entropy

MDPI

Article

Looking into the Market Behaviors through the Lens of Correlations and Eigenvalues: An Investigation on the Chinese and US Markets Using RMT

Yong Tang [1,2,*], Jason Xiong [3], Zhitao Cheng [4], Yan Zhuang [1], Kunqi Li [5], Jingcong Xie [6] and Yicheng Zhang [2]

1 School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China; tukouacademic@gmail.com
2 Department of Physics, University of Fribourg, Chemin du Musée 3, 1700 Fribourg, Switzerland; yi-cheng.zhang@unifr.ch
3 Walker College of Business, Appalachian State University, Boone, NC 28608, USA; xiongjj@appstate.edu
4 School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China; 2020110801009@std.uestc.edu.cn
5 Department of Electrical and Computer Engineering, State University of New York at Stony Brook, Stony Brook, NY 11794, USA; kunqi.li@stonybrook.edu
6 Terry College of Business, University of Georgia, Athens, GA 30602, USA; jx32164@uga.edu
* Correspondence: tangyong@uestc.edu.cn

Abstract: This research systematically analyzes the behaviors of correlations among stock prices and the eigenvalues for correlation matrices by utilizing random matrix theory (RMT) for Chinese and US stock markets. Results suggest that most eigenvalues of both markets fall within the predicted distribution intervals by RMT, whereas some larger eigenvalues fall beyond the noises and carry market information. The largest eigenvalue represents the market and is a good indicator for averaged correlations. Further, the average largest eigenvalue shows similar movement with the index for both markets. The analysis demonstrates the fraction of eigenvalues falling beyond the predicted interval, pinpointing major market switching points. It has identified that the average of eigenvector components corresponds to the largest eigenvalue switch with the market itself. The investigation on the second largest eigenvalue and its eigenvector suggests that the Chinese market is dominated by four industries whereas the US market contains three leading industries. The study later investigates how it changes before and after a market crash, revealing that the two markets behave differently, and a major market structure change is observed in the Chinese market but not in the US market. The results shed new light on mining hidden information from stock market data.

Keywords: financial big data; stock market modeling; random matrix theory; eigenvalue analysis

check for updates

Citation: Tang, Y.; Xiong, J.; Cheng, Z.; Zhuang, Y.; Li, K.; Xie, J.; Zhang, Y.-C. Looking into the Market Behaviors through the Lens of Correlations and Eigenvalues: An Investigation on the Chinese and US Markets Using RMT. *Entropy* **2023**, *25*, 1460. https://doi.org/10.3390/e25101460

Academic Editor: José Roberto Iglesias

Received: 10 July 2023
Revised: 8 October 2023
Accepted: 11 October 2023
Published: 18 October 2023

1. Introduction

Thanks to the availability of financial data in a wide range of frequencies from tick to daily, it is possible to apply data mining and knowledge discovery methods beyond traditional finance but from data science, network analysis, and even physics, etc. The asset prices in the markets result from complicated dynamics of spreading and reacting to market signals and information. The market structures are embedded in the price movements, which are normally correlated with each other. As a starting point for the underlying cornerstones of finance theories like modern portfolio theory (MPT) [1] and capital asset pricing model [2], the correlation information of assets prices is always at heart for theoretical studies and finance industrial practices in portfolio management and risk management, etc.

For a portfolio of N stocks, we need a correlation matrix with $N \times N$ elements to describe the pairwise relationships. With the increase of N, the number of possible relationships snowballs, making it difficult and challenging to calculate or analyze. To extract the hidden structure and essential information, it is necessary to simplify the network by filtering the less important elements to make it feasible to analyze portfolios even with a very large N. In the past few years, we see some methods have been introduced to simplify the stock matrices. To study the correlation behaviors of the financial markets, a correlation matrix is constructed from the price time series before we apply methods and techniques such as principal component analysis [3–6], multidimensional scaling [7], factor analysis [4], minimum spanning tree [8,9], hierarchical clustering [8,10,11], and singular value decomposition [12].

Simplification of the correlation matrix requires validation, which statistically validates the matrix and keeps those validated elements to achieve a simple matrix with fewer noises and ease of analysis. The validations provide statistical confidence in the results or insights extracted from the validated matrices. The underlying idea of design validation is to compare the empirical matrices with random ones generated from the same distributions, random shuffles, or statistical tests with which the null hypothesis is set up to be tested with empirical data. Any deviations from these *benchmarks* are considered noises and should be filtered. Similarly, given an empirical correlations matrix (and the derived distance matrices for the networks), we can consider a random matrix with the same size. A null hypothesis can be introduced to test the statistical validation of each element of the original empirical matrix by comparing the distributions. The basic idea is that any deviations from the random distribution are believed as validated with genuine information from the system. In contrast, those falling within the random distribution are pure random noises that contain no system information.

Specifically, in this study, based on a dataset covering nine years of stock prices, we systematically investigate the stock markets of China and the US using random matrix theory (RMT) to study and compare the correlation properties and the dynamics of eigenvalues and eigenvectors. The findings revealed that the two stock markets are both similar and different in many ways. The results add new insights into market behaviors with implications for finance applications such as portfolio management and optimization, market risk monitoring, and trading strategy design. Meanwhile, this study also serves as a framework for data mining and knowledge in financial big data using RMT.

This work is organized as follows. First, we review the literature in Section 2. The methodology is introduced in Section 3. In Section 4 we describe the dataset of two markets and the properties of correlation matrices. Using RMT, in Section 5, the properties and behaviors of eigenvalues and eigenvectors are analyzed with an investigation of a market switch study. Finally, Section 6 presents conclusions, discussions, and limitations.

2. Literature Review

In this section, we introduce literature from three aspects. First, RMT and its applications are introduced in Section 2.1, representative studies of applying RMT in analyzing financial markets are described in Section 2.2, as well as recent studies focusing on comparing different stock markets, especially the US and Chinese markets are discussed in Section 2.3.

2.1. RMT and Its Applications

Originating in mathematical physics, RMT was first introduced by physicists to study nuclear activities back in the 1950s [13]. Eugene Wigner used RMT to model the excited states of nuclei in reactions which was hard to obtain by using traditional methods. Instead, he proposed to analyze the eigenvalues and their spacing of a random matrix [14]. The basic idea is to analyze the statistical properties of eigenvalues of the random counterpart whereas it is practically impossible to analyze the individual eigenvalues of the original complex system. RMT provides a powerful toolbox to reveal properties of matrices whose elements are sampled from randomness, usually based on certain probability distributions. Soon, RMT was proved an efficient tool for many challenges in physics and beyond. Before long, RMT attracted significant interest from scholars in various fields with wide-spreading applications like physics, mathematics, biology, engineering, computer science, and social science. After decades of development, RMT has become an important research field with rich theoretical implications and real applications in a variety of disciplines, such as spectrum analysis and filter in information processing, signal detection and channel estimation in wireless communication, data analysis in high dimensional space, and optimization in machine learning. Interested readers should refer to works by Potters and Tao for details on theories and applications of RMT [15,16].

2.2. Applying RMT Approaches in Financial Markets

In finance, RMT was first introduced into the study of financial markets by [17], and more recently, there are significant advances in applying RMT in finance studies and applications [18–20]. In one study, RMT is applied to analyze stock market behaviors [21]. In another study, the world stock market is analyzed with RMT [22]. Recent works also investigated various financial markets using RMT [23–25].

