UNIVERSITAT DE BARCELONA

Dependence and Systemic Risks in Financial Markets: Spatial and Upper Tail Analysis

Carlos Alberto Acuña González

RSITAT DE
ELONA

2022

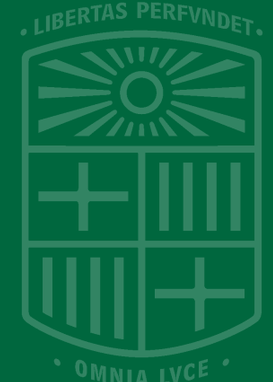
PhD in Business | Carlos Alberto Acuña

UNIVERSITAT DE
BARCELONA

PhD in Business

Dependence and Systemic Risks in Financial Markets: Spatial and Upper Tail Analysis

Carlos Alberto Acuña



UNIVE
BARC

PhD in Business

Thesis title:

Dependence and Systemic Risks in Financial Markets: Spatial and Upper Tail Analysis

PhD student:

Carlos Alberto Acuña

Advisors:

Catalina Bolancé
Salvador Torra

Date:
September 2022

UNIVERSITAT DE BARCELONA

Acknowledgment

I would like to express my gratitude to my supervisors, Dr. Catalina Bolancé and Dr. Salvador Torra, for their patience, guidance and intense dedication throughout the preparation of this thesis. In addition, I extend my thanks to the Riskcenter research group and to the Department of Econometrics, Statistics and Applied Economics for their support. Finally, I especially want to thank my parents for their unconditional support.

Contents

List of Figures

List of Tables

Introduction

This thesis focuses on systemic risk. According to Wikipedia this risk is defined as "financial system instability, potentially catastrophic, caused or exacerbated by idiosyncratic events or conditions in financial intermediaries". Over a number of decades, the financial system has suffered various situations that can be classified as systemic risk. This thesis analyses the most recent financial crises since 2008, with the bankruptcy of *Lehman Brothers*, until the 2020s, with the COVID-19 Pandemic which also affected the financial markets. The main objective is to study the financial market linkages using spatial dependence statistics and upper tail dependence deduced from a new non-parametric estimator of copulas.

In general, this thesis proposes alternative procedures for finding the relationship between systemic risk and spatial dependence between market returns, losses or alternative financial risk measures. Furthermore, the last chapter proposes a new inference using copulas which allows us to test the positive dependence between the most extreme positive losses between different financial markets, i.e., the upper tail dependence.

In the first chapter the feasibility and benefits of using neighbourhood relations between stock markets based on time criteria are analysed, such as the time differences between country capitals where each financial market operates and the simultaneous opening hours between these markets. These criteria are compared with the distance in kilometres between country capitals. The objective is to find clusters between neighbouring stock indices that are associated with dependent financial markets. We apply the idea of spatial dependence between markets and use the Moran's I statistic proposed by Moran (1950), calculated monthly for the period between January 2000 and December 2015, in order to analyse the spatial dependencies between market log-returns. The results show that the criterion based on simultaneous opening hours provides more relationships between neighbourhood markets. In addition, particularly between European markets, neighbourly relations were more intense during the 2008 financial crisis generated by the fall of *Lehman Brothers*. As a result of the events that occurred during this financial crisis, financial institutions were anxious to detect possible neighbour relationships between markets for systemic reasons in order to identify possible new sources of risk based on spatial dependence between indices. A Spanish version of this chapter was published in Acuña et al. (2018), with the title "*Análisis de la dependencia espacial*

entre índices bursátiles".

In Chapter 2 of this thesis the period analysed is extended until March 2021, incorporating Brexit and the COVID-19 pandemic in the study. This same period between January 2000 and March 2021 is also used in Chapters 3 and 4.

Chapter 2 analyses two new dynamic distance criteria applied to stock markets based on exogenous criteria; the well-known World Uncertainty Index (WUI) and the proposed Google Trends Uncertainty Index (GTUI). The chapter discusses the feasibility and benefits of these dynamic distances compared to alternative hour-based criteria. Using the new distance criterion to obtain the Moran's Index, the spatial dependence between the financial indices losses of 46 stock markets is analysed.

Specifically, Chapter 2 focuses on learning more about the possible relation between systemic risk and spatial dependence. This systemic risk is related to the most important financial crises of the last 17 years: the bankruptcy of *Lehman Brothers*, the sub-prime mortgage crisis, the crisis of European debt, Brexit and the COVID-19 pandemic, the latter also affecting the financial markets. This second chapter was published in Acuña et al. (2021), with the title "Dynamic distances between stock markets: use of uncertainty indices measures".

The aim of Chapter 3 is to study the spatial dependence between the risk measures associated with the financial indices losses, specifically the variance (volatility) and the Value-at-Risk (VaR). The distribution associated with these risk measures has a strong right skewness, i.e., a long and heavy right tail, so it is important to analyse how this can affect the inference based on the asymptotic normality of the global and local dependency tests based on Moran's statistic. With this aim in mind, in Chapter 3, the finite sample properties of inference based on global and local Moran's I statistics are analysed through a simulation study that assumes that the data are generated from distributions with different shapes (symmetric, asymmetric and heavy tailed distributions). Furthermore, we propose an alternative bootstrap based inference that improves Type I and Type II errors of asymptotic inference.

The spatial dependency between stock market risks has been discussed by using the definition of neighbour based on exogenous criterion derived from the Google Trends Dynamic Uncertainty Index (GTUI) proposed in Chapter 2. We show the impact of systemic risk on spatial dependency between risk measures related to the most significant financial crises since 2005: the *Lehman Brothers* bankruptcy, the

sub-prime mortgage crisis, the European debt crisis, Brexit and the COVID-19 pandemic, the latter also affecting the market economy. The risks are measured using the monthly variance or volatility and the monthly VaR of the filtered losses associated with the analysed stock indices. Specifically, the global spatial dependence between the risk measures of 46 stock markets and the local spatial dependence for 10 world reference stock markets are analysed. Chapter 3 was published in Bolancé et al. (2022), with the title "Non-Normal market losses and spatial dependence using uncertainty indices".

Chapter 4 focuses on the use of a proposed new kernel estimator of the copula to analyse the upper tail dependence between stock indices losses. A copula is a multivariate cumulative distribution function with marginal distributions $Uniform(0,1)$. This fact means that a classical kernel estimator does not work and this estimator must be corrected at bounds, which increases the difficulty of the estimation and, in practice, the bias correction limit may not provide the desired improvement. A quantile transformation of marginals is a way of improving the classical kernel approach. A first study using the standard normal quantile transformation was presented by Omelka et al. (2009). The objective in the development of this chapter lies in showing that a Beta quantile transformation is optimal, and a kernel estimator based on this transformation is analysed. In addition, the basic properties that allow the new estimator to be used for inference on extreme value copulas are tested. The results of a simulation study show how the new nonparametric estimator improves the alternative kernel estimators of copulas. We illustrate our proposal with an analysis of financial data. The application shows the Spanish index (IBEX 35) has upper tail dependence with European neighbour markets as well as with other markets such as those in UK, USA and Hong Kong. Chapter 4 was published in Bolancé and Acuña (2021), with the title "A new kernel estimator of copulas based on Beta quantile transformations".

In general, this thesis analyses the dependence between the financial stocks indices from different perspectives, taking into account the spatial dependence between financial returns, financial losses and financial risk measures. Finally, the thesis presents an innovative analysis of upper tail dependence between Spanish financial index losses against the losses of other market indices which are more or less close. Taken together, all this indicates that strong dependence is related to certain systemic risks caused by particular financial crises, but the intensity of the dependency changes from one financial crisis to another. Throughout the period analysed, the sub-prime crisis is revealed as the one with the greatest spatial dependence between financial markets. With regard to the other crises, the results depend on the analysed

variable, loss or risk measure. When the focus is on the risk measures, Brexit is revealed as a source of dependency between financial markets. If we analyse losses, it is the European debt crisis that reflects a strong spatial dependency between markets.

In addition to the four chapters described above, this thesis contains a final section of conclusions and an Appendix. The latter includes some complementary results in the form of tables and figures, the results of a procedure linked to the calculation of the statistic proposed in Chapter 4, and the R programs used to obtain the results of Chapters 3 and 4, in which the main methodological contributions of this thesis are presented.

Chapter 1

Analysis of spatial dependence between stock market indices

1.1 Introduction

The bankruptcy of *Lehman Brothers* in 2008 generated a global financial crisis, prompting the financial and/or insurance entities to seek to improve their understanding of the different types of relationships between markets, in order to be able to create more efficient strategies of investment from the perspective of the assumed risk. In practice, if managers have indices in their portfolios whose behaviour is affected by the crisis or by any other circumstance, knowing the spatial relationships, in addition to the temporary dependence, would help detect the level of contagion between neighbouring markets. This type of information would enable the generation of better portfolios than those obtained from the classic Mean-Variance perspective, i.e., with higher average earnings and lower risk.

However, as this chapter goes on to analyse, the concept of proximity between financial markets does not have to be linked to a geographical distance. Here, we focus on spatial relationships by comparing different ways of measuring proximity or distance between stock markets, with the aim of analysing which is the most appropriate measure to detect the spatial relationships between financial markets. Both types of measures, distance and similarity, can be used interchangeably and both are related, i.e., $D_{ij} = 1 - S_{ij}$, where D_{ij} and S_{ij} are, respectively, the distance and similarity between two financial markets i and j.

In the literature, geographical distance has been the most used measure to analyse spatial dependence between regions. However, this type of relationship loses validity when the aim is to analyse the simultaneous behaviour of stock indices, because they are not linked to a geographical point but rather to an hourly behaviour related to the opening and closing moment of the stock exchanges in each market. In this

chapter, we propose relationships between indices based on hourly criteria, such as the number of common opening hours of the financial markets associated with the different stock indices or the time differences between countries.

To infer the spatial dependence significance in the behaviour of different regions, spatial econometrics offers alternative statistics, among which is the global Moran's I statistic (see Moran, 1950) and its local version proposed by Anselin (1995). These statistics are based on the design of a spatial weight matrix which is created from the distances or similarities between the regions, in our case between the analysed stock markets. Spatial dependence analysis has more commonly been carried out for regional or urban studies; in contrast, their use in the financial field is scarce. However, this type of tool in the financial field helps detect if there are linkages between the different market clusters and when these linkages are produced.

The global Moran's I statistic indicates whether a variable is distributed completely randomly over the space or, conversely, whether there is a significant association of similar values between neighbouring regions. This statistic is characterized by contrasting the presence of a spatial joint dependency scheme, but it is not able to detect the existence of clusters of regions or stock markets with similar values. With this aim in mind, the use of a local Moran's I statistic is proposed that, as its name suggests, is used to infer stock market by stock market the spatial dependence each has with its neighbours.

The study of the relationships between stock indices has been approached in the literature from alternative perspectives. Some papers have focused on analysing possible interactions between stock indices, by implementing models that are based on the representation of the correlation matrix between indices without taking into account the distance between markets (see, for example, Hamao et al., 1990; Karolyi and Stulz, 1996; Longin and Solnik, 1995; Asgharian and Bengtsson, 2006; Martin and Dungey, 2007; Asgharian and Nossman, 2011). Alternatively, geographic and economic relationships are factors that have been used to explain the joint movements of the stock markets. For example, using gravity models, observed that geographic distances served to establish relationships between equity markets. However, using a function that describes the spatial correlation between the regions in an observed point pattern, called the "mark correlated function", Eckel et al. (2011) found that geographic proximity is irrelevant and does not influence the correlations of stock returns when distances go beyond 50 miles. From the semi-variogram, which helps to reflect the maximum distance of a variable in a region in the form of a point and the way this variable influences another point at different distances, Avilés et al. (2012) concluded that the dependencies between market returns are not related to geographic proximity but are strongly linked to foreign direct investment. Fernandez (2011) explored the notion of spatial depen-

dence by formulating a spatial version of the capital asset price model (S-CAPM), which allows alternative measures of distance between companies to be considered, such as those based on market capitalization. The results show that there are spatial effects between the markets. Asgharian et al. (2013) used a Durbin spatial model to investigate the extent to which countries' economic and geographic relationships affect their co-movements in their stock markets: here bilateral trade turns out to be the most adequate to capture co-variations in the returns. Arnold et al. (2013) modified the autoregressive spatial models and adapted them to the analysis of financial returns in order to compare three different types of spatial dependency: global dependencies, dependencies within industrial branches and local dependencies. Their results show that their approach can lead to more accurate forecasts for risk measures than standard approaches based on factoring the matrix of variances and covariances.

Recently, Weng and Gong (2016) propose a new approach, which is based on the use of copulas to define the matrix of spatial weights. Their model collects the spatiotemporal dependencies and, in addition, they show how conditional volatility is an important factor in determining the returns of the shares.

In a similar way to Avilés et al. (2012) and Asgharian et al. (2013), in this study we use spatial statistic methods to analyse how the links of stock indices depend on their relative distance or closeness. We compare two alternative distance measures and a similarity measure: the geographical distance between country capitals, the time difference between country capitals and the number of hours that two financial markets are open jointly. Specifically, in this chapter we analyse how the distance between markets should be measured to detect possible spatial relationships between a country and its closest neighbours. One of the alternative criteria to geographic distance is the number of common opening hours of the financial markets, which is a proximity measure. It is a measure of trade synchronization and it has already been used by Flavin et al. (2002) as a proxy to measure the ease of trade between two markets, interpreting that the greater the overlap in the opening hours of the markets, the greater their relationships. Alternatively, we propose a second measure of distance between markets that consists of calculating the time differences between the country capitals that host the analysed stock markets. This difference is based on time zones. A time zone is a region that shares a uniform standard time for legal, business and social purposes. Time in any world time zone can be expressed as a difference from Coordinated Universal Time (UTC). Finally, it should be noted that the three relationships used in this paper are exogenous to stock market indices, thus avoiding endogeneity problems that are presented in other relationships proposed in the literature, which are based on macroeconomic variables.

Chapter 1 is organised as follows. In Section 1.2 we define three types of relationship between indices: the geographical distance based on the kilometres between capitals, the time difference between cities and, lastly, the number of common opening hours in the markets; we then define the global and local Moran's I statistics. In Section 1.3 we describe the data and the results of the analysis carried out. Finally, Section 1.4 summarises the main conclusions of this chapter.

1.2 Methodology

1.2.1 Relationships between stock market indices

With the aim of capturing the interrelationship between stock indices, we use three criteria: two of distances and one of similarity between financial markets. These are formally defined as:

- $D_{ij}^1 =$ Geographical distance between the cities i and j to which the stock market indices belong.
- $D_{ij}^2 =$ Time distance (in hours) between the cities i and j to which the stock market indices belong.
- $S_{ij} =$ Similarity measured as common opening hours in stock markets i and j.

By depending on a georeferenced location, the distances D_{ij}^1 and D_{ij}^2 are considered physical variables. In contrast, the similarity S_{ij} is a relationship that does not depend on a geographical location. It is evident that the number of hours of joint opening in the stock markets is a measure of the commercial synchronicity, so that the greater S_{ij} the greater the proximity between the markets.

Once the distances or similarities relationships are defined, we obtain three $N \times N$ matrices, where N is the total number of stock markets analysed. When different units of measurement are used -hours and kilometres- it is necessary to use a transformation that makes them comparable and rescales them on the interval $[0, 1]$. Asgharian et al. (2013) propose a transformation that lies in the construction of contiguity matrices, that are denoted as C_{ij}^k, where $(k = 1, 2)$, and whose elements are defined as:

- Let S_{ij} be a similarity relationship, the elements of C_{ij}^1 are obtained as:

$$C_{ij}^1 = 1 - \frac{\max_j S_{ij} - S_{ij}}{\max_j S_{ij} - \min_j S_{ij}} \quad \forall \ i \neq j. \tag{1.1}$$

- Let D_{ij} be a distance relationship, the elements of C_{ij}^2 are obtained as:

$$C_{ij}^2 = 1 - \frac{D_{ij} - \mathrm{m}in_j D_{ij}}{\mathrm{m}ax_j D_{ij} - \mathrm{m}in_j D_{ij}} \quad \forall \ i \neq j. \tag{1.2}$$

This contiguity matrix ensures that the relationships between markets are between $[0, 1]$, such that, if C_{ij}^k is close or equal to 1 the markets i and j are close and if C_{ij}^k is close or equal to 0 the markets i and j are distant. It is important to mention that the contiguity matrix is not symmetrical, i.e., market i may be close (distant) to market j, but this does not necessarily imply that market j is close (distant) to market i.

1.2.2 Spatial dependency

We describe the statistics used to test whether there is spatial dependence between financial markets; specifically, between the logarithm of the returns (log-returns) of the stock market indices. Spatial dependency or spatial autocorrelation is the functional relationship between a point in space and what happens in other nearby points. In other words, the value of a variable in one area is not only explained by its internal condition but also by what occurs in neighbouring areas. This type of behaviour can be positive or negative, that is, if the studied phenomenon that affects an area occurs in the rest of the surrounding areas, generating a grouping, we have positive spatial dependence; on the other hand, negative spatial dependence occurs when the phenomenon studied in one region is presented in the opposite way in its neighbouring areas. In our case, when dealing with equity markets, we work with returns or losses, therefore, positive spatial dependence will imply that if a country has a stock market index that makes profits, the indexes of its neighbouring countries will also obtain profits, and the same is true of losses. Negative spatial dependence will imply greater diversification between neighbouring countries and, therefore, the losses of some will be offset by the benefits of others. To test the statistical significance of the spatial dependence, Moran (1950) proposes a statistic called Moran's I.

Let $P_{i,t}$ be the closing price of market i at period t and let $r_{i,t} = \log\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$ be the corresponding log-returns, the Moran's I statistic at period t is:

$$I_t = \frac{N}{S_0} \frac{\sum_{i=1}^{N} \sum_{i \neq j}^{N} W_{ij}(r_{i,t} - \bar{r}_t)(r_{j,t} - \bar{r}_t)}{\sum_{i=1}^{N}(r_{i,t} - \bar{r}_t)^2} \ \forall\, i \neq j. \tag{1.3}$$

where N is the number of stock indices, W_{ij} is the element in row i and column j

of the spatial weights matrix W. This matrix is defined by standardizing the rows of the contiguity matrix C^k, $k = 1, 2$, i.e., the sum of each row of C^k is 1; S_0 is the sum of all the elements of W and $\bar{r}_t$ is the mean of the log-returns at time t. In this chapter, the I_t defined in (1.3) is used to test the null hypothesis of spatial independence versus the alternative hypothesis of positive spatial dependence. Asymptotically, the Moran's I statistic follows a Normal distribution with known mean and variance (see Moran, 1950). This implies that the inference based on Normal distribution works when N is large.

Although global spatial dependence was observed for the entire sample of size N, it is difficult to deduce results that help reach a clear decision. For this reason, it is essential to locate which indices cause the existence of the global spatial dependence. With this aim in mind, Anselin (1995) defines the local Moran's index that allows us to carry out the individual analysis for each stock index i. The local spatial dependence index is:

$$I_{i,t} = \frac{(r_{i,t} - \bar{r}_t) \sum_{i \neq j}^{N} W_{i,j} (r_{j,t} - \bar{r}_t)^2}{\sum_{j=1}^{N} (r_{j,t} - \bar{r}_t)^2 / N}. \tag{1.4}$$

In the same way as the global index, the asymptotic distribution of the local index is Normal and, therefore, we are able to perform the contrast defined above individually for each stock index. [1]

The inference made from the global and local Moran's I statistics is based on the assumption that the data are independent and identically distributed (iid) as a Normal. However, in many analyses, such as the one proposed in this paper, the data is not normally distributed. Therefore, analysis of the effect of the non-normality of the data on the inference based on Moran's I is essential. Regarding this aspect, Griffith (2010) concludes that the assumption of normality is not essential for asymptotic inference based on Normal distribution. [2]

1.3 Empirical analysis

1.3.1 The data

In the analysis of spatial dependence between financial markets, which we show below, we study the stock market behaviour around the world before and during the

[1] In Chapter 3 we delve into the expressions of the mean and variance of the spatial dependence indices.

[2] In Chapter 3 we carry out the analysis of finite sample based inference with alternative distributions very different to the Normal.

2008 financial crisis. To carry out this analysis, we use monthly information on the returns for 46 stock market indices, between January 2000 and December 2015. We establish the beginning of the crisis as being in September 2008, coinciding with the collapse of *Lehman Brothers*[3]. Therefore, we have information on log-returns for the two sub-periods that comprise a total of 104 months before the crisis (B.C.) and a total of 88 months during the crisis (D.C.). We analyse 46 indices that are distributed in five geographical areas: America, Europe, Asia, Oceania and North Africa.[4]

To calculate the monthly log-returns $r_{i,t} = log\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$ we use the closing price on the last day of each month $(P_{i,t})$ and compare it with the same price of the previous month $(P_{i,t-1})$. The Table 1 of Section A-1 of the Appendix at the end of this thesis shows the maximum, minimum and average of each index for the two sub-periods analysed, i.e., before and during the *Lehman Brothers* financial crisis.

Although, as is justified at the end of Section 1.2, the normality of the data is not fundamental to guaranteeing the properties of the inference by using the Moran's I statistic, we have estimated skewness and kurtosis in order to have more information about the behaviour of our data and we have made individual normality tests for the log-returns of each country before and after the last financial crisis. We used the *Shapiro-Wilk* normality test and the results are shown in Table 2 of Section A-1 of the Appendix at the end of this thesis. In all cases we conclude that the analysed log-returns are not normally distributed. Using the *Jarque - Bera* statistic for testing the null hypothesis of skewness equal to zero and kurtosis equal to three, which are those of the Normal distribution, the results also reject this null hypothesis.

1.3.2 Distances and similarities between stock indices

In Section 1.2 the spatial dependence is defined as the correlation between stock markets indices taking into account the distances between the financial markets. Therefore, the definition and calculation of these distances is essential for the estimation of the Moran's I statistics for the spatial dependency test. In Figures 1.1, 1.2 and 1.3, we represent in gray scale the three weight matrices W obtained from the three criteria that we analyse in this study: distances in kilometres, time distance in hours and number of hours with common opening. Comparing the three criteria we see how the results in Figure 1.3 show a better separation between groups of stock markets, in contrast to Figures 1.1 and 1.2 where the gray scale tends to be

[3]The results do not change when the beginning of the crisis is established in August 2007, when the governments and central banks of some countries responded to the collapse of the economy with unprecedented fiscal stimuli, the expansion of monetary policy and institutional bailouts.

[4]The data have been extracted from Datastream, Yahoo Finance and Investing.

more progressive. Therefore, this first analysis shows us how the criterion based on the common opening hours of the markets is the one that most clearly defines the neighbourhood groups.

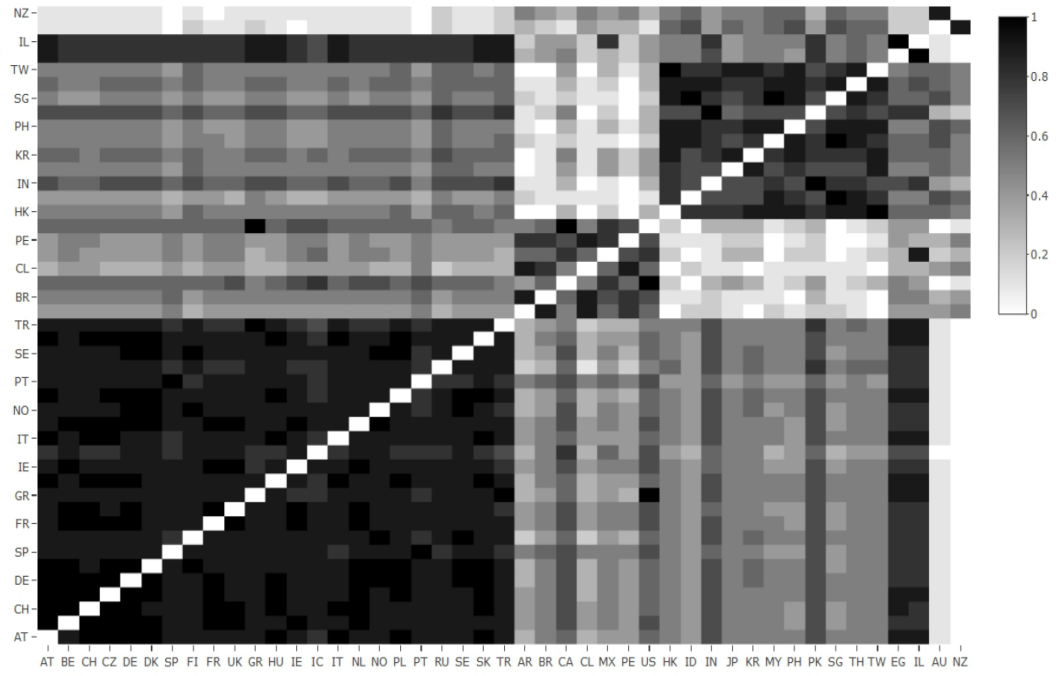


Figure 1.1: Weights between stock markets obtained from distances in kilometres.

We carried out a hierarchical cluster analysis for each of the three weight matrices represented in Figures 1.1, 1.2 and 1.3. As expected, with the criteria based on distances, whether they are hourly or in kilometres, a clear grouping of countries according to their neighbourhood is not obtained. The criterion based on common opening hours between markets provides a clear grouping of countries into three groups of neighbours. The resulting dendrogram is shown in Figure 1.4 (the dendrograms obtained using kilometre and hour distances are shown in Figures 1 and 2 of Section A-2 of the Appendix at the end of this thesis). Table 1.1 shows the neighbourhood groups obtained. The first group contains, almost entirely, countries on the European continent, along with Israel, Egypt and Turkey. In the second group are the countries of the American continent and, finally, in the third group are the Asian countries along with Oceania.

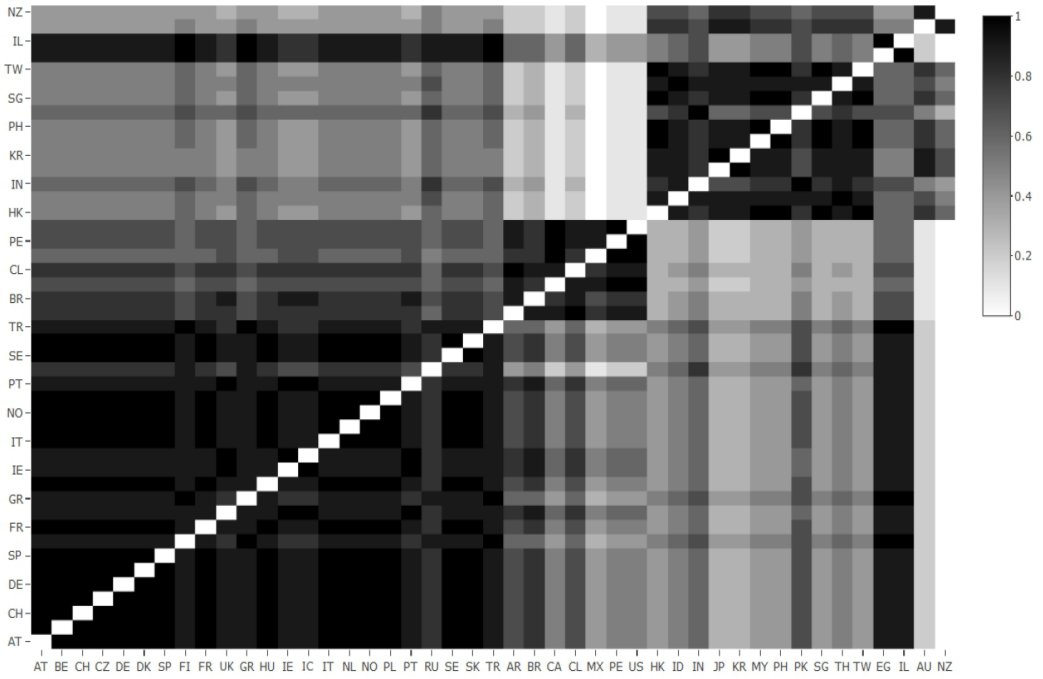


Figure 1.2: Weights between stock markets obtained from distances in hours.

1.3.3 Spatial dependency analysis

We show the results of the spatial dependence analysis between stock markets. For this purpose, we use the weights obtained from the common opening hours between markets criterion. We focus on positive spatial dependence, although the analysis could also be extended to cases where the dependence is negative.

The global Moran's I statistic is calculated for each month, the results of the inference at 5% significance level are shown in Figure 1.5. Analysing the results by month, June tends to accumulate more years with statistically significant spatial dependence, followed by February, July, August and November. If we compare

Table 1.1: Neighbour groups obtained from the dendrogram in Figure 1.4.

GROUP 1	Germany, Austria, Belgium, Denmark, Egypt, Slovakia, Spain, Finland, France, Greece, Holland, Hungary, Ireland, Iceland, Israel, Italy, Norway, Poland, Portugal, United Kingdom, Czech Republic, Russia, Sweden, Switzerland, Turkey
GROUP 2	Argentina, Brazil, Canada, Chile, United States (DJ), United States (SP), United Kingdom, Czech Republic, Russia, Sweden, Switzerland, Turkey Mexico, Peru
GROUP 3	Australia, South Korea, Philippines, Hong Kong, India, Indonesia, Japan, Malaysia, New Zealand, Pakistan, Singapore, Thailand, Taiwan

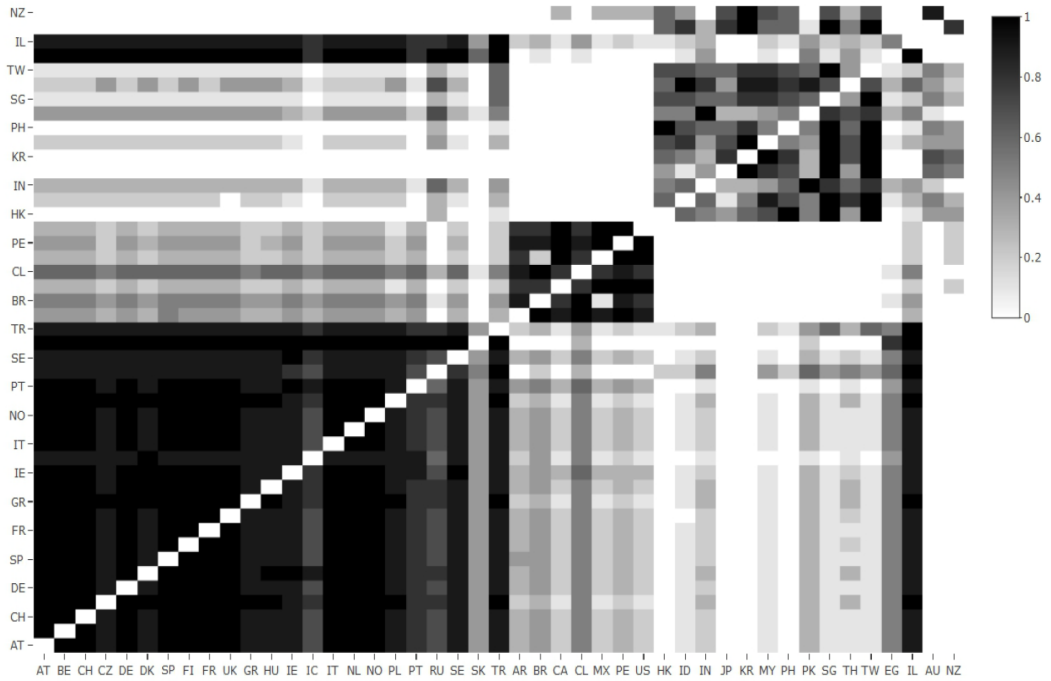


Figure 1.3: Weights between stock markets obtained from the common opening hours between markets.

the results annually, 2010 is the year with the most months with significant global spatial dependence. Also, at first glance, it seems that during the crisis (from 2008) the number of periods with spatial dependence increases significantly.

Analysis of global spatial dependence has given us an initial idea of the situation regarding the joint movements of the markets over time. However, in order to detect where these movements occur, local analysis of spatial dependence is necessary. With this in mind, for each index and each month, the local Moran's I statistic is calculated, which allows us to infer the spatial dependence between each stock index and its closest neighbours. Table 1.2 shows some results obtained from the local analysis. Specifically, we have estimated the linear correlation between each monthly local spatial index and global index, the number of months with positive spatial dependence and relative frequency of these months. Furthermore, in the last column of Table 1.2 we have included the p-value associated with the test of difference between proportions, with the aim of testing if before the crisis (B.C.) and during the crisis (D.C.) the proportion of months with positive spatial dependence is the same. The rows corresponding to the financial markets where it is concluded that there are statistically significant differences between proportions at 5% and 10% significance level are indicated in bold and italics, respectively. This result indicates that the frequency of times in which the Moran's I statistic showed statistically

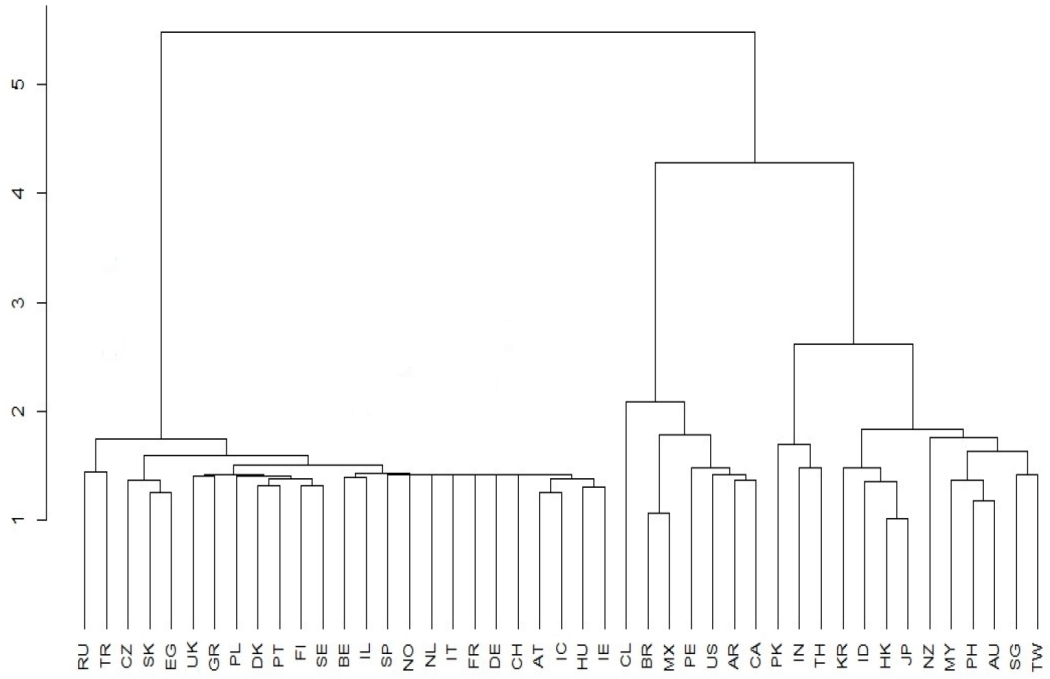


Figure 1.4: Dendogram from the weight matrix obtained from the criterion of common opening hours between markets.

significant positive spatial dependence was low in all the countries analysed.[5]

The first column after the country name in Table 1.2 corresponds to the Pearson correlation coefficient between the local and global Moran's I statistics. The highest correlation occurs between the local index for Italy and the global index, followed by the Philippines and France local indices, in all three cases a value of the Pearson correlation coefficient greater than 0.5 is obtained. The following columns of Table 1.2 summarise the results of local spatial inference. For the two sub-periods analysed - before and during the last financial crisis - the number of months in which the local Moran's I statistic indicates that spatial dependence is statistically significant at 5% before and during the crisis is counted. In general, the number of periods in which the local test concludes that the positive spatial dependence is statistically significant is not very high and, as observed in Figure 1.5, it tends to coincide with the central months of the year (June, July and August).

