# Financial crisis network effects on developing economies: the role of 'too big to fail' countries

Mauricio Rada-Orellana[1]

[1] *Universidad del Pacifico, Lima, Peru*

**Abstract**

On September 16, 2008, following the implosion of the credit crisis, world stock markets, led by NYSE, just crashed. Likewise, at the end of the year, GPD variation of the countries diminished to around 0% or negative numbers. Nevertheless, many developing countries somehow managed to have positive variations and were not practically affected. For instance, Peruvian GDP grew 9.1% in 2008. This study aims to understand how the cascading effects happen along the economic system with a focus on the spread in developing countries. We hypothesize that centrality properties play an essential role in network effects. Here, we characterize a network based on the world stock exchange. Then, a community detection method was applied to identify whether the communities relate to the type of economy. Finally, we identify the crash propagation between the economies and level of shock. Preliminary, the results show that no matter the size or the geographical location; the network effects have a differential impact on developing economies, or, paradoxically, in developed ones as Japan which suffered a high impact unlike its neighbor China in 2008. This study offers a new perspective for policymakers to define patterns of the crash within the same type of economies.

## 1 Introduction

Why do a crisis that occurs in Turkey can extend to a far away country like Bolivia, and how does it happen? In 1776, Adam Smith stated that a variety of joint labor is needed to form a unit product using the example of the production of a woolen coat, thus he argued that *"without the assistance and co-operation of many thousands, the very meanest person in a civilized country could not be provided"* [Smith, 1776]. That description identifies the interconnection of elements at the micro level, likewise, at the macro level, countries' economies are not totally independent of one other due to increasing globalization. The interconnection between economies are part of the current economic system, and as other complex systems reveal emergent behaviors.

One of these behaviors is the networks effects that take place in financial markets as the stock exchanges of each country. This work aims to outline the network's pattern as the cascading effects on the markets. Here, we hypothesize that centrality properties play an essential role in the development of the crash propagation. This study builds the topology of the stock market of 74 countries around the world to then analyze its network properties using cross-correlations. Moreover, this topology was analyzed with community detection techniques to determine an optimum threshold correlation for obtaining modular and coherent communities. Finally, the characteristics of the network served to model the failure propagation of economic crises until 2017 and describe its pattern on developing countries.

The driving reason behind this work is to help policymakers to understand how variations in the capitalization of some markets (i.e., 'too big to fail' countries) alter the whole world network

and harm other markets–whether from developed, developing or frontier countries. In addition, it contributes to the growing body of academic knowledge with applications of network techniques to financial markets–either in static and dynamic phases.

The remainder of this article is organized as follows: The literature review section analyzes current studies and establishes the network theory framework. The methods section structures the steps to be taken in network characterization and subsequent analysis. The results section shows the findings of the experiment, pointing out the detected communities and epidemic model. The discussion section sums up the whole study.

## 2 Literature review

### 2.1 Financial markets as complex systems

The study of complex systems has benefited from the interdisciplinary approaches by the areas of mathematics, physics, sociology, and other fields. These types of systems are comprised of different elements which are connected by different interactions, broadly known as networks. Financial markets are an example of a complex network and it is important to understand its structure since it affects its functionality [Strogatz, 2001]. The first works in this area started with Mantegna (1999), where he finds the topological structure of portfolios of US stocks using the correlation matrix [Mantegna, 1999], and then, he introduces new conceptual approaches from the physics applied to economic problems [Mantegna and Stanley, 1999]. From there, these techniques have been widely used in many other stock markets like in China [Huang et al., 2009], Brazil [Tabak et al.,

2010], Greece [Dimitrios and Vasileios, 2015], Korea [Kim et al., 2007] and United States [Boginski et al., 2005] [Sun et al., 2015]. Similarly to these studies, this paper started characterizing the network with a correlation matrix and calculating the statistical properties, however, it also explores the topic of failure propagation using epidemic modeling. With respect to the analyzed data, this paper is not constrained to a certain country, since it includes the aggregation of market capitalization in 74 economies.

### 2.2 Network methodological issues

Before beginning with the methods section, it is necessary to define the established network framework which is mainly based on graph theory. This subsection is divided into two parts:

1. Definition of statistical properties

   To identify the most important nodes, centrality is widely one of the most used methods [Newman, 2018]. There are many definitions of centrality, this work used three of them. These measures can vary depending on the threshold ($\theta$) for the correlation.

   (a) Degree centrality It is the building block for more complex measures since it is simply the degree of each node. Given a node $i$, the degree centrality is the total number of edges $j$, which is represented in the adjacency matrix $A_{ij}$

   $$C_{ki} = \sum_{j=1}^{n} A_{ij}$$

   (b) Closeness centrality It is inverse of the average of the shortest paths from the node to all nodes in the network. Given the node $i$ with the shortest path or geodesic distance

to node $j$, since it is a distance the centrality measure is the inverse of that value.

$$C_{ci} = \frac{n}{\sum_j d_{ij}}$$

(c) Betweenness centrality It measures the importance of the node to connect other ones. Given the number of geodesic distance from nodes $s$ and $d$ that pass through $i$ ($d^i_{st}$), and total geodesic distance between those nodes.

$$C_{bi} = \sum_{st} \frac{d^i_{st}}{d_{st}}$$

2. Definition of community detection metrics

Community detection is the process to naturally divide a network into groups of nodes with the goal of have many edges within a group and few between the groups, with the objective of understanding the structure of the network [Newman, 2018]. There are three definitions of community: (1) cliques refer to groups of nodes that are a complete subgraph, (2) strong communities are connected subgraphs whose nodes have more edges within the community than outside of it, and (3) weak communities that are subgraphs that overall have more edges within that outside the group [Barabási et al., 2016]. After performing a community detection technique is necessary to evaluate the quality of the subgraphs, there are three measures to consider:

(a) Modularity: It is the most popular measure, with the goal of finding the maximum value as an expression of the goodness of the partitions [Newman and Girvan, 2004]. Given as a proportion of correct edges, where $m$ is the number of edges and the sum aggregates all pairs of vertices $i$ and $j$, $A_{ij}$ is an element of the adjacency matrix, $P_{ij}$ is

the number of edges between nodes $i$ and $j$ in the null model, and $\delta$ is a binary variable that represents the value of 1 only if nodes $i$ and $j$ are part of the same subgraph.

$$M(Q) = \frac{1}{2m} \sum_{ij} (A_{ij} - P_{ij}) \delta(C_i, C_j)$$

(b) Coverage: It represents the proportion between the number of intra-community edges with the total number of edges [Fortunato, 2010]. The goal is to maximize the value, then all the edges will be only located inside its respective communities, it is measured by $C(Q)$.

(c) Performance: It represents the sum of the number of pair of nodes $i$ and $j$ that are in the same community $C_i = C_j$ and are connected $(i, j) \in E$, and the pair of nodes that are separated $C_i \neq C_j$ and not connected $(i, j) \notin E$, with respect of the total number of edges [Fortunato, 2010]. Where $E$ is the set of edges and $n$ the number of nodes with the goal to maximize the value.

$$P(Q) = \frac{|(i, j) \in E, C_i = C_j| + |(i, j) \notin E, C_i \neq C_j|}{n(n-1)/2}$$

## 3 Methods

The methodology comprises the stages of treatment, summary, characterization, and analysis of the stock market valuation from 1974 to 2017. The whole description of the methodology is depicted in Figure 2.

### 3.1 Data description

The data collected for this study comes from the open data of The World Bank whose source is the World Federation of Exchange Database, a group of yearly time series of 136 entities (that includes countries and groups of countries) from 1975 to 2017 [World Bank, 2017]. Nevertheless, due to the lack of data for some countries, the data was filtered to include only 74 entities. In other words, the only included entities are the ones with at least 12 registered years–which is the 25$^{th}$ percentil. It is important to note that each data point comes from the official report of each national stock market and is aggregated on a yearly basis. The market capitalization is represented as the share price times the number of shares for listed domestic companies. The entity profile includes the name of the entity, the region, the income level–according to World Bank classification, and market classification–according to the S&P Dow Jones Indices Country Classification of 2012 [Standard and Poors, 2012].

### 3.2 Network characterization and measurement

The stock markets were defined as the nodes of the network since the purpose is to analyze the effects between the variations among them. On the other hand, the edges come from the adjacency matrix of zero lag cross-correlation between the market's capitalization. Let $x_{i(t)}$ be the value of the market entity $i$ at the time $t$. The time lag is used when the data points have a short period of aggregation–i.g., daily basis–and the effect is not instantaneously absorbed between the markets. Since this data is yearly aggregated, we used zero lag series. The dataset contains the total capital value, while the correlation matrix should contain the values of influence i.e., the returns or first differences. The returns of a market $i$ in a time $t$ with respect to time $t-1$ is defined

as $s_{i(t)} = ln(x_{i(t)} - x_{i(t-1)})$. Then, the cross-correlation between the returns of market $i$ and the returns of the market $j$ is calculated as:

$$\rho_{ij} = \frac{\sum_t (s_{i(t)} - \bar{s_i})(s_{j(t)} - \bar{s_j})}{\sqrt{\sum_t (s_{i(t)} - \bar{s_i})^2}\sqrt{\sum_t (s_{j(t)} - \bar{s_j})^2}}$$

The value of $\rho_{ij}$ range from -1 to 1, depending on how strong is the cross-anticorrelation or cross-correlation, respectively [Mantegna and Stanley, 1999]. The correlation matrix corresponds to the adjacency matrix, where the edge weight between $x_i$ and $x_j$ is the $\rho_{ij}$ value. For the 97 markets of world stock exchange, there are $n(n-1)/2 = (74*73)/2 = 5402$ $\rho_{ij}$. The weight used in the network was defined as $|\rho_{ij}|$, to simplify the analysis and only gather the influence whether it is positive or negative.

Subsequently, the graph $G(E, V, W)$ is formed by a finite set of $n$ nodes $V$, a set of undirected edges $E \subset VxV$ where $(x_i, x_j)$ represents an edge between $x_i \in V$ and $x_j \in V$, and a set of weights $w_e \geq \theta$.

### 3.3 Community detection

To perform a community clustering is necessary to have a sparse matrix or weighted links so edge density is sufficiently heterogeneous to form groups [Fortunato, 2010]. Although the stock market is not a sparse network, it is comprised by weighted edges ($|\rho_{ij}|$).

The graph $G(E, V, W)$ was analyzed by performing a qualitative and a quantitative analysis. Qualitative analyses have used the advantages of visualization to identify communities in networks

of 50 nodes or less, however, this technique is not helpful for larger networks. To find the communities, we use the weighted degree and type of economy as parameters. While the first shows the importance of the stock market and is represented with the node size; the second shows which type of economy each stock market belongs to, and is represented by the node color. To have a better visualization, the graph is positioned with a layout algorithm that clusters the nodes with higher relationships between them, here we use Fruchterman Reingold [Fruchterman and Reingold, 1991].

