

Luu, Duc Thi; Yanovski, Boyan; Lux, Thomas

**Working Paper**

## An analysis of systematic risk in worldwide economic sentiment indices

Economics Working Paper, No. 2018-03

**Provided in Cooperation with:**

Christian-Albrechts-University of Kiel, Department of Economics

*Suggested Citation:* Luu, Duc Thi; Yanovski, Boyan; Lux, Thomas (2018) : An analysis of systematic risk in worldwide economic sentiment indices, Economics Working Paper, No. 2018-03, Kiel University, Department of Economics, Kiel

This Version is available at:

<https://hdl.handle.net/10419/174870>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

C | A | U

Christian-Albrechts-Universität zu Kiel

Department of Economics

Economics Working Paper  
No 2018-03

# An Analysis of Systematic Risk in Worldwide Economic Sentiment Indices

by Duc Thi Luu, Boyan Yanovski and Thomas Lux

issn 2193-2476



# An Analysis of Systemic Risk in Worldwide Economic Sentiment Indices

DUC THI LUU<sup>†</sup>, BOYAN YANOVSKI<sup>†</sup> and THOMAS LUX<sup>†</sup>

<sup>†</sup>*Department of Economics, University of Kiel, 24118 Kiel, GERMANY*

*(e-mail: d.t.luu@economics.uni-kiel.de)*

*(e-mail: boyan.yanovski@gmail.com)*

*(e-mail: lux@economics.uni-kiel.de)*

(Dated: Ver. January 2, 2018)

## Abstract

In this paper, we investigate the temporal dynamics of correlations between sentiment indices worldwide. Employing the tools of Random Matrix Theory (RMT) and Principal Component Analysis (PCA), our paper aims to extract latent information embedded in the interactions between economic and business sentiment indices around the world. We find that: (i) The dynamics of the sentiment indices across countries can be well explained by the evolution of a single factor (the “market mode”); (ii) during most periods, some groups of countries exhibit sentiment dynamics less associated with (or divergent from) the market mode, while (iii) during the financial crisis, no country or group of countries has been able to escape the market mode, which accounts for almost all movements in the indices. We argue that strong “global” information signals, like the collapse of the US housing market in 2007, can lead to a homogenization of the expectation structure around the world, as such information can provide a coordination signal for a global phase of low confidence.

*JEL classification numbers: C19, E30, E71, G01*

*Keywords: Sentiment Index; Correlations; Financial Crisis; Random Matrix Theory; Principal Component Analysis*

## I. INTRODUCTION

The analysis of business cycles in different countries is one of the fundamental issues in international economics. So far, much of the analysis is often directed at investigating the synchronization and convergence of “tangible” macroeconomic variables like GDP growth rates, unemployment rates, and so forth (e.g. Bordo and Hebling, 2003; Bordo and Hebling, 2011; Artis et al., 2011; Kose et al., 2012; Ferroni and Klaus, 2015). However, up until now, issues related to the correlations between the expectation structures across different countries have been receiving less attention.

Expectations are a key driver of fluctuations in economic activity since most economically relevant decisions have a strong inter-temporal component (e.g. investment, consumption or saving decisions). This was emphasized, in particular, by Keynes (1936), and later by Minsky (1977) and Akerlof and Shiller (2009). Empirically, such claims are supported by studies by authors like Santero and Westerlund (1996), Howrey (2001), Taylor and McNabb (2007), Carriero and Marcellino (2011), Milani (2011), van Aarle and Kappler (2012), or Milani (2014) in which the structure of the expectations is measured by sentiment or confidence indices. The expectations themselves are formed on the basis of past experience or on currently incoming information signals from the economy. We argue that “global” information signals, like the collapse of the US housing market in 2007, can lead to a homogenization of the expectation structure around the world, as such information can provide a coordination signal for a global phase of pessimistic expectations. Here we confine ourselves to the phenomenological analysis of coordination of expectations. Whether this synchronization is justified in fundamental terms by the spillovers between countries in real economic activity, or whether it constitutes another, psychological factor of contagion, should be investigated in subsequent research.

This study contributes to the understanding of cross- correlations between economic and business sentiment indices worldwide. We aim to answer three main research questions: (i) how many statistically significant common factors can we extract from the joint dynamics of the sentiment indices worldwide; (ii) how well do these common factors account for the dynamics of the individual indices; and (iii) how does the weight of these factors change over time?

We analyze two data sets, i.e. the Business Confidence Index (BCI) collected by the

OECD and the Economic Sentiment Indices (ESI) collected by the Eurostat.<sup>1</sup> In terms of methods, instead of using traditional approaches based on econometric models, we employ Random Matrix Theory (see, for example, Laloux et al., 1999; Bouchaud and Potters, 2009) and Principal Component Analysis (see, for example, Jolliffe, 1986; Billio et al., 2012; Wang et al., 2011) to investigate the dynamics of the correlation matrix of country-specific sentiment/confidence indices. We extract the hidden factors encoded in the empirical correlations across countries by analyzing the group of eigenvalues (and their corresponding eigenvectors) deviating from the random bulk. In this way, we can capture the evolution of the statistically significant factors underlying the dynamics of the correlation matrix. The extent to which different countries are affected by these factors can be thought of as the risk of sentiment contagion that the individual countries are facing during a particular period.

This paper is structured as follows. In Sec. II we briefly describe the data and methods employed in our study. Sec. III reports our main findings. Discussions and concluding remarks are found in Sec. IV.

## II. DATA AND METHODS

### A. Data

We consider two data sets containing country-specific sentiment indices. The first data set is collected by the OECD, which consists of all the OECD members and several other countries including China, Russia, India, Turkey, and Brazil. We name this data set OECD<sup>+</sup>. The data set captures the Business Confidence Index (BCI) measured monthly for each country. The index is based on the entrepreneurs' assessments of their current production, orders and stocks, as well as on their expectations for the immediate future (e.g. OECD, 2016).<sup>2</sup> To avoid the problem of missing data in some reported countries, we confine our analysis to the period from January 2002 until the end of 2015. This gives us data on the monthly business confidence indices in 33 countries.

The second data set reports the Economic Sentiment Indices (ESI) of Eurozone members

---

<sup>1</sup> See the next section for a more detailed description of the two data sets.

<sup>2</sup> See OECD. (2016). Business Confidence Index (indicator). doi:10.1787/3092dc4f-en (Accessed on 29 January 2016).

and other European countries. The ESI summarizes consumer confidence, as well as the developments and expectations in the other surveyed sectors, i.e. industry (manufacturing), services, retail trade and construction sectors (e.g. EC, 2016).<sup>3</sup> In our analysis, we name this group of countries EU<sup>+</sup>. We use this data set for the period from January 1997 to December 2015, which gives us 24 monthly economic sentiment indices associated with 24 European countries.

## B. Methods

### Correlation matrix

Given the reported  $N$  indices for every month  $\{SI_{i,t}\}_{i=1:N}$  from time  $t=1$  to  $t=T$ , we apply a standard normalization procedure to the data.<sup>4</sup> First, we consider the difference in logs across periods

$$I_{i,t} = \ln(SI_{i,t+1}) - \ln(SI_{i,t}). \quad (1)$$

As a second step, we define the normalized log-sentiment index for the time horizon  $T$  as

$$X_{i,t} = \frac{I_{i,t} - \langle I_{i,t} \rangle}{\sigma_{i,t}}, \quad (2)$$

where  $\langle I_i(t) \rangle$  and  $\sigma_i(t)$  are respectively the time average and the standard deviation of  $I_i(t)$  over the time horizon  $T$ . Now we have  $\langle X_i \rangle = 0$  and  $\text{Var}(X_i) = 1$ . Next, we consider the rectangular matrix  $X = \{X_{i,t}\}_{N \times T}$  and the associated correlation matrix of the  $N$  normalized log-sentiment indices

$$C = \{C_{ij}\}_{N \times N} = \frac{1}{T} X X^T, \quad (3)$$

where the notation  $X^T$  stands for the matrix transposition of  $X$ . The value of  $C_{ij}$  denotes the correlation between country index  $i$  and  $j$ , where  $-1 < C_{ij} < 1$ , for  $1 \leq i < j \leq N$ . Note that, for any  $i$  we always have  $C_{ii} = 1$ . In case  $C_{ij} > 0$  ( $< 0$ ) the two countries  $i$  and  $j$  are positively (negatively) empirically correlated, while  $C_{ij} = 0$  indicates a lack of any correlation.

### Similarity matrix

<sup>3</sup> See EC. (2016). Economic Sentiment Index (ESI). [http://ec.europa.eu/economy\\_finance/db\\_indicators/surveys/index\\_en.htm](http://ec.europa.eu/economy_finance/db_indicators/surveys/index_en.htm) (Accessed on 29 January 2016).

<sup>4</sup> In our study, we choose  $T=36$  (months), which satisfies the condition that  $T > N$ .

One of the methods we use, to study the central question of the evolution of the sentiment correlation matrix over time, is the method proposed in Münnix et al. (2012), which is often applied when identifying states of stock markets.<sup>5</sup> The main idea is to come up with a measure of the similarity between correlation matrices from different periods. Suppose we observe two correlation matrices  $C(t_1)$  and  $C(t_2)$  associated with the two distinct periods  $t_1$  and  $t_2$  from the sample  $\{1, 2, \dots, \mathcal{T}\}$ , then the similarity  $S$  between those two matrices is defined as

$$S_{t_1, t_2} = \langle |C(t_1) - C(t_2)| \rangle, \quad (4)$$

where  $|\dots|$  is the notation for the absolute value. Note that a higher value of  $S_{t_1, t_2}$  indicates that the “distance” between two correlation matrices is higher; in contrast, a smaller value of  $S_{t_1, t_2}$  reveals a higher level of similarity between the two matrices.

### Random Matrix Theory

RMT, which was originally developed in nuclear physics by Wigner and Dirac to explain complex quantum systems, has emerged as one of the most important techniques for extracting latent information embedded in empirical correlations from the financial sector (e.g. Laloux et al., 1999; Laloux et al., 2000; Plerou et al., 2002; Kim and Jeong, 2005; Meng et al., 2014; Jiang et al., 2014; Uechi et al., 2015; MacMahon and Garlaschelli, 2015). Surprisingly, the applications of RMT in macroeconomic time series are very limited.<sup>6</sup> Only a few studies, such as those by Ormerod and Mounfield (2002) and Ormerod (2008), have employed that technique to investigate the phenomenon of business cycle synchronization over time.

