

Relationships Between World Stock Market Indices: Evidence from Economic Networks

Vojtěch Fučík

Abstract The main objective of this paper is to analyse the relationships between world stock market indices. Classical minimum spanning tree approach is used, supported by edge stability analysis via bootstrap technique. Centrality measures analysis is applied for finding hubs in the network. Stock market indices are found to cluster according to the geographical continent they belong to. Hang Seng Index and CAC 40 are argued to be the most influential and important indices in the world. Evidence of large intra-continental and lower intercontinental stock market integration is observed.

Keywords Minimum spanning tree • Clustering • Stock market indices

Introduction

The study of topological properties of complex systems such as economic networks has become very popular. This growing interest may be observed not only in economics and finance but also in natural sciences or medicine.

Our main aim in this paper is to investigate the similarities among world stock market indices. Each index is represented as a node in the network. Edge weights in this network are then connected to pairwise cross-correlations. This network, however, contains too much information. For this purpose, we need to apply an algorithm filtering the network. By filtering we mean the removal of the weakest links. As an algorithm for filtering the network, we utilize the minimum spanning tree (MST) approach. After construction of the MST, we determine stability of its links. In the end, our task is to determine which node is central in the MST network, or in other words, we are trying to answer the question which stock market index plays a role of a hub in the network.

V. Fučík (✉)

University of Economics, Prague, Department of Banking and Insurance, Nam. W. Churchilla
4, 130 67 Prague, Czech Republic
e-mail: Fucik.V@seznam.cz

© Springer International Publishing AG 2018

D. Procházka (ed.), *The Impact of Globalization on International Finance and Accounting*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-319-68762-9_5

43

In this paper, we utilize the methodology provided by Mantegna (1999) for the construction of the MST. The MST construction is then accompanied by link stability analysis described in Tumminello et al. (2005) in more detail.

A milestone in the research of topological properties of stock markets is represented by an early work of Mantegna (1999), where he analyses a hierarchical structure of stocks traded on the US stock exchanges. Mantegna analyses portfolios of stocks used to compute S&P 500 and DJIA indices. He finds out that the investigated US stocks cluster according to the industry sector they belong to.

Bonanno et al. (2000) study the links between different world economies through the analysis of time series of the sets of stock market indices. Bonanno et al. (2001) find out that the pairwise cross-correlations between stock returns vary with the changing time horizon.

Although the vast majority of studied stock markets are American or British, there also exist papers focusing on different markets. For instance, Sandoval (2012) studies topological properties of Brazilian stock market, and Situngkir and Surya (2005) researched into the consequences caused by different lengths of the MSTs. Situngkir and Surya (2005) claim that the total length of the MST varies over time. They also find out that a higher length of the MST indicates stock market stabilization after previous monetary crisis.

Tabak et al. (2010) are engaged in the investigation of topological properties of commodities markets. They find out that the commodities form clusters according to the sector they belong to. Methodology similar to the one used in Mantegna (1999) is also often applied to the study of volatility of stock returns, e.g. Micciche et al. (2003).

The paper is structured as follows. In Data and Methodology we present a brief review of methodology utilized in the paper. Data and Methodology also provides data description and subsequent data transformations and statistical tests. We provide the results of our analysis in Results and Discussion. Finally Conclusion concludes the paper.

Data and Methodology

We analyse the price evolutions of 21 most important global stock market indices. Analysed indices named with their tick symbols are presented in Table 1. All data were downloaded from Yahoo Finance.

Our data span the period from January 2007 to February 2017, altogether 2405 daily observations.

Let us denote $P_i(t)$ the adjusted closing price of the i -th index at time t , $t = 1, \dots, T$, where T corresponds to the last observed time period. We are able to express logarithmic daily returns as follows:

Table 1 Analysed world stock market indices

| Europe | USA | Americas | Asia/Pacific |
|--------|--------|----------|--------------|
| ATX | DJI | MERV | AORD |
| BFX | IXIC | BVSP | SSEC |
| FCHI | GSPC | MXX | HSI |
| GDAXI | GSPTSE | | BSESN |
| SSMI | | | JKSE |
| | | | KLSE |
| | | | N225 |
| | | | KS11 |
| | | | TWII |

Source: Yahoo Finance

$$r_i(t) \equiv \frac{P_i(t) - P_i(t-1)}{P_i(t-1)}, \quad t = 2, 3, \dots, T. \quad (1)$$

Since for our subsequent analysis we need to work with stationary data, we cannot use highly nonstationary adjusted closing prices but rather logarithmic returns defined in Eq. 1 as is common in scientific literature.

