



OPEN

Loss of structural balance in stock markets

Eva Ferreira¹, Susan Orbe¹, Jone Ascorbebeitia², Brais Álvarez Pereira³ & Ernesto Estrada^{4,5,6}✉

We use rank correlations as distance functions to establish the interconnectivity between stock returns, building weighted signed networks for the stocks of seven European countries, the US and Japan. We establish the theoretical relationship between the level of balance in a network and stock predictability, studying its evolution from 2005 to the third quarter of 2020. We find a clear balance–unbalance transition for six of the nine countries, following the August 2011 Black Monday in the US, when the Economic Policy Uncertainty index for this country reached its highest monthly level before the COVID-19 crisis. This sudden loss of balance is mainly caused by a reorganization of the market networks triggered by a group of low capitalization stocks belonging to the non-financial sector. After the transition, the stocks of companies in these groups become all negatively correlated between them and with most of the rest of the stocks in the market. The implied change in the network topology is directly related to a decrease in stock predictability, a finding with novel important implications for asset allocation and portfolio hedging strategies.

The complex nature of the stock market is condensed in the phrase: “Economists are as perplexed as anyone by the behavior of the stock market”¹. Understanding this complex behavior—characterized by a combination of saw-tooth movements and low-frequency upward and downward switches²—is a major goal in finance. However, the importance of stock markets goes beyond finance to impact macroeconomic modeling and policy discussion³. In fact, it has been recognized that the stock market is a good predictor of the business cycle and of the components of gross national product of a given country³. Special emphasis has been placed on understanding the role of risk and uncertainty in stock markets from different perspectives^{4–16}, with a focus on security pricing and corporate investment decisions. At the end of the day, as said by Fischer and Merton³: “in the absence of uncertainty, much of what is interesting in finance disappears”.

As the majority of complex systems, stock markets have an exoskeleton formed by entities and complex interactions, which give rise to the observed patterns of the market behavior^{17–20} (see Ref.²¹ for a review of the recent research in financial networks and their practical applicability). For the same set of entities, there are different forms to define the connectivity between them. A typical way of connecting financial institutions is by means of borrowing/lending relations^{21–24}. However, in the case of stock markets, where stocks represent the nodes of the network, the inter-stock connectivity is intended to capture the mutual trends of stocks over given periods of time. Mantegna²⁵ proposed to quantify the degree of similarity between the synchronous time evolution of a pair of stock prices by the correlation coefficient ρ_{ij} between the two stocks i and j . Then, this correlation coefficient is transformed into a distance using: $d_{ij} = \sqrt{2(1 - \rho_{ij})}$. Mantegna’s approach has been widely extended in econophysics²⁶ where it is ubiquitous nowadays as a way to define the connectivity between stocks. This approach has been used for instance for the construction of networks to analyze national or international stock markets^{27–32}. A characteristic feature of this approach is that although $\rho_{ij} \in [-1, 1]$, the distance d_{ij} used as a weight of the links between stocks is nonnegative. This allows the use of classical network techniques for their analysis (see Ref.³³ for a review). To avoid certain loss of information inherent to the previous approach³⁴, Tse et al.³⁵ decide that an edge exists between a pair of stocks only if $\rho_{ij} > |z|$ for a given threshold z , which however retains the nonnegativity of the edge weights.

¹Department of Quantitative Methods, University of the Basque Country UPV/EHU, Avda. Lehendakari Aguirre 83, 48015 Bilbao, Spain. ²Department of Economic Analysis, University of the Basque Country UPV/EHU, Avda. Lehendakari Aguirre 83, 48015 Bilbao, Spain. ³Nova School of Business and Economics (Nova SBE) NOVAFRICA, and BELAB, Carcavelos, Portugal. ⁴Institute of Mathematics and Applications, University of Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain. ⁵ARAID Foundation, Government of Aragon, Zaragoza, Spain. ⁶Institute for Cross-Disciplinary Physics and Complex Systems (IFISC, UIB-CSIC) Campus Universitat de les Illes Balears, 07122 Palma de Mallorca, Spain. ✉email: estrada66@unizar.es

The importance of allowing explicit distinction between positive and negative interdependencies of stocks, which is lost in the previous approaches, was recently remarked by Stavroglou et al.³⁶ They built separated networks based on positive and negative interdependencies. This separation may still hide some important structural characteristics of the systems under analysis. Fortunately, there is an area of graph theory that allows studying networks with the simultaneous presence of positive and negative links. These graphs, known as signed graphs, are known since the end of the 1950s³⁷ (see Ref.³⁸ for a review), but only recently have emerged as an important tool in network theory^{39–43}. A focus on signed networks has been the theory of social balance^{37,44}. It basically states that social triads where the three edges are positive (all-friends) or where two are negative and one is positive—the enemy of my enemy should be my friend—are more abundant in social systems than those having all-negative, or two positive and one negative ones. The first kind of networks are known as balanced, and the second ones as unbalanced (see Refs.^{39–43} for applications).

Here we introduce the use of signed networks to represent stock markets of nine developed countries and to study the degree of balance. First note that in finance, the time varying behavior of networks is an important characteristic, since the time varying comovements become a risk factor⁴⁵. Hence, we analyze the time variation of the degree of balance on stock networks between January 2005 and September 2020, thus addressing a gap pointed out in a recent study²¹. To build the graphs we use the rank correlations between stocks, which provide a measure of similarity better suited for financial distributions. We find a balance-unbalance transition (BUT) in six of the countries studied. The BUTs occurred around September/October 2011 for the US, Greece, Portugal, Ireland, and Spain, and later in France, which took place just after the Black Monday in August 2011. Neither Germany, nor Italy, nor Japan showed clear signs of this BUT. We discover that the observed BUTs are mainly triggered by a reorganization of the topology of the stock market networks. They consist of the movement of a few low capitalization stocks from the periphery to the center of the networks by forming cliques of fully-negative interdependencies among them and with most of the rest of the stocks in the market. These transitions impact directly on the naive predictability of stock prices from pairwise rank correlations. This is a novel finding with a direct impact on optimal asset allocation and hedging, which opens promising avenues for future research.

Theoretical approaches

Correlations and predictability of interrelated stock prices. To study the similarity between stock price changes we consider the time series of the log returns of stocks, $Y_i(t) = \log [P_i(t)/P_i(t-1)]$, where $P_i(t)$ is the daily adjusted closing price of the stock i at time t for $t = 1 \dots S$. Let us consider the log return of three stocks forming the vectors Y_1, Y_2 and Y_3 , such that $Y_i = [Y_i(1), \dots, Y_i(S)]^T$. The similarity between the orderings of the log returns Y_A and Y_B of two stocks A and B can be captured by the Kendall's tau

$$\tau = 2\mathcal{P}((Y_A - Y'_A)(Y_B - Y'_B) > 0) - 1, \quad (1)$$

where Y'_ℓ is an independent copy of the vector Y_ℓ , $\ell = A, B$, and $\mathcal{P}(X)$ is the probability of the event X . If two rankings are concordant, $\tau > 0$, if two rankings are independent, $\tau = 0$, and if the two rankings are discordant, $\tau < 0$.

