

A clustering approach to examine the dynamics of the NASDAQ topology in times of crisis

Salim Lahmiri*

Department of Computer Science, University of Quebec at Montreal, Montreal, Canada

ARTICLE INFO

Article history:

Received May 15, 2012

Accepted 8 June 2012

Available online

June 9 2012

Keywords:

Financial crisis

Market dynamics

Wavelet analysis

Time series clustering

ABSTRACT

This paper investigates the dynamics of the NASDAQ topology before, during, and after 2008 financial crisis. First, multiresolution analysis by virtue of wavelet transform is employed to denoise each NASDAQ sector return series. Second, the correlation matrix of sectors is built and analyzed in each time period to view comovements of sectors. Third, hierarchical clustering trees are constructed in each time period to find out how the structure of the NASDAQ market evolves through time. Our results suggest that interrelationships between sectors become stronger in times of crisis and especially in post-crisis period. In addition, some markets tend to form the same cluster in all time periods; for instance the Industrial and Bank sectors and the Telecommunication and Computer sectors. However, the general topology of the NASDAQ market has been considerably changed over periods. In sum, the complex structure of the NASDAQ market is dynamic and is more integrated after 2008 financial crisis. This result indicates that there are less diversification opportunities in the post-crisis period in comparison with pre-crisis period. These empirical findings are important for the development of subsequent portfolio strategies.

© 2012 Growing Science Ltd. All rights reserved.

1. Introduction

There is a growing literature devoted to the study of the effect of 2008 financial crisis on US financial markets and economy; including the effect on bank lending (Ivashina and Scharfstein, 2010), corporate spending (Campello et al., 2010), federal fund market (Afonso et al., 2011), corporate governance (Erkins et al., 2012), and hedge fund trading (Ben-David et al., 2012). The purpose of this study is to examine the evolution of the topology of the stock market by studying the interconnections between sectors before, during and after 2008 financial crisis. Indeed, stock market changes affect portfolio value and this could be very harmful for investor's wealth if the portfolio is not suitably balanced; especially in times of financial crisis and economic depression. For instance, the 2008 financial crisis is one of the most serious financial crisis since the Great Depression and caused great financial instability in the US stock markets and decrease in portfolio wealth. In this context, it is important to study the dynamics of the stock market sectors before, during and after the most recent

* Corresponding author.

E-mail addresses: lahmiri.salim@courrier.uqam.ca (S. Lahmiri)

financial crisis since a well diversified stock portfolio should hold stocks across most of sectors in the financial market. The stock market is a complex network where price changes have complex interrelations. Consequently, connections among stock market sectors are important in studying the dynamics of industrial networks in the context of active portfolio management. In other words, it is important for both portfolio managers and investors to analyze the dynamics of the stock market topology through time in order to adjust portfolio weights to minimize risk. In particular, if a high correlation between industrial sectors exists before, during, and after financial crisis than one may conclude that those markets are exposed to common factors.

In this study, the dynamics of the interconnections of the stock market sectors are analyzed by virtue of hierarchical trees (Liao, 2005). A hierarchical tree is a clustering method where data objects are grouped into a tree of clusters. The goal of clustering is to determine the structure (topology) in an unlabeled data. In particular, data is organized into homogeneous groups where the within-group-object similarity is minimized and the between-group-object dissimilarity is maximized. As mentioned previously, the stock market is a complex system. Thus, hierarchical trees are appropriate to study the dynamics of the interconnections of the stock market sectors. Indeed, a quantitative description of hierarchies of a system is a key to model complex systems (Anderson, 1972). In this paper, the wavelet transform (Percival and Walden, 2000) is employed to denoise the returns series of each sector in the stock market. As a result, the denoised return series represent the trend properties of the original return time series. Then, clustering is performed using the denoised data.

The paper is organized as follows. Section 2 presents the methodology, whilst Section 3 provides data and show empirical results. Finally, Section 4 presents the main conclusions.