On the other hand, quantitative detection aims to reach sets of communities by optimizing a certain quality value. It is important to note that, as opposed to partitioning and traditional data clustering, community detection techniques has to perform without a previously defined number or size of groups [Barabási et al., 2016] [Newman, 2018]. We used the fast greedy algorithm [Clauset et al., 2004], which is based on the optimization of the modularity. Accordingly, it tries to find the sets of communities that maximize that modularity quality function. First, the algorithm was performed for the full stock market network. Then the graph was filtered to include only edges with correlations more or equal than a threshold ($\theta$), which is the calculated minimum value whose greedy application maximizes the modularity, a graph $G(E, V, W)$ where $w_e \geq \theta$. Higher values decrease the number of edges and nodes, while lower values deteriorate the modularity and other performance metrics. Previously, other algorithms were tested (like as Kernighan-Lin [Newman, 2006] [Kernighan and Lin, 1970], Spin Glass [Reichardt and Bornholdt, 2006] or label propagation [Raghavan et al., 2007]). Unlike the greedy algorithm, other ones lead to very small communities and gathered a worse modularity.

### 3.4 Epidemics analysis

A key advantage of characterizing the network is that we can identify and model network effects as cascading behavior. Cascading failures and propagation of information in the network are effects quite common in technological, biological and social networks, for instance, blackouts in power grids [Dobson et al., 2007]. In the case of financial crisis, a cascading effect can indicate the propagation of downturn between economic agents of the financial network—i.g., bankrupting of banks or companies in a cascade.

Similarly, as the community detection process, the network was analyzed with qualitative and quantitative techniques. In particular for epidemic analysis, a new attribute was introduced to the nodes, which is the state of the stock market at the time $t$. Here, the property of compartmentalization is introduced, it refers to the classification of the stage of the agent—i.e., the state—during the epidemic spread. Along economic processes, stock markets can, broadly, be classified as *prosper*, in *crisis* or *stable*.

The qualitative analysis used the graph $G(E, V, W)$ as a network base and the state attribute to visually differentiate the state of the agents. The prosper markets are comprised by all the stock exchanges with returns equal or greater than 10%, stable market are comprised by all the stock exchanges with returns from -10% to 10% and crisis markets are comprised by all the stock exchanges with returns equal or less than 10%. All returns are computed for the average of each year from 1975 to 2017 with a time window of 3 years since we want to capture a long-lasting crisis. The output was an iteration of the dynamic process of the stock market network, whenever

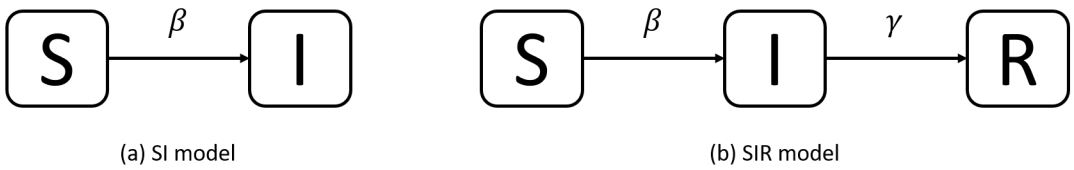


Figure 1: Epidemiological models

a crisis ignites, prosper and stable markets start to change its steady state to crisis state.

Quantitative analysis modeled the epidemic process following the epidemiology convention and the well-established mathematical framework [Diekmann and Heesterbeek, 2000]. To simplify the model, both *prosper* and *stable* states are grouped in only one state called *susceptible*, correspondingly, *crisis* state is renamed as *infected*. During the process, the susceptible individuals change their state by getting infected. Furthermore, in many biological disease process, some of the infected individuals eventually change their state as they recover or die. In the case of the financial crisis process, some of the infected markets eventually recover, a state called *removed*. This structure is called compartmental hypothesis. These three states build two different models: (1) the SI model where the markets can only be susceptible (S) or infected (I)–after being susceptible, and (2) the SIR model where the markets can also be removed (R) (recover)–after being infected. These flows are depicted in Figure 1.

Although the mathematical modeling of the epidemiological diseases is out of the scope of this work, we used the terminology and calculated some of the ratios to have a better understanding of the crisis of 2008. Either in an SI or SIR model, the susceptible turn into infected with a rate of infection $\beta$, however, it can turn into recovered with a rate of recovery $\gamma$. To calculate the

marginal rate of infected elements, it is necessary to multiply the rate of infection by the number of individuals infected and susceptible, and then subtract from this the number, the marginal rate of recovered (multiplication of rate of recovery by the number of infected).

$$\frac{\Delta I_{t+\Delta t}}{\Delta t} = \beta S_t I_t - \gamma I_t$$

A positive marginal rate of infected means that the infection is increasing, but if it is negative, means that eventually, the infection will disappear. The latter case happens when $R_0$ or basic reproductive number is less than 1, or the recovery rate is greater than the infection rate. $\gamma$ is the inverse of the recovery time.

$$\frac{\Delta I_{t+\Delta t}}{\Delta t} < 0 \rightarrow \beta S_t I_t - \gamma I_t < 0$$