When considering the correlations between three stocks forming a triad there are four cases that can emerge: (i) the three pairs of stocks are correlated: $\tau(Y_1, Y_2) > 0$, $\tau(Y_1, Y_3) > 0$ and $\tau(Y_2, Y_3) > 0$; (ii) one pair of stocks is correlated and two pairs are anticorrelated: $\tau(Y_1, Y_2) > 0$, $\tau(Y_1, Y_3) < 0$ and $\tau(Y_2, Y_3) < 0$; (iii) two pairs of stocks are correlated and one pair is anticorrelated: $\tau(Y_1, Y_2) > 0$, $\tau(Y_1, Y_3) > 0$ and $\tau(Y_2, Y_3) < 0$; (iv) the three pairs of stocks are anticorrelated: $\tau(Y_1, Y_2) < 0$, $\tau(Y_1, Y_3) < 0$ and $\tau(Y_2, Y_3) < 0$.

Let us focus on estimating the trend of one of the stocks from another in the correlation triad considering a time varying regression model (see Appendix B in the Supplementary Information (SI)). Let us write $\hat{Y}_i(t)(Y_{j \neq i}(t))$ for the nonparametric smoothed estimate of Y_i from Y_j : $\hat{Y}_i(t) = \hat{m}_i(Y_j(t))$, where $m_i(\cdot)$ is a non-specified unknown function. Then, for any stock, e.g., Y_1 , we can make two predictions from another stock, e.g., Y_2 . One of them is simply $\hat{Y}_1(t)(Y_2(t))$, and the other is using first Y_2 to estimate Y_3 , and then estimate Y_1 from the last, i.e., $\hat{Y}_1(t)(\hat{Y}_3(Y_2(t)))$, that is $\hat{Y}_1(t) = \hat{m}_1(\hat{Y}_3(t))$, where $\hat{Y}_3(t) = \hat{m}_3(Y_2(t))$. In the cases (i) and (ii) the trends of Y_1 estimated by $\hat{Y}_1(t)(Y_2(t))$ and by $\hat{Y}_1(t)(\hat{m}_3(Y_2(t)))$ are the same. In the case (iii) $\hat{Y}_1(t)(Y_2(t))$ estimates an increase of Y_1 with the increase of Y_2 , while $\hat{Y}_1(t)(\hat{m}_3(Y_2(t)))$ estimates a decrease of Y_1 with the increase of Y_2 . Therefore, observing the trend evolution of Y_2 does not clearly estimate the trend of Y_1 independently of the quality of the pairwise rank correlations. In other words, the trend of Y_1 is unpredictable from Y_2 . A similar situation occurs for the case (iv).

In closing, the “predictability” of the trend of a stock Y_i from that of Y_j increases when the signs of the two estimations coincide. Otherwise, such predictability drops as a consequence of the different trends predicted by $\hat{Y}_i(t)(Y_j(t))$ and $\hat{Y}_i(t)(\hat{m}_k(Y_j(t)))$. If we represent the three stocks at the vertices of a triangle and the signs of the estimates as the corresponding edges, we have that the four cases analyzed before can be represented as in Fig. 1. Because the magnitude of $\tau_{ij}(t)$ quantifies the quality of the estimation for each t , we suggest to replace the edges of the triangles in Fig. 1 by the corresponding values of Kendall's tau instead of their estimates signs (for more details see Appendix B in the SI). In this case, a measure of the predictability of the stocks in a given triad is given by $\tilde{K}(t) = \tau_{ij}(t)\tau_{ik}(t)\tau_{jk}(t)$, with $\tilde{K}(t) \rightarrow +1$ corresponding to larger predictability and with $\tilde{K}(t) \rightarrow -1$ to poorer predictability at time t . Obviously, we should extend this measure beyond triangles, which is what we do in the next Section.

A triangle like those illustrated in Fig. 1 is known as a signed triangle. A signed triangle for which the product of its edge signs is positive is known as a balanced triangle (cases (i) and (ii) in Fig. 1). Otherwise, it is said to be unbalanced (cases (iii) and (iv) in Fig. 1). These concepts are naturally extended to networks of any size by means of the concept of signed graphs (see next Section). A signed graph where all the triangles (more generally

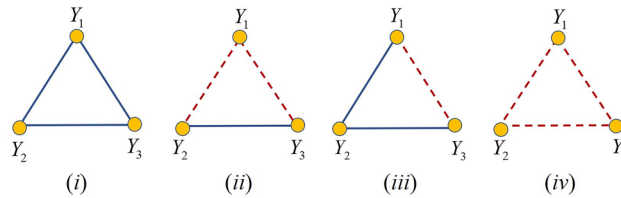


Figure 1. Graphical representation of the interdependencies between stocks Y_1, Y_2 and Y_3 in a triad. Stocks are represented as the vertices of a triangle and edges represent the correlation between them as accounted for by Kendall's tau. Positive are represented as solid (blue) lines between the corresponding stocks, and negative are represented as dashed (red) lines.

all its cycles) are balanced is said to be balanced. Therefore, we move our scenario of stock predictability to that of the analysis of balance in signed stock networks.

Balance in weighted signed stock networks. Let the stocks in a given market be represented as a set of nodes V in a graph $G = (V, E, W)$, where $E \subseteq V \times V$ is a set of edges which represent the correlations between pairs of stocks, and $W : E \rightarrow [-1, +1]$ is a mapping that assigns a value between -1 and $+1$ to each edge. We call these graphs weighted signed stock networks (WSSNs). The edges of the WSSN are defined on the basis of time varying Kendall's tau between the log returns of the stocks.

For any cycle in a WSSN we will say that it is positive if the product of all Kendall's taus forming the edges of the cycle is positive. A WSSN is balanced if all its cycles are positive. However, the question is not to reduce the problem to a yes or not classification but to quantify how close or far a WSSN is from balance. To do so, we consider a hypothetical scenario in which a given WSSN $G = (V, E, W)$ has evolved from a balanced network $G' = (V, E, |W|)$. That is, we are interested in quantifying the departure of a given WSSN from a balanced version of itself. We define the equilibrium constant K for the equilibrium $G' \rightleftharpoons G$ as $K = \exp(-\beta \Delta F)$, where β denotes the inverse temperature of the network and ΔF is the change of the Gibbs free energy of the system (for more details about the components of the equilibrium constant see Ref.⁴¹). The equilibrium constant is bounded as $0 < K \leq 1$, with the lower bound indicating a large departure of the WSSN from its balanced analogous and reaching $K = 1$ when it is balanced. We prove in Appendix C of the SI that this K is related to \bar{K} , which was defined in the previous section to account for the product of the signs of all edges in the cycles of the WSSN.