For a general analysis of the results, the relative frequencies, with respect to the total months in both periods, where the positive spatial dependence is not rejected are shown. Furthermore, the p-value associated with the test of differences between

[5]The same results were also obtained using the two weight matrices obtained from the criteria of distances in kilometres and hours. In general, from the results the same is deduced as described below, but less apparent.

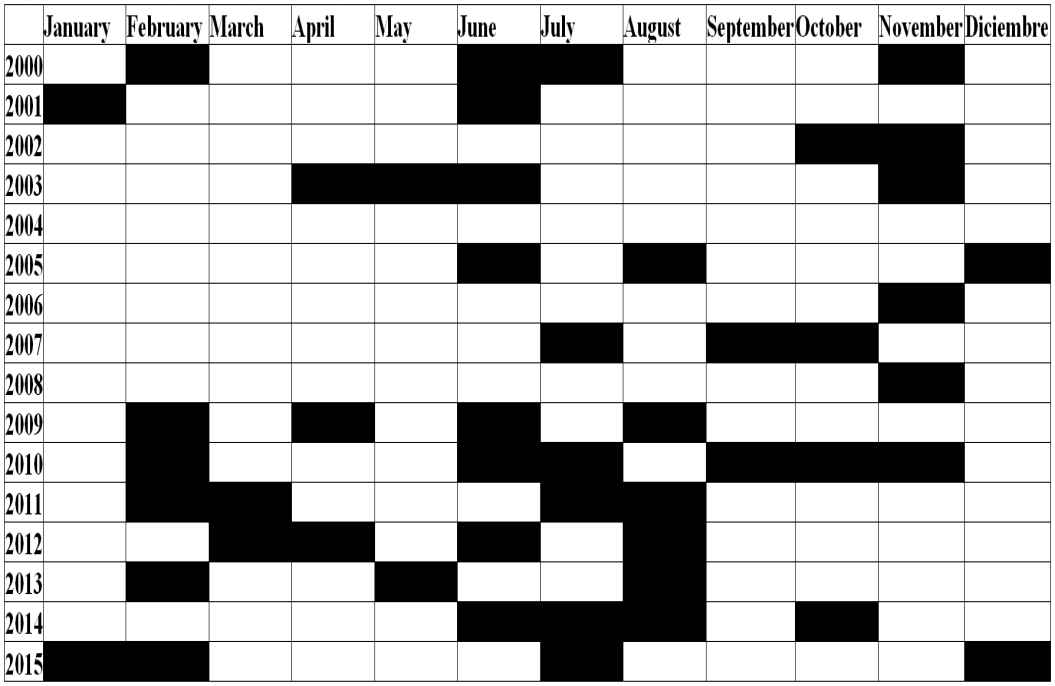


Figure 1.5: Global spatial dependency. The months in which the statistic is significant at 5% are shown in black.

proportions of month before and during the crisis appears in the last column of Table 1.2. In general, from the results of Table 1.2 it is the first group, basically formed by the countries of the European continent, where there are more rows indicating differences between the behaviours of the financial markets before and during the crisis. In this first group, in those cases in which the difference between proportions is statistically significant (13 markets out of the 25), with the exception of Sweden, it can be concluded that spatial dependence is stronger during the crisis. The countries where these frequencies are highest during the crisis are Greece, Italy and Portugal (in that order). However, in the second and third groups, the cases in which there are significant differences between proportions before and after the crisis are less in relative terms and, furthermore, when this occurs, the highest proportion occurs before the crisis.

The possible differences in the temporal behaviour of the spatial dependence found between the different groups of neighbours, together with the evolution of the log-returns, both on average and in variance, should be taken into account in short and medium-term investment decisions.

Table 1.2: Results of the frequency of months with local spatial dependence between stock market log-returns. The rows corresponding to the financial markets. Where it can be concluded that there are statistically significant differences between proportions at 5% and 10% significance level are indicated in bold and italics, respectively.

Countries	Correlation with global Moran's I	Months B.C	Relative frequency B.C.	Months D.C	Relative frequency D.C.	p-value of the test of differences between proportions
			GROUP 1			
Austria	**0.189**	**2**	**0.019**	**9**	**0.103**	**0.006**
Belgium	0.305	6	0.057	6	0.069	0.367
Czech Rep.	0.361	8	0.076	11	0.126	0.123
Denmark	**0.224**	**1**	**0.010**	**6**	**0.069**	**0.014**
Egypt	*0.313*	*5*	*0.048*	*9*	*0.103*	*0.069*
Finland	0.282	6	0.057	5	0.058	0.495
France	*0.523*	*4*	*0.038*	*8*	*0.092*	*0.062*
Germany	0.295	7	0.067	5	0.058	0.397
Greece	**0.362**	**5**	**0.048**	**20**	**0.230**	**0.000**
Hungary	**0.356**	**2**	**0.019**	**11**	**0.126**	**0.002**
Iceland	-0.001	7	0.067	4	0.046	0.270
Ireland	**0.344**	**3**	**0.029**	**10**	**0.115**	**0.009**
Israel	-0.095	3	0.029	1	0.012	0.204
Italy	**0.551**	**4**	**0.038**	**17**	**0.195**	**0.000**
Netherlands	0.359	6	0.057	2	0.023	0.120
Norway	0.097	2	0.019	4	0.046	0.142
Poland	0.188	4	0.038	7	0.081	0.104
Portugal	**0.401**	**6**	**0.057**	**12**	**0.138**	**0.028**
Russia	**0.161**	**4**	**0.038**	**10**	**0.115**	**0.021**
Slovakia	-0.042	6	0.057	8	0.092	0.177
Spain	**0.313**	**3**	**0.029**	**10**	**0.115**	**0.009**
Sweden	**0.040**	**11**	**0.105**	**3**	**0.035**	**0.031**
Swiss	0.121	4	0.038	6	0.069	0.169
Turkey	0.113	8	0.076	6	0.069	0.424
UK	**0.191**	**1**	**0.010**	**5**	**0.058**	**0.029**
			GROUP 2			
Argentina	0.340	7	0.067	9	0.103	0.180
Brazil	0.172	4	0.038	5	0.058	0.263
Canada	0.182	0	0.000	1	0.012	0.135
Chile	**-0.039**	**6**	**0.057**	**0**	**0.000**	**0.012**
Mexico	*0.222*	*1*	*0.010*	*4*	*0.046*	*0.057*
Peru	*0.206*	*3*	*0.029*	*6*	*0.069*	*0.094*
USA (DJ)	0.194	3	0.029	3	0.035	0.408
USA (SP)	0.261	3	0.029	3	0.035	0.408
			GROUP 3			
Australia	**0.375**	**1**	**0.010**	**5**	**0.058**	**0.029**
Hong Kong	0.418	5	0.048	8	0.092	0.111
India	**0.362**	**14**	**0.133**	**5**	**0.058**	**0.040**
Indonesia	0.473	13	0.124	8	0.092	0.241
Japan	0.141	4	0.038	2	0.023	0.275
Malaysia	0.411	7	0.067	5	0.058	0.397
New Zealand	0.198	4	0.038	7	0.081	0.104
Pakistan	0.248	8	0.076	7	0.081	0.456
Philippines	0.549	9	0.086	10	0.115	0.250
Singapore	0.465	5	0.048	7	0.081	0.174
South Korea	0.352	5	0.048	4	0.046	0.479
Taiwan	0.257	8	0.076	4	0.046	0.195
Thailand	0.457	13	0.124	9	0.103	0.329

Note: B.C. (before the crisis) and D.C. (during the crisis)

1.4 Conclusions

In this chapter we have shown the importance of the neighbourhood criterion when analysing the spatial dependence between stock markets. Compared with the criterion based on kilometric distances, widely used in other research areas, and with the hourly distances criterion, the similarity criterion obtained from accounting for common opening hours between markets provides much more visible and obvious neighbourhood relationships. This can be translated into an improvement of the inference of the existence of spatial dependence between the stock returns of neighbouring countries.

The results on spatial dependence show clear differences between the behaviour of European countries and the remaining countries. Specifically, European countries showed a stronger relationship during the crisis than before, which is not seen in the other countries. In other words, in a hypothetical portfolio made up of European market indices the systematic risk may be greater and reduces the risk reduction capacity that comes with diversification.

Chapter 2

Dynamic distances between stock markets: Use of uncertainty indices measures

2.1 Introduction

We propose a monthly Uncertainty Index for stock markets using the Google Trends tool, that is called Google Trend Uncertainty Index (GTUI). We use this index to calculate dynamic distances between stock markets and analyse the spatial dependence with the global Moran's I statistic proposed by Moran (1950). Our aim is to analyse, in different financial crisis periods, if systemic risk is reflected in an increase in the spatial dependence between market losses calculated as the negative logarithm of stock returns. Market losses were calculated based on the loss function defined in McNeil et al. (2015).[6]

Historically, geographical distances have been the most widely used measure for calculating the spatial dependence between regions but, as shown in Chapter 1 (see also Acuña et al., 2018), these distances are not valid for analysing spatial dependence between stock markets. An alternative method as proposed by Flavin et al. (2002) as a proxy for the ease of trading was to use the number of overlapping operating hours of stock markets as a measure of trading synchronisation. These authors found that the more hours of common trading, the greater the degree of equity price co-movement. The Chapter 1 of this thesis showed that overlapping operating hours criterion improves the spatial dependence results obtained using geographical and time distances.

Uncertainty Indexes have been used recently in the literature because they reflect, either through official reports or internet searches, the concerns of economic and financial agents as well as the general public about events that affect the behaviour of a country's economy. Ahir et al. (2019) obtained a quarterly index of

[6] In Chapter 4 we will show an example of analysis using extreme value dependence with copulas that was published in Bolancé and Acuña (2021).

uncertainty, called World Economic Uncertainty Index (WUI), which was based on counting the number of times that the word "uncertainty" and its variants appeared in the Economist Intelligence Unit (EIU, https://www.eiu.com/n/) for 143 countries. These authors concluded that the level of uncertainty is significantly higher in developing countries, is positively associated with economic policy uncertainty and stock market volatility and negatively with GDP growth. Baker et al. (2016) calculated a monthly index of Global Economic Policy Uncertainty (GEPU) that was based on the raw count of terms in three categories (economy, policy and uncertainty) divided by the total number of articles in the newspapers of 16 countries that included these terms, the searches being done in the respective native language. Using a similar process, Ghirelli et al. (2019) obtained their specific Uncertainty Index for Spain. By using Google Trends tool, Weinberg (2020) proposed an Economic Policy Uncertainty Index for the largest economies in the European Union (Germany, France, Italy and Spain). Previously, Castelnuovo and Tran (2017) obtained an economic Google Trend Uncertainty Index for the United States and Australia.

In this chapter, the period of analysis is extended to March 2021. Now, based on the WUI and our proposed GTUI, we first calculate two types of dynamic distances between stock markets and analyse their evolution over the last 17 years; specifically, we study how these distances change in the financial crisis periods as systemic risk proxy. Secondly, we use the Global Moran's I statistic to analyse the changes of spatial dependence between international stock markets that have been induced by the various global financial crises over the last 17 years. We identify the month with significant spatial dependence throughout the analysed period and study if during the financial crisis periods of the Lehman Brothers bankruptcy, the sub-prime mortgage crisis, the European debt crisis, Brexit and the COVID-19 pandemic, spatial dependence was more frequent compared to non-crisis periods.

The remainder of the chapter is organised as follows. Section 2.2 presents our distance measures and spatial dependence test. In Section 2.3 it is described the data and empirical analysis of uncertainty indices and global spatial dependence. Section 2.4 offers the conclusions.

2.2 Dynamic distances and spatial dependence

We propose the calculation of two alternative distances between countries based on two uncertainty indices: the WUI and the GTUI. The former is a quarterly index that is calculated for 143 countries by counting the number of times the words "uncertain", "uncertainty" and "uncertainties" are mentioned in the EIU country reports (see Ahir et al., 2019). The latter is a monthly index that is specifically designed for

our study.

The EIU reports discuss the main political and economic developments in each of the 143 countries analysed, together with an analysis of the political and economic conditions. A team of analysts from each country and a central editorial team of the EIU are employed to do this. For the WUI to be comparable between different countries, the frequency of words by country and quarterly report is divided by the total number of words in this report and multiplied by 1000.

The highest peaks of the WUI coincide with the following events: the second Gulf war (2003), the attacks of September 11 (2011), the European debt crisis (2009), the referendum of the United Kingdom in favour of Brexit and the presidential election in the United States (2016), the border control crisis in Europe and the ARS-CoV-2 outbreak (2019) and with the weather phenomenon El Niño. These peaks of uncertainty tend to be more synchronised between developed economies and between economies whose countries have closer commercial and financial ties. Furthermore, a lower uncertainty has been observed in developed economies.

In our financial context, the WUI has some limitations. Firstly, we only have quarterly information. Secondly, this index only works with the word "uncertainty" and some variants of it. Using the Google Trends tool, we propose an alternative index that can be obtained monthly and, in addition to the word "uncertainty" used in the WUI, it includes other relevant words related to the financial markets trends.

2.2.1 Google Trends Uncertainty Index

The Google Trends is a Google Labs tool that shows the most popular searched terms using the Google search engine. This tool gives information related to the frequency with which a search for a particular term is carried out in various regions of the world and in various languages. The available data range from 2004 to date.

The Google Trends Uncertainty Index (GTUI) is based on the idea that economic agents, represented by internet users, search for information online when they are not sure. This implies that the frequency of searching for terms that may be associated with future and possible bad events is high when the level of uncertainty is high. To obtain the specific index for each country, we select a broad set of keywords that are often cited in the Federal Reserve Beige Book for the U.S. and the Reserve Bank Statement on Monetary Policy. English is chosen as the common language since it is the mostly widely used language in the world. The words of interest that are selected, related to financial markets and the crisis events that have occurred in recent years, are the following: "austerity", "bankruptcy", "dollar", "financial crisis", "recession", "risk", "stock exchange", "share price", "stock market" and "uncertainty". The Google Trends tool enables us to find out the percentage search

frequency of each of these words by country and period. In our study, 45 countries and 46 stock indices (2 for USA) are analysed monthly. These countries and stock indices are listed, among others, in the Table 4 of the Section A-1 of the Appendix at the end of this thesis. Each country is labelled using the two first digits notation. The frequency of words per country and period are added to obtain our proposed monthly GTUI, i.e, the monthly GTUI is the sum of the frequencies of words in each country i at month t.[7]

2.2.2 Global Moran's I statistic

Since this study focuses on spatial dependence measures to analyse how markets are interconnected, the specification of the linkages matrices for each period t between stock markets is important; these matrices are called W_t and in our case are 46×46. For calculating the elements of W_t we use the criteria defined by Asgharian et al. (2013) based on constructing a contiguity matrix C_t between markets. Each element of this matrix (C_{ijt}) indicates how contiguous is each market j to each market i, $\forall i \neq j$, at period t, according to a measure of distance (or similarity) between both countries. We then define the matrix C_t using distance criteria as proposed in this study.

Let D_{ijt} be a measure of distance between countries i and j at period t, the elements C_{ijt} in C_t are given by:

$$C_{ijt} = 1 - \frac{D_{ijt} - \mathrm{m}in_j D_{ijt}}{\mathrm{m}ax_j D_{ijt} - \mathrm{m}in_j D_{ijt}} \qquad \forall \ i \neq j \ t = 1, ..., T.$$

This definition of contiguity ensures that all elements of C_t lie between zero and one, if C_{ijt} is near 1 the longest distance is from country i to country j and C_{ijt} is near 0 if the shortest distance is between country i and country j. Moreover, C_{ijt} is not necessarily symmetric, it can be occur that country j is an important neighbor for country i (i.e., C_{ijt} is close to zero) but country i may be unimportant for country j (i.e., C_{ijt} is close to one). The linkages matrix or spatial weights in the matrices W_t are obtained from C_{ijt} through row standardization such that, for each row i, $\sum_j W_{ijt} = 1$.

Through the GTUI we can obtain a monthly distance matrix between countries. Let $GTUI_{it}$, $i = 1, ..., 46$ be the value of the uncertainty index for the stock market i in the period t, the distances distances between pairs of stock markets are defined as:

$$D_{ijt} = |GTUI_{it} - GTUI_{jt}|, \ \forall i \neq j \ t = 1, ..., T. \tag{2.1}$$

[7] In Chapter 3 the GTUI will be described with more detail.

Alternatively, the WUI can also be used to calculate distances between stock markets. However, this uncertainty index is available quarterly and to calculate monthly spatial weights we assume that the WUI is constant throughout the three months of each quarter.

The Moran's I statistic at period t is defined as Moran (1950):

$$I_t = \frac{N}{S_{0t}} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} W_{ijt} \left(l_{it} - \bar{l}_t\right) \left(l_{jt} - \bar{l}_t\right)}{\sum_{i=1}^{N} \left(l_{it} - \bar{l}_t\right)^2}, \tag{2.2}$$

where $N = 46$, S_{0t} is the sum of all the spatial weights W_{ijt}, $\forall i \neq j$ and l_{it} is the loss of the index i in period t, that are equal to negative log-returns, i.e., $l_{it} = -log\left(\frac{P_{i,t+1}}{P_{i,t}}\right)$, where $P_{i,t}$ is the value of stock index i in period t. In alternative studies the spatial weights matrix is discretised, only those areas whose continuity coefficients are greater than the mean or median of all these coefficients are considered neighbors (see Asgharian et al., 2013). In this paper, taking into account that financial markets have a global behavior and that systemic risk has global effect, we decided to use continuous weights matrix, that give more weight to those more similar markets, according to the distance criterion. In addition to global index defined in expression (2.2), we have obtained the results for analysing local spatial dependence index defined in Chapter 1 for log-returns. In this case, the results that we obtained with continuous matrices of weight were hardly statistically significant.[8]

The asymptotic distribution of I_t is Normal with mean and variance known (Moran, 1950) and we can therefore perform the test of positive global spatial dependence for each period t.

The inference suggested by the global Moran's I statistic is based on the assumption that the data are independent and identically distributed (iid). Regarding non-normality, Griffith (2010) concludes that the assumption of normality is not essential for the asymptotic properties of the Moran's I statistic.

2.3 Empirical Analysis

We analyse the spatial dependence between monthly losses of 46 stock indices of 45 countries, which are listed and described in Table 3 of the Section A-1 of the Appendix at the end of this thesis, and which shows that USA is the only analysed country with two stock indices. The data cover the period from January 2004 to March 2021 and take into account different events that carried a systemic risk:

[8] The discretised matrix will be used in Chapter 3 for global and local positive spatial dependence analysis.

Lehman Brothers bankruptcy in September 2008, the sub-prime mortgage crisis between 2007 and 2009, the European debt crisis (Euro debt) since the end of 2009 until mid 2014, Brexit in June of 2016 and the COVID-19 pandemic since March 2020. In Table 2.1 we show the beginning and end of these crisis periods. Similarly to Chapter 1 but for a longer period, for calculated the monthly losses we use the closing price on the last day of months t and $t-1$.

Table 2.1: Crisis periods.

	Sub-prime	**Euro-debt**	**Brexit**	**COVID-19**
Beginning	31/08/2007	30/06/2010	30/06/2016	29/02/2020
End	30/06/2009	30/06/2014	31/01/2020	31/03/2021

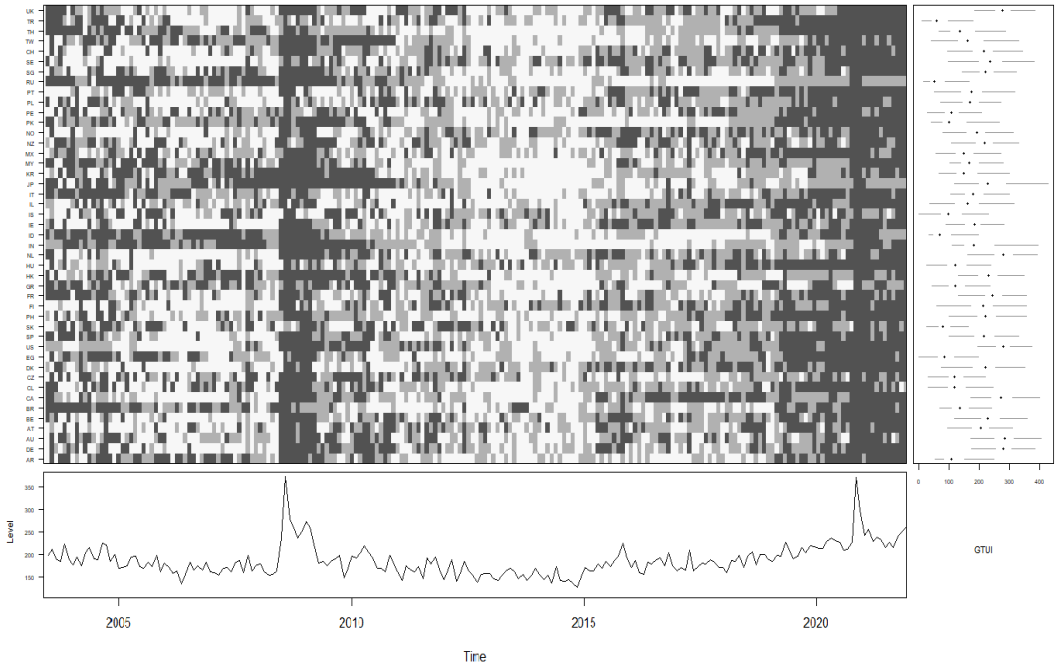


Figure 2.1: Values of the Google Trend Uncertainty Index for each country (at top) - the darker the shading, the higher the value - and plot of the mean index (at bottom). The box-plots of the GTUI for each country are shown on the right.

Before beginning the analysis, we plotted the uncertainty index GTUI in Figure 2.1, the plot of the WUI is shown in Figure 3 of Section A-2 of the Appendix at the end of this thesis. The top part of both figures represents the values of the indices; the darker the shading, the higher the value. In the bottom part of both figures the

mean of the index for all countries is plotted and the box-plots of the Uncertainty Index for each country are shown on the right. For the WUI we only have data up to the last quarter of 2020 and we assume that for each observed quarter the index value is constant overall three months. At first glance, it is evident that both indices behave differently. We compare the GTUI in Figure 2.1 with the filtered series of losses using the ARMA-GARCH models shown in Table 4 of Section A-1 of the Appendix at the end of this thesis. The filtered series are free of the temporal component, i.e. they take independent values, which guarantees the properties of the inference that is presented at the end of this section and compares different periods. For each month, we test the spatial dependence between the random components of the series, this dependence implies a greater systemic risk.

The filtered series are plotted in Figure 2.2 and, comparing with Figure 2.1 and Figure 3 of Section A-2 of the Appendix at the end of this thesis, it can be seen that our GTUI captures market uncertainty better than the WUI. In Figure 2.3 we plot the GTUI for each country with the overall mean of its confidence interval at 95% confidence level, showing a similar behavior in all analyzed countries, especially in the periods of most uncertainty.

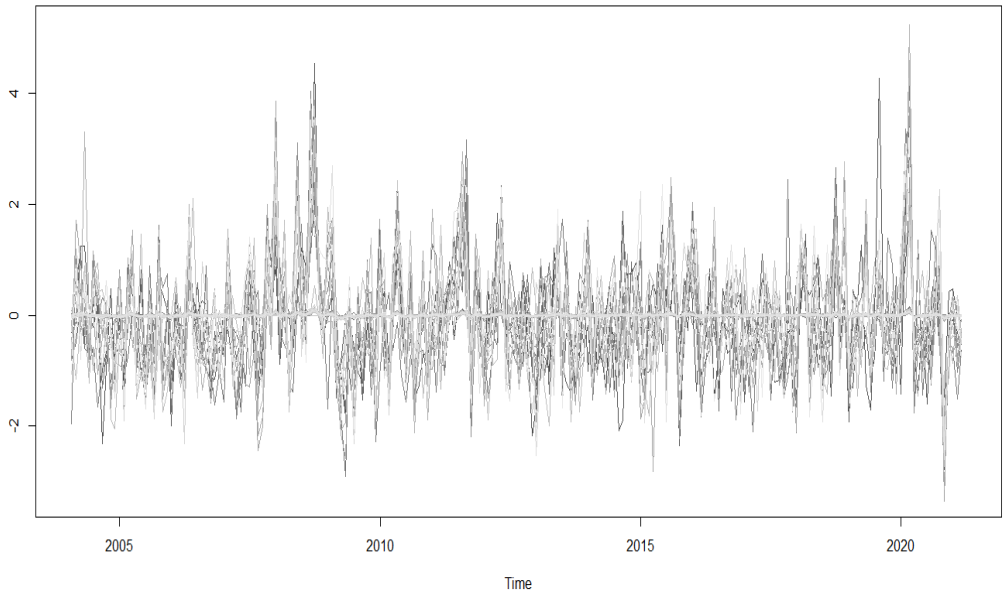


Figure 2.2: Filtered losses for the 46 stock indices.

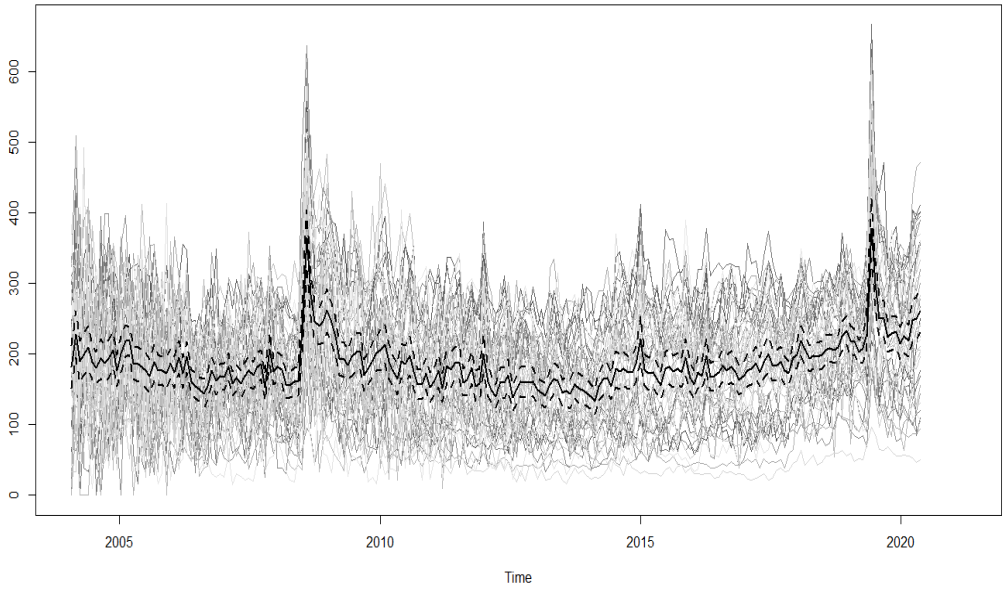


Figure 2.3: GTUI for the 45 countries, mean (the thickest line) and 95% confidence intervals of the mean (dashed lines).

Focusing on the right part of Figure 2.1, it can be seen that the countries that reached highest GTUI values are Japan followed by Australia. The periods showing the highest GTUI values are at the time of the sub-prime mortgage crisis and the COVID-19 pandemic.

With the aim of analyzing how these distances change in the different crisis periods we carry out a hierarchical cluster using complete linkage with the distance matrix obtained at the beginning and end of each crisis period; the dendograms are shown it Section A-2 of the Appendix at the end of this thesis (Figures 4, 5, 6 and 7). We see that, in all cases, using a "Height" between 0.10 and 0.20, three groups can be formed that change at different moments of time. The summary of the groups is shown in Table 2.2. At the beginning of the sub-prime mortgage and Euro debt crises Japan was isolated from the other countries. The same situation occurred for United Kingdom at the beginning of the Brexit period. At the end of the Brexit period and in the last month analyzed during the COVID-19 pandemic the isolated countries were Russia and Netherlands.

Table 2.2: Groups of countries obtained with hierarchical cluster with distances based on GTUI.

	Sub-prime	**Euro debt**	**Brexit**	**COVID-19**
	Beginning of crisis period			
Group 1	HU,EG,RU,CZ SK,PE,TR	IS,BR,EG,HU ID,PK,PE,RU TR	AR,BR,CL,CZ GR,HU,IN,MY MX,PE,KR,TW TH	GR,IS,ID,RU SK,TR
Group 2	AR,AT,BE,BR CA,CL,DK,FI FR,GR,IS,ID IE,IL,IT,JP MY,MX,NL,NZ NO,PK,PL,PT RU,KR,SP,SE TW,TH	AU,AT,CA,DK FR,UK,DE,NL NZ,NO,SG,SP TW,US	EG,ID,PK,RU SK,TR	AU,CA,DK,DE NL,SE,UK,US
Group 3	AU,DE,HK,IN PH,SG,CH,UK US	AR,BE,CH,CL CZ,FI,GR,HK IN,IE,IL,IT MY,MX,PH,PL PT,SK,KR,SE CH,TH	AU,AT,BE,CA DK,FI,FR,DE HK,IS,IE,IL IT,JP,NL,NZ NO,PH,PL,PT SG,SP,SE,CH US	AR,AT,BE,BR CL,CZ,EG,FI FR,HK,HU,IN IE,IL,IT,JP MY,MX,NZ,NO PK,PE,PH,PL PT,SG,KR,SP CH,TW,TH
No Group	JP	JP	UK	
	End of crisis period			
Group 1	AU,IN,JP,NL US	AR,DE,EG,GR IS,MX,NL,PE PK,RU,TH	AU,BE,CA,DK FI,FR,DE,NL PH,SP,SE,CH TH,UK,US	BR,CL,CZ,EG GR,IS,ID,PE SK,TR
Group 2	AR,BE,CA,DK FI,FR,GR,HK NZ,PK,PH,SG KR,SP,SE,CH UK	AT,AU,BE,CA FI,FR,UK,NZ SE,US	AR,AT,BR,CL EG,HK,HU,IN IE,IL,IT,JP MY,MX,NZ,NO PK,PE,PL,PT	AR,SP,HU,IL IT,IE,PL,TH TW,JP,MY,KR MX,PK
Group 3	AT,BR,CL,CZ DK,EG,HU,ID IE,IL,IT,IS MY,MX,NO,PE PL,PT,RU,SK TW,TH,TR	BR,CH,CL,CZ DK,SP,HK,HU IE,IL,IN,IT JP,KR,MY,NO PH,PL,PT,SG SK,TW	CZ,GR,IS,ID SK,TR	AT,AU,BE,CA CH,DE,DK,FI FR,UK,HK,IN NO,NZ,PH,PT SE,SG,US
No Group			RU	NL,RU

Based on GTUI we calculate a dynamic distance between countries using expression (2.1) to estimate spatial dependence. In the Figure 2.4 we show the monthly time series of the Moran's I statistic for spatial dependency, that was defined in expression (2.2). This index detects greater spatial dependence throughout the crisis periods (see Table 2.1), that are also associated with the most similar behavior between markets according to the GTUI.

In Figure 2.5 we summaries the results of the monthly Moran's I test of spatial dependence based on GTUI distances. We compare the crisis periods (in grey and framed with a thick black line) and non-crisis periods (in white). For each crisis and non-crisis period the months with significant spatial dependence at 10% significance level are marked in black. The same figures using WUI distances and overlapping hours criterion are shown in Section A-2 of the Appendix at the end of this thesis (Figures 8 and 9). These figures shows that spatial dependence is more significant using the GTUI distances.

Using GTUI distances, in Table 2.3 we compare the proportions of months with spatial dependence in crisis and non-crisis periods and test if these proportions are larger for a crisis period than for a non-crisis period. The last column in Table 2.3 shows the p-values associated with the statistic for testing equality of the proportions in crisis and non-crisis periods against the alternative - that in a crisis period this proportion is greater than in a non-crisis period. The results shows that, taking into account all crises periods, spatial dependence tends to increase (p-value=0.08). The sub-prime is the crisis period with the most spatial dependence, i.e., in these periods there was more contagion between markets and, therefore, the systemic risk increases.

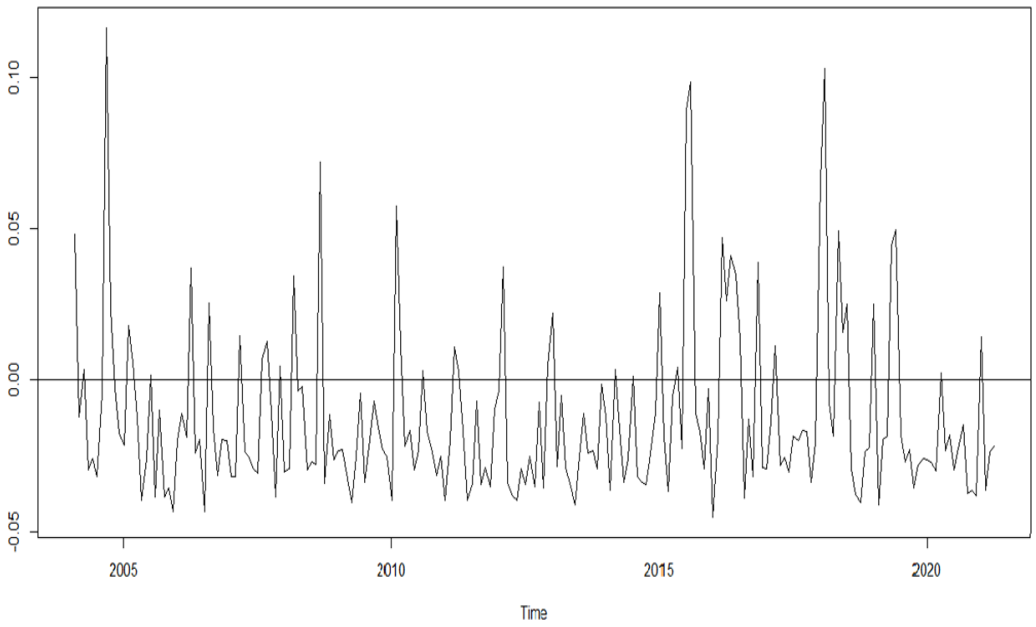


Figure 2.4: Moran's I statistic for spatial dependency.