Define  $\{\lambda_i\}_{i=1}^{i=N}$  to be the eigenvalues of the correlation matrix  $C$  and consider the probability density function of these eigenvalues

$$\rho_C(\lambda) = \frac{dn(\lambda)}{d(\lambda)}, \quad (5)$$

where  $n(\lambda)$  is the number of eigenvalues of  $C$  less than  $\lambda$ .

According to RMT, if all  $X_{it} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$ , for  $N, T \rightarrow \infty$  and  $Q = \frac{T}{N} \rightarrow a = \text{constant} > 1$ , the probability density function  $\rho_C(\lambda)$  of eigenvalue  $\lambda$  will follow the so-called Marchenko-

<sup>5</sup> One can also use other similarity measures such as the one proposed in Münnix et al. (2010).

<sup>6</sup> We suggest the readers to, for instance, Bouchaud and Potters (2009) for a more detailed review of the financial applications of RMT.

Pastur (M-P) law

$$\rho_C(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_{max}^{RMT} - \lambda)(\lambda - \lambda_{min}^{RMT})}}{\lambda}, \text{ for } \lambda_{min}^{RMT} \leq \lambda \leq \lambda_{max}^{RMT}, \quad (6)$$

and  $\rho_C(\lambda) = 0$  elsewhere, with  $\lambda_{max}^{RMT}$  and  $\lambda_{min}^{RMT}$  being, respectively, the upper and lower bounds of eigenvalues associated with a random correlation matrix with the variance  $\sigma^2$  and  $Q = \frac{T}{N}$ . According to RMT these bounds are given by

$$\lambda_{max}^{RMT} = \sigma^2(1 + \sqrt{1/Q})^2, \text{ and } \lambda_{min}^{RMT} = \sigma^2(1 - \sqrt{1/Q})^2. \quad (7)$$

We are interested in the latent information encoded in the eigenvectors corresponding to the largest eigenvalues deviating from the bulk of eigenvalues associated with a random correlation matrix with the same variance and the same  $Q$ . Suppose  $\lambda_1 > \dots > \lambda_k > \lambda_{max}^{RMT} > \dots > \lambda_N$  are the eigenvalues of the empirical correlation matrix  $C$  in descending order and their corresponding eigenvectors are  $u_1, u_2, \dots, u_N$ . Then, eigenvalues outside the support of the M-P law would indicate factors of significant correlation across the country specific indices as these eigenvalues lie outside the range of observations that could be attributed to random sampling. As a limit law, the M-P result applies to broad classes of random matrices. In particular, it applies to general matrices with identically and independently drawn entries with variance 1 for which all the moments are finite. It also holds asymptotically for random variables with fat tails (hyperbolic decay of their density functions) as long as the fourth and higher moments exist. However, preasymptotically, for finite  $N$  and  $T$ , large realizations could lead to spurious large eigenvalues outside the M-P spectrum (cf. Birole et al. (2007) and literature therein). In the Appendix, we provide additional statistical results that indicate that such spurious eigenvalues do not occur in our data. The elements of the eigenvector  $u_1$  can be interpreted as the effect of the strongest common factor (extracted from the correlation matrix) on all country-specific indices (see, for example, Plerou et al., 2002). In the following, we will be referring to this factor as the “market mode” or the “market factor”. In our study, we will investigate the temporal dynamics of the largest eigenvalues (larger than  $\lambda_{max}^{RMT}$ ) and their corresponding eigenvectors, in order to identify periods with distinct cross-country sentiment correlation structures, as well as, to quantify the systemic risk associated with these periods (see, for example, Billio et al., 2012; Zheng et al., 2012; Meng et al., 2014; Nobi and Lee, 2016).

### Decomposition and noise filtering



Note that we can diagonalize the correlation matrix  $C$  as

$$C = U\Lambda U^T, \quad (8)$$

where  $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_N\}$  and the matrix  $\{U\}_{N \times N}$  is orthonormal, whose  $i^{\text{th}}$  column is the normalized eigenvector  $u_i$  associated with  $\lambda_i$ . From Eq. (8) we have

$$\lambda_i = u_i^T C u_i = u_i^T \text{Cov}(X_t) u_i = \text{Var}(u_i^T X_t). \quad (9)$$

The total variance of  $X_t$  is then

$$\sum_{i=1}^N \text{Var}(X_{i,t}) = N = \sum_{i=1}^N \lambda_i = \sum_{i=1}^N \text{Var}(u_i^T X_t). \quad (10)$$

Now we can see that  $\lambda_i$  indicates the portion of total variance of  $X_t$  contributed by the principal component  $y_{i,t} = u_i^T X_t$  (e.g. Jolliffe, 1986).

We can rewrite Eq. (8) as

$$C = \sum_{i=1}^{i=N} \lambda_i u_i u_i^T. \quad (11)$$

The expression  $C^m = \lambda_1 u_1 u_1^T$  represents the part of the sentiment correlation structure accounted for by the market mode (recall that the eigenvalues are indexed in descending order). We can filter the market mode away from  $C$ . Following Kim and Jeong (2005), we define the filtered correlation matrix

$$M = C - C^m. \quad (12)$$

From Eq. (11), we can express  $M$  in the following way

$$M = C - \lambda_1 u_1 u_1^T = \sum_{i=2}^{i=N} \lambda_i u_i u_i^T. \quad (13)$$

The latent information encoded in the eigenvectors of the second largest eigenvalue can also be useful if it is still large enough not to fall within the random bulk (i.e. if  $\lambda_2 > \lambda_{max}^{RMT}$ ). In general, information embedded in other eigenvalues larger than  $\lambda_{max}^{RMT}$  is associated with important factors other than the market mode. In that case, the correlation matrix can be decomposed as

$$C = C^m + C^g + [C - C^m + C^g], \quad (14)$$

where  $C^g$  accounts for correlations captured by the second most important factor. For instance, in the analysis of stock markets, it has been repeatedly suggested that the sectoral

component can be captured by the eigenvectors associated with the second largest eigenvalues. In our study, we can think of the cultural and economic peculiarities of particular countries or groups of countries (e.g. emerging markets) as being such a factor.

### Absorption ratio

From Eq. (8) and Eq. (9), the absorption ratios are given by

$$E_i = \sum_{j=1}^i \lambda_j / N \quad (\text{for } i = 1, 2, \dots, N). \quad (15)$$

$E_i$  represents the fraction of the total variance of  $X_t$  explained by the first  $i$  principal components (since  $\sum_{j=1}^N \lambda_j / N = \frac{\text{trace}(C)}{N} = 1$ , we always have  $E_N = 1$ ). What we are interested in, are comparisons between  $E_1, \dots, E_k$  and  $E_N = 1$  (i.e.  $E_N$  stands for 100% of the variance as shown in Eq. (10)), where  $k$  is the largest integer for which  $\lambda_k > \lambda_{max}^{RMT}$  is true. Besides using the average of correlations, the absorption ratios can be used to infer the systemic risk in the market (see, for example, Pukthuanthong and Roll, 2009; Billio et al., 2012; Zheng et al., 2012; Meng et al., 2014). For instance, high values of  $E_1$  associated with a high  $\lambda_1$  signal a strong co-movement of the individual sentiment indices, which implies a high systemic risk.

### Inverse Partition Ratio

The inverse of the Inverse Partition Ratio (IPR) measures the number of eigenvector components (i.e. countries) strongly associated with a particular factor (the market mode, for example). It is defined as

$$\text{IPR}(i) = \sum_{j=1}^N u_i(j)^4. \quad (16)$$

Recalling that the elements of each eigenvector are normalized, i.e.  $\sum_{j=1}^N u_i(j)^2 = 1$  ( $\forall i = 1, 2, \dots, N$ ), it is easy to show that for all  $i = 1, 2, \dots, N$ , we have

$$\frac{1}{N} \leq \text{IPR}(i) \leq 1, \quad (17)$$

where  $\text{IPR}(i) = \frac{1}{N}$  if and only if  $u_i(j)^2 = \frac{1}{N}$  for all  $j = 1, 2, \dots, N$ ; while  $\text{IPR}(i) = 1$  if and only if only one element of the eigenvector  $u_i$  is different than zero, which implies that only this element (country) contributes to this particular factor. Overall, the inverse of the IPR indicates the number of eigenvector components that contribute significantly to that eigenvector. More specifically, a low IPR indicates that countries contribute more equally. In contrast, a large IPR would imply that the factor is driven by the dynamics of a small number of countries.

