (January 2005–July 2007), GFC (July 2007–September 2009) and post-GFC periods (September 2009–December 2013). The results suggest that the dependencies change in a complex manner and there is evidence of greater reliance on the Student t copula in the copula choice within the tree structures for the GFC period which is consistent with the existence of larger tails to the distributions of returns. One of the attractions of this approach to risk-modelling is the flexibility available in the choice of distributions used to model co-dependencies. We demonstrated the calculation of portfolio VaR on the basis of these dependency measures and the method appears to work well on the basis of coverage ratio tests, which do not reject the null hypthesis in back-tests. This contrasts with the results on simulations to the same data set based on a GARCH(1, 1) model and the Gaussian distribution.

The main limitation is the static nature of the approach and dynamic applications are in the process of development. Breckmann and Czado [16] have proposed a COPAR model which provides a vector autoregressive VAR model for analysing the non-linear and asymmetric co-dependencies between two series. A more dynamic approach will be the subject of future work.

Author Contributions: Allen, McAleer and Singh, conceived the paper, Singh under took the analysis. Allen, McAleer and Singh wrote up the paper.

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

References

1.   Joe, H. *Multivariate Models and Dependence Concepts*; Chapman & Hall: London, UK, 1997.
2.   Nelsen, R. *An Introduction to Copulas*, 2nd ed.; Springer: New York, NY, USA, 2006.
3.   Joe, H. Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. *Lect. Notes-Monogr. Ser.* **1996**, *28*, 120–141.
4.   Bedford, T.; Cooke, R.M. Probability density decomposition for conditionally dependent random variables modeled by vines. *Ann. Math. Artic. Intell.* **2001**, *32*, 245–268.
5.   Bedford, T.; Cooke, R.M. Vines—A new graphical model for dependent random variables. *Ann. Stat.* **2002**, *30*, 1031–1068.
6.   Kurowicka, D.; Cooke, R.M. *Uncertainty Analysis with High Dimensional Dependence Modelling*; John Wiley: Chichester, UK, 2006.
7.   Aas, K.; Czado, C.; Frigessi, A.; Bakken, H. Pair-copula constructions of multiple dependence. *Insurance Math. Econ.* **2009**, *44*, 182–198.
8.   Schirmacher, D.; Schirmacher, E. Multivariate dependence modeling using pair-copulas. In Proceedings of the Society of Actuaries: 2008 Enterprise Risk Management Symposium, Chicago, IL, USA, 14–16 April 2008. Available online: http://www.soa.org/library/monographs/other-monographs/2008/april/2008-erm-toc.aspx (accessed on 1 September 2017).
9.   Chollete, L.; Heinen, A.; Valdesogo, A. Modeling international financial returns with a multivariate regime switching copula. *J. Financ. Econ.* **2009**, *7*, 437–480.
10. Heinen, A.; Valdesogo, A. Asymmetric CAPM dependence for large dimensions: The canonical vine autoregressive model. In *CORE Discussion Papers 2009069*; Universite catholique de Louvain: Louvain-la-Neuve, Belgium, 2009.

11. Berg, D. Copula goodness-of-fit testing: An overview and power comparison. *Eur. J. Financ.* **2009**, *15*, 675–701.

12. Min, A.; Czado, C. Bayesian inference for multivariate copulas using pair-copula constructions. *Accept. Publ. J. Financ. Econom.* **2010**, *8*, 511.

13. Smith, M.; Min, A.; Czado, C.; Almeida, C. Modeling longitudinal data using a pair-copula decomposition of serial dependence. *J. Am. Stat. Assoc.* **2010**, *492*, 1467–1479.

14. Allen, D.E.; Ashraf, A.; McAleer, M.; Powell, R.J.; Singh, A.K. Financial dependence analysis: Applications of vine copulas. *Stat. Neerlandica* **2013**, *67*, 403–435.

15. Patton, A.J. Copula based models for financial time series. In *Handbook of Financial Time Series*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 767–785.

16. Brechmann, E.C.; Czado, C. COPAR—Multivariate time-series modelling using the COPula AutoRegressive model. In *Working Paper*; Technical University of München: München, Germany, 2012.

17. Sklar, A. *Fonctions de Repartition a n Dimensions et Leurs Marges*; Publications de l'Institut de Statistique de L'Universite de Paris: Paris, France, 1959; Volume 8, pp. 229–231.

18. Joe, H.; Li, H.; Nikoloulopoulos, A. Tail dependence functions and vine copulas. *J. Multivar. Anal.* **2010**, *101*, 252–270.

19. Geidosch, M.; Fischer, M. Application of vine copulas to credit portfolio risk modeling. *J. Risk Financ. Manag.* **2016**, *9*, 4, doi:10.3390/jrfm9020004.

20. Fischer, M.; Kraus, D.; Pfeuffer, M.; Czado, C. Stress testing German industry sectors: Results from a vine copula based quantile regression. *Risks* **2017**, *5*, 38, doi:10.3390/risks5030038.

21. Panagiotelis, A.; Czado, C.; Joe, H.; Stöber, J. Model selection for discrete regular vine copulas. *Comput. Stat. Data Anal.* **2017**, *106*, 138–152.

22. Bedford, T.; Daneshkhah, A.; Wilson, K.J. Approximate uncertainty modeling in risk analysis with vine copulas. *Risk Anal.* **2016**, *36*, 792–815, doi:10.1111/risa.12471.

23. Scheffer, M.; Weiss, G.N.F. Smooth nonparametric Bernstein vine copulas. *Quant. Financ.* **2017**, *17*, 139–156.

24. Aas, K. Pair-copula constructions for financial applications: A review. *Econometrics* **2016**, *4*, 43, doi:10.3390/ econometrics4040043.

25. Fermanian, J.D. Recent developments in copula models. *Econometrics* **2017**, *5*, 34, doi:10.3390/econometrics5030034.

26. Kurowicka, D.; Cooke, R.M. A parametrization of positive definite matrices in terms of partial correlation vines. *Linear Algebra Its Appl.* **2003**, *372*, 225–251.

27. Cooke, R.M.; Joe, H.; Aas, K. Vines Arise Chapter 3. In *DEPENDENCE MODELING Vine Copula Handbook*; Kurowicka, D., Joe, H., Eds.; World Scientific Publishing Co.: Singapore, 2011.

28. Dissman, J.F. Statistical Inference for Regular Vines and Application. Master's Thesis, Technische Universitat München, München, Germany, 2010.

29. Brechmann, E.C.; Schepsmeier, U. Modeling Dependence with C- and D-vine Copulas. The R-package CDVine. Available online: http://cran.r-project.org/web/packages/CDVine/vignettes/CDVine-package.pdf (accessed on 1 September 2012).

30. Morales-N'apoles, O.; Cooke, R.; Kurowicka, D. About the Number of Vines and Regular Vines on n Nodes. Available online: https://repository.tudelft.nl/islandora/object/uuid:912abf55-8112-48d2-9cca-323f7f6aecc7?collection=research (accessed on 10 September 2017).

