An Analysis of Herding in the Korean Stock Market Using Network Theory

Soosung Hwang*
Sungkyunkwan University

Young-Il Kim#
Sungkyunkwan University

Jinho Shin♦
Sungkyunkwan University

Abstract

We investigate whether herd behavior in the equity market is led by 'core' stocks or by 'peripheral' stocks connected to core stocks, which we identify with a technique from network theory. Using non-securities stocks listed in the Korea Exchange from January 2005 to December 2015, we find strong evidence of herding in the Korean stock market, as in previous studies on herding. Herding arises only when the market is in stress: during bear states, core stocks herd toward the market portfolio and peripheral stocks herd toward core stocks in their clusters. During bull markets, however, adverse herding arises mainly driven by securities stocks, and thus cross-sectional dispersion in returns increases. Core stocks are not necessarily the stocks whose market values are large but instead are mid-sized stocks.

Keywords: Herd behavior, Cross-sectional standard deviation, Network analysis

JEL codes: G02, D85

* College of Economics, Sungkyunkwan University, 25-2 Sungkyunkwan-Ro, Jongno-gu, Seoul 110-745, South Korea. Tel: +82 (0)2 760 0489, Fax: +82 (0)2 744 5717, Email: shwang@skku.edu.

College of Economics, Sungkyunkwan University, 25-2 Sungkyunkwan-Ro, Jongno-gu, Seoul 110-745, South Korea. Tel: +82 (0)2 760 0489, Fax: +82 (0)2 744 5717, Email: kyizxc1@naver.com.

♦ Corresponding author: College of Economics, Sungkyunkwan University, 25-2 Sungkyunkwan-Ro, Jongno-gu, Seoul 110-745, South Korea. Tel: +82 (0)2 3788 2334, Fax: +82 (0)2 3788 2350, Email: jinho.shinn@gmail.com.

1. Introduction

Herding is an important element of behavior in financial markets as it can distort asset prices, leading to market inefficiency. Empirical studies have suggested some evidence of herding by market experts such as analysts or institutional investors from their clustering behavior (Welch, 2000; Barber, Odean, and Zhu, 2009; Choi and Sias, 2009; Hirshleifer and Teoh, 2009). These studies, however, do not necessarily indicate that asset prices are biased such that the efficient allocation of assets is disturbed. Other studies investigate the effects of herding on asset prices using cross-sectional dispersion of returns or betas (Christie and Huang, 1995; Chang, Cheng, and Khorana, 2000; Hwang and Salmon, 2004). They test if the cross-sectional dispersion of returns or betas decreases when the market is under stress and thus herding arises.

Herding may be more prominent within industries rather than in the entire equity market because signals and recommendations by financial analysts or decisions by business managers are often at the industrial level (Choi and Sias, 2009; Bikhchandani and Sharma, 2001; Yao, Ma, and He, 2014; Gebka and Wohar, 2013; Demirer, Lien, and Zhang, 2015). Although the connection between individual firms identified by industries is intuitively appealing, firms are connected for other reasons such as ownership connections (Anton and Polk, 2014), connections in trading (Shleifer and Vishny, 1992; Coval and Stafford, 2007), or pairs by co-integrated prices (Gatev, Goetzmann and Rouwenhorst, 2006). They may be connected because of their vertical relationships or because they belong to the same business family. Firm characteristics, e.g., size, book-to-market, liquidity, and growth (Harvey, Liu and Zhu, 2015), can also connect stocks for which investors face similar pricing problems.

In this study, we identify connections using network theory to investigate herding in the stock market. If connections identified by network theory can group stocks better than industries, the effects of herding on stock returns are more likely to be observed in connected stocks than in stocks grouped by industries. For this purpose, we reduce the complexity of financial dependencies between individual stocks using the minimum spanning tree (MST) proposed by Mantegna (1999). If market and industry are the only two connections that explain individual stocks, herding at the market and industry levels should represent irrational price distortion during market stress (Bikhchandani and Sharma, 2001; Demirer, Lien, and Zhang, 2015). However, if there are other types of connections such as those discussed above, herding at the market or industry level may not capture investor herd behavior in equity markets.

The purpose of this study is to investigate whether herd behavior in the equity market is led by a small number of 'core' stocks or by the 'peripheral' stocks connected to the core stocks, which we identify using the MST. Analyzing 533 non-securities stocks listed in the Korea Exchange from January 2005 to December 2015, we identify 36 core stocks. The top three core stocks, Keyang Electric Machinery, Hyundai BNG Steel, and Hanjin Heavy Industries and Construction, are connected to 50, 44, and 45 peripheral stocks, respectively. It is noteworthy that the largest two firms, Samsung Electronics and Korea Electric Power Corporation, are not identified as core stocks. When securities firms are included in the analysis, approximately half of the core stocks are in the securities sector. These securities firms hold a large amount of shares listed in the Korea Exchange, and thus their stock returns are closely connected to stocks in other sectors.

Using cross-sectional dispersion in returns as a herd measure (Christie and Huang, 1995; Chang, Cheng, and Khorana, 2000), we find strong evidence of herding when market returns are extreme. When the market is in stress, investors behave irrationally and cross-sectional dispersion in returns decreases, i.e., the returns of core stocks come closer to the market return, and those of peripheral stocks also approach the returns of core stocks in their clusters.

These results are different from herding decomposed by industries. As in Chang, Cheng, and Khorana (2000), we find evidence of herding for the entire sample period. However, when market states (bull and bear periods) are considered, we do not find herding in bear markets; we only find evidence of adverse herding in bull markets within-industry (cross-sectional dispersion of individual stocks with respect to their industry) and cross-industry (cross-sectional dispersion of industries with respect to the market). During bull markets, cross-sectional dispersion in returns increases and investors do not follow the movements of the market or of core stocks. However, evidence of adverse herding is found only when securities stocks are included in the analysis.

Our contribution to the literature can be summarized as follows. First, stocks can be grouped in an effective way using network theory to identify the characteristics and the behavioral patterns of independent entities – such as people, groups, and objects – through understanding the network structure. Many attempts have been made for equities, and the recent surge in social network analysis makes it possible to analyze the diverse channels through which researchers approach the topic. For the proponents of network analysis, the equity market is a complex network, and we explore this topic for a bias in investor behavior.

Second, this paper contributes to the existing research on herding by studying connections between individual stocks. Prior studies on herd behavior use various connections, including investor entities (i.e., individuals, foreigners, and institutions), the aggregate market, and the industrial level. For example, Christie and Huang (1995) investigate herding at the market level whereas Bikhchandani and Sharma (2001), Choi and Sias (2009), Yao, Ma, and He (2014), Gebka and Wohar (2013), and Demirer, Lien, and Zhang (2015) analyze herding at the industry level. Chen (2013) and Chang and Lin (2015) study herding behavior at the international level. On the other hand, some studies investigate herding for groups that are sorted by market capitalization (Chang, Cheng, and Khorana, 2000; Kim, 2013). We use connections identified by networks, which we believe describe price co-movements in the equity market better than industries or sizes.

This paper is organized as follows. In the following section, we describe how to construct the MST using Kruskal's (1956) algorithm and how to test herding using the network identified by the algorithm. In Section 3, we present the properties of core and peripheral stocks and report the empirical results for herding. Section 4 concludes our paper.

2. Networks in the Stock Market and Herding

In order to investigate herd behavior in networks, we first explain how to identify core and peripheral stocks using networks in the stock market and then propose testable models for the analysis of herd behavior of these two groups.

2.1. Analysis of Network and Clusters

Stocks are often grouped by industries because signals that investors receive, recommendations by financial analysts, and business decisions by managers are often at the industry level (Choi and Sias, 2009; Bikhchandani and Sharma, 2001; Yao, Ma, and He, 2014; Gebka and Wohar, 2013; Demirer, Lien, and Zhang, 2015).[1] However, industry is not the only way to group stocks. There are different types of connections between stocks that belong to different industries. Some examples of connections that are known to affect asset prices are

---

[1] Others investigate herding at the international level because of the globalization of financial markets (Gebka and Wohar, 2013; Chen, 2013; Chang and Lin, 2015).

ownership connections (Anton and Polk, 2014), connections in trading (Shleifer and Vishny, 1992; Coval and Stafford, 2007), or pairs by co-integrated prices (Gatev, Goetzmann and Rouwenhorst, 2006). Connections may also arise between firms that have a vertical relationship or firms that are owned by the same business family. When connections are identified by firm characteristics, e.g., size, book-to-market, liquidity, and growth, these characteristics can be used to form groups of stocks for which investors face similar pricing problems (Harvey, Liu and Zhu, 2015).

In this study, we use network theory to summarize these complex dimensions of connections in the stock market. A stock market network is constructed such that stocks in the market can be grouped into two groups, core stocks and peripheral stocks. Following Mantegna (1999), we use the distance measure to generate the minimum spanning tree (MST). The distance measure is calculated as follows using a Spearman's rank correlation coefficient (ρ_{ij}):[2]

$$d_{ij} = 1 - |\rho_{ij}|, \tag{1}$$

where i and j denote individual stocks i and j, respectively. The distance measure ranges from 0 to 1 and shows less correlation as its value approaches 1. When there are N individual stocks, $N(N-1)/2$ distances are calculated.

The distances are then used to construct the MST using Kruskal's (1956) algorithm. Kruskal's algorithm finds a subset of the distances and forms a tree that includes every stock, where the total weight of all the distances in the tree is minimized. More specifically, the MST method forms a network by sequentially selecting non-circular links with the shortest distance among $N(N-1)/2$ number of links. The MST method has an advantage in that it efficiently utilizes information by conserving most network properties (Cormen, Leiserson, Rivest, and Stein, 2009). With N stocks in the market, $N(N-1)/2$ correlations or distances can be reduced to $N-1$ links that have the shortest distance. For example, when N=1,000, we have approximately half a million links (correlations) to be analyzed, but using the MST algorithm, we only have 999 connections.

Kruskal's algorithm allows us to divide individual stocks into a certain number of coherent groups so that the minimum distance between stocks in different groups is maximized. There are no specific criteria for grouping and we use the following heuristic method for clustering.

[2] Spearman correlations are used in this study instead of Pearson correlations because of the non-normality of stock returns.