2.2.1. RMT in Financial Correlation Analysis

Rooted in the correlation analysis, RMT offers a new look into the structures and behaviors of the financial markets. Applying RMT to financial markets is closely related to the analysis of correlation matrices and network structures [26–30]. The market is full of noises, and the useful information in correlation matrices built from price data might be covered by the noises and make correlation analysis less meaningful [31]. To quantify the validations of correlations, recently, there are many works applying RMT into the studies of the correlation matrices of financial markets [17,31–37]. Recently, there is an emergence of research using RMT in financial markets to filter noises and reveal embedded market properties. The cross-correlations of stock prices are studied using RMT to identify correlated relationships [38]. Furthermore, free random variables are applied in RMT analysis in financial time series [39]. RMT has also been applied to return estimation and asset allocation in Markowitz mean-variance optimization [40].

2.2.2. RMT in Eigenvalue Analysis

RMT provides a powerful tool for eigenvalue analysis in financial markets. Using time-shifted series, the lagged correlation matrices are studied from the RMT approach to compute eigenvalue density and identify deviations [41]. It has been verified that the largest eigenvalue λ_{max} is a good estimator of the average correlation of the correlation matrices constructed from a sliding window approach [42]. The same results are also reported, revealing that the average correlation co-moves with the largest eigenvalue for the component stocks of S&P500 [36]. For normalized eigenvectors, the value of S_{ij} ranges from 0 to 1. In other words, the two eigenvectors change from orthogonal to exactly the same. One study reports that the effect of noises on the risk becomes insignificant in measure of the fixed portfolio while remaining important for an optimized portfolio for small values of N/L [31]. This indicates that the correlation matrix can still be valid in traditional risk management and portfolio optimization; noises cover even most information. Using simulation methods, many correlation matrix filtering approaches are tested, and the approaches based on random matrix theory are found to perform consistently well in all cases [43]. The eigenvalue distribution of the emerging stock market is different from developed markets though correlation distributions and other properties are similar. Methods based on clustering for portfolio optimization and effective size determination are proposed. The results are found to be improved compared to RMT approaches [37], which indicates that RMT might be further combined with other methods in filtering matrix and optimizing a portfolio [27].

It was found that the average of correlations in the correlation matrix can be well estimated from the largest eigenvalue as

$$\lambda_{max}/N \sim \langle C_{ij} \rangle. \quad (1)$$

Following the RMT approach, the largest eigenvalues are found to be responsible for the market mode. By removing this, the correlation matrix is cleaned to reveal the topological structures [28]. The details of the residual noise part for a market are studied, revealing that the noise band is composed of more sub-bands [11]. Using RMT, the Chinese stock market is studied [18], a similar anti-correlation relationship between sub-sectors is studied [44], and the results show that the prominent sector structure exists. The distribution of eigenvalues also reveals that the market is likely to be influenced by the Chinese government's global financial crisis and policies. In a further study on the sub-sectors of a stock market, local interaction structures are found to change during financial crises [19]. The sign information of components in eigenvectors is again used to detect the sub-sector anti-correlations [45]. Focusing on how the credit market and stock market behave before and after a financial crisis, RMT is applied and finds that the largest eigenvalue of the credit market precedes that of the stock market [46]; this indicates that the pattern changes of eigenvalues have potential implications in the understanding of interrelationships between different markets. Market contagion is also investigated from financial network analysis and, naturally, RMT. Market contagion is an important indicator of market stability. By looking into the structural changes in networks and properties revealed by RMT, one can identify and predict the market contagion and thus major market switches [47–51].

2.2.3. RMT in Eigenvalue Distributions

Since the introduction of RMT into the study of financial markets, much literature investigated different markets. An earlier study points out that the lower bound is positive and no eigenvalues fall between 0 and λ_{min} also vanish above λ_{max} [17]. Since the empirical values of N and L are limited far from ∞, the edges are blurred with some eigenvalues falling beyond the bounds [32]. The distribution of the spacings of eigenvalues $s \equiv \lambda_{i+1} - \lambda_i$ are found to agree with a Wigner distribution of the energy spacing levels [34]. This provides evidence indicating that the empirical correlation matrix is consistent with its random matrix counterpart. Many empirical studies reveal that only a small fraction of eigenvalues and their corresponding eigenvectors contain system information while most are embedded in noises [17,52–54]. It has been reported that the portion of the largest eigenvalues deviating from the theoretical prediction of the counterpart random matrix is 6% [17], 4.7% [54], 2% [34], 11% [53], and 1% [35].

Furthermore, the study of [55] adds new evidence that not all eigenvalues that fall into the theoretical interval predicted by the random matrix are purely random noise but still carry some information. Derived from the eigenvector-eigenvalue identity, a study showed that dominant eigenvalues, super eigenvalues, and maximum eigenvalues could help to analyze the spectrum of the financial correlation matrix in depth [56]. In computational results and applications in financial markets, one study reviewed the previous works, including some real-world applications, and presented promising analytical techniques from random matrix theory [26,57]. Another study proposed general, exact formulas for the overlaps between the eigenvectors of large correlated random matrices with noises [58]. Besides the intro-relationship of stock markets, another study revealed a deep relationship between news and world financial indices using tools of random matrix theory [59]. Economic policy is another field that has a significant influence on the stock market, and [60] analyzed the correlation matrix and different stock network structures to reveal the implication of the correlation matrix components. The work of [61] fused previous models, which made predictions based on the arbitrarily long time horizon and introduced an ensemble of random rectangular matrices from the observations of independent Lévy processes over a fixed-time horizon. To summarize and compose a benchmark for the study of correlated time-series signals, ref. [62] used supersymmetric theory to generate the statistics of eigenvectors of the cross-correlations of correlated time-series. Another study investigated the correlations of Chinese stocks before and during the 2008 crisis based on the random matrix analysis [63].

2.3. Comparative Studies on Different Markets

In another thread, there are abundant studies dedicated to the comparative studies on the two major stock markets, namely the Chinese and US markets [29,64–67]. Although RMT has been applied to stock markets, there is still a lack of comprehensive studies using RMT to analyze the Chinese and US markets. How the correlations and eigenvalues behavior are related to the market switches between bull and bear markets is still not sufficiently investigated. There is a thread of literature on comparing the dynamics of markets in different countries [29,30]. Considering the signs of eigenvector components, sub-sectors of positive and negative signs can be derived from sectors in anti-correlation. The sub-sectors are detected with strong appearances in the Chinese stock market but weaker in the US stock market [44]. US and British stock exchanges are studied by using RMT on the asymmetric correlation matrix with a lag of τ [3]. One work revealed the different strengths of correlations between stocks, especially the oil sector and banking stocks in the Nigerian Stock Market (NSM) and Johannesburg Stock Exchange (JSE), for the period of 2009 to 2013, using random matrix theory [68]. Comparative analyses on two different stock markets—the S&P 500 (USA) and Nikkei 225 (JPN) via the power mapping method from the random matrix theory, and found strong consistency between the states of the two stock markets as well as the feasibility to predict critical state (market crash) [69].

Particularly, some works investigated the markets of the US and China [64,66]. According to the strong connection between financial assets and institutions and the diversity as well as the localization of the stock market, one study previously analyzed the topological structure of financial networks of two major markets of China and the US with complex network theory [29]. Several studies investigated the two markets from aspects of comovement [70], impacts of trade conflicts and pandemic [65,67,71], and conditional correlations [72]. These studies revealed the different behaviors of the two major stock markets. However, there is still a lack of comprehensive studies on the Chinese and US stock markets from the perspective of RMT. In this sense, this work aims to fill this gap by systematically investigating the two markets using correlation analysis and RMT.