31. Vuong, Q.H. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* **1989**, *57*, 307–333.

32. Kurowicka, D. Optimal truncation of vines. In *Dependence Modeling: Handbook on Vine Copulae*; Kurowicka, D., Joe, H., Eds.; World Scientific Publishing Co.: Singapore, 2011.

33. Czado, C.; Schepsmeier, U.; Min, A. Maximum likelihood estimation of mixed C-vines with application to exchange rates. *Stat. Model.* **2011**, *12*, 229–255.

34. Mendes, B.V.D.M.; Semeraro, M.M.; Leal, R.P.C. Pair-copulas modeling in finance. *Financ. Mark. Portf. Manag.* **2010**, *24*, 193–213.

35. Brechmann, E.C.; Czado, C. Risk Management with High-Dimensional Vine Copulas: An Analysis of the Euro Stoxx 50. *Stat. Risk Model.* **2013**, *30*, 307–342.

36. Fink, H.; Klimova, Y.; Czado, C.; Stober, J. Regime switching vine copula models for global equity and volatility indices. *Econometrics* **2017**, *5*, 3, doi:10.3390/econometrics5010003.

37. Beil, M. Modeling Dependencies among Financial Asset Returns Using Copulas. Master's Thesis, Technische Universitt München, München, Germany, 2013.

38. Christoffersen, P. Evaluating Interval Forecasts. *Int. Econ. Rev.* **1998**, *39*, 841–862.

39. Christoffersen, P.; Hahn, J.; Inoue, A. Testing and Comparing Value-at-Risk Measures. *J. Empir. Financ.* **2001**, *8*, 325–342.

Adoption of Falsified Medical Products in a Low-Income Country: Empirical Evidence for Suriname

Philip Hans Franses [1,*] and Madesta Lede [2]

[1]   Econometric Institute, Erasmus School of Economics, 3062PA Rotterdam, The Netherlands
[2]   Department of Economics, Anton de Kom University, Leysweg 86, P.O. Box 9212, Paramaribo, Suriname; madestaml@hotmail.com
*   Correspondence: franses@ese.eur.nl

Abstract: Based on detailed shipping figures for Suriname's main harbour in Paramaribo, we estimate the total shipments (in kilograms) of original and falsified medical products for 1996–2008 across five product categories. Using various time series techniques and diffusion models, we document that total cumulative shipments of falsified products make about 40% of total shipments. We observe that there are apparently two distinct sets of consumers for original and for falsified products. Subsequently, we survey more than 300 citizens of Suriname from various demographics and ask questions about their potential adoption of falsified medicines. We find that income, age, and family size have no correlation, while the way people are insured does. Hence, the two sets of consumers can roughly be identified and clear-cut policy suggestions are presented. *"The World Health Organization (WHO) estimates that up to 1% of medicines available in the developed world is likely to be counterfeited. This figure rises to 10% globally, although in some developing countries they estimate one third of medicines are counterfeit"* (Various internet sites consulted January 2010 and the best estimate we have).

Keywords: falsified products; medication; health risk; low-income country

1. Introduction

The diffusion of pharmaceutical products has received some attention in the recent literature; see for example [1]. It has been documented that the diffusion patterns show various similarities with the diffusion patterns of durable consumer goods, see for example [2]. As pharmaceutical products may experience the impact of regulatory regimes, diffusion patterns may change direction and slope due to these regimes, see for example [3]. In all studies available, there is a strong focus on the diffusion patterns in western countries, usually driven by data availability, although [2] also include data for various low-income countries. A common assumption across all studies is that the focus is on the producer of the pharmaceutical products, usually the US or European companies.

In this paper we also address the diffusion patterns of pharmaceutical products, where we relegate the focus on the actual shipments of such products in a low-income country. This change of focus then naturally has to include the fact that in low-income countries often many pharmaceutical products are potentially falsified products. The exact available amount of these falsified drugs is unknown, and therefore we will first provide a method to estimate the fraction of falsified products within a single low-income country.

Exact data on falsified drugs are usually not available, and certainly not in a low-income country. One way to estimate the size of the falsified market is to try to disentangle the falsified diffusion from the total diffusion, as is proposed in [4] for software products. An alternative would be to estimate it directly from the shipments data, if these are available, see for example [5].

The focus in the present paper is the diffusion of original and falsified medical products in the South American country Suriname. Based on detailed data on shipments of products in five distinct categories, we estimate the diffusion of original and falsified products for the sample from 1996 to 2008. With these data, we can analyze the correlations between the two diffusion patterns. Also, we examine the lead and lag structure across original and falsified products. Finally, we estimate the total cumulative shipments using the familiar Bass model, and we compare the total levels for each of the categories. So, we provide estimates of the total amount that in the end will have found its way to consumers, here in Suriname. Note that we are not interested in the actual figures, but merely in the fraction of falsified products in the total amount.

One key main finding in this paper is that falsified drugs and other medical products eventually make up around 40% of the total shipments in this particular country of Suriname. Note that this fraction exceeds the estimates of the WHO. A second finding is that the two diffusion processes are correlated contemporaneously, reacting similarly to the same outside shocks, but that there is no significant lead or lag relationship. This suggests that the original and falsified products each address a distinct segment of the market. This finding has led us to survey many individuals in Suriname and to ask a range of questions. We aimed to find the distinctive features of the two groups of consumers, and it seems that the key distinctive feature concerns insurance. Basically, poor insurance means that people are more likely to opt for fake medication.

The outline of the paper is as follows. First, in Section 2 we describe the first part of the data collection which entails the creation of the two series for original and for falsified products. In Section 3 we analyze these data using basic time series techniques and familiar diffusion models. Section 4 continues with a discussion of the outcomes of a survey amongst more than 300 citizens of Suriname. Section 5 concludes with a discussion of the main findings and their implications.

2. Data Collection on Adoption over Time

We have collected data for the South American country of Suriname. It is a small and open economy, with a single large (maritime) harbour in the capital city of Paramaribo. The country has many natural resources. It used to be a colony of the Netherlands, before it became independent in 1975. Economic growth slowed down since the beginning of the 1980s, but since 2000 there has been steady growth in GDP and an increase in welfare. The country is sparsely populated, with most of the citizens living in the coastal area. Although average welfare is on the rise, the income inequalities in Suriname are quite large. The country boasts a non-negligible group of multi-millionaires, but on the other hand it is estimated that more than 60% of the population is below international poverty levels.

Sales records for medical products simply do not exist. Hence, we will have to estimate the diffusion patterns of these products using alternative methods. We consulted the General Bureau of Statistics (GBS) in Suriname and we were able to collect annual data on the imports of products in five categories for the period 1996 to and including 2008. The product categories are: Pharmaceutical items; Medicine, for sale small scale; Medicine, not for sale small scale; Wound-covering materials; and Blood items. These imports are measured in kilograms and in total value (in US dollars). We decided to take the weights as the measurement unit to avoid correction for inflation in the exporting countries. Crucial for our purposes is that we have information on the countries of origin of the shipments. In Appendix A, we give the countries of origin, and we indicate from which countries which products in the five categories originate. Clearly, Suriname imports from a long list of countries, and there is also variation across the categories.