The similarities in a group of stock market indices are standardly expressed using pairwise cross-correlations. The pairwise cross-correlation coefficient ρ_{ij} for a pair of indices i and j with logarithmic returns $r_i(t)$ and $r_j(t)$, respectively, $t = 2, \dots, T$, can be calculated as follows:

$$\rho_{ij} = \frac{\langle r_i(t)r_j(t) \rangle - \langle r_i(t) \rangle \langle r_j(t) \rangle}{\sqrt{\langle r_i^2(t) - \langle r_i(t) \rangle^2 \rangle \langle r_j^2(t) - \langle r_j(t) \rangle^2 \rangle}} \quad (2)$$

where symbol $\langle \cdot \rangle$ stands for the average over the studied period T . The correlation coefficients range between -1 (perfect anti-correlation) and 1 (perfect correlation). When $\rho_{ij} = 0$, then a pair of indices is uncorrelated. Since we assume n stock market indices, we finally obtain a square correlation matrix $\mathbf{C} = (\rho_{ij})_{i,j=1, \dots, n}$ of size n .

Using a nonlinear transformation proposed by Mantegna (1999), cross-correlation coefficients are then transformed into the distance coefficients d_{ij} as follows:

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \quad (3)$$

The distance coefficients range between 0 (perfect correlation) and 2 (perfect anti-correlation). When $d_{ij} = \sqrt{2}$ then a pair of indices (i, j) is uncorrelated. Finally we may arrive at a square symmetric distance matrix $\mathbf{D} = (d_{ij})_{i,j=1, \dots, n}$ with zeros on diagonal.

Now we are able to construct a graph $G = (V, E)$, where V is a set of stock market indices and E is a set of edges between these vertices. The weights of the edges are

already provided in the distance matrix **D**. The graph **G**, however, contains too much information for an easy interpretation. There are $\frac{n(n-1)}{2}$ unique edges between pairs of indices in **G**.

To address this problem, we utilize Kruskal's algorithm for finding the MST of **G**. The main idea of the MST procedure is that it filters the edges between pairs of indices such that only the most important ones prevail.

The algorithm proceeds as follows. First it finds an edge with a minimum weight and marks it. It then searches for edges with minimum weights among all the unmarked edges that do not produce a loop with the already marked edges. The algorithm stops when it succeeds in finding a spanning tree, i.e. when the set of marked edges contains $n - 1$ elements. Kruskal's algorithm produces a MST which is not unique. A number k of different MST produced by the algorithm may exist, but the sum of weights of the edges in the MST is minimized and unique, or

$$\sum_{e \in MST_1} w(e) = \sum_{e \in MST_2} w(e) = \dots = \sum_{e \in MST_k} w(e) \quad (4)$$

where $w(e): E_i \mapsto [0, 2]$ is a weight function that assigns a weight (distance) to every edge e from the set of edges $E_i \subset E$ corresponding to the $MST_i, i = 1, 2, \dots, k$.

After the MST procedure is completed, we arrive at a connected graph that contains only $n - 1$ unique edges between pairs of indices and does not contain any loop. Now we can carry out centrality measures analysis. The property of centrality of a particular node coincides with a size of influence this node has in the whole network.

We will assess node's centrality according to the following four criteria: (1) degree, (2) closeness, (3) betweenness and (4) eigenvector centrality. The proper definitions of the above-mentioned measures can be found in Sandoval (2012). For our purposes it, however, suffices to state very informal definitions. *Degree* of a node is the number of neighbours this node has in the network. The higher the degree, the more influence the particular node has in the network. *Closeness* of a node measures the average distance from it to all other nodes in the network. The higher the closeness, the more influence in the network the particular node has. *Betweenness* centrality answers the question whether it is necessary to involve a particular node in communication between any other two nodes, in other words whether the particular node is transitory between other pairs of nodes in the network. If it is, then the particular node is important. *Eigenvector* centrality boosts the strength of degree centrality by taking into account not just the number of neighbours but also their own influence in the network.