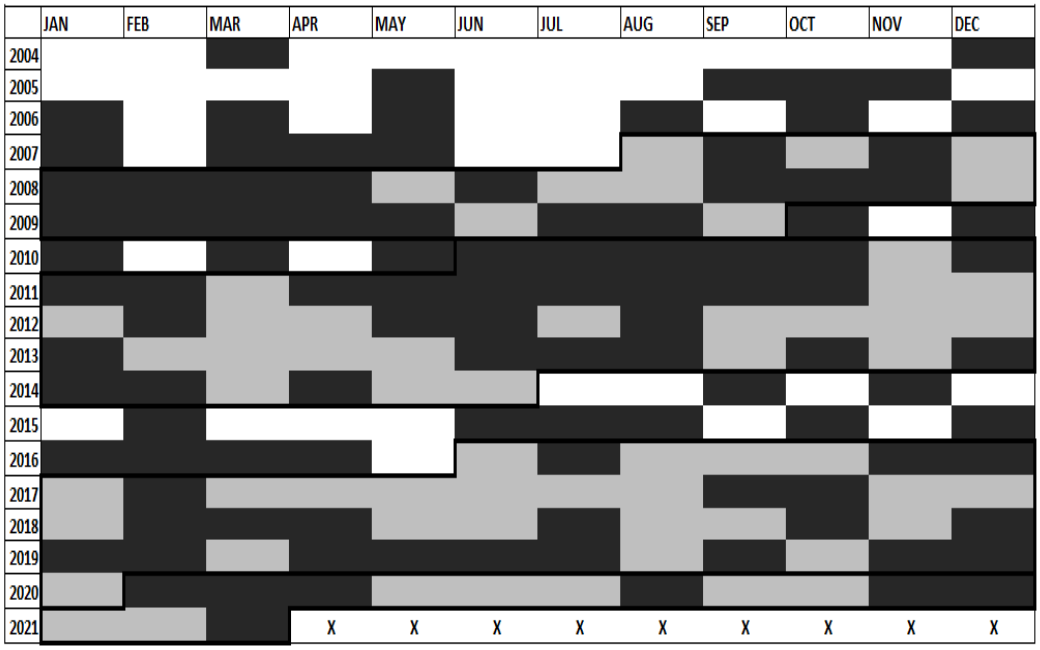


Figure 2.5: Monthly results for Moran's I test of spatial dependence with GTUI distances. The crisis period is in gray, the non-crisis period in white and significant spatial dependence at 10% level in black (the months outside the analyzed period are marked with x).

Table 2.3: Frequencies and proportion of month with spatial dependence and p-values associated with the test of equality of proportion against greater proportion in crisis periods.

	Frequency	**Proportion**	**p-value**
Sub-prime	17	0.65	0.03
Euro debt	28	0.57	0.09
Brexit	21	0.48	0.37
COVID-19	7	0.50	0.35
Total crisis	73	0.55	0.08
Total non-crisis	33	0.45	

2.4 Conclusions

We have shown that our proposed Google Trend Uncertainty Index is a good indicator of financial instability. Periods with higher values of GTUI are associated with financial crisis periods and with higher losses. Dynamic distances based on GTUI allow us to improve the statistical significance of the Moran's I statistic for testing spatial dependence. We observe how events related to systemic risk increase the global spatial dependence between the analysed 46 stock markets, i.e., the dependence is stronger between stock markets with similar GTUI than between markets with different GTUI and, furthermore, this dependence increases in crisis periods when the values of the GTUI are higher. This behaviour cannot be detected with the WUI based distances neither with hours overlapping criterion.

Chapter 3

Non-Normal Market Losses and Spatial Dependence using Uncertainty Indices

3.1 Introduction

Our aim in this chapter is to analyse if systemic risk is reflected in an increase in the spatial dependence between market risks, i.e., whether the detrimental and/or favourable effects of systemic risk are similar between markets with certain neighbourhood characteristics related to similar economic uncertainty. In our analysis, the market risk is approximated by the variance (volatility) and by the metric given by the Value-at-Risk (VaR) associated with the potential losses of stock markets. The losses are calculated from the negative logarithm of stock returns (hereinafter log-returns) (based on the loss function defined in McNeil et al., 2015). The differences between the volatility and the VaR as measures of risk is that the former takes into account both tails of the distributions, i.e., losses at right tail and profits at left tail, while the latter focuses on the right tail of the distribution, i.e., on the losses.

The Global Moran's I statistic Moran (1950) and its local version proposed by Anselin (1995) allow us to carry out inference on global and local spatial dependence, respectively. Both statistics are based on the assumption that the data are normal, independent and identically distributed and they have a known asymptotic normal distribution that is used to test the positive statistical significance of the spatial dependence, understood in our case as having similar behaviour between financial markets. Regarding non-normality, Griffith (2010) concludes that the assumption of normality is not essential for the asymptotic properties of the global Moran's I statistic. However, for the local statistic, the non-normality of the data causes errors in the inference based on the normal distribution. Given that our risk data clearly have of non-normal distribution, they are right skewed and have extreme values, a simulation study is carried out to check how the global and local

inference on spatial dependence is affected. The simulation study provides new results regarding the global Moran's I test based on normal distribution when the data have heavy-tailed distribution, i.e. the study proves as spatial econometrics can be applied to the special nature of financial data.

Various studies have used bootstrap inference for testing global spatial dependence. In relation to the analysis that we present here, we highlight some papers that give robustness to our results. For example, in a regression context Yang (2015) studied the consistency of the inference through a set of LM (Lagrange Multiplier) statistics using the residual-based bootstrap methods for testing spatial dependence, a special case being the global Moran's I statistic used in this paper. This author proves the consistency theoretically as well as through a simulation study where results are obtained for normal, normal mixture and log-normal. Jin and Lee (2015) also analysed the consistency of bootstrap inference of global Moran's I statistic in the spatial econometric model context, and carried out a simulation study for normal and chi-square distributions. Focusing on local spatial dependence, Mei et al. (2020) proposed a bootstrap method to approximate the distribution under the null hypothesis of the local Moran statistic proposed by Anselin (1995). These authors demonstrate that the asymptotic normal approximation sometimes fails.

Another important consideration of this paper is the definition of neighbourhood between financial markets. Geographical distances have been the most widely used measure for calculating the spatial dependence between regions but, as shown by Acuña et al. (2018), this criterion is not valid for determining the neighbours of each stock market. An alternative criterion consists of using the number of overlapping operating hours of stock markets as a measure of trading synchronisation, as proposed by Flavin et al. (2002) as a proxy for the ease of trading. Acuña et al. (2018) showed that overlapping operating hours criterion improves the spatial dependence results obtained using geographical distances. However, the latter is a static criterion and it is conceivable that, when the evolution of financial markets is analysed, the neighbours may change over time depending on the expectations and on the positioning of investors. For this reason, we propose the use of uncertainty indices that are exogenous to the markets themselves to give dynamism to our analysis.

Uncertainty indices have been used recently in the literature because, either through official reports or internet searches, they reflect the concerns of economic and financial agents as well as the general public about events that affect the behaviour of the country's economy. Ahir et al. (2019) obtained a quarterly index of uncertainty, called the World Economic Uncertainty Index (WUI), which was based on counting the number of times that the words "uncertainty" and its variants appeared in the Economist Intelligence Unit (EIU, https://www.eiu.com/n/) for 143 countries. These authors concluded that the level of uncertainty is significantly

higher in developing countries, and it is positively associated with economic policy uncertainty and stock market volatility and negatively with GDP growth. Baker et al. (2016) calculated a monthly index of Global Economic Policy Uncertainty (GEPU) that was based on the raw count of terms in three categories (economy, policy and uncertainty) divided by the total number of articles in the newspapers of 16 countries that included these terms, the searches being done in the respective native language. Using a similar process, Ghirelli et al. (2019) obtained their specific Uncertainty Index for Spain. By using the Google Trends tool, Weinberg (2020) proposed an Economic Policy Uncertainty Index for the largest economies in the European Union (Germany, France, Italy and Spain). Previously, Castelnuovo and Tran (2017) obtained an economic Google Trend Uncertainty Index for the United States and Australia.

In this paper, we use an ad-hoc Google Trend Uncertainty Index (GTUI) to select the monthly neighbours of the 46 stock indices and to carry out the spatial dependence analysis. First, we use the Global Moran's I statistic to analyse the changes of spatial dependence between stock markets. Second, we carry out a local spatial dependence, focusing on the following countries: Spain, Germany, France, Italy, UK, US, Argentina, Brazil, Japan and Hong Kong. In both analyses we identify the months with significant spatial dependence throughout the analysed period and study if during the financial crisis periods of the Lehman Brothers bankruptcy, the US sub-prime mortgage crisis, the European debt crisis, Brexit and the COVID-19 pandemic spatial dependence was more frequent compared to non-crisis periods. The proposed global spatial dependence index is an indicator of market linkages; taking this into account, our analysis answers the following three questions:

1. Are months with risk positive spatial dependence during financial crisis periods more frequent than during the non-financial crisis period?

2. Is positive spatial dependence just as common in all periods of crisis?

3. Are there differences between volatilities with positive spatial dependence and VaRs with positive spatial dependence?

Furthermore, the analysis of local spatial dependence allows one to determine which countries have more weight in contagions during the analysed period and, therefore, which countries are the causes of them.

There are many studies that analyse market linkages, in terms of contagion effects between markets, during crisis periods and using different statistic methodologies. Below, we summarise some examples in relation to the markets that are studied in this paper. Related to the global financial crisis, Dimitriou et al. (2013) used a multivariate time series model for the mean, variance and correlation of log-returns

called "Multivariate AR(1)-FIAPARCH-DCC process" (FIAPARCH-fractionally integrated asymmetric power ARCH and DCC-dynamic conditional correlation) to investigate the contagion effects between the five largest emerging equity markets, Brazil, Russia, India, China and South Africa (BRICS) and the US, through different phases of the crisis, and they conclude that this contagion is increased from early 2009 onwards and is greater during bull periods. In Lien et al. (2018) the indirect effects of volatility between the stock markets of the US and eight East Asian countries were analysed before and during the Asian currency crisis and the sub-prime credit crisis. Among other results, these authors show how the US market is the transmitter and its volatility spills over to other markets during both crisis periods. However, between the East Asian countries, Japan and Hong Kong are markets in which volatility spills over from multiple markets during the sub-prime credit crisis period, but not during the Asian currency crisis. Mohti et al. (2019) used copula models that were fitted to the ARMA-GARCH filtered log-returns to investigate the contagion effects of the sub-prime financial crisis in 18 frontier markets: in relation to countries analysed in this paper these authors found that for the US and Argentina the effects were more pronounced during booms than during busts. Tilfani et al. (2021), using log-returns, analysed the time cross-correlations between the US and eight other stock markets (the rest of the G7 plus China and Russia) before, during and after the financial crisis (2007-2008) and found, among other results, that in the period immediately before the crisis the levels of correlation with the US stock market increased, which could be understood as an overheating of the markets or perhaps an increase in contagion due to systemic risk. After the crisis the results point to a contagion effect. The effect of the European debt crisis on different stock markets around the word was analysed by Samarakoon (2017), and showed that the Asian markets do not present pervasive evidence of contagion from the European debt crisis Samitas and Tsakalos (see 2013, for an analysis on European economy).

In relation to the impact of Brexit, Breinlich et al. (2018) used the abnormal returns to analyse the stock market reaction to the outcome of the 2016 UK referendum on EU membership, and showed that the impact would depend on the nature of post-Brexit UK-EU relations. Alternatively, using the log-returns data, Ameur and Louhichi (2021) analysed the impact of Brexit on the dependency between UK, France and Germany stock markets and found that volatility and the total spillover effects increased in line with Brexit press releases regarding negotiations on the future relationship between the EU and the UK. A similar analysis is presented in Li (2020) in which Italy, Poland and Ireland are also included. Burdekin et al. (2018) presented an extended analysis of the Brexit effect, worldwide, using an econometric model based on log-return and abnormal log-returns to quantify the negative impact of Brexit on different stock markets. Their results showed that the Eurozone

was the hardest hit, and the impact was felt the most by the so-called PIIGS group (Portugal, Ireland, Italy, Greece and Spain) due to their poor fiscal positions. On the contrary, the BRICS nations (Brazil, Russia, India, China and South Africa) fared much better than average, experiencing positive abnormal returns despite negative gross returns. In addition to these studies, it is clear that the subsequent COVID-19 pandemic has affected the development of events and has made it difficult to analyse the final effect of Brexit on financial markets.

More recently, various studies have analysed the impact of the COVID-19 pandemic on financial markets. For example, Chopra and Mehta (2022) used a multivariate DCC-GARCH model of log-returns to compare the presence of contagion for the Asian stock markets during the four main financial crises; the Asian financial crisis, the US sub-prime crisis, the Eurozone debt crisis and the COVID-19 pandemic. The results showed that the US sub-prime crisis was the most contagious for the Asian stock markets, while the impact of the COVID-19 pandemic was the least contagious. Li et al. (2020) investigated whether the uncertainty index proposed by Baker et al. (2020), referred to as IDEMV (Infectious Disease Equity Market Volatility), had additional predictive power for stock market volatility in France, Germany and the UK during the COVID-19 pandemic. The results showed that the IDEMV had stronger predictive power for French and UK stock market volatility during the COVID-19 pandemic; however, the VIX (Volatility Index) had superior predictive power for the three European stock markets. Using a VARMA(1,1)-DCC-GARCH model for the log-return, Akhtaruzzaman et al. (2021) analysed the financial contagion between China and the G7 countries. The results showed that China and Japan appeared to be transmitters during the COVID-19 pandemic.

In general, the studies that have analysed the contagion between financial markets used alternative multivariate time series models with log-returns series. In this study, we present an alternative analysis focused on the risk and use spatial dependence statistics to analyse linkages between the financial markets volatility and its VaR metric, considering distances between the markets uncertainty levels measured by the GTUI instead of geographical distances.

The remainder of the paper is organised as follows. Section 3.2 presents the procedure to test global and spatial dependence using asymptotic normal distribution and bootstrap method. In Section 3.3 the results of a simulation study are presented that compare the asymptotic inference with bootstrap inference carried out with the global Moran's I statistic and local Moran statistic and assuming different distributions. In Section 3.4 we describe the data and the spatial dependence results. In Section 3.5 we summarise the main conclusions.

3.2 The Spatial Dependence Model and Statistics for Testing

Let $\mathbf{Y}_t = (Y_{1t}, ..., Y_{nt})'$ be a vector with data of n countries at period $t = 1, ..., T$. It is assumed that $\mathbf{Y}_1, ..., \mathbf{Y}_T$ are independent and identically distributed (iid). The spatial autoregressive (SAR) model at period t is defined as:

$$\mathbf{Y}_t = \mu_t + \rho_t W_t \mathbf{Y}_t + \epsilon_t, \tag{3.1}$$

where μ_t is a vector with deterministic means that can be estimated separately, ρ_t is the spatial autocorrelation at period t and ϵ_t is a vector with n Normal iid errors with mean 0 and variance $\sigma^2_{\epsilon_t}$. The weights matrix at period t, W_t, is $n \times n$ and identifies the neighbours for each $i = 1, ..., n$. In financial analysis, a fundamental component of the model defined in (3.1) is the weights matrix W_t, given that the geographical distance criterion does not work. So, a specific dynamic criteria based on internet searches from economic agents is defined.

3.2.1 Google Trends Uncertainty Index

Google Trends is a Google Lab that uses the Google search engine to find information related to the frequency with which a search for a particular term is carried out in various regions of the world and in various languages. The available monthly data range is from 2004 to present.

The GTUI is based on the idea that economic agents, represented by internet users, search for information online when they are not sure. This implies that the frequency of searching for terms that may be associated with future and possible bad events is high when the level of uncertainty is high. To obtain the specific index for each country, we select a broad set of keywords that are often cited in the Federal Reserve Beige Book for the US and the Reserve Bank Statement on Monetary Policy. English is chosen as the common language since it is the mostly widely used in the world. A total of 10 economic terms of interest are selected: "austerity", "bankruptcy", "dollar", "financial crisis", "recession", "risk", "stock exchange", "share price", "stock market" and "uncertainty". These terms are related to financial markets and the crisis events that have occurred in recent years and the Google Trends tool enables us to find the frequency of searches using each of them by country and month.

We selected the 10 economic terms mentioned above based on the dictionary proposed by Castelnuovo and Tran (2017). Others studies that used similar indices are Weinberg (2020) and Baker et al. (2016). The selected words are the most com-

mon among the dictionaries proposed for each measure of uncertainty mentioned by the previous authors, since these are the words that best reflect the interest of the internet users in response to an important economic world event.

Let f_{hit} be the frequency of terms h in country i at period t, so the uncertainty index is defined as:

$$GTUI_{it} = \sum_{h=1}^{10} f_{hit}. \tag{3.2}$$

It must be remembered that the vocabulary determines the construction of the uncertainty index.

3.2.2 Weights Matrix Definition

To calculate the elements of W_t we use the criteria defined by Asgharian et al. (2013) based on constructing a contiguity matrix C_t between markets. This matrix indicates how contiguous market i to market j is at period t, according to a measure of distance (or similarity) between both countries. We then define the matrix C_t using the following distance criterion between uncertainty index values:

$$D_{ijt} = |GTUI_{it} - GTUI_{jt}|, \ \forall i \neq j \ t = 1, ..., T. \tag{3.3}$$

Let C_{ijt} be the element of C_t that is the contiguity measure between country i (row) and country j (column), which is given by:

$$C_{ijt} = 1 - \frac{D_{ijt} - \min_j D_{ijt}}{\max_j D_{ijt} - \min_j D_{ijt}} \quad \forall \ i \neq j \ t = 1, ..., T.$$

This definition of contiguity ensures that all elements of C_t lie between 0 and 1; if C_{ijt} is near 0 the longest distance is from country i to country j and if C_{ijt} is near 1 the shortest distance is between country i and country j. Moreover, C_{ijt} is not necessarily symmetric (i.e. $C_{ijt} \neq C_{jit}$); it could be that country j is an important neighbour for country i (i.e., C_{ijt} is close to 1) but country i may be unimportant for country j (i.e., C_{ijt} is close to 0). The linkages matrix or spatial weights in the matrices W_t are obtained from C_{ijt} through row standardisation. By construction, there are zeros on the diagonal of W_t; a market can not be a neighbour to itself. Then:

$$W_{ijt} = \frac{C_{ijt}}{\sum_{j=1}^{n} C_{ijt}}, \ \forall i \neq j \ t = 1, ..., T, \tag{3.4}$$

such that, for each row i, $\sum_{j=1}^{n} W_{ijt} = 1$.

In order to determine with more precision which markets are neighbours, the continuity matrix is discretised, i.e. values 0 or 1 are assigned based on whether

two elements are considered neighbours or not. The criterion can be based on a value c of the continuity matrix C, for example, c can be equal to the median or to a quantile. Let C_t^* be the discretised continuity matrix, so $C_{ijt}^* = 1$ if $C_{ijt} \geq c$ and $C_{ijt}^* = 0$ to the contrary. In practice, the weight matrix is obtained from the row standardisation of C_t^*.

Taking into account that the sum of the rows of C_t^* is the number of neighbours n_{it} for each market, each row in the weight matrix has element equal to $\frac{1}{n_{it}}$ or 0. Considering the SAR model defined in (3.1), the coefficients associated with the neighbors of the market i are $\rho_t \frac{1}{n_{it}}$. Therefore, the larger the number of neighbors is, the weaker the spatial dependency relation between markets.

3.2.3 Global Moran's I Statistic

The Moran's I statistic at period t is defined as Moran (see 1950):

$$I_t = \frac{n}{S_{0t}} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ijt} \left(Y_{it} - \bar{Y}_t\right) \left(Y_{jt} - \bar{Y}_t\right)}{\sum_{i=1}^{n} \left(Y_{it} - \bar{Y}_t\right)^2}, \tag{3.5}$$

where $\bar{Y}_t$ is the sample mean and $S_{0t} = \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ijt}$. Note that for standardised row weight matrix $S_{0t} = n$. Hereafter, the sub-index t is eliminated to simplify notation. Using matrix notation:

$$I = \frac{n}{S_{0t}} \frac{\tilde{\mathbf{Y}}' W \tilde{\mathbf{Y}}}{\tilde{\mathbf{Y}}' \tilde{\mathbf{Y}}}, \tag{3.6}$$

where $\tilde{\mathbf{Y}} = \mathbf{Y} - \bar{Y}\mathbf{1}_n$ is a column vector with the n centred data, where $\mathbf{1}_n$ is a column vector with n ones. The asymptotic distribution of Moran's I statistic is Normal. Under the no spatial autocorrelation null hypothesis the expectation is:

$$E(I) = -\frac{1}{n-1}. \tag{3.7}$$

The variance can be calculated under normality assumption and under unknown distribution. For the former the result is:

$$V_N(I) = \frac{1}{(n-1)(n+1)S_0^2} \left(n^2 S_1 - n S_2 + 3S_0^2\right) - \left(E(I)\right)^2 \tag{3.8}$$

and for the latter it is:

$$V(I) = \frac{n\left[\left(n^2 - 3n + 3\right) S_1 - n S_2 + 3S_0^2\right] - k\left[\left(n^2 - n\right) S_1 - 2n S_2 + 6S_0^2\right]}{(n-1)^{(3)} S_0^2} - \left(E(I)\right)^2, \tag{3.9}$$

where k is the sample kurtosis coefficient, $S_1 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(W_{ij}+W_{ji})^2$,

$$S_2 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{j=1}^{n}W_{ij}+\sum_{j=1}^{n}W_{ji}\right)^2$$

and $(A)^{(b)} = A(A-1)...(A-b+1)$.

The inference suggested by the global Moran's I statistic is based on the assumption that the data are independent and identically distributed (iid).

Inference on positive global spatial dependence can be based on the statistic $Z_I = \frac{I-E(I)}{V(I)} \sim Normal(0,1)$, for a given significance level p, which is a value near 0, the null hypothesis of no spatial dependence is rejected if $P(Z > Z_I) \leq p$, where Z is a standard normal random variable.

Regarding non-normality, Griffith (2010) concluded that this is not essential for the asymptotic properties of the Moran's I statistic. So, the inference based on the asymptotic normal distribution of the global Moran's I statistic works. We analyse to what extent this property is true by comparing normal based inference with the non-parametric inference based on bootstrap samples.

Let $\tilde{\mathbf{Y}}^*_{(1)},...,\tilde{\mathbf{Y}}^*_{(n_B)}$ be a set of n_B bootstrap random samples of size n that are selected with replacement. For each bootstrap sample the Moran's I statistic is calculated as:

$$I^*_{(l)} = \frac{n}{S_0}\frac{\tilde{\mathbf{Y}}^{*\prime}_{(l)}W\tilde{\mathbf{Y}}^*_{(l)}}{\tilde{\mathbf{Y}}^{*\prime}_{(l)}\tilde{\mathbf{Y}}^*_{(l)}},\ l = 1,...,n_B. \tag{3.10}$$

The inference on positive global spatial dependence can be based on the empirical distribution of bootstrap samples; the null hypothesis of no spatial dependence is rejected if $\frac{\sum_{l=1}^{n_B} i\left(I^*_{(l)} > I\right)}{n_B} \leq p$, where $i(a) = 1$ if condition a between parentheses is true. Jin and Lee (2015) proved the consistency of bootstrap inference based on Moran's I statistic.

3.2.4 Testing Local Spatial Dependence

The global spatial dependence analysis indicates whether there are linkages between all markets, or not, but it does not allow us to identify which markets are linked or which have spatial dependence with their neighbours and, therefore, links in terms of similar risk. With the aim of analysing the local spatial dependence we use the local Moran test proposed by Anselin (1995), defined as:

$$I_i = \frac{n\left(Y_i-\bar{Y}\right)}{\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2}\sum_{j=1}^{n}W_{ij}\left(Y_j-\bar{Y}\right). \tag{3.11}$$

Note that $I = \sum_{i=1}^{n} I_i$. The expectation of local Moran statistic is:

$$E(I_i) = -\frac{W_i}{n-1}, \tag{3.12}$$

where $W_i = \sum_{j=1}^{n} W_{ij}$. The variance is:

$$V(I_i) = \frac{W_i^{(2)}(n-k)}{n-1} + \frac{2W_i^{(hm)}(2k-n)}{(n-1)(n-2)} - (E(I_i))^2, \tag{3.13}$$

where $W_i^{(2)} = \sum_{i \neq j} W_{ij}^2$ and $2W_i^{(hm)} = \sum_{h \neq i} \sum_{m \neq i} W_{ih} W_{im}$. Although, asymptotically the distribution of I_i is normal, in practice the exact distribution of this statistic is unknown and normal based inference does not work. Furthermore, given that local inference implies carrying out multiple tests using the same sample we will need to modify the significance level, for example, using Bonferroni correction. For a given significance level p, if the number of multiple tests are r the true significance level will be p/r. Similarly to the global Moran's I test, we compare normal inference with bootstrap inference using the n_B bootstrap samples defined above. Mei et al. (2020) proved the consistency of bootstrap inference for local Moran statistic.

3.3 Simulation Study

In this section we analyse the results of the inference based on the global Moran's I and the local Moran statistics in finite sample. We simulate the values C_{ij} of the continuity matrix to obtain the weight matrix W, and we also simulate the values of the random variable Y in the SAR model defined in (3.1). We obtain $1,000$ samples of sizes $n = 50$ and $n = 200$, respectively. To simulate the values in W, we analyse the behaviour of the C_{ijt} used in the application presented in this paper and we see that these values have a behaviour similar to a random variable with distribution $Beta(4.7, 3)$. To simulate the values of the random variable Y we use the following results from the SAR model:

$$\mathbf{Y} = (I_n - \rho W)^{-1} (\boldsymbol{\mu} + \boldsymbol{\epsilon}), \tag{3.14}$$

where I_n is the identity matrix of order n. To generate data from (3.14) we assume $\boldsymbol{\mu} = \mathbf{0}$ and the value of $\boldsymbol{\epsilon}$ are generated following alternative distributions that have different shapes and tail behaviour. These distributions are: normal with parameters $\mu = 0$ and $\sigma = 0.25$; a Student's t with 3 degree of freedom, $\mu = 0$ and $\sigma = \frac{0.25}{\sqrt{3}}$; a log-normal with $\mu = 0$ and variance 0.5; and a log-logistic with $\mu = 0$

and variance 0.5. Note that these distributions have alternative shapes that can be found when risk variables are analysed.The normal and the Student's t are symmetric, the second having heavier tails than the first. The log-normal and log-logistic are right skewed and the second has a heavier right tail than the first. Furthermore, log-logistic distribution is asymptotically Pareto-type right-tailed distribution. For each distribution the values $\rho = 0.9, 0.5, 0.1$ are used.

In addition to the alternative distributions, the simulation study also analyses the effect on the test results depending on the number of neighbours. With this aim the continuity matrix is discretised using different criteria for obtaining c (remember $C^*_{ij} = 1$ if $C_{ij} > c$ and $C^*_{ij} = 0$ on the contrary): the median (second quartile) of the values C_{ij}, their quantile at 75% confidence level (third quartile) and their quantile at 90% confidence level. Note that the higher the quantile the smaller the number of neighbours. So, spatial dependence is stronger as ρ increases and the number of neighbours decreases.

In Table 3.1 the results of inference at 5% significance level using the global Moran's I statistic are shown. On the $1,000$ replicates of each sample, we calculate the percentage of rejection of the null hypothesis of spatial independence from the alternative hypothesis of positive spatial dependence. For every replicate, the test is carried out using asymptotic inference based on normal distribution (N) and with the finite inference based on $1,000$ bootstrap random samples (B) with replacement and the same size of the original samples. For normal distributions, Student's t and log-normal, the results with N and B are similar in all cases. When the spatial dependence is clear, i.e. $\rho = 0.9$ and the neighbourhood criterion is based on the quantile at 90% confidence level the percentage of rejections is practically equal to 1 in all cases. When the number of neighbours increases this percentage decreases and similarly when the value ρ decreases. For the log-logistic distribution, again, the results for N and B are very similar when the neighbourhood criterion is based on the quantile at 90% confidence level and $\rho = 0.9$ or $\rho = 0.5$. However, when the number of neighbours is at its highest and $\rho < 0.9$ the percentage of rejection with N is greater than that obtained with B. Compared with the alternative distributions, for the log-logistic, that is Pareto tailed, the asymptotic inference based on Normal distribution will have larger type I error, i.e. the null hypothesis could be rejected with more probability when this is true.

Tables 3.2, 3.3 and 3.4 show the results for local spatial inference using the local Moran statistic. The results for three observations with different number of neighbours (maximum, median and minimum) are, respectively, analysed. In general, as expected, the results of local inference indicate that the percentage of rejections of the null hypothesis of spatial independence is much lower than in the global test, this finding having already been presented by Anselin (1995). However, in our simulation study we find some novel results, described below.

In the same way as for the global Moran's I, with the local Moran statistic we also observe that as the number of neighbours increases the null hypothesis is not rejected with more frequency. Furthermore, as Table 3.2 shows, concerning the results for the case with the maximum number of neighbours with each criterion, by using asymptotic inference (N) we obtain a higher percentage of rejection than with bootstrap (B). In contrast, in Tables 3.3 and 3.4, where the cases with median and minimum neighbours are analysed, respectively, the percentage of rejection tends to be the highest with B, i.e. bootstrap inference has clearly more power than asymptotic inference based on normal distribution. In other words, fewer errors are made when rejecting the null hypothesis of independence.

Analysing the results of local inference for alternative distributions, we observe that the differences between N and B are the lowest for normal and Student's t distributions. For log-normal and log-logistic distributions the asymptotic inference barely detects spatial dependence in those cases where it could be stronger, i.e. minimum number of neighbours with neighbourhood criteria based on the quantile at 90% confidence level and with $\rho = 0.9$. In these cases the bootstrap inference considerably improves normal based inference.

Table 3.1: Simulation results for global Moran's I statistic using asymptotic inference (N) and bootstrap inference (B), both at 5% significance level.

$\alpha = 0.1$			Normal		Student's t		Log-normal		Log-logistic	
n	Quantile	ρ	N	B	N	B	N	B	N	B
		0.9	0.461	0.456	0.543	0.527	0.536	0.520	0.728	0.663
	50	0.5	0.268	0.275	0.306	0.290	0.298	0.275	0.495	0.299
		0.1	0.126	0.121	0.133	0.127	0.116	0.108	0.176	0.094
		0.9	0.806	0.808	0.879	0.874	0.877	0.868	0.956	0.952
50	75	0.5	0.456	0.457	0.521	0.507	0.504	0.474	0.737	0.635
		0.1	0.133	0.131	0.155	0.149	0.151	0.140	0.228	0.108
		0.9	0.999	0.999	0.996	0.996	0.997	0.997	1.000	0.999
	90	0.5	0.815	0.813	0.850	0.846	0.858	0.844	0.957	0.936
		0.1	0.198	0.206	0.193	0.186	0.196	0.183	0.293	0.156
		0.9	0.498	0.503	0.498	0.507	0.536	0.528	0.710	0.665
	50	0.5	0.285	0.286	0.298	0.297	0.300	0.297	0.501	0.347
		0.1	0.134	0.133	0.131	0.126	0.145	0.140	0.203	0.100
		0.9	0.838	0.840	0.833	0.830	0.858	0.857	0.951	0.937
200	75	0.5	0.495	0.502	0.496	0.483	0.504	0.488	0.684	0.596
		0.1	0.141	0.143	0.160	0.154	0.162	0.157	0.222	0.104
		0.9	0.993	0.993	0.998	0.998	0.998	0.998	1.000	1.000
	90	0.5	0.842	0.844	0.818	0.814	0.829	0.820	0.961	0.952
		0.1	0.191	0.192	0.211	0.203	0.209	0.200	0.294	0.149
$\alpha = 0.05$			Normal		Student's t		Log-normal		Log-logistic	
n	Quantile	ρ	N	B	N	B	N	B	N	B
		0.9	0.341	0.341	0.402	0.384	0.403	0.368	0.640	0.482
	50	0.5	0.168	0.172	0.189	0.183	0.187	0.163	0.381	0.174
		0.1	0.069	0.069	0.067	0.067	0.061	0.058	0.109	0.048
		0.9	0.705	0.704	0.776	0.765	0.790	0.759	0.923	0.900
50	75	0.5	0.318	0.316	0.361	0.344	0.383	0.319	0.611	0.354
		0.1	0.075	0.077	0.087	0.078	0.093	0.074	0.161	0.050
		0.9	0.993	0.991	0.992	0.992	0.994	0.993	0.999	0.999
	90	0.5	0.711	0.708	0.760	0.740	0.758	0.705	0.887	0.787
		0.1	0.109	0.108	0.118	0.116	0.116	0.095	0.222	0.073
		0.9	0.353	0.355	0.380	0.375	0.410	0.394	0.610	0.525
	50	0.5	0.182	0.192	0.178	0.178	0.209	0.201	0.393	0.202
		0.1	0.068	0.068	0.072	0.069	0.067	0.060	0.108	0.036
		0.9	0.726	0.724	0.708	0.710	0.752	0.748	0.901	0.867
200	75	0.5	0.350	0.342	0.350	0.344	0.370	0.349	0.569	0.336
		0.1	0.086	0.087	0.084	0.085	0.096	0.088	0.158	0.052
		0.9	0.987	0.986	0.993	0.993	0.995	0.994	1.000	1.000
	90	0.5	0.987	0.986	0.709	0.700	0.750	0.728	0.892	0.817
		0.1	0.100	0.099	0.127	0.121	0.121	0.112	0.215	0.063

Table 3.2: Simulation results for local Moran's I statistic with maximum number of neighbours using asymptotic inference (N) and bootstrap inference (B), both at 10% significance level.

$\alpha = 0.1$			Normal		Student's t		Log-normal		Log-logistic	
n	Quantile	ρ	N	B	N	B	N	B	N	B
		0.9	0.144	0.143	0.152	0.153	0.141	0.153	0.204	0.148
	50	0.5	0.119	0.119	0.124	0.124	0.120	0.123	0.184	0.129
		0.1	0.099	0.099	0.105	0.097	0.098	0.101	0.161	0.098
		0.9	0.181	0.186	0.217	0.204	0.190	0.193	0.229	0.175
50	75	0.5	0.132	0.140	0.157	0.156	0.137	0.139	0.141	0.139
		0.1	0.091	0.100	0.127	0.107	0.092	0.087	0.083	0.080
		0.9	0.304	0.289	0.318	0.275	0.309	0.282	0.272	0.155
	90	0.5	0.203	0.197	0.200	0.203	0.192	0.195	0.181	0.173
		0.1	0.132	0.122	0.112	0.109	0.085	0.105	0.063	0.080
		0.9	0.114	0.108	0.110	0.120	0.110	0.092	0.126	0.141
	50	0.5	0.103	0.101	0.102	0.112	0.097	0.079	0.116	0.113
		0.1	0.097	0.088	0.094	0.096	0.089	0.073	0.119	0.099
		0.9	0.134	0.131	0.132	0.137	0.122	0.136	0.163	0.137
200	75	0.5	0.114	0.108	0.113	0.117	0.100	0.115	0.102	0.116
		0.1	0.096	0.087	0.099	0.094	0.085	0.086	0.081	0.081
		0.9	0.186	0.190	0.171	0.187	0.175	0.183	0.185	0.193
	90	0.5	0.137	0.154	0.131	0.132	0.132	0.140	0.127	0.130
		0.1	0.102	0.118	0.099	0.096	0.098	0.101	0.053	0.060
$\alpha = 0.05$			Normal		Student's t		Log-normal		Log-logistic	
n	Quantile	ρ	N	B	N	B	N	B	N	B
		0.9	0.071	0.076	0.077	0.084	0.079	0.089	0.103	0.078
	50	0.5	0.051	0.064	0.064	0.065	0.061	0.065	0.087	0.065
		0.1	0.040	0.048	0.054	0.045	0.052	0.051	0.071	0.041
		0.9	0.108	0.111	0.125	0.116	0.101	0.109	0.106	0.102
50	75	0.5	0.059	0.078	0.089	0.078	0.064	0.065	0.062	0.070
		0.1	0.040	0.049	0.060	0.057	0.040	0.040	0.033	0.036
		0.9	0.205	0.195	0.210	0.183	0.208	0.178	0.191	0.155
	90	0.5	0.125	0.127	0.110	0.112	0.103	0.107	0.120	0.077
		0.1	0.074	0.067	0.050	0.064	0.029	0.046	0.032	0.033
		0.9	0.054	0.053	0.048	0.071	0.054	0.050	0.054	0.071
	50	0.5	0.049	0.049	0.044	0.065	0.046	0.045	0.048	0.056
		0.1	0.049	0.042	0.043	0.058	0.041	0.039	0.049	0.043
		0.9	0.066	0.075	0.066	0.075	0.072	0.081	0.062	0.069
200	75	0.5	0.054	0.060	0.060	0.061	0.053	0.067	0.039	0.055
		0.1	0.048	0.045	0.054	0.050	0.045	0.049	0.030	0.039
		0.9	0.096	0.121	0.095	0.098	0.102	0.101	0.126	0.107
	90	0.5	0.071	0.093	0.067	0.068	0.066	0.075	0.087	0.054
		0.1	0.057	0.062	0.040	0.047	0.042	0.045	0.022	0.030

Table 3.3: Simulation results for local Moran's I statistic with medium number of neighbours using asymptotic inference (N) and bootstrap inference (B), both at 10% significance level.