$$R_0 = \frac{\beta}{\gamma} < \frac{1}{S_t}$$

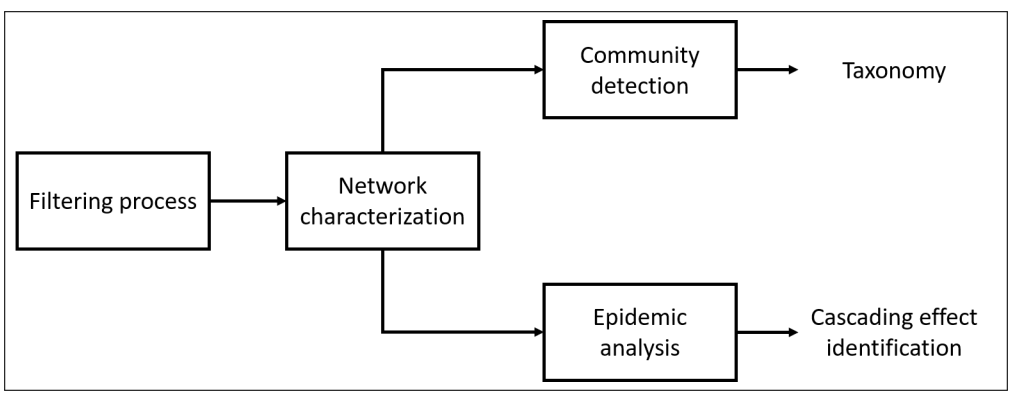


Figure 2: Methodology model

## 4 Results

### 4.1 Network topology

Since this is a weighted network, the data included both positive and negative correlations as

shown in Figure 3. It has a shape similar to a normal distribution with a mean of 0.35. Most of the values are larger than 0, which tells that the number of stocks that variate in the same direction is much higher than others. Figure 4 shows the degree distribution of the network, which is, in fact, right-skewed and similar to a power-law distribution, so most of the nodes have a low degree while just a few (hub nodes) concentrate the highest degrees. It is important to note that this pattern (called scale-free network) only intensifies when $\theta > 0.50$. Table 1 displays the results for each threshold. Figure 5 represents the stock markets plotted within two centralities, it shows that India, Brazil, Hungary, and Italy are the most central nodes according to both measures. The graph also shows a positive correlation between the closeness and betweenness. It is interesting that these markets are developing countries, moreover, from Figure 4, these are part of the 'too big to fail' nodes. This phenomenon will be clarified in the subsection 4.3.

Table 1: Topology table with different thresholds.

| Correlation | $\geq 0.1$ | $\geq 0.2$ | $\geq 0.3$ | $\geq 0.4$ | $\geq 0.5$ | $\geq 0.6$ | $\geq 0.7$ | $\geq 0.8$ | $\geq 0.9$ |
|---|---|---|---|---|---|---|---|---|---|
| Nodes | 74 | 74 | 74 | 74 | 74 | 73 | 69 | 54 | 28 |
| Edges | 2320 | 1984 | 1636 | 1255 | 909 | 601 | 368 | 171 | 45 |
| Average degree | 62.7 | 53.6 | 44.2 | 33.9 | 24.6 | 16.5 | 10.7 | 6.3 | 3.2 |
| Diameter | 2 | 2 | 2 | 3 | 4 | 5 | 6 | 5 | n.d. |
| Isolates | 0 | 0 | 0 | 0 | 0 | 1 | 5 | 20 | 46 |

### 4.2 Community structure

Figure 7 represents how strongly connected markets tend to group in sections of the graph. The graph is colored with red (high income), blue (upper middle income) and green (low middle in-

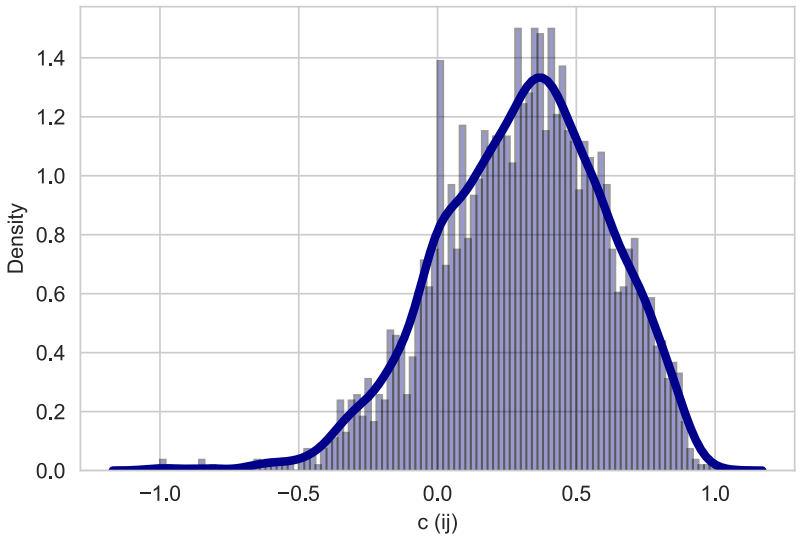


Figure 3: Distribution of correlation coefficients between the stock markets.

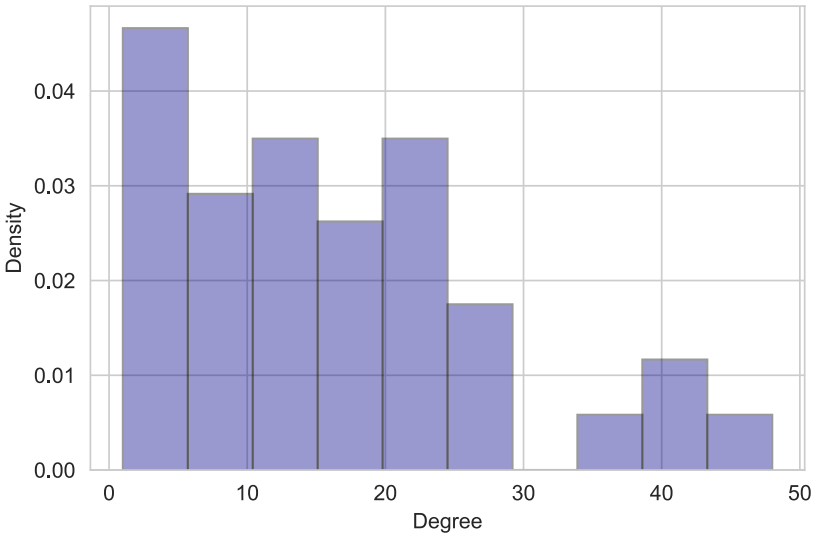


Figure 4: Distribution of network degree using a 0.60 threshold.