To estimate the shipments of original products and falsified products, we use the software piracy index, which can be obtained from www.nationmaster.com. This website presents a list of countries and gives an estimate of the fraction of products (here: computer software) that are most likely falsified. In Appendix B we provide a list of relevant countries for our product categories and the corresponding percentages. Like this list, we shall take it as likely that 23% of the pharmaceutical products that are imported from Japan amount to falsified products, and that as much as 82% of these products

originating from China are falsified. Of course, we shall never be certain whether these percentages also hold for our product categories, nor if this figure changes over time or amounts to a biased estimate, and hence the subsequent data should not be considered as exact amounts, but merely as estimates. At the same time, and for our purposes, the absolute numbers do not matter, as only their time series properties do. In order to allow the reader to verify the computations about which we report in the next section, we present our data in Appendix C. Graphical details and other aspects of these series will be discussed in the next section.

3. Analysis of Shipments Data

This section deals with the statistical analysis of the two diffusion series. We analyze the levels of the shipments, as in Appendix C, and the cumulated shipments. The latter are used to estimate the eventual total size of the market. The first pairs of series are analyzed to see if there are leading or lagging diffusion processes amongst the two series. We discuss the methodology, first for cumulative shipments and then for shipments themselves, and we discuss the results in each of the five categories.

3.1. Methodology for Cumulative Shipments

We denote the level of the original products shipments as O_t and the level of the shipments of falsified product as C_t where $t = 1996, 1997, \ldots, 2008$. We also compute the cumulative shipments, and label these as CO_t and CC_t. Graphs of the series for the five categories appear in Figures 1–5. The second panel of each of these figures suggests that the cumulative shipments obey a product life cycle that can also be observed in the diffusion of durable products. Hence, the first part of our methodology involves considering diffusion models to estimate the inflection points and the total levels of these cumulative series.
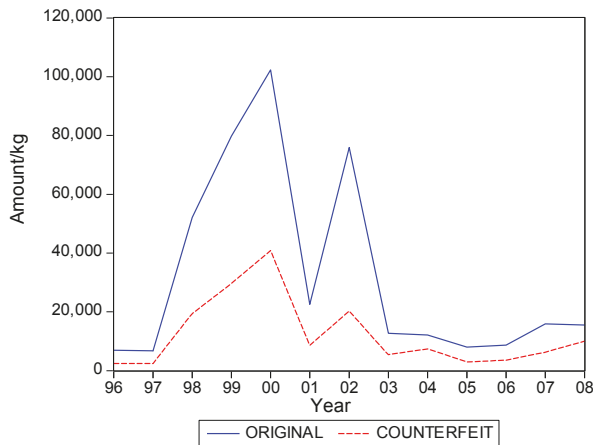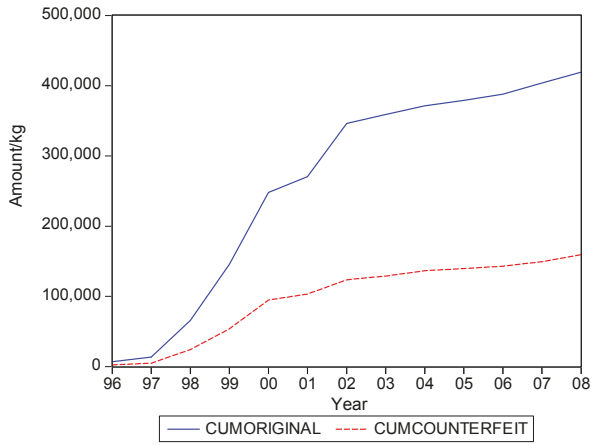

Figure 1. *Cont.*

Figure 1. Shipments and cumulative shipments of original and falsified products: Pharmaceutical items.
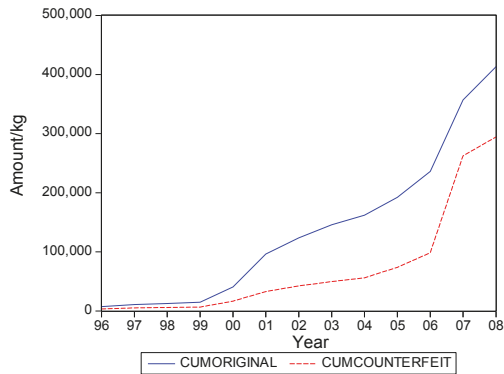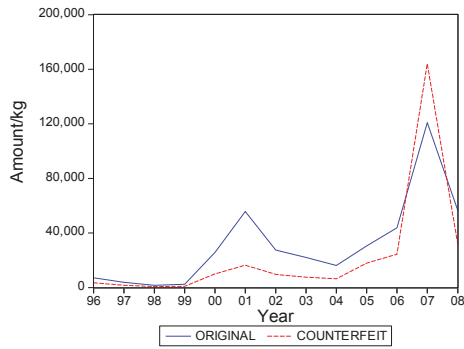

Figure 2. Shipments and cumulative shipments of original and falsified products: Medicine, for sale small scale. Vertical axis is the amount and horizontal axis is the year.
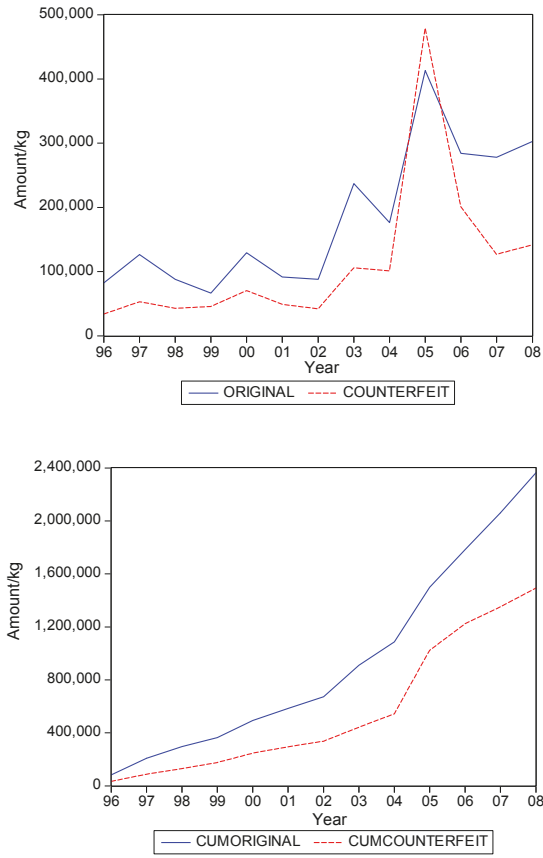
Figure 3. Shipments and cumulative shipments of original and falsified products: Medicine, not for sale small scale. Vertical axis is the amount and horizontal axis is the year.
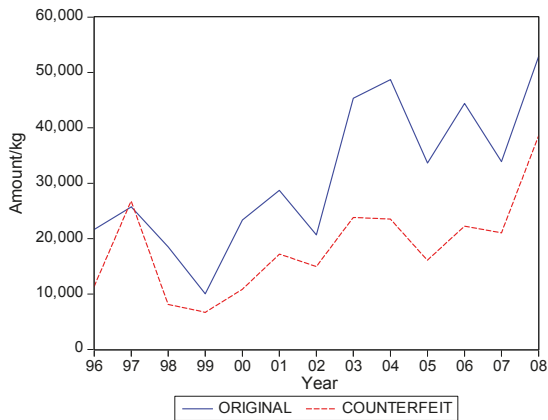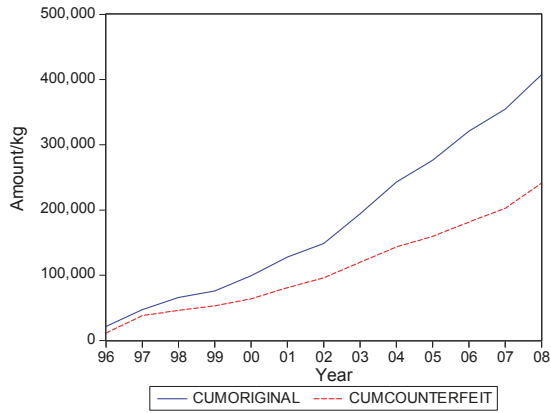