2. Methodology

Wavelet decomposition

Wavelet analysis is a mathematical method widely used in science and engineering problems for signal processing. The discrete wavelet decomposition of a given signal (market time series) $f(t)$ is given by:

$$f(t) = \sum_{k=-\infty}^{+\infty} a_k \phi_k(t) + \sum_{k=-\infty}^{+\infty} \sum_{j \geq 0} d_{j,k} \psi_{j,k}(t), \quad (1)$$

where a_k and $d_{j,k}$ are respectively the low pass and high pass coefficients of the wavelet transform, j and k are real and positive scaling parameters, and $\phi(t)$ and $\psi(t)$ are respectively the father wavelet and the mother wavelet which satisfy:

$$\int \phi dt < \infty, \quad \int \psi dt < 0, \quad \int \phi^* \psi dt = 0, \quad \int |\phi|^2 dt = \int |\psi|^2 dt = 1, \quad \text{and} \quad \psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k). \quad (2)$$

The father wavelet allows analyzing the smooth part of the signal to obtain its low frequency component (denoised signal), whereas the mother wavelet analyzes the signal to obtain the high-frequency (detail) parts of it. The approximation (smooth) coefficients a_k (denoised signal) and the details coefficients $d_{j,k}$ are given by:

$$a_{j,k} = \sum_n \ell(n - 2k) a_{j+1,n}, \quad (3)$$

$$d_{j,k} = \sum_n h(n - 2k) a_{j+1,n}, \quad (4)$$

where ℓ and h are the low-pass and high-pass filter respectively.

Hierarchical clustering

The agglomerative hierarchical clustering which is used in this study is a bottom-up partitioning algorithm that starts in treating each financial time series as a cluster. Then at each higher level, a new cluster is generated by joining the two closest clusters at the lower level. As a result of the clustering algorithm, a hierarchical tree of the elements of the system which is composed of stock market sectors is obtained. In this paper, the correlation distance is used to evaluate the distance between two different time series. The clustering algorithm is applied to the correlation matrix between pairs of elements as a similarity measure. The correlation distance metric $corr$ between wavelet approximation coefficients (a) of two different financial time series p and q is given by:

$$corr(a_p, a_q) = 1 - \frac{(a_p - \bar{a}_p)(a_q - \bar{a}_q)^T}{\left((a_p - \bar{a}_p)(a_q - \bar{a}_q)^T \right)^{0.5} \left((a_p - \bar{a}_p)(a_q - \bar{a}_q)^T \right)^{0.5}}, \quad (5)$$

where \bar{a} is the average and T is the transpose operator.

Based on distances between wavelet approximation coefficients, a linkage method is used to find the distance between two wavelet approximation coefficients clusters. The Ward's linkage (Ward, 1963) method is used in this study. For two clusters c_1 and c_2 , the Ward's linkage is given by:

$$d(c_1, c_2) = \sqrt{\frac{2n_1n_2}{n_1 + n_2}} \|\bar{c}_1 - \bar{c}_2\| \quad (6)$$

where n and \bar{c} are respectively the number of samples and the centroid of a cluster.

3. Data and empirical results

In this study, daily closing price for the NASDAQ industrial sectors are collected for the period from January 2006 to December 2011 corresponding to approximately 1492 data points for each sector. For instance, the NASDAQ indices considered in this paper (see Figure 1) are Biotechnology, Financial 101, Industrial, Bank, Telecommunication, Computer, Insurance, Other Finance, and Transport sector. The overall sample is divided into pre-crisis (January 3rd 2006 to June 15th 2007), crisis (June 16th 2007 to March 9th 2009), and post-crisis (March 10th 2009 to December 30th 2011). As observed graphically in Figure 1, there was a significant upward, downward, and upward respectively during pre-crisis, crisis, and post-crisis period. Figure 2 shows the wavelet approximation (denoised signal) and detail coefficients of the returns for each sector. The returns are computed as first difference of the logarithmic of the index price level. Tables 1, 2, and 3 provide the correlation matrix (The heatmap of bivariate correlations) of the NASDAQ sectors during pre-crisis, crisis, and post-crisis period respectively.