To calculate balance in a signed network, Faccchetti et al.⁴⁰ used an intensive computational technique that assigns a $+1$ or a -1 to all the nodes so as to minimize the energy functional

$$h(\sigma) := - \sum_{(v,w) \in E} J_{vw} \sigma_v \sigma_w, \tag{2}$$

where $\sigma_k \in \{+1, -1\}, k = 1, \dots, n$ with n equal to the number of nodes and J_{vw} accounts for the sign of the edge in the signed graph. Therefore, $h(\sigma) := - \sum_{v,w} A_{vw} \sigma_v \sigma_w$, where A_{vw} are the entries of the adjacency matrix of a signed graph. The energy functional $h(\sigma)$ is a quadratic form of the Hamiltonian: $h(\sigma) = \bar{1}^T \hat{\mathcal{H}}(G) \bar{1}$, where $\bar{1}$ is an all-ones vector. Here we use a different approach, which avoids the minimization of that energy functional, and which is based on the free energies $F(G)$ and $F(G')$ appearing in $\Delta F = F(G) - F(G')$. In this case we notice that the energy functional (2) is minimized if $J_{vw} \sigma_v \sigma_w > 0$ and maximized if $J_{vw} \sigma_v \sigma_w < 0$. Therefore, we identify this term with the Kendall's tau: $J_{vw} \sigma_v \sigma_w := \tau_{vw}$, which means that $\hat{\mathcal{H}}(G) = -A(G)$, such that correlated pairs of stocks contribute to the minimization of the energy functional and anticorrelated ones contribute in the opposite direction. Then, we have $\Delta F = -\beta^{-1} \ln(Z(G)/Z(G'))$, where $Z(G) = \text{tr}(\exp(-\beta \hat{\mathcal{H}}(G)))$.

For the current work, we consider the Economic Policy Uncertainty (EPU) index⁴⁶ as a proxy for the inverse temperature β . EPU is compiled for every country on a monthly basis and represents a level of "agitation" of the market at a given date. We then define a relative inverse temperature as $\beta_{rel} = EPU / \max(EPU)$, where $\max(EPU)$ is the maximum EPU reported for that country in the period of analysis. In this case, the risk is high when the "temperature" of the system at a given time is approaching the maximum temperature reached by that country in the whole period, $\beta_{rel} \rightarrow 1$. On the other hand, the risk is minimum when the temperature at a given time is much smaller than the maximum temperature of the period, $\beta_{rel} \rightarrow 0$.

This finally gives our measure of structural balance for a WSSN^{41,47}:

$$K = \frac{Z(G)}{Z(G')} = \frac{\text{tr}(\exp(-\beta_{rel} \hat{\mathcal{H}}(G)))}{\text{tr}(\exp(-\beta_{rel} \hat{\mathcal{H}}(G')))} = \frac{\text{tr}(\exp(\beta_{rel} A(G)))}{\text{tr}(\exp(\beta_{rel} A(G')))} \tag{3}$$

In the Supplementary Information (Appendix C) we prove that $K = 1$ if and only if the WSSN is balanced. The departure of K from unity characterizes the degree of unbalance that the WSSN has.

Weighted signed stock networks. In order to define the weight of the edges in the WSSNs we use Kendall's tau rank correlation coefficients. Since we are interested in the time evolution of the stock networks, the edges are determined in terms of time varying rank correlations. The dependence between the stock log returns

$(Y_A, Y_B)^T$ is estimated with the following nonparametric time varying Kendall's tau estimator proposed by Ascorbebeitia et al.⁴⁵:

$$\hat{\tau}_{A,B}(t) = \frac{4}{1 - \sum_{s=1}^S w_h(t-s)^2} \sum_{s,r=1}^S w_h(t-s)w_h(t-r)\mathbb{I}\{Y_A(s) < Y_A(r), Y_B(s) < Y_B(r)\} - 1 \quad (4)$$

for $t = 1, \dots, S$, where $w_h(t-s) = (Sh)^{-1}k((t-s)/(Sh))$ is a sequence of kernel weights that smooths over the time space, $h > 0$ is the bandwidth that regulates the degree of smoothness, and $\mathbb{I}\{\cdot\}$ is the indicator function. In all our calculations we consider the Epanechnikov kernel $k(x) = \frac{3}{4}(1-x^2)\mathbb{I}\{|x| < 1\}$. This kernel assigns a higher weight to those values of the indicator function that are close in time and lower weights to observations of the indicator farther away from t . The smoothing parameter h is selected minimizing the mean squared error of the Kendall's tau estimator in (4) (for more details see Appendix A in the SI and Ref.⁴⁵).

Let us now fix t and calculate $\hat{\tau}_{ij}(t)$ for every pair of stocks (i, j) in a given market and define the rank correlations matrix $M(t)$ whose entries are the values of $\hat{\tau}_{ij}(t)$. Let $\varepsilon \in \mathbb{R}^+$ be a given threshold. We create the matrix $A(t)$ whose entries are $A_{ij}(t) = \hat{\tau}_{ij}(t)\mathbb{I}\{|\hat{\tau}_{ij}(t)| \geq \varepsilon\}$. The threshold determines the value below which the correlation is considered to be negligible. Increasing ε decreases the number of connections in the corresponding network. We have considered $\varepsilon = 0.3$ a reasonable value. Moreover, the main findings are robust to alternative specifications, such as $\varepsilon = 0.2$ and $\varepsilon = 0.4$. The matrix $A(t)$ is then the adjacency matrix of a WSSN at time t . Here we construct weighted adjacency matrices based on daily returns from January 2005 until September 2020. Therefore, we have a set of WSSNs obtained with matrices $\mathcal{A} = \{A(t = 1), \dots, A(t = S)\}$.