Figure 4. *Cont.*

Figure 4. Shipments and cumulative shipments of original and falsified products: Wound covering materials. Vertical axis is the amount and horizontal axis is the year.
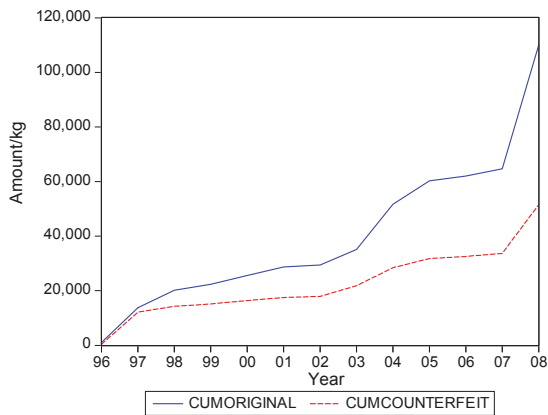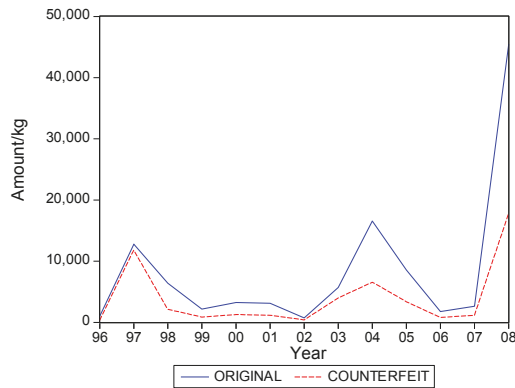

Figure 5. Shipments and cumulative shipments of original and falsified products: Blood items. Vertical axis is the amount and horizontal axis is the year.

For the CO_t and CC_t variables, we consider the familiar Bass [6] model, which in our notation can be written in regression format as

$$O_t = p_o m_o + (q_o - p_o)CO_{t-1} - \frac{q_o}{m_o}CO_{t-1}^2 + \varepsilon_t \tag{1}$$

and

$$C_t = p_c m_c + (q_c - p_c)CC_{t-1} - \frac{q_c}{m_c}CC_{t-1}^2 + v_t \tag{2}$$

for original and falsified products, respectively. The parameters p and q characterize the shape of the diffusion pattern, and determine the location of the inflection point I which can be calculated as

$$I = \frac{1}{p+q} \log\left(\frac{q}{p}\right) \tag{3}$$

The parameters in the Bass model may be difficult to estimate in case the data does not yet include the inflection point. If this happens, we replace the Bass model by the symmetric logistic curve, given by

$$OC_t = \frac{m_o}{1 + \exp(-\gamma_o(t - I_o))} \tag{4}$$

and

$$CC_t = \frac{m_c}{1 + \exp(-\gamma_c(t - I_c))} \tag{5}$$

for original and falsified products, respectively. Here the parameter γ characterizes the shape of the curve, and again I corresponds with the timing of the inflection point. Due to the imposed symmetric nature of the logistic curve, one can still reasonably adequately estimate the location of the inflection point, even when it is not included in the sample. The estimation routine is Nonlinear Least Squares. In both cases, that is the Bass model and the logistic curve, our interest lies in the parameters m, which is the ultimate level of total shipments, and I, which is the location of the inflection point.

In Table 1 we present the present the estimates of the total level. We also compute the fraction of m_c over $m_o + m_c$. As said, this latter fraction is most relevant for our purposes. Table 1 suggests that it can range from about 0.3 to 0.5, with an average close to 0.4. In words, about 40% of the eventual total shipments in medical products in Suriname could concern falsified products. It is interesting to see that the variation of the estimate of the fraction across the five product categories is quite small.

Table 1. Fraction of falsified shipments over total shipments, based on estimates of the eventual total level.

	Variable	Estimated Total Level		
		Original	Falsified	Fraction
(A)	Pharmaceutical items	396,066	148,209	0.272
(B)	Medicine, for sale small scale	1,283,915	1,262,614	0.496
(C)	Medicine, not for sale small scale	4,463,491	2,128,846	0.323
(D)	Wound-covering materials	582,769	386,618	0.399
(E)	Blood items	372,778	267,351	0.418
Average				0.382

Note: The Bass model is used for (A), while the logistic model is used for (B), (C), (D) and (E).

Table 2 gives the estimates of the inflection points. For two of the five categories, the years with these points coincide. For Medicine, not for sale small scale, the inflection point of falsified products occurred earlier, suggesting that the total level for this series is attained earlier than for the original products. For two categories the falsified series peak later, with Blood Items meaning that in the next

few years after 2008 still a substantial amount of falsified products will be imported in Suriname, at least, if the current situations persists.

Table 2. Estimated inflection points for original and falsified shipments.

Variable		Estimated Inflection Point	
		Original	Falsified
(A)	Pharmaceutical items	1999	1999
(B)	Medicine, for sale small scale	2010	2010
(C)	Medicine, not for sale small scale	2008	2006
(D)	Wound-covering materials	2005	2006
(E)	Blood items	2013	2018

Note: The Bass model is used for (A), while the logistic model is used for (B), (C), (D) and (E).

3.2. Methodology for Shipments

The graphs in Figures 1–5 seem to suggest that shipments of original and of falsified products show similar diffusion patterns. It is of interest to see if one of the two series is leading (or lagging).