The simplicity of the MST methodology is redeemed by its serious drawback. Sometimes we cannot be sure whether a particular edge in the constructed MST is present because of its relevance for the network or only by coincidence. This edge stability problem can be analysed using bootstrap technique described in Tumminello et al. (2007).

In this method we first construct the original MST. Then we resample or bootstrap the original time series such that the length of the series is fixed. Some observations may be repeated, and some may not be present at all in the bootstrapped sample. Then, we construct the MST from the bootstrapped series, and we record the links. Such procedure is repeated up to 1000 times. The stability of a particular link in the original MST is then determined as a proportion of number of occurrences of a particular link in the bootstrapped MSTs and the number of bootstrap samples 1000.

Results and Discussion

In this section we present and comment on the results of the MST for the studied network of stock market indices. In Fig. 1 we illustrate the MST for stock market indices together with stability of the links.

We detected three main clusters of stock market indices:

1. Asia/Pacific
2. Europe
3. USA + Americas

Therefore, we may argue that the stock market indices do cluster according to the geographical location, which is the main finding of this paper. We also detected that a large cluster of Asia/Pacific can be further divided into two smaller sub-clusters. N225, AORD, TWII and KS11 are Japanese, Australian, Taiwanese and South Korean indices, respectively. We may also argue that this subcluster is composed of indices of more “western-style-oriented” countries, although the definition of “western-style” country is not very rigorous.

Regarding stability of the edges, there are several findings worth noting. First, the stabilities among stock market indices in all three clusters are very high, often as high as 100%. This again proves large interdependence between indices situated in similar geographical locations. Second, the stability of the transitory edges between these three main clusters, i.e. HSI – ATX and GDAXI – GSPC, is very poor. Thus we may argue that relations and similarity between indices from particular geographical areas/cluster are not that high. Based on this and above-mentioned findings, we may argue that the globalization or integration is very high on the local regional level (in terms of continents), but there exist further possibilities for enlarged stock market integration between individual continents (with the exception of USA + Americas). This might suggest a possibility for international diversification while investing in the stocks, i.e. the similarity between stock market indices and different clusters/regions has not reached the levels of similarity that are observable between indices in the particular clusters/regions.

In order to determine the most influential nodes in the network (MST), we utilized centrality measures described in Data and Methodology. The results are depicted in Table 2 in Appendix. Based on degree centrality, there is always one