3. Methodology

3.1. Construction of Correlation Matrices

For an empirical correlation matrix C of size $N \times N$ generated from N returns series of length L, we can construct the elements as

$$C = \frac{1}{L} M M^T, \tag{2}$$

where M is a $N \times L$ matrix with normalized return $y_i(t)$ for each stock at every time t, where

$$y_i(t) = \frac{Y_i(t) - \langle Y_i(t) \rangle}{\sigma_i}, \tag{3}$$

where $Y_i(t)$ stands for the return at time t.

The study of [73] provides a study of the eigenvalues spectrum for the Chinese stock market with a sliding window approach. The inverse participation ratio is defined as

$$I^k = \sum_{l=1}^{N} \left[u_l^k \right]^4, \tag{4}$$

where u_l^k is the components of eigenvector v^k, to measure the deviation degree of eigenvectors [53]. A criterion of fractional Gaussian noise (fGn) is used to evaluate the autocorrelation matrix of stocks showing agreement with fGn, though the stock returns are non-Gaussian [20].

3.2. Eigenvalue Analysis Using RMT

RMT is a powerful tool in the analysis of eigenvalues of noisy data in various fields [15,16,74,75]. According to RMT [15,16], the eigenvalue distribution of a pure random matrix C_{random} with the same size of C follows

$$p(\lambda)_{random} = \frac{Q}{2\pi}\frac{\sqrt{(\lambda_{max}-\lambda)(\lambda-\lambda_{min})}}{\lambda}, \tag{5}$$

where λ_{min} and λ_{max} are the theoretical minimum and maximum eigenvalue bounds of random matrix, the Q is the ratio of L/N satisfying the requirement that $Q > 1$, $L \to \infty$, and $N \to \infty$ [17]. Using the empirical data, we can also get the empirical distribution as

$$p(\lambda)_{random} = \frac{1}{N}\frac{dn(\lambda)}{d\lambda}. \tag{6}$$

Theoretically, with the knowledge of Q, we can determine the theoretical eigenvalue bounds as

$$\lambda_{min,max} = 1 + \frac{1}{Q} \mp 2\sqrt{\frac{1}{Q}}. \tag{7}$$

With these calculations, we can construct and determine the theoretical distribution of a null hypothesis random matrix. The empirical eigenvalues that fall within the interval of $[\lambda_{min}, \lambda_{max}]$ are pure random noises, and those that fall beyond the interval are the validated eigenvalues carrying true information of the system. In this way, we also get the validated corresponding eigenvectors for those validated eigenvalues. Also, we can go further to investigate the statistical validation of the eigenvectors. The distribution of the eigenvector components in v_i for eigenvalue λ_i follows the Porter-Thomas distribution [17] as

$$P(v_i) = \frac{1}{\sqrt{2\pi}}e^{-\frac{v_i^2}{2}}, \tag{8}$$

with which we can validate the eigenvector components by comparing the distributions. It has been reported that the distribution of eigenvector components of the largest eigenvalues shows a great difference from the theoretical predictions [17].

In short, we first construct the correlation matrix for the N stocks and calculate the corresponding theoretical bounds of eigenvalues predicted with RMT, and analyze the eigenvalues with special attention to the largest and second largest eigenvalues.

4. Data and Correlation Matrices

4.1. Data

In this paper, we study the stock markets of China and the United States. There are three considerations in choosing these two markets. First, both markets are major stock markets in the world with tremendous total market scales and a large number of stocks that are actively traded. Second, the two markets both experienced major market shifts between bull and bear markets demonstrating rich market dynamics and behaviors. Third, the US market and Chinese market are representatives of a much-matured market and still-developing market, respectively. We collected the daily price data of the components of the China Securities Index 300 (CSI300) and Standard & Poor's 500 (S&P500) between 4 January 2007 and 6 November 2015. In total, the dataset covers 2149 trading days for CSI300, and 2228 trading days for S&P500. The data of CSI300 are retrieved from the CSMAR Solution Database of Shenzhen GTA Education Tech. Ltd. The data of S&P500 are extracted from Yahoo Finance service. We further selected 163 stocks from CSI300 with at least 2000 trading dates without continuous 100 non-trading dates, whereas we selected 468 stocks with at least 2100 trading dates from S&P500. Later, we refer to the screened stocks as CSI163 and S&P468, respectively [29].

4.2. Correlation Matrices

In a stock market, the prices of stocks fluctuate constantly showing complex behaviors. It is important to investigate the performance of individual stocks as well as the interactive behaviors among stocks. To evaluate the interactive co-movement behaviors among the prices of assets, the correlation is a fundamental concept widely used in studies of price dynamics and is used in traditional theories. When the correlation is considered, in traditional theories, like in MPT where the correlation matrices are actually inputs for the portfolio optimization [1], the correlation is assumed as fixed. Still, in the real world, the correlations fluctuate and demonstrate some collective behaviors in market crashes. As a starting point for studying the structure and behavior of markets, correlation analysis is found to be useful not only in theory but also in practices of portfolio risk estimation and optimization [33,76]. Especially during periods of crisis, highly collected co-movements of the stocks are very likely to cause significant losses for a portfolio, so it is necessary to watch the portfolio's correlations. Also, to understand the market structure and the dynamics, it is interesting to investigate the correlations [8,33,35,77–79].

Following the definition and notation widely used in the literature, the Pearson correlation coefficient [8]

$$\rho_{ij} = \frac{\langle Y_i Y_j \rangle - \langle Y_i \rangle \langle Y_j \rangle}{\sqrt{\left(\langle Y_i^2 \rangle - \langle Y_i \rangle^2 \right) \left(\langle Y_j^2 \rangle - \langle Y_j \rangle^2 \right)}} \tag{9}$$

can be calculated for each stock pair of s_i and s_j using the logarithmic return

$$Y_i = \ln P_i(t) - \ln P_i(t-1). \tag{10}$$

The value of ρ_{ij} ranges from -1 to -1, indicating a dynamic relationship for the two stocks from a complete anti-correlation to a complete correlation. For a perfect uncorrelated pair, $\rho_{ij} = 0$ by definition. If there are N stocks in consideration, then there will be N^2 correlation coefficients fitting into a $N \times N$ correlation matrix. Correlation analysis has been applied in the study of market structures [8,28,80] and portfolio optimization [31,37,43,54].

In the RMT approach, the statistics of the eigenvalues distribution and the deviation between empirical distribution and the distribution generated from a random fashion are discussed to describe the information contribution of these deviated eigenvalues and the corresponding components of the eigenvectors. But first, the empirical results are tested against a random matrix case [31].

5. Eigenvalues and Eigenvectors for CSI163 and S&P468

5.1. Eigenvalues

Based on the correlation matrices we built in the previous section, we are ready to investigate the eigenvalues and eigenvectors of both markets. First, we use all the logged daily returns data of both two markets, CSI163 and S&P468 over the whole study period, which is 4 January 2007 and 6 November 2015 covering 2149 trading dates for the former and 2228 trading dates for the latter. We present the probability density distributions (PDF) of eigenvalues from the empirical correlation matrix and theoretically predicted by using random matrix theory for CSI163 in Figure 1 and for S&P468 in Figure 2, respectively. For both markets, we find that most empirical eigenvalues are within the RMT predicted interval with some exceptions. As shown in Figure 1, for CSI163, the theoretical predicted eigenvalues bounds are $\lambda_{min} = 0.5250$ and $\lambda_{max} = 1.6267$. We see that there are 7 eigenvalues are larger than the largest eigenvalue predicted by RMT, i.e., 4.29% of all eigenvalues fall beyond the interval. The largest eigenvalue $\lambda_1 = 60.2252$ is nearly 37 times the predicted largest eigenvalue, i.e., $\lambda_1/\lambda_{max} = 37.0238$. For S&P468, as shown in Figure 2, the largest eigenvalue $\lambda_1 = 189.5698$ which is almost 89 times the bound predicted by RMT.