To examine leads and lags relations, we estimate a vector autoregression of order 1 (VAR(1)), which consists of the following two equations

$$O_t = \mu_1 + \beta_1 C_t + \rho_1 O_{t-1} + \delta_1 C_{t-1} + \varepsilon_t \tag{6}$$

and

$$C_t = \mu_2 + \beta_2 O_t + \rho_2 O_{t-1} + \delta_2 C_{t-1} + \nu_t \tag{7}$$

Note that a genuine VAR(1) model would not include current C_t and O_t, but here it allows for a partial correction for current effects. The test of interest concerns the joint significance of ρ_1, δ_1 in (6) and of ρ_2, δ_2 in (7).

In Table 3 we present the *p* values of the F tests. Except for the category Medicine, not for sale small scale, we find that these parameters are not significantly different from zero. For the exceptional category, we learn that lagged shipments of originals leads current shipments of originals. So, in sum, there is no leading or lagging variable across originals and falsified products.

Table 3. *p*-values of the F test for the joint relevance of one-year lagged shipments of original products and one-year lagged shipments of falsified products (Each equation contains an intercept and current sales of the other variable).

Variable	Dependent Variable	
	Original	Falsified
Pharmaceutical items	0.933	0.806
Medicine, for sale small scale	0.119	0.382
Medicine, not for sale small scale	0.038	0.325
Wound covering materials	0.389	0.840
Blood items	0.653	0.529

Table 4 gives the contemporaneous correlation between the two series, and it is quite evident that this correlation is high.

Hence, we find that original and falsified medical products follow similar diffusion patterns, and that no series is leading or lagging the other series. When it is assumed that all individuals in Suriname are equally likely to need medical products, our findings suggest that the shipments apparently address two distinct segments of consumers. There is a demand for original products

and there is a demand for falsified products, with the relative demand for originals in the end being 1.5 times as large as demand for falsified products, as the fraction of falsified products in the total shipments is 40%.

Table 4. Correlation between shipments of original products and of falsified products.

Variable	
Pharmaceutical items	0.973
Medicine, for sale small scale	0.914
Medicine, not for sale small scale	0.871
Wound covering materials	0.835
Blood items	0.932
Average	0.905

4. Health Care in Suriname

We now proceed with an in-depth analysis of who the two distinct segments of consumers might be, that is, which characteristics qualify these two groups?

For that purpose we set out a survey amongst Surinamese individuals in Suriname. Suriname is a country with a low-income economy, which has a population of approximately 534,000 inhabitants (GBS 2013). It is estimated that more than 60% of the Surinamese population is below international poverty levels, while there also is a non-negligible group of multi-millionaires in this country. Although average welfare has been reported to be on the rise, the income inequalities in Suriname are quite large.

As is the case in countries all over the world, in Suriname the main responsibility concerning public health lies in the hands of the Surinamese government. Its Ministry of Health (MOH) is the government agency in Suriname which is responsible for ensuring the quality of health care and the formulation of pharmaceutical policy, including monitoring its quality. There is a National Medicines Policy 2004–2008 (NMP), which depicts the policy of the MOH with respect to the supply of medicines in Suriname. Important aspects of the NMP include ensuring the availability and accessibility of the selected essential medicines for the entire population, strengthening regulation bodies, and establishing a national quality control laboratory [7]. With regard to the quality control of medicines, the MOH reported that it has arranged and monitored the efficacy and safety of the drugs regulation in Suriname in accordance with the WHO recommendations [8], [9].

In order to safeguard quality control, the MOH has regulation bodies (Registration Authority, Pharmaceutical Inspection) in place and it is reported that the MOH works closely with other regulatory and enforcement institutions such as the Inspection of Customs and Excise, the departments of the Ministry of Trade and Industry and the Police Corps of Suriname [9]. In addition, the MOH has partnerships with the professional association of Pharmacists in order to determine the Surinamese Pharmacy Standard as a guide for self-regulation and other matters relating to the profession and service. MOH also partners with other countries in their region, with which it exchanges information about the quality of imported medicines as well as with producers to detect any wrongdoing [9].

In summary, the [9] reports that the quality of drugs which enter Suriname through official channels (particularly pharmacies) and which are distributed in this country can be deemed as good, due to the fact that: (1) all of these drugs are recorded in the register of packaged drugs; (2) all of these drugs are subjected to a quality control; and (3) in the entire distribution and sales channels these drugs are under the supervision of qualified pharmacists.

In spite of the abovementioned claims on the part of the MOH, one of the leading newspapers in Suriname reported in 2011 that no less than 85% of total sales in the legal circuit in the country concerned fake (SFFC) medicines (newspaper De Ware Tijd 17 September 2011, p. 1). This figure was adopted from an investigation on the prevalence of fake (SFFC) medicines in Suriname, which was reported to have been done by a private investigation agency on behalf of Pfizer (one of the largest pharmaceutical companies in the world). Despite the press release issued in their defence suggesting

that the claim was incorrect and that there was no such evidence, the [9] to date has seemed to fail to present solid documentation to underpin their arguments.

An investigation regarding the strengthening of the pharmaceutical quality assurance and legislation (from 2008 to 2010), which was commissioned by the MOH and which was done by Health Research For Action (HERA), shed light on some key concerns [8]. These concerns are that, although quality is central in the formulated NMP and even though this policy was prepared in accordance with the WHO recommendations on how to develop and implement a national drug policy and which was approved by the Government of Suriname in 2005, the applicable laws and regulations in the field of medicinal products are still obsolete. The current legislation appears incomplete and relies on the Medicines Act dating from 1896.

Also, while the MOH is able to guaranty the quality of the drugs that are formally admitted and legally imported, the quality of all other drugs outside these official processes cannot be guaranteed. It was also reported that there is little control on processes in the informal market. There is also a lack of routine sampling on the market, and the only quality control laboratory in Suriname is that of the Drug Supply Company Suriname (DSCS) ("Bureau Geneesmiddelen Voorziening Suriname"). This laboratory occasionally performs tests at the request of the Pharmaceutical Inspection, even though the DSCS laboratory has limited capacity and seemed to have performed without (adequate) professional guidance for years. Furthermore, DSCS itself is one of the largest importers of drugs in Suriname and therefore may not be seen as an independent institute. So, although the MOH attached importance to establishing a national quality control laboratory in Suriname (as stated in the NMP), to date there is no independent quality control laboratory for drugs in the country.

In addition, [8] shows that even though the MOH has several partnerships, there is insufficient professional consultation between the authorities concerning registration, inspection, pharmacovigilance (drug safety), the MOH and Council for the Essential Medicines Program ("Raad voor het Essentieel Geneesmiddelenprogramma"), and as a result the regulatory system cannot guarantee the quality of all drugs on the market. There is also no structural exchange of information between countries with stringent medicine regulatory systems in order to optimize the system. Also, the database of registered medicines is not accessible online, it is also inadequately automated, it is only accessible for professionals, and the applications that the professionals submit are only processed by hand.