Results

We study the stock markets of France, Germany, Greece, Italy, Ireland, Japan, Portugal, Spain, and the US during the period between January 2005 and September 2020 (see “Methods” section for details about data). The first interesting observation is the large balance observed for all markets between January 2005 and August 2007, when all of them display $K \geq 0.98$. However, in five countries there are sudden drops of balance around September/October 2011. This transition from highly balanced to poorly balanced markets occurs in Ireland in September 2011 and in the US, Portugal, Greece, and Spain in October 2011. The transitions are observed by naked eye in the plots of the temporal evolution of balance (see top panels of Fig. 2a–e). For a quantitative analysis of these BUTs as well as the identification of the dates at which they occurred, we use the detrended cumulative sums (DCS) of the balance (see Fig. 2). In these five markets, the general trend is a decay of the balance from 2005 to 2020. Therefore, the DCS increases for those periods in which the balance does not follow the general trend, i.e., when it does not decay with time. A negative slope of the DCS in some periods indicates that the balance drops more abruptly during this period than the general decreasing trend. Then, we can observe that DCS has a positive slope in all five markets with BUT between January 2005 and September/October 2011. At this point, the DCS changes its slope indicating an abrupt decay of balance. We select the point in which this change of slope occurs as the date marking the BUT. We should notice that there are some differences in the behavior of the DCS in each specific market. The US, Portugal and Greece display an initial increasing period followed by a continuous decay one. Ireland displays a short period in which DCS has slope slightly positive but close to zero between September 2011 and January 2016 when it definitively starts to decay. The market in Spain interrupted the decay of its balance on April 2017 when it starts to recover balance at a rate similar to the one of the period between January 2005 to October 2011. The value of the balance drops dramatically again after the crisis produced by COVID-19 as can be seen in Fig. 2e. Finally, we have included the stock market of France in Fig. 2f because it displays a behavior similar to that of Spain, although the BUT occurs significantly later with respect to the other countries in this group, i.e., on September 2012.

The markets of Germany, Italy, and Japan display very constant values of their balance across the whole period of analysis. Although there are some oscillations at certain specific dates, their DCS display an almost zero slope confirming the constancy of the balance of these markets (see Fig. 3).

Discussion

Among the factors that may influence the BUT in six out of nine countries studied, we should mention the global level of risk at which a given market is exposed to at a given time. We account for this factor through the parameter $0 < \beta_{rel} \leq 1$, which is based on the relative EPU index. Following our theoretical model, when the level of global risk is very low, $\beta_{rel} \rightarrow 0$, we have that $K \rightarrow 1$ and the corresponding WSSN tends to be balanced independently of its topology and edge weights. The analysis of the relative EPU values for all countries indicates that the level of risk accounted for by this indicator was relatively low from January 2005 to August 2007. This is exactly the same period for which a high balance was observed in all markets with $K \geq 0.97$. However, the relative EPU indices alone are not able to explain the BUTs observed in most markets around September/October 2011. For instance, the relative EPU index peaked at different values for different markets: Portugal (June 2016), the US (May 2020), Spain (October 2017), France (April 2017), Greece (November 2011) and Ireland (May 2020). It has been found that the US EPU can achieve better forecasting performance for foreign stock markets than the own national EPUs. This is particularly true during very high uncertainty episodes⁴⁸ (see Appendix F in the SI for more details and references). Following this empirical observation, we can speculate that a plausible cause that triggered the BUT in five European countries were the events of August 2011, when the US EPU index rocketed to its highest value between 2005 and the COVID-19 crisis.

The topological analysis of BUT should necessarily start by considering the role of negative interdependencies. It is obvious that they are necessary for unbalance, i.e., a totally positive network cannot be unbalanced. However, it has been widely demonstrated in the mathematical literature that this is not a sufficient condition for unbalance,