- Criterion 1: A stock that has at least K directly linked peripheral stocks.
- Criterion 2: A stock that has at least one link to another core stock.
- Criterion 3: A bridge stock (that exists between two core stocks) that has at least K directly or indirectly linked peripheral stocks.

The minimum number K of peripheral stocks linked to a core stock needs to be defined considering the number of clusters (the number of core stocks out of the total number of stocks). If K is too large, clusters may include less connected stocks and thus may not show investor herding by connection. On the other hand, if K is too small, the number of clusters increases too much and connected stocks may belong to different clusters. Criterion 2 explains that there should be only one link between two core stocks because the MST method requires that every stock must be linked, and thus, a single link between the clusters is considered as being little correlated. Criterion 3 assigns a bridged core stock and its peripheral stocks into a separate cluster when the bridged core stock, which serves as a connection between two core stocks, has at least K links to peripheral stocks.

2.2. Herd Measure and Testable Models

Various measures have been proposed to investigate herd behavior in financial markets. Lakonishok, Shleifer, and Vishny (1992) base their criterion on the trades conducted by a subset of market participants over a period of time. Wermers (1999) proposes a portfolio-change measure designed to capture both the direction and intensity of trading by investors. However, these measures do not directly show the effects of herding on asset prices. Christie and Huang (1995) argue that the magnitude of cross-sectional dispersion of individual stock returns decreases during large price changes when investors imitate the observed decisions of others in the market rather than follow their own beliefs and information.

In this study we investigate herding between connected stocks under the assumption that stocks with close connections are more affected by investor herding than those grouped by industries. If investors observe and follow movements of closely connected stocks, the prices of connected stocks may co-move via investors' herd behavior. Suppose the cross-sectional variance (CSV) in returns is:

$$\mathrm{CSV} = E[(r_{it} - r_{mt})^2], \tag{2}$$

where r_{it} and r_{mt} denote returns of stock i and the market at time t, respectively. The CSV can be decomposed into CSVs in core and peripheral stocks, as follows:

$$\begin{aligned} \text{CSV} &= E[(r_{it} - r_{mt})^2] \\ &= E[(r_{it} - r_{cit} + r_{cit} - r_{mt})^2] \\ &= E[(r_{it} - r_{cit})^2] + E[(r_{cit} - r_{mt})^2] \\ &= CSV^P + CSV^C, \end{aligned} \tag{3}$$

assuming $E[(r_{it} - r_{cit})(r_{cit} - r_{mt})] = 0$, where CSV^P is the CSV of peripheral stocks with respect to core stocks (r_{cit}) and CSV^C is the CSV of core stocks with respect to the market (r_{mt}).

In our study, we use cross-sectional standard deviations rather than cross-sectional variance for consistency with other previous studies. Cross-sectional dispersions (CSDs) are defined as follows:[3]

$$\text{CSD}_t = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(r_{it} - r_{mt})^2}, \tag{4}$$

$$\text{CSD}_t^P = \sqrt{\sum_{ci=1}^{N_c} w_{ci}\frac{1}{N_{ci}}\sum_{i=1}^{N_{ci}}(r_{it} - r_{cit})^2}, \tag{5}$$

$$\text{CSD}_t^C = \sqrt{\sum_{ci=1}^{N_c} w_{ci}(r_{cit} - r_{mt})^2}, \tag{6}$$

where N_c and N_{ci} represent the numbers of core stocks and their peripheral stocks linked to a core stock c, respectively, and $w_{ci} = \frac{N_{ci}}{N}$. See the Appendix for the details of the equations. When industry is used for grouping, r_{cit} is replaced with equally weighted industry returns.

2.3. Empirical models for testing herd behavior in the equity market

If investors' tendency to follow the market consensus increases during large market movements (Christie and Huang, 1995; Chang, Cheng, and Khorana, 2000), the cross-sectional dispersion decreases with the market volatility. To investigate this, Christie and Huang (1995) regress cross-sectional dispersions in returns on an intercept and two dummy variables designed to capture extreme positive and negative market returns. Negative coefficients on the dummy variables can be interpreted as evidence of herding.

In this study, we test this type of herding using the following regression:

$$\text{CSD}_t = \gamma_0 + \gamma_1^+|r_{mt}|I_{r_{mt}\geq 0} + \gamma_1^-|r_{mt}|\left(1 - I_{r_{mt}\geq 0}\right)$$

[3] See the Appendix for the details of the equations. Note that $\text{CSD}_t \neq \text{CSD}_t^P + \text{CSD}_t^C$.

$$+\gamma_2^{+} r_{mt}^2 I_{r_{mt}\geq 0} + \gamma_2^{-} r_{mt}^2 (1 - I_{r_{mt}\geq 0}) + \varphi \mathrm{CSD}_{t-1} + \varepsilon_t, \tag{7}$$

where ε_t is an error term, and $I_{r_{mt}\geq 0}$ equals one when the market return is positive or zero and zero otherwise. The lagged CSD_{t-1} is used as an explanatory variable because of the persistence of CSD_t. The coefficients on the absolute market return are expected to be positive, i.e., $\gamma_1^{+} > 0$ and $\gamma_1^{-} > 0$, because of a close association between market volatility and cross-sectional dispersion in returns (Hwang and Satchell, 2005). In this regression, we expect both γ_2^{+} and γ_2^{-} to be negative if investors follow others during large market movements. In particular, if investors follow others at large and negative market returns, we expect $\gamma_2^{-} < \gamma_2^{+} < 0$.

Herding may increase when markets are in stress (Christie and Huang, 1995). To investigate herding during periods of market stress, we test herding in different market states, i.e., bull and bear states. Motivated by the regime switching literature (e.g., Hamilton, 1989), we identify bull and bear states using the following simple regime switching model:

$$r_{mt} = \mu_1 S_{1t} + \mu_2 S_{2t} + \sigma_t \varepsilon_t, \tag{8}$$
$$\sigma_t = \sigma_1 S_{1t} + \sigma_2 S_{2t},$$

where r_{mt} is the market return, μ_i and σ_i are the expected market return and volatility of regime $i = 1, 2$, respectively, and the dummy (state) variable S_{it} is one when regime i is selected and zero otherwise. As in Hamilton (1989), the state variables are assumed to be governed by a first-order Markov chain. The regime switching model is estimated using the Bayesian Markov Chain Monte Carlo Gibbs sampling estimation. Once the two states are identified, they are named 'bull' and 'bear' states according to the characteristics of the expected market return and volatility.

The difference in herding between bull and bear states can be tested using the following regression equation:

$$\begin{aligned} \mathrm{CSD}_t = {} & \gamma_0 + \gamma_{1u}^{+} |r_{mt}| I_{r_{mt}\geq 0} I_{ut} + \gamma_{1u}^{-} |r_{mt}| \left(1 - I_{r_{mt}\geq 0}\right) I_{ut} \\ & + \gamma_{1d}^{+} |r_{mt}| I_{r_{mt}\geq 0} (1 - I_{ut}) + \gamma_{1d}^{-} |r_{mt}| \left(1 - I_{r_{mt}\geq 0}\right)(1 - I_{ut}) \\ & + \gamma_{2u}^{+} r_{mt}^2 I_{r_{mt}\geq 0} I_{ut} + \gamma_{2u}^{-} r_{mt}^2 \left(1 - I_{r_{mt}\geq 0}\right) I_{ut} \\ & + \gamma_{2d}^{+} r_{mt}^2 I_{r_{mt}\geq 0} (1 - I_{ut}) + \gamma_{2d}^{-} r_{mt}^2 \left(1 - I_{r_{mt}\geq 0}\right)(1 - I_{ut}) + \varphi \mathrm{CSD}_{t-1} + \varepsilon_t, \end{aligned} \tag{9}$$

where I_{ut} equals one in the bull state and zero otherwise. In general, negative coefficients on r_{mt}^2 suggest herding. If herding intensifies when the market goes down in bear states, we expect a larger negative coefficient γ_{2d}^{-}. Equations (7) and (9) are used for CSD_t, CSD_t^P, and CSD_t^C for herding in the entire market, peripheral stocks, and core stocks, respectively.

3. Empirical Analysis

We investigate the herd behavior of Korean stocks using the network structure. Daily returns of 558 common stocks listed in the Korea Exchange are used for the sample period from January 2005 to December 2015. For robustness of our results, we use three different types of grouping methods: networks (clusters) estimated with all stocks, networks (clusters) estimated with stocks excluding securities firms, and 24 industries classified by the Korea Exchange using the Global Industry Classification Standard (GICS).[4] Our data source is Datastream. Equal weight is used to calculate the market and the index returns because CSDs are not value weighted.

3.1. Network structure of the Korean stock market

We first estimate a correlation matrix of 558 stock returns, and then, determine the network of the Korean stocks. The network is composed of 558 nodes and 557 links. In Figure 1 we visualize networks using a program called Pajek for three cases: a network under the assumption that stock returns are randomly correlated (panel A), a network with all stocks in the market (panel B), and a network with non-securities stocks (panel C). The network of random correlation generated by Pajek spans equally among stocks and has no pattern. On the contrary, both the networks with all stocks and without securities stocks are distinct from the random network in panel A because they visualize many core stocks. The network with all stocks shows a concentration of connections to a smaller number of core stocks.

We create clusters with K=6 in the first and the third criteria of the heuristic method for clustering so that at least six peripheral stocks are connected to a core stock. The number of core stocks identified by these criteria is 5−6% of all stocks. Table 1 shows clusters and their core stocks sorted by the numbers of links in the clusters. When all stocks are included in the network analysis, there are 28 clusters, 11 of which are securities firms whose performance depends on that of other stocks in the equity market.[5] The top five clusters include 233 stocks,

---

[4] We also test 17 industries that have at least five stocks. The results are not different from those reported with the 24 industries.

[5] This result is consistent with the literature on the Korean stock market network, for example, Lee and Woo (2013) who find that the top four out of 15 stocks that have a large influence in the Korean stock market are securities

and the 28 clusters include 530 stocks. When the securities firms are excluded, more clusters, 36, are found but the number of peripheral stocks in each of the clusters decreases so that 497 stocks are included in the 36 clusters.[6]

These results are summarized in Figure 2, where the connections between stocks using the MST are visualized. The first figure for the network with all stocks shows that Dongbu Securities and KDB Daewoo Securities are the cores of the two largest clusters, which include 74 and 64 stocks, respectively (Table 1). The second figure for the network with non-securities stocks shows that concentration to the largest few clusters is less severe.