There are 12 eigenvalues that are larger than the bound, i.e., 3.56% are beyond the interval and carry real market information.

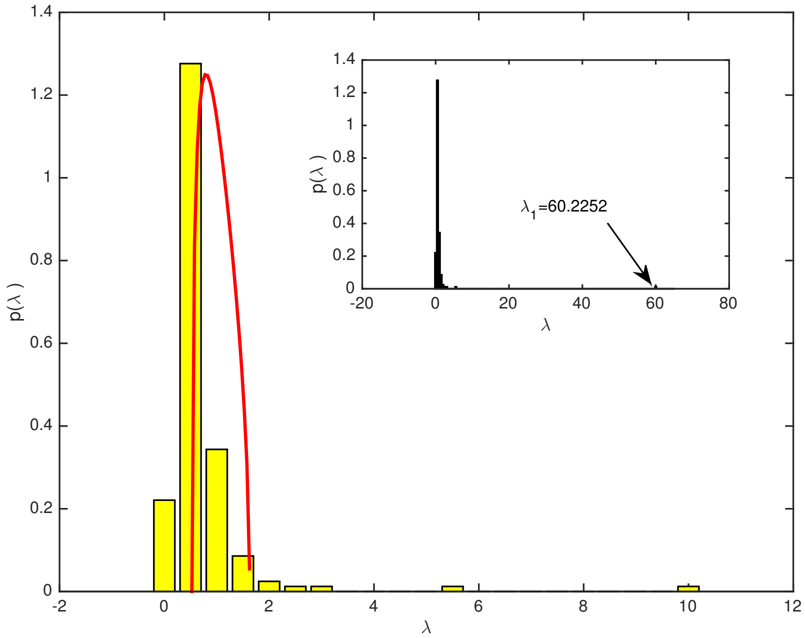


Figure 1. The eigenvalue distributions for CSI163 correlation matrix over the whole study period. The yellow bars are distributions of all eigenvalues calculated from the empirical correlation matrix of 163 daily log return time series, and the red curve is the theoretical distribution predicted from the random matrix theory by using a random matrix of the same size as the empirical correlation matrix. The upper bound is $\lambda_{max} = 1.6267$. The inset is a plot of all empirical eigenvalues including the largest eigenvalue $\lambda_1 = 60.2252$.

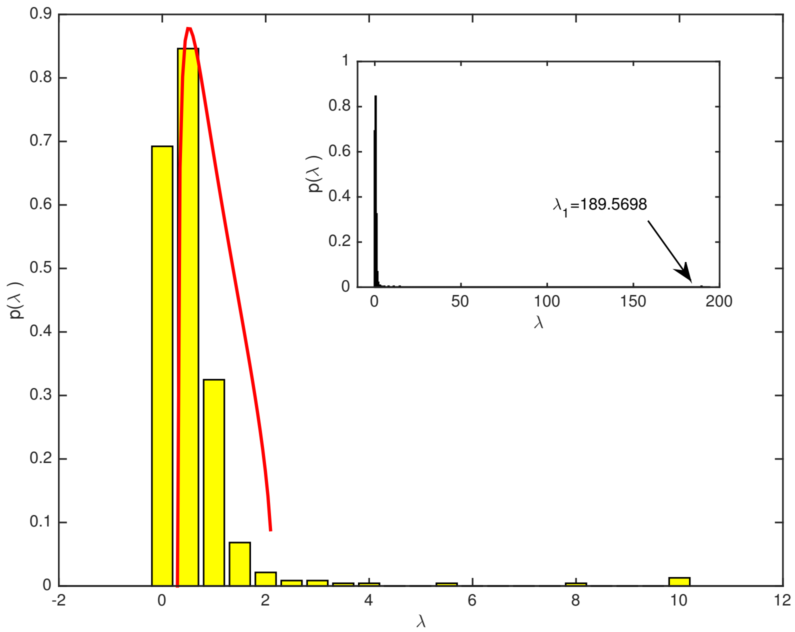


Figure 2. The eigenvalue distributions for S&P468 correlation matrix over the whole study period. The yellow bars are distributions of all eigenvalues calculated from the empirical correlation matrix of 468 daily logged return series, and the red curve is the theoretical distribution predicted from the random matrix theory by using a random matrix of the same size as the empirical correlation matrix. The upper bound is $\lambda_{max} = 2.130$. The inset is a plot of all empirical eigenvalues, including the largest eigenvalue $\lambda_1 = 189.5698$.

Using the sliding window approach, we can investigate the dynamic properties of eigenvalue distributions. For CSI163 and S&P468, we use the window size $L_{csi163} = 170$ and $L_{S\&P468} = 500$, respectively, to satisfy the requirement of $Q = L/N > 1$. In choosing the window sizes, basically, we desire a window that is large enough to cover significant market periods. A shorter window might lead to short-term noises that do not reflect the fundamental dynamics of the markets. Furthermore, the window moves at a step of one trading date; this allows our sliding windows to move smoothly with the finest possible granularity and capture detailed market behaviors. For each sliding window, we use the data of N stocks to calculate the pairwise correlation matrix C, from which we further calculate the λ_1/N and average correlation $\langle C_{ij} \rangle = \sum C_{ij}/N^2$. As shown in Figure 3a,b, we see that for both markets, the values of λ_1/N and the average correlation $\langle C_{ij} \rangle$ correlated very well over the whole study period indicating that λ_1/N is a good estimator of the average correlation $\langle C_{ij} \rangle$ as we have introduced previously.

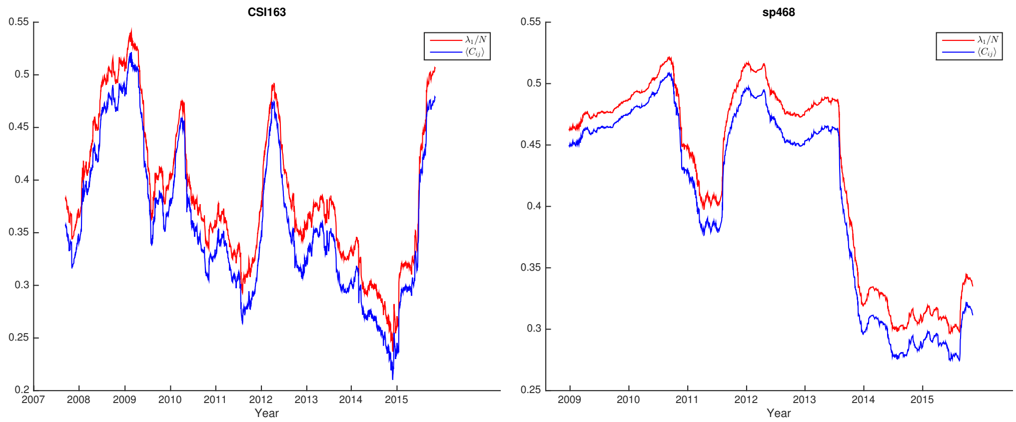


Figure 3. The largest eigenvalue λ_1/N and the average correlation $\langle C_{ij} \rangle$ for all sliding windows of CSI163 (**a**) and S&P468 (**b**). We see that the two curves fit very well.