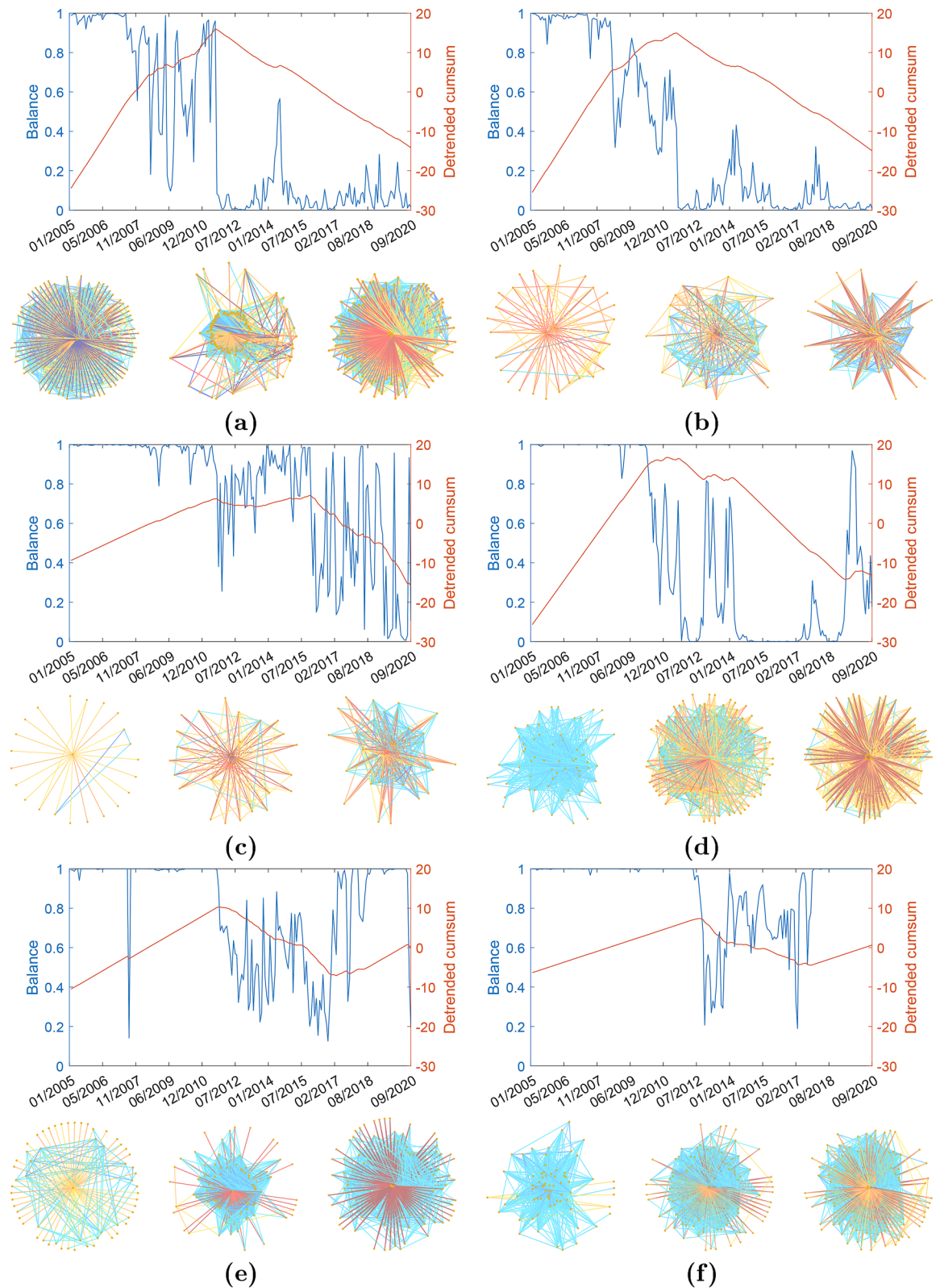


Figure 2. Balance degree evolution. (a) The US. (b) Portugal. (c) Ireland. (d) Greece. (e) Spain. (f) France. (Top panels) Illustration of the evolution of the balance for the period between January 2005 and September 2020 in the WSSN (blue line) and of its detrended cumulative sum (red line). (Bottom panels) Illustration of three snapshots of representative networks at different times (before, at, and after the BUT). The WSSNs are illustrated using the degree of the nodes as a proxy for the location of the nodes. The most central nodes are at the centre and the low degree nodes are located at the periphery of the graph. The colors of the edges correspond to the Kendall's tau estimates, with links going from dark red for the positive ones to blue for the negative ones. In the Supplementary Information we provide Matlab images of all the networks displayed in the figure where node labels are incorporated. The reader can zoom in these networks for details.

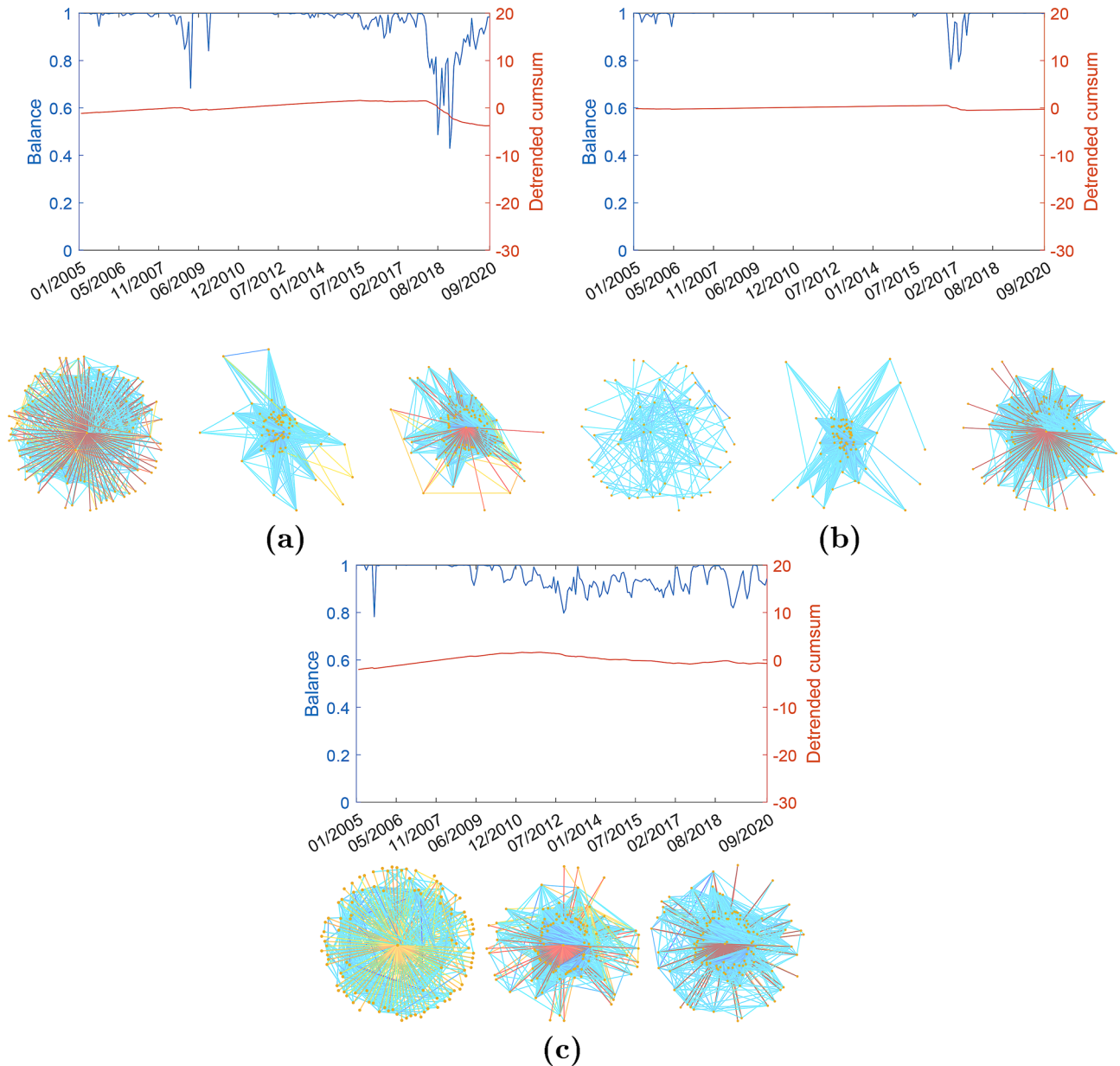


Figure 3. Balance degree evolution. **(a)** Germany. **(b)** Italy. **(c)** Japan. (Top panels) Illustration of the evolution of the balance for the period between January 2005 and September 2020 in the WSSN (blue line) and of its detrended cumulative sum (red line). (Bottom panels) Illustration of three snapshots of representative networks at different times (before, at, and after the BUT). The WSSNs are illustrated using the degree of the nodes as a proxy for the location of the nodes. The most central nodes are at the center and the low degree nodes are located at the periphery of the graph. The colors of the edges correspond to the Kendall's tau estimates, with links going from dark red for the most negative to blue for the positive ones. In the Supplementary Information we provide Matlab images of all the networks displayed in the figure where node labels are incorporated. The reader can zoom in these networks for details.

i.e., there are networks with many negative edges which are balanced^{42,47}. As an illustrative example let us consider the WSSN of the US of September 3, 2010, which has 950 negative edges and a balance of $K \approx 0.245$ indicating its lack of balance. However, the US WSSN of February 18, 2011 also has 950 negative edges, but it is balanced with $K \approx 0.946$. Similarly, the WSSN of Portugal on July 6, 2012 ($K \approx 0.700$) and on May 3, 2019 ($K \approx 0.407$), both have 720 negative edges. The ratio of negative to total edges is neither a sufficient condition for balance. For instance, in Spain on August 12, 2005 the market was perfectly balanced ($K = 1$), while on September 2007 it was unbalanced ($K \approx 0.142$), although both WSSNs have the same proportion of negative to total edges, i.e., 1.4%.