### III. FINDINGS

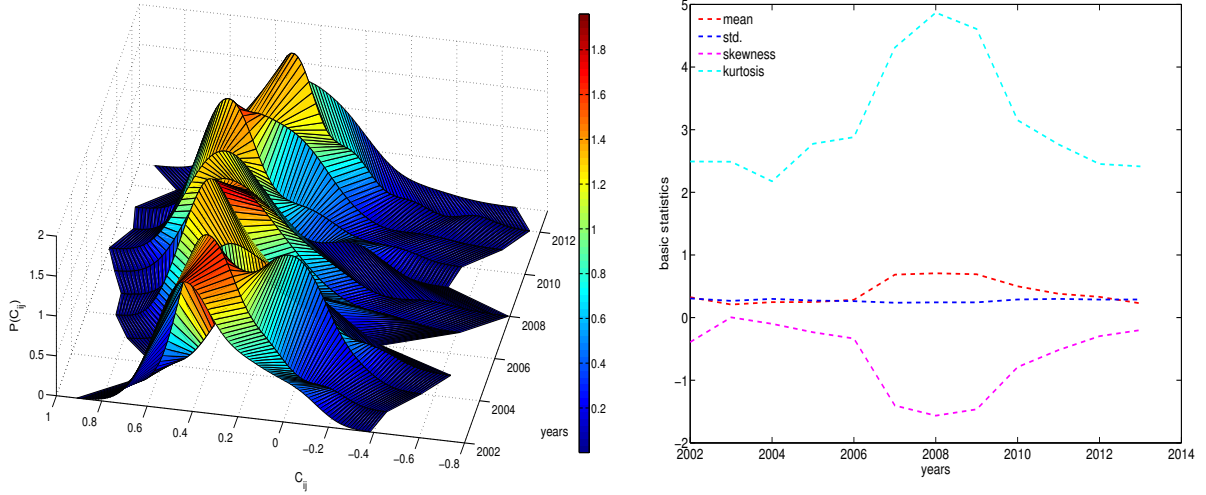
We report the temporal dynamics of the distribution of correlations between sentiment indices in Figure (1) for the BCI data (OECD<sup>+</sup> group) and in Figure (2) for the ESI data (EU<sup>+</sup> group).<sup>7</sup> Our first observation is that the distribution of  $C_{ij}$  is generally asymmetric, and its shape is not stable over time. A noticeable change can be easily detected for the period of the financial crisis (2007 to 2009). More specifically, the average of correlations and the kurtosis increase during that time, while the skewness decreases significantly. In addition, we find that for all years, in the case of the BCI data, the average of correlations is always positive and the distribution always is left-skewed, signaling that the mass of the distribution of correlations is concentrated on the positive side. This implies that, overall, countries tend to be more positively than negatively correlated (see, for example, Plerou et al., 2002). A similar observation can be made for the ESI data, except for several years, during which a positive skewness is observed (in particular, around the period when the Eurozone was implemented). We provide the following potential explanation for the increased number of negative correlations during that period. Before the introduction of the Euro the interest rates in the “periphery” (Spain, Italy, Portugal, etc.) were much higher than those in the “core” (Germany, Netherlands, France, etc.) of the monetary union. Thus, during the implementation of the Eurozone, the sentiment in the “periphery” of the union might have been positively affected by the convergence of the interest rates across the Eurozone members, while the effect on the sentiment in the “core” might have been negative.<sup>8</sup> In other words, the convergence implies an increase of interest rates in the “core” and a fall in the “periphery”, which might result in opposing sentiment dynamics in the two areas of the union. In this rare case, the global information signal might have had implications that differ across countries.

Since we observe that the sentiment correlation matrix is not stable over time, the question of how to identify the different states of  $C$  comes to the fore. In the previous section, we have introduced a method to quantify the similarity between correlation matrices (see Eq. (4)). This method allows us to identify particular states of the sentiment correlation matrix.

---

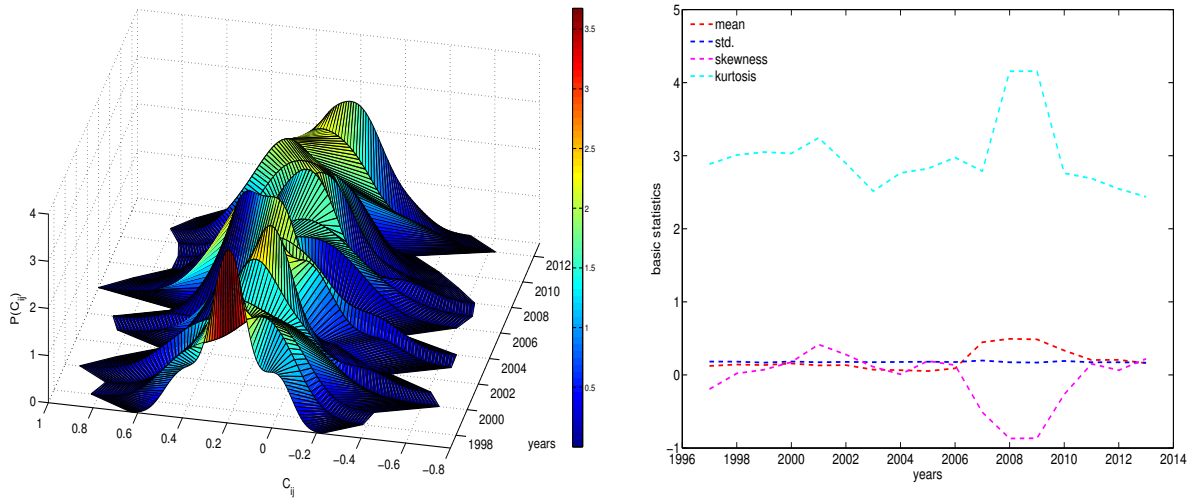
<sup>7</sup> The entries of the correlation matrix for each year have simply been pooled, after which a kernel density estimator has been used to arrive at a distribution in a particular year.

<sup>8</sup> For a discussion of the interest rate convergence in Eurozone see, for example, Arghyrou et al., 2009.



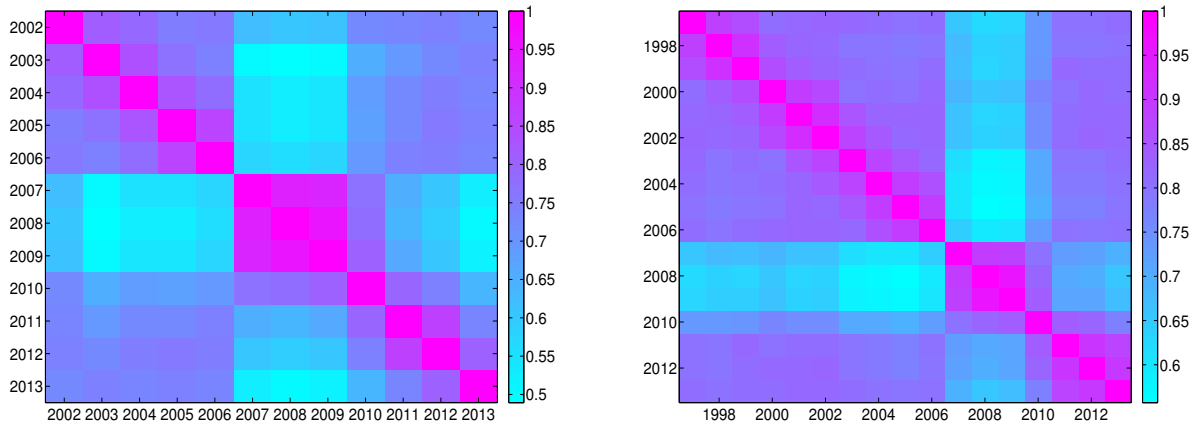
(a) Distribution of  $C$ , BCI data, OECD<sup>+</sup> group      (b) Statistics of  $C$ , BCI data, OECD<sup>+</sup> group

FIG. 1: Evolution of the distribution of the elements of  $C$ , for BCI data in the OECD<sup>+</sup> group. Panel (a) shows the distribution of the elements of  $C$  from 2002 to 2013. Panel (b) reports the basic statistical indicators of the elements of  $C$  including the mean, standard deviation (std.), skewness, and kurtosis.



(a) Distribution of  $C$ , ESI data, EU<sup>+</sup> group      (b) Statistics of  $C$ , ESI data, EU<sup>+</sup> group

FIG. 2: Evolution of the distribution of the elements of  $C$ , for ESI data in the EU<sup>+</sup> group. Panel (a) shows the distribution of the elements of  $C$  from 1997 to 2013. Panel (b) reports the basic statistical indicators of the elements of  $C$  including the mean, standard deviation, skewness, and kurtosis.



(a) Similarity among C, BCI data, OECD<sup>+</sup> group (b) Similarity among C, ESI data, EU<sup>+</sup> group

FIG. 3: Identifying states of correlation matrix using similarity-based analysis. Panel (a) shows the similarity among correlation matrices C from 2002 to 2013, for BCI data in the OECD<sup>+</sup> group. Panel (b) shows the similarity among correlation matrices C from 1997 to 2013, for ESI data in the EU<sup>+</sup> group.

Figure (3) shows the similarity between the temporal correlation matrices for the BCI data and for the ESI data, respectively in panels (a) and (b). Three states can be identified from these panels, i.e. before 2007, from 2007 to 2009, and after 2009. We can see that the period from 2007 to 2009 is very homogeneous with respect to the correlation structures observed. The correlation matrices from the years 2007, 2008 and 2009 are very similar to each other compared to matrices from other periods. This is consistent with what we have found during our analysis of the evolution of the basic statistics of the correlation matrices over time. In addition, in panel (b) we see that the correlation matrices for the EU<sup>+</sup> group associated with the period of the European debt crisis (2011 to 2013) exhibit strong similarities relative to correlation matrices from other periods. We can thus conclude that for the EU<sup>+</sup> we can detect an additional distinct state of the correlation matrix associated with the time of the debt crisis in Europe. In the following, we are going to look more closely at what these distinct states are characterized by.

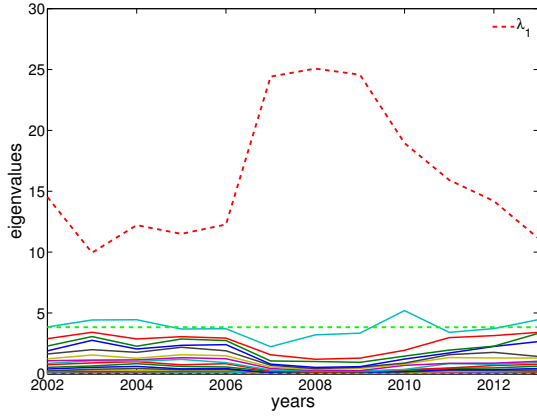
We start by investigating the spectrum of the correlation matrix and its evolution over time. In Figure (4) and Figure (5), panel (a) we see that the largest eigenvalue  $\lambda_1$  is typically more than three times larger than the upper bound  $\lambda_{max}^{RMT}$  for the OECD<sup>+</sup> group, and more than 1.3 times larger for the EU<sup>+</sup> group. In all years,  $\lambda_1$  always deviates from the random

bulk associated with the M-P law. Figure (6) shows the distribution of the eigenvalues of a random correlation matrix (with the same variance and the same  $Q$ ) compared to the actual distribution of the eigenvalues in 2007 for both groups of countries. For the EU<sup>+</sup> group, during the whole sample period, only the first eigenvalue  $\lambda_1$  is larger than the upper bound  $\lambda_{\max}^{RMT}$ , while for the OECD<sup>+</sup> group, in some years, a second eigenvalue  $\lambda_2$  crosses this upper bound.<sup>9</sup> The second factor may be interpreted as a group factor. On some rare occasions (e.g. like in the years 2003, 2004 and 2010), particular countries (including some “emerging markets”) can have sentiment dynamics opposing those of the rest of the world. We can detect this by looking at the eigenvector elements associated with countries like Mexico, Turkey, Slovakia, Russia, China in some years and recognizing that these elements have the opposite sign of the elements associated with countries from the rest of the world. Some developed countries like New Zealand and Australia for instance, also show a similar behavior. All this suggests that the cross-country sentiment dynamics are driven primarily by a single factor (the market mode) and only on rare occasions does a second factor become marginally significant. In the following, we will thus be concentrating on the market mode and on the relationship that countries or groups of countries have with it.