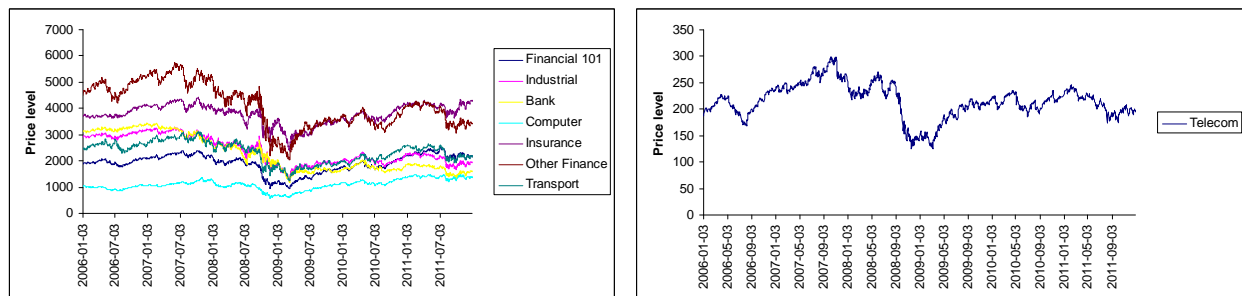


Fig. 1. NASDAQ sectors from January 2006 to December 2011

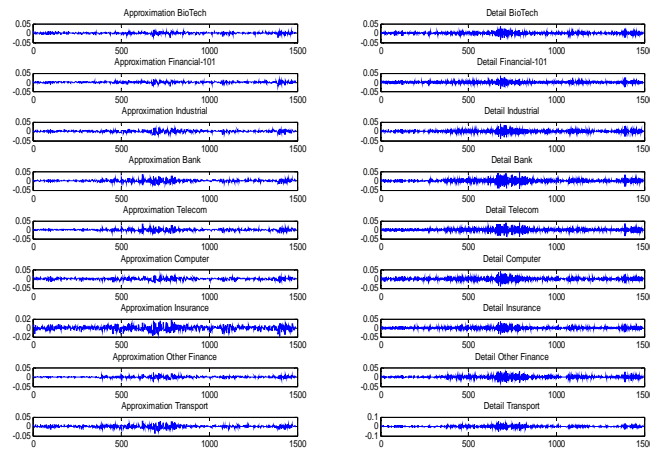


Fig. 2. Wavelet analysis of NASDAQ indices from January 2006 to December 2011

Table 1

Pre-crisis correlation heatmap

	Biotech	Financial 101	Industrial	Bank	Telecom	Computer	Insurance	Other Finance	Transport
Biotech		0.7867	0.6491	0.5741	0.6512	0.7122	0.6248	0.7095	0.6550
Financial 101			0.8075	0.7216	0.8044	0.8495	0.7690	0.8634	0.8324
Industrial				0.9570	0.6376	0.6756	0.8486	0.9047	0.7099
Bank					0.5420	0.5669	0.8113	0.7759	0.6347
Telecom						0.8349	0.5932	0.7167	0.6646
Computer							0.6272	0.7688	0.7329
Insurance								0.7634	0.6515
Other Finance									0.7625
Transport									

correlation > 0.90
0.9 > correlation > 0.8
0.8 > correlation > 0.7
correlation < 0.7

Table 2

Crisis correlation heatmap

	Biotech	Financial 101	Industrial	Bank	Telecom	Computer	Insurance	Other Finance	Transport
Biotech		0.7721	0.6330	0.5913	0.6648	0.6766	0.6972	0.6450	0.6851
Financial 101			0.8363	0.7653	0.8905	0.8955	0.8044	0.8804	0.8770
Industrial				0.9716	0.7530	0.7899	0.8715	0.9435	0.8223
Bank					0.6763	0.7111	0.8526	0.8532	0.7749
Telecom						0.9063	0.6926	0.8156	0.7754
Computer							0.7138	0.8397	0.7818
Insurance								0.8036	0.7591
Other Finance									0.8275
Transport									

correlation > 0.90
0.9 > correlation > 0.8
0.8 > correlation > 0.7
correlation < 0.7