$\alpha = 0.1$			Normal		Student's t		Log-normal		Log-logistic	
n	Quantile	ρ	N	B	N	B	N	B	N	B
		0.9	0.148	0.170	0.166	0.161	0.153	0.161	0.169	0.160
	50	0.5	0.123	0.129	0.139	0.130	0.123	0.122	0.133	0.126
		0.1	0.103	0.098	0.114	0.106	0.100	0.090	0.099	0.084
		0.9	0.177	0.186	0.161	0.188	0.182	0.207	0.212	0.174
50	75	0.5	0.139	0.144	0.144	0.137	0.129	0.151	0.130	0.117
		0.1	0.110	0.111	0.117	0.094	0.102	0.103	0.057	0.073
		0.9	0.277	0.286	0.254	0.290	0.267	0.293	0.206	0.275
	90	0.5	0.203	0.201	0.164	0.199	0.162	0.194	0.143	0.190
		0.1	0.128	0.132	0.099	0.101	0.083	0.121	0.051	0.069
		0.9	0.118	0.128	0.105	0.112	0.128	0.139	0.112	0.129
	50	0.5	0.114	0.117	0.096	0.106	0.118	0.118	0.104	0.108
		0.1	0.105	0.105	0.092	0.092	0.113	0.105	0.104	0.078
		0.9	0.119	0.121	0.135	0.155	0.127	0.150	0.149	0.120
200	75	0.5	0.108	0.109	0.110	0.125	0.107	0.129	0.094	0.087
		0.1	0.095	0.082	0.088	0.102	0.093	0.093	0.049	0.048
		0.9	0.171	0.190	0.154	0.171	0.161	0.180	0.131	0.187
	90	0.5	0.141	0.147	0.136	0.137	0.118	0.148	0.095	0.126
		0.1	0.112	0.119	0.099	0.107	0.084	0.103	0.027	0.048

$\alpha = 0.05$			Normal		Student's t		Log-normal		Log-logistic	
n	Quantile	ρ	N	B	N	B	N	B	N	B
		0.9	0.081	0.095	0.091	0.091	0.083	0.087	0.070	0.090
	50	0.5	0.064	0.066	0.073	0.071	0.059	0.064	0.048	0.066
		0.1	0.048	0.051	0.055	0.051	0.041	0.050	0.034	0.042
		0.9	0.101	0.114	0.100	0.101	0.112	0.113	0.108	0.084
50	75	0.5	0.075	0.085	0.081	0.077	0.069	0.081	0.066	0.063
		0.1	0.052	0.053	0.057	0.049	0.046	0.052	0.021	0.037
		0.9	0.167	0.189	0.151	0.192	0.154	0.198	0.137	0.169
	90	0.5	0.107	0.124	0.101	0.111	0.093	0.117	0.113	0.121
		0.1	0.066	0.073	0.059	0.050	0.050	0.059	0.033	0.040
		0.9	0.064	0.063	0.047	0.057	0.064	0.070	0.046	0.070
	50	0.5	0.059	0.056	0.044	0.051	0.061	0.064	0.042	0.054
		0.1	0.053	0.045	0.042	0.043	0.059	0.048	0.037	0.034
		0.9	0.056	0.064	0.071	0.080	0.053	0.084	0.091	0.056
200	75	0.5	0.048	0.056	0.056	0.065	0.048	0.056	0.054	0.040
		0.1	0.039	0.041	0.047	0.049	0.037	0.036	0.021	0.026
		0.9	0.091	0.110	0.093	0.104	0.093	0.115	0.091	0.111
	90	0.5	0.077	0.087	0.068	0.076	0.060	0.088	0.076	0.072
		0.1	0.059	0.066	0.045	0.053	0.037	0.057	0.017	0.025

Table 3.4: Simulation results for local Moran's I statistic with minimum number of neighbours using asymptotic inference (N) and bootstrap inference (B), both at 10% significance level.

$\alpha = 0.1$			Normal		Student's t		Log-normal		Log-logistic	
n	Quantile	ρ	N	B	N	B	N	B	N	B
		0.9	0.168	0.180	0.173	0.173	0.176	0.198	0.195	0.180
	50	0.5	0.133	0.152	0.139	0.139	0.141	0.152	0.111	0.123
		0.1	0.106	0.125	0.113	0.108	0.116	0.120	0.081	0.069
		0.9	0.184	0.212	0.210	0.246	0.172	0.228	0.167	0.237
50	75	0.5	0.130	0.161	0.161	0.182	0.121	0.164	0.119	0.164
		0.1	0.094	0.110	0.121	0.110	0.075	0.107	0.041	0.056
		0.9	0.212	0.309	0.147	0.300	0.099	0.313	0.066	0.252
	90	0.5	0.175	0.224	0.125	0.209	0.078	0.222	0.064	0.156
		0.1	0.123	0.111	0.086	0.118	0.046	0.116	0.043	0.062
		0.9	0.122	0.141	0.147	0.119	0.125	0.153	0.127	0.143
	50	0.5	0.110	0.126	0.129	0.107	0.119	0.132	0.095	0.114
		0.1	0.099	0.114	0.111	0.091	0.107	0.112	0.085	0.080
		0.9	0.131	0.160	0.120	0.135	0.147	0.180	0.139	0.180
200	75	0.5	0.119	0.139	0.104	0.115	0.130	0.145	0.088	0.110
		0.1	0.105	0.122	0.091	0.094	0.110	0.114	0.033	0.055
		0.9	0.179	0.193	0.163	0.183	0.141	0.227	0.095	0.203
	90	0.5	0.145	0.155	0.128	0.138	0.100	0.162	0.077	0.141
		0.1	0.114	0.118	0.100	0.093	0.071	0.105	0.023	0.044
$\alpha = 0.05$			Normal		Student's t		Log-normal		Log-logistic	
n	Quantile	ρ	N	B	N	B	N	B	N	B
		0.9	0.097	0.123	0.102	0.101	0.101	0.123	0.069	0.091
	50	0.5	0.077	0.090	0.085	0.077	0.059	0.094	0.044	0.052
		0.1	0.052	0.064	0.065	0.056	0.049	0.066	0.026	0.032
		0.9	0.099	0.142	0.125	0.165	0.081	0.134	0.120	0.152
50	75	0.5	0.064	0.096	0.092	0.106	0.052	0.092	0.102	0.088
		0.1	0.043	0.062	0.068	0.067	0.033	0.050	0.028	0.029
		0.9	0.115	0.220	0.077	0.176	0.058	0.181	0.058	0.142
	90	0.5	0.082	0.134	0.077	0.134	0.055	0.124	0.053	0.088
		0.1	0.059	0.064	0.056	0.058	0.033	0.058	0.038	0.045
		0.9	0.062	0.071	0.073	0.065	0.055	0.089	0.040	0.072
	50	0.5	0.055	0.065	0.059	0.057	0.048	0.077	0.025	0.048
		0.1	0.052	0.058	0.051	0.052	0.044	0.060	0.022	0.024
		0.9	0.066	0.102	0.067	0.077	0.081	0.115	0.106	0.088
200	75	0.5	0.055	0.086	0.055	0.059	0.068	0.092	0.070	0.047
		0.1	0.045	0.071	0.052	0.046	0.055	0.072	0.016	0.026
		0.9	0.092	0.123	0.106	0.102	0.086	0.140	0.074	0.119
	90	0.5	0.072	0.088	0.076	0.076	0.054	0.083	0.068	0.085
		0.1	0.048	0.057	0.058	0.054	0.031	0.046	0.021	0.019

3.4 Data Analysis

In our study, 45 countries and 46 stock indices (USA has two: Standard & Poor's 500 and Dow Jones) are analysed monthly. These countries and stock indices are listed in Table 3 of Section A-1 of the Appendix at the end of this Thesis. Each country is labelled using the two digits notation. The frequencies of words per country and month were added to obtain our proposed GTUI. The period analysed is from January 2004 to March 2021. The four sub-periods distinguished, containing the most important financial crises in the 21st century to date, are as follows: the US sub-prime period between 31 August, 2007 and 30 June, 2009; the Euro debt crisis between 30 June, 2010 and 30 June, 2014; Brexit between 30 June, 2016 and 31 January, 2020; and the COVID-19 pandemic between 29 February, 2020 and 31 March, 2021. To obtain the results of spatial dependence the criterion for defining the neighbours was the median. The criteria based on the most extreme quantiles used in the simulation study of Section 3.3 lead to a very small number of neighbours, possibly even equal to zero, in some periods which causes difficulties in calculating the statistics for testing spatial dependency.

The initial data are the monthly value of market indices MI_{ti}, $t = 1, ..., 207$ and $i = 1, ..., 46$. We used monthly data because the information from Google Trends necessary for the construction of the uncertainty index is available monthly. The monthly losses are $l_{ti} = -log\left(\frac{MI_{ti}}{MI_{(t-1)i}}\right)$, where for $t = 1$ we obtain the value of $t - 1 = 0$. In total we have 46 series of losses, described in Table 3 of Section A-1 of the Appendix at the end of this Thesis. The Shapiro-Wilk test for small samples and the Kolmogorov-Smirnov test for large samples are used to study the normality of the risk series, for all the series the normality assumption is rejected. For each monthly loss series, we study its stationarity in mean and variance and filter the series with the fitted ARMA-GARCH models that are shown in Table 4 of Section A-1 of the Appendix. The series of losses called l_{ti} and the standardised residuals r_{ti} of ARMA-GARCH models, called filtered series, are plotted in Figure 3.1. Three main positive peaks are prominent in the plot on the left. The first corresponds to October 2010 in the middle of the Euro debt crisis, where Iceland reached losses of more than 50%, followed by Peru, Argentina and Russia, whose indices lost 20% of their value. The second peak is in August 2019 and corresponds to the economic crisis in Argentina, a country whose stock index lost more than 23% of its value. The third is in March 2020, coinciding with the health crisis of the COVID-19 pandemic which has affected the whole world and during which, for example, Austria, Spain, Italy and Greece in the Eurozone lost more than 10% of their stock index value. In the plot on the right, showing the standardised residuals of ARMA-GARCH fitted model, the four crisis periods are reflected with greater

instability in the series.

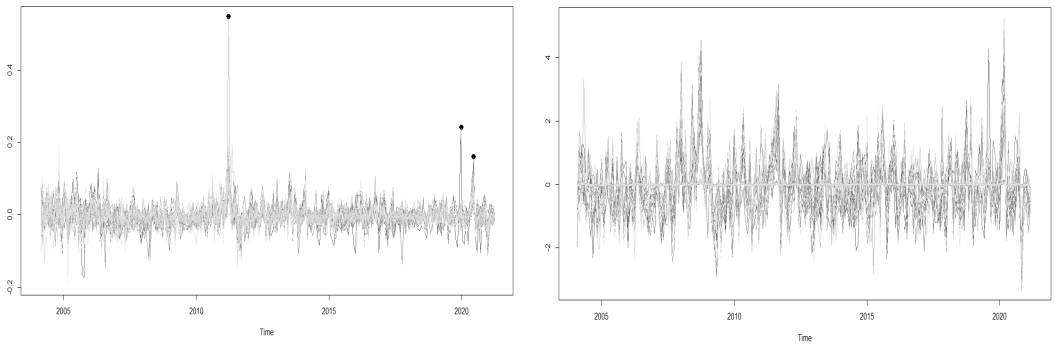


Figure 3.1: Losses (left) and standardised residuals of filtered losses (right) for the 46 stock indices.

With the filtered data we estimate the monthly risk with two alternative measures, the volatility and the VaR at 99% confidence level, using the rolling method with a window of 12 months. This window is selected considering that we work with long monthly series and, in addition to the volatility, where a window of 6 is very common, we must estimate the VaR at 99%, in this case a window of 12 provides less sensitive results. The first year of data was not available, leaving us with information for 195 months, from January 2005 to March 2021. Let r_{it} be the standardised residual of stock market i and month t, the volatility is estimated with the known formula for the variance, $\hat{\sigma}_{it}^2 = \frac{\sum_{s=t-12}^{t}(ris-\bar{r}_{it})^2}{12}$, where $\bar{r}_{it} = \frac{\sum_{s=t-12}^{t} r_{is}}{12}$. The VaR is estimated using the formula that takes into account the deviations from the normal distribution, i.e the modified VaR MVaR). Due to the non-normality of the data and the diversity of the 46 analysed series, parametric and Monte Carlo approaches do not make sense, the MVaR is the most consistent estimation in our case, and it is:

$$MVaR_{it} = \bar{l}_{it} + Z_{CF} \times \hat{\sigma}_{it}.$$

In the MVaR formula the term

$$Z_{CF} = Z_\alpha + \frac{\left(Z_\alpha^2 - 1\right) S}{6} + \frac{\left(Z_\alpha^3 - 3Z_\alpha\right) K}{24} - \frac{\left(2Z_\alpha^3 - 5Z_\alpha\right) S^2}{36}$$

is the Cornish Fisher approximation of the α quantile, where S and $K = k - 3$ refer to skewness and the excess of kurtosis of the data. Our variable, therefore, in formulas (3.5) and (3.11) for global and local Moran statistics are $Y_{it} = \hat{\sigma}_{it}^2$ and $Y_{it} = MVaR_{it}$. In Figure 3.2 both risk measures for the 46 indices are plotted,

including their means (magenta long dashed line) and medians (yellow short dashed line) throughout the analysed period. In all the figures that are shown in this section, the crisis periods are shaded.

Figure 3.2 shows that throughout the analysed period the mean of the risk variable is greater than its median, which reflects the skewness of the data due to the presence of extreme risks, especially during the sub-prime period. In the exercise of simulation of Section 3.3, Table 3.1 shows that in these cases, when spatial dependence is not significant, normal inference tends to reject the null hypothesis more frequently than bootstrap inference, i.e. normality assumption increases the error type I. This result is reflected in our analysis in Figure 3.3 where, along with the value of the global Moran's I statistic, its upper limits at 95% confidence level are plotted, estimated with normal distribution (thin dashed line) and bootstrap (thick dashed line), with the former always below the latter.

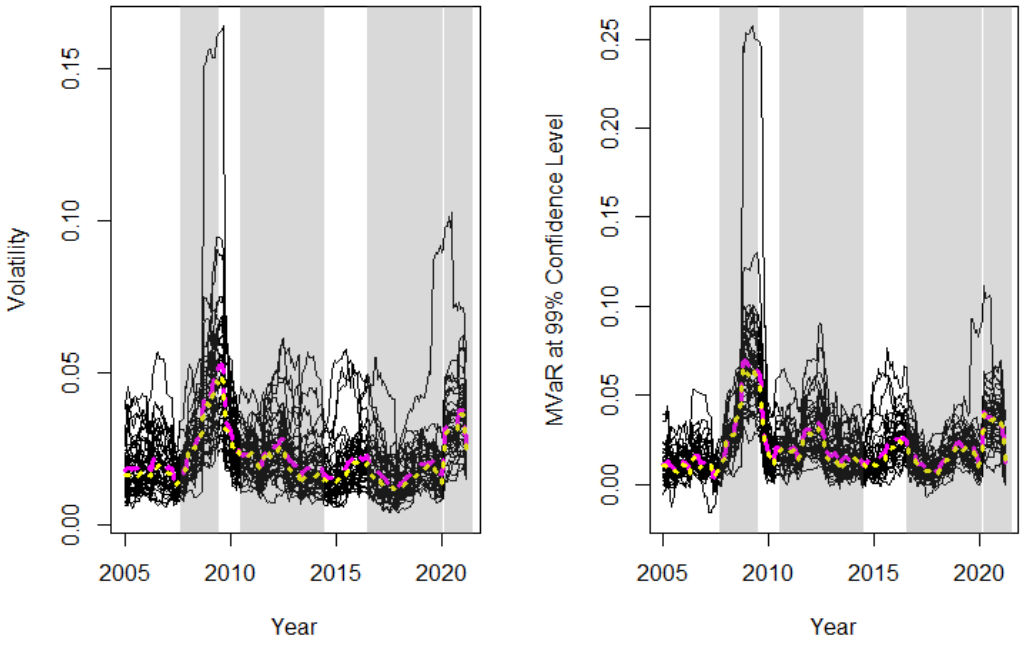


Figure 3.2: Volatility (left) and MVaR at 99% confidence level. The crisis periods are shaded. Mean values plotted in magenta long dashed line and median in yellow short dashed line .

In Figure 3.3, we observe that positive spatial dependence for volatility is more frequent than for MVaR. This is justified given that the former takes into account both tails of the loss distribution and the latter is focused on the right tail. So, on the right we plot positive spatial dependence between extreme losses. For the volatility, the period with more positive spatial dependence is the sub-prime crisis, followed by the Euro debt crisis; however, for the MVaR, the sub-prime period remains the one that causes the greatest positive spatial dependence between extreme losses, but is followed by Brexit. Similar results on US sub-prime crisis have been obtained in recent work by Chopra and Mehta (2022), these authors showed that the sub-prime crisis was the most contagious for the Asian stock markets. In addition, previous works also showed the contagion of this crisis in different financial markets around the world Lien et al. (2018); Mohti et al. (2019).

To complete the results of Figure 3.3, in Table 3.5 we show the test of differences of proportions of months with significant positive spatial dependence between each crisis period and the non-crisis period. The results of Table 3.5 allow us to answer the 3 questions posed in the Introduction of this paper. For volatility, in the four crisis periods the proportion of the months with positive spatial dependence is greater than for the non-crisis period. However, the differences are significant for the sub-prime and Euro debt crises. In contrast, for the spatial dependence estimate with the MVaR the proportion of months for Euro debt crisis is lower than that for the non-crisis period, although the difference is not significant at 5% significance level. In this case, the Brexit period reflects a stronger spatial dependence than the Euro debt period. In relation to the COVID-19 pandemic the results with volatility and MVaR do not show significant differences compared to the non-crisis period. In response to the first question, we can affirm that the volatility spatial dependence, that takes into account the two tails of the distribution (losses and profits), is clearly more frequent in periods of crisis, i.e. spatial dependence occurs in boom and bust periods. Regarding the questions second and third, we observe that there are clear differences between the spatial dependence detected in the different periods of crisis and between spatial dependence criteria (volatility and MVaR).

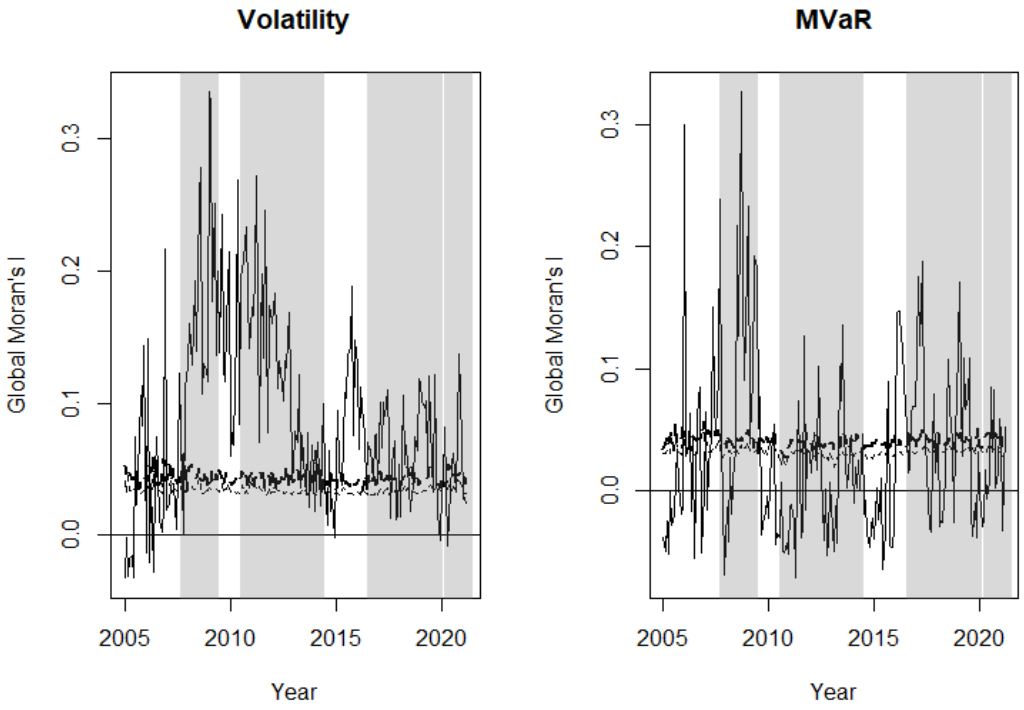


Figure 3.3: Global Moran's I statistic for Volatility (left) and for MVaR at 99% confidence level (right). The crisis periods are shaded. Upper limits at 95% confidence level that have been estimated with normal distribution (thin dashed line) and bootstrap (thick dashed line).

Table 3.5: Frequencies and proportion of months with significant positive global spatial dependence and test at 5% significance level of difference between proportion of months with significant positive spatial dependence in each crisis period with respect to the non-crisis period.

	Volatility			MVaR		
Volatility	Frequency	Proportion	p-value	Frequency	Proportion	p-value
Sub-prime	22	0.957	0.001	13	0.565	0.010
Euro Debit	42	0.857	0.002	12	0.245	0.287
Brexit	30	0.682	0.239	22	0.500	0.014
COVID-19	6	0.429	0.901	5	0.357	0.316
Total crisis	100	0.769	0.012	52	0.400	0.070
Total no crisis	40	0.615		19	0.292	

Next, we analyse the local spatial dependency for the market indices of the fol-

lowing countries: Spain, Germany, France, Italy, UK, US, Argentina, Brazil, Japan and Hong Kong, in total 11 indices given that the US has two. The upper confidence limits are carried out with Bonferroni correction for multiple null hypothesis (11 in our case). In Figure 3.4 the monthly number of neighbours for different groups of countries is plotted. It is in Japan and Hong Kong where the trend shows an increase. In the EU, Italy has the greatest number of neighbours throughout almost all of the period, followed by Spain, while Germany is the EU country with the fewest neighbours of uncertainty. Argentina is below Brazil and the UK has a more unstable behaviour in terms of number of neighbours than the US.

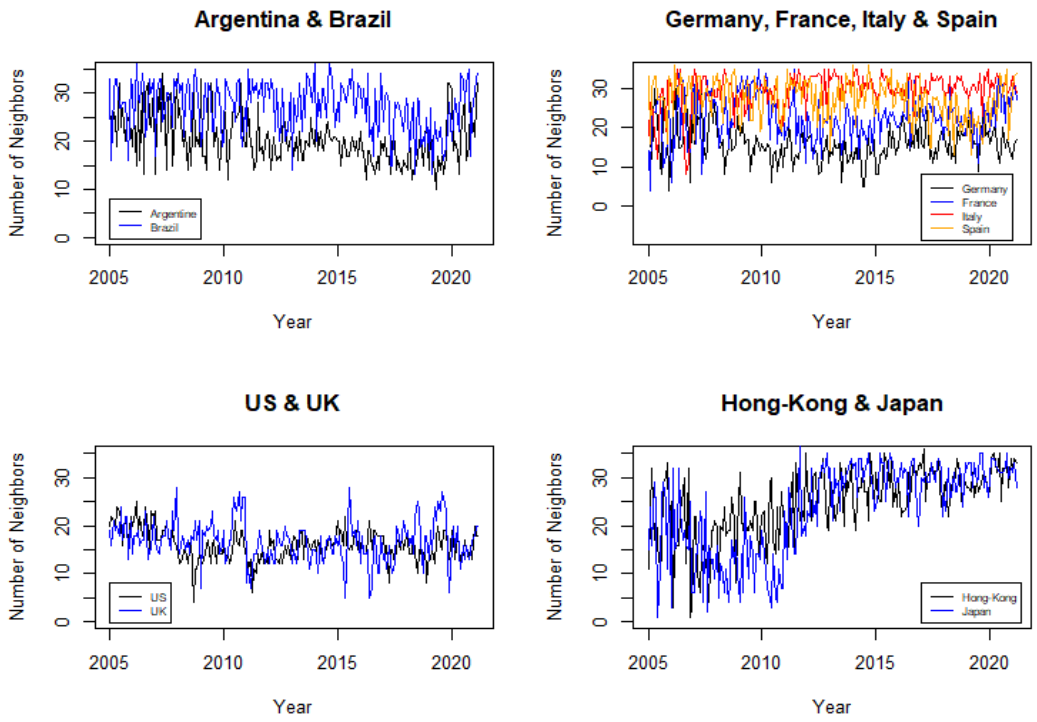


Figure 3.4: Monthly number of neighbours.

The figures from the 3.5 to the 3.9 plot the local spatial dependence statistics (solid line) and the bootstrap upper limits at 95% confidence level (dashed line) for testing significant positive local spatial dependence. These figures are obtained focusing on right tail of the loss distribution, i.e. using MVaR for the 11 indices. Specifically, Figure 3.5 shows the results for Argentina and Brazil, Figure 3.6 for the four the EU markets, Figure 3.7 for UK, Figure 3.8 for US and, finally, the local spatial dependence results for Japan and Hong Kong are shown in Figure 3.9. The results obtained with the volatility are omitted; similar to global spatial dependence,

these show a somewhat stronger dependence. A significant local spatial dependence implies that the country spreads its situation of extreme losses to its neighbouring countries, i.e., neighbours with similar uncertainty. These linkages between stock market losses are apparent for some countries and for certain specific periods. Figure 3.5 shows a significant local positive spatial dependence for Argentina just before the Brexit referendum and at the end of this period. With regard to the EU countries, Figure 3.6 shows that it is the French stock market that shows a significant spatial dependence before and during the sub-prime period, as well as some months with strong local dependency during the Brexit and COVID-19 period. The UK shows similar results to France, although the period of strong local spatial dependence before and after the Brexit referendum is particularly prominent. The US indices are those that show more frequent local spatial dependence in all the crisis periods apart from during the COVID-19 pandemic. Finally, the two Asian stock markets have significant local spatial dependence in the sub-prime period, and more so in Japan.

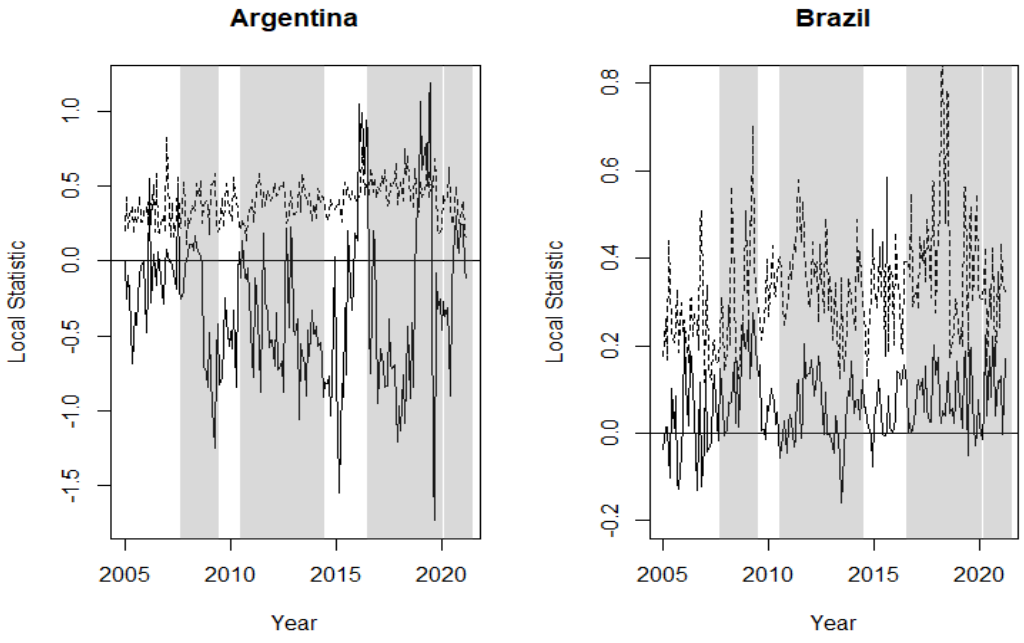


Figure 3.5: Local Moran's I statistic for MVaR at 99% confidence level (solid line) and bootstrap upper limits at 95% confidence level (dashed line). The crisis periods are shaded. Results for Argentine and Brazil stock indices.

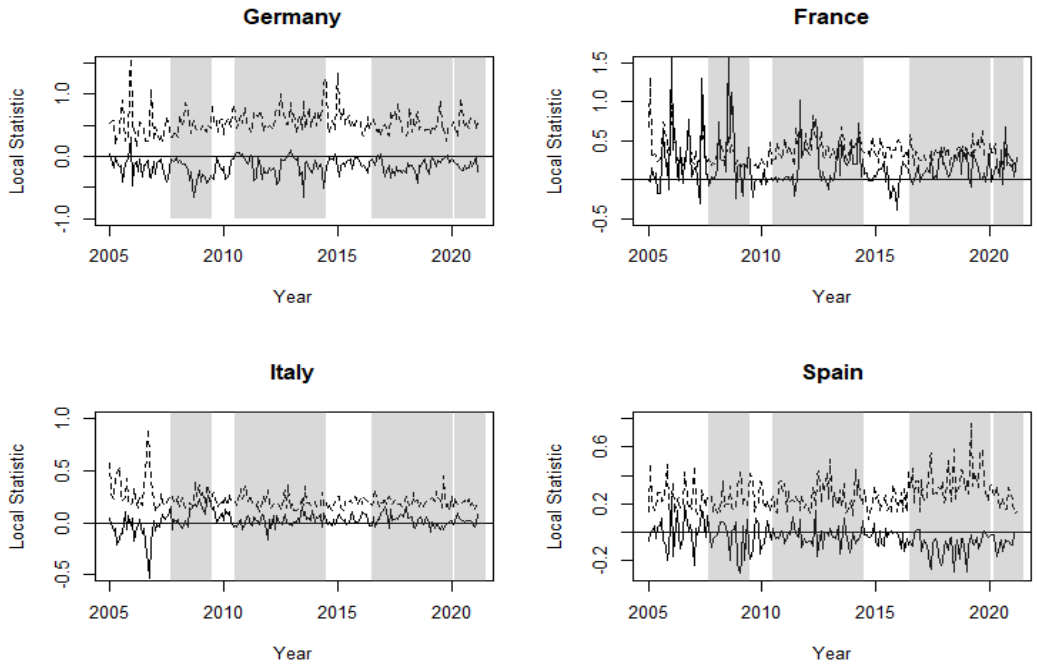


Figure 3.6: Local Moran's I statistic for MVaR at 99% confidence level (solid line) and bootstrap upper limits at 95% confidence level (dashed line). The crisis periods are shaded. Results for EU stock indices.

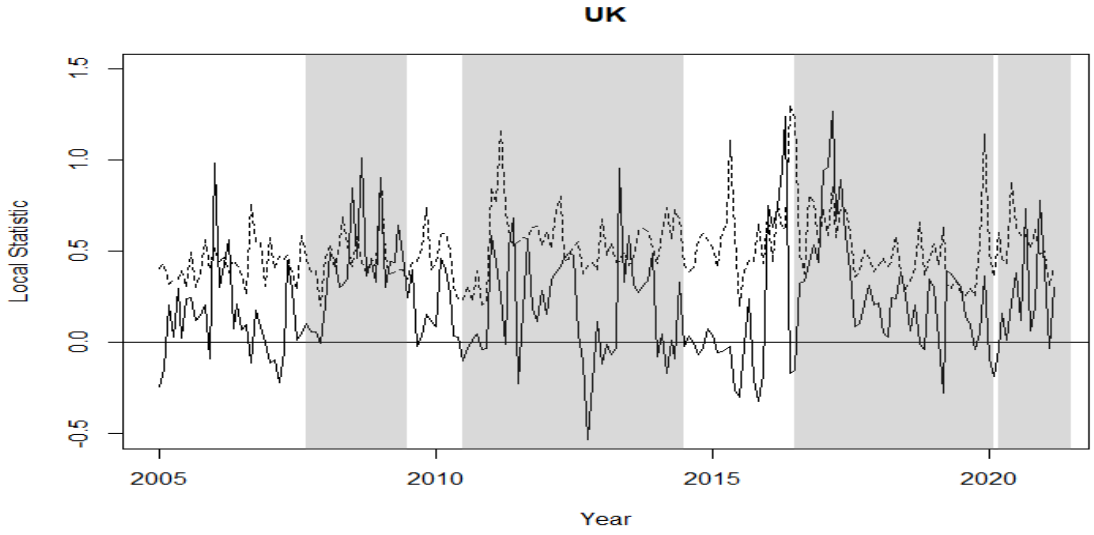


Figure 3.7: Local Moran's I statistic for MVaR at 99% confidence level (solid line) and bootstrap upper limits at 95% confidence level (dashed line). The crisis periods are shaded. Results for UK stock index.

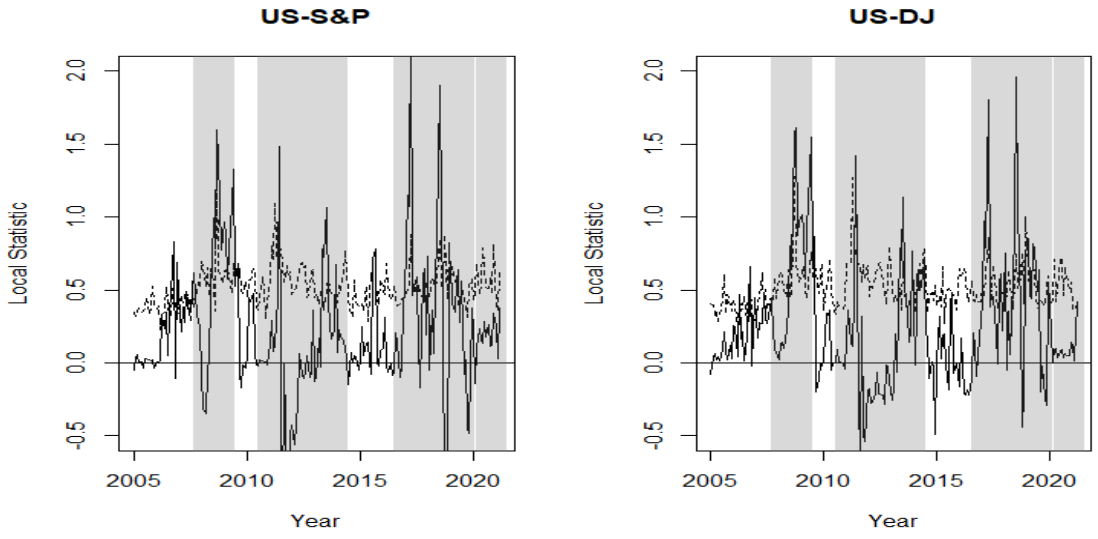


Figure 3.8: Local Moran's I statistic for MVaR at 99% confidence level (solid line) and bootstrap upper limits at 95% confidence level (dashed line). The crisis periods are shaded. Results for US stock indices.