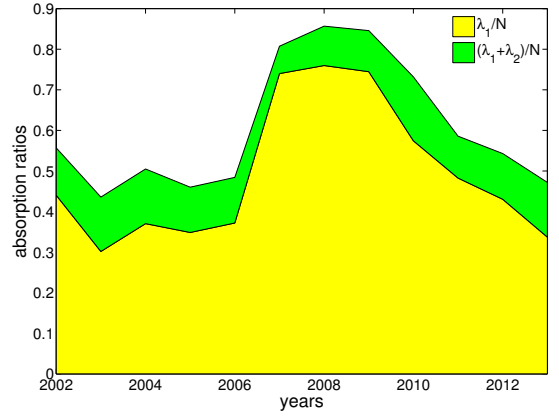
The evolution of the importance of the market factor for the cross-country sentiment dynamics can also be observed in Figure (4) and Figure (5). We see that during the financial crisis the importance of the market factor becomes overwhelming since both the largest eigenvalue and the associated absorption ratio jump by approximately 100%. Since the largest eigenvalue and the associated absorption ratio increase together, we can say that the systemic worldwide component of sentiment was high during the period 2007 to 2009. We can also see that, for the EU<sup>+</sup> group, the absorption ratios after 2009 are still higher than during the period before the financial crisis. The perceived threat to the Eurozone’s stability stemming from the risk of sovereign default of some member states might have prevented individual sentiment indices in Europe from diverging from the market mode.

---

<sup>9</sup> We also observe that  $\lambda_1$  is always close to  $N\langle C_{ij} \rangle$  (where  $\langle C_{ij} \rangle$  stands for the pooled average of  $C$ ), which supports the presence of one common factor affecting all indices.

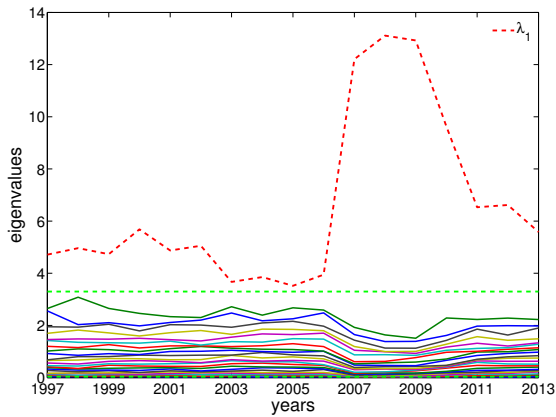


(a) Eigenvalues

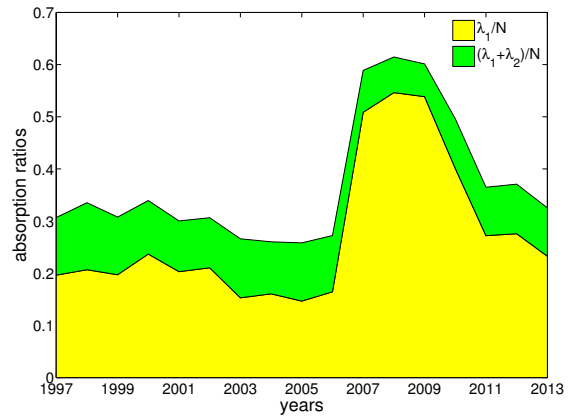


(b) Absorption ratios

FIG. 4: Evolution of eigenvalues and absorption ratios for BCI data in the OECD<sup>+</sup> group. In panel (a), the red dashed line stands for the largest eigenvalue, the green dashed lines stand for the interval  $[\lambda_{min}^{RMT}, \lambda_{max}^{RMT}]$  explained by RMT. Panel (b) shows the absorption ratios associated with the first and the second largest eigenvalues.

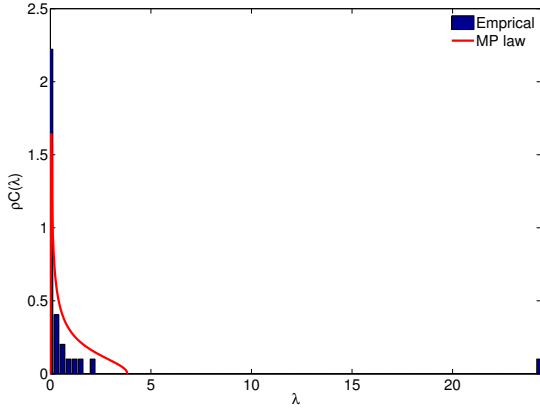


(a) Eigenvalues

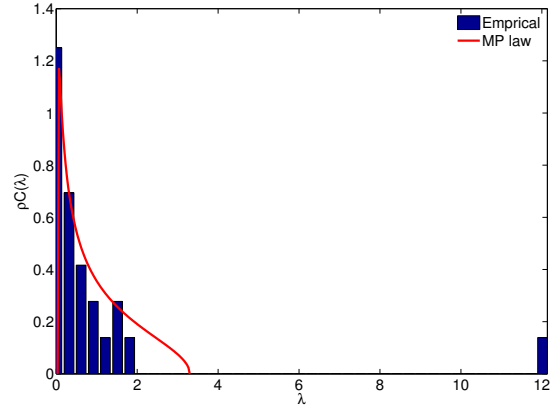


(b) Absorption ratios

FIG. 5: Evolution of eigenvalues and absorption ratios for ESI data in the EU<sup>+</sup> group. In panel (a), the red dashed line stands for the largest eigenvalue, the green dashed lines stand for interval  $[\lambda_{min}^{RMT}, \lambda_{max}^{RMT}]$  explained by RMT. Panel (b) shows the absorption ratios associated with the first and the second largest eigenvalues.



(a) Dist. of eigenvalues, BCI data



(b) Dist. of eigenvalues, ESI data

FIG. 6: Distribution of eigenvalues of  $C$  in 2007, compared with RMT. Panel (a) for BCI data in the  $OECD^+$  group. Panel (b) for ESI data in the  $EU^+$  group.

Next, let us look in more detail at the components of the eigenvector  $u_1$  associated with the market mode. We find that in the period of financial crisis (2007 to 2009), the components of  $u_1$  become more homogeneous, evidencing the synchronization of the sentiment indices around the world (see Figures (7), (8), and (9)). This result holds true for both the  $EU^+$  and  $OECD^+$  groups. Still, a few countries like China, South Africa or New Zealand seem to be somewhat less influenced by the market mode during the crisis. During normal times, much more divergent behavior is observed. More specifically, in case of the  $OECD^+$  group, countries like Italy, France, Belgium, Hungary, U.K., Austria, Slovenia, Denmark, Chile, Netherlands or Germany contribute the most to the market mode, while the sentiment dynamics in other countries like Finland, Korea, Mexico, New Zealand, Turkey, Slovakia, Australia, Russia, China or India can be divergent to a certain extent. For the  $EU^+$  group, countries like Belgium, Denmark, Germany, France, Italy, Netherlands or Sweden contribute the most to the market mode, while the sentiment dynamics in other countries like Bulgaria, Latvia, Poland, Romania, Slovakia or Finland can be divergent to a certain extent.



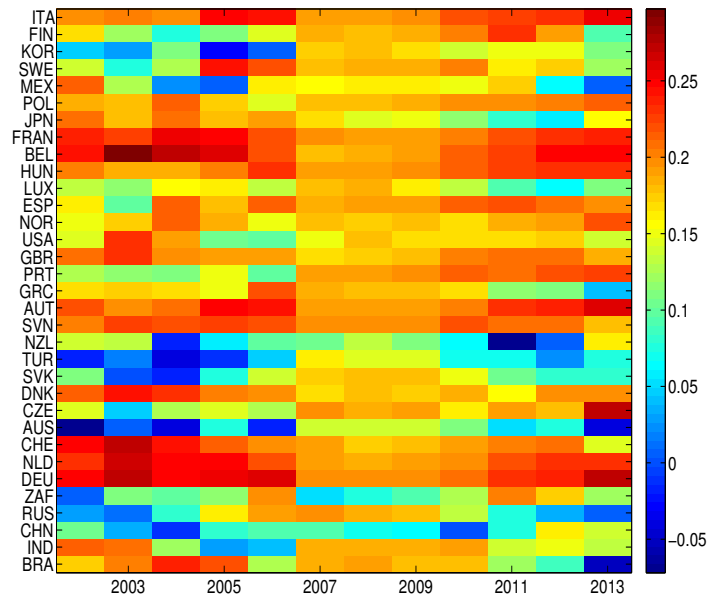


FIG. 7: Evolution of the eigenvector components of the largest eigenvalue ( $\lambda_1$ ), for BCI data in the OECD<sup>+</sup> group. Without loss of generality, we assume that the sign of the eigenvector element that has the largest absolute value is non-negative. The financial crisis from 2007 to 2009 is captured by drastic changes in the the largest eigenvector components, i.e. they become more homogeneous during that period.