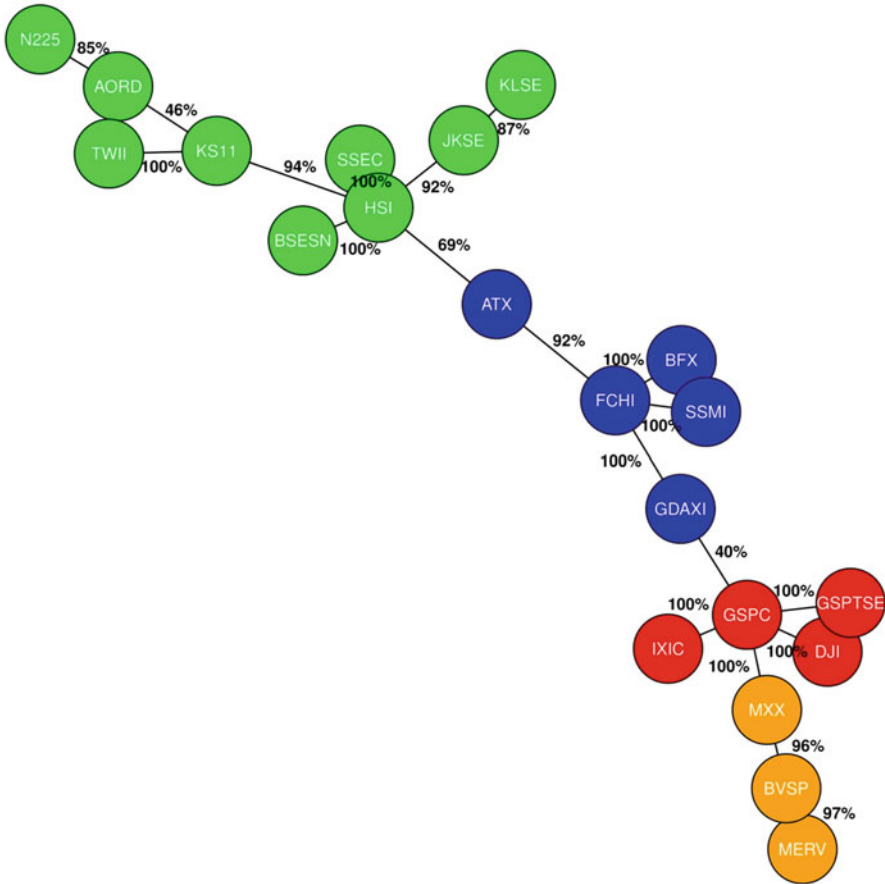


Fig. 1 MST of stock market indices network (Source: Author’s computations)

central stock market index for every cluster (geographic area): for Asia/Pacific it is HSI (Hang Seng Index), for Europe it is FCHI (CAC 40), and for USA + Americas it is GSPC (S&P 500). We can also observe the relative importance of KS11 (KOSPI Composite Index) in the Asia/Pacific cluster which again relates to the possible existence of two main subclusters in the overall Asia/Pacific region.

Using betweenness centrality we can draw several conclusions. There are two most transitory indices in the network – FCHI and HSI which further proves the hypothesis about these two indices being the most important in the network. Relative importance of ATX (Austrian Traded Index), GDAXI (DAX 30) and GSPC is also obvious from Table 2. However, importance of KS11 as a means of communication between other indices was not supported by betweenness centrality. Therefore, we can think of KS11 only as a local centroid not the global one.

Finally eigenvector centrality further supports the local dominance of ATX in the Europe cluster (apart from FCHI). HSI can be thought of as an index with the largest number of the most influential neighbours in the network. The results of closeness centrality further indicate most dominant role of FCHI and HSI indices in the network.

To sum up the centrality measures analysis, we found that

1. Every cluster/region has its own hub: FCHI (Europe), HSI (Asia/Pacific) and GSPC (USA + Americas).
2. FCHI and HSI can be thought of as hubs for the entire network, i.e. these are the most influential indices in the network (world).
3. American indices (mainly GSPC) do not play such an overwhelming role as is usually believed (based on our analysis).
4. European indices (ATX, FCHI, GDAXI) are transitory indices, i.e. they are very open and needed in the communication between other global stock market indices.
5. Indices from Americas region are situated in the periphery of the network without much communication and influence of global importance.

Conclusion

We analysed the relationships between world stock market indices in the period of 2007–2017 using the MST approach. We managed to detect three large clusters of indices based on the geographical location of underlying indices. We proved the integration of stock markets on the level of continents but stated that there are possibilities for global stock market integration enlargement still present. This might suggest a possibility for international diversification when investing in stocks. The clustering into three large clusters may not be based solely on geographic location but also on culture and investment style in these locations. For the question of whether this is the case, however, our analysis cannot provide relevant answers.

We managed to detect large interdependence between US and Americas stock market indices, but we argued that the global importance of Americas indices is low.

Using the centrality measures analysis, we argued that every region/cluster has its own hub: HSI for Asia/Pacific, FCHI for Europe and GSPC for the USA. We also argued that based on our analysis the most important stock market indices in the global scale are HSI and FCHI.

The added value of this paper for the investment society is straightforward. Since we are able to describe both visually and quantitatively the hierarchical structure of the stock market indices network, we are also able to detect those indices situated in the periphery of this network. Following the results of Pozzi et al. (2013), we may thus claim that there can be built a well-diversified portfolio that effectively reduces the investment risk.

There exist many possibilities for further research. Mainly we need to mention the MST dynamics. It requires the observation of the changing MSTs with respect to different time windows. This possibility is motivated by our assumption that the MSTs would differ, if we analysed only precrisis period, crisis period or postcrisis period. This kind of research could suggest whether any particular index had set for different trajectory of closing price evolution owing to the postcrisis stock market consolidation.