come). This shows that income levels are representative for the degree of absolute correlation, especially for high-income countries. These countries are grouped in the lower middle of the figure, however, there are cases like Norway that are placed outside its income level group. The middle-income countries overall tend to position in the periphery of the graph. Likewise, Figure 8 represent the economy type with orange (developed), green (emerging), blue (frontier) and grey (no data). In this case, while emerging countries tend to position in the center of the graph, geographic proximity play a role for some countries. For instance, Peru, Chile, and Brazil tend to be close, as well as Philippines, Thailand and Indonesia. An interesting case is Argentina, that is between the emerging and developed groups. Although currently is in the frontier group, it was downgraded from emerging after the last S&P classification.

For the full sample, the fast greedy algorithm identified three communities as represented in Table 2. The first one is mainly composed of developed countries whose average market capitalization is the highest among the communities (4,152 billion of dollars). Paradoxically, this community is also comprised of frontier countries at a low proportion, however, these countries are either located close to developed economies or were one of the richest countries in the past (Argentina). Although community 2 is comprised of a mix of developed and emerging countries, it contains most of the latter group, so can be considered as a transition community between 1 and 3. Significantly, community 3 is the groups of the frontier economies whose average capitalization is 25 billion of dollars, which is 166 times lower than the developed countries from the first group.

Before performing the last community detection, it is necessary to filter the edge weights with

the threshold ($\theta$). Figure 6 graphs the variation of the index depending on the quality measure. Although the measures are not monotonous functions, modularity has an upward trend until it reaches 0.88 ($M(Q) = 0.43$), however, coverage reaches a maximum at 0.85 ($C(Q) = 0.75$). From 0.82 to 0.85, the measures have approximately the similar values, since we prefer the lowest $\theta$ that maximizes the quality, we choose $\theta = 0.82$. Then we obtain the graph $G(E, V, W)$ where $w_e \geq 0.82$; $M(Q) = 0.38$, $C(Q) = 0.76$ and $P(Q) = 0.73$. The results are depicted in Table 3. Both the full graph and subgraph results clearly show that stock markets form modular communities that mainly match their income level and type of economy.

Table 2: Full community structure. Expressed in billions of US$.

| Communities | Community 1 | | | Community 2 | | | Community 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Economy type | % | Max | Mean | % | Max | Mean | % | Max | Mean |
| Developed | 71.4 | 32100 | 4152 | 40 | 6220 | 1778 | 0 | n.d. | n.d. |
| Developing | 7.1 | 201 | 201 | 56 | 8710 | 1159 | 8 | 51 | 51 |
| Frontier | 21.4 | 109 | 39 | 4 | 19 | 19 | 92 | 86 | 25 |

Table 3: Subgraph community structure. Expressed in billion of US$

| Communities | Community 1 | | | Community 2 | | | Community 3 | | | Community 4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Income level | % | Max | Mean | % | Max | Mean | % | Max | Mean | % | Max | Mean |
| High income | 100.0 | 32100 | 3117 | 44.4 | 4350 | 1237 | 28.6 | 6 | 5 | 100.0 | 3 | 3 |
| Lower middle income | 0.0 | n.d. | n.d. | 16.7 | 2330 | 1047 | 42.9 | 86 | 51 | 0.0 | n.d. | n.d. |
| Upper middle income | 0.0 | n.d. | n.d. | 38.9 | 1230 | 485 | 28.6 | 8710 | 4416 | 0.0 | n.d. | n.d. |

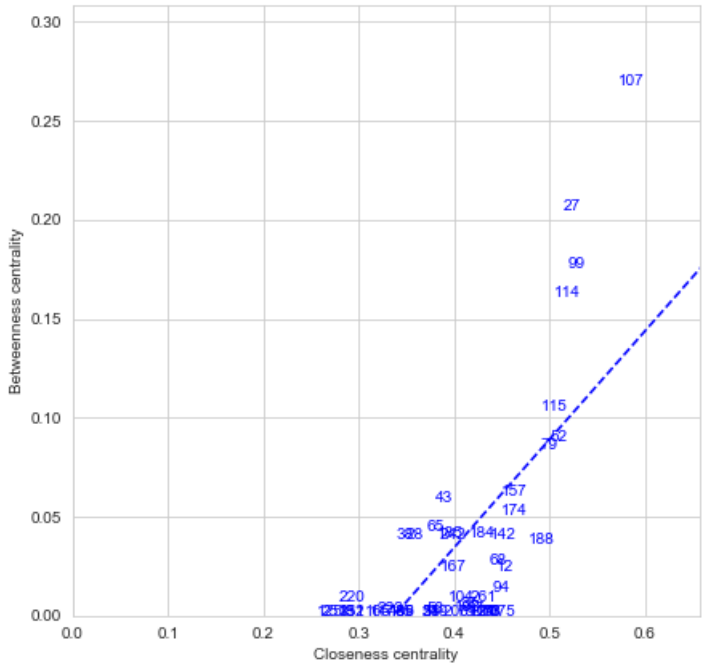


Figure 5: Centrality measures by stock market.