Hence, taking all the above together, it seems that the Surinamese government apparently has a long way to go before they will be able to ensure that the quality of drugs on the Surinamese market meets the requirements set by the WHO (that is, in the formal system, and also reducing the size of the informal system).

With regard to the availability of medicines in Suriname, it is the case that some medicines are only available in pharmacies on a doctor's prescription, while other medications can be purchased without such a doctor's prescription, in either a pharmacy, a drugstore or sometimes even in shops and supermarkets. Concerning the medicines that are available to the Surinamese people, it is their health insurance company (or insurer, which could be the government again) that determines which medicines (the variety, the amounts, the brands and therefore the quality of medicines) pharmacies can provide to the clients on doctor's prescription. Although health insurance companies have to comply with the NMP, each health insurance company (insurer) has its own medicines register containing those medicines of which the cost are reimbursed by them, provided that these medicines are prescribed by a doctor. Bear in mind that HERA reports that there is an inadequately functioning (quality) control mechanism for pharmaceutical products. Below we will see that insurance is apparently a key factor.

Regarding the coverage of medical costs of the people in Suriname, the Pan American Health Organization (2010, p. 6) reported that the medical costs of approximately 34% of the Surinamese population are covered by private health insurance, company plans, or paid out-of-pocket. The State Health Insurance Fund (SHI) ("Stichting Staatsziekenfonds") covers the medical costs of 26% of the Surinamese population. The majority of individuals who are insured by this health insurance company

are employees of the Surinamese government, who are obliged to purchase their health insurance from this health insurer. The Ministry of Social Affairs and Housing (MSAH) covers the medical costs of 31% of the Surinamese population, who are coined as the under-privileged. The medical costs of the remaining 8% of the Surinamese population are covered by the Foundation Medical Mission Primary Health Care Suriname (FMM) ("Stichting Medische Zending Primary Health Care Suriname"), which is a private, non-profit, primary health care organization which focuses on communities in the interior part of the country and which is subsidized by the MOH for a great part.

5. Collecting Survey Data

In order to select the sample from the population of Surinamese people from the age of 18 and living in Paramaribo, we employed a multistage cluster sampling technique. GBS drew 30 clusters in Paramaribo (the capital city of Suriname) from which 20 households were selected. Each cluster had a pointer address providing the interviewers the first household where a respondent should be interviewed. The 19 remaining households in every cluster were selected by using a specific method. This method entails that the total of prospective 600 respondents from the 30 clusters (each with 20 households) were selected using the 'birthday rule'.

Data collection lasted about six weeks in February and March 2013 and the survey response rate was 52%, which amounts to 311 respondents. Most of the 298 prospective respondents who did not cooperate either refused to participate, of whom, interestingly enough, the majority of these respondents were people living in elite neighbourhoods, or they were not available for an interview even after repeated visits by the interviewers. It might be that the survey response is influenced by the fact that there have been several survey studies on other topics prior to this study, particularly in Paramaribo. Consequently, various prospective respondents had the impression that they were approached to partake in a study in which they had already participated. The majority of these respondents did not even give the interviewers the opportunity for further explanation.

6. Results from the Survey

When we asked 311 individuals how often they have, consciously, purchased falsified medicines, we had 283 valid cases with 71% who said "never", and hence 29% said "not never". 13% of the respondents answered with "Rarely", 10% with sometimes, 4% with "often" and 6 individuals (2%) said "very often". Of the 311 respondents who participated in this survey, 58% were females and 42% were males.

40% of all respondents had a monthly income with a maximum of Surinamese Dollar SRD800 (approximately Euro183 or USD 246). The monthly income for the 2 subsequent categories (between SRD800 and SRD1200 and between SRD1201 and SRD2000) was about 25% for each category. Only 11% of all respondents had a monthly income of at least SRD 2001 (approximately Euro616 or USD470). A histogram is provided in Figure 6.

Even though, in general, not more than 40% of the respondents had a monthly income of SRD800 or less, among those who are subsidized by the government this number was 64%. Irrespective of whether the costs of their health insurance were co-paid or fully paid by the employer (a distinction which is further ignored), 69% of all the participating respondents had some kind of health insurance. Among the remaining 31% of the respondents, only 8% were fully uninsured and paid their medical bills themselves (out-of-pocket). The other 23% of this group of respondents was either subsidized by the government through MSAH or through FMM.
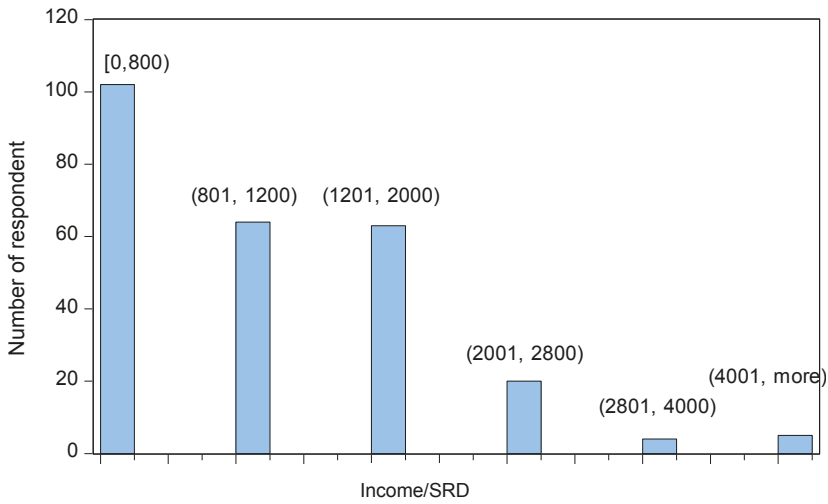
Figure 6. Income distribution of respondents, amounts in SRD.

Among those respondents who paid for their health insurance (the insured who either paid their health insurance themselves or whose employer co-paid or fully paid the health insurance, a distinction which is also further ignored in this investigation), 77% of them were insured by the SHIF, 22% were insured by a private health insurance company, and just 1% of all insured respondents were insured by both SHIF and a private health insurance company.

In our following analyses the respondents with a SHIF health insurance as well as one by a private insurance company are categorized under those who are insured by a private insurance company, as this group accounts for only 1% of all respondents. This is because these insurance plans are often perceived as being more favorable than those of SHIF, which is because private insurance companies often offer more elaborate insurance packages.

With regard to the respondents who are insured, only 36% indicated to have always had all the costs for medications reimbursed. In this category, those who are only insured by a private health insurance company (not the SHIF) represent the majority (62%). Just 27% of the respondents who were only insured by SHIF indicated to have always had all the costs for medications reimbursed.