Figure 3 depicts the number of peripheral stocks included in each of the clusters. The shape of this link distribution suggests a power law distribution and is consistent with the previous studies that stock markets belong to a scale-free network (Garlaschelli, Battiston, Castri, Servedio, and Caldarelli, 2005). For example, the results with non-securities stocks show that most stocks have weak relations with others because 336 out of 533 (63%) stocks have a single link and 90 (17%) stocks have two links, whereas the top three clusters have 139 stocks.

It is interesting that the core stocks identified with non-securities stocks do not include the largest stocks such as Samsung Electronics or Korea Electric Power Corporation. Our results indicate that these largest stocks are not connected with other stocks in the market despite their importance (weights) in market return. In fact, the network analysis shows us that medium stocks such as Keyang Electric Machinery, Hyundai BNG Steel, and Hanjin Heavy Industries and Construction are the top three core stocks that have 139 stocks in their clusters. Although we cannot conclude that these results show any lead-lag relationship between stock returns in the market, it is surprising to find that mid-size stocks are more linked to other stocks.

3.2. Estimation of Market States and Properties of Cross-sectional Dispersion

In this subsection, using the core and peripheral stocks identified in the previous subsection, we investigate the properties of cross-sectional dispersions in different market states.

Herding arises when financial markets are in stress and it becomes difficult for investors to process information rationally (Schwert, 1990; Christie and Huang, 1995; Chang, Cheng,

---

firms.

[6] When 25 securities stocks are excluded, the total number of stocks becomes 533.

and Khorana, 2000; Brunnermeier, 2001). To investigate herding when markets are in stress, we identify market states using the regime switching model in (8).[7] Figure 4 reports the smoothed probabilities of the two market regimes that we estimate using equally weighted market returns without securities stocks.[8] Bear states are identified during the financial crisis in 2008 and 2009, the late 2011, and intermittently in 2006, 2007, and 2015. The number of days in bear states (when the smoothed probabilities of bear states are larger than 0.5) is 512, and the average daily return and standard deviation of the market return are -0.26% and 5.11%, respectively. In general, bull periods are far more frequent: the number of days in bull states is 2,218. The average daily return and standard deviation of the market return during the bull state are 0.11% and 0.76%, respectively. Markets are in stress when market returns are negative and volatility is high (in bear states).

For comparison purposes, we also calculate the cross-sectional dispersion of industry returns with respect to market returns, and cross-sectional dispersion of individual stock returns with respect to their industry returns, which are also denoted as CSD_t^C and CSD_t^P, respectively. When the CSD is estimated using industry classifications as in Chang, Cheng, and Khorana (2000), Park (2011), Kim and Choe (2012), and Kim (2013), our measure of herding at the industry level, CSD_t^P, can be regarded as the aggregated CSD of all industries at the industry level:

$$\mathrm{CSD}_t^P = \sqrt{\sum_{ci=1}^{N_c} w_{ci} \frac{1}{N_{ci}} \sum_{i=1}^{N_{ci}} (r_{it} - r_{cit})^2} = \sqrt{\sum_{ji=1}^{N_j} w_{ji} CSV_t^{P_j}},$$

where $CSV_t^{P_j} = \frac{1}{N_{ji}} \sum_{i=1}^{N_{ji}} \left(r_{it} - r_{jit}\right)^2$ for industry j and N_{ji} is the number of stocks in industry j. As in Yao, Ma, and He (2014), Gebka and Wohar (2013), and Demirer, Lien, and Zhang (2015), if herding arises at the industry level, we observe herding in CSD_t^P.

Table 2 reports the basic statistical properties of CSD_t, CSD_t^C, CSD_t^P, and the stock market returns (r_{mt}), whose dynamics are shown in Figure 5. There is little difference in the properties of cross-sectional dispersions between when all stocks are used and when securities stocks are excluded. In panel A of Table 2, when securities stocks are excluded, daily averages

[7] The standard conjugate Gaussian distribution and the inverted gamma distribution are used for μ_i and σ_i, respectively. We estimate the transition probabilities using conjugate beta priors, but use weak priors for the transition probabilities in order to avoid frequent changes in regimes. The results are generated with 10,000 iterations after 10,000 burn-in iterations. For detailed explanations, see Kim and Nelson (1999) and Hwang and Satchell (2010).

[8] There is little difference in the smoothed probabilities between the two market returns (equally weighted market returns with all stocks and without securities stocks).

of the cross-sectional dispersions of core stocks (CSD_t^C) and peripheral stocks (CSD_t^P) are 2.44% and 3.67%, respectively. The average CSD_t^C and CSD_t^P are 2.32% and 3.51% in bull states, but increase to 2.98% and 4.37% in bear states, respectively. Thus, the cross-sectional dispersions of core stocks and peripheral stocks increase during bearish markets.

These results indicate that core stocks are less dispersed than peripheral stocks, and that the dispersion increases when the market is in stress. Panel C shows similar patterns in CSD_t^C and CSD_t^P for industry-sorted groups, but CSD_t^C is much smaller than CSD_t^P because equally weighted industry returns are used rather than returns of a core stock. However, the difference in the unconditional cross-sectional dispersions does not indicate herding during bull markets, which we test in the following subsection.

3.3. Herd Behavior Investigated with the Networks

We now investigate herding in cross-sectional stock returns using Equation (7). If herd behavior occurs, then the coefficients on r_{mt}^2 should be negative, as the CSD decreases when investors irrationally follow returns of core stocks at extreme market movements. The effects of herding on the coefficients of r_{mt}^2 would increase further when markets are in stress.

Table 3 reports the regression results of CSDs for the entire, bull and bear periods, using the clusters with all stocks, non-securities stocks, and industries. Bull and bear states are identified by the smoothed probability in Figure 4 ($\mathrm{prob}(S_{it}) \geq 0.5$). As expected, the coefficients on the absolute market return are all positive and significant. This result is consistent with a close association of market volatility and cross-sectional dispersion in returns (Hwang and Satchell, 2005). In the regression of the CSD, all coefficients, γ_2s, are negative and significant regardless of core or peripheral stocks in the entire period. This result confirms that herd behavior exists in the Korean stock market as in Chang, Cheng, and Khorana (2000), Park (2011), Kim and Choe (2012), and Kim (2013).

However, these results with the entire period are misleading because asymmetric responses of CSDs to r_{mt}^2 in different market states are disregarded. Panel C of Table 3 shows no statistical evidence of herding in bull or in bear states when the industry is used to group stocks: the coefficients on r_{mt}^2 are not negative at the 5% significance level. It is only when market states are disregarded that the results for the entire period show evidence of herding.

Moreover, adverse herding is observed during bull states for industry and a network with all stocks (panels A and C, respectively). This means that core stock returns or industry index returns are less likely to follow the market consensus during large market movements in bull periods. However, when securities stocks are excluded from the network, we do not find any statistical evidence of adverse herding (Panel B). Therefore, the results indicate that adverse herding arises when securities stocks perform as core stocks in the network.

Herding is observed only during bear states for the two cases where networks are used. When non-securities stocks are used to form core and peripheral stocks, we find evidence of herding in bear markets in core stocks and peripheral stocks. As the clusters estimated by the MST directly measure connections in price movements, the evidence of herding suggests co-movement in returns in bear markets. The network identified with all stocks may not represent connections between non-securities stocks, because it is dominated by securities stocks (Table 1).

Finally, evidence of adverse herding during bull states and herding during bear states is more clear when market returns are positive rather than negative. When herding intensifies due to investors' panic behavior, we expect severe herding when market returns are negative in bear states. Our results show that herding arises in bear states, but only when market returns are positive.

3.4. Robustness of Results

The robustness of our results are tested using Equation (9). The results in Table 4 are consistent with the findings in Table 3. Herding occurs in bear states between stocks that are closely correlated, whereas adverse herding is observed in bull markets. The difference in coefficients between bull and bear states is significant in all cases: the null hypothesis H_0: $\gamma_{2u}^{+} = \gamma_{2d}^{+}$ is rejected at the 5% significance level.

Our results are robust to different minimum numbers of peripheral stocks connected to a core stock. We set *K*= 5 and 7 instead of 6, and investigate herding for core and peripheral stocks, as described above. For example, when the minimum number of peripheral stocks connected to a core stock is set to 7, the numbers of clusters decrease to 18 and 20 from 28 and 36 for all stocks and non-securities stocks, respectively. The results of Equation (7) when *K*= 7 in Table 5 are consistent with those in Table 3. Herding arises in bear states and adverse herding is observed only when securities stocks are included. Otherwise, we do not find evidence of adverse herding.

4. Conclusions

In this study, we analyze networks in the Korean stock market using the minimum spanning tree algorithm, and then, investigate if herd behavior is led by a small sample of 'core' stocks or by 'peripheral' stocks during bear states in the stock market. Using cross-sectional dispersions of core stocks and of peripheral stocks, we show that herding arises for both core stocks and peripheral stocks during bear states.

We also find a few interesting asymmetric features of herding during bull and bear market states. First, during bull states, we find adverse herding, i.e., that the CSDs increase at extreme market movements. Adverse herding appears to be mainly driven by securities firms because it is significant only when networks with all stocks or industry are used for grouping. Second, both core stocks and peripheral stocks exhibit herding in bear market states. However, it is noteworthy that herding exhibited in bear states is significant when the stock market rises.

Our study suggests that co-movements in asset returns should be analyzed using networks identified with connections rather than the conventional grouping method such as industries. This is because stock returns in an industry are not necessarily closely connected with each other. The patterns of return co-movements show us a different story when the connections are identified with correlations and analyzed using network theory.