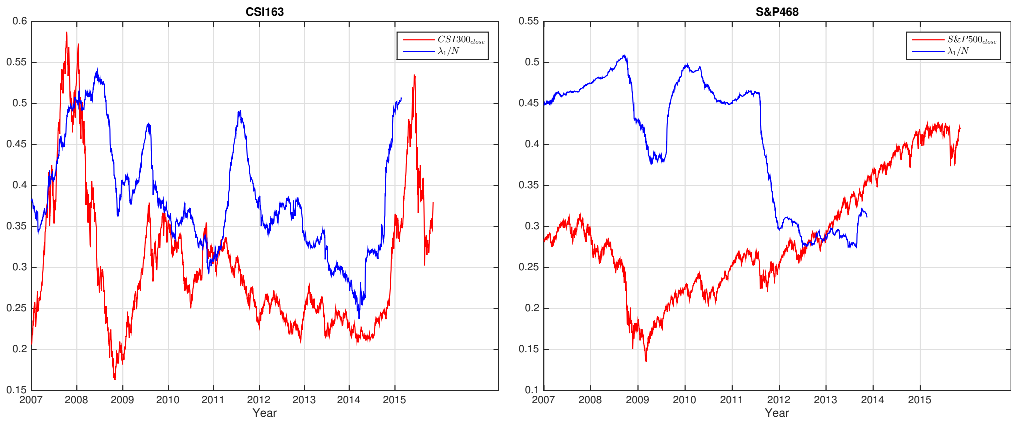
In Figure 4a,b, we plot the largest eigenvalue λ_1/N and the index close prices of CSI300 (a) and S&P500 (b). After the left shifting, we find that λ_1/N and the index itself show similar trends. This shows that λ_1/N is also an indicator of the index itself. For CSI163, the trend similarity is relatively more obvious than that of S&P468. If we do not perform left shifts, we find that λ_1/N is anti-co-move with the index showing that during market crashes, the λ_1/N (also the average correlation $\langle C_{ij} \rangle$) becomes larger, i.e., the stocks of the market are correlated, whereas during calm periods, the λ_1/N becomes small indicating fewer correlations among stocks.

To see how the eigenvalues distributed in the whole study period. In Figures 5 and 6, we plot the distributions of the eigenvalues (excluding the largest eigenvalue) of all sliding windows over the study periods CSI163 and S&P468, respectively. As the figures show, most eigenvalues are very small. Though many eigenvalues are within the bounds of prediction based on RMT, we also observe some eigenvalues are larger than the upper bound $\lambda_{max} = 3.9172$ for CSI163 and $\lambda_{max} = 3.8709$ for S&P468. We define the fraction of eigenvalues that are larger than the predicted λ_{max} using RMT as

$$p^d = \frac{|\lambda > \lambda_{max}|}{N}, \tag{11}$$

i.e., the ratio of the number of eigenvalues deviated beyond λ_{max} to the total number of eigenvalues N. Since the eigenvalues carry meaningful information about the market, this ratio can be employed as an indicator describing how much information is embedded in the distribution of the empirical eigenvalues. Using the sliding window approach, we calculate the fraction for each window and plot with the index close price for CSI163 in Figure 7 and S&P468 in Figure 8, respectively.

For better visualizations, we shrink the index values of 200,000 times for $CSI163_{close}$ and 100,000 times for $S\&P468_{close}$, respectively. As we can see, the values of p^d stay unchanged between sudden changes, so the curves of p^d show a shape of discrete stages with ups and downs. More interestingly, we find that the changes of p^d coincide with the changes in index closing prices. As shown in Figure 7 for CSI163 and Figure 8 for S&P468, the changing points of the p^d precisely mark out the local minimums (marked with yellow dots) and local maximums (marked with red dots) of the index itself.



(**a**) (**b**)

Figure 4. The largest eigenvalue λ_1/N and the index close price of CSI300 (**a**) and S&P500 (**b**). The largest eigenvalue λ_1/N curves are left shifted 170 trading dates for CSI300 and 500 trading dates for S&P500, for the window size is 170 for CSI163 and 500 for S&P468. For better visualizations, we shrink the indices of CSI300 and S&P500 10,000 times and 5000 times, respectively. We see that the shifted curves of λ_1/N are similar to the indices.

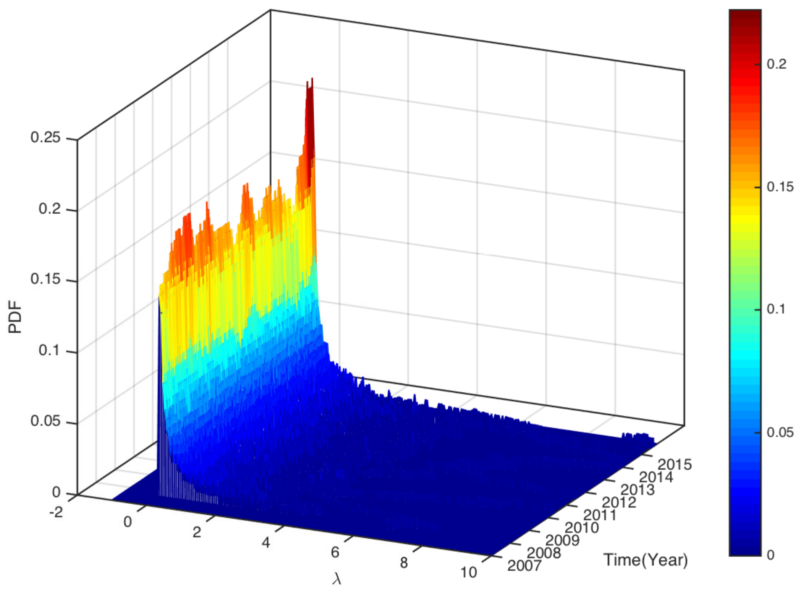


Figure 5. The PDF of all eigenvalues (excluding the largest eigenvalue λ_1) distribution for all sliding windows over the study period of CSI163.

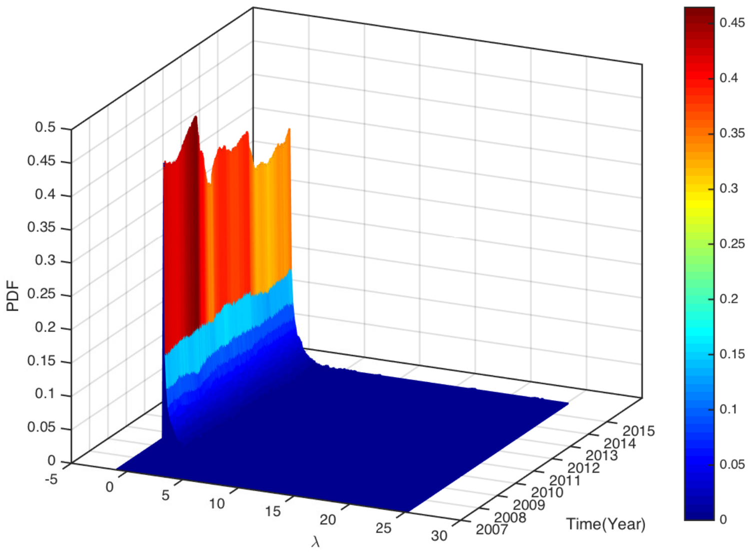


Figure 6. The PDF of all eigenvalues (excluding the largest eigenvalue λ_1) distribution for all sliding windows over the study period of S&P468.