A detailed exploration of the WSSNs with low balance occurring after BUT (see Fig. 2) revealed the existence of small fully-negative cliques (FNC) formed by a group of s stocks (see Appendix E in the SI for a greater evaluation of networks snapshots). A FNC is highly unbalanced. That is, the balance of an FNC of s stocks is $K = (\exp(-\beta_{rel}(s-1)) + (s-1)e^{\beta_{rel}}) / (\exp(\beta_{rel}(s-1)) + (s-1)e^{-\beta_{rel}})$, which for a fixed value of β_{rel} clearly tends

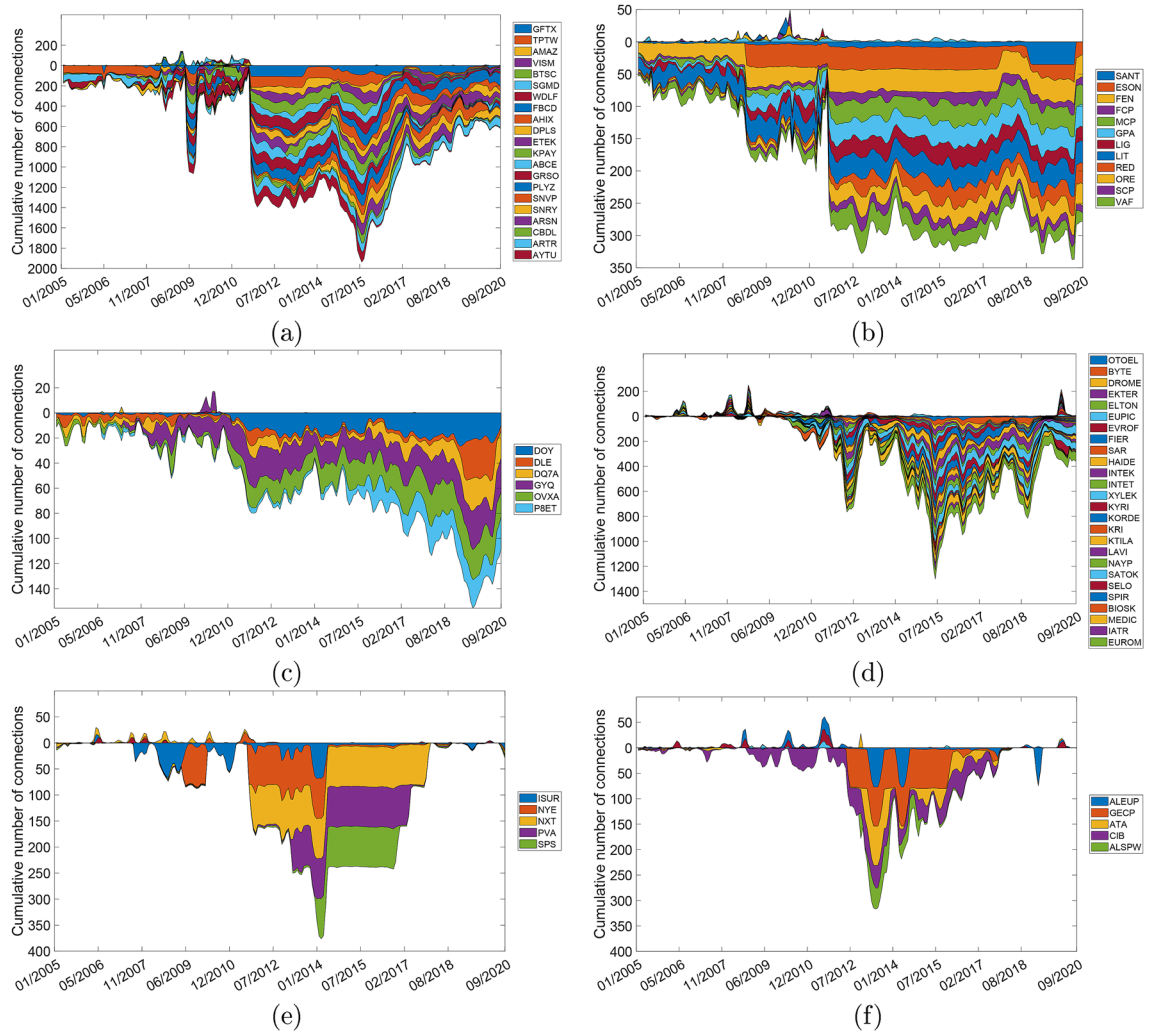


Figure 4. Positive and negative cumulative centrality degrees. **(a)** The US. **(b)** Portugal. **(c)** Ireland. **(d)** Greece. **(e)** Spain. **(f)** France. Temporal evolution of the cumulative number of connections (degrees) of every stock in the FNC of the six countries where BUT occurred. The cumulative number of positive connections with other stocks are shown over the horizontal line, and the cumulative number of negative connections are shown below that line.

to zero as $s \rightarrow \infty$. Therefore, the unbalanced nature of the post-transition structure of stock markets is mainly due to the presence of these cliques of mutually anti-correlated stocks. Additionally, each of the stocks in the FNC increases significantly its number of negative connections with the rest of stocks after BUT (see Fig. 4), being also negatively connected to almost every other stock in the WSSN. This subgraph resembles a kind of graph known as *complete split graph* (CSG)⁴⁹.

To explore the determinants of the lack of balance in these networks we simulate the structure of the networks displaying BUT by means of a quasi-CSG-WSSN (see “Methods” section for details) and compare it with a simulated random WSSN. The quasi-CSG-WSSNs have the same number of nodes and edges as the WSSNs but the size of the central clique is variable. We determine the “optimal” value of the sizes in the simulated networks, s_{opt} , by minimizing the root mean square error (RMSE) between the spectrum of the real WSSN and that of the quasi-CSG-WSSN. In SI Appendix G, Table S1, we report that the RSME of the random model is about twice bigger than that of the quasi-CSG for the six countries where BUT occurs. This clearly indicates that the topological organization, more than the number of positive/negative connections, is what determines the lack of balance in these networks.

The size of the FNC correlates very well with the intensity of the drop in the mean values of K before and after the BUT ($r \approx 0.857$). Two other important characteristics of these FNC are that: (i) they are mainly formed by non-financial entities (see SI Appendix G, Table S1), and (ii) they are composed by small and micro-caps with the exception of SAR in Greece which is a mid-cap. We have shown (see Appendix H in the SI) that if we split the WSSN by financial (F) and non-financial (NF) sectors, then the BUT is observed mainly for the networks containing NF-NF interactions.