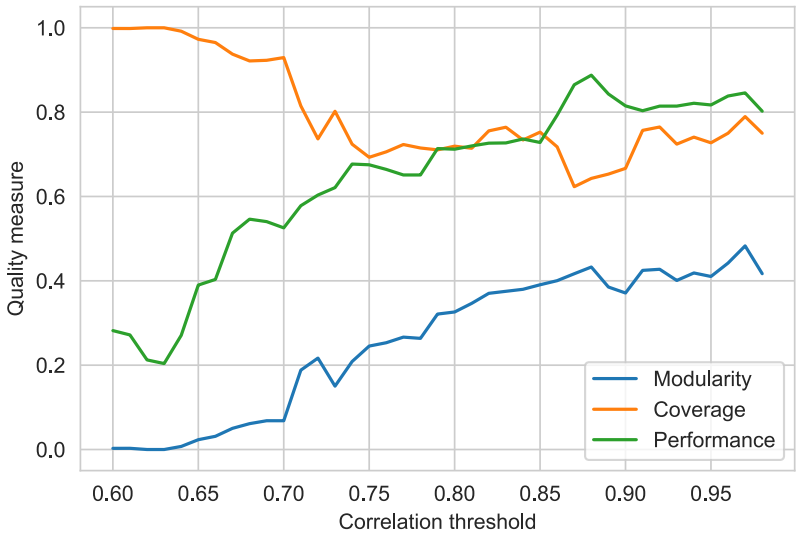


Figure 6: Value per quality measure.

### 4.3 Virulence and propagation of crisis

In order to visualize the propagation of the market crisis in 2008 (Figure 9), a dynamical transformation was performed, showing the pattern of how it extended with a climax in 2010. Nodes colored with green are prospering markets, blue nodes represent stable markets, red ones represent markets in crisis, and nodes that did not report data for that period are colored with white. In 2007, almost all markets reported returns of more than 10%, however at the end of 2008, most of the developed ones (lower in the left of the graph) were in crisis. Even though emerging (center) and frontier (upper right) economies were also infected, the proportion was significantly less. For instance, Brazil, Chile, and Peru did not reach the crisis state, and only Chile reached from prospering to stable. Mexico, a US-dependent economy, actually only reached to the stable state. By comparison, Japan suffered the downturn each year, as opposed to their neighbors Singapur and

Hong Kong.

The Figure 9 shows a pattern that the crisis started in developed economies (at $t = 2008$) and propagated to emerging and then frontier ones (at $t = 2009$). And this is the reason why most of the most important nodes (high centrality) are developing economies: they connect the developed groups with the frontier and other emerging countries. At the end of 2008 and 2009, the number of infected markets was approximately 30, however, in 2010, when India got infected, the number increased to 37. In 2011, only Italy was infected and the infection reduced to 12. It also important to note that, only when Brazil is infected, its neighborhood also does it as it happened after the Chinese stock market crash of 2015.

Finally, the average recovery time per country was of 2.14 years, which means that $\gamma$ is 0.467. Given the crisis was not a complete epidemic, $R_0 < 1$ or $\gamma > \beta$. So, the infection rate of that 2008 crash was a positive value less than 0.467.

## 5 Discussion

In this work, we characterize the stock market network to describe the topology of the graph, perform community detection techniques to prove that market communities follow their classification as a developed, developing or frontier economy, and build a dynamical representation of the failure propagation during the 2008 crisis. In addition, this study analyzed the data from 74 countries to find more general patterns than country-specific ones.

These findings can help policymakers to be more prepared for the forthcoming financial crisis by identifying the community of the economy and the impact that the infected markets can generate. Furthermore, the threshold value analyzed here can help future studies to define a limit for the correlations and get better results, as Huang et al. (2009) did with the Chinese market [Huang et al., 2009]. The community detection has shown the communities that exist in the current world stock market, while the epidemic modeling has explored how the crisis propagated during a certain crisis. Even though this study addressed many issues, one of the limitations is the size of the data since the original set had to be cut due to incomplete fields and did not include monthly values. Future research can explore in testing different centrality measures as eigenvector centrality, besides that, the epidemic modeling of various economic crisis can reveal new patterns.

## 6 References


Barabási et al., 2016. Barabási, A.-L. et al. (2016). *Network science*. Cambridge university press.

Boginski et al., 2005. Boginski, V., Butenko, S., and Pardalos, P. M. (2005). Statistical analysis of financial networks. *Computational statistics & data analysis*, 48(2):431–443.

Clauset et al., 2004. Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Physical review E*, 70(6):066111.

Diekmann and Heesterbeek, 2000. Diekmann, O. and Heesterbeek, J. A. P. (2000). *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*, volume 5. John Wiley & Sons.

Dimitrios and Vasileios, 2015. Dimitrios, K. and Vasileios, O. (2015). A network analysis of the greek stock market.

Dobson et al., 2007. Dobson, I., Carreras, B. A., Lynch, V. E., and Newman, D. E. (2007). Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(2):026103.

Fortunato, 2010. Fortunato, S. (2010). Community detection in graphs. *Physics reports*, 486(3-5):75–174.

Fruchterman and Reingold, 1991. Fruchterman, T. M. and Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164.

Huang et al., 2009. Huang, W.-Q., Zhuang, X.-T., and Yao, S. (2009). A network analysis of the Chinese stock market. *Physica A: Statistical Mechanics and its Applications*, 388(14):2956–2964.

Kernighan and Lin, 1970. Kernighan, B. W. and Lin, S. (1970). An efficient heuristic procedure for partitioning graphs. *The Bell system technical journal*, 49(2):291–307.

Kim et al., 2007. Kim, K., Kim, S. Y., and Ha, D.-H. (2007). Characteristics of networks in financial markets. *Computer physics communications*, 177(1-2):184–185.

Mantegna, 1999. Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, 11(1):193–197.