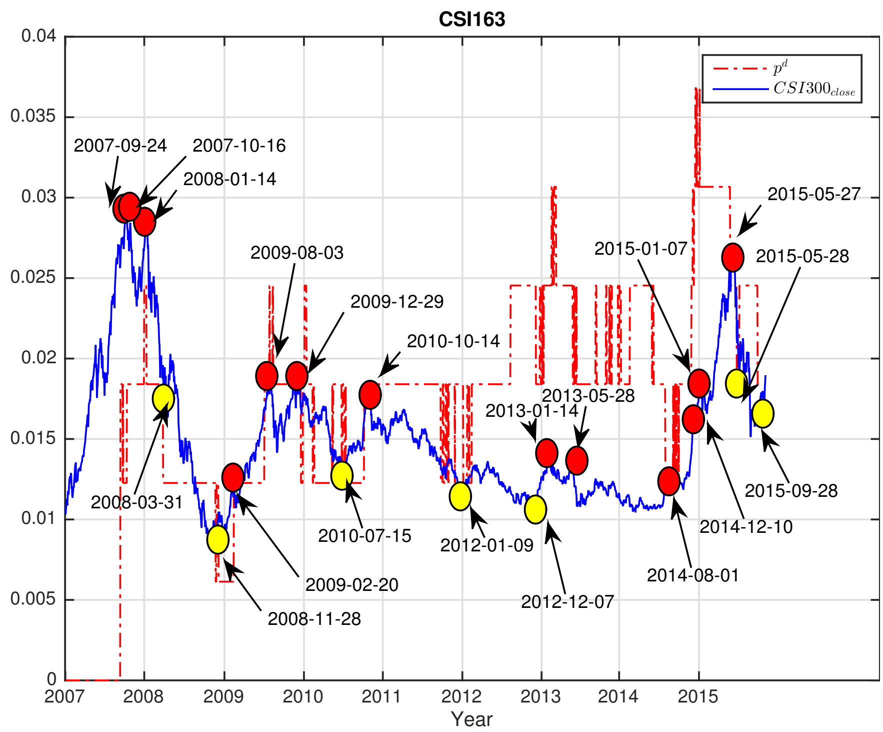


Figure 7. The fraction p^d of eigenvalues beyond the predicted largest eigenvalue versus the index close price for CSI163 over the study period. For better visualization, we rescale the index values by shrinking 200,000 times. The coincidences of changes of fraction p^d and the index closing price are marked out in red dots for local maximums and yellow dots for local minimums on the price curve with dates.

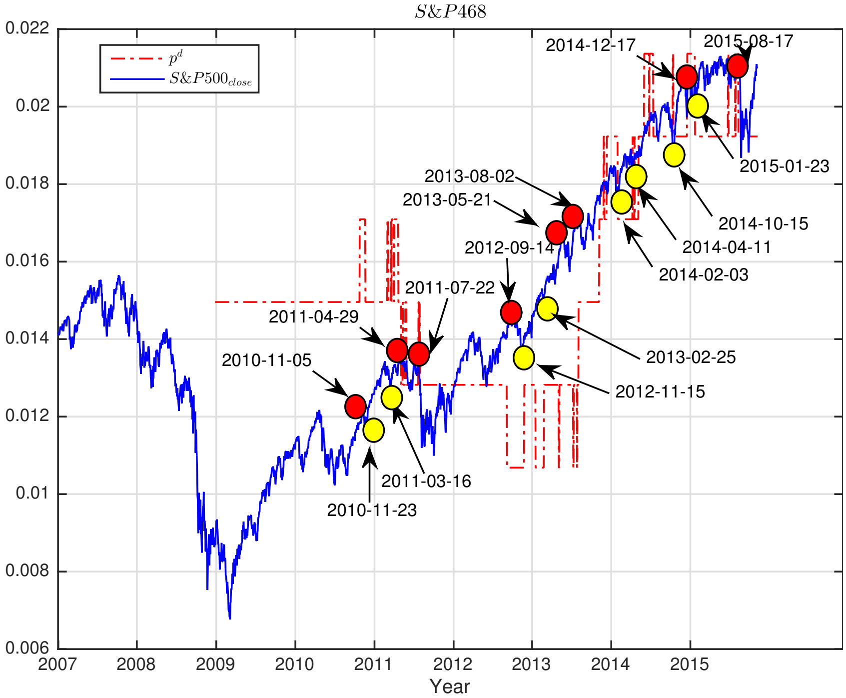


Figure 8. The fraction p^d of eigenvalues beyond the predicted largest eigenvalue versus the index close price for S&P468 over the study period. For better visualization, we rescale the index values by shrinking 100,000 times. The coincidences of changes of fraction p^d and the index closing price are marked out in red dots for local maximums and yellow dots for local minimums on the price curve with dates.

We see that p^d is relatively stable with many fixed periods, but the changes of p^d can match with the significant market changes in index closing prices. Some of them are even leading the index for several days. This observation indicates that p^d has the potential to monitor the market situation. Once the p^d changes value, investors must be cautious and pay particular attention to the market fluctuations both of surges and crashes. This information might also be useful in designing trading strategies to catch major market mode switches.

In Table 1, we summarize the properties of eigenvalues that deviate beyond the λ_{max}. We see that only a very small fraction of eigenvalues is larger than the theoretically predicted eigenvalue. On average, only 3.0268 eigenvalues for CSI163 and 7.2250 eigenvalues for S&P468 are beyond the bounds. The average fraction is $\langle p^d \rangle = 0.0186$ for CSI163 and $\langle p^d \rangle = 0.0154$ for S&P468, respectively.

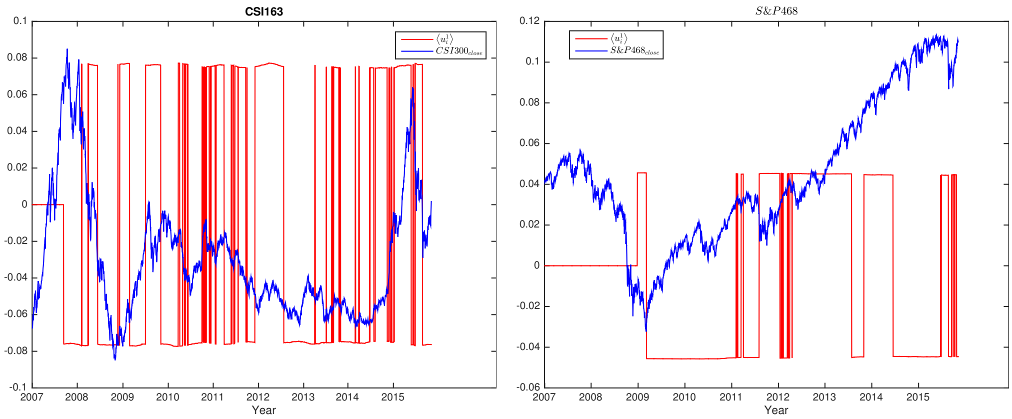
Table 1. Properties of eigenvalue deviation fraction p^d for CSI163 and S&P468. The avg. number is the average number of eigenvalues deviated beyond the predicted upper bounds λ_{max}.

Market	Avg. Number	p^d_{min}	p^d_{max}	$\langle p^d \rangle$
CSI163	3.0268	0.0061	0.0368	0.0186
S&P468	7.2250	0.0107	0.0214	0.0154

5.2. Largest Eigenvalue

To study the eigenvector u^1 corresponding to the largest eigenvalue λ_1, we take an average of all eigenvector components. Since the λ_1 stands for the whole market, we expect that the average components are related to the index. We plot the $\langle u_i^1 \rangle$ with the index close prices of both markets for each sliding window in Figure 9a,b. As shown in the figures, the value of $\langle u_i^1 \rangle$ changes happened on the dates or periods of major market changes.