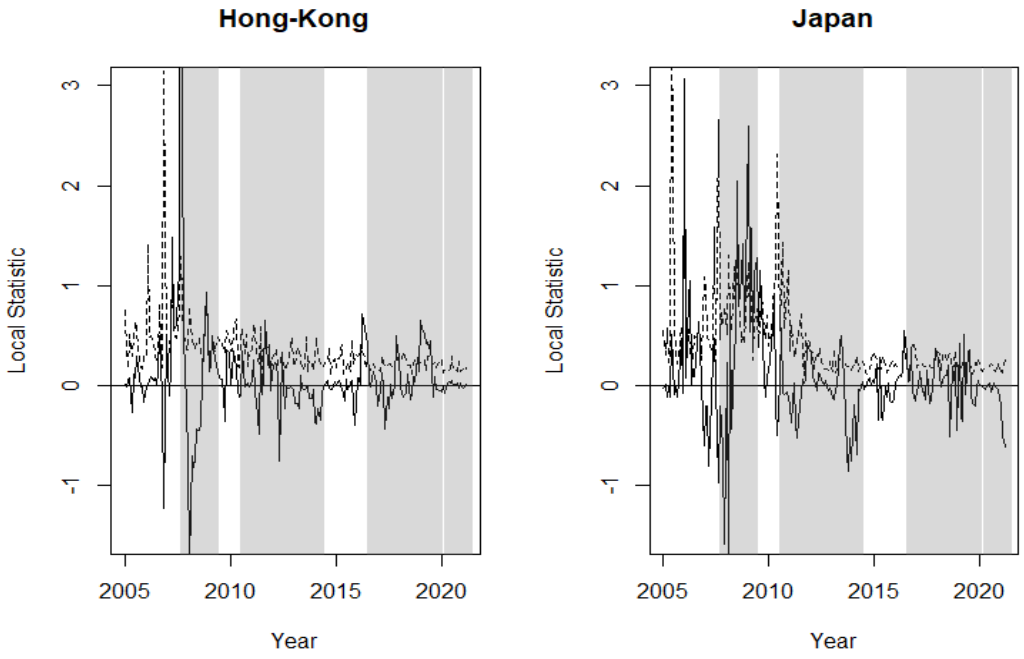


Figure 3.9: Local Moran's I statistic for MVaR at 99% confidence level (solid line) and bootstrap upper limits at 95% confidence level (dashed line). The crisis periods are shaded. Results for Japan and Hong Kong stock indices.

We observe that between non-EU countries, Argentina has a certain level of contagion both in the sub-prime crisis and due to Brexit, but it is relatively lower compared to the EU countries analysed. It is interesting to note that countries such as Germany and Spain have financial markets relatively isolated from the others, since they do not have relevant spatial correlation.

With respect to the other countries analysed, there are certain relevant impulses: the US indices clearly felt an impact from the Brexit crisis due to their interconnection with the UK economy; France, curiously, has connections for all crises except Brexit, since it has somehow benefited from it.

With regard to the possibility of detecting crises in advance, through relevant impacts on the spatial correlation, we see that the Euro debt crisis does not show any sign of impact, nor does Brexit, but the sub-prime crisis does show in almost all countries an increase in the relevant spatial correlation. We could perhaps consider that the type of crisis can influence contagion before or during it and that the markets adjust their expectations differently depending on what they are like.

3.5 Conclusions

This chapter presents a complete study of the analysis of global and local spatial dependence with global and local Moran's I statistics in the context of financial markets and their uncertainty.

A simulation study is presented which shows how, when there are extreme values in the right tail of the distribution, the inference with global and local Moran's I statistics based on the normal distribution increases the type I error. The same tests are also done based on the bootstrap resampling technique.

Following the analysis of the data, there are some interesting findings. The period of the global financial crisis of the sub-prime is the one that caused more linkages between the extreme losses of the analysed stock indices. We can therefore conclude that systemic risk in this period caused more losses than in the others periods. The Euro debt period is the one with less global spatial dependence between extreme losses, followed by the COVID-19 pandemic period. Between the 10 countries whose local spatial dependences were analysed, those that were contagious for their neighbours of uncertainty in certain months were Argentina, France, the US, the UK, Japan and Hong Kong.

The impact of the financial crisis of the sub-prime is the most visible in our study such that it presents a greater presence of joint or systemic risk, where the most extreme losses are greater. However, the opposite is seen in two other crises analysed, Euro debt and COVID-19. We consider that a possible justification for this comes from the fact that their impact was more indirect on the markets and more adequate management tools were created. With regard to Brexit, which was consolidated little by little, we see that there are no signs of contagion before it, but there are during it in the countries with more interconnected economies: the US, Hong Kong, Japan and, of course, the UK. Finally, we would like to mention a degree of isolation of certain countries in terms of contagion -Germany, Spain, Italy and Brazil- this may because these countries have different structures than those that are more closely related. The greater variability of the local Moran indicator itself could be an early indicator of the crises themselves, and detecting which markets are more sensitive to them could be a way of preparing the latter for any adjustments needed.

Chapter 4

A new kernel estimator of copulas based on Beta quantile transformations

4.1 Introduction

Based on the kernel method and transformations, we present a new nonparametric estimator of a multivariate copula that improves the empirical copula and the most prominent kernel estimators (see Omelka et al., 2009, for a detailed review). We use the new estimator to analyse and test the extreme value dependence between the losses in the Spanish stock market index and different stock market indices of Europe, USA and China.

The copula model allows us to represent the dependence structure of a multivariate random vector of continuous variables $\mathbf{X} = (X_1, ..., X_J)'$, which combines with marginal distributions to give the multivariate distribution. This idea was established in the fundamental theorem proposed by Sklar (1959). This theorem shows that a multivariate cumulative distribution function (cdf) H of the random vector $\mathbf{X}$, with marginal distributions functions $F_1, ..., F_J$, has associated a copula C, so that:

$$H(x_1, ..., x_J) = C(F_1(x_1), ..., F_J(x_J)). \tag{4.1}$$

In practice, the dependence structure and marginal distributions are unknown and both will need to be fitted. We assume that marginal cdfs can be easily adjusted using parametric distributions or nonparametric methods and we focus on the fitting of dependence structure using a copula. It is often difficult by visualizing the data to select the appropriate dependency structure and, therefore, the right copula model. Alternatively, a nonparametric estimation of a copula can be obtained whose results can be used for estimating joint probabilities or for testing the adequacy of a copula family, for example, the extreme value copula family. In this paper, these two aims of our new nonparametric estimator are analysed through a simulation study.

Because there are a lot of dependence structures represented by different copulas families, specific tests for choosing the best copula are useful. The approach for developing a test for the adequacy of copulas takes its lead from, for example, the proposal of Genest and Rivest (1993) for bivariate Archimedean copulas; the test of Scaillet (2005) on inference for the positive quadrant dependence hypothesis; the test for equality between two copulas of Rémillard and Scaillet (2009) or the test of symmetry for bivariate copulas of Genest et al. (2012).

On inference for extreme value copulas, alternative types of tests have been proposed, among which the most well known are the test of Genest et al. (2011) based on a Cramér-von Mises statistic, the test analysed by Ghorbal et al. (2009) based on an U-statistic and the test of Kojadinovic et al. (2011) that uses the $max - stable$ property and is also based on a Cramér-von Mises statistic (see also Bahraoui et al., 2014, for complete properties of the test based on $max - stable$ property).

The test proposed by Kojadinovic et al. (2011) is based on the empirical copula that is equivalent to the multivariate empirical distribution. However, the empirical copula is inefficient for certain shapes of distribution, for example, when the marginal cdfs are associated with extreme value distributions. Alternatively, Omelka et al. (2009) analyse how testing extreme value copula can be based on different kernel estimators. The main difficulty of a classical kernel estimator is its bias on the boundaries when the function values at these points are positive. Based on this concern, Chen and Huang (2007) analyse the kernel copula estimator with local linear boundary correction which the authors proved reduces bias and variance. Alternatively, Omelka et al. (2009) propose the transformation of a kernel copula estimator based on standard normal inverse distribution function transformations, which is very easy to implement and has the same weak convergence properties as the previous proposal. In this paper, an improved transformed kernel estimator is proposed that has the same weak convergence properties and is useful for the inference on extreme value copulas. The theoretical results are shown for the bivariate case, but they are easily extrapolated to the multivariate case.

In Section 4.2, we present the background on kernel estimation of copulas, the new estimator and its theoretical asymptotic properties for testing the max-stable property of extreme value copulas. The Section 4.3 presents the simulation results that allow us to analyse finite sample properties and inference errors Type 1 and Type 2. As an illustration, in Section 4.4, a financial risk analysis is carried out where the extreme value copula family hypothesis between the Spanish stock market index and different neighbouring and non-neighbouring countries is tested. Finally, we conclude in Section 4.5.

4.2 Kernel estimation of copulas

Let $(X_{i1}, X_{i2})'$, $\forall\ i = 1, ..., n$, be a sample of n independent and identically distributed (i.i.d.) bivariate data, the product kernel estimator of the bivariate cdf can be expressed as:

$$\widehat{H}(x_1, x_2) = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{x_1 - X_{i1}}{b_1}\right) K\left(\frac{x_2 - X_{i2}}{b_2}\right), \tag{4.2}$$

where K is the cdf associated with the kernel function k, that is a bounded or asymptotically bounded and symmetric probability density function (pdf) (see Liu and Yang, 2008; Wang et al., 2013, for a review on kernel estimation of the multivariate distribution function). Examples of such functions are the Epanechnikov and the Gaussian kernels Silverman (1986). The parameters $b_1 > 0$ and $b_2 > 0$, known as the bandwidths or smoothing parameters, control the smoothness of the estimation. Thus, the larger the value of b_1 and b_2, the smoother the resulting function. Their values depend on the sample size n - the biggest sample size n, the lower the smoothing parameters - but obtaining optimal values for these smoothing parameters is one of the greatest difficulties posed by the kernel estimation.

Based on Slark's theorem, from (4.2) we specify the kernel estimator of copula as:

$$\tilde{C}\left(\widehat{F}_1(x_1), \widehat{F}_2(x_2)\right) = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{x_1 - X_{i1}}{b_1}\right) K\left(\frac{x_2 - X_{i2}}{b_2}\right), \tag{4.3}$$

where $\widehat{F}_j(x_j)$, $j = 1, 2$, are estimators of the marginal cdfs that, in practice, can be obtained based on a parametric distribution or with a non-parametric estimator. Given that the copulas allow us to separate dependence structure from marginal distribution, we focus on estimating the first; so, the aim is to estimate a multivariate cdf with $Uniform(0, 1)$ marginal distributions, whose kernel estimator for bivariate case is expressed as:

$$\widetilde{C}(u_1, u_2) = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{u_1 - U_{i1}}{b}\right) K\left(\frac{u_2 - U_{i2}}{b}\right), \tag{4.4}$$

where, unlike (4.2), given that the marginal distributions are $Uniform(0, 1)$, we assume $b_1 = b_2 = b$ and $b \to 0$ as $n \to \infty$, taking into account the relationship between b and n, hereinafter we denote it as b_n. In practice, we need to define observations $(U_{i1}, U_{i2})'$, $\forall\ i = 1, ..., n$, the values of the marginal empirical distributions $U_{ij} = \frac{1}{n} \sum_{k=1}^{n} I\left(X_{kj} \leq X_{ij}\right)$, $j = 1, 2$ and $i = 1, ..., n$, are a natural choice. However, it is known that empirical distribution takes value 1 at the maximum value observed and most of the commonly used copulas (Gumbel,

Clayton, Gaussian and Student's t) are not finite derivatives (copula density values) at corners $(0,0),(1,1),(1,0),(1,1)$; then, these empirical distributions are replaced by corrected versions that are known as pseudo-data and that can be defined as: $\hat{U}_{ij} = \frac{1}{n+1}\sum_{k=1}^{n} I\left(X_{kj} \leq X_{ij}\right)$ or, as Chen and Huang (2007) suggested, $\hat{U}_{ij} = \frac{1}{n}\sum_{k=1}^{n} I\left(X_{kj} \leq X_{ij}\right) - \frac{1}{2n}$. So, the kernel estimator of a copula is defined as:

$$\widehat{C}\left(u_1,u_2\right) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{u_1-\hat{U}_{i1}}{b_n}\right) K\left(\frac{u_2-\hat{U}_{i2}}{b_n}\right). \tag{4.5}$$

To obtain the estimator defined in (4.5) a kernel function, K, needs to be selected that will have minimal effect on the results obtained, and to calculate the bandwidth b_n, whose value will have an important effect on the estimated copula. The bandwidth b_n can be calculated using some cross-validation or plug-in method or using the rule-of-thumb proposed by Silverman (1986) for the kernel estimator of pdf adapted to the kernel estimator of cdf Hill (1985); Liu and Yang (2008).

The properties of a kernel estimator depend on some smoothness characteristics of the cdf; in our context in particular, it is a requirement that the first two derivatives take finite values different from zero. Furthermore, when the distribution has a bounded domain and the density at boundary takes positive values, as in the case of the bivariate copula with domain on $[0,1]^2$, the estimator defined in (4.4) has boundary bias. This means that the kernel estimator at boundary is not consistent (see Wand and Jones, 1995, pp. 46 and 47, for a clear description in the kernel density estimator context). This is problematic since our aim is to test if our data is generated by an extreme value copula. There are three alternative proposals to achieve consistency at boundary of a kernel estimator of a copula. Boundary kernel methods are the most common techniques proposed in the context of kernel regression and density estimation Gasser and Müller (1979); Gasser et al. (1985), the main difficulty with the use of this type of kernel being that it does not integrate one which, in practice, could be inconvenient. Chen and Huang (2007) proposed a kernel estimator of copulas with linear boundary correction, the weakness of their method is that for many common families of copulas (e.g., Clayton, Gumbel, Gaussian and Student's t) the bias at some of the corners of the unit square is only of order $O(b_n)$, versus the $O(b_n^2)$ that is reached in the central values of the domain, where $O(\cdot)$ is the asymptotic order operator. Another way to correct boundary bias is using the mirror-reflection kernel estimator, this method being proposed by Gijbels and Mielniczuk (1990) to estimate the density of the copula. In all cases, the main difficulty of a kernel estimator with or without boundary bias is calculating the smoothing parameter whose value will greatly affect the results.

An alternative strategy to avoid boundary bias and to calculate the smoothing pa-

rameter easily is to transform $Uniform(0,1)$ marginal distributions of the copula so that the kernel estimator of the new marginal distributions does not have boundary bias and their shapes allow us to minimise the bias of the kernel estimator. This idea also addresses the problems of the estimator defined in (4.3) based on the original scale of the data. On the one hand, although the marginal distributions are not uniform, they can have shapes that could also be subject to inconsistency at the boundaries, i.e. the distribution could have bounded domain on one or both sides with positive density. On the other hand, the problems associated with the kernel estimator defined in (4.3) are widely known when the distribution to estimate has one or two long tails (see Buch-Larsen et al., 2005; Alemany et al., 2013; Bolancé and Guillen, 2021).

The transformed estimator of the copula is based on the equality:

$$C\left(u_1,u_2\right)=C^T\left(T(u_1),T(u_2)\right),$$

i.e., the values of the copula function C evaluated on original $Uniform(0,1)$ scale are equal to the values of function C^T evaluated on transformed scale. So, the transformed kernel estimator (TKE) of a copula is defined as:

$$\begin{aligned}\widehat{C}^T\left(u_1,u_2\right) &= \frac{1}{n}\sum_{i=1}^{n}K\left(\frac{T(u_1)-T(\hat{U}_{i1})}{b_n}\right)K\left(\frac{T(u_2)-T(\hat{U}_{i2})}{b_n}\right)\\ &= \widehat{\widehat{C}}\left(u_1,u_2\right),\end{aligned} \tag{4.6}$$

where $T(\cdot)$ is a transformation which is equal to the inverse of a given continuous cdf. The estimator defined in (4.6) has a fundamental advantage over the kernel estimator defined in (4.5) and its versions that incorporate boundary bias reduction; given that the function $T(\cdot)$ is the inverse of a given cdf, we know the marginal distributions of C^T and the bandwidth can be calculated based on these distributions.

Omelka et al. (2009) proposed that $T=\Phi^{-1}$, where Φ is the cdf of the standard normal distribution. This standard normal transformation is based on the idea that the normal distribution does not have boundary bias problems and it can be estimated easily using a classical kernel estimator. This transformed estimator is called Gaussian transformed kernel estimator and is defined as:

$$\widehat{C}^G\left(u_1,u_2\right)=\frac{1}{n}\sum_{i=1}^{n}K\left(\frac{\Phi^{-1}(u_1)-\Phi^{-1}(\hat{U}_{i1})}{b_n}\right)K\left(\frac{\Phi^{-1}(u_2)-\Phi^{-1}(\hat{U}_{i2})}{b_n}\right). \tag{4.7}$$

In practice, in this case the value of bandwidth can be calculated using the idea of rule-of-thumb of Silverman (1986) applied to the standard normal marginal cdfs,

that is $b_n = 3.572n^{-\frac{1}{3}}$. In the simulation study presented in Section 4.3, we show the difference between the mean integrated squared error of a copula, $MISE = \int_0^1 \int_0^1 \left[\hat{C}(u_1,u_2) - C(u_1,u_2)\right]^2 du_1 du_2$, using optimal b_n and using the proposed rule-of-thumb.

We propose an alternative estimator to the one defined in (4.7) using a transformation T that is better than Φ^{-1}. Our proposal is based on the second-order approximation properties of univariate kernel estimator of marginal distributions. When $b_n \to 0$ as $n \to \infty$, f is a continuous pdf and the first derivative f' exists, the bias and variance of kernel estimator of cdf are (see Reiss, 1981; Azzalini, 1981; Hill, 1985):

$$E\left[\widehat{F}(y)\right] - F(y) = \frac{1}{2}b^2 f'(y)\int_{-1}^{1} t^2 k(t)\,dt + o\left(b^2\right) \tag{4.8}$$

and

$$V\left[\widehat{F}(y)\right] = \frac{F(y)\left[1-F(y)\right]}{n} - f(y)\frac{b}{n}\int_{-1}^{1} K(t)\left[1-K(t)\right]dt + o\left(\frac{b}{n}\right). \tag{4.9}$$

By addition of the integrated variance and the integrated squared bias, we can approximate the MISE of the kernel estimation of marginal distributions as:

$$\begin{aligned} MISE\left[\widehat{F}(y)\right] &= \int \frac{F(y)\left[1-F(y)\right]}{n}dy - \frac{b}{n}\int_{-1}^{1} K(t)\left[1-K(t)\right]dt \\ &+ \frac{1}{4}b^4 \int \left[f'(y)\right]^2 dy \left(\int_{-1}^{1} t^2 k(t)\,dt\right)^2 \\ &+ o\left(b^4\right) + o\left(\frac{b}{n}\right), \end{aligned} \tag{4.10}$$

where the integral limits are given by the domain of argument variable Y. From expression (4.10) it is easy to deduce that the distribution that minimises MISE also minimises the functional $\int \left[f'(y)\right]^2 dt = \int \left[F''(y)\right]^2 dt$. Terrell (1990) found the pdf family that minimises the functionals of type $\int \left[f^{(p)}(y)\right]^2 dy$, where p is the order of the derivative. This principle was applied to cdf and quantile kernel estimation by Alemany et al. (2013), who showed how the $Beta(3,3)$, whose pdf and cdf are:

$$\begin{aligned} m(t) &= \frac{15}{16}\left(1-t^2\right)^2, -1 \le t \le 1 \text{ and} \\ M(t) &= \frac{3}{16}t^5 - \frac{5}{8}t^3 + \frac{15}{16}t + \frac{1}{2}, \end{aligned} \tag{4.11}$$

minimises the functional $\int \left[F_j''(t_j)\right]^2 dt_j$, $j = 1,2$, and therefore minimises the in-

tegrated bias of the classical kernel estimator of a cdf. Subsection 4.2.1 includes the theoretical results on testing extreme value copulas, and Theorem 4.1 shows as the cdf M has the properties that allow us to conclude that the kernel estimator of $Beta(3,3)$ does not have boundary bias (see Omelka et al., 2009). So, the Beta transformed kernel estimator of a copula is:

$$\begin{aligned} &\widehat{C}^B(u_1,u_2) = \\ = \ &\frac{1}{n}\sum_{i=1}^{n} K\left(\frac{M^{-1}(u_1)-M^{-1}(\hat{U}_{i1})}{b_n}\right) K\left(\frac{M^{-1}(u_2)-M^{-1}(\hat{U}_{i2})}{b_n}\right) \end{aligned} \tag{4.12}$$

where b_n can be calculated using rule-of-thumb applied to the $Beta(3,3)$ marginal distributions, that is $b_n = 3^{\frac{1}{3}} n^{-\frac{1}{3}}$. In the simulation study shown in Section 4.3, the MISE calculated with this bandwidth is compared to the one that minimises MISE.

Next, we present some theoretical results related to the weak convergence to a Gaussian process $\mathbb{G}$ of the estimator defined in (4.12) and the max-stable property for testing extreme value copulas.

4.2.1 Theoretical results

We use the result from Fermanian et al. (2004) for the weak convergence of the kernel estimator of a copula defined in (4.5) to a Gaussian process $\mathbb{G}$ in the space of all bounded real-valued functions on $[0,1]^2$, i.e. $l^\infty([0,1]^2)$, which is expressed as follows:

$$\begin{aligned} &\sqrt{n}\left(\widehat{C}(u_1,u_2)) - C(u_1,u_2)\right) \longmapsto \mathbb{G}(u_1,u_2) = \\ = \ &\mathbb{B}(u_1,u_2) - \partial_1 C(u_1,u_2)\mathbb{B}(u_1,1) - \partial_2 C(u_1,u_2)\mathbb{B}(1,u_2), \end{aligned} \tag{4.13}$$

where $\partial_j C(u_1,u_2), j=1,2$, are the partial derivatives of the function C with respect to u_j, $\longmapsto$ indicates weak convergence and $\mathbb{B}$ is a Brownian bridge on $[0,1]^2$ with covariance function:

$$E[\mathbb{B}(u_1,u_2)\mathbb{B}(u_1',u_2')] = C(u_1 \wedge u_1', u_2 \wedge u_2') - C(u_1,u_2)C(u_1',u_2'),$$

where $\wedge$ is the minimum.

The weak convergence defined in (4.13) requires that the copula has continuous partial derivatives. Furthermore, Omelka et al. (2009) proved the weak convergence of local linear, mirror reflection and Gaussian transformed kernel estimators of copula. These authors remark that it is sufficient to assume that the first partial derivatives are continuous on $(0,1)^2$, i.e. we can eliminate the corners. This is an

important result, given that most of the commonly used copulas (Clayton, Gumbel, Normal and Student's t) do not have finite partial derivatives at the corners.

The weak convergence of our Beta transformed kernel estimator is defined in the following theorem.

Theorem 4.1. *Let us suppose a continuous copula C, with continuous first order partial derivatives and bounded second order partial derivatives on $(0,1)^2$ that satisfies the following asymptotic properties: $\partial_1^2 C(u_1,u_2) = O\left(\frac{1}{u_1(1-u_1)}\right)$, $\partial_2^2 C(u_1,u_2) = O\left(\frac{1}{u_2(1-u_2)}\right)$ and $\partial_1\partial_2 C(u_1,u_2) = O\left(\frac{1}{\sqrt{u_1u_2(1-u_1)(1-u_2)}}\right)$. If $b_n = O\left(n^{-\frac{1}{3}}\right)$ the Beta transformed kernel estimator $\widehat{C}^B$ meets the weak convergence defined in (4.13).*

Proof. **Proof of Theorem 4.1**

Let $J(t) = T^{-1}(t)$ be the inverse transformation function in $\widehat{C}^T$, the proof of Theorem 4.1 comes directly from the results of Theorem 2 in Omelka et al. (2009), who proved that, if the first derivative $J'(t)$ and $\frac{[J'(t)]^2}{J(t)}$ are bounded, then $\widehat{C}^T$ converge weakly to the Gaussian process $\mathbb{G}$. For $\widehat{C}^B$ we have that $J(t) = M(t) = \frac{3}{16}t^5 - \frac{5}{8}t^3 + \frac{15}{16}t + \frac{1}{2}$, $J'(t) = m(t) = \frac{15}{16}(1-t^2)^2$ and $\frac{\left(M'(t)\right)^2}{M(t)} = \frac{\left(\frac{15}{16}(1-t^2)^2\right)^2}{\frac{3}{16}t^5 - \frac{5}{8}t^3 + \frac{15}{16}t + \frac{1}{2}}$, $|t| < 1$. Directly, we know that the pdf m of the $Beta(3,3)$ is bounded. Moreover, if the quotient $\frac{\left(M'(t)\right)^2}{M(t)}$ is analysed, the maximum is approximately found at $t \approx -0.45332$. □

The weak convergence of Theorem 4.1 allows us to use $\widehat{C}^B$ for the inference on copulas. We focus on an extreme value copula test based on the proposal of Kojadinovic et al. (2011), that analyses the $max-stable$ property associated with this family of copulas (see, for example, Segers, 2012). A copula is $max-stable$ if $\forall r > 0$ and $\forall u_1, u_2$ in $[0,1]$ the null hypothesis $H_0^r : C(u_1,u_2) = C^r(u_1^{1/r}, u_2^{1/r})$ is not rejected from the alternative $H_1^r : C(u_1,u_2) \neq C^r(u_1^{1/r}, u_2^{1/r})$. In practice, we test the $max-stable$ hypothesis using some values of $r \geq 1$ (see Kojadinovic et al., 2011),

$$H_0 : \bigcap_{r\geq 1} H_0^r$$
$$H_1 : \bigcup_{r\geq 1} H_1^r.$$

To test the previous hypotheses we propose estimating $\mathbb{D}^r(u_1,u_2) = \sqrt{n}\left(C(u_1,u_2) - C^r\left(u_1^{1/r}, u_2^{1/r}\right)\right)$ using the Beta transformed kernel estimator of the copula, i.e. $\widehat{\mathbb{D}}^r(u_1,u_2) = \sqrt{n}\left(\widehat{C}^B(u_1,u_2) - \widehat{C}^{B^r}\left(u_1^{1/r}, u_2^{1/r}\right)\right)$.

Proposition 4.2. *If the partial derivatives of the copula $C(u_1, u_2)$ are continuous then for any $r > 0$ we have:*

$$\widehat{\mathbb{D}}^r(u_1,u_2) - \mathbb{D}^r(u_1,u_2) \longmapsto \mathbb{C}^r(u_1,u_2) = \\ rC^{r-1}(u_1^{1/r},u_2^{1/r})\mathbb{G}(u_1^{1/r},u_2^{1/r}) - \mathbb{G}(u_1,u_2), \tag{4.14}$$

in $l^\infty([0,1]^2)$.

Proof. **Proof of Proposition 4.2**

The result in Proposition 4.2 is obtained from:

$$\widehat{\mathbb{D}}^r(u_1,u_2) - \mathbb{D}^r(u_1,u_2) = \\ \sqrt{n}\left[\left(\widehat{C}^B(u_1,u_2) - C(u_1,u_2)\right) - \left(\widehat{C}^{B^r}(u_1^{1/r},u_2^{1/r}) - C^r(u_1^{1/r},u_2^{1/r})\right)\right].$$

Using the convergence of Theorem 4.1:

$$\sqrt{n}\left(\widehat{C}^B(u_1,u_2) - C(u_1,u_2)\right) \longmapsto \mathbb{G}(u_1,u_2).$$

We now need to prove the weak convergence of $\sqrt{n}\left(\widehat{C}^{B^r}(u_1^{1/r},u_2^{1/r}) - C^r(u_1^{1/r},u_2^{1/r})\right)$. To this end, we use the result of Kojadinovic et al. (2011), that proved the weak convergence of this difference for empirical copula (see also Bahraoui et al., 2014). In general, this result can be directly extrapolated to the kernel estimator and, in particular, to the Beta transformed kernel estimator, considering that $\widehat{\widehat{C}}^r\left(u_1^{1/r},u_2^{1/r}\right) = \widehat{C}^{B^r}\left(M^{-1}(u_1^{1/r}),M^{-1}(u_2^{1/r})\right)$. Then, under H_0, $\mathbb{D}^r(u_1,u_2) = 0$, $\widehat{\mathbb{D}}^r(u_1,u_2)$ it weakly converges to process (4.14). □

For hypothesis testing given a fixed r, we use a Cramér-von Mises statistic:

$$\widehat{S}^r = \int_0^1\int_0^1 \left(\widehat{\mathbb{D}}^r(u_1,u_2)\right)^2 du_1 du_2 \tag{4.15}$$

and for a range of values $r_1,...,r_t$, the following statistic can be considered:

$$\widehat{S}^{r_1,...,r_t} = \sum_{i=1}^{t} \widehat{S}^{r_i}. \tag{4.16}$$

For implementing the test based on $\widehat{S}^{r_1,...,r_t}$, we use the numerical approximation proposed by Kojadinovic et al. (2011), replacing the empirical copula by a Beta transformed kernel estimator of the copula. The procedure is as follows:

1. The statistics $\widehat{S}^{r_l}$ are approximated using a uniformly spaced grid (u_{j1}, u_{j2}), $j = 1, ..., m$, of points on $(0,1)^2$, i.e. $\widehat{S}^{r_l} \approx \frac{1}{m}\sum_{j=1}^{m}\left(\widehat{\mathbb{D}}^{r_l}(u_{j1}, u_{j2})\right)^2$.

2. R independent copies of $\widehat{\mathbb{D}}^{r_l}$, $\widehat{\mathbb{D}}^{r_l,(1)}, \ldots, \widehat{\mathbb{D}}^{r_l(R)}$ are generated, such that

$$\left(\widehat{\mathbb{D}}^{r_l}, \widehat{\mathbb{D}}^{r_l,(1)}, \ldots \widehat{\mathbb{D}}^{r_l,(R)}\right) \longmapsto \left(\mathbb{D}^{r_l}, \mathbb{D}^{r_l,(1)}, \ldots \mathbb{D}^{r_l,(R)}\right),$$

   where $\mathbb{D}^{r_l,(1)}, \ldots, \mathbb{D}^{r_l,(R)}$ are independent copies of $\mathbb{D}^{r_l}$. The process of obtaining these independent copies of $\widehat{\mathbb{D}}^{r_l}$ is described in Section A-3 of the Appendix at the end of this Thesis.

3. To calculate the copies of $\widehat{S}^{r_l}$ as $\widehat{S}^{r_l,(k)} = \frac{1}{m}\sum_{j=1}^{m}\widehat{\mathbb{D}}^{r_l,(k)}(u_{j1}, u_{j2})$ and to obtain the p-value of the statistics as:

$$\frac{1}{R}\sum_{s=1}^{R}\boldsymbol{I}(\widehat{S}^{r_l,(s)} \geq \widehat{S}^{r_l}).$$

4.3 Simulation Study

We summarise the results of our simulation study, we aim to evaluate the finite sample properties of our Beta transformed kernel estimator in (4.12), and compare it with the empirical copula, with the classical kernel estimator in (4.5) and with the Gaussian transformed kernel estimator in (4.7). We also obtain some results using boundary kernel, but the computational times are longer and in our simulation study we do not achieve better results than those obtained with classical kernel.

We show two types of results; in the former, the errors between the estimations and true copulas are compared and, in the latter, the differences between the extreme value copula tests obtained with the empirical copula and with the Beta transformed kernel estimator are analysed.

4.3.1 Analysing the errors of kernel estimators

To carry out the study, we simulate 500 samples of size $n = 50$ and $n = 500$ from different family and parameters of copulas that are indicated in the tables with the simulation results shown in this section. The alternative estimators are compared approximating the $MISE = \int_0^1\int_0^1\left[\hat{C}(u_1,u_2) - C(u_1,u_2)\right]^2 du_1 du_2$ using a grid uniformly spaced in 99×99 points on $(0,1)^2$. The Epanechnikov kernel is used in all cases. Furthermore, bandwidth b_n needs to be calculated, its value has an important impact on the results. Sometimes, the calculation of b_n requires long optimization processes based on leave-one out estimators (see Wand and Jones, 1995,

for a review on kernel density estimation). Alternatively, the rule-of-thumb similar to the proposal of Silverman (1986) can be used; however, to calculate the rule-of-thumb smoothing parameter we would need to use a parametric copula with a given parameter. A direct alternative, based on the product kernel estimator, consists of using the rule-of-thumb based on independent marginal reference distribution. The difficulty with $Uniform(0,1)$ marginal is that $\int \left[F_j''(t)\right]^2 dt = 0$, $j = 1,2$, and the rule-of-thumb smoothing parameter based on the proposal of Silverman (1986) can not be calculated. In this case, to use the standard normal bandwidth is an easy solution (see Buch-Larsen et al., 2005, for an example on kernel density estimation of $Uniform(0,1)$ transformed data). In our simulation study, two types of results are shown, on the one hand, the obtained with the rule-of-thumb bandwidth based on standard normal distribution for kernel and Gaussian transformed kernel estimator and based on $Beta(3,3)$ for Beta transformed kernel estimator. On the other hand, the obtained optimizing the approximate $MISE$ on a grid the values for b_n.

To facilitate the interpretation of the results, we calculate the quotient between the MISE obtained for each kernel estimator and the one obtained using the empirical copula. The reference values of MISE for the empirical copula are shown in Table 5 in Section A-1 of the Appendix at the end of this Thesis. Tables 4.1 and 4.2 contain, respectively, the quotients for the analysed elliptical and archimedean copulas. These results show how, using the adequate smoothing parameter, the analysed kernel estimators improve the empirical copula. Focusing on archimedean copulas in Table 4.2, it can be seen that, if the optimal smoothing parameter is used, in all cases the best results are obtained with the Beta transformed kernel estimator ($\widehat{C}^B$). With the optimal bandwidth, the Gaussian transformed kernel estimator ($\widehat{C}^G$) only slightly improves the classical kernel estimator ($\widehat{C}$) for Gumbel with dependence parameter equal to 3 and 4, i.e. for the most extreme value dependence copulas. In Table 4.1, for elliptical copulas, the results with optimal smoothing parameter are similar, $\widehat{C}^B$ is the best and $\widehat{C}^G$ improves $\widehat{C}$ when the data is generated by the most extreme value copulas, the Student's t with dependence parameter equal to 0.9.

In practice, we will not know what the optimal smoothing parameter is, and having estimators that allow this parameter to be obtained in a direct and simple way is essential. As shown in Tables 4.1 and 4.2, $\widehat{C}^B$ and $\widehat{C}^G$ have this characteristic; in both cases the results with the rule-of-thumb smoothing parameter are near the optimal results and the lowest MISEs are obtained with $\widehat{C}^B$.

Table 4.1: Quotient between the approximate MISE of kernel estimators of the copula (numerator) and approximate MISE of the empirical copula (denominator) for elliptical copulas (* indicates optimal bandwidth).