All in all, these BUTs change the predictability of the stocks in the way we have explained in “Correlations and predictability of interrelated stock prices” section from more predictable markets to more unpredictable ones.

Conclusions

Considering stock-stock correlations in stock markets as signed graphs allowed us to introduce the concept of balance into stock market networks for the first time. We related the level of balance in these networks with stock predictability and identified a previously unknown transition between balanced markets to unbalanced ones in six out of nine countries studied. These balance-unbalance transitions occur in the US, Greece, Portugal, Spain, and Ireland around September 2011, following the Black Monday, and later on in France. No transition is observed in the stock markets of Germany, Italy, and Japan for the same period of time. The balance-unbalance transition is driven by a reorganization of the stock-stock correlations of a group of low capitalization stocks, mostly of non-financial entities which collapse into a fully-negative clique of anticorrelated pairs of stocks. Further studies are needed to understand the reasons and implications of this reorganization of non-financial entities in a given number of markets and the associated loss of network balance and stock predictability.

Methods

Data. We construct stock networks for some European countries, the US, and Japan. The seven selected European countries include Greece, Italy, Ireland, Portugal, and Spain (GIIPS), due to their financial vulnerability since the last global financial crisis. For comparison purpose, we have selected two other countries from the European Union: Germany and France. Finally, the US and Japan are included as non-European important drivers of the world economy. To construct the networks we use equities' daily closing price data from 01-03-2005 to 09-15-2020, obtained from the Morningstar database. Details about the cleaning process of the data are given in the SI, Appendix D. This data set is available from the authors on reasonable request.

As the inverse temperature, β of the networks we use monthly EPU index data⁴⁶ for each country from January 2005 to September 2020. Portugal is an exception since there is no specific EPU index available for it. Therefore, the European EPU index is considered instead. EPU index data is publicly available in the Economic Policy Uncertainty website⁵⁰.

Quasi-CSG-WSSN. Let n , m_- and m_+ be the number of nodes, of negative and positive edges in a real WSSN, respectively. We create a CSG with a clique of $s < n$ nodes and $s(2n - s - 1) < m_-$ edges. We complete the quasi-CSG-WSSN by adding randomly and independently $m_- - s(2n - s - 1)/2$ negative and m_+ positive edges among the nodes not in the clique. Additionally, we create the random-WSSN by adding randomly and independently m_- and m_+ edges among n nodes a la Erdős-Rényi⁵¹. A Matlab code for building such graphs (Algorithm 1) and an example of quasi-CSG-WSSN are given in Appendix G of the SI.

We then test the following two hypotheses:

1. Unbalanced WSSNs depend only on n , m_- and m_+ and not on any specific structure.
2. Unbalanced WSSNs depend on the existence of a specific quasi-CSG structure.

In the first case, the real-world WSSN would be more “similar” to its random version than to the quasi-CSG-WSSN and in the second case, it would be more similar to the quasi-CSG-WSSN. For the “similarity” between the different networks we focus here on the spectrum of their adjacency matrices. The reason for that is that the eigenvalues of this matrix determine the balance of the network as we have seen from its definition before. For determining the value of s we explore all possible values and find the root mean square error (RMSE) between the spectrum of the real WSSN and that of the quasi-CSG-WSSN. The smallest RMSE, s_{opt} , determines the “best” value of s .

Received: 19 April 2021; Accepted: 19 May 2021

Published online: 09 June 2021

References

1. Hall, R. E. Struggling to understand the stock market. *Am. Econ. Rev.* **91**, 1–11 (2001).
2. Preis, T., Schneider, J. J. & Stanley, H. E. Switching processes in financial markets. *Proc. Natl. Acad. Sci.* **108**, 7674–7678 (2011).
3. Fischer, S. & Merton, R. C. Macroeconomics and finance: The role of the stock market. In *Carnegie-Rochester conference series on public policy*, vol. 21, 57–108 (Elsevier, 1984).
4. Asgharian, H., Christiansen, C. & Hou, A. J. Economic policy uncertainty and long-run stock market volatility and correlation. Available at SSRN. <https://doi.org/10.2139/ssrn.3146924> (2019).
5. De Bondt, W. F. & Thaler, R. Does the stock market overreact?. *J. Financ.* **40**, 793–805 (1985).
6. Fama, E. F. & French, K. R. Common risk factors in the returns on stocks and bonds. *J. Financ. Econ.* **33**, 3–56 (1993).
7. Loretan, M. & English, W. B. Evaluating correlation breakdowns during periods of market volatility. Available at SSRN. <https://doi.org/10.2139/ssrn.231857> (2000).
8. Solnik, B., Boucrelle, C. & Le Fur, Y. International market correlation and volatility. *Financ. Anal. J.* **52**, 17–34 (1996).
9. Spelta, A., Flori, A., Pecora, N., Buldyrev, S. & Pammolli, F. A behavioral approach to instability pathways in financial markets. *Nat. Commun.* **11**, 1–9 (2020).
10. Su, Z., Fang, T. & Yin, L. Understanding stock market volatility: What is the role of us uncertainty?. *N. Am. J. Econ. Finance* **48**, 582–590 (2019).
11. Wen, X., Wei, Y. & Huang, D. Measuring contagion between energy market and stock market during financial crisis: A copula approach. *Energy Econ.* **34**, 1435–1446 (2012).
12. Gjerstad, S. D., Porter, D., Smith, V. L. & Winn, A. Retrading, production, and asset market performance. *Proc. Natl. Acad. Sci.* **112**, 14557–14562 (2015).
13. Smith, A., Lohrenz, T., King, J., Montague, P. R. & Camerer, C. F. Irrational exuberance and neural crash warning signals during endogenous experimental market bubbles. *Proc. Natl. Acad. Sci.* **111**, 10503–10508 (2014).