Irrespective of how the medical expenses of the respondents are covered, about 80% of them indicated they have never suspected in the past that medication they obtained from a pharmacy, thus through the official channel (on doctor's prescription), could have been an SFFC medication. Only 26% of the respondents who suspected to have obtained SFFC medication via the official channel did not use these (supposed SFFC) medications. Almost one third of all respondents (27%) used these supposed SFFC medications often or more frequently after obtaining them.

When we evaluate the correlations across the answers to the question of whether one consciously purchased falsified medicines (with answers Never, Rarely, Occasionally, Often and Very often) with various demographics in Table 5, we see that the only significant correlation appears for the way people are insured. Hence, none of the variables of gender, income, age, or size of household associates with the conscious use of falsified products.

Table 5. Pearson Chi-square *p*-values of correlation between "conscious purchases of falsified medicines and demographics.

The way health insurance is organized	0.001
Which health insurer?	0.731
Gender	0.636
Age	0.334
Personal monthly income	0.214
Household monthly income	0.555
Number of adults in household	0.327
Number of children in household	0.369
Total number of household members	0.136

When we zoom in on the insurance variable in Table 6, which portrays the results to the question on the frequency of the consumers' use of falsified medicine, we see that the answer "Never" is given more often than expected by individuals who: (D) have a health insurance at a private company; (E) have insurance via relatives (for example parents); and (F) who state that their employer takes care of their insurance. At the same time, the answers "Rarely" and "Occasionally" are given more than expected by people who answer: (A) "I have no health insurance"; and (B) "the Ministry pays for my insurance due to my lack of income".

Table 6. Type of insurance against frequency of use of falsified medicine.

Answer to Question:					
Never	**Rarely**	**Occasionally**	**Often**	**Very Often**	**Insurance**
A	18	6 [+]	0	2	0
B	34	11 [+]	14 [+]	1	1
C	1	0	2	0	0
D	31 [+]	3	1	3 [+]	0
E	28 [+]	3	3	1	0
F	80 [+]	13	7	4	4 [+]
G	1	0	1	0	0
H	2	0	0	0	1 [+]

[+] Fraction more than expected; A: I have no health insurance; B: The Ministry pays for my insurance due to my lack of income; C: FMM pays for my insurance; D: I have a health insurance at a private company; E: I have insurance via relatives (for example parents); F: My employer takes care of my insurance; G: I am insured via my employer and via someone else; H: I am insured via my employer and at a private company.

These results clearly indicate that the way people are insured marks a dichotomy. Poor insurance associates with more conscious purchases and use of falsified medication.

7. Discussion of Results and Implications

In addition to the available global estimates of the WHO, we provided detailed estimates of shipments of original and falsified medical products in the South American country of Suriname. The precise numbers of these estimates are not particularly relevant, as we are interested in the relative numbers and the current and dynamic correlations across the two series. When we analyze the diffusion patterns, we see that total shipments (to be observed many years from now) will cover about 40% falsified products. Moreover, we see that current correlation is high, and that there are no relevant leads or lags. Hence, the two types of shipments seem to address two distinct clusters of consumers. And, indeed, when we collect survey data for Surinamese individuals we see that these clusters seem to associate with the way the individuals are insured. Properly insured people have a smaller tendency to purchase consciously falsified medicine. Here we foresee an important role for the government.

For Suriname we emphasize that the battle against falsified pharmaceutical products is one that asks much effort on the part of the Surinamese government, in particular in relation to insurance. As a

member of the WHO, this country needs to make greater effort to adopt the WHO schemes and to make more use of the information, instructions, and opportunities presented by this organization's global coalition of stakeholders, that is, the International Medical Products Anti-Counterfeiting Taskforce (IMPACT). The Surinamese government needs to tighten the relevant legislation and regulations so that substantial fines and even more appropriate prison sentences can be declared against those persons and organizations who engage in the import and sale of falsified pharmaceuticals. Given our results, we also recommend that the control with regard to the market of medical products, not only in the official distribution channel, is increased and regulations with respect thereto is tightened. Most certainly, control in the official distribution channel for medical products must be intensified.

At the same time, the Surinamese government needs to raise the awareness of the citizens. As it is assumed that the majority of the citizens have no background information on the pharmaceutical products they use, let alone on how to detect falsified versions among the original pharmaceuticals (in those cases where possible), the government must take the leading role in empowering the citizens by adequately informing them about how to determine the originality of pharmaceuticals. The Surinamese government must also inform the inhabitants about the fatal consequences of falsified versions of pharmaceutical products, and thus the great health risks these products may inflict on those who use them. The tightening of the government's policy in the battle against falsified pharmaceuticals is very likely to result in a significant reduction of the trade in these products in Suriname, and hence reduce the risk of exposure to the dangers of these products. As mentioned earlier, not only legislation and regulation policy needs to be tightened, but also the policy regarding control and the effective implementation thereof, needs to be taken into account.

Additional to the Surinamese government, there are at least three stakeholders that can also play a role in this battle. The first stakeholder is customs. Since it is suspected that large portions of falsified pharmaceuticals are smuggled into the country, this stakeholder will need to improve its control concerning the import of these products. Stricter control, combined with tight legislation and regulation will enable the custom officials to increasingly contribute to reducing the import of and the trade in falsified pharmaceuticals.

The second stakeholder, MOH, and more particular the Pharmaceutical Inspection, the control mechanism of the Surinamese government with regard to the pharmaceutical products, is responsible for inspecting and controlling the pharmaceutical market. This department is supported by the Office for Pharmaceutical Supply in Suriname. The latter focuses on the quality controls of locally produced and imported drugs. The Pharmaceutical Inspection, small in size, may lack the capability to optimally inspect and control the pharmaceutical business in Suriname. We suggest that the government increases its support to these organizations.

The final stakeholder is the consumer. As the consumer (thus the individuals, but also hospitals and similar institutions) ultimately determine the extent to which the sale of products increase or decrease, this stakeholder also needs to participate in reducing the trade in falsified pharmaceuticals in Suriname. Of course, before consumers are able to consciously choose to purchase and use original pharmaceutical products, first and foremost they will need to be adequately informed about the disadvantages of falsified pharmaceuticals. Furthermore, they will need to be informed about how to detect the falsified pharmaceuticals from the original versions, in those instances where this is possible.

Appendix A. Import from Countries

Table A1. Imports from countries.