For the eigenvector u^1, we also confirm that all components have the same sign, either positive or negative [81], i.e., all stocks contribute to the movement of the market on the eigenvector u^1 in the same direction; they either climb or fall. It is worth noting that, in practice, one might choose to remove the market mode of the largest eigenvalue before analyzing the eigenvalues. Here, we directly analyze the second-largest eigenvalue and the corresponding eigenvector for simplicity.



(**a**) (**b**)

Figure 9. The average of eigenvector components corresponding to the largest eigenvalue $\langle u_i^1 \rangle$ and the index close price of both CSI300 (**a**) and S&P500 (**b**). For better visualizations, we shrink the index close price by 25,000 times and 10,000 times for CSI300 and S&P500, respectively. We see that the changes of $\langle u_i^1 \rangle$ happen on the dates when the markets change.

5.3. Second Largest Eigenvalue

It is believed that the largest eigenvalue λ_1 stands for the market mode itself, whereas the second largest λ_2 eigenvalue and its corresponding eigenvector u^2 contain more information about the market. Now, we focus on the distribution of the components in u^2. As we know, the values of components in eigenvectors represent the weights for the corresponding eigenvector; the best idea to allocate investment in portfolio management is that we long the assets with positive signs and short the assets with negative signs. The eigenportfolio based on eigenvector u^j is given as:

$$P_j = \sum_{i=1}^{N} \frac{1}{\sqrt{\lambda_j}} \frac{u_i^j}{\sigma_i} Y_i, \tag{12}$$

where N is the number of assets, U_i^j is the component for asset s_i in eigenvector u_j, and Y_i is the return for asset s_i. This indicates that larger eigenvalues λ_i bring fewer weights for assets in a risky portfolio, whereas smaller eigenvalues bring smaller risks with greater weights on the assets.

For industry I_i, the contribution of I_i is defined as

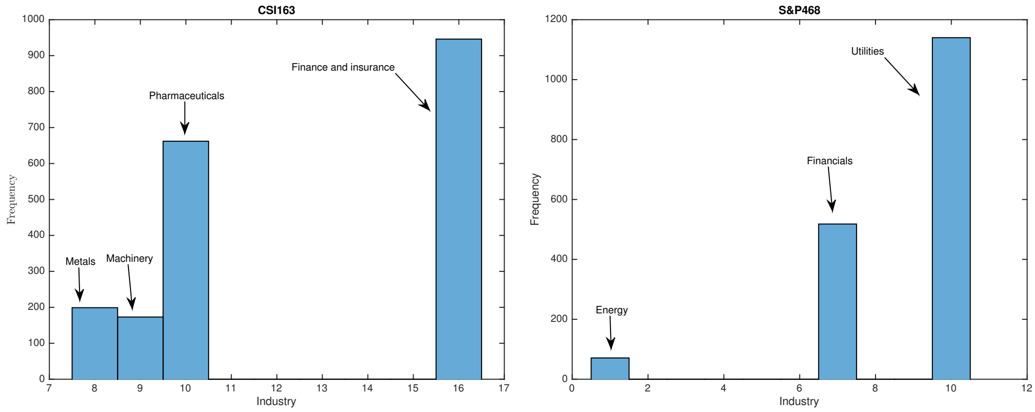
$$I_i^j(t) = \sum_{k \in I_i} \left(u_k^j \right)^2, \tag{13}$$

where u_k^j is the value of the stock belonging to industry I_i. By dividing over the total values of all industries, we get the normalized contribution for industry I_i

$$\bar{I}_i^j(t) = \frac{I_i^j(t)}{\sum_i I_i^j(t)}. \tag{14}$$

Compared with another approach [73], the normalized values allow comparison between any two industries, thus making the ranking of industries possible. Of course, Equation (14) also indicates that $\sum_i \overline{I}_i^j(t) = 1$.

Using Equations (13) and (14), we calculate and rank all industries in all sliding windows for both CSI163 and S&P468. For a given date, we can get the contributions from all sectors to the eigenvector components for the second largest eigenvalue u^2. We investigate which industries appear in the components with the largest values. In Figure 10a,b, we plot the histograms for industries that appeared for CSI163 and S&P468, respectively. We find that four industries appeared for CSI163, which are finance and insurance, pharmaceuticals, machinery, and metals, whereas for S&P468, we find only three industries appeared, which are utilities, financials, and energy. This reveals the leading industrial sectors for the two markets over the whole study period.



(**a**) (**b**)

Figure 10. The frequencies of industries appearing in the largest values of eigenvector components of CSI163 (**a**) and S&P468 (**b**). For CSI163, four industries appear in the largest eigenvector components, whereas there are three industries for that of S&P468.

5.4. Market Switching

Both the Chinese and US markets experienced significant fluctuations during our study period covering some major market mode changes of bull markets and bear markets. In our study period between 4 January 2007 and 6 November 2015, the Chinese stock market enjoyed a bull market period from 2007 to 2008 and surged to its historical height in 2008 but soon suffered a major crash and only partially recovered in the middle of 2008 and stepped into bear market mode before long. This bear market mode lasted for almost seven years, and only finished in 2015, being replaced by a rocket bull market mode. Unfortunately, the 2015 bull market was very short and tumbled greatly into bear market mode again with huge drops. For S&P500, the US stock markets also suffered a great market crisis in 2008, but the market changed into a very long climbing bull market in 2009.

To investigate how the u^2 changes before and after a market crash, we choose a case study period between 24 July 2008 and 16 February 2009 for CSI300 centering with a market turning point on 4 November 2008, covering 135 trading days and a period between 26 December 2008 and 2 June 2009, and for S&P500 centering with a market turning point on 9 March 2009, covering 108 trading days. We denote the ranking for stock s_i at time t as $R_i(t)$ according to the normalized values. For a period of $[t_s, t_e]$ of length $L_{s,e}$, the averaged ranking for s_i is

$$\langle R_i(t) \rangle = \frac{1}{L_{s,e}} \sum_{t=t_s}^{t_e} R_i(t), \tag{15}$$

where $L_{s,e} = t_e - t_s + 1$ is the number of trading dates in the period. By calculating all averaged rankings for all of the stocks in both periods before and after the market crash, we can get the top and bottom 10 stocks for CSI163 and S&P468. The top and bottom 10 stocks according to the averaged ranking for CSI163 in the *Fall* stage and *Climb* stage are presented in Tables 2 and 3, respectively. The same lists are presented in Tables 4 and 5 for the Fall and Climb stages of S&P468.

Table 2. The top ten and bottom ten stocks of the second largest eigenvalue u^2 of CSI163 ranked by the average u^2 components values in the Fall stage between 24 July 2008 and 4 November 2008.