Appendix

Using core and peripheral stocks, we decompose the CSV into two parts, cross-sectional variance of core stocks and cross-sectional variance of peripheral stocks, as follows:

$$CSV_t = \frac{1}{N}\sum_{i=1}^{N}(r_{it} - r_{mt})^2$$
$$= \frac{1}{N}\sum_{i=1}^{N}(r_{it} - r_{cit})^2 + \frac{1}{N}\sum_{i=1}^{N}(r_{cit} - r_{mt})^2 + \frac{2}{N}\sum_{i=1}^{N}(r_{it} - r_{cit})(r_{cit} - r_{mt})$$
$$= \frac{1}{N}\sum_{ci=1}^{N_c}\sum_{i=1}^{N_{ci}}(r_{it} - r_{cit})^2 + \frac{1}{N}\sum_{ci=1}^{N_c} N_{ci}(r_{cit} - r_{mt})^2$$
$$= \sum_{ci=1}^{N_c} w_{ci}\frac{1}{N_{ci}}\sum_{i=1}^{N_{ci}}(r_{it} - r_{cit})^2 + \sum_{ci=1}^{N_c} w_{ci}(r_{cit} - r_{mt})^2,$$

assuming $\frac{2}{N-1}\sum_{i=1}^{N}(r_{it} - r_{cit})(r_{cit} - r_{mt}) = 0$, where r_{cit} denotes a core stock return, N_c and N_{ci} represent the numbers of core stocks and their peripheral stocks linked to core stock *c*, respectively, and $w_{ci} = \frac{N_{ci}}{N}$. The first component represents the weighted average cross-sectional variance of peripheral stocks linked to core stocks, whereas the second component represents the weighted cross-sectional variance of core stocks to the market.

References

Anton, M. and C. Polk, 2014, Connected Stocks, Journal of Finance 69(3):1099-1127.

Barber, B.M, T. Odean, and N. Zhu, 2009, Do Retail Trades Move Markets?, Review of Financial Studies 22:151-186.

Bikhchandani, S. and S. Sharma, 2001, Herd Behavior in Financial Markets, IMF Staff Paper 47(3):279-310.

Chang, C.H., and S.J. Lin, 2015, The effect of national culture and pitfalls on investors' decision-making: Herding behavior in international stock market, International Review of Economics and Finance 37:380-392.

Chang, E.C., J.W. Cheng, and A. Khorana, 2000, An examination of herd behavior in equity markets: an international perspective, Journal of Banking and Finance 24:1651-1679.

Chen, T., 2013, Do Investors Herd in Global Stock Markets?, Journal of Behavioral Finance 14(3): 230-239.

Choi, N. and R.W. Sias, 2009, Institutional industry herding, Journal of Financial Economics 94(3):469-491.

Christie, W.G. and R.D. Huang, 1995, Following the pied piper: Do individual returns herd around the market?, Financial Analysts Journal 51(4): 31-37.

Cormen, T.H., C.E. Leiserson, R.L. Rivest, and C. Stein, 2009, Introduction to Algorithms, MIT Press.

Coval, J. and E. Stafford, 2007, Asset fire sales (and purchases) in equity markets, Journal of Financial Economics 86:479-512.

Demirer, R., D. Lien, and H. Zhang, 2015, Industry herding and momentum strategies, Pacific-Basin Finance Journal 32: 95-110.

Garlaschelli, D., S. Battiston, M. Castri, V. Servedio, and G. Caldarelli, 2005, Statistical Mechanics and its Applications 350(2):491-499.

Gatev, E., W.N. Goetzmann, and K.G. Rouwenhorst, 2006, Pairs Trading: Performance of a Relative Value Arbitrage Rule, Yale ICF Working Paper No.08-03.

Gebka, B. and M.E. Wohar, 2013, International herding: Does it differ across sectors? Journal of International Markets, Institutions & Money 23: 55-84.

Hamilton, J.D., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Econometrica 57:357-384.

Harvey, C.R., Y. Liu, and H. Zhu, 2015, …and the Cross-Section of Expected Returns, Working Paper.

Henry, O. T., 2009. Regime switching in the relationship between equity returns and short-term interest rates in the UK. Journal of Banking & Finance 33(2): 405-414.

Hirshleifer, D. and S.H. Teoh, 2009, The Psychological Attraction Approach to Accounting and Disclosure Policy, Contemporary Accounting Research 26(4):1067-1090.

Hwang, S. and M. Salmon, 2004, Market stress and herding, Journal of Empirical Finance 11:585-616.

Hwang, S. and M. Salmon, 2009, Sentiment and beta herding, Working paper.

Hwang, S. and S.E. Satchell, 2005, Valuing Information Using Utility Functions, European Journal of Finance 11(1):1-16.

Hwang, S. and S.E. Satchell, 2010, How Loss Averse Are Investors in Financial Markets?, Journal of Banking and Finance 34:2425-2438.

Kim, C. and C.R. Nelson, 1999, State-space models with regime switching, MIT Press.

Kim, J.H. and H. Choe, 2012, The herding behavior of professionals and its impact in the Korean stock market, Asian Review of Financial Studies 41(3):463-495.

Kim, S.H., 2013, Empirical Tests of Herding Behavior in the Korean Stock Market, Hankook Gyungje Youngu 31(3):117-144.

Kruskal, J.B., 1956, On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem, Proceedings of the American Mathematical Society 7(1):48-50.

Lakonishok, J., A. Shleifer, and R.W. Vishny, 1992, The impact of institutional trading on stock prices, Journal of Financial Economics 32(1):23-43.

Lee, Y.J. and G. Woo, 2013, Analysis of the Stock Market Network for Portfolio Recommendation, Journal of Digital Contents Society 13(11):48-58.

Mantegna, R.N., 1999, Hierachical structure in financial markets, European Physical Journal B 11(1):193-197.

Park, B.J., 2011, Herd Behavior and Volatility in Financial Markets, Journal of the Korean data and information science society 22(6):1199-1215.

Shleifer, A. and R.W. Vishny, 1992, Liquidation Value and Debt Capacity: A Market Equilibrium Approach, Journal of Finance 47(4):1343-1366.

Welch, I., 2000, Herding among security analysts, Journal of Financial Economics 58:369-396.

Wermers, R., 1999, Mutual Fund Herding and the Impact on Stock Prices, Journal of Finance 54, 581-622.

Yao, J., C. Ma, and W. Peng He, 2014, Investor herding behavior of Chinese stock market, International Review of Economics and Finance 29: 12-29.

Table 1 Clusters and Core Stocks

This table shows the core stocks and their links in each clusters identified by the MST and the heuristic method that requires at least 6 peripheral stocks connected to a core stock. Panel A reports 28 clusters identified using all stocks and Panel B shows 36 clusters when securities firms are excluded.

	Core Stocks	Number of directly linked peripheral stocks	Number of total peripheral stocks
A. Network with All Stocks			
Cluster1	Dongbu Securities Co., Ltd.	50	74
Cluster2	Kdb Daewoo Securities Co.	33	64
Cluster3	Sk Securities Company Limited	24	36
Cluster4	Hyundai Securities Company Limited	20	26
Cluster5	Hanwha Investment&Securities Company	19	33
Cluster6	Hyundai Bng Steel Co Ltd	17	20
Cluster7	Samsung Securities Company Limited	13	24
Cluster8	Yuanta Securities Korea Co., Ltd	11	19
Cluster9	Keyang Electric Machinery Company	10	12
Cluster10	Hanjin Heavy Ind & Const Holdings	9	16
Cluster11	Doosan Infracore Company Limited	8	16
Cluster12	Korea Investment Holdings Company	8	13
Cluster13	Hmc Investment Securities Company	7	6
Cluster14	Daishin Securities Company Limited	7	6
Cluster15	Nh Investment & Securities Co Ltd	7	12
Cluster16	Seoyon Co Ltd	7	13
Cluster17	Hyundai Steel Co	6	15
Cluster18	Hyundai Motor Company Limited	6	7
Cluster19	Taekwang Industrial Company	6	13
Cluster20	Daelim Industrial Company Limited	6	8
Cluster21	Chongkundang Holdings Corp	6	10
Cluster22	Tongyangmoolsan Co Ltd	6	6
Cluster23	Dong Wha Pharm Company Limited	5	14
Cluster24	Yungjin Pharmaceutical Company	5	11
Cluster25	Gs Engineering & Construction Corp	5	12
Cluster26	Ni Steel Company Limited	4	17
Cluster27	Lotte Chemical Corp	5	20
Cluster28	Hanyang Securities Co., Ltd.	6	7

B. Network with Non-securities Stocks			
Cluster1	Keyang Elec.Mch.	32	50
Cluster2	Hyundai Bng Steel	29	44
Cluster3	Hanjin Hvind.& Con.Hdg.	20	45
Cluster4	Hansol Logistics	12	23
Cluster5	Daou Technology	10	17
Cluster6	Hankuk Carbon	10	9
Cluster7	Doosan Infracore	9	15
Cluster8	Doosan Engr.& Con.	8	11
Cluster9	Seoyeon	7	9
Cluster10	Dong Wha Pharm.	7	17
Cluster11	Daelim Industrial	7	11
Cluster12	Gs Engr. & Con.	7	14
Cluster13	Tong Yang Moolsan	7	6
Cluster14	Hanwha	7	8
Cluster15	Hwashin	6	8
Cluster16	Kb Financial Group	6	15
Cluster17	Chongkundang	6	7
Cluster18	Lg Life Sciences	6	6
Cluster19	Samsung C & T	6	14
Cluster20	Hyundai Heavy Industries	6	6
Cluster21	K C Tech	6	8
Cluster22	Sam Young Eltn.	6	13
Cluster23	Hanwha Chemical	6	13
Cluster24	Lotte Chemical	6	21
Cluster25	Taekwang Indl.	6	9
Cluster26	Hyundai Steel	6	12
Cluster27	Moonbae Steel	6	6
Cluster28	Posco	6	12
Cluster29	Ni Steel	4	11
Cluster30	Hyundai Marine & Fire In.	5	7
Cluster31	Willbes & Company	5	10
Cluster32	Lotte Chilsung	5	6
Cluster33	Bukwang Pharmaceutical Ind	5	9
Cluster34	Kwang Dong Pharm.	5	11
Cluster35	Hansol Technics	5	8
Cluster36	Mirae	5	6

Table 2 The Properties of Cross-sectional Dispersion of returns

The table report the properties of cross-sectional dispersion of returns for the two different clustering methods and industries. Each clusters identified by the MST and the heuristic method requires at least 6 peripheral stocks connected to a core stock. Bull and bear states are estimated with smoothed probabilities of the regime switching model in (8).