Table 3

Post-crisis correlation heatmap

	Biotech	Financial 101	Industrial	Bank	Telecom	Computer	Insurance	Other Finance	Transport
Biotech		0.8423	0.7616	0.7152	0.7400	0.7862	0.7387	0.7841	0.7636
Financial 101			0.8847	0.8283	0.8846	0.9348	0.8319	0.9155	0.9078
Industrial				0.9768	0.7870	0.8228	0.9061	0.9579	0.8718
Bank					0.7317	0.7610	0.8810	0.8852	0.8298
Telecom						0.8834	0.7653	0.8136	0.8107
Computer							0.7668	0.8672	0.8478
Insurance								0.8658	0.8261
Other Finance									0.8800
Transport									

correlation > 0.90
0.9 > correlation > 0.8
0.8 > correlation > 0.7
correlation < 0.7

The heatmaps of bivariate correlations indicate a strong relationship between the Industrial and Bank sectors and the Industrial and Other Finance sectors in the pre-crisis period. These correlations increased further during time of crisis. In addition, the correlation between the Telecommunication and Computer sectors increased significantly from 0.8349 in pre-crisis period to achieve 0.9063 in the crisis period. However, this correlation decreased to 0.8834 in the post-crisis period. Many other relationships between sectors became strongly pronounced in the post-crisis period. For instance, the correlations between Financial 101 and Computer, Financial 101 and Other Finance, Financial 101 and Transport, Industrial and Bank, Industrial and Insurance, and Industrial and Other Finance. The correlations between Other Finance and Bank, Other Finance and Telecommunications, Other Finance and Computer, Other Finance and Insurance were between 0.7167 and 0.7759 in the pre-crisis period and increased in the crisis period (between 0.8036 and 0.8532) and also in the post-crisis period (between 0.8136 and 0.8852). Similarly, Telecommunication, Computer, and Transport sectors became strongly correlated with the other sectors in times of crisis and particularly after the crisis. The bottom line is that the correlations between NASDAQ sectors are clearly dynamic; hence not static; and that the relationships between sectors become stronger in the post-crisis period. These findings show strong evidence that the topology of the stock market varies through time.

Cluster analysis by means of hierarchical trees is used to view how the structure of the stock market varies during pre-crisis, crisis, and post-crisis period. The agglomerative clustering results (dendrogram) based on the nine NASDAQ indices are given in Figure 3. As it is shown, the agglomerative cluster analysis separates the NASDAQ indices into groups of similar markets. NASDAQ markets progress from being individual and isolated at the bottom of the hierarchical clustering tree to being all clustered in a single group at the top of the hierarchical clustering tree. For instance, there are obvious divisions between the NASDAQ markets on the left hand and the right hand side of the tree in pre-crisis, crisis and post-crisis periods. The five main findings follow. First, Industrial, Bank, and Insurance sectors are in the same cluster in all subsamples. Moreover, Industrial and Bank sectors form altogether one distinctive group in all times. Similarly, Telecommunication and Computer were closely relation in pre-crisis, crisis, and post-crisis periods. In particular, they formed one distinctive group during the first two periods. Third, Financial 101 and Other Finance markets are closely related to form one cluster in pre-crisis and crisis period, but are completely separated in post-crisis period. Indeed, the first one belongs to the main group in the right hand of the clustering tree, and the second one belongs to the main group in the left hand of the clustering tree. Fourth, the Biotechnology sector is being individual and isolated in all periods as shown in the right hand of all hierarchical clustering trees. Finally, by comparing the groups that form the left hand and the right hand of the clustering trees in all periods it is found that the topology of the NASDAQ stock market has been structurally changed over time.