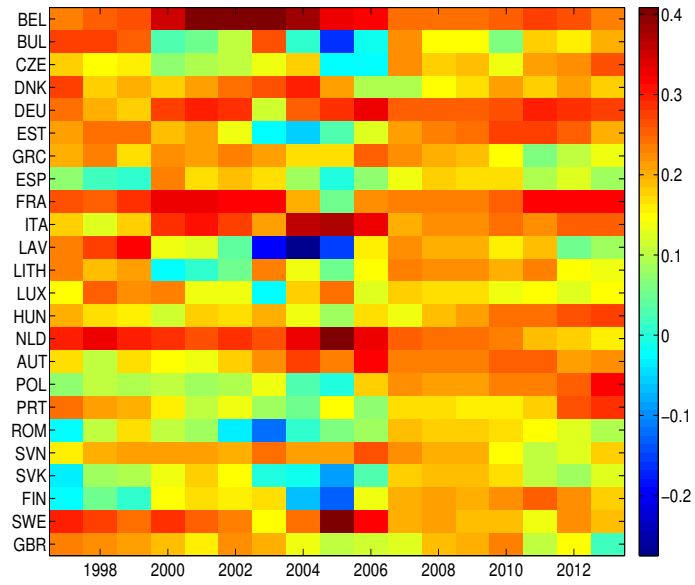
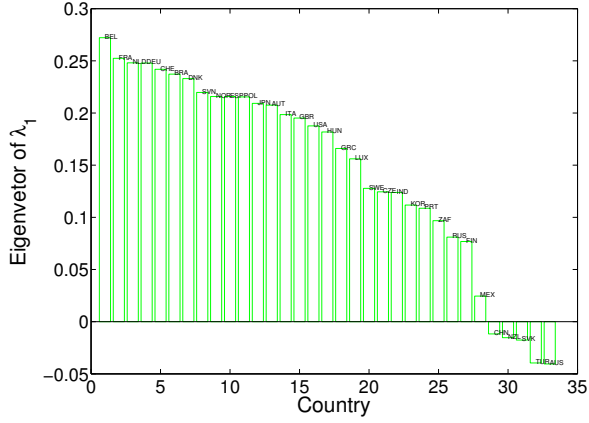
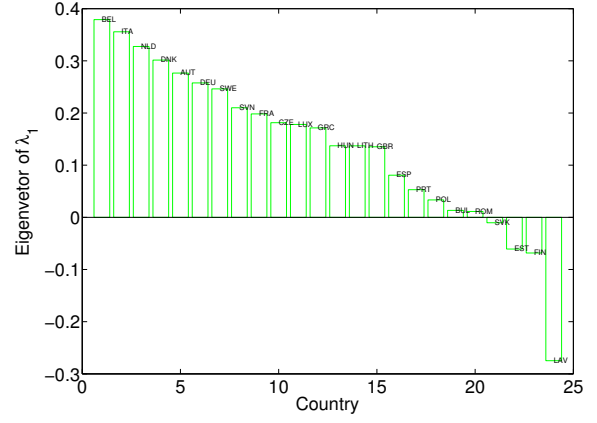


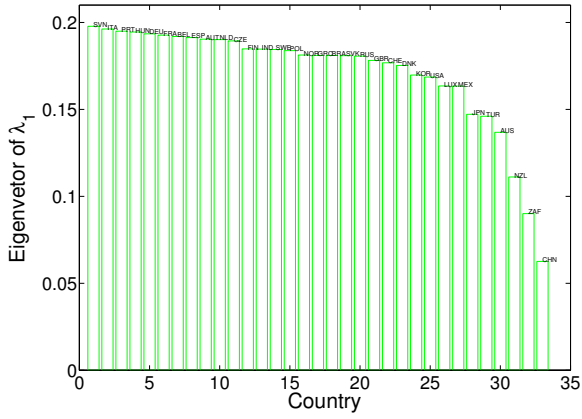
FIG. 8: Evolution of the eigenvector components of the largest eigenvalue ( $\lambda_1$ ), for ESI data in the EU<sup>+</sup> group. Without loss of generality, we assume that the sign of the eigenvector element that has the largest absolute value is non-negative. The financial crisis from 2007 to 2009 is captured by drastic changes in the the largest eigenvector components, i.e. they become more homogeneous during that period.



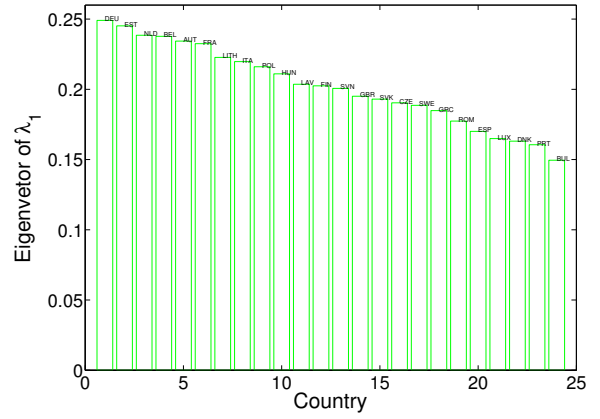
(a) Eivec1 in 2003, BCI



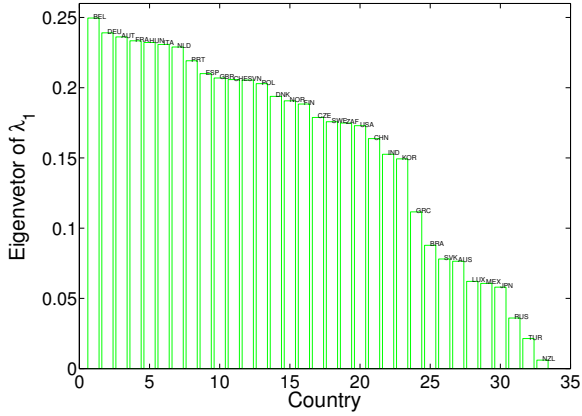
(b) Eivec1 in 2003, ESI



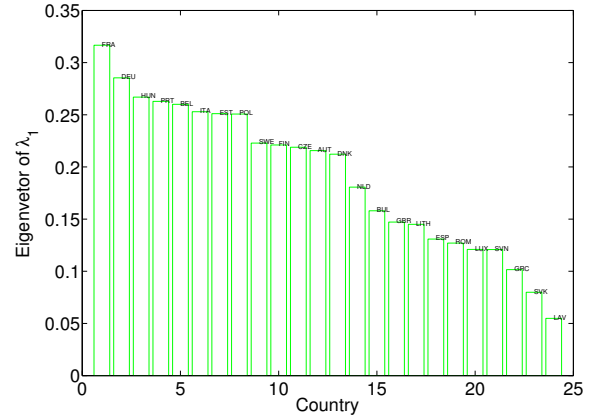
(c) Eivec1 in 2007, BCI



(d) Eivec1 in 2007, ESI



(e) Eivec1 in 2011, BCI



(f) Eivec1 in 2011, ESI

FIG. 9: Eigenvector components of  $\lambda_1$  (Eivec1). Without loss of generality, we assume that the sign of the eigenvector element that has the largest absolute value is non-negative. The three left panels (a), (c), and (e) are for BCI data in the OECD<sup>+</sup> group, in 2003, 2007, and 2011. The three right panels (b), (d), and (f) are for ESI data in the EU<sup>+</sup> group, in 2003, 2007, and 2011. The country code associated with each eigenvector component is also reported.

We further analyze the eigenvectors by calculating the IPR. Figure (10) shows the IPR versus the corresponding eigenvalues for the years 2003, 2007, and 2011 as examples. Panels (a) and (b) do this for the OECD<sup>+</sup> and for the EU<sup>+</sup> groups, respectively. We find that for those years, in which if the largest eigenvalue strongly deviates from the random bulk, the associated eigenvector also exhibits the largest inverse of IPR, meaning that the sentiment dynamics in the majority of the countries is influenced by the market mode.

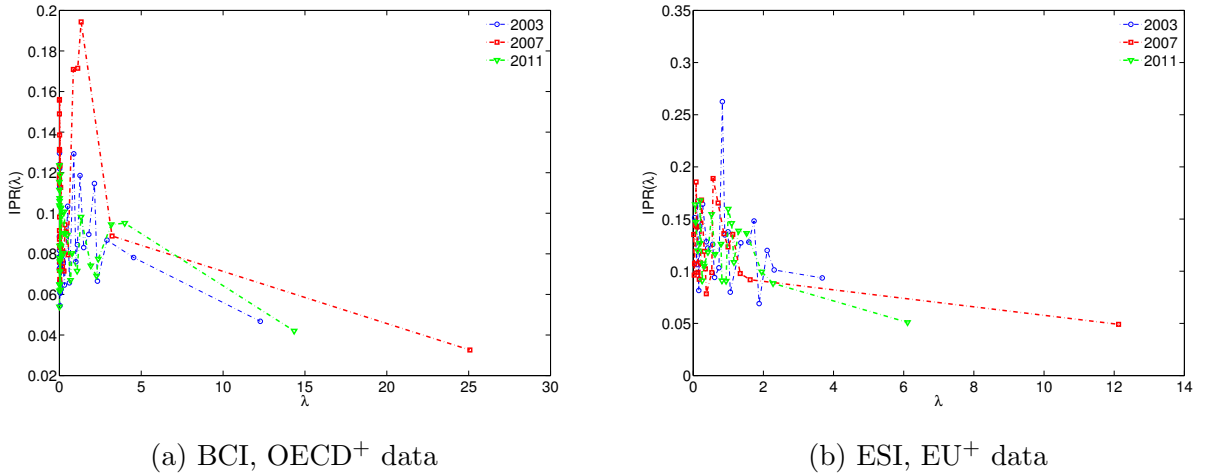
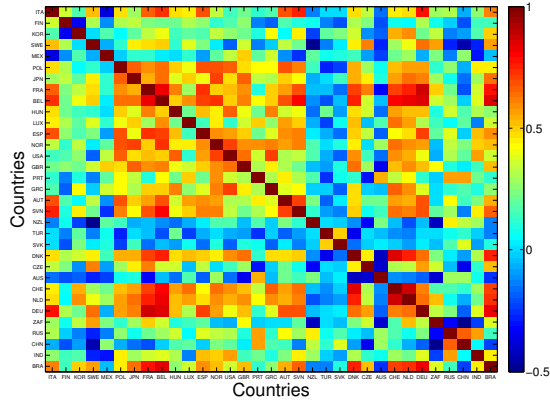


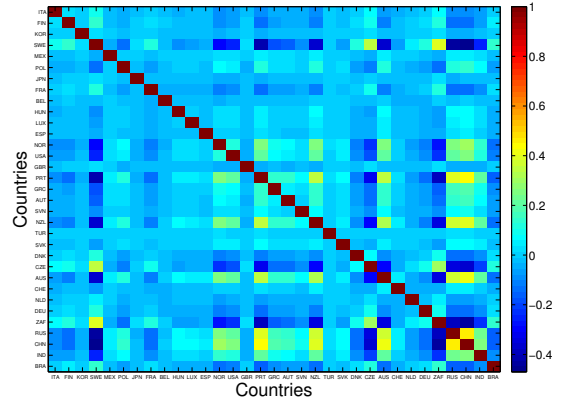
FIG. 10: Comparison of Inverse Participation Ratios of the eigenvectors of the correlation matrix for periods before the financial crisis, during the financial crisis, and after the financial crisis. Panel (a) shows the IPR for BCI in the OECD<sup>+</sup> group, panel (b) shows the IPR for ESI in the EU<sup>+</sup> group. In both panels, three years, (i) 2003, (ii) 2007, and (iii) 2011, are chosen as the examples.