A. Network with All Stocks

Market States		Mean(%)	Median(%)	S.D.(%)	Skewness	Kurtosis	Observation
Entire	CSD_t	2.8304	2.7266	0.6183	1.9722	12.1348	2730
	CSD_t^C	2.2923	2.0292	1.0655	1.9079	8.5038	2730
	CSD_t^P	3.6135	3.3793	1.0248	1.9426	9.0125	2730
	r_{mt}	0.0375	0.162	1.2248	-1.5736	17.2866	2730
Bull States	CSD_t	2.7101	2.6352	0.5233	2.1256	16.7386	2230
	CSD_t^C	2.1402	1.9235	0.95	2.1563	10.9747	2230
	CSD_t^P	3.432	3.2478	0.8765	2.096	10.8448	2230
	r_{mt}	0.0983	0.1831	0.7867	-0.7357	5.0758	2230
Bear States	CSD_t	3.3667	3.2159	0.7173	1.7767	8.5571	500
	CSD_t^C	2.9705	2.7149	1.27	1.2819	4.8851	500
	CSD_t^P	4.4231	4.2181	1.2276	1.5316	6.2307	500
	r_{mt}	-0.234	0.0495	2.3128	-0.8366	6.6792	500

B. Network with Non-Securities Stocks

Market States	Variables	Mean(%)	Median(%)	S.D.(%)	Skewness	Kurtosis	Observation
Entire	CSD_t	2.848	2.746	0.6193	2.017912	12.65938	2730
	CSD_t^C	2.4406	2.2725	0.8438	1.863395	11.2467	2730
	CSD_t^P	3.6703	3.4967	0.8817	1.851314	10.50398	2730
	r_{mt}	0.0378	0.1702	1.2052	-1.6176	17.63215	2730
Bull States	CSD_t	2.7291	2.6601	0.5276	2.236779	18.13494	2218
	CSD_t^C	2.3154	2.1802	0.7463	2.09964	16.14319	2218
	CSD_t^P	3.5072	3.3815	0.7529	1.926021	13.22963	2218
	r_{mt}	0.1057	0.1906	0.7639	-0.74393	5.075484	2218
Bear States	CSD_t	3.3631	3.2159	0.7166	1.747149	8.406452	512
	CSD_t^C	2.9831	2.8432	1.0116	1.241222	5.466647	512
	CSD_t^P	4.377	4.2145	1.0376	1.610029	7.634917	512
	r_{mt}	-0.2564	0.008	2.2624	-0.83456	6.784008	512

C. Grouping by Industry

Market States	Variables	Mean(%)	Median(%)	S.D.(%)	Skewness	Kurtosis	Observation
Entire	CSD_t	2.8304	2.7266	0.6183	1.9722	12.1348	2730
	CSD_t^C	0.7273	0.6719	0.2672	2.3075	15.5731	2730
	CSD_t^P	2.7295	2.6296	0.5855	2.0388	13.2437	2730
	r_{mt}	0.0375	0.162	1.2248	-1.5736	17.2866	2730
Bull States	CSD_t	2.7101	2.6352	0.5233	2.1256	16.7386	2230
	CSD_t^C	0.6874	0.6451	0.2291	2.5895	24.7884	2230
	CSD_t^P	2.616	2.5526	0.4998	2.3171	19.6243	2230
	r_{mt}	0.0983	0.1831	0.7867	-0.7357	5.0758	2230
Bear States	CSD_t	3.3667	3.2159	0.7173	1.7767	8.5571	500
	CSD_t^C	0.9052	0.8176	0.3426	1.54	5.972	500
	CSD_t^P	3.2354	3.0934	0.667	1.7762	8.7321	500
	r_{mt}	-0.234	0.0495	2.3128	-0.8366	6.6792	500

Table 3 The Effects of Market Volatility on Cross-sectional Dispersion

The table report the regression results of the following equation:

$$\mathrm{CSD}_t = \gamma_0 + \gamma_1^+|r_{mt}|I_{r_{mt}\geq 0} + \gamma_1^-|r_{mt}|(1 - I_{r_{mt}\geq 0}) + \gamma_2^+ r_{mt}^2 I_{r_{mt}\geq 0} + \gamma_2^- r_{mt}^2(1 - I_{r_{mt}\geq 0}) + \varphi\mathrm{CSD}_{t-1} + \varepsilon_t,$$

where CSD_t is estimated using all stocks, core stocks, and peripheral stocks, and $I_{r_{mt}\geq 0}$ is an indicator variable that is one when $r_{mt} \geq 0$ and zero otherwise. For the results in panel C, core stocks are represented by industry returns, and peripheral stocks are stocks included in each of the industries. Each clusters identified by the MST and the heuristic method requires at least 6 peripheral stocks connected to a core stock. Bull and bear states are estimated with smoothed probabilities of the regime switching model in (8). The numbers in the round brackets represent heteroskedasticity robust t-statistics.

A. Network with All stocks

Entire Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.011	**0.179**	**0.282**	**-1.505**	-0.368	0.555	0.549
	(10.472)	(5.823)	(9.360)	(-2.338)	(-0.734)	(14.862)	
CSD_t^C	0.011	**0.561**	**0.430**	**-4.138**	-1.208	0.367	0.300
	(14.932)	(7.244)	(7.452)	(-2.076)	(-1.386)	(11.643)	
CSD_t^P	0.016	**0.500**	**0.483**	**-3.801**	-1.156	0.451	0.432
	(14.657)	(7.282)	(9.290)	(-2.183)	(-1.47)	(14.794)	

Bull Period

CSD_t	0.012	**0.039**	**0.167**	**6.007**	3.560	0.535	0.371
	(8.773)	(0.815)	(3.698)	(2.043)	(2.077)	(10.871)	
CSD_t^C	0.012	0.000	**0.236**	**31.909**	2.995	0.355	0.203
	(15.598)	(0.002)	(2.426)	(3.166)	(0.670)	(9.681)	
CSD_t^P	0.017	0.031	**0.282**	**26.159**	4.440	0.440	0.296
	(14.912)	(0.280)	(3.441)	(3.087)	(1.206)	(12.650)	

Bear Period

CSD_t	0.012	**0.171**	**0.272**	-1.572	-0.384	0.538	0.676
	(6.011)	(3.337)	(5.282)	(-1.746)	(-0.519)	(9.506)	
CSD_t^C	0.013	**0.546**	**0.394**	**-4.519**	-1.002	0.349	0.329
	(6.585)	(4.356)	(3.882)	(-2.072)	(-0.773)	(5.495)	
CSD_t^P	0.020	**0.470**	**0.441**	-3.772	-0.868	0.410	0.463
	(6.029)	(4.021)	(4.673)	(-1.745)	(-0.709)	(5.890)	

B. Network with Non-Securities Stocks

Entire Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.011 (10.389)	**0.168** (5.528)	**0.283** (9.385)	**-1.410** (-2.375)	-0.371 (-0.747)	0.548 (14.167)	0.534
CSD_t^C	0.013 (17.870)	**0.363** (6.073)	**0.317** (7.316)	**-2.671** (-2.076)	-0.280 (-0.436)	0.380 (13.708)	0.299
CSD_t^P	0.016 (15.244)	**0.321** (6.650)	**0.400** (9.658)	**-2.652** (-2.964)	-0.724 (-1.19)	0.504 (18.291)	0.469
Bull Period							
CSD_t	0.012 (8.757)	0.044 (0.846)	**0.180** (3.666)	5.376 (1.664)	3.120 (1.630)	0.525 (10.365)	0.354
CSD_t^C	0.013 (16.507)	0.133 (1.468)	**0.239** (3.257)	8.537 (1.407)	0.168 (0.063)	0.371 (10.867)	0.177
CSD_t^P	0.016 (13.618)	0.131 (1.649)	**0.284** (3.994)	7.924 (1.522)	2.628 (0.972)	0.490 (14.311)	0.311
Bear Period							
CSD_t	0.012 (6.216)	**0.157** (3.237)	**0.268** (5.312)	-1.487 (-1.817)	-0.343 (-0.464)	0.545 (9.946)	0.672
CSD_t^C	0.015 (7.974)	**0.407** (4.600)	**0.297** (4.164)	**-3.482** (-2.472)	-0.161 (-0.179)	0.339 (6.484)	0.356
CSD_t^P	0.018 (6.769)	**0.309** (3.999)	**0.364** (5.051)	**-2.703** (-2.187)	-0.465 (-0.491)	0.469 (8.355)	0.538

C. Grouping by Industry

Entire Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.011 (10.472)	**0.179** (5.823)	**0.282** (9.360)	**-1.505** (-2.338)	-0.368 (-0.734)	0.555 (14.862)	0.549
CSD_t^C	0.004 (18.671)	**0.144** (7.054)	**0.128** (8.446)	-0.733 (-1.247)	-0.040 (-0.158)	0.313 (12.430)	0.350
CSD_t^P	0.011 (9.672)	**0.149** (5.461)	**0.261** (9.335)	**-1.300** (-2.582)	-0.393 (-0.881)	0.557 (13.545)	0.536
Bull Period							
CSD_t	0.012 (8.773)	0.039 (0.815)	**0.167** (3.698)	**6.007** (2.043)	**3.560** (2.077)	0.535 (10.871)	0.371
CSD_t^C	0.004 (21.031)	0.023 (0.741)	**0.083** (3.417)	**5.810** (2.607)	0.828 (0.840)	0.335 (13.112)	0.195
CSD_t^P	0.012 (8.179)	0.038 (0.826)	**0.151** (3.499)	4.453 (1.573)	**3.494** (2.144)	0.530 (9.843)	0.358
Bear Period							
CSD_t	0.012 (6.011)	**0.171** (3.337)	**0.272** (5.282)	-1.572 (-1.746)	-0.384 (-0.519)	0.538 (9.506)	0.676
CSD_t^C	0.005 (7.556)	**0.176** (5.634)	**0.137** (5.494)	-1.099 (-1.853)	-0.121 (-0.349)	0.227 (3.729)	0.467
CSD_t^P	0.011 (5.944)	**0.129** (2.842)	**0.247** (5.211)	-1.282 (-1.716)	-0.394 (-0.592)	0.558 (9.860)	0.671

Table 4 The Effects of Market Volatility on Cross-sectional Dispersion in Bull and Bear States