Top 10 Rank	**Tick**	**Stock Name**	**Industry**
1	2007	Hualan Biological Engineering Inc.	Pharmaceuticals
2	600,867	Star Lake Bioscience Co., Inc.	Pharmaceuticals
3	600,085	Beijing Tongrentang Co., Ltd.	Pharmaceuticals
4	963	Huadong Medicine Co., Ltd.	Wholesale
5	600,332	Sichuan Hongda Co., Ltd.	Metals
6	600,108	Gansu Yasheng Industrial (Group) Co., Ltd.	Agriculture
7	600,535	Nanjing Chixia Development Co., Ltd.	Real estate
8	600,277	Jiangsu Hengrui Medicine Co., Ltd.	Pharmaceuticals
9	600,089	TBEA Co., Ltd.	Machinery
10	999	Sanjiu Medical & Pharmaceutical Co., Ltd.	Pharmaceuticals
Bottom 10 Rank	**Tick**	**Stock Name**	**Industry**
154	46	Oceanwide Construction Group Co., Ltd.	Real estate
155	601,988	China Construction Bank	Finance
156	2	China Vanke Co., Ltd.	Real estate
157	600,048	Poly Real Estate Group Co., Ltd.	Real estate
158	601,398	Guangshen Railway	Transportation
159	600,016	China Minsheng Banking Corp. Ltd.	Finance
160	600,015	Hua Xia Bank Co., Ltd.	Finance
161	1	Shenzhen Development Bank Co., Ltd.	Finance
162	600,036	China Merchants Bank Co., Ltd.	Finance
163	600,000	Shanghai Pudong Development Bank	Finance

The tables reveal some exciting results. In Table 2, we see that stocks of finance and real estate occupy the bottom ten while stocks of pharmaceuticals dominate the top 10 in the Fall stage of CSI300, and this phenomenon remains unchanged during the Climbing stage after the market turning point. This indicates and confirms again that financials are not the only dominating players in the Chinese stock market. In the Climbing stage, as shown in Table 2, stocks of pharmaceuticals still dominate the top 10, and the stocks of finance remain at the bottom part. This shows that the internal structure of the CSI300 market remains almost unchanged before and after the market crashes.

Being opposite to CSI163, S&P468 demonstrates a different behavior before and after the crash period. As shown in Table 4, stocks of financials dominate the top positions with the smallest rankings; in other words, stocks of financials play significant roles in the Fall stage; however, stocks of energy collectively occupy the bottom 10. When the market entered the Climb stage, passing the turning point, the whole rankings reversed with stocks of energy becoming the top stocks whereas the financials stocks fell to the bottom, as shown in Table 5.

Table 3. The top ten and bottom ten stocks of the second largest eigenvalue u^2 of CSI163 ranked by the average u^2 components values in the Climb stage between 4 November 2008 and 16 February 2009.

Top 10 Rank	Tick	Stock Name	Industry
1	999	Sanjiu Medical & Pharmaceutical Co., Ltd.	Pharmaceuticals
2	2007	Hualan Biological Engineering Inc.	Pharmaceuticals
3	629	Panzhihua New Steel & Vanadium Co., Ltd.	Metals
4	600,089	TBEA Co., Ltd.	Machinery
5	600,085	Beijing Tongrentang Co., Ltd.	Pharmaceuticals
6	538	Yunnan Baiyao Industry Co., Ltd.	Pharmaceuticals
7	963	Huadong Medicine Co., Ltd.	Wholesale
8	729	Beijing Yanjing Brewery Co., Ltd.	Food & Beverage
9	600,535	Nanjing Chixia Development Co., Ltd.	Real estate
10	600,332	Sichuan Hongda Co., Ltd.	Metals
Bottom 10 Rank	**Tick**	**Stock Name**	**Industry**
459	157	Changsha Zoomlion Heavy Industry	Machinery
460	600,030	CITIC Securities Co., Ltd.	Finance
461	600,585	Jiangsu Changjiang Electronics Technology	Electronics
462	601,988	China Construction Bank	Finance
463	601,398	Guangshen Railway	Transportation
464	1	Shenzhen Development Bank Co., Ltd.	Finance
465	600,015	Hua Xia Bank Co., Ltd.	Finance
466	600,016	China Minsheng Banking Corp. Ltd.	Finance
467	600,036	China Merchants Bank Co., Ltd.	Finance
468	600,000	Shanghai Pudong Development Bank Co., Ltd.	Finance

Table 4. The top ten and bottom ten stocks of the second largest eigenvalue u^2 of S&P468 ranked by the average u^2 components values in the Fall stage between 26 December 2008 and 9 March 2009.

Top 10 Rank	Tick	Stock Name	Industry
1	STI	SunTrust Banks	Financials
2	ZION	Zions Bancorp	Financials
3	MTB	M&T Bank Corp.	Financials
4	CMA	Comerica Inc.	Financials
5	WFC	Wells Fargo	Financials
6	BBT	BB&T Corporation	Financials
7	JPM	JPMorgan Chase & Co.	Financials
8	RF	Regions Financial Corp.	Financials
9	LEN	Lennar Corp.	Consumer Discretionary
10	PNC	PNC Financial Services	Financials
Bottom 10 Rank	**Tick**	**Stock Name**	**Industry**
459	EOG	EOG Resources	Energy
460	MUR	Murphy Oil	Energy
461	OXY	Occidental Petroleum	Energy
462	HP	Helmerich & Payne	Energy
463	NBL	Noble Energy Inc.	Energy
464	XEC	Cimarex Energy	Energy
465	APC	Anadarko Petroleum Corp.	Energy
466	DO	Diamond Offshore Drilling	Energy
467	DVN	Devon Energy Corp.	Energy
468	APA	Apache Corporation	Energy

Table 5. The top ten and bottom ten stocks of the second largest eigenvalue u^2 of S&P468 ranked by the average u^2 components values in the Climb stage between 10 March 2009 and 2 June 2009.

Top 10 Rank	**Tick**	**Stock Name**	**Industry**
1	APA	Apache Corporation	Energy
2	DVN	Devon Energy Corp.	Energy
3	ETR	Entergy Corp.	Utilities
4	DO	Diamond Offshore Drilling	Energy
5	NBL	Noble Energy Inc.	Energy
6	APC	Anadarko Petroleum Corp.	Energy
7	FE	FirstEnergy Corp.	Utilities
8	OXY	Occidental Petroleum	Energy
9	MUR	Murphy Oil	Energy
10	XOM	Exxon Mobil Corp.	Energy
Bottom 10 Rank	**Tick**	**Stock Name**	**Industry**
459	USB	US Bancorp	Financials
460	JPM	JPMorgan Chase & Co.	Financials
461	RF	Regions Financial Corp.	Financials
462	WFC	Wells Fargo	Financials
463	BBT	BB&T Corporation	Financials
464	PNC	PNC Financial Services	Financials
465	ZION	Zions Bancorp	Financials
466	CMA	Comerica Inc.	Financials
467	MTB	M&T Bank Corp.	Financials
468	STI	SunTrust Banks	Financials

6. Conclusions and Discussion

In this study, we applied random matrix theory to study the eigenvalues and their eigenvectors of the US and Chinese stock markets. The correlation properties are studied, and some eigenvalues of the correlation matrices beyond the predicted bounds are observed in both markets. The largest eigenvalues λ_1 are dozens of times larger than the predicted λ_{max}. They are found to be potential market indicators. Eigenvalue deviation fractions beyond the predicted largest eigenvalue are observed to pinpoint market turning points. For the two markets, the most influential industry sectors are identified. They behave differently when the market crashes. These findings provide information on the dynamics of eigenvalues and eigenvectors. This is useful for investors and regulators to monitor the markets. On the other hand, the eigenvalues are related to factor models. The largest eigenvalue stands for the market itself and the corresponding eigenvector has impacts on most stocks, described as the single factor model for stock s_i: $r_i = \beta_i r_M + e_i$, where r_M is the market return, for N stocks, the correlation matrix has one dominant eigenvalue. The CAMP is a special case of a single factor model. However, other eigenvalues are beyond the predicted λ_{max}. It is natural to model the returns in multi-factors as proposed in arbitrage pricing theory (APT), $r_i = \sum \beta_{ki} f_k + e_