		$n = 50$					
Copula	Parameter	$\widehat{C}^B$	$\widehat{C}^{B*}$	$\widehat{C}^G$	$\widehat{C}^{G*}$	$\widehat{C}$	$\widehat{C}^*$
Gaussian	0.9	0.7727	0.6910	0.7861	0.7423	3.8676	0.7079
	0.5	0.6649	0.5879	0.6795	0.6993	2.7843	0.6108
	0.3	0.6394	0.5889	0.6553	0.6572	2.5860	0.6046
Student's t	0.9	0.8186	0.6516	0.8350	0.7120	3.8347	0.6626
(d.f.=1)	0.5	0.7211	0.6294	0.7368	0.6734	2.9403	0.6679
	0.3	0.7005	0.6793	0.7176	0.7254	2.9272	0.6467
Student's t	0.9	0.7907	0.6405	0.8055	0.6769	3.6564	0.6650
(d.f.=2)	0.5	0.7022	0.6338	0.7163	0.6569	2.8777	0.6432
	0.3	0.6592	0.6676	0.6746	0.7255	2.8678	0.6519
Student's t	0.9	0.7962	0.6671	0.8112	0.6782	3.8211	0.7376
(d.f.=3)	0.5	0.6912	0.6354	0.7083	0.6506	2.8528	0.5949
	0.3	0.6626	0.6243	0.6790	0.6636	2.6940	0.6084
		$n = 500$					
Copula	Parameter	$\widehat{C}^B$	$\widehat{C}^{B*}$	$\widehat{C}^G$	$\widehat{C}^{G*}$	$\widehat{C}$	$\widehat{C}^*$
Gaussian	0.9	0.9099	0.8032	0.9199	0.8347	6.3718	0.8162
	0.5	0.8445	0.7300	0.8532	0.7915	3.8651	0.7603
	0.3	0.8298	0.7316	0.8388	0.7905	3.4954	0.7757
Student's t	0.9	0.9232	0.8285	0.9337	0.8479	6.0640	0.8557
(d.f.=1)	0.5	0.8513	0.7903	0.8602	0.8267	4.1395	0.8244
	0.3	0.8361	0.7852	0.8454	0.8201	3.8260	0.8224
Student's t	0.9	0.9129	0.8110	0.9231	0.8371	6.1449	0.8448
(d.f.=2)	0.5	0.8445	0.7733	0.8532	0.8229	3.8651	0.8109
	0.3	0.8298	0.7710	0.8388	0.8172	3.4954	0.8123
Student's t	0.9	0.9289	0.8202	0.9389	0.8438	6.2509	0.8560
(d.f.=3)	0.5	0.8529	0.7772	0.8619	0.8286	3.8559	0.8145
	0.3	0.8384	0.7748	0.8479	0.8243	3.4502	0.8147

Table 4.2: Quotient between the approximate MISE of kernel estimators of the copula (numerator) and approximate MISE of the empirical copula (denominator) for archimedean copulas (* indicates optimal bandwidth).

		n=50					
Copula	Parameter	$\widehat{C}^B$	$\widehat{C}^{B*}$	$\widehat{C}^G$	$\widehat{C}^{G*}$	$\widehat{C}$	$\widehat{C}^*$
Frank	1	0.6469	0.1641	0.6620	0.5961	2.3603	0.5979
	2	0.6710	0.5689	0.6858	0.6412	2.5716	0.5896
	3	0.6784	0.5684	0.6925	0.6099	2.5097	0.5584
Clayton	1	0.6859	0.5703	0.7012	0.7032	2.4038	0.6370
	2	0.7645	0.5924	0.7803	0.7263	2.8428	0.5983
	3	0.7839	0.6616	0.7968	0.7124	3.1101	0.6493
Gumbel	2	0.6531	0.2448	0.6645	0.2707	2.3911	0.2469
	3	0.7739	0.5976	0.7862	0.6517	3.6354	0.5600
	4	0.7781	0.6332	0.7959	0.6510	3.8202	0.6145
		n=500					
Copula	Parameter	$\widehat{C}^B$	$\widehat{C}^{B*}$	$\widehat{C}^G$	$\widehat{C}^{G*}$	$\widehat{C}$	$\widehat{C}^*$
Frank	1	0.8117	0.7460	0.8216	0.8067	3.0013	0.7950
	2	0.8204	0.7417	0.8304	0.8096	3.1943	0.7854
	3	0.8298	0.7596	0.8401	0.8117	3.5081	0.7807
Clayton	1	0.8393	0.7561	0.8484	0.8253	3.3645	0.8131
	2	0.8640	0.7674	0.8724	0.8291	4.3028	0.8225
	3	0.8844	0.7811	0.8934	0.8303	5.0583	0.8290
Gumbel	2	0.8577	0.7698	0.8662	0.8180	4.8203	0.8009
	3	0.8996	0.7982	0.9102	0.8286	6.1713	0.8298
	4	0.9294	0.8148	0.9421	0.8357	6.9994	0.8444

4.3.2 Test for extreme value copula

We show the results of a reduced simulation study (to avoid long computing periods) that allows us to compare Type 1 and Type 2 errors of the extreme value copula test on smaller-sized samples, calculated from the empirical copula (proposed by Kojadinovic et al., 2011) and from the Beta transformed kernel estimator ($\widehat{C}^B$) proposed in this paper, where the optimal bandwidth is used. For this experiment, we use 100 samples of size $n = 50$. Both tests are implemented using a uniformly spaced 99×99 points on $(0,1)^2$ for calculating $\widehat{S}^r$ and with $K = 100$ estimated copies of $\mathbb{D}^{r_l,(k)}$, $k = 1, ..., 100$. The results for type 1 errors are shown in Table 4.3 for theoretical extreme value copulas with which H_0 is true. Table 4.4 shows

the type 2 errors for copulas with non-dependence in extreme. They indicate that using $\widehat{C}^B$ we reduce Type 1 error at the cost of increasing Type 2 error, although for $n = 50$ this error is already high for the test based on the empirical copula. In general, a larger-size sample is required for reducing Type 2 error.

The results in Table 4.3 imply that, if the data is generated by an extreme value copula, $\widehat{C}^B$ based test practically ensures a correct result. This is fundamental in the context of risk quantification, given that the consequences of not detecting extreme dependency could be more serious than those of not detecting the opposite, i.e. non-dependence in extreme.

Table 4.3: Error Type 1 calculated with different significant levels α.

	Empirical Copula							
	Gumbel		Student's t (d.f.=1)		Student's t (d.f.=2)		Student's t (d.f.=3)	
Parameter	2	4	0.9	0.3	0.9	0.3	0.9	0.3
$\alpha = 0.10$	0.08	0.02	0.25	0.38	0.59	0.86	0.70	0.56
$\alpha = 0.05$	0.01	0.00	0.16	0.22	0.39	0.71	0.48	0.36
$\alpha = 0.01$	0.00	0.00	0.03	0.07	0.14	0.39	0.23	0.13
	$\widehat{C}^B$							
	Gumbel		Student's t (d.f.=1)		Student's t (d.f.=2)		Student's t (d.f.=3)	
Parameter	2	4	0.9	0.3	0.9	0.3	0.9	0.3
$\alpha = 0.10$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\alpha = 0.05$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\alpha = 0.01$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.4: Error Type 2 calculated with different significant levels α.

	Empirical Copula					
	Gaussian		Frank		Clayton	
Parameter	0.9	0.3	1	3	1	3
$\alpha = 0.10$	0.50	0.18	0.34	0.32	0.70	0.69
$\alpha = 0.05$	0.61	0.37	0.54	0.53	0.80	0.87
$\alpha = 0.01$	0.81	0.63	0.73	0.79	0.98	0.95
	$\widehat{C}^B$					
	Gaussian		Frank		Clayton	
Parameter	0.9	0.3	1	3	1	3
$\alpha = 0.10$	0.67	0.50	0.6	0.62	0.80	0.81
$\alpha = 0.05$	0.85	0.77	0.80	0.79	0.95	0.98
$\alpha = 0.01$	0.92	0.90	0.82	0.78	1.00	1.00

4.4 Data Analysis

For illustrating the usefulness of our proposed estimator $\widehat{C}^B$, we analyse the dependence between the Spanish stock market index (IBEX35) and the stock market indexes of some neighbouring European countries, namely Germany (DAX) , France (CAC40), Italy (FTSE MIB), Portugal (PSI20) and United Kingdom (FTSE100) as well as the two principal stock market indexes of the USA (DOWJONES and S&P500) and the Hong Kong stock market index (HANG SENG) (see Hussain and Li, 2018, for an analysis of extreme dependence between markets).

Two types of results are shown:

1. The fit of non parametric copulas to estimate the probability that the observed losses of two stock market indexes together exceed some percentiles, i.e. we estimate the value of $1-C(q,q)$, $q=0.9,0.95,0.99,0.995$, with the analysed kernels estimators.

2. The test to analyse if the data is generated by an extreme value copula.

To carry out the analysis we use a database of the monthly losses of the stock market indexes from January 2000 to March 2021. These losses are calculated from the quotes of the analysed indexes that are public and can be downloaded, for example, from *Investing.com*. Throughout the period analysed, three major events influenced market performance leading to higher losses than in periods of stability: the Lehman Brothers crisis that began in September 2008, the referendum on Brexit on June 23, 2016, and the ongoing COVID-19 crisis which started in March 2020. The three events are considered systemic risks that affect all markets and, if this effect is simultaneous, the data should be generated by an extreme value copula. In Table 4.5, the main descriptive statistics of the losses in percentage are shown. Furthermore, normality tests and a positive skewness test are carried out and, in all cases, normality hypothesis is rejected and skewness greater than zero can not be rejected, i.e., in absolute value, positive losses are bigger that negative ones.

Table 4.5: Descriptive statistics of monthly losses in percentage: Means, Standard Deviation (STD), Minimum (Min), First Quantile (Q25), Median, Third Quantile (Q75), Maximum (Mas), Skewness and Kurtosis.

	Means	STD	Min	Q25	Median	Q75	Max	Skewness	Kurtosis
Spain	0.0565	2.5931	-9.7537	-1.3088	-0.2354	1.3759	10.9100	0.3025	2.0327
Germany	-0.1196	2.6469	-8.4139	-1.7341	-0.3275	1.1770	12.7390	0.8978	2.9100
France	0.0043	2.2675	-7.9611	-1.3871	-0.3142	1.2895	8.3501	0.5444	1.4149
Italy	0.1044	2.6915	-8.9727	-1.5668	-0.2481	1.5654	11.0363	0.4579	1.5547
Portugal	0.1564	2.3768	-7.2752	-1.2973	0.0391	1.4856	10.1398	0.5301	1.6257
UK	0.0080	1.7670	-5.0582	-1.1605	-0.3326	0.9555	6.4533	0.6996	1.1175
USA (DJ)	-0.1711	1.8644	-4.8586	-1.2329	-0.3472	0.7277	6.5807	0.6714	1.3456
USA (SP)	-0.1650	1.9120	-5.1864	-1.2883	-0.4313	0.7868	8.0621	0.7009	1.4097
Hong Kong	-0.0919	2.6021	-6.8459	-1.6894	-0.4514	1.2787	11.0508	0.5713	1.1150

The losses of the Spanish stock market index are plotted together with the indexes of the countries listed for comparison. In Figure 4.1, we compare Spain (in blue) with four countries that also currently belong to the European Union (in black) and, Figure 4.2, the comparison is made with the other countries (in black).

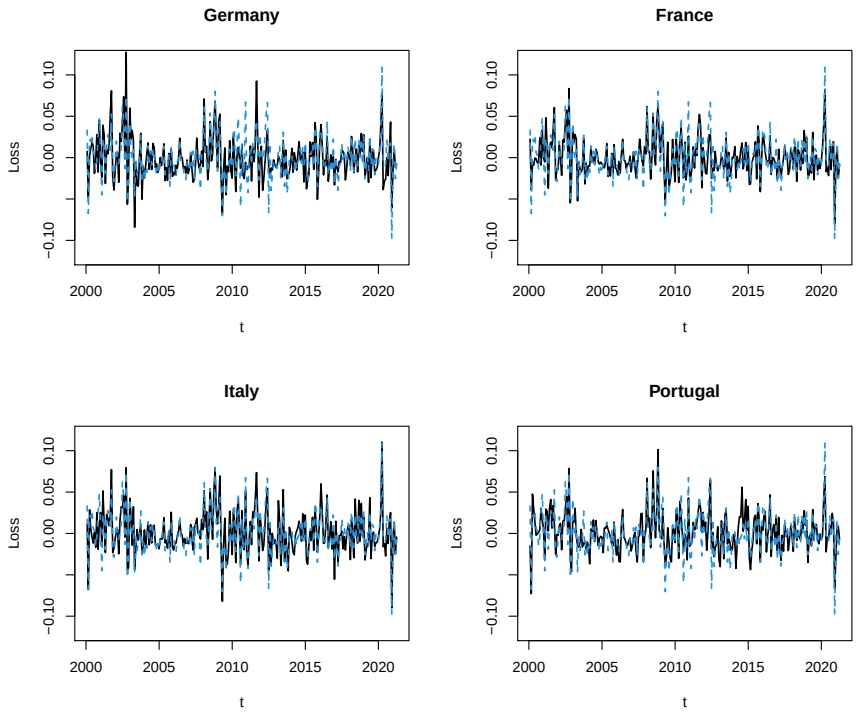


Figure 4.1: Losses of Spanish stock market index (dashed line in blue) and stock market indexes of four countries in the European Union (solid line in black).

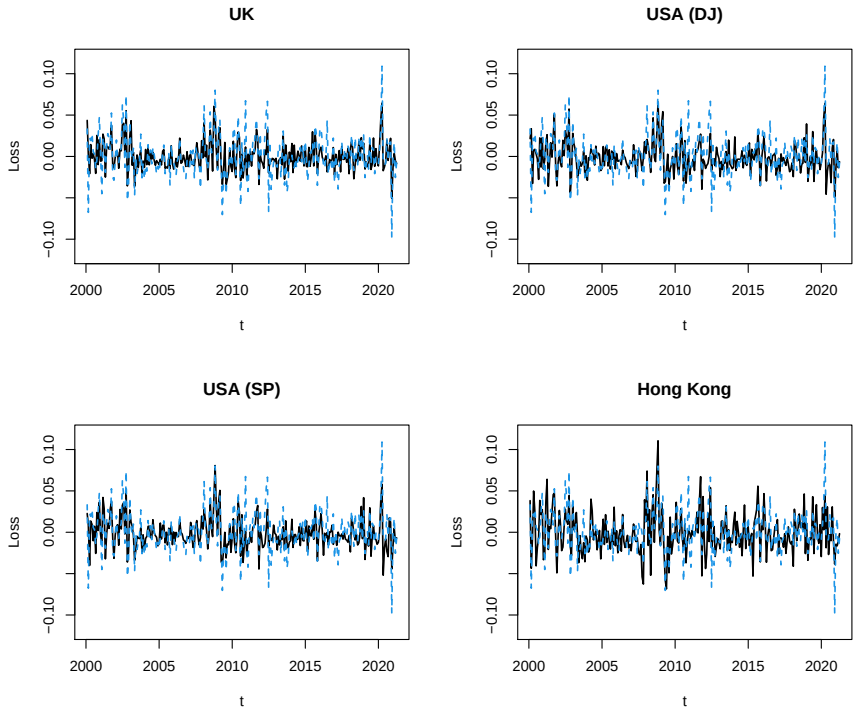


Figure 4.2: Losses of Spanish stock market index (dashed line in blue) and stock market indexes of four countries outside of the European Union (solid line in black).

To obtain the Beta transformed kernel estimator we need the data to be i.i.d. so, with this in mind, we analyse if the the monthly losses of the stock market indexes have some kind of time dependence on the mean or on the variance. The simple and partial autocorrelation functions of the series and the square series allows us to find the $ARMA(p,q)-GARCH(P,Q)$ model used to filter series and to get i.i.d. data (see, for example, Francq and Zakoïan, 2004). The filter models used are shown in the Table 4 of Section A-1 of Appendix at the end of this Thesis.

In Table 4.6, we show the results of $1-C(q,q)$ for $q=0.9,0.95,0.99,0.999$ estimated with $\widehat{C}^B$, i.e. the probability of jointly exceeding a given extreme quantile. The upper tail dependence can be approximated as $\frac{\widehat{C}^B(q,q)}{q}$. In Table 5 of Section A-1 of Appendix at the end of this Thesis, we show that the results obtained with the empirical copula and $\widehat{C}^G$ provide lower values than $\widehat{C}^B$. It should be noted that the empirical copula tends to underestimate the probability of the tail when extreme values exist. Furthermore, in the simulation study $\widehat{C}^B$ improves $\widehat{C}^G$ for all the compared copulas.

Table 4.6: Values of $1-\widehat{C}^B(q,q)$ for Spain and the countries analysed.

q	0.9	0.95	0.99	0.995
Germany	0.1533	0.0848	0.0246	0.0151
France	0.1500	0.0853	0.0244	0.0150
Italy	0.1497	0.0827	0.0234	0.0143
Portugal	0.1583	0.0858	0.0234	0.0145
UK	0.1585	0.0879	0.0253	0.0155
USA (DJ)	0.1583	0.0862	0.0242	0.0149
USA (SP)	0.1579	0.0880	0.0246	0.0149
Hong Kong	0.1662	0.0928	0.0255	0.0155

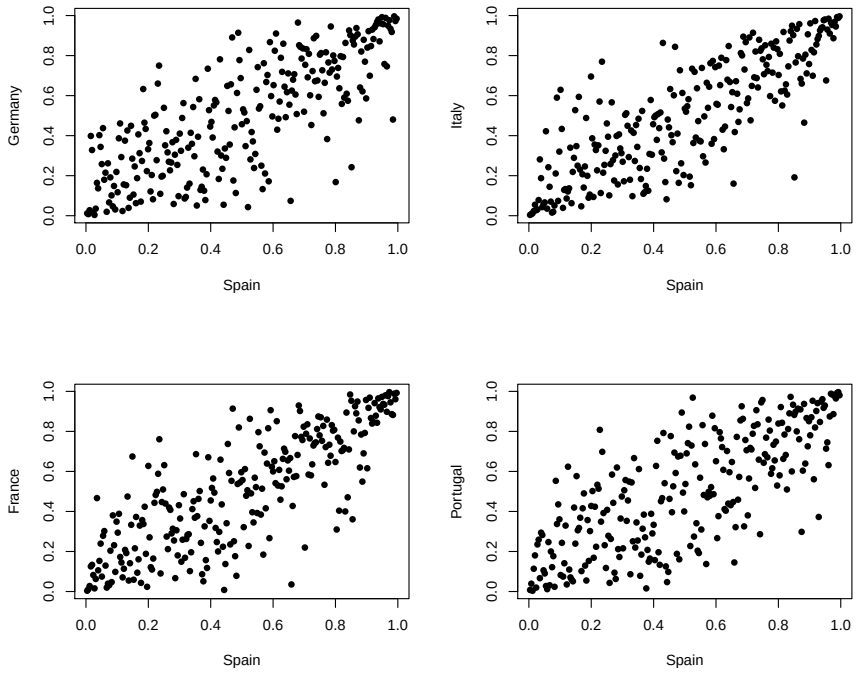


Figure 4.3: Pseudo-data for each bivariate copula between Spain and Germany, Italy, France and Portugal.

We obtain the results of the extreme value copula test of Kojadinovic et al. (2011) based on the empirical copula, and the same test based on the Beta transformed kernel estimator that is analysed in this paper, using the asymptotically optimal smoothing parameter $b_n = 3^{\frac{1}{3}} n^{-\frac{1}{3}}$ and a grid of 4 values around it. As expected, all the results indicate that all the analysed bivariate data have a dependence structure generated by an extreme value copula. This behaviour has been accentuated by the

COVID-19 crisis, which has led to greater losses and a systemic risk that lasts over time (see Ashraf, 2020; Zhang et al., 2020, for a review on effect of COVID-19 on markets returns and volatility). In Figures 4.3 and 4.4, pairs of pseudo-data are plotted; in all cases some accumulation of points is detected near the corner $[1, 1]$, which is an indicator of extreme value dependence.

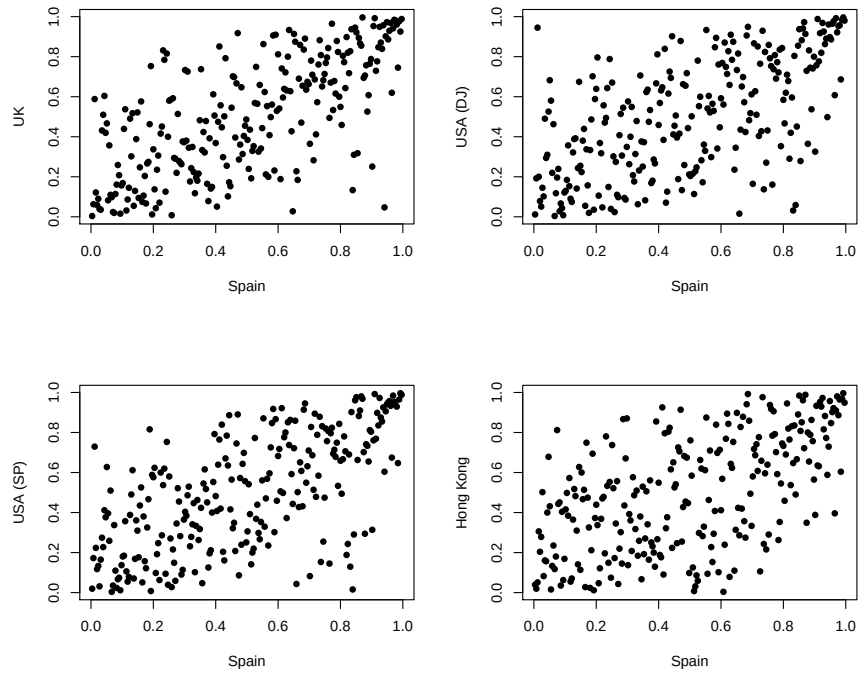


Figure 4.4: Pseudo-data for each bivariate copula between Spain and UK, USA (DJ and SP) and Hong Kong.

4.5 Conclusions

A new kernel estimator of a copula based on a transformation is analysed. The asymptotic theoretical properties that allow it to be used for inference are proved. A simulation study shows that the proposed estimator improves the alternative estimator in the most common copulas.

The new estimator allows us to reduce the Type 1 error associated with the extreme value copula test, while the Type 2 error increases slightly. A future line of work would be to investigate how to reduce the Type 2 error with a small sample size of the tests based on the $max - stable$ property.

The financial data analysis shows that the new estimator is useful for the risk analysis linked to the upper tail dependence between stock markets.

General Conclusions

In this thesis we addressed three very important issues related to systemic risk and linkages between stock markets. Firstly, we analysed alternative distances and similarities between the stock markets using alternative static criteria (distances in kilometres and hours and similarity of overlapping operating hours between financial markets) and dynamic criteria (distances between the comments of the economic and financial agents). Secondly, the spatial dependence between financial markets was analysed by using global and local Moran's I statistics; both are tools widely used in the field of spatial econometrics. In our case, we used them in a novel way, with the aim of studying the relation between the periods with significant spatial dependence between financial markets and those considered as periods of systemic risk that are related to the main most recent financial crises: the Lehman Brothers bankruptcy, the sub-prime mortgage crisis, the European debt crisis, Brexit and the COVID-19 pandemic, the latter also affecting the financial markets. Finally, a new Kernel estimator of a copula is proposed to analyse the upper tail dependence between market losses.

Throughout the thesis, the analyses were carried out using different financial variables: log-returns, losses, volatility and Value-at-Risk, all of them calculated from the stock market indices associated with the stock markets that were included in the studies. Using different financial variables has enabled us to achieve different results, depending on the type of variable used. For example, using log-return or losses the greatest spatial dependence occurs during the sub-prime and Euro debt periods; however, using Value-at-Risk, the same result is seen for sub-prime and Brexit periods and not for Euro debt.

Specifically, in Chapter 1, a comparison was made between static distance criteria between financial markets. We compared the classical distances in kilometres and in hours with a criterion of similarity based on common opening hours between markets, which is interpreted as a measure of synchronisation between financial markets. We concluded that this synchronisation measure shows a greater reflection of the stock market linkages, given that neighbourhoods between markets are detected with more visibility.

Chapter 1 focused on the Lehman Brothers bankruptcy. By using the common opening hours similarity criterion, analysis of the results obtained with global Moran's

I statistic showed a greater spatial dependency between the log-returns of the stock indices in the European financial markets, specially in the financial crisis period. These results can help financial actors, such as banks, insurance companies or investors, to create their portfolios taking into account the effect of systemic risk on stock market linkages.

The local spatial dependence analysis that was carried out in Chapter 1 showed less obvious results. showed less obvious results. However, it did show that the European markets were those with the most intensive links during the financial crisis caused by the *Lehman Brothers* bankruptcy. Most markets in this group showed greater dependence intensity during the crisis.

In Chapter 2 we presented the spatial dependence results using a dynamic distance. In this case, the spatial dependence between the losses of the stock markets indices was analysed. With this aim in mind, we used dynamic weight matrices calculated monthly for estimating the monthly global Moran's I statistic. The monthly weight matrices are obtained from the distance matrices between financial stock markets based on the differences between the values of the monthly Google Trend Uncertainty Indices (GTUI) by country. Focusing on the GTUI, its greater values are given in periods of financial crises. In concordance with these greater values, the global Moran's I statistic indicates greater spatial dependence between markets with similar GTUI and this occurs during periods of crisis where the GTUI has high values, especially during the period of the sub-prime mortgage crisis and the European debt crisis. We concluded that the distances based on GTUI better reflect the behaviour of the stock markets and the possible approximations between them, given that the GTUI is an index that takes into account the interests of financial actors.

Chapter 3 presents a similar analysis to Chapter 2 but, instead of between losses, it focuses on spatial dependence between alternative monthly risk measures as the variance (volatility) and the Value-at-Risk calculated on the losses. In the presence of extreme values in the risk measures, the distribution in finite sample of the global and local Moran's I statistics was analysed and showed that inference based on asymptotic normality does not work. Alternative bootstrap inference improved the Type 1 and Type 2 errors of the tests on global and local spatial dependence.

In the analysis of the impact of the great crises between 2004 and 2021 on the stock markets, a greater presence of systemic risk was observed during the the sub-prime mortgage crisis, where the risk was at its most extreme. However, this systemic risk, linked to a greater spatial dependence between market risks, was not so evident

during the European debt crisis and the COVID-19 pandemic. In the case of Brexit, as was expected, before this period there were no signs of contagion between stock markets but significant signs were observed during the rollout of Brexit. In general, the analysis gives us a useful tool, not only to detect possible crises, but also to detect which markets may be more sensitive to these possible crises.

Finally, in Chapter 4 we analysed the upper tail dependence between the losses of the Spanish stock market index (IBEX35) and the same losses of five benchmark stock market indices in European continent: Germany (DAX), France (CAC40), Italy (FTSE MIB), Portugal (PSI20) and United Kingdom (FTSE100), as well as the two principal stock market indices of the USA (DOWJONES and S&P500) and the Hong Kong stock market index (HANG SENG). Between the eight benchmark indices we included four neighbouring European stock indices and four more distant financial markets. To carry out the study we proposed a new kernel estimator of copula based on the inverse Beta cumulative distribution function transformation and deduced a new statistic for testing the positive upper tail dependence. In this way, we observed how the proposed test statistic based on the transformed kernel estimator reduces the Type 1 error. We conclude that all the bivariate losses between those of the Spanish index and each of the eight analysed indices are generated by an extreme value copula, i.e., there is significant upper tail dependence in all cases and the larger dependences are given between Spain with Hong Kong followed by UK. The significant positive upper tail dependence implies that in loss periods we observe linkages between the Spanish index and the market indices compared with it.

The thesis presents an extensive study of the dependence between financial markets from different points of view. This has given us a variety of results that can be useful for making investment decisions based on different financial indicators, whether they are returns, losses or risks of losses in the financial markets.

A-1 Complementary Tables

Table 1: Descriptive statistics of the monthly log-returns analysed during the two sub-periods.

			B.C.			D. C.		
Country	**Label**	**Index**	**Min**	**Media**	**Max**	**Min**	**Media**	**Max**
Argentina	AR	Merval	-0.1184	0.0044	0.1723	-0.1990	0.0099	0.1066
Australia	AU	S&P-ASX200	-0.0476	0.0015	0.0318	-0.0469	0.0008	0.0301
Austria	AT	ATX	-0.1158	0.0035	0.0540	-0.1416	-0.0007	0.0588
Belgium	BE	BEL20	-0.0798	-0.0009	0.0587	-0.1046	0.0016	0.0481
Brazil	BR	BOVESPA	-0.0818	0.0044	0.0716	-0.1238	-0.0007	0.0628
Canada	CA	S&PTSX	-0.0688	0.0014	0.0422	-0.0806	0.0005	0.0461
Chile	CL	IPSA	-0.0633	0.0037	0.0620	-0.0437	0.0015	0.0648
Denmark	DK	OMX	-0.0820	0.0013	0.0526	-0.0904	0.0053	0.0804
Egypt	EG	EGX 30	-0.1001	0.0074	0.1467	-0.1505	0.0000	0.1083
Finland	FI	OMXH25	-0.0938	-0.0020	0.0782	-0.0585	0.0026	0.0986
France	FR	CAC40	-0.0835	-0.0016	0.0547	-0.0631	0.0007	0.0514
Germany	DE	DAX	-0.1274	-0.0007	0.0841	-0.0926	0.0031	0.0673
Greece	GR	ATH	-0.0937	-0.0027	0.0858	-0.1419	-0.0075	0.0862
Hong Kong	HK	HANG SENG	-0.0740	0.0003	0.0626	-0.1105	0.0010	0.0685
Hungary	HU	BUX	-0.0735	0.0032	0.0726	-0.1452	0.0012	0.0654
Iceland	IS	ICEX	-0.0828	0.0031	0.0574	-0.5453	-0.0044	0.0714
India	IN	BSE Sensex 30	-0.0862	0.0039	0.0635	-0.1186	0.0035	0.1081
Indonesia	ID	IDX	-0.0726	0.0041	0.0615	-0.1638	0.0046	0.0797
Ireland	IE	ISEQ 20	-0.1024	-0.0014	0.0429	-0.0768	0.0032	0.0774
Israel	IL	TA100	-0.0843	0.0021	0.0609	-0.0809	0.0025	0.0602
Italy	IT	FTSE MIB	-0.0795	-0.0021	0.0679	-0.0773	-0.0009	0.0821
Japan	JP	NIKKEI 225	-0.0648	-0.0022	0.0437	-0.1182	0.0026	0.0525
Malaysia	MY	KLCI	-0.0479	0.0009	0.0551	-0.0717	0.0025	0.0552
Mexico	MX	IPC	-0.0674	0.0052	0.0666	-0.0854	0.0027	0.0476
Netherland	NL	AEX	-0.0982	-0.0029	0.0633	-0.0928	0.0014	0.0460
New Zealand	NZ	S&PNZX10	-0.0559	0.0032	0.0563	-0.0413	0.0036	0.0329
Norway	NO	OSEAX	-0.1188	0.0031	0.0499	-0.1052	0.0027	0.0609
Pakistan	PK	KARACHI100	-0.0957	0.0078	0.1047	-0.1949	0.0064	0.0784
Peru	PE	IGBVL	-0.0750	0.0075	0.0815	-0.2026	-0.0007	0.1413
Philippines	PH	PSEI	-0.0834	0.0008	0.0667	-0.1196	0.0050	0.0606
Poland	PL	WIG20	-0.0752	0.0012	0.0885	-0.1159	-0.0012	0.0754
Portugal	PT	PSI20	-0.0786	-0.0017	0.0728	-0.1014	-0.0021	0.0437
Rep. Czech	CZ	PX	-0.0933	0.0037	0.0672	-0.1374	-0.0012	0.0743
Russia	RU	RTSI	-0.1330	0.0080	0.1324	-0.1951	-0.0024	0.1159
Singapore	SG	STI	-0.0810	-0.0003	0.0492	-0.1188	0.0008	0.0838
Slovakia	SK	SAX	-0.0427	0.0073	0.1263	-0.0889	-0.0022	0.0549
South Korea	KR	KOSPI200	-0.0763	0.0014	0.0880	-0.1143	0.0015	0.0551
Spain	SP	IBEX35	-0.0745	-0.0002	0.0675	-0.0800	-0.0007	0.0702
Sweden	SE	OMXS30	-0.0719	-0.0018	0.0698	-0.0802	0.0032	0.0681
Swiss	CH	SMI	-0.0610	-0.0005	0.0461	-0.0522	0.0014	0.0419
Taiwan	TW	TWII	-0.0934	-0.0016	0.0974	-0.0698	0.0019	0.0607
Thailand	TH	SET	-0.1058	0.0009	0.0921	-0.1560	0.0042	0.0568
Turkey	TR	BIST100	-0.1897	0.0036	0.1880	-0.1124	0.0034	0.0894
UK	UK	FTSE100	-0.0606	-0.0014	0.0361	-0.0492	0.0012	0.0352
USA	US-DJ	DOWJONES	-0.0573	-0.0002	0.0438	-0.0658	0.0024	0.0396
USA	US-SP	S&P 500	-0.0506	-0.0010	0.0401	-0.0806	0.0028	0.0444

Note: B.C. (before the crisis) and D.C. (during the crisis).

Table 2: Normality tests and shape measures of the distribution of log-returns.