Above, we have detected a common factor underlying the dynamics of the sentiment indices. Now, we will compare the sentiment correlation matrix over time before and after filtering the effect of that factor. The results are shown in Figures (11) and (12), respectively for the OECD<sup>+</sup> and EU<sup>+</sup> groups. Overall, the raw correlations are significantly reduced after the information encoded in the largest eigenvalue and its corresponding eigenvector is subtracted. In addition, we can see that at the time of financial crisis (exemplified here by the graph for 2007, which is the same for 2008 and 2009), the raw correlations between countries increase but their filtered counterparts exhibit reduced correlations. In other words, the increase in the raw correlations were accompanied by an increase in the fraction of the correlations associated with the market mode. Note that for the OECD<sup>+</sup> group, some significant correlations still appear after the filtering in some years of the sample period.

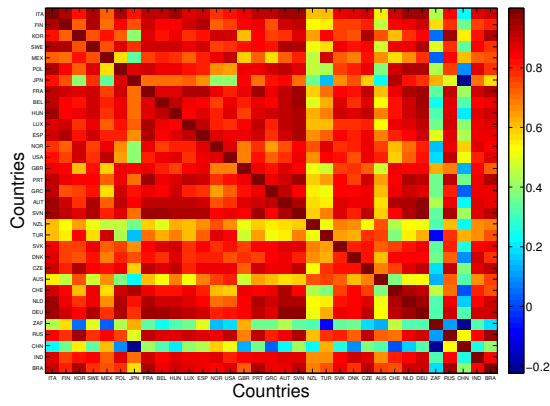
They are actually mainly contributed by emerging markets like China or India, for which the informational signal associated with the collapse of the US housing market might have been less relevant due to their limited exposure to toxic securities.



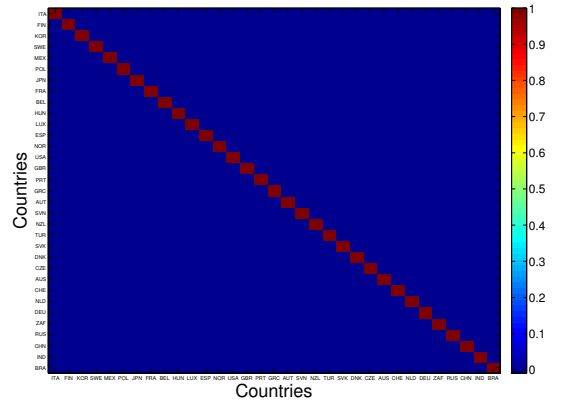
(a) C in 2003



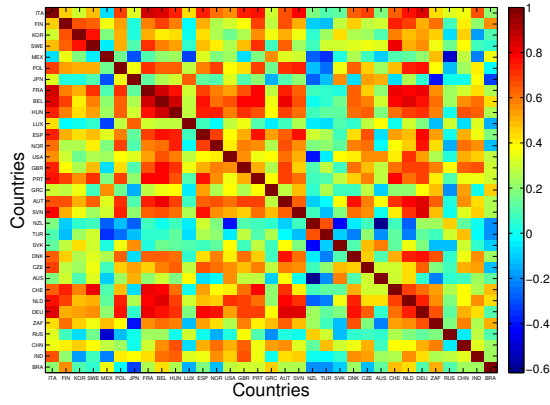
(b) M in 2003



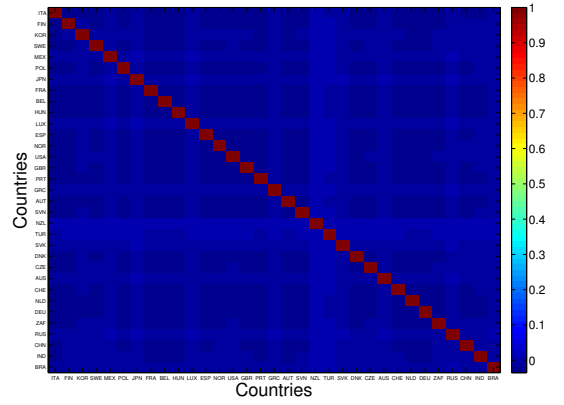
(c) C in 2007



(d) M in 2007

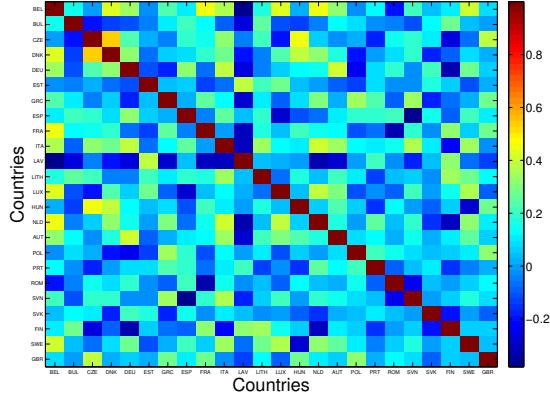


(e) C in 2011

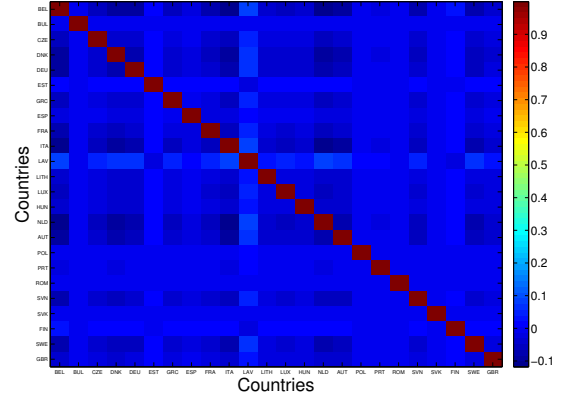


(f) M in 2011

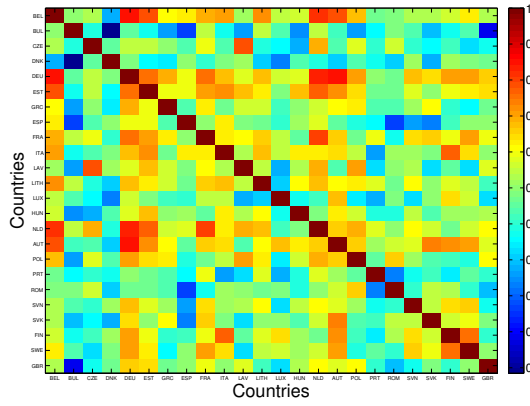
FIG. 11: Correlations between countries, for BCI data in the OECD<sup>+</sup> group. The three left panels (a), (c), and (e) are the raw correlation matrices in 2003, 2007, and 2011. The three right panels (b), (d), and (f) are the correlation matrices filtered by the RMT method in 2003, 2007, and 2011.



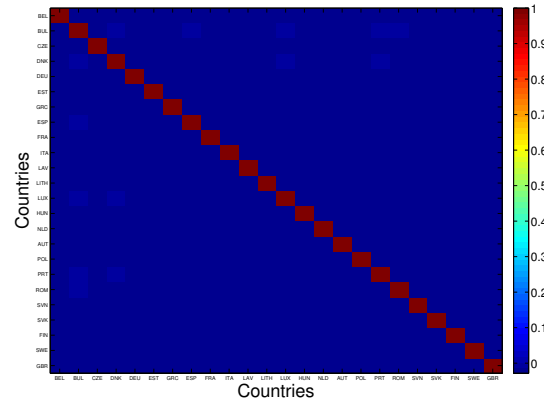
(a) C in 2003



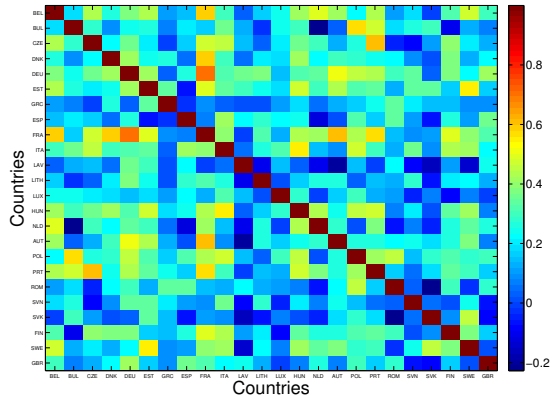
(b) M in 2003



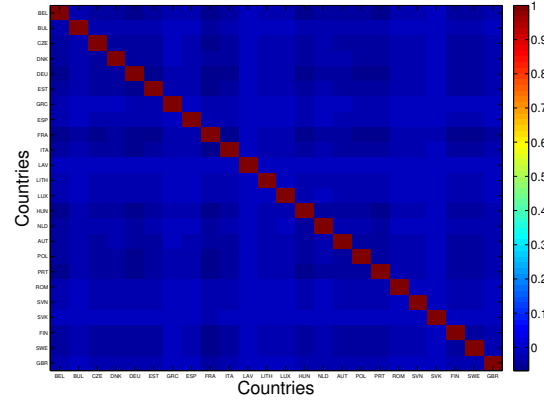
(c) C in 2007



(d) M in 2007



(e) C in 2011



(f) M in 2011

FIG. 12: Correlations between countries, for ESI data in the  $EU^+$  group. The three left panels (a), (c), and (e) are the raw correlation matrices in 2003, 2007, and 2011. The three right panels (b), (d), and (f) are the correlation matrices filtered by the RMT method in 2003, 2007, and 2011.