Mantegna and Stanley, 1999. Mantegna, R. N. and Stanley, H. E. (1999). *Introduction to econophysics: correlations and complexity in finance*. Cambridge university press.

Newman, 2018. Newman, M. (2018). *Networks*. Oxford university press.

Newman, 2006. Newman, M. E. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical review E*, 74(3):036104.

Newman and Girvan, 2004. Newman, M. E. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113.

Raghavan et al., 2007. Raghavan, U. N., Albert, R., and Kumara, S. (2007). Near linear time algorithm to detect community structures in large-scale networks. *Physical review E*, 76(3):036106.

Reichardt and Bornholdt, 2006. Reichardt, J. and Bornholdt, S. (2006). Statistical mechanics of community detection. *Physical Review E*, 74(1):016110.

Smith, 1776. Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations.* William Strahan and Thomas Cadell.

Standard and Poors, 2012. Standard and Poors (2012). 2012 country classification review. Data retrieved from S&P Dow Jones Indices,`https://us.spindices.com/search/?ContentType=CountryClassifications`.

Strogatz, 2001. Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410(6825):268.

Sun et al., 2015. Sun, W., Tian, C., and Yang, G. (2015). Network analysis of the stock market.

Tabak et al., 2010. Tabak, B. M., Serra, T. R., and Cajueiro, D. O. (2010). Topological properties of stock market networks: The case of Brazil. *Physica A: Statistical Mechanics and its Applications*, 389(16):3240–3249.

World Bank, 2017. World Bank (2017). Market capitalization of listed domestic companies (current US$). Data retrieved from World Federation of Exchanges database,`https://data.worldbank.org/indicator/CM.MKT.LCAP.CD`.

## 7 Appendices

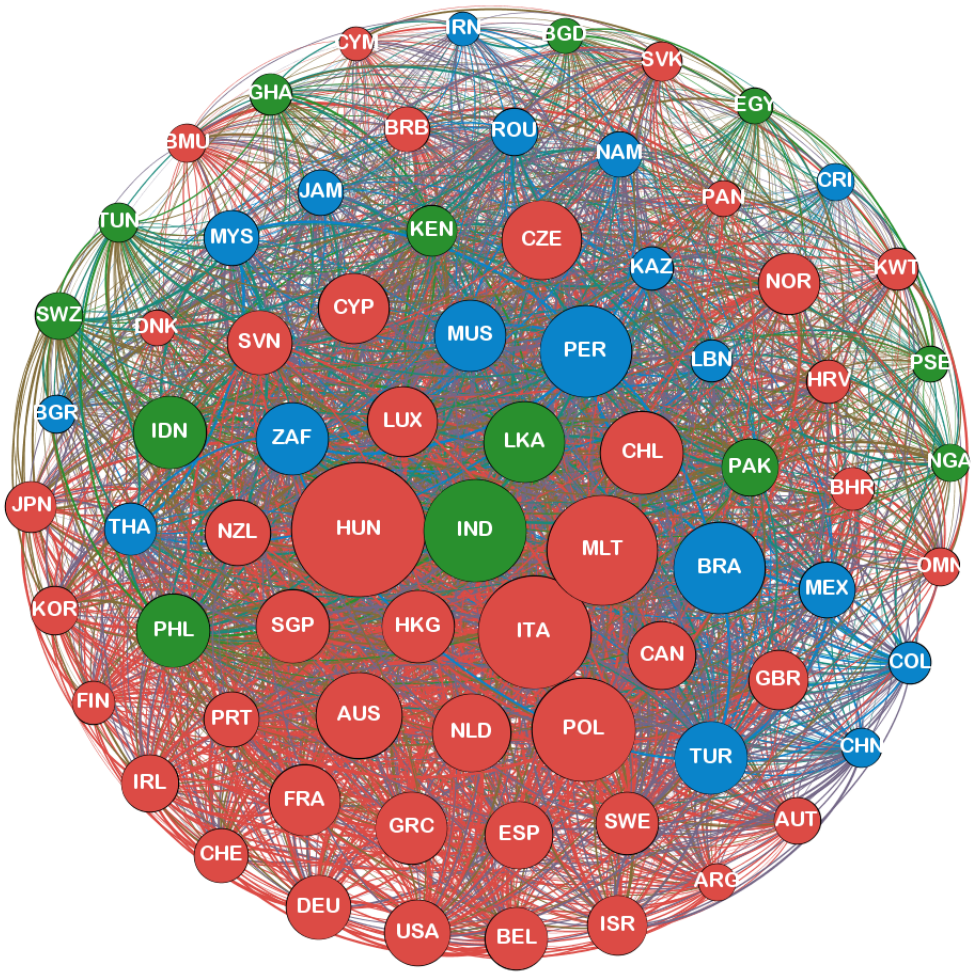


Figure 7: Visualization of network of stocks. Nodes colored by country income level and size represents the weighted degree.