	B.C. *Shapiro-Wilk*	**p-value**	**Kurtosis**	**Skewness**	**D.C. *Shapiro-Wilk***	**p-value**	**Kurtosis**	**Skewness**
DE	0.9576	0.0020	3.2713	-0.9971	0.9401	0.0006	2.0824	-0.7381
AR	0.9499	0.0006	2.5180	0.5705	0.9519	0.0027	2.7224	-0.8454
AU	0.9549	0.0013	0.8223	-0.8365	0.9494	0.0019	-0.2953	-0.4856
AT	0.9507	0.0007	5.2294	-1.4799	0.9375	0.0004	4.9409	-1.2813
BE	0.9509	0.0007	3.0205	-1.4216	0.9487	0.0017	5.6482	-1.5042
BR	0.9538	0.0011	-0.4236	-0.3198	0.9438	0.0009	3.3814	-0.7920
CA	0.9439	0.0002	1.7018	-1.0347	0.9492	0.0019	5.0198	-1.1947
CL	0.9421	0.0002	0.8446	-0.2952	0.9551	0.0043	0.6040	0.3630
KR	0.9555	0.0014	-0.2170	-0.1807	0.9392	0.0005	6.8118	-1.2957
DK	0.9476	0.0004	1.1702	-0.8278	0.9525	0.0029	3.3909	-0.6845
US-DJ	0.9458	0.0003	0.9779	-0.4476	0.9561	0.0049	1.6993	-0.8725
US.SP	0.9478	0.0004	0.3079	-0.4328	0.9557	0.0047	2.4913	-1.0242
EG	0.9499	0.0006	0.8486	0.2173	0.9531	0.0032	1.6098	-0.5648
SK	0.9433	0.0002	5.0367	1.5182	0.9553	0.0044	3.3027	-0.8743
SP	0.9586	0.0024	1.0510	-0.3983	0.9382	0.0004	0.5774	-0.1924
PH	0.9370	0.0001	-0.0365	-0.2064	0.9608	0.0099	7.5716	-1.5079
FI	0.9453	0.0003	1.1449	-0.3646	0.9427	0.0008	1.7190	0.2174
FR	0.9505	0.0006	1.5357	-0.8102	0.9533	0.0033	-0.2724	-0.3429
GR	0.9647	0.0068	0.5327	-0.4436	0.9256	0.0001	0.2355	-0.4019
NL	0.9474	0.0004	2.1926	-1.0827	0.9468	0.0013	2.4998	-0.8380
HK	0.9434	0.0002	0.0438	-0.4350	0.9411	0.0006	2.6283	-0.7180
HU	0.9536	0.0011	-0.0928	-0.3664	0.9480	0.0016	4.2610	-1.0179
IN	0.9508	0.0007	-0.0303	-0.6241	0.9481	0.0016	5.1530	-0.2549
ID	0.9577	0.0020	-0.4293	-0.3256	0.9417	0.0007	12.9596	-2.1595
IE	0.9297	0.0000	2.0141	-1.1876	0.9707	0.0456	1.9123	-0.6827
IS	0.9600	0.0030	0.9163	-0.6570	0.9390	0.0005	54.4042	-6.8424
IL	0.9523	0.0008	0.7061	-0.4537	0.9496	0.0020	2.1036	-0.4647
IT	0.9623	0.0044	1.6352	-0.6329	0.9283	0.0001	0.1563	-0.2597
JP	0.9528	0.0009	-0.4607	-0.3617	0.9375	0.0004	3.5214	-1.1571
MY	0.9389	0.0001	-0.0373	-0.1997	0.9459	0.0012	7.3785	-0.9563
MX	0.9261	0.0000	-0.0513	-0.3817	0.9634	0.0146	3.2187	-0.7942
NO	0.9329	0.0001	2.8666	-1.2418	0.9659	0.0215	5.2470	-1.1854
NZ	0.9395	0.0001	1.2163	-0.5159	0.9549	0.0042	1.9973	-0.8922
PK	0.9360	0.0001	0.4982	-0.0751	0.9579	0.0065	18.3701	-3.0105
PE	0.9290	0.0000	0.2372	-0.1851	0.9595	0.0081	8.0159	-0.5263
PL	0.9554	0.0014	0.1819	0.0597	0.9350	0.0003	3.3268	-0.6374
PT	0.9557	0.0015	1.7170	-0.4837	0.9437	0.0009	1.4706	-0.7192
UK	0.9542	0.0012	1.2166	-0.9628	0.9385	0.0004	-0.1230	-0.3474
CZ	0.9555	0.0014	1.6585	-0.8533	0.9461	0.0012	6.2766	-1.0760
RU	0.9570	0.0018	0.6968	-0.5025	0.9403	0.0006	2.8577	-0.6337
SG	0.9453	0.0003	1.5370	-1.0137	0.9449	0.0010	7.1591	-0.8638
SE	0.9485	0.0005	0.4866	-0.3267	0.9543	0.0038	2.9475	-0.5981
CH	0.9450	0.0003	0.6987	-0.8130	0.9515	0.0026	0.5769	-0.4330
TH	0.9424	0.0002	1.2456	-0.4204	0.9533	0.0033	11.7078	-2.2853
TW	0.9553	0.0014	0.9149	-0.0168	0.9483	0.0017	0.8398	-0.0692
TR	0.9476	0.0004	1.5144	-0.1442	0.9485	0.0017	0.9018	-0.2765

Note: B.C. (before the crisis) and D.C. (during the crisis).

Table 3: Descriptive statistics of stock market losses.

Country	Label	Index	Mean	STD	Min	Median	Max	Skew	Kurtosis
Argentina	AR	MERVAL	-0.80%	4.66%	-10.66%	-0.58%	23.28%	0.952	3.988
Australia	AU	S&P-ASX200	-0.15%	1.76%	-4.12%	-0.46%	10.34%	1.467	5.467
Austria	AT	ATX	-0.15%	2.96%	-9.44%	-0.47%	14.38%	1.420	5.567
Belgium	BE	BEL20	-0.11%	2.14%	-8.10%	-0.47%	10.46%	1.155	4.372
Brazil	BR	BOVESPA	-0.34%	2.95%	-6.81%	-0.35%	15.43%	1.009	4.153
Canada	CA	S&PTSX	-0.17%	1.70%	-4.61%	-0.38%	8.48%	1.480	6.016
Chile	CL	IPSA	-0.24%	2.05%	-6.48%	-0.18%	7.27%	0.043	0.727
Czech Rep	CZ	PX	-0.10%	2.62%	-7.43%	-0.28%	13.74%	1.276	5.344
Denmark	DK	OMX	-0.37%	2.09%	-8.04%	-0.48%	9.04%	0.819	3.385
Egypt	EG	EGX 30	-0.48%	3.96%	-13.54%	-0.47%	15.05%	0.281	1.842
Finland	FI	OMXH25	-0.24%	2.27%	-9.86%	-0.47%	8.10%	0.413	2.400
France	FR	CAC40	-0.10%	2.10%	-7.96%	-0.38%	8.20%	0.535	1.768
Germany	DE	DAX	-0.27%	2.27%	-6.73%	-0.57%	9.25%	0.822	2.233
Greece	GR	ATH	0.21%	3.91%	-11.19%	-0.21%	14.19%	0.661	1.242
Hong Kong	HK	HANG SENG	-0.18%	2.54%	-6.85%	-0.52%	11.05%	0.687	1.855
Hungary	HU	BUX	-0.33%	2.83%	-7.97%	-0.52%	14.52%	0.871	3.816
Iceland	IC	ICEX	0.00%	4.64%	-7.14%	-0.43%	54.52%	8.187	90.454
India	IN	BSE SENSEX 30	-0.45%	2.82%	-10.81%	-0.50%	12.92%	0.957	4.191
Indonecia	ID	IDX	-0.47%	2.53%	-7.97%	-0.73%	16.38%	1.655	9.189
Ireland	IE	ISEQ 20	-0.09%	2.48%	-7.74%	-0.37%	10.24%	1.099	2.886
Israel	IL	TA35	-0.24%	2.04%	-4.74%	-0.47%	8.72%	0.985	2.479
Italy	IT	FTSE MIB	0.03%	2.63%	-8.97%	-0.32%	11.04%	0.500	1.908
Japan	JP	NIKKEI 225	-0.21%	2.39%	-6.09%	-0.48%	11.82%	0.902	2.505
Malaysia	MY	KLCI	-0.14%	1.51%	-5.52%	-0.28%	7.17%	0.513	2.941
Mexico	MX	IPC	-0.35%	2.09%	-5.38%	-0.39%	8.54%	0.718	2.012
Netherlands	NL	AEX	-0.14%	2.12%	-5.50%	-0.49%	9.54%	1.192	3.654
New Zeland	NZ	S&PNZX10	-0.34%	1.48%	-3.64%	-0.46%	6.05%	1.088	2.725
Norway	NO	OSEAX	-0.37%	2.42%	-6.09%	-0.60%	11.88%	1.432	5.196
Pakistan	PK	KARACHI100	-0.49%	3.09%	-8.78%	-0.79%	19.49%	1.707	8.665
Peru	PE	IGBVL	-0.47%	3.57%	-14.13%	-0.43%	20.26%	0.597	6.057
Philippines	PH	PSEI	-0.33%	2.33%	-6.06%	-0.57%	11.96%	1.192	4.825
Poland	PL	WIG20	-0.05%	2.64%	-8.18%	-0.21%	11.59%	0.455	1.706
Portugal	PT	PSI20	0.07%	2.32%	-6.71%	-0.14%	10.14%	0.714	1.893
Russia	RU	RTSI	-0.20%	4.08%	-11.59%	-0.41%	19.51%	0.815	2.730
Singapore	SG	STI	-0.11%	2.18%	-8.38%	-0.39%	11.88%	1.019	5.996
Slovakia	SK	SAX	-0.15%	2.16%	-12.63%	-0.22%	8.89%	-0.685	6.303
South Korea	KR	KOSPI200	-0.28%	2.29%	-5.80%	-0.39%	11.43%	0.743	3.155
Spain	SP	IBEX35	-0.02%	2.49%	-9.75%	-0.31%	10.91%	0.395	2.892
Sweden	SE	OMXS30	-0.25%	2.00%	-6.81%	-0.37%	8.02%	0.817	2.266
Swiss	CH	SMI	-0.14%	1.54%	-4.19%	-0.38%	5.22%	0.536	0.560
Taiwan	TW	TWII	-0.21%	2.24%	-6.07%	-0.42%	9.06%	0.689	1.876
Thailand	TH	SET	-0.14%	2.50%	-7.13%	-0.41%	15.60%	1.384	7.275
Turkey	TR	BIST100	-0.44%	3.31%	-8.94%	-0.75%	11.42%	0.417	0.500
UK	UK	FTSE100	-0.08%	1.68%	-5.06%	-0.35%	6.45%	0.730	1.604
USA	US	DOWJONES	-0.23%	1.76%	-4.86%	-0.36%	6.58%	0.806	2.064
USA	US	S&P 500	-0.26%	1.82%	-5.19%	-0.52%	8.06%	0.892	2.487

Table 4: Fitted ARMA-GARCH models to the monthly losses series.

Country	Label	Index	Model
Argentina	AR	MERVAL	*ARMA(0, 0) - GARCH(1. 1)*
Australia	AU	S & P-ASX200	*ARMA(0, 0) - GARCH(1, 1)*
Austria	AT	ATX	*ARMA(1, 0) - GARCH(0, 0)*
Belgium	BE	BEL20	*ARMA(1, 0) - GARCH(0, 0)*
Brazil	BR	BOVESPA	*ARMA(0, 0) - GARCH(0, 0)*
Canada	CA	S&PTSX	*ARMA(0, 0) - GARCH(1, 1)*
Chile	CL	IPSA	*ARMA(0, 0) - GARCH(1, 1)*
Czech Rep	CZ	PX	*ARMA(1, 0) - GARCH(0, 0)*
Denmark	DK	OMX	*ARMA(1, 0) - GARCH(0, 0)*
Egypt	EG	EGX30	*ARMA(1, 0) - GARCH(0, 0)*
Finland	FI	OMXH25	*ARMA(1, 0) - GARCH(0, 0)*
France	FR	CAC40	*ARMA(0, 0) - GARCH(1, 1)*
Germany	DE	DAX	*ARMA(0, 0) - GARCH(0, 0)*
Greece	GR	ATH	*ARMA(1, 0) - GARCH(0, 0)*
Hong Kong	HK	HANG SENG	*ARMA(0, 0) - GARCH(1, 1)*
Hungary	HU	BUX	*ARMA(1, 0) - GARCH(0, 0)*
Iceland	IS	ICEX	*ARMA(1, 0) - GARCH(0, 0)*
India	IN	BSE SENSEX30	*ARMA(0, 0) - GARCH(1, 1)*
Indonecia	ID	IDX	*ARMA(1, 0) - GARCH(0, 0)*
Ireland	IE	ISEQ20	*ARMA(1, 0) - GARCH(0, 0)*
Israel	IL	TA35	*ARMA(0, 0) - GARCH(1, 1)*
Italy	IT	FTSE MIB	*ARMA(0, 0) - GARCH(0, 0)*
Japan	JP	NIKKEI225	*ARMA(0, 0) - GARCH(1, 1)*
Malaysia	MY	KLCI	*ARMA(0, 0) - GARCH(0, 0)*
Mexico	MX	IPC	*ARMA(0, 0) - GARCH(0, 0)*
Netherlands	NL	AEX	*ARMA(0, 0) - GARCH(1, 1)*
New Zeland	NZ	S&PNZX10	*ARMA(0, 0) - GARCH(0, 0)*
Norway	NO	OSEAX	*ARMA(1, 0) - GARCH(0, 0)*
Pakistan	PK	KARACHI100	*ARMA(0, 0) - GARCH(0, 0)*
Peru	PE	IGBVL	*ARMA(1, 0) - GARCH(0, 0)*
Philippines	PH	PSEI	*ARMA(0, 0) - GARCH(0, 0)*
Poland	PL	WIG20	*ARMA(0, 0) - GARCH(1, 1)*
Portugal	PT	PSI20	*ARMA(1, 0) - GARCH(0, 0)*
Russia	RU	RTSI	*ARMA(1, 0) - GARCH(0, 0)*
Singapore	SG	STI	*ARMA(0, 0) - GARCH(0, 0)*
Slovakia	SK	SAX	*ARMA(1, 0) - GARCH(0, 0)*
South Korea	KR	KOSPI200	*ARMA(0, 0) - GARCH(0. 0)*
Spain	SP	IBEX35	*ARMA(0, 0) - GARCH(0, 0)*
Sweden	SE	OMXS30	*ARMA(0, 0) - GARCH(0, 0)*
Swiss	CH	SMI	*ARMA(0, 0) - GARCH(1, 1)*
Taiwan	TW	TWII	*ARMA(1, 0) - GARCH(0, 0)*
Thailand	TH	SET	*ARMA(1, 0) - GARCH(0, 0)*
Turkey	TR	BIST100	*ARMA(0, 0) - GARCH(0, 0)*
UK	UK	FTSE100	*ARMA(0, 0) - GARCH(1, 1)*
USA	US	DOWJONES	*ARMA(0, 0) - GARCH(1, 1)*
USA	US	S&P 500	*ARMA(0, 0) - GARCH(1, 1)*

Results of simulation study of Chapter 4. Table 5 shows the MISE estimated with the empirical copula.

Table 5: Approximate MISE for the empirical copula $\times 1,000$ (d.f. indicates degree of freedom).

Copula	Gaussian			Student's t (d.f.=1)			Student's t (d.f.=2)			Student's t (d.f.=3)		
Parameter	0.9	0.5	0.3	0.9	0.5	0.3	0.9	0.5	0.3	0.9	0.5	0.3
n=50	3.2473	3.0454	2.9277	3.3064	3.1751	2.9481	3.4585	3.1341	2.8430	3.3167	3.0923	2.9834
n=500	0.3152	0.3104	0.2984	0.3356	0.3101	0.2967	0.3272	0.3104	0.2984	0.3276	0.3098	0.2971
Copula	Frank			Clayton			Gumbel					
Parameter	1	2	3	1	2	3	2	3	4			
n=50	3.1141	3.0819	3.3664	3.2126	3.3294	3.4487	7.8489	3.4600	3.4147			
n=500	0.2873	0.2966	0.3040	0.3089	0.3201	0.3238	0.3081	0.3058	0.3029			

Table 6: Values of $1 - C(q, q)$ for Spain and th countriese analysed, estimated with the empirical copula and Gaussian transformed kernel estimator.

	Empirical Copula				$\widehat{C}^G$			
q	0.9	0.95	0.99	0.995	0.9	0.95	0.99	0.995
Germany	0.1294	0.0706	0.0157	0.0078	0.1490	0.0794	0.0176	0.0081
France	0.1339	0.0709	0.0118	0.0079	0.1469	0.0807	0.0171	0.0077
Italy	0.1294	0.0667	0.0118	0.0067	0.1464	0.0774	0.0150	0.0039
Portugal	0.1417	0.0669	0.0118	0.0079	0.1566	0.0797	0.0162	0.0076
UK	0.1378	0.0709	0.0157	0.0079	0.1549	0.0820	0.0184	0.0082
USA (DJ)	0.1417	0.0630	0.0118	0.0079	0.1546	0.0807	0.0169	0.0078
USA (SP)	0.1378	0.0709	0.0118	0.0079	0.1538	0.0836	0.0165	0.0075
Hong Kong	0.1378	0.0709	0.0118	0.0079	0.1538	0.0836	0.0179	0.0075

A-2 Complementary Figures

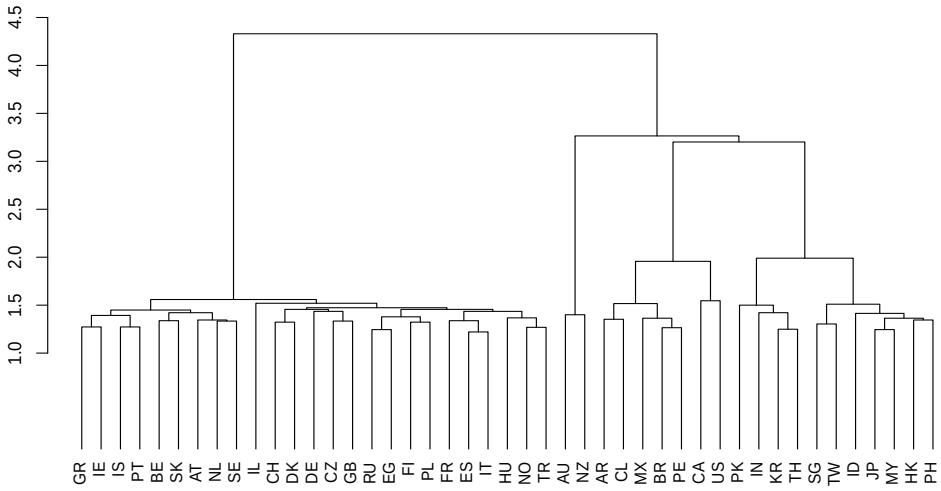


Figure 1: Dendrogram obtained from the weight matrix obtained from the criterion of kilometres distance between markets.

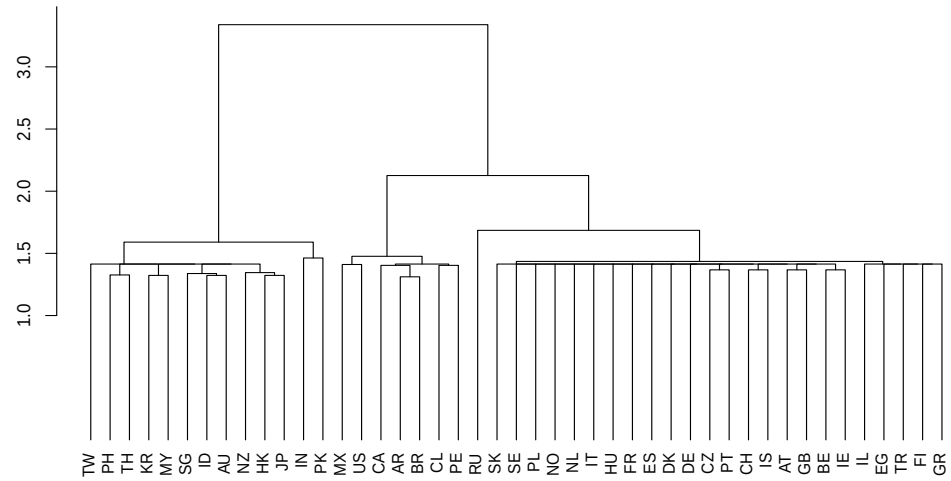


Figure 2: Dendrogram obtained from the weight matrix obtained from the criterion of hours distance between markets.

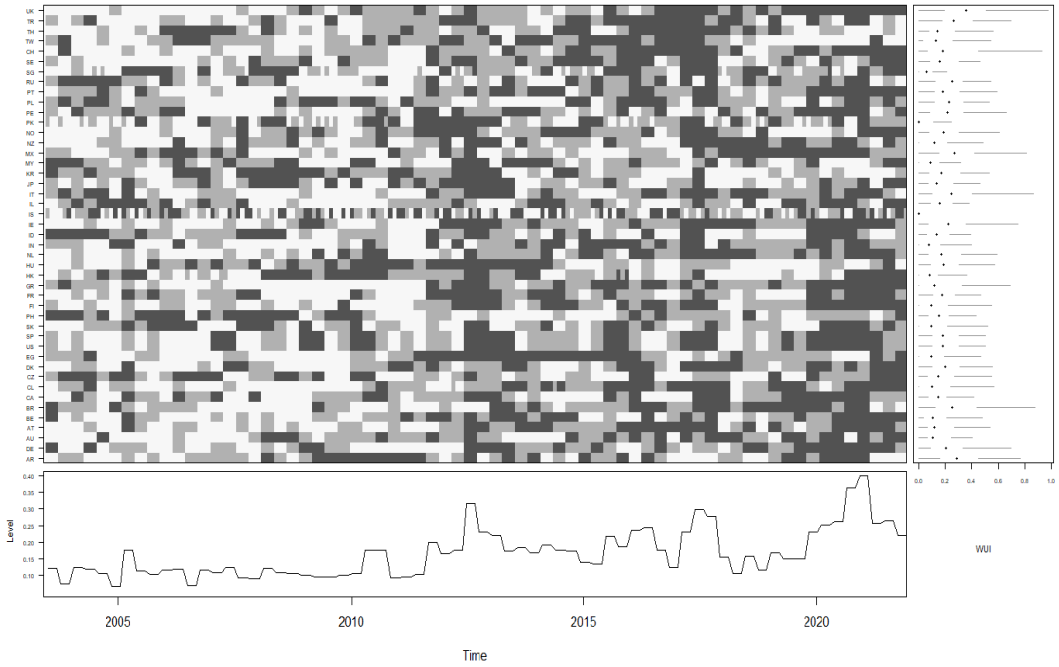


Figure 3: Values of the World Uncertainty Index (at top) - the darker the shading, the higher the value - and plot of the mean index (at bottom). The box-plots of the WUI for each country are shown on the right.

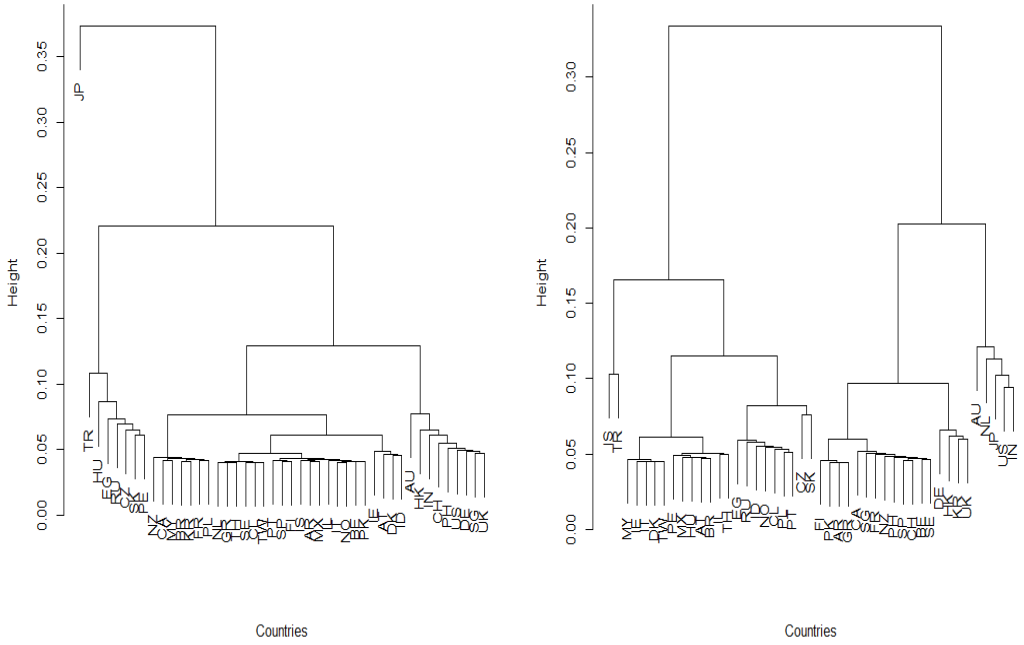


Figure 4: Dendogram at beginning of sub-prime mortgage crisis (left) and end of sub-prime mortgage crisis (right).

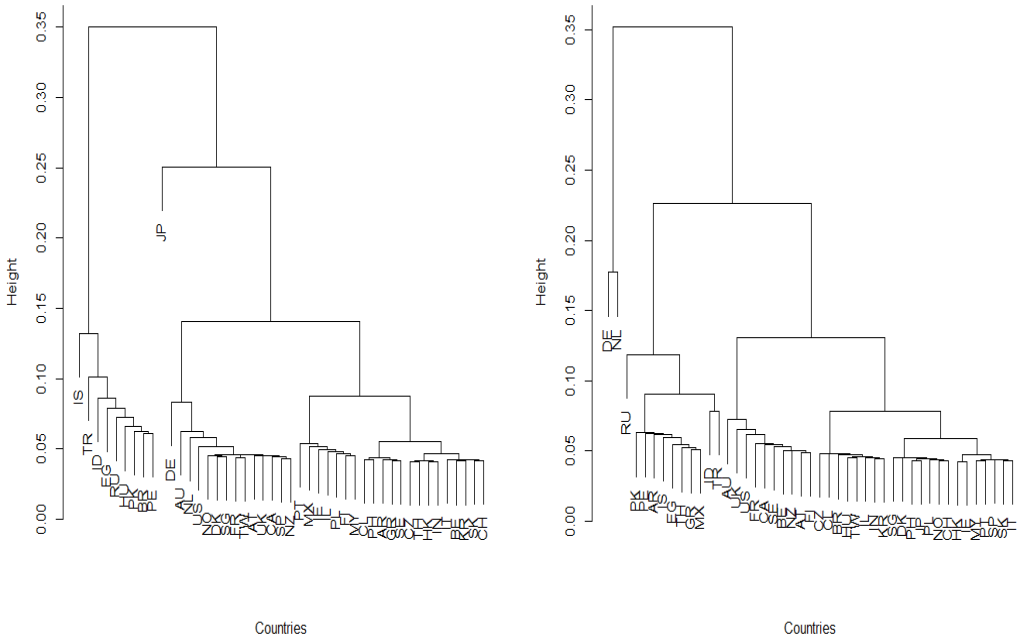


Figure 5: Dendogram at beginning of Euro debt (left) and end of Euro debt (right).

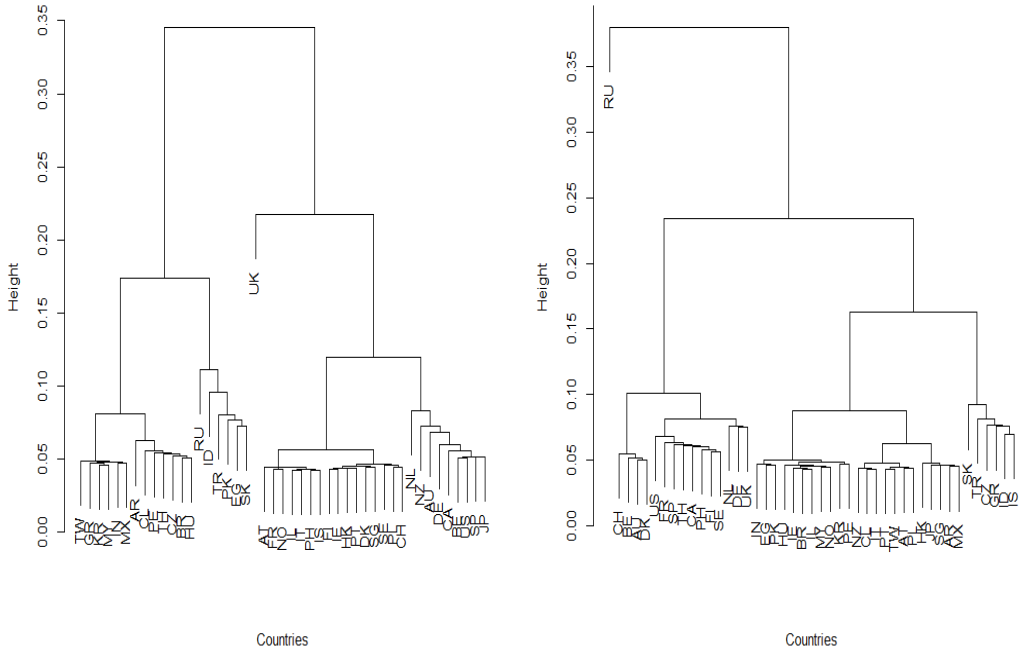


Figure 6: Dendogram at beginning of Brexit (left) and end of Brexit (right).

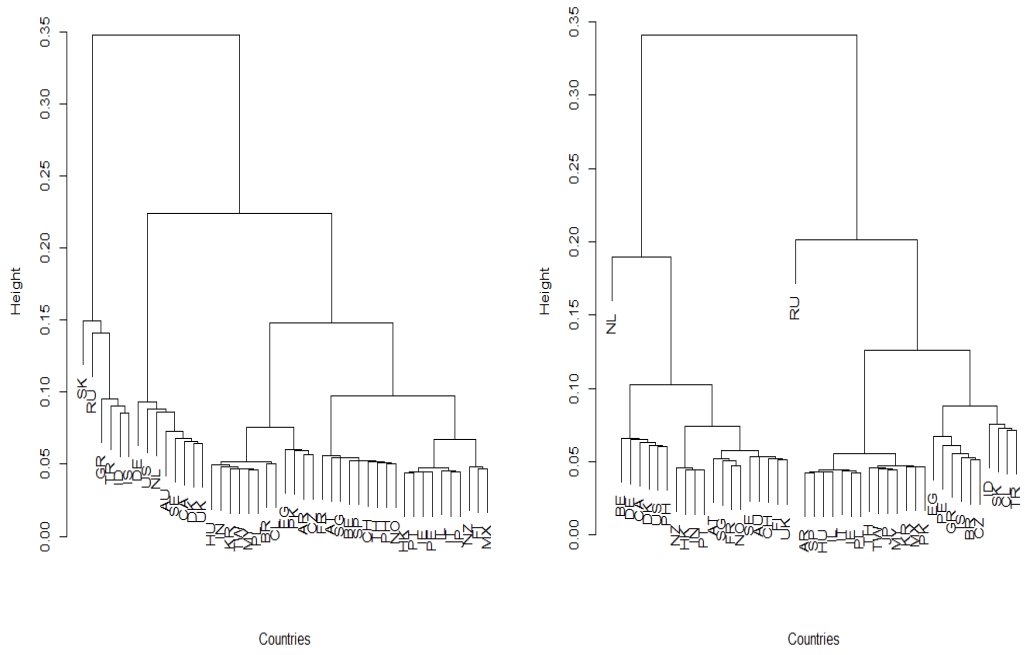


Figure 7: Dendogram at beginning of COVID-19 (left) and end of COVID-19 (right).

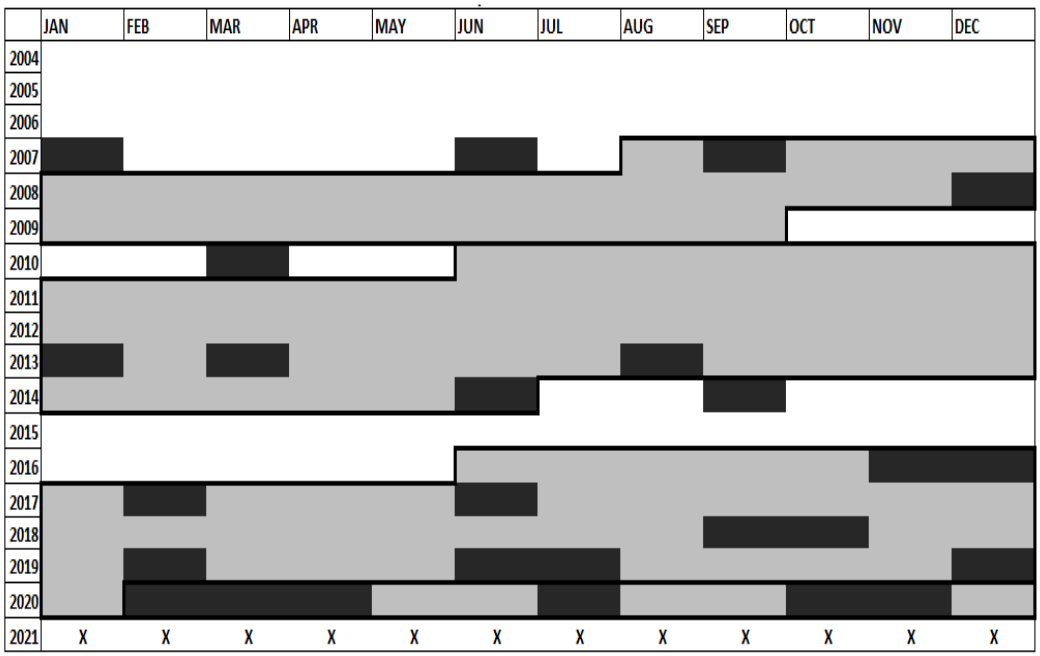


Figure 8: Monthly results for Moran's I test of spatial dependence with WUI distance. The crisis period is in gray, the non-crisis period in white and significant spatial dependence at 10% level in black (the months outside the analyzed period are marked with x).

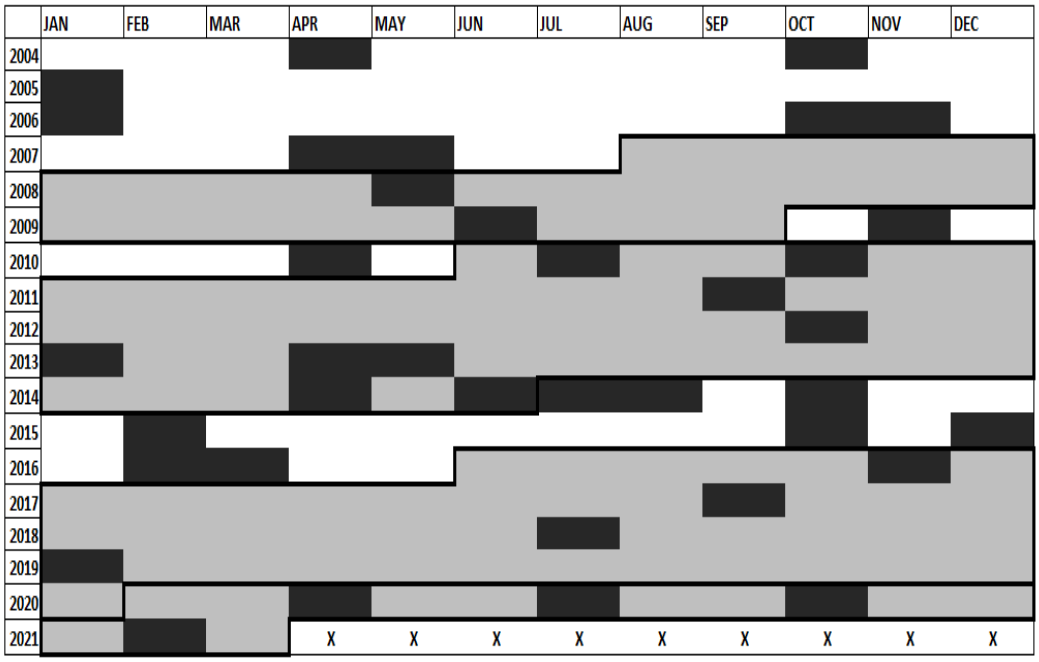


Figure 9: Monthly results for Moran's I test of spatial dependence with hours overlapping criterion. The crisis period is in gray, the non-crisis period in white and significant spatial dependence at 10% level in black (the months outside the analyzed period are marked with x).