## IV. CONCLUSIONS

In this paper, we have analyzed the evolution of the empirical correlations between the macroeconomic sentiment indices in different countries. Overall, we observe different states in the correlation structure, associated with a varying importance of the common factor (“market mode”). The correlations between indices are significantly reduced after the effects of that common factor are subtracted.

It would seem that many of the information signals worldwide have a common component, since generally sentiment indices tend to comove. During normal times, however, the sentiment in various countries or groups of countries can “resist” the common factor or can even, on rare occasions, “swim against the tide”. This is the case for some emerging markets like China, Turkey and other countries like Australia, New Zealand. However, in the presence of strong global information signals, we observe a strong synchronization of the sentiment dynamics all over the world. We consider the collapse of the US housing market (2007-2009) as an example of such global signals. In the case of the Eurozone debt crisis (2011-2013), the sentiment synchronization is high only within Europe, which can be interpreted as an indication that the Eurozone debt crisis is not perceived as a global information signal in countries outside of Europe. Information signals can also cause the sentiments to diverge, if the information has different implications for particular countries or groups of countries. We consider the interest rate convergence associated with the establishment of the Eurozone around the year 2000 to be an example of such an effect.

Overall, we believe that RMT and principal component analysis of the ensemble of worldwide or regional sentiment data can reveal important information on the correlations between business and consumer sentiment in different countries. The tools and results presented in this paper should provide relevant input for business cycle forecasts and the analysis of international co-movements of macroeconomic activity.

## V. REFERENCES

- Akerlof G. A., Shiller R. J. 2009. *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism*. Princeton University Press.
- Argyrou M. G., Gregoriou A. and Kontonikas A. 2009. Do real interest rates converge?



Evidence from the European union. *Journal of International Financial Markets, Institutions and Money* 19 (3), pp. 447- 460.

Artis M., Chouliarakis G., Harischandra P. K. G. 2011. Business cycle synchronization since 1880. *The Manchester School* 79 (2), pp. 173-207.

Billio M., Getmansky M., Lo A. W., Pelizzon L. 2012. Econometric measures of systemic risk in the finance and insurance sectors. *Journal of Financial Economics* 104 (3), pp. 535-559.

Biroli G., Bouchaud J. -P., Potters M. 2007. On the top eigenvalue of heavy-tailed random matrices. *Europhysics Letters*, 78 (1).

Bouchaud J. -P., Potters M. 2009. Financial applications of random matrix theory: a short review. Working Paper. Available at: [arXiv:0910.1205](https://arxiv.org/abs/0910.1205).

Bordo M. D., Helbling T. 2003. Have national business cycles become more synchronized?. NBER Working Paper 10130. Available at: <http://www.nber.org/papers/w10130>.

Bordo M. D., Helbling T. 2011. International business cycle synchronization in historical perspective. *The Manchester School* 79 (2), pp. 208-238.

Carriero A., Marcellino M. 2011. Sectoral survey-based confidence indicators for Europe. *Oxford Bulletin of Economics and Statistics* 73 (2), pp. 175-206.

EC. 2016. Economic sentiment index (indicator). Available at: [http://ec.europa.eu/economy\\_finance/db\\_indicators/surveys/index\\_en.htm](http://ec.europa.eu/economy_finance/db_indicators/surveys/index_en.htm) (Accessed on 29 January 2016).

Ferroni F., Klaus B. 2015. Euro area business cycles in turbulent times: Convergence or decoupling. ECB Working Paper 1819. Available at: <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1819.en.pdf?8aacf9b049f3a7e6360b69d13e435844>.

Hill B. M. 1975. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* (3), pp.1163-1174

Howrey E. P. 2001. The predictive power of the index of consumer sentiment. *Brookings Papers on Economic Activity* 2001 (1), pp. 175-207.

Jiang X. F., Chen T. T., Zheng B. 2014. Structure of local interactions in complex financial dynamics. *Scientific Reports* 4.

Jolliffe I.T. 1986. *Principal component analysis*. Springer-Verlag, New York.

Keynes J. M. 1936. *The general theory of employment, interest and money*. New York:

Harcourt, Brace & World.

Kim D. -H., Jeong H. 2005. Systematic analysis of group identification in stock markets. *Physical Review E* 72 (4).

Kose M. A., Otrok C., Prasad E. 2012. Global business cycles: Convergence or decoupling?. *International Economic Review* 53 (2), pp. 511-538.

Laloux L., Cizeau P., Bouchaud J. -P., Potters M. 1999. Noise dressing of financial correlation matrices. *Physical Review Letters* 83 (7), pp.1467-1470.

Laloux L., Cizeau P., Potters M., Bouchaud J. -P. 2000. Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance* 3 (3), pp. 391-397.

MacMahon M., Garlaschelli D. 2015. Community detection for correlation matrices. *Physical Review X* 5 (2).

Milani F. 2011. Expectation shocks and learning as drivers of the business cycle. *The Economic Journal* 121 (552), pp. 379-401.

Milani F. 2014. Sentiment and the U.S. business cycle. Working Paper. Available at: <http://www.economics.uci.edu/files/docs/workingpapers/2014-15/14-15-04-1.pdf>.

Minsky H. 1977. The financial instability hypothesis: An interpretation of Keynes and an alternative to standard theory. *Nebraska Journal of Economics and Business* 16 (1), pp. 5-16.

Meng H., Xie W. -J., Jiang Z. -Q., Podobnik B., Zhou W. -X., Stanley H. E. 2014. Systemic risk and spatiotemporal dynamics of the US housing market. *Scientific Reports* 4.

Münnix M. C., Shimada T., Schäfer R., Leyvraz F., Seligman T. H., Guhr T., Stanley H. E. 2012. Identifying states of a financial market. *Scientific Reports* 2.

Münnix M. C., Schäfer R., Grothe O. 2010. Estimating correlation and covariance matrices by weighting of market similarity. Working Paper. Available at: [arXiv:1006.5847](https://arxiv.org/abs/1006.5847).

Nobi A., Lee J. W. 2016. State and group dynamics of world stock market by principal component analysis. *Physica A: Statistical Mechanics and its Applications* 450, pp. 85-94.

OECD. 2016. Business confidence index (indicator). Available at: <https://data.oecd.org/leadind/business-confidence-index-bci.htm#indicator-chart> (Accessed on 29 January 2016).

Ormerod P., Mounfield C. 2002. The convergence of European business cycles 1978-2000.

Physica A: Statistical Mechanics and its Applications 307 (3-4), pp. 494-504.

Ormerod P. 2008. Random matrix theory and macro-economic time-series: An Illustration using the evolution of business cycle synchronisation, 18862006. Economics: The Open-Access, Open-Assessment E-Journal 2, pp. 1-10.

Plerou V., Gopikrishnan P., Rosenow B., Amaral L. A. N., Guhr T., Stanley H. E. 2002. Random matrix approach to cross correlations in financial data. Physical Review E 65 (6).

Pukthuanthong K., Roll R. 2009. Global market integration: An alternative measure and its application. Journal of Financial Economics 94 (2), pp. 214-232.

Santero T., Westerlund N. 1996. Confidence indicators and their relationship to changes in economic activity. OECD Economics Department Working Papers, No. 170, OECD Publishing. Available at: <http://dx.doi.org/10.1787/537052766455>.

Taylor K., McNabb R. 2007. Business cycles and the role of confidence: Evidence for Europe. Oxford Bulletin of Economics and Statistics 69 (2), pp. 185-208.

Uechi L., Akutsu T., Stanley H. E., Marcus A. J., Kenett D. Y. 2015. Sector dominance ratio analysis of financial markets. Physica A: Statistical Mechanics and its Applications 421, pp. 488509.

van Aarle B., Kappler M. 2012. Economic sentiment shocks and fluctuations in economic activity in the euro area and the USA. Intereconomics 47 (1), pp. 44-51.

Wang D., Podobnik B., Horvatic D., Stanley H. E. 2011. Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices. Physical Review E 83 (4).

Zheng Z., Podobnik B., Feng L., Li B. 2012. Changes in cross-correlations as an indicator for systemic risk. Scientific Reports 2.

## VI. APPENDIX

In this section we will examine what happens to the upper bound of the largest eigenvalue if the random variables  $X_{i,t}$  still have zero mean and unit variance but may have power law tails with exponent  $\mu$ . In such a case large changes in the sentiment indices may cause spurious apparent correlations and substantial overestimation of the (theoretical) largest eigenvalue of the sample correlation matrix  $C = \frac{1}{T}XX^T$  (e.g. see Biroli et al. (2007) and Bouchaud and Potters (2009)).

Denote  $S$  the largest element of  $|X_{i,t}|$  (for all  $i = 1, 2, \dots, N$  and all  $t$  in the considered time window of length  $T$ ). According to Biroli et al. (2007), whenever  $S \leq (NT)^{1/4}$ , the upper bound for the largest eigenvalue is  $\lambda_{max} = (1 + \sqrt{Q})^2$ , and when  $S > (NT)^{1/4}$ , the largest eigenvalue becomes  $\lambda_{max} = (1/Q + S/T^2)(1 + T/S^2)$ . In addition, if  $\mu > 4$ , the largest element of  $|X_{i,t}|$  is order of  $(NT)^\mu$ , and it is smaller than  $(NT)^{1/4}$ . In this case,  $\lambda_{max}^{RMT} = (1 + \sqrt{Q})^2$  can be still used as the upper bound for the largest eigenvalue of  $C$ .

To check whether these conditions hold for our data sets, we first use the standard Hill estimator (e.g. Hill (1975)), to examine the behavior of the exponent  $\mu$  in the tails of  $|X_{i,t}|$ . After that, we compare  $S$ , the largest value of  $X_{i,t}$  (in absolute terms), with  $(NT)^{1/4}$  to see whether we can still use  $\lambda_{max}^{RMT} = (1 + \sqrt{Q})^2$  as the upper bound of the eigenvalues of the null-model.