14. Engle, R. F. & Ruan, T. Measuring the probability of a financial crisis. *Proc. Natl. Acad. Sci.* **116**, 18341–18346 (2019).
15. Jackson, M. O. & Pernoud, A. Systemic risk in financial networks: A survey. Available at SSRN. <https://doi.org/10.2139/ssrn.3651864> (2020).
16. Buldyrev, S. V., Flori, A. & Pammolli, F. Market instability and the size-variance relationship. *Sci. Rep.* **11**, 1–8 (2021).
17. Haldane, A. G. Rethinking the financial network. In *Fragile Stabilität–Stabile Fragilität*, 243–278 (Springer, 2013).
18. Bougheas, S. & Kirman, A. Complex financial networks and systemic risk: A review. *Compl. Geogr. Econ.* **19**, 115–139 (2015).
19. Grilli, R., Iori, G., Stamboglis, N. & Tedeschi, G. A networked economy: A survey on the effect of interaction in credit markets. In *Introduction to Agent-Based Economics*, 229–252 (Elsevier, 2017).
20. Iori, G. & Mantegna, R. N. Empirical analyses of networks in finance. In *Handbook of Computational Economics*, vol. 4, 637–685 (Elsevier, 2018).
21. Bardoscia, M. *et al.* The Physics of Financial Networks. arXiv:2103.05623 (2021).
22. Battiston, S., Gatti, D. D., Gallegati, M., Greenwald, B. & Stiglitz, J. E. Default cascades: When does risk diversification increase stability?. *J. Financ. Stab.* **8**, 138–149 (2012).
23. Battiston, S., Puliga, M., Kaushik, R., Tasca, P. & Caldarelli, G. Debrank: Too central to fail? Financial networks, the fed and systemic risk. *Sci. Rep.* **2**, 1–6 (2012).
24. Battiston, S. *et al.* Complexity theory and financial regulation. *Science* **351**, 818–819 (2016).
25. Mantegna, R. N. Hierarchical structure in financial markets. *Eur. Phys. J. B-Condens. Matter Compl. Syst.* **11**, 193–197 (1999).
26. Kutner, R. *et al.* Econophysics and sociophysics: Their milestones & challenges. *Phys. A* **516**, 240–253 (2019).
27. Birch, J., Pantelous, A. A. & Soramäki, K. Analysis of correlation based networks representing DAX 30 stock price returns. *Comput. Econ.* **47**, 501–525 (2016).
28. Brida, J. G., Matesanz, D. & Seijas, M. N. Network analysis of returns and volume trading in stock markets: The Euro Stoxx case. *Phys. A* **444**, 751–764 (2016).
29. Kauê Da'Maso Peron, T., da Fontoura Costa, L. & Rodrigues, F. A. The structure and resilience of financial market networks. *Chaos Interdiscip. J. Nonlinear Sci.* **22**, 013117 (2012).
30. Wang, G.-J., Xie, C. & Stanley, H. E. Correlation structure and evolution of world stock markets: Evidence from pearson and partial correlation-based networks. *Comput. Econ.* **51**, 607–635 (2018).
31. Zhao, L. *et al.* Stock market as temporal network. *Phys. A* **506**, 1104–1112 (2018).
32. Guo, X., Zhang, H. & Tian, T. Development of stock correlation networks using mutual information and financial big data. *PLoS ONE* **13**, e0195941 (2018).
33. Tumminello, M., Lillo, F. & Mantegna, R. N. Correlation, hierarchies, and networks in financial markets. *J. Econ. Behav. Organ.* **75**, 40–58 (2010).
34. Heiberger, R. H. Stock network stability in times of crisis. *Phys. A* **393**, 376–381 (2014).
35. Chi, K. T., Liu, J. & Lau, F. C. A network perspective of the stock market. *J. Empir. Financ.* **17**, 659–667 (2010).
36. Stavroglou, S. K., Pantelous, A. A., Stanley, H. E. & Zuev, K. M. Hidden interactions in financial markets. *Proc. Natl. Acad. Sci.* **116**, 10646–10651 (2019).
37. Harary, F. *et al.* On the notion of balance of a signed graph. *Mich. Math. J.* **2**, 143–146 (1953).
38. Zaslavsky, T. A mathematical bibliography of signed and gain graphs and allied areas. *Electron. J. Combin.* **8**, DS8 (2012).
39. Leskovec, J., Huttenlocher, D. & Kleinberg, J. Signed networks in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* **1361–1370** (2010).
40. Facchetti, G., Iacono, G. & Altafini, C. Computing global structural balance in large-scale signed social networks. *Proc. Natl. Acad. Sci.* **108**, 20953–20958 (2011).
41. Estrada, E. & Benzi, M. Walk-based measure of balance in signed networks: Detecting lack of balance in social networks. *Phys. Rev. E* **90**, 042802 (2014).
42. Kirkley, A., Cantwell, G. T. & Newman, M. Balance in signed networks. *Phys. Rev. E* **99**, 012320 (2019).
43. Shi, G., Altafini, C. & Baras, J. S. Dynamics over signed networks. *SIAM Rev.* **61**, 229–257 (2019).
44. Heider, F. Attitudes and cognitive organization. *J. Psychol.* **21**, 107–112 (1946).
45. Ascorbebeitia Bilbatua, J., Ferreira García, E. & Orbe Mandaluniz, S. The effect of dependence on European market risk. A non-parametric time varying approach. *J. Bus. Econ. Stat.* **1**, 2. <https://doi.org/10.1080/07350015.2021.1883439> (2021).
46. Baker, S. R., Bloom, N. & Davis, S. J. Measuring economic policy uncertainty. *Q. J. Econ.* **131**, 1593–1636 (2016).
47. Estrada, E. Rethinking structural balance in signed social networks. *Discret. Appl. Math.* **268**, 70–90 (2019).
48. Bijsterbosch, M. & Guérin, P. Characterizing very high uncertainty episodes. *Econ. Lett.* **121**, 239–243 (2013).
49. Estrada, E. & Benzi, M. Core-satellite graphs: Clustering, assortativity and spectral properties. *Linear Algebra Appl.* **517**, 30–52 (2017).
50. Economic policy uncertainty index. <https://www.policyuncertainty.com/index.html>.
51. Erdős, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–60 (1960).

Acknowledgements

E.F., S.O., and J.A. were supported by the Spanish Ministry of the Economy and Competitiveness under Grant ECO2014-51914-P; the UPV/EHU under Grants BETS-UFI11/46, MACLAB-IT93-13 and PES20/44; and the Basque Government under BiRTE-IT1336-19. J.A. also acknowledges financial support under PIF16/87 from UPV/EHU. E.E. thanks partial financial support from Ministerio de Ciencia, Innovación y Universidades, Spain, Grant PID2019-107603GB-I00.

Author contributions

E.F. and E.E. designed and E.E. directed the investigation. E.F., S.O., and J.A. constructed the data sets, and defined and constructed the adjacency matrices. B.A. and E.E. defined, constructed, and analyzed the networks. E.E. wrote the paper with inputs from all the authors. All the authors contributed to preparing the SI.

Funding

Brais Álvarez Pereira's work on this study was funded by Fundação para a Ciência e a Tecnologia (UIDB/00124/2020, UIDP/00124/2020 and Social Sciences Datalab – PINFRA/22209/2016), POR Lisboa and POR Norte (Social Sciences DataLab, PINFRA/22209/2016).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-91266-4>.

Correspondence and requests for materials should be addressed to E.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021, corrected publication 2022