The table report the regression results of the following equation:

$$\mathrm{CSD}_t = \gamma_0 + \gamma_{1u}^{+}|r_{mt}|I_{r_{mt}\geq 0}I_{\mathrm{ut}} + \gamma_{1u}^{-}|r_{mt}|(1 - I_{r_{mt}\geq 0})I_{\mathrm{ut}} + \gamma_{1d}^{+}|r_{mt}|I_{r_{mt}\geq 0}I_{\mathrm{dt}} + \gamma_{1d}^{-}|r_{mt}|(1 - I_{r_{mt}\geq 0})(1 - I_{\mathrm{ut}})$$
$$+\gamma_{2u}^{+}r_{mt}^{2}I_{r_{mt}\geq 0}I_{\mathrm{ut}} + \gamma_{2u}^{-}r_{mt}^{2}(1 - I_{r_{mt}\geq 0})I_{\mathrm{ut}} + \gamma_{2d}^{+}r_{mt}^{2}I_{r_{mt}\geq 0}I_{dt} + \gamma_{2d}^{-}r_{mt}^{2}(1 - I_{r_{mt}\geq 0})(1 - I_{\mathrm{ut}}) + \varphi\mathrm{CSD}_{t-1} + \varepsilon_t,$$

where CSD_t is estimated using all stocks, core stocks, and peripheral stocks, $I_{r_{mt}\geq 0}$ is an indicator variable that is one when $r_{mt} \geq 0$ and zero otherwise, and I_{dt} is an indicator variable that is one when the smoothed probability of the bull regime is larger than 0.5 and zero otherwise. Each clusters identified by the MST and the heuristic method requires at least 6 peripheral stocks connected to a core stock. The smoothed probability of bull and bear states is estimated using the regime switching model in (8). For the results in panel C, core stocks are represented by industry returns, and peripheral stocks are stocks included in each of the industries. The numbers in the round brackets represent heteroskedasticity robust t-statistics.

	A. Network with All Stocks			**B. Network with Non- Securities Stocks**			**C. Grouping by Industry**		
	CSD_t	CSD_t^C	CSD_t^P	CSD_t	CSD_t^C	CSD_t^P	CSD_t	CSD_t^C	CSD_t^P
γ_0	**0.012**	**0.012**	**0.018**	**0.012**	**0.014**	**0.017**	**0.012**	**0.004**	**0.011**
	(10.652)	(16.800)	(15.909)	(10.521)	(18.366)	(15.594)	(10.652)	(20.686)	(9.743)
γ_{1u}^{+}	0.009	-0.036	-0.020	0.009	0.088	0.065	0.009	0.032	0.000
	(0.178)	(-0.280)	(-0.182)	(0.162)	(0.972)	(0.793)	(0.178)	(1.026)	(-0.001)
γ_{1u}^{-}	**0.144**	**0.209**	**0.244**	**0.153**	**0.206**	**0.236**	**0.144**	**0.090**	**0.121**
	(3.200)	(2.147)	(2.935)	(3.186)	(2.802)	(3.294)	(3.200)	(3.636)	(2.848)
γ_{1d}^{+}	**0.201**	**0.581**	**0.518**	**0.191**	**0.445**	**0.365**	**0.201**	**0.163**	**0.169**
	(5.041)	(6.585)	(6.346)	(4.834)	(6.432)	(6.237)	(5.041)	(6.971)	(4.695)
γ_{1d}^{-}	**0.298**	**0.426**	**0.485**	**0.298**	**0.332**	**0.417**	**0.298**	**0.127**	**0.281**
	(7.894)	(5.849)	(7.527)	(8.066)	(6.356)	(8.162)	(7.894)	(6.835)	(8.105)
γ_{2u}^{+}	**7.269**	**33.534**	**28.539**	**6.819**	10.652	**10.925**	**7.269**	**5.528**	**5.986**
	(2.324)	(3.417)	(3.375)	(2.008)	(1.739)	(2.023)	(2.324)	(2.500)	(1.986)
γ_{2u}^{-}	**4.204**	3.779	5.551	**3.895**	1.151	4.073	**4.204**	0.652	**4.329**
	(2.440)	(0.858)	(1.511)	(2.046)	(0.427)	(1.484)	(2.440)	(0.662)	(2.668)
γ_{2d}^{+}	**-1.885**	**-4.928**	**-4.434**	**-1.805**	**-3.967**	**-3.394**	**-1.885**	**-1.074**	**-1.649**
	(-3.225)	(-2.871)	(-2.878)	(-3.366)	(-3.742)	(-4.275)	(-3.225)	(-2.155)	(-3.626)
γ_{2d}^{-}	-0.633	-1.313	-1.323	-0.622	-0.528	-1.009	-0.633	-0.051	-0.710
	(-1.148)	(-1.297)	(-1.496)	(-1.148)	(-0.737)	(-1.480)	(-1.148)	(-0.174)	(-1.480)
φ	**0.540**	**0.356**	**0.436**	**0.535**	**0.365**	**0.490**	**0.540**	**0.300**	**0.542**
	(14.094)	(11.032)	(14.106)	(13.498)	(12.965)	(17.512)	(14.094)	(11.964)	(12.797)
Adj. R^2	0.552	0.308	0.438	0.537	0.304	0.472	0.552	0.355	0.539
$\gamma_{2u}^{+} - \gamma_{2d}^{+}$	9.154	38.461	32.974	8.624	14.619	14.319	9.154	6.602	7.635
chi-square	9.397	15.760	15.848	6.975	5.897	7.383	9.397	9.313	6.881
$\gamma_{2u}^{-} - \gamma_{2d}^{-}$	4.837	5.092	6.874	4.517	1.679	5.082	4.837	0.702	5.039
chi-square	8.328	1.376	3.666	5.934	0.400	3.584	8.328	0.526	10.167

Table 5 The Effects of Market Volatility on Cross-sectional Dispersion When the Minimum Peripheral Stocks Connected to a Core Stock Is Seven

Table 3 is replicated with core and peripheral stocks identified by the MST and the heuristic method that requires at least 7 peripheral stocks connected to a core stock.

A. Network with All Stocks

Entire Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.011	**0.179**	**0.282**	**-1.505**	-0.368	0.555	0.549
	(10.472)	(5.823)	(9.360)	(-2.338)	(-0.734)	(14.862)	
CSD_t^C	0.011	**0.581**	**0.422**	-3.884	-0.932	0.316	0.240
	(15.748)	(6.663)	(6.650)	(-1.737)	(-1.036)	(10.320)	
CSD_t^P	0.017	**0.546**	**0.503**	-3.970	-1.038	0.413	0.388
	(15.590)	(6.862)	(8.618)	(-1.875)	(-1.170)	(13.802)	

Bull Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.012	0.039	**0.167**	**6.007**	**3.560**	0.535	0.371
	(8.773)	(0.815)	(3.698)	(2.043)	(2.077)	(10.871)	
CSD_t^C	0.012	-0.010	**0.225**	**34.735**	3.791	0.305	0.158
	(16.727)	(-0.060)	(2.047)	(2.733)	(0.729)	(8.695)	
CSD_t^P	0.018	0.014	**0.287**	**30.250**	5.322	0.406	0.263
	(16.725)	(0.105)	(3.141)	(2.900)	(1.292)	(12.406)	

Bear Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.012	0.171	**0.272**	-1.572	-0.384	0.538	0.676
	(6.011)	(3.337)	(5.282)	(-1.746)	(-0.519)	(9.506)	
CSD_t^C	0.013	**0.537**	**0.354**	-4.038	-0.401	0.302	0.268
	(6.673)	(3.775)	(3.171)	(-1.603)	(-0.292)	(4.704)	
CSD_t^P	0.021	**0.507**	**0.441**	-3.858	-0.529	0.366	0.421
	(6.139)	(3.758)	(4.178)	(-1.479)	(-0.383)	(5.062)	

B. Network with Non-Securities Stocks

Entire Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.011	**0.168**	**0.283**	**-1.410**	-0.371	0.548	0.534
	(10.389)	(5.528)	(9.385)	(-2.375)	(-0.747)	(14.167)	
CSD_t^C	0.014	**0.402**	**0.348**	**-2.671**	-0.610	0.297	0.224
	(20.358)	(6.079)	(7.291)	(-1.983)	(-0.868)	(10.996)	
CSD_t^P	0.018	**0.361**	**0.414**	**-2.619**	-0.557	0.448	0.416
	(17.274)	(6.601)	(9.141)	(-2.526)	(-0.775)	(16.784)	

Bull Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.012	0.044	**0.180**	5.376	3.120	0.525	0.354
	(8.757)	(0.846)	(3.666)	(1.664)	(1.630)	(10.365)	
CSD_t^C	0.015	0.160	**0.258**	7.966	-0.652	0.284	0.110
	(18.990)	(1.617)	(3.179)	(1.182)	(-0.219)	(8.504)	
CSD_t^P	0.019	0.125	**0.297**	9.743	2.167	0.431	0.249
	(15.927)	(1.443)	(3.930)	(1.670)	(0.759)	(13.191)	

Bear Period

	γ_0	γ_1^+	γ_1^-	γ_2^+	γ_2^-	φ	Adj R^2
CSD_t	0.012	**0.157**	**0.268**	-1.487	-0.343	0.545	0.672
	(6.216)	(3.237)	(5.312)	(-1.817)	(-0.464)	(9.946)	
CSD_t^C	0.016	**0.467**	**0.345**	**-3.813**	-0.703	0.270	0.307
	(8.970)	(4.846)	(4.523)	(-2.729)	(-0.736)	(5.409)	
CSD_t^P	0.021	**0.370**	**0.377**	**-3.031**	-0.283	0.413	0.511
	(7.516)	(4.465)	(4.877)	(-2.389)	(-0.265)	(7.449)	

Figure 1 Network Visualization

The network in the stock market is visualized with Pajek, a program for large network analysis. The first figure shows a network when individual stocks are not correlated. The second and third figures represent networks of all stocks and stocks excluding securities firms.