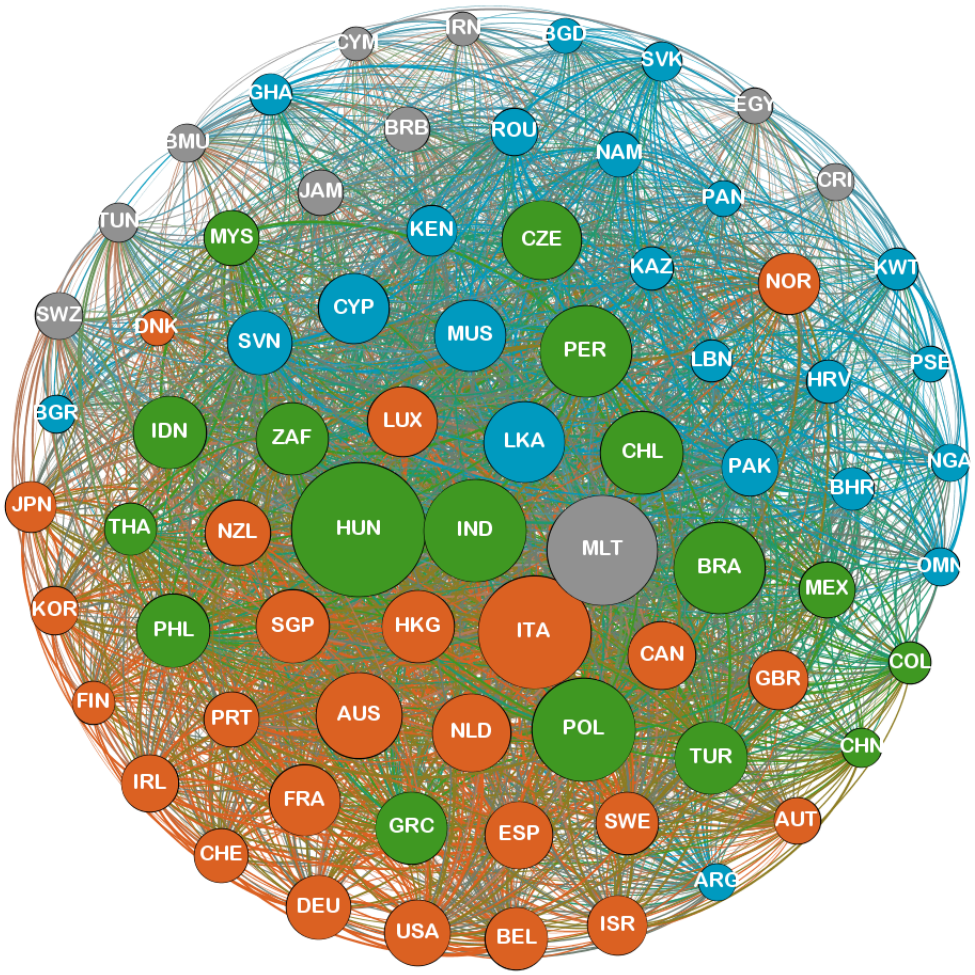


Figure 8: Visualization of network of stocks. Nodes colored by country economy type and size represents the weighted degree.

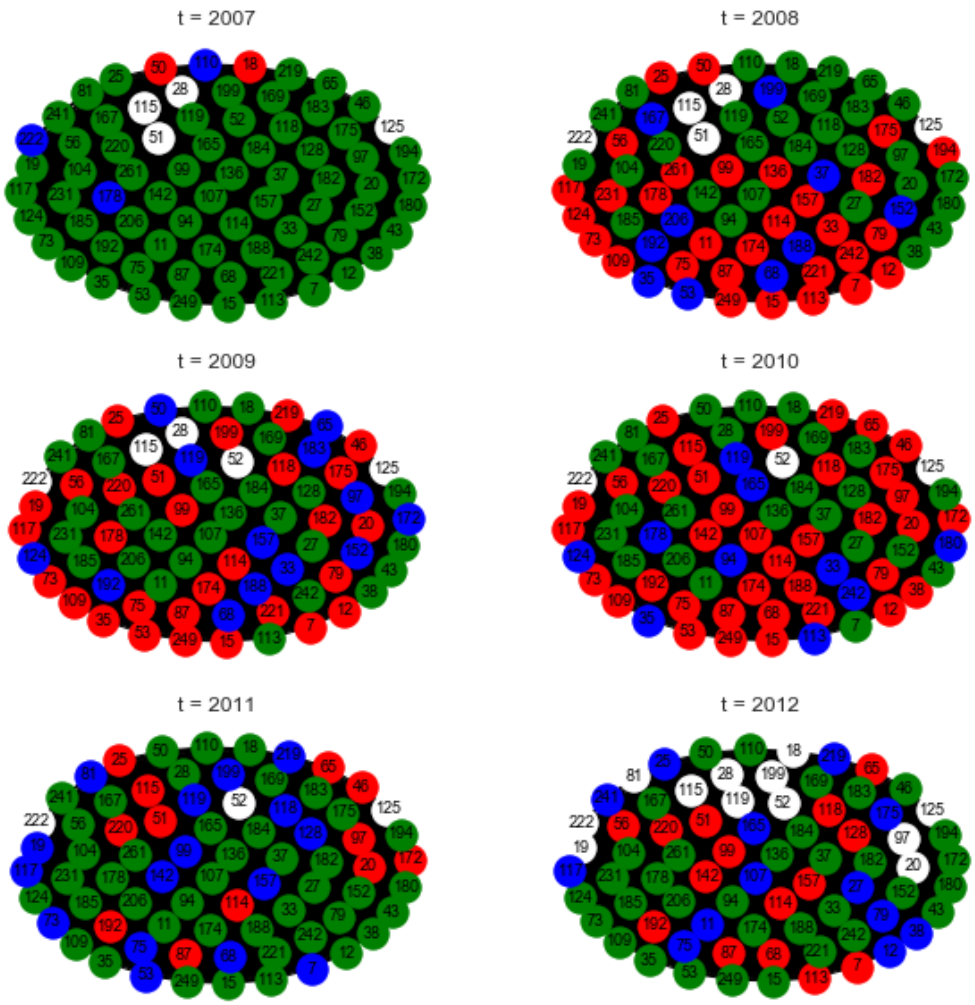


Figure 9: Visualization of the market crises from 2007 to 2012. The year 2012 shows an increase of infected nodes due to the European debt crisis.