In sum, a relatively high degree of commonality is found for some of the NASDAQ sectors although there are still noticeable divisions between most of sub-groups of the NASDAQ markets. In particular, Industrial, Bank, and Insurance sectors, and Telecommunication and Computer sectors are unchanged through periods but the other NASDAQ sectors cluster differently. The evolution of the hierarchical clustering through periods indicates that certain homogeneity exists between some markets; for instance the Industrial and Bank sectors and the Telecommunication and Computer sectors. Thus, the components of each of these groups of sectors are driven by the same factors. In addition, the evolution of the hierarchical clustering tree through periods suggests that the Biotechnology sector could be an interesting alternative investment when seeking diversification; especially in the pre-crisis and crisis samples. All in all, the topology of the NASDAQ market is dynamic and is significantly different from a period to another. In addition, the relationships between sectors become stronger in the post-crisis period. In the context of portfolio optimization, these findings are important to be taken into account by portfolio managers to construct hierarchical structures between different sectors when investing in the NASDAQ market.

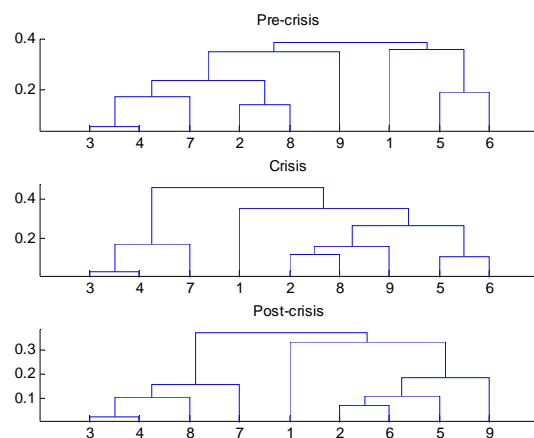


Fig. 3. Evolution of the hierarchical clustering tree of NASDAQ markets. Sector are Biotechnology (1), Financial 101 (2), Industrial (3), Bank (4), Telecommunication (5), Computer (6), Insurance (7), Other Finance (8), and Transport (9).

4. Conclusion

A clustering approach is used to examine the comovements of the NASDAQ sectors before, during and after 2008 financial crisis. Knowledge of these common relationships could help identify potential strategies for assets allocation and portfolio management. The wavelet transform is applied to return series for denoising. Then, an agglomerative hierarchical clustering tree is applied to the denoised series in each time period. It is found that the correlations between sectors become stronger in times of crisis and particularly in post-crisis period. Thus, common factors drive many components of the NASDAQ in the last period. In addition, the general structure of the NASDAQ market has been considerably changed over periods. Thus, the topology of the NASDAQ market is clearly dynamic. In addition, it is more integrated in post-crisis period. These empirical findings should be taken into account by investment managers when selecting portfolios. For future work, based on these findings portfolios will be simulated and profits evaluated in each period. In addition, formal statistical tests may be conducted to examine in depth the degree of cointegration among NASDAQ sectors.

References

- Afonso, G., Kovner, A., & Schoar, A. (2011). Stressed, not frozen: the federal funds market in the financial crisis. *The Journal of Finance*, 66 (4), 1109–1139.
- Anderson, P.W. (1972). More is different. *Science*, 177, 393–396.
- Ben-David, I., Franzoni, F., & Moussawi, R. (2012). Hedge fund stock trading in the financial crisis of 2007–2009. *The Review of Financial Studies*, 25(1), 1–54.
- Campello, M., Graham, J.R., & Harvey, C.R. (2010). The real effects of financial constraints: Evidence from a financial crisis. *Journal of Financial Economics*, 97, 470–487.
- Erkens, D.H., Hung, M., & Matos, P. (2012). Corporate governance in the 2007–2008 financial crisis: Evidence from financial institutions worldwide. *Journal of Corporate Finance*, 18, 389–411.
- Ivashina, V., & Scharfstein, D. (2010). Bank lending during the financial crisis of 2008. *Journal of Financial Economics*, 97, 319–338.
- Liao, T.W. (2005). Clustering of time series data—a survey. *Pattern Recognition*, 38, 1857–1874.
- Percival, D.B. & Walden, A.T. *Wavelet methods for time series analysis*. Cambridge, New York: Cambridge University Press, 2000.
- Ward, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 236–244.