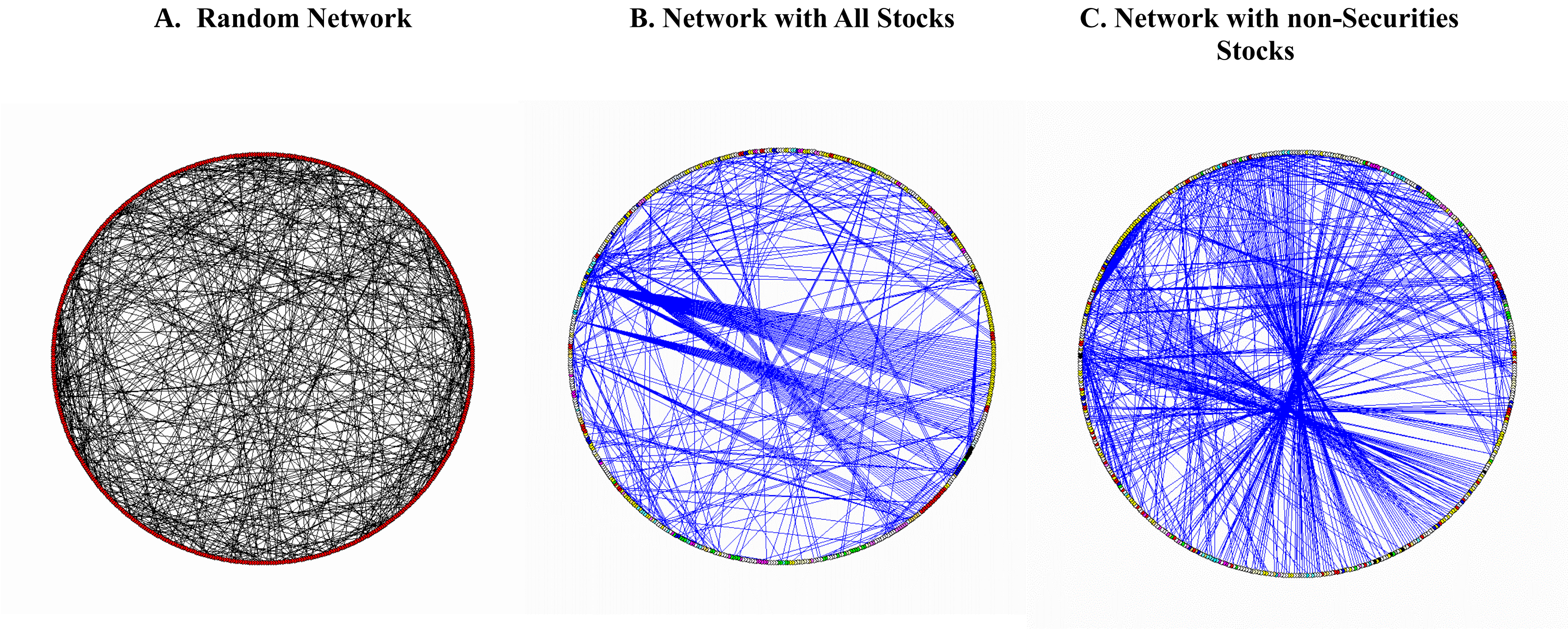
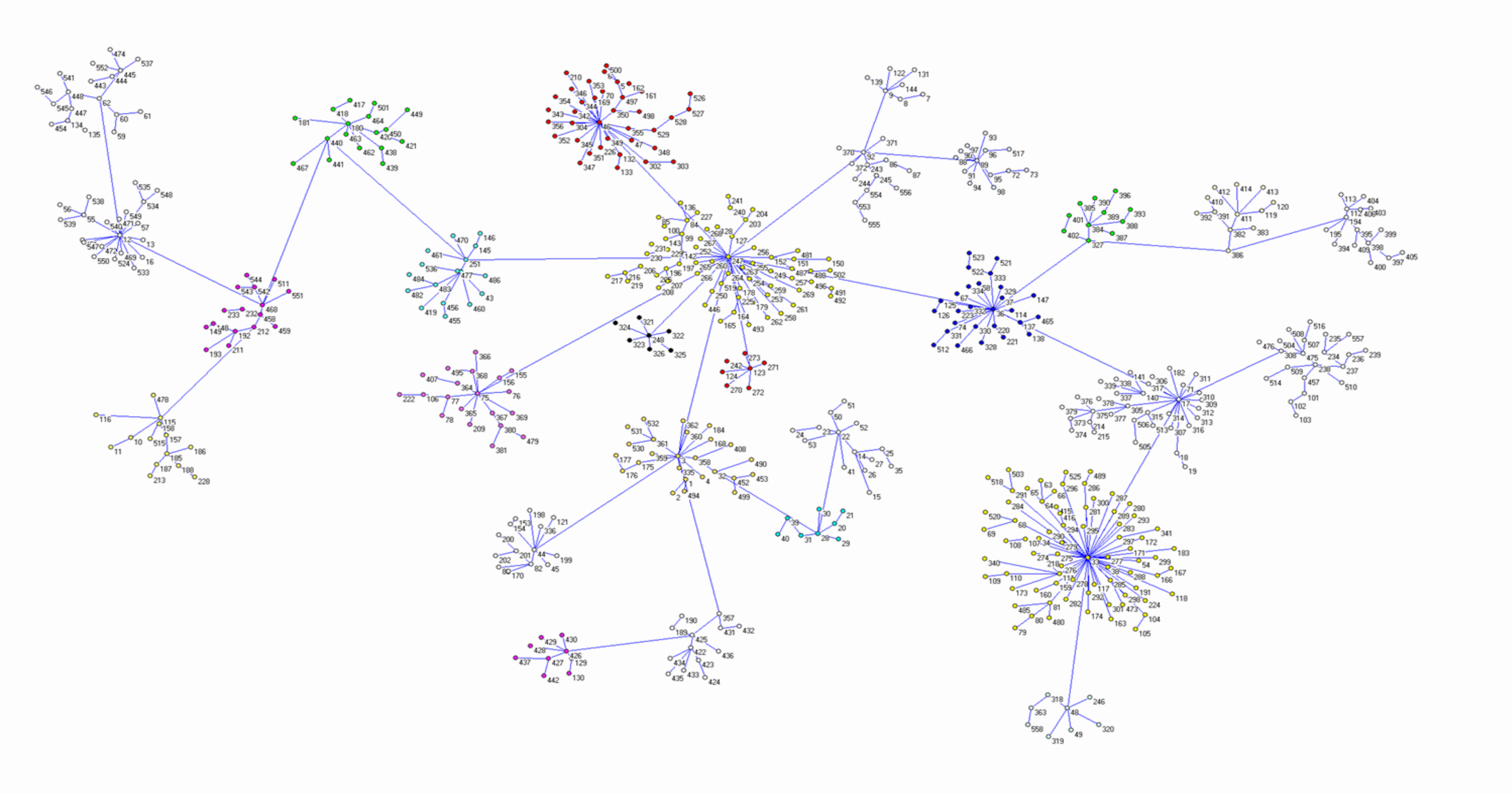


Figure 2 Network

The figures show network identified with the Minimal Spanning Tree and the heuristic method for clustering (A core stock has at least 6 directly linked peripheral stocks, a core stock that has at least one link to another core stock, and a bridge stock (that exists between two core stocks) that has at least 6 directly or indirectly linked to peripheral stocks)

A. Network with All Stocks



B. Network with non-Securities Stocks

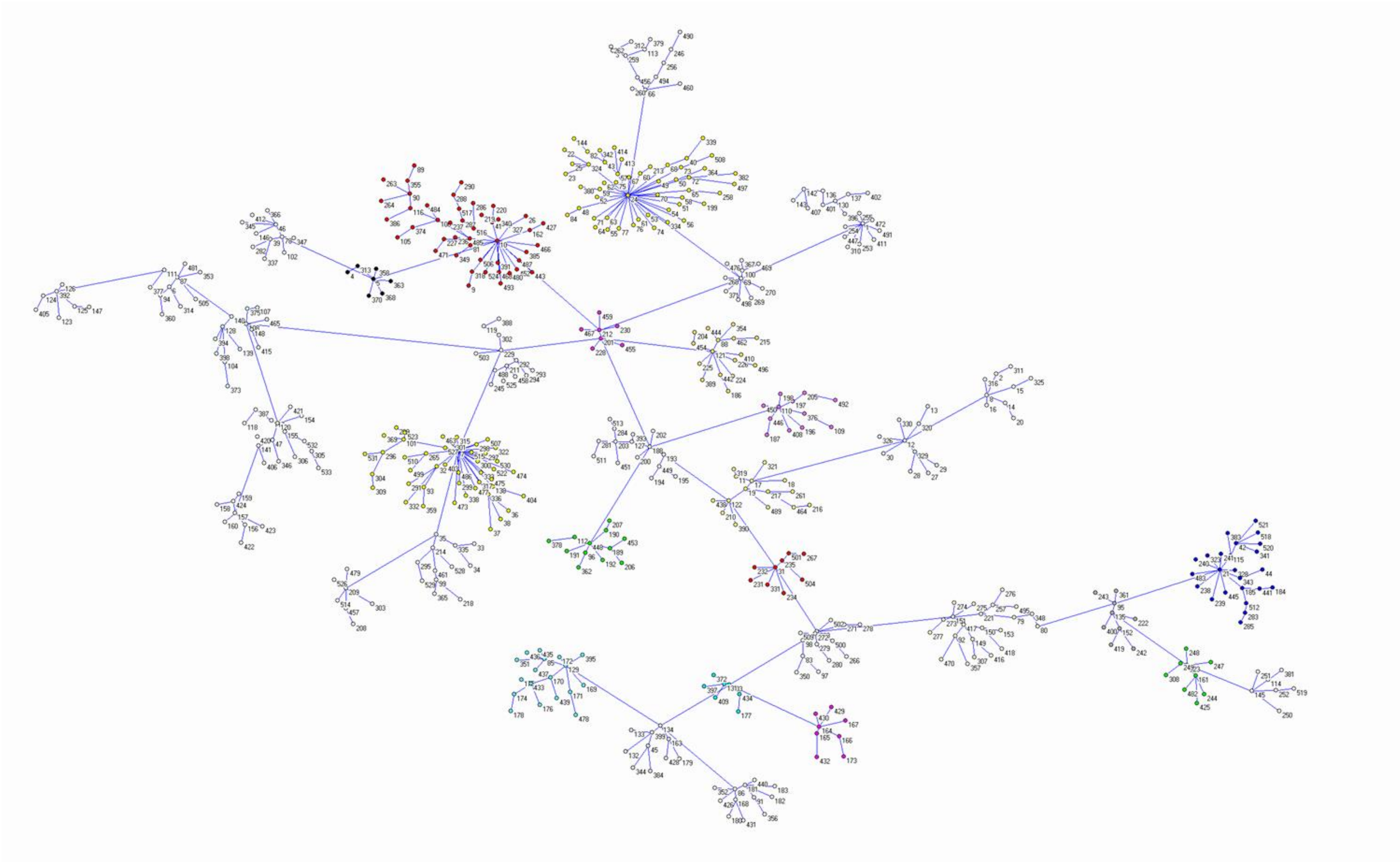
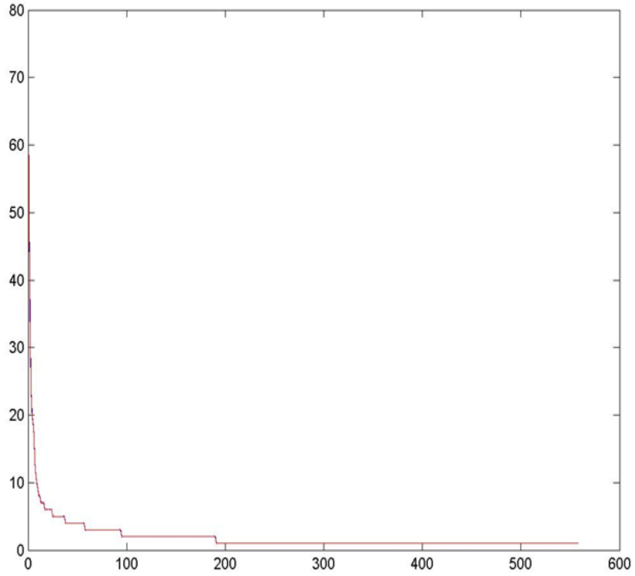