A-3 Procedures

Procedure to obtain independent copies

To generate independent copies of $\widehat{\mathbb{D}}^{r_l}$ we use the result of Proposition 4.2 and under H_0 we define $\widehat{\mathbb{D}}^{r}(u_1,u_2) \longmapsto rC^{r-1}(u_1^{1/r},u_2^{1/r})\mathbb{G}(u_1^{1/r},u_2^{1/r}) - \mathbb{G}(u_1,u_2)$. Following Kojadinovic et al. (2011), replacing the empirical copula by a Beta transformed kernel estimator of copula, the Gaussian process is estimated as:

$$\widehat{\mathbb{G}}^{(s)}(u_1,u_2) = \hat{\alpha}^{(s)}(u_1,u_2) - \frac{\partial \widehat{C}^{B}(u_1,u_2)}{\partial u_1}\hat{\alpha}^{(s)}(u_1,1) - \frac{\partial \widehat{C}^{B}(u_1,u_2)}{\partial u_2}\hat{\alpha}^{(s)}(1,u_2), \tag{A-1}$$

where $s = 1,...R$ is a large number of copies. In expression (A-1) we have that $\hat{\alpha}^{(s)}(u_1,u_2) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(Z_i^{(s)} - \bar{Z}^{(s)}\right)K\left(\frac{M^{-1}(u_1)-M^{-1}(\hat{U}_{i1})}{b_n}\right)K\left(\frac{M^{-1}(u_2)-M^{-1}(\hat{U}_{i2})}{b_n}\right)$, where $Z_i^{(s)}$ are i.i.d. random variables with mean 0 and variance 1, so that $\int_0^{\infty} Pr\left(|Z_i^{(s)}| > t\right)dt < \infty$, $Pr(\cdot)$ is the probability function. Finally, we obtain the copies of the statistic for the extreme value copula test as:

$$\widehat{\mathbb{D}}^{r_l,(s)} = r_l\left\{\widehat{C}^{B}(u_1,u_2)\right\}^{r_l-1}\widehat{\mathbb{G}}^{(s)}(u_1^{\frac{1}{r_l}},u_2^{\frac{1}{r_l}}) - \widehat{\mathbb{G}}^{(s)}(u_1,u_2). \tag{A-2}$$

A-4 R programs

A-4.1 Inference with asymptotic and bootstrap global Moran's I statistic distribution

```
# Load packages (install them if it is necessary)

library(spdep)   ### Moran's I
library(sp)      ###
library(MASS)

#########################################################
##### Read GTUI data.                              ####
##### The first column indicates period and the rest
##### stock market indices
#########################################################


dat<-read.csv(file="GTUI.csv", header=T, dec=",", sep=";")


#####################################
#       Contiguity matrices         #
#####################################

nt<-195  ### Number of periods
# We have 46 stock markets
Dall<-array(0,dim = c(46,46,nt))
colnames(dat[2:47])->colnames(Dall)


for(j in 1:nt)
{
  D<-dat[dat$X==j,]
  data<-D[,c(2:47)]

  comb<-combn(data,2,simplify=FALSE)

  diff1<-vector("list")
```

```
  for (i in 1:length(comb))
  {
    frame<-as.data.frame(comb[i])
    diff1[[i]]<-abs(frame[,1]-frame[,2])
  }

  diff1<-as.data.frame(diff1)

    td<-t(colSums(diff1))
  t<-as.vector(td)

  ### A symmetric matrix
  SIMTRICA<-vec2mat(data=t,nrow=46)

  HASGARIAN<-apply(SIMTRICA, MARGIN = 2,
  FUN = function(X)(1-(X-min(X))/(max(X)-min(X))))

  diag(HASGARIAN) <- 0

  EST<-(HASGARIAN)
  Dall[,,j]<-EST
}


# Read financial data: filtered returns or losses
# or risks measures.
# First column is market label, second the month, third
# the numbered index and
# the forth the financial value to analyse
dades<-read.table(file="fd.csv", header=TRUE, dec=",",
sep=";")


n=46 # number of markets
nb=1000
nrep<-195 # numbers of periods

# For boostrap values
Igbot<-matrix(0,nb,nrep)
```

```
# For test results
test<-matrix(0,nrep,9)

for(r in 1:nrep){
C<- Dall[,,r]

m1<- C+(max(C)+10)*diag(diag(matrix(1,n,n)))

vec<-as.vector(c(m1))

vec<-vec[(vec<(max(C)+10))]

# medi<-mean(vec)
# medi<-median(vec)
# medi<-quantile(vec,0.75)
 medi<-quantile(vec,0.90)

C<-matrix(as.numeric(C>=medi),n,n)

W<-C/rowSums(C)

which(is.na(W))
W[is.na(W)] <- 0

rowSums(C)
neib<- mat2listw(W,style="M")

Y<-dades[dades$X==r,]$Vol
# Y<-dades[dades$X==r,]$Mod.VaR95
# Y<-dades[dades$X==r,]$Mod.VaR99

# Global Moran's statistic for original data
Ig<-moran.test(Y, neib, zero.policy=TRUE)

test[r,1]<-Ig$estimate[1] #valor estadarizado
# sin estandarizar
test[r,2]<-Ig$estimate[2]-qnorm(0.95)*sqrt(Ig$estimate[3])
test[r,3]<-Ig$estimate[2]+qnorm(0.95)*sqrt(Ig$estimate[3])
```

```
# sin estandarizar
test[r,4]<-Ig$estimate[2]-qnorm(0.90)*sqrt(Ig$estimate[3])
test[r,5]<-Ig$estimate[2]+qnorm(0.90)*sqrt(Ig$estimate[3])

# Bootstrap samples
set.seed(123456)

for(i in 1:nb){
  Yb<-sample(Y,n,replace = T)
  Yb<-as.vector(Yb)
  # Global Moran's statistic for bootstrap sample
  g<-moran.test(Yb, neib, zero.policy=TRUE)
  Igbot[i,r]<-g$estimate[1]
}

Igb<-as.vector(unlist(Igbot[,r], use.names=FALSE))
test[r,6]<-quantile(Igb,0.05)
test[r,7]<-quantile(Igb,0.95)
test[r,8]<-quantile(Igb,0.10)
test[r,9]<-quantile(Igb,0.90)
print(r)
}

write.table(test, file="Global_Vol_cuantil90.csv",
sep=";",dec=",")
```

A-4.2 Inference with asymptotic and bootstrap local Moran's I statistic distribution

```
# We analyse 11 indices

# For boostrap values
Igbot<-array(0,dim = c(nb,nrep,11))
# For test results
test<-array(0,dim = c(nrep,5,11))
# For the number of neighbors
Cm<-matrix(0,nrep,10)

# The number labels of the analysed indices
```

```
nneigh<-matrix(0,nrep,11)
i1=1
i2=2
i3=6
i4=12
i5=13
i6=14
i7=18
i8=20
i9=28
i10=29
i11=46

indneign<-c(i1,i2,i3,i4,i5,i6,i7,i8,i9,i10,i11)
tini<-Sys.time()
for(r in 1:nrep){
C<- Dall[,,r]

m1<- C+(max(C)+10)*diag(diag(matrix(1,n,n)))

vec<-as.vector(c(m1))

vec<-vec[(vec<(max(C)+10))]

# Change the median for quantiles...
 medi<-median(vec)
# medi<-quantile(vec,0.75)
# medi<-quantile(vec,0.90)

C<-matrix(as.numeric(C>=medi),n,n)

W<-C/rowSums(C)
sumrC<-rowSums(C)
nneigh[r,]<-sumrC[indneign]

neib<- mat2listw(W,style="M")

Y<-dades[dades$X==r,]$Vol
```

```
Ig<-localmoran(Y, neib, zero.policy=NULL, na.action=na.fail,
conditional=TRUE,alternative = "greater",
p.adjust.method="none", mlvar=TRUE,
spChk=NULL, adjust.x=FALSE)

test[r,1,1]<-Ig[i1,1]
test[r,2,1]<-Ig[i1,2]+qnorm(0.99)*sqrt(Ig[i1,3])
test[r,3,1]<-Ig[i1,2]+qnorm(0.95)*sqrt(Ig[i1,3])
test[r,1,2]<-Ig[i2,1]
test[r,2,2]<-Ig[i2,2]+qnorm(0.99)*sqrt(Ig[i2,3])
test[r,3,2]<-Ig[i2,2]+qnorm(0.95)*sqrt(Ig[i2,3])
test[r,1,3]<-Ig[i3,1]
test[r,2,3]<-Ig[i3,2]+qnorm(0.99)*sqrt(Ig[i3,3])
test[r,3,3]<-Ig[i3,2]+qnorm(0.95)*sqrt(Ig[i3,3])
#
test[r,1,4]<-Ig[i4,1]
test[r,2,4]<-Ig[i4,2]+qnorm(0.99)*sqrt(Ig[i4,3])
test[r,3,4]<-Ig[i4,2]+qnorm(0.95)*sqrt(Ig[i4,3])
test[r,1,5]<-Ig[i5,1]
test[r,2,5]<-Ig[i5,2]+qnorm(0.99)*sqrt(Ig[i5,3])
test[r,3,5]<-Ig[i5,2]+qnorm(0.95)*sqrt(Ig[i5,3])
test[r,1,6]<-Ig[i6,1]
test[r,2,6]<-Ig[i6,2]+qnorm(0.99)*sqrt(Ig[i6,3])
test[r,3,6]<-Ig[i6,2]+qnorm(0.95)*sqrt(Ig[i6,3])
#
test[r,1,7]<-Ig[i7,1]
test[r,2,7]<-Ig[i7,2]+qnorm(0.99)*sqrt(Ig[i7,3])
test[r,3,7]<-Ig[i7,2]+qnorm(0.95)*sqrt(Ig[i7,3])
test[r,1,8]<-Ig[i8,1]
test[r,2,8]<-Ig[i8,2]+qnorm(0.99)*sqrt(Ig[i8,3])
test[r,3,8]<-Ig[i8,2]+qnorm(0.95)*sqrt(Ig[i8,3])
test[r,1,9]<-Ig[i9,1]
test[r,2,9]<-Ig[i9,2]+qnorm(0.99)*sqrt(Ig[i9,3])
test[r,3,9]<-Ig[i9,2]+qnorm(0.95)*sqrt(Ig[i9,3])
#
test[r,1,10]<-Ig[i10,1]
test[r,2,10]<-Ig[i10,2]+qnorm(0.99)*sqrt(Ig[i10,3])
test[r,3,10]<-Ig[i10,2]+qnorm(0.95)*sqrt(Ig[i10,3])
test[r,1,11]<-Ig[i11,1]
```

```
test[r,2,11]<-Ig[i11,2]+qnorm(0.99)*sqrt(Ig[i11,3])
test[r,3,11]<-Ig[i11,2]+qnorm(0.95)*sqrt(Ig[i11,3])

for(i in 1:nb){
  Yb<-sample(Y,n,replace = T)
  Yb<-as.vector(Yb)

  g<-localmoran(Yb, neib, zero.policy=NULL,
  na.action=na.fail, conditional=TRUE,
  alternative = "greater", p.adjust.method="none",
  mlvar=TRUE,spChk=NULL, adjust.x=FALSE)

  Igbot[i,r,1]<-g[i1,1]
  Igbot[i,r,2]<-g[i2,1]
  Igbot[i,r,3]<-g[i3,1]

  Igbot[i,r,4]<-g[i4,1]
  Igbot[i,r,5]<-g[i5,1]
  Igbot[i,r,6]<-g[i6,1]

  Igbot[i,r,7]<-g[i7,1]
  Igbot[i,r,8]<-g[i8,1]
  Igbot[i,r,9]<-g[i9,1]

  Igbot[i,r,10]<-g[i10,1]
  Igbot[i,r,11]<-g[i11,1]
  }

test[r,4,1]<-quantile(Igbot[,r,1],0.99)
test[r,5,1]<-quantile(Igbot[,r,1],0.95)

test[r,4,2]<-quantile(Igbot[,r,2],0.99)
test[r,5,2]<-quantile(Igbot[,r,2],0.95)

test[r,4,3]<-quantile(Igbot[,r,3],0.99)
test[r,5,3]<-quantile(Igbot[,r,3],0.95)

test[r,4,4]<-quantile(Igbot[,r,4],0.99)
test[r,5,4]<-quantile(Igbot[,r,4],0.95)
```

```
test[r,4,5]<-quantile(Igbot[,r,5],0.99)
test[r,5,5]<-quantile(Igbot[,r,5],0.95)

test[r,4,6]<-quantile(Igbot[,r,6],0.99)
test[r,5,6]<-quantile(Igbot[,r,6],0.95)

test[r,4,7]<-quantile(Igbot[,r,7],0.99)
test[r,5,7]<-quantile(Igbot[,r,7],0.95)

test[r,4,8]<-quantile(Igbot[,r,8],0.99)
test[r,5,8]<-quantile(Igbot[,r,8],0.95)

test[r,4,9]<-quantile(Igbot[,r,9],0.99)
test[r,5,9]<-quantile(Igbot[,r,9],0.95)

test[r,4,10]<-quantile(Igbot[,r,10],0.99)
test[r,5,10]<-quantile(Igbot[,r,10],0.95)
test[r,4,11]<-quantile(Igbot[,r,11],0.99)
test[r,5,11]<-quantile(Igbot[,r,11],0.95)

print(r)
}
write.table(test, file="Local_ModVaR99_Median.csv",
sep=";",dec=".")
```

A-4.3 Testing extreme value copulas

```
library(rugarch)
library(qrmtools)
library(copula)
library(timeSeries)


loss<-read.csv("losses.csv")
nr<-nrow(loss)
nr
mcopa<-loss[,c("Alemania","España")]
mcopi<-loss[,c("España","Italia")]
```

```
mcopf<-loss[,c("España","resFr")]
mcopp<-loss[,c("España","resPor")]
mcopu1<-loss[,c("España","resUK")]
mcopu2<-loss[,c("España","resUSADJ")]
mcopu3<-loss[,c("España","resUSASP")]
mcoph<-loss[,c("España","resHK")]

mcop<-pobs(mcopa)

plot(mcop[,1],mcop[,2])

summary(mcop)

ini<-Sys.time()
r<-c(3,4,5)
###FUNCTIONS
## pdf beta(3,3)
dbeta33 = function(x){
  rr<-((15/16)*(1-x^2)^2)*(abs(x)<=1)
  return(rr)
}

## cdf of beta(3,3)
pbeta33 = function(x){
  if((abs(x)<=1)){
    rr<-(3/16)*x^5-(5/8)*x^3+(15/16)*x+0.5}
  if((x<=-1)){
    rr<-0}
  if((x>=1)){
    rr<-1}
  return(rr)
}

## cdf for quantile estimation beta(3,3)
pbeta33_p = function(x){
  if((abs(x)<=1)){
    rr<-(3/16)*x^5-(5/8)*x^3+(15/16)*x+0.5-p}
  if((x<=-1)){
    rr<-0-p}
```

```
  if((x>=1)){
    rr<-1-p}
  return(rr)
}

  m2<-mcop
  for(i in 1:nr){
    p <- m2[i,1]
    m2[i,1] <- uniroot(pbeta33_p,c(-1,1))$root
    p <- m2[i,2]
    m2[i,2] <- uniroot(pbeta33_p,c(-1,1))$root
  }
  tmcop<-m2

 rej <- matrix(seq(0.001,0.999,0.005), 2, 200,
  byrow = TRUE)
 rej2 <- matrix(seq(0.001,0.999,0.005), 2, 200,
  byrow = TRUE)

 rej<-t(rej)
 rej2<-t(rej2)
rej2_r<-cbind(rej2,rej2,rej2)
ngrid<-nrow(rej) # number of poing the the grid

# Transformed grid

for (i in 1:ngrid){
  p <- rej2[i,1]
  rej2[i,1] <- uniroot(pbeta33_p,c(-1,1))$root
  p <- rej2[i,2]
  rej2[i,2] <- uniroot(pbeta33_p,c(-1,1))$root
  for(ri in 1:3){
    p <- rej2_r[i,2*(ri-1)+1]**(1/r[ri])
    rej2_r[i,2*(ri-1)+1] <- uniroot(pbeta33_p,c(-1,1))$root
    p <- rej2_r[i,2*(ri-1)+2]**(1/r[ri])
    rej2_r[i,2*(ri-1)+2] <- uniroot(pbeta33_p,c(-1,1))$root

  }
}
```

```
## pdf Kernel Epanecnikov
kintd<-function(u){
  rr <- 0.75*(1-u**2)
}

## cdf Kernel Epanecnikov
Kintd<-function(u){
  rr<--0.25*((1+u)**2)*(u-2);
  return (rr)}

## call pdf kernel
pdfhatz<-function(z,v,nr,b){
  u = (z-v)/b
  #z is a scalar, v is a scalar
  rr<-kintd(u)
  return(rr)}

## call cdf kernel
cdfhatz<-function(z,v,nr,b){
  u = (z-v)/b
  #z is a scalar, v is a scalar
  rr<-Kintd(u)
  return(rr)}


Dif_Inv<- function(res){
  if(abs(res)<=1){
    rr <- (15/16)*(1-(res)^2)^2
  }else{
    rr <- 0
  }
  return(1/rr)
}

###kernel estimation of cdf and derivatives

FUN_x<- function(z){
```

```
  ind12n<-((z[1]-v[,1])/b<=(-1))*((z[2]-v[,2])/b<=(-1))
  suma0<-ind12n*0
  ind12p<-((abs(z[1]-v[,1])/b<1)*(abs(z[2]-v[,2])/b<1))
  suma1<-(1/nr)*cdfhatz(z[1],v[,1],nr,b)*
  cdfhatz(z[2],v[,2],nr,b)*ind12p
  ind1n2p<-((z[1]-v[,1])/b>=1)*((z[2]-v[,2])/b<1)
  suma2=((1/nr)*cdfhatz(z[2],v[,2],nr,b))*ind1n2p
  ind1p2n<-((z[1]-v[,1])/b<1)*((z[2]-v[,2])/b>=1)
  suma3<-cdfhatz(z[1],v[,1],nr,b)*(1/nr)*ind1p2n
  ind1n2n<-((z[1]-v[,1])/b>=1)*((z[2]-v[,2])/b>=1)
  suma4<-(1/nr)*ind1n2n
  rr=sum(suma0)+sum(suma1)+sum(suma2)+sum(suma3)+sum(suma4)
  return(rr)
}

FUNd1_x<- function(z){
  #z es un vector de dim 2, v es una matriz
  ind12n<-((z[1]-v[,1])/b<=(-1))*((z[2]-v[,2])/b<=(-1))
  suma0<-ind12n*0
  ind12p<-((abs(z[1]-v[,1])/b<1)*(abs(z[2]-v[,2])/b<1))
  suma1<-(1/(nr*b*dbeta33(z[1])))*pdfhatz(z[1],v[,1],nr,b)*
  cdfhatz(z[2],v[,2],nr,b)*ind12p
  ind1n2p<-((z[1]-v[,1])/b>=1)*((z[2]-v[,2])/b<1)
  suma2=ind1n2p*0
  ind1p2n<-((z[1]-v[,1])/b<1)*((z[2]-v[,2])/b>=1)
  suma3<-(1/(nr*b*dbeta33(z[1])))*pdfhatz(z[1],v[,1],nr,b)*
  ind1p2n
  ind1n2n<-((z[1]-v[,1])/b>=1)*((z[2]-v[,2])/b>=1)
  suma4<-ind1n2n*0
  rr=sum(suma0)+sum(suma1)+sum(suma2)+sum(suma3)+sum(suma4)
  return(rr)
}


FUNd2_x<- function(z){
  ind12n<-((z[1]-v[,1])/b<=(-1))*((z[2]-v[,2])/b<=(-1))
  suma0<-ind12n*0
  ind12p<-((abs(z[1]-v[,1])/b<1)*(abs(z[2]-v[,2])/b<1))
  suma1<-(1/(nr*b*dbeta33(z[2])))*cdfhatz(z[1],v[,1],nr,b)*
```

```
  pdfhatz(z[2],v[,2],nr,b)*ind12p
  ind1n2p<-((z[1]-v[,1])/b>=1)*((z[2]-v[,2])/b<1)
  suma2=(1/(nr*b*dbeta33(z[2])))*pdfhatz(z[2],v[,2],nr,b)*
  ind1n2p
  ind1p2n<-((z[1]-v[,1])/b<1)*((z[2]-v[,2])/b>=1)
  suma3<-ind1p2n*0
  ind1n2n<-((z[1]-v[,1])/b>=1)*((z[2]-v[,2])/b>=1)
  suma4<-ind1n2n*0
  rr=sum(suma0)+sum(suma1)+sum(suma2)+sum(suma3)+sum(suma4)
  return(rr)
}


CKernelBeta<-array(0, dim=c(ngrid,ngrid))
CKernelBeta_r<-array(0, dim=c(ngrid,ngrid,3))
dCKernelBeta_r<-array(0, dim=c(ngrid,ngrid,3))
d1CKernelBeta<-array(0, dim=c(ngrid,ngrid))
d2CKernelBeta<-array(0, dim=c(ngrid,ngrid))
d1CKernelBeta_r<-array(0, dim=c(ngrid,ngrid,3))
d2CKernelBeta_r<-array(0, dim=c(ngrid,ngrid,3))


## Tail dependence Empririca copula

rej1<-c(0.99,0.99)
    AUX1 <- mcop[,1]<=rej1[1]
    AUX2 <- mcop[,2]<=rej1[2]
    CEmpirica<-sum(AUX1 == TRUE & AUX2 == TRUE)/nr
    res_Emp<-1-CEmpirica
    res_Emp

## Tail dependence classical kernel copula

    # Definimos la ventana optima para este caso
    b <- 3.572*nr^(-1/3)
    v <- mcop
        CKernelClasico <- FUN_x(c(rej1[1],rej1[2]))
        res_ker<-1-CKernelClasico
        res_ker
```

```
## Tail dependence transformed Gaussian kernel copula
          m2<-mcop
          for(i in 1:nr){
            p <- m2[i,1]
            m2[i,1] <- qnorm(p)
            p <- m2[i,2]
            m2[i,2] <- qnorm(p)}
          rej31<-rej1
          p<-rej1[1]
          rej31[1]<-qnorm(p)
          p<-rej1[2]
          rej31[2]<-qnorm(p)

    b <- 3.572*nr^(-1/3)
    v <- m2
     CKernelNormal<- FUN_x(c(rej31[1],rej31[2]))
     res_Gtker<-1-CKernelNormal
     res_Gtker

 ## Tail dependence transformed beta kernel copula
     m1<-mcop
     for(i in 1:nr){
       p <- m1[i,1]
       m1[i,1] <- uniroot(pbeta33_p,c(-1,1))$root
       p <- m1[i,2]
       m1[i,2] <- uniroot(pbeta33_p,c(-1,1))$root}
     rej21<-rej1
     p<-rej1[1]
     rej21[1]<-uniroot(pbeta33_p,c(-1,1))$root
     p<-rej1[2]
     rej21[2]<-uniroot(pbeta33_p,c(-1,1))$root

     b <- (3)^(1/3)*nr^(-1/3)
     v <- m1
     CKernelBeta<- FUN_x(c(rej21[1],rej21[2]))
     res_Btker<-1-CKernelBeta
     res_Btker
```

```
# Test with empirical copula
evTestC(mcop)

# Test with beta transformed kernel copula

# Dn<-matrix(0,ngrid,ngrid)
Sr<-matrix(0,1,3)
b <- 3**(1/3)*nr^(-1/3)

  v <- tmcop
  for(i in 1:ngrid){
    for(j in 1:ngrid){
      CKernelBeta[i,j]<- FUN_x(c(rej2[i,1],rej2[j,2]))
      d1CKernelBeta[i,j]<-FUNd1_x(c(rej2[i,1],rej2[j,2]))
      d2CKernelBeta[i,j]<-FUNd2_x(c(rej2[i,1],rej2[j,2]))
      for(ri in 1:3){
        # v <- tmcopr[,(2*(k-1)+1):(2*k),ri]
        CKernelBeta_r[i,j,ri]<-FUN_x(c(rej2_r[i,2*
        (ri-1)+1],rej2_r[i,2*(ri-1)+2]))**(r[ri])
        dCKernelBeta_r[i,j,ri]<-r[ri]*FUN_x(c(rej2_r[i,2*
        (ri-1)+1],rej2_r[i,2*(ri-1)+2]))**(r[ri]-1)
        d1CKernelBeta_r[i,j,ri]<-FUNd1_x(c(rej2_r[i,2*
        (ri-1)+1],rej2_r[i,2*(ri-1)+2]))*((1/r[ri])*
        rej[i,1]**((1-r[ri])/r[ri]))
        d2CKernelBeta_r[i,j,ri]<-FUNd2_x(c(rej2_r[i,2*
        (ri-1)+1],rej2_r[i,2*(ri-1)+2]))*((1/r[ri])*
        rej[j,2]**((1-r[ri])/r[ri]))
        Sr[1,ri]<-(nr/(ngrid*ngrid))*
        sum((CKernelBeta_r[,,ri]-CKernelBeta)**2)
      }
    }
  }

Sr345<-rowSums(Sr)
Sr345
###WE CAN USE BOOTSTRAP TO CALCULATE P-VALUE,
###THE PREDURE OF THE TEST IS LONG AND THE RESULTS
###ARE SIMILAR TO EMPIRICAL COPULA
fin<-Sys.time()
```

Bibliography

Acuña, C. A., Bolancé, C., and Torra, S. (2018). Análisis de la dependencia espacial entre índices bursátiles. *Anales del Instituto de Actuarios Españoles*, 24:79–97.

Acuña, C. A., Bolancé, C., and Torra, S. (2021). Dynamic distances between stock markets: use of uncertainty indices measures. *Anales del Instituto de Actuarios Españoles*, 27:55–73.

Ahir, H., Bloom, N., and Furceri, D. (2019). The World Uncertainty Index. Technical Report 19–27, Stanford Institute for Economic Policy Research.

Akhtaruzzaman, M., Boubaker, S., and Sensoy, A. (2021). Financial contagion during COVID-19 crisis. *Finance Research Letters*, 38:101604.

Alemany, R., Bolancé, C., and Guillen, M. (2013). A nonparametric approach to calculating value-at-risk. *Insurance: Mathematics and Economics*, 52:255–262.

Ameur, H. B. and Louhichi, W. (2021). The Brexit impact on European market co-movements. *Annals of Operations Research*, 303:1387–1403.

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis*, 27:93–115.

Arnold, M., Stahlberg, S., and Wied, D. (2013). Modeling different kinds of spatial dependence in stock returns. *Empirical Economics*, 44:761–774.

Asgharian, H. and Bengtsson, C. (2006). Jump Spillover in International Equity Markets. *Journal of Financial Econometrics*, 4:167–203.

Asgharian, H., Hess, W., and Liu, L. (2013). A spatial analysis of international stock market linkages. *Journal of Banking & Finance*, 37(12):4738–4754.

Asgharian, H. and Nossman, M. (2011). Risk contagion among international stock markets. *Journal of International Money and Finance*, 30:22–38.

Ashraf, B. N. (2020). Stock markets' reaction to COVID-19: Cases or fatalities? *Research in International Business and Finance*, 54:101249.

Avilés, G. F., Montero, J., and Orlov, A. (2012). Spatial modeling of stock market comovements. *Finance Research Letters*, 9:202–212.

Azzalini, A. (1981). A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, 68:326–328.

Bahraoui, Z., Bolancé, C., and Pérez-Marín, A. (2014). Testing extreme value copulas to estimate the quantile. *SORT-Statistics and Operations Research Transactions*, 38:89–102.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Baker, S. R., Bloom, N., Davis, S. J., Kost, K., Sammon, M., and Viratyosin, T. (2020). The Unprecedented Stock Market Reaction to COVID-19. *The Review of Asset Pricing Studies*, 10:742–758.

Bolancé, C. and Acuña, C. A. (2021). A new kernel estimator of copulas based on Beta quantile transformations. *Mathematics*, 9:1078.

Bolancé, C., Acuña, C. A., and Torra, S. (2022). Non-normal market losses and spatial dependence using uncertainty indices. *Mathematics*, 10:1317.

Bolancé, C. and Guillen, M. (2021). Nonparametric Estimation of Extreme Quantiles with an Application to Longevity Risk. *Risks*, 9(4).

Breinlich, H., Leromain, E., Novy, D., Sampson, T., and Usman, A. (2018). The Economic Effects of Brexit: Evidence from the Stock Market. *Fiscal Studies*, 39:581–623.

Buch-Larsen, T., Guillen, M., Nielsen, J., and Bolancé, C. (2005). Kernel density estimation for heavy-tailed distributions using the Champernowne transformation. *Statistics*, 39:503–518.

Burdekin, R. C. K., Hughson, E., and Gu, J. (2018). A first look at Brexit and global equity markets. *Applied Economics Letters*, 25:136–140.

Castelnuovo, E. and Tran, T. D. (2017). Google It Up! A Google Trends-based Uncertainty Index for the United States and Australia. *Economics Letters*, 161:149–53.

Chen, S. and Huang, T. (2007). Nonparametric estimation of copula functions for dependence modelling. *The Canadian Journal of Statistics*, 35:265–282.

Chopra, M. and Mehta, C. (2022). Is the COVID-19 pandemic more contagious for the Asian stock markets? A comparison with the Asian financial, the US subprime and the Eurozone debt crisis. *Journal of Asian Economics*, 79:101450.

Dimitriou, D., Kenourgios, D., and Simos, T. (2013). Global financial crisis and emerging stock market contagion: A multivariate FIAPARCH-DCC approach. *International Review of Financial Analysis*, 30:46–56.

Eckel, S., Löffler, G., Maurer, A., and Schmidt, V. (2011). Measuring the effects of geographical distance on stock market correlation. *Journal of Empirical Finance*, 18:237–247.

Fermanian, J. D., Radulović, D., and Wegkamp, M. H. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10:847–860.

Fernandez, V. (2011). Spatial linkages in international financial markets. *Quantitative Finance*, 11:237–245.

Flavin, T., Hurley, M. J., and Rousseau, F. (2002). Explaining Stock Market Correlation: A Gravity Model Approach. *Manchester School*, 70:87–106.

Francq, C. and Zakoïan, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10:605–637.

Gasser, T., G., M. H., and Mammitzsch, V. (1985). Kernels for Nonparametric Curve Estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47:238–252.

Gasser, T. and Müller, H. G. (1979). Kernel estimation of regression functions. In Gasser, T. and Rosenblatt, M., editors, *Smoothing Techniques for Curve Estimation*, pages 23–68, Berlin, Heidelberg. Springer.

Genest, C., Kojadinovic, I., Nešleová, J., and Yan, J. (2011). A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli*, 17:253 – 275.

Genest, C., Nešleová, J., and Quessy, J. F. (2012). Tests of symmetry for bivariate copulas. *Bernoulli*, 64:811 – 834.

Genest, C. and Rivest, L. P. (1993). Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association*, 88:1034–1043.

Ghirelli, C., Pérez, J. J., and Urtasun, A. (2019). A new economic policy uncertainty index for Spain. *Economics Letters*, 182:64–67.

Ghorbal, B., Genest, C., and Nešleová, J. (2009). On the ghoudi, khoudraji, and rivest test for extreme-value dependence. *The Canadian Journal of Statistics*, 37:534–552.

Gijbels, I. and Mielniczuk, J. (1990). Estimating the density of a copula function. *Communications in Statistics - Theory and Methods*, 19:445–464.

Griffith, D. (2010). The Moran Coefficient for Non-Normal Data. *Journal of Statistical Planning and Inference*, 140:2980–2990.

Hamao, Y., Masulis, R., and Ng, V. (1990). Correlations in Price Changes and Volatility across International Stock Markets. *Review of Financial Studies*, 3:281–307.

Hill, P. D. (1985). Kernel estimation of a distribution function. *Communications in Statistics - Theory and Methods*, 14:605–620.

Hussain, S. I. and Li, S. (2018). The dependence structure between chinese and other major stock markets using extreme values and copulas. *International Review of Economics and Finance*, 56:421–437.

Jin, F. and Lee, L. F. (2015). On the bootstrap for Moran'I test for spatial dependence. *Journal of Econometrics*, 184(2):295–314.

Karolyi, G. and Stulz, R. (1996). Why do markets move together? an investigation of u.s.-japan stock return comovements. *Journal of Finance*, 51:951–86.

Kojadinovic, I., Segers, J., and Yan, J. (2011). Large-sample tests of extreme-value dependence for multivariate copulas. *The Canadian Journal of Statistics*, 39:703–720.

Li, H. (2020). Volatility spillovers across European stock markets under the uncertainty of Brexit. *Economic Modelling*, 84:1–12.

Li, Y., Liang, C., Ma, F., and Wang, J. (2020). The role of the IDEMV in predicting european stock market volatility during the COVID-19 pandemic. *Finance Research Letters*, 36:101749.

Lien, D., Lee, G., Yang, L., and Zhang, Y. (2018). Volatility spillovers among the US and Asian stock markets: A comparison between the periods of Asian currency crisis and subprime credit crisis. *The North American Journal of Economics and Finance*, 46:187–201.

Liu, R. and Yang, L. (2008). Kernel estimation of multivariate cumulative distribution function. *Journal of Nonparametric Statistics*, 20:661–677.

Longin, F. and Solnik, B. (1995). Is the correlation in international equity returns constant: 1960 - 1990? *Journal of International Money and Finance*, 14:3 – 26.

Martin, V. and Dungey, M. (2007). Unravelling financial market linkages during crises. *Journal of Applied Econometrics*, 22:89–119.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools Revised edition*. Princeton University Press, USA.

Mei, C. L., Xu, S. F., and Chen, F. (2020). Consistency of bootstrap approximation to the null distributions of local spatial statistics with application to house price analysis. *Journal of Inequalities and Applications*, 217:2–19.

Mohti, W., Dionísio, A., Ferreira, P., and Vieira, I. (2019). Contagion of the Subprime Financial Crisis on Frontier Stock Markets: A Copula Analysis. *Economies*, 7.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37:17–23.

Omelka, M., Gijbels, I., and Veraverbeke, N. (2009). Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing. *The Annals of Stadistics*, 37:3023–3058.

Reiss, R.-D. (1981). Nonparametric estimation of smooth distribution functions. *Scandinavian Journal of Statistics*, 8:116–119.

Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100:377–386.

Samarakoon, L. P. (2017). Contagion of the Eurozone debt crisis. *Journal of International Financial Markets, Institutions and Money*, 49:115–128.

Samitas, A. and Tsakalos, I. (2013). How can a small country affect the European economy? the Greek contagion phenomenon. *Journal of International Financial Markets, Institutions and Money*, 25:18–32.

Scaillet, O. (2005). A Kolmogorov-smirnov type test for positive quadrant dependence. *The Canadian Journal of Statistics*, 33:415–427.

Segers, J. (2012). Max-Stable models for multivariate extremes. *REVSTAT-Statistical Journal*, 10:61–82.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London.

Sklar, A. (1959). Fonctions de répartition á n dimensions et leurs marges. publications de l'institut de statistique de l'université de paris. *The Annals of Stadistics*, 8:229–231.

Terrell, G. R. (1990). The maximal smoothing principle in density estimation. *Journal of the American Statistical Association*, 85:270–277.

Tilfani, O., Ferreira, P., and Boukfaoui, M. Y. E. (2021). Dynamic cross-correlation and dynamic contagion of stock markets: a sliding windows approach with the DCCA correlation coefficient. *Empirical Economics*, 60:1127–1156.

Wand, M. P. and Jones, M. C. (1995). *Kernel Smoothing*. Chapman & Hall, London.

Wang, J., Cheng, F., and Yang, L. (2013). Smooth simultaneous confidence bands for cumulative distribution functions. *Journal of Nonparametric Statistics*, 25:395–407.

Weinberg, L. (2020). The Google Trends Uncertainty (GTU) Index: A measure of economic policy uncertainty in the EU using Google Trends. *Undergraduate Economic Review*, 17:Article 2.

Weng, Y. and Gong, P. (2016). Modeling spatial and temporal dependencies among global stock markets. *Expert Systems with Applications*, 43:175–185.

Yang, Z. (2015). LM tests of spatial dependence based on bootstrap critical values. *Journal of Econometrics*, 185(1):33–59.

Zhang, D., Hu, M., and Ji, Q. (2020). Financial markets under the global pandemic of covid-19. *Finance Research Letters*, 36:101528.