Acknowledgements This article was supported by IGA of the Faculty of Finance and Accounting, VŠE, (Project No. IG102027, Pokročilé metody risk managementu s použitím umělé inteligence, 59/2017) and by the Czech Science Foundation (Project No. GA 17-02509S).

Appendix

Table 2 Centrality measures

| Index | Degree | Closeness | Betweenness | Eigenvector |
|--------|--------|-----------|-------------|-------------|
| ATX | 2 | 0.024 | 99 | 0.514 |
| BFX | 1 | 0.020 | 0 | 0.035 |
| FCHI | 4 | 0.024 | 117 | 0.172 |
| GDAXI | 2 | 0.023 | 91 | 0.037 |
| SSMI | 1 | 0.019 | 0 | 0.046 |
| DJI | 1 | 0.019 | 0 | 0.002 |
| IXIC | 1 | 0.018 | 0 | 0.002 |
| GSPC | 5 | 0.020 | 96 | 0.019 |
| GSPTSE | 1 | 0.016 | 0 | 0.006 |
| MERV | 1 | 0.011 | 0 | 0.001 |
| BVSP | 2 | 0.014 | 19 | 0.003 |
| MXX | 2 | 0.017 | 36 | 0.007 |
| AORD | 2 | 0.014 | 19 | 0.221 |
| SSEC | 1 | 0.015 | 0 | 0.462 |
| HSI | 5 | 0.022 | 117 | 1.000 |
| BSESN | 1 | 0.016 | 0 | 0.407 |
| JKSE | 2 | 0.017 | 19 | 0.490 |
| KLSE | 1 | 0.013 | 0 | 0.203 |
| N225 | 1 | 0.012 | 0 | 0.083 |
| KS11 | 3 | 0.018 | 53 | 0.505 |
| TWII | 1 | 0.014 | 0 | 0.176 |

Source: Author's computations

References

- Bonanno G, Lillo F, Mantegna RN (2001) High-frequency cross-correlation in a set of stocks. *Quant Financ* 1:96–104. <https://doi.org/10.1080/713665554>
- Bonanno G, Vandewalle N, Mantegna RN (2000) Taxonomy of stock market indices. *Phys Rev E* 62(6):7615–7618. <https://doi.org/10.1103/PhysRevE.62.R7615>.
- Mantegna RN (1999) Hierarchical structure in financial markets. *Eur Phys J B* 11:193–197. <https://doi.org/10.1007/s100510050929>
- Micciche S, Bonanno G, Lillo F, Mantegna RN (2003) Degree stability of a minimum spanning tree of price return and volatility. *Phys A* 324:66–73. [https://doi.org/10.1016/S0378-4371\(03\)00002-5](https://doi.org/10.1016/S0378-4371(03)00002-5)
- Pozzi F, Matteo TD, Aste T (2013) Spread of risk across financial markets: better to invest in the peripheries. *Sci Rep* 3. <https://doi.org/10.1038/srep01665>
- Sandoval LJ (2012) A map of the Brazilian stock market. *Adv Complex Syst* 15(5). <https://doi.org/10.1142/S0219525912500427>
- Situngkir H, Surya Y (2005) On stock market dynamics through ultrametricity of minimum spanning tree. *Tech Rep Surya Res Inter*. <https://dx.doi.org/10.2139/ssrn.768204>.
- Tabak B, Serra T, Cajueiro D (2010) Topological properties of commodities networks. *Eur Phys J B* 74(2):243–249. <https://doi.org/10.1140/epjb/e2010-00079-4>
- Tumminello M, Aste T, Matteo TD, Mantegna RN (2005) A tool for filtering information in complex systems. *Proc Natl Acad Sci U S A* 102(30):10421–10426. <https://doi.org/10.1073/pnas.0500298102>
- Tumminello M, Coronello C, Lillo F, Micciche S, Mantegna RN (2007) Spanning trees and bootstrap reliability estimation in correlation based networks. *Int J Bifurcation Chaos* 17(7):2319–2329. <https://doi.org/10.1142/S0218127407018415>