Figure 3 The distribution of links

The network distributions represent the number of peripheral stocks included in each of the clusters.

A. Network with All Stocks



B. Network with non-Securities Stocks

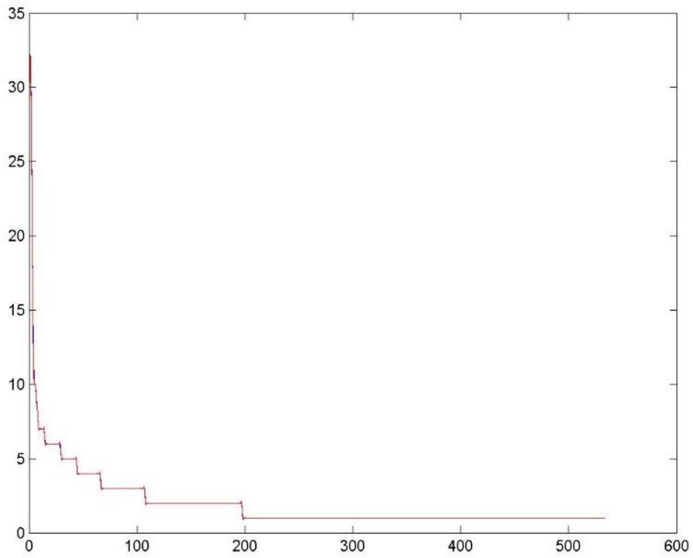


Figure 4 Probability of Regimes

We identify bull and bear markets using the following simple regime switching model:

$$r_{mt} = \mu_1 S_{1t} + \mu_2 S_{2t} + \sigma_t \varepsilon_t,$$

$$\sigma_t = \sigma_1 S_{1t} + \sigma_2 S_{2t},$$

where r_{mt} is the market return, μ_i and σ_i are the expected market return and volatility of regime $i = 1, 2$, respectively, and the dummy (state) variable, S_{it}, is one when regime i is selected, and zero otherwise. As in Hamilton (1989), the state variables are assumed to be governed by a first-order Markov chain. The regime switching model is estimated using the Bayesian Markov Chain Monte Carlo Gibbs sampling estimation. The standard conjugate Gaussian distribution and the inverted gamma distribution are used for μ_i and σ_i, respectively. We estimate the transition probabilities using conjugate beta priors, but use weak priors for the transition probabilities in order to avoid frequent changes in regimes. The results are generated with 10,000 iterations after 10,000 burn-in iterations. Once the two states are identified, they are labelled according to the characteristics of the expected market return and volatility.

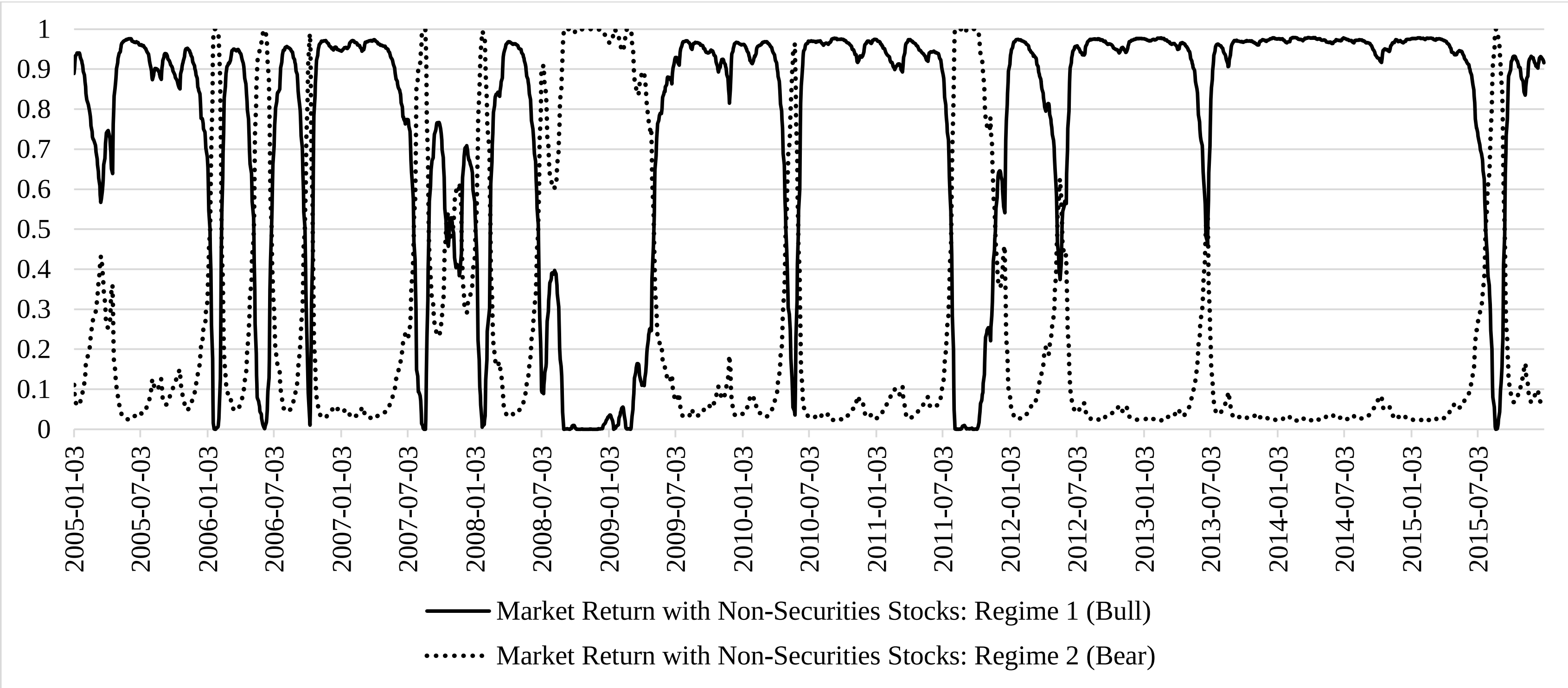
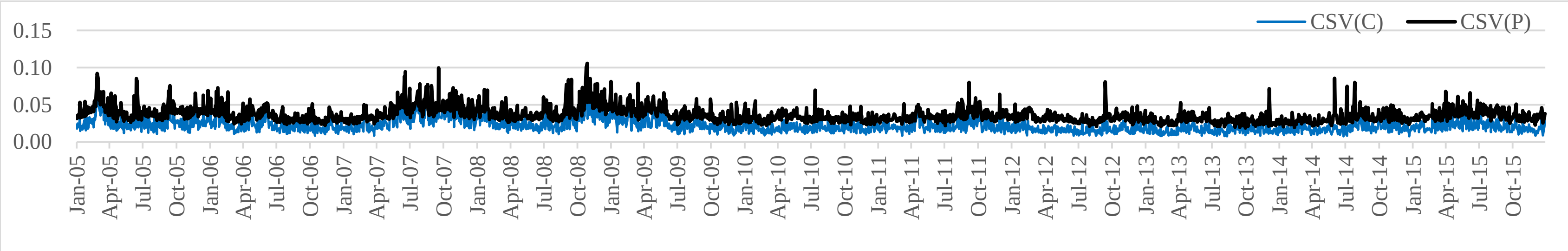


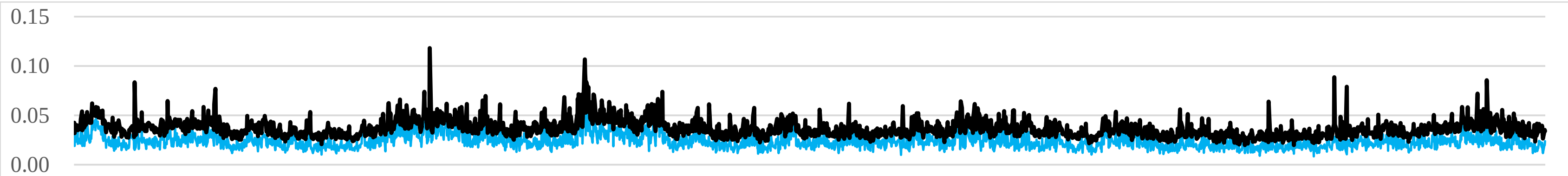
Figure 5 Dynamics of the Cross-sectional Dispersions

The figure shows the dynamics of cross-sectional dispersions as in equations (4)-(6). CSV(P) represents cross-sectional dispersion of peripheral stock returns with respect to their core stock returns and CSV(C) represents cross-sectional dispersion of core stock returns with respect to the market return.

A. Cross-sectional Dispersion with Network with All Stocks



B. Cross-sectional Dispersion with Non-securities Stocks



C. Cross-sectional Dispersion by Industry