Countries	Product Category				
	A	**B**	**C**	**D**	**E**
Antigua			x	x	
Dutch Antilles	x		x	x	x
Aruba				x	
Australia	x		x		
Austria			x	x	
Bahamas	x				
Barbados	x	x	x		
Belgium	x	x	x	x	x
Brazil	x	x	x		x
Canada	x	x	x	x	x
China	x	x	x	x	x
Colombia	x		x	x	
Costa Rica	x		x		
Cyprus	x		x		
Denmark	x		x		x
Egypt			x		
El Salvador			x		
United Emirates				x	
French Guyana	x	x		x	x
France	x	x	x	x	x
Germany	x	x	x	x	
Greece			x		
Guatemala	x		x		
Guyana	x		x	x	
Hungary					x
India		x	x	x	
Indonesia		x	x	x	
Israel			x		
Ireland	x	x			
Italy					x
Jamaica	x				x
Japan	x		x	x	
South Korea	x	x	x	x	x
St Lucia	x				
Mexico			x		x
Namibia	x				
Netherlands	x	x	x	x	x
New Zealand			x		x
Niger	x				
Norway					x
Panama	x	x	x	x	
Portugal		x		x	
Puerto Rico	x	x	x	x	
Russia			x		
San Marino			x		
South Africa	x				
Spain	x				
Sweden	x		x		
Switzerland	x	x		x	x
Taiwan			x	x	
Thailand		x	x	x	x
Trinidad and Tobago	x	x	x	x	
United Kingdom	x	x	x	x	
United States	x	x	x	x	x
Venezuela	x				

(A) Pharmaceutical items; (B) Medicine, for sale small scale; (C) Medicine, not for sale small scale; (D) Wound covering materials and (E) Blood items.

Appendix B. Percentage of Products that Is Falsified (Based on the Software Piracy Rate, www.nationmaster.com, Consulted December 2009)

Table A2. Countries and fraction of falsified products.

Antigua	50	Dutch Antilles	50
Aruba	50	Australia	28
Austria	25	Bahamas	50
Barbados	50	Belgium	25
Brazil	59	Canada	33
China	82	Colombia	58
Costa Rica	61	Cyprus	50
Denmark	25	Egypt	60
El Salvador	81	United Emirates	35
French Guyana	42	France	42
Germany	27	Greece	58
Guatemala	80	Guyana	50
Hungary	42	India	69
Indonesia	84	Israel	32
Ireland	34	Italy	49
Jamaica	50	Japan	23
South Korea	43	St Lucia	50
Mexico	61	Namibia	50
Netherlands	28	New Zealand	22
Niger	50	Norway	29
Panama	74	Portugal	43
Puerto Rico	44	Russia	73
San Marino	50	South Africa	34
Spain	43	Sweden	25
Switzerland	25	Taiwan	40
Thailand	78	Trinidad and Tobago	50
United Kingdom	26	United States	20
Venezuela	87		

Note: When a country does not appear on the list, we use the score 50.

Appendix C. The Data

Table A3. Shipments of original and falsified products: Pharmaceutical items, general (in kilograms).

Year	Original	Falsified
1996	6984.570	2437.430
1997	6726.670	2404.330
1998	52,109.53	19,346.47
1999	79,784.10	29,644.90
2000	102,338.1	40,856.93
2001	22,542.07	8584.930
2002	75,868.91	20,278.09
2003	12,747.33	5462.670
2004	12,117.31	7397.690
2005	8025.620	2938.380
2006	8639.090	3554.910
2007	15,909.47	6238.530
2008	15,565.72	10,033.28

Table A4. Shipments of original and falsified products: Medicine, for sale small scale (in kilograms).

Year	Original	Falsified
1996	7140.120	3402.880
1997	3721.980	1606.020
1998	1631.590	673.4100
1999	2360.470	818.5300
2000	25,662.14	9946.860
2001	55,775.76	16,261.24
2002	27,512.91	9549.090
2003	22,004.05	7445.950
2004	16,083.69	6314.310
2005	30,502.89	17,907.11
2006	43,727.94	24,332.06
2007	120,884.8	164,090.2
2008	56,326.23	32,061.77

Table A5. Shipments of original and falsified products: Medicine, not for sale small scale (in kilograms).

Year	Original	Falsified
1996	82,410.42	34,027.58
1997	126,437.3	53,107.74
1998	87,937.55	42,676.45
1999	66,519.33	45,686.67
2000	129,504.3	70,367.69
2001	91,668.09	48,919.91
2002	88,133.77	42,033.23
2003	237,083.1	105,841.9
2004	176,400.2	101,030.8
2005	413,139.4	479,416.6
2006	284,124.9	200,984.2
2007	278,071.7	127,044.3
2008	302,822.2	141,556.8

Table A6. Shipments of original and falsified products: Wound covering materials (in kilograms).

Year	Original	Falsified
1996	21,680.89	11,375.11
1997	25,731.19	26,771.81
1998	18,485.60	8128.400
1999	10,025.28	6677.720
2000	23,374.10	10,868.90
2001	28,716.47	17,209.53
2002	20,663.89	14,954.11
2003	45,351.56	23,804.44
2004	48,713.34	23,512.66
2005	33,667.00	16,122.00
2006	44,422.59	22,254.41
2007	33,909.00	21,029.00
2008	52,893.00	38,540.00

Table A7. Shipments of original and falsified products: Blood items (in kilograms).

Year	Original	Falsified
1996	1007.850	364.1500
1997	12,778.71	11,796.29
1998	6391.800	2106.200
1999	2160.360	838.6400
2000	3251.140	1259.860
2001	3112.780	1131.220
2002	731.4600	370.5400
2003	5704.680	3990.320
2004	16,571.03	6558.970
2005	8564.730	3352.270
2006	1759.190	777.8100
2007	2634.270	1126.730
2008	45,435.90	17,764.10

References

1. Staake, T.; Thiesse, F.; Fleisch, E. The emergence of counterfeit trade: A literature review. *Eur. J. Mark.* **2009**, *43*, 320–349. [CrossRef]
2. Desiraju, R.; Nair, H.; Chintagunta, P. Diffusion of new pharmaceutical drugs in developing and developed nations. *Int. J. Res. Mark.* **2004**, *21*, 341–357. [CrossRef]
3. Stremersch, S.; Lemmens, A. Sales growth of new pharmaceuticals across the globe: The role of regulatory regimes. *Mark. Sci.* **2009**, *28*, 690–708.
4. Givon, M.; Mahajan, V.; Muller, E. Software piracy: Estimation of lost sales and the impact of software diffusion. *J. Mark.* **1995**, *59*, 29–37. [CrossRef]
5. Franses, P.H.; Lede, M. *Diffusion of Original and Counterfeit Products in a Developing Country*; Econometric Institute Report 2010-08; Erasmus School of Economics: Rotterdam, The Netherlands, 2010.
6. Bass, F.M. A new product growth model for consumer durables. *Manag. Sci.* **1969**, *15*, 215–227. [CrossRef]
7. Ministry of Health. *Fake Medicine Briefing Ministry of Health (Dutch: Nepmedicijnen brieving Ministerie van Volksgezondheid)*; Media Release: Paramaribo, Suriname, 2011.
8. Health Research for Action (HERA). *Strengthening of Pharmaceutical Quality Assurance and Legislation*; Final Report, Part I—Main Report; Ministry of Health: Paramaribo, Suriname, 2010.
9. Ministry of Health. *National Medicines Policy Suriname 2005–2008 (Dutch: Het Nationaal Geneesmiddelenbeleid 2005–2008)*; Council for the National Vital Medicines Program, Government Publishing Service: Paramaribo, Suriname, 2005.

MDPI