The Hill estimator for different lengths of the tail of  $|X_{i,t}|$  in terms of a percentage of the sample containing the largest observations in  $|X_{i,t}|$  (across all years and all countries in the respective groups) is shown in figure (13) for the ESI data and in figure (14) for the BCI data. We typically find that  $|X_{i,t}|$  have heavy tails with an exponent  $\mu$  in the range from 2.7 to 7. In addition, when considering only the top 1% to 5% largest observations,  $\mu$  is larger than 3.

Furthermore, in order to have a more comprehensive assessment on the behavior of the tail exponent, we decompose the entire observation period into separate time windows and then estimate  $\mu$  for each window. As shown in figure (15), with the top 5% largest elements of  $|X_{i,t}|$ ,  $\mu$  is larger than 4 for most of the windows. However, in the case of the ESI data (EU<sup>+</sup> group), for the time window associated with the time of the financial crisis (2007-2009), we observe that  $\mu$  is in the range 3–4, implying the presence of large fluctuations in sentiment indices during that time.

While the existence of the fourth moment of the theoretical distribution of  $|X_{i,t}|$  seems not to be guaranteed for all sub-samples, the crucial inequality  $S \leq (NT)^{1/4}$  holds for all considered time windows in both data sets (see Figure (16)). Taken together, these results suggest that the relevant limit for the (theoretical) largest eigenvalue of  $C$  remains  $\lambda_{max}^{RMT} = (1 + \sqrt{Q})^2$ .

Hill estimator for different percentages of sample size over the entire observation period

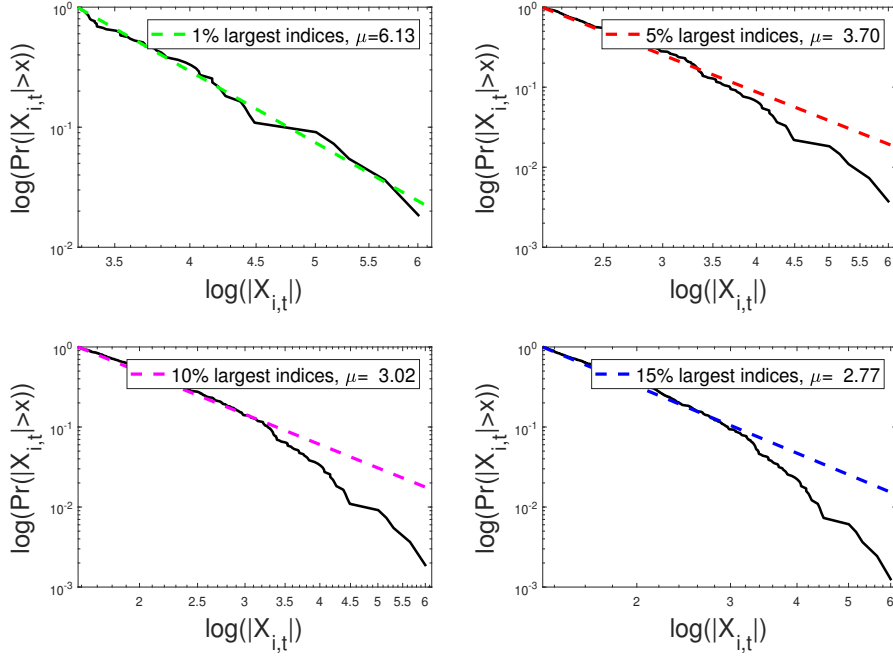


FIG. 13: Hill estimates of the tail exponent of the distribution of  $|X_{i,t}|$  based on different lengths of the tail over the entire observation period for the ESI data (EU<sup>+</sup> group). Panels (a), (b), (c), and (d) show the results for tails defined as 1%, 5%, 10%, and 15% of the largest observations in  $|X_{i,t}|$ , respectively. In each panel, we plot the empirical complementary cumulative distribution function (CCDF) of  $|X_{i,t}|$  on a log-log scale. The solid line depicts the empirical CCDF, while the dashed line represents the power law  $\Pr(|X_{i,t}| > x) \sim x^{-\mu}$  using the respective Hill estimate for  $\mu$ .

Hill estimator for different percentages of sample size over the entire observation period

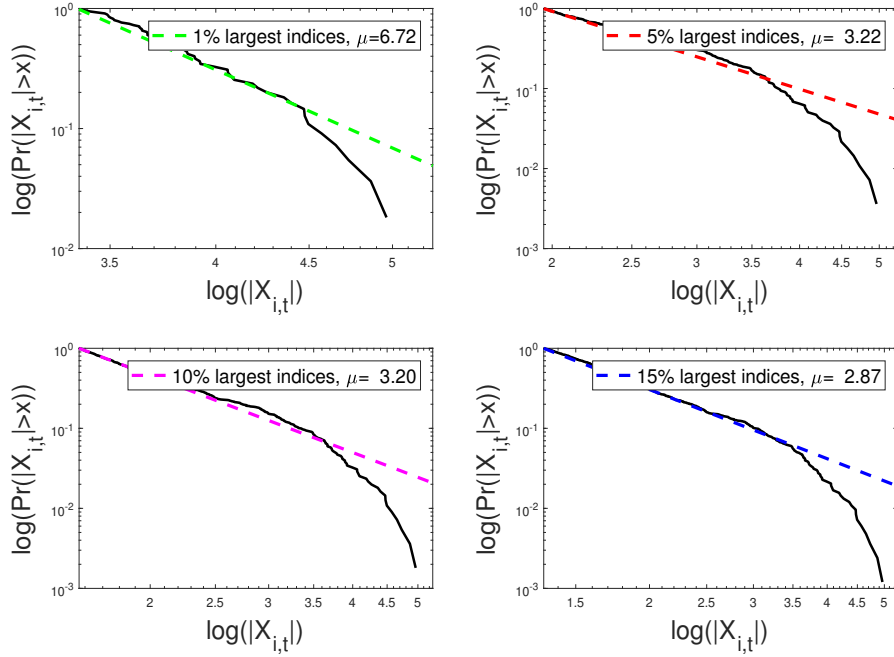
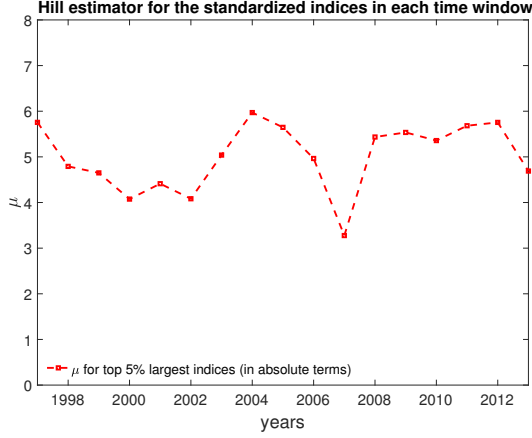
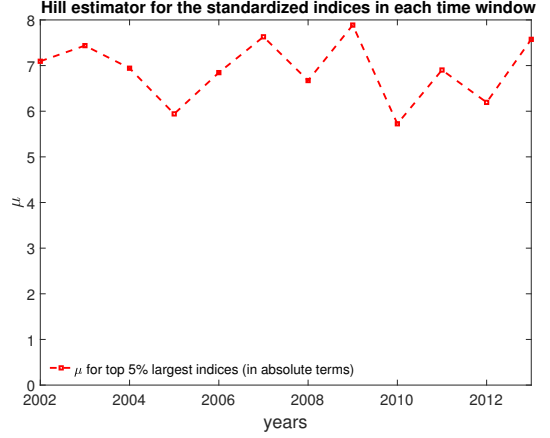


FIG. 14: Hill estimates of the tail exponent of the distribution of  $|X_{i,t}|$  based on different lengths of the tail over the entire observation period, for the BCI data (OECD<sup>+</sup> group). Panels (a), (b), (c), and (d) show the results for tails defined as 1%, 5%, 10%, and 15% of the largest observations in  $|X_{i,t}|$ , respectively. In each panel, we plot the empirical complementary cumulative distribution function (CCDF) of  $|X_{i,t}|$  on a log-log scale. The solid line depicts the empirical CCDF, while the dashed line represents the power law  $\Pr(|X_{i,t}| > x) \sim x^{-\mu}$  using the respective Hill estimate for  $\mu$ .

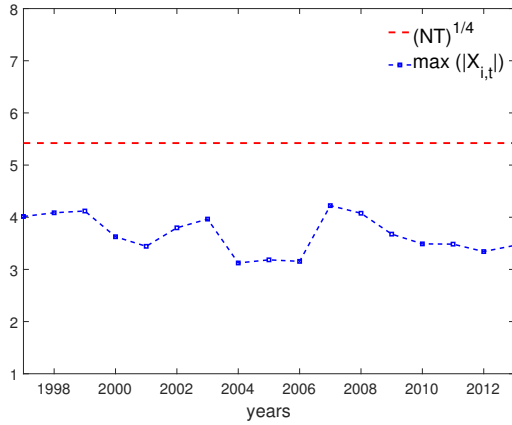


(a) for ESI data, EU<sup>+</sup> group

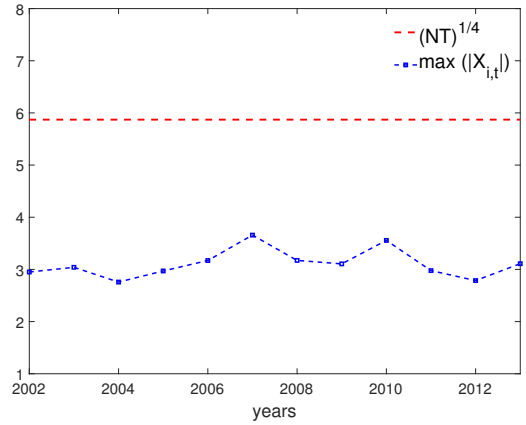


(b) for BCI data, OECD<sup>+</sup> group

FIG. 15: The tail exponent  $\mu$  of the distribution of  $|X_{i,t}|$  in each time window, computed using the Hill estimator on the 5% largest observations in the respective window.



(a) for ESI data, EU<sup>+</sup> group



(b) for BCI data, OECD<sup>+</sup> group

FIG. 16: The largest element of  $|X_{i,t}|$  compared to  $(NT)^{1/4}$  in each time window. We always observe that  $\max(|X_{i,t}|) < (NT)^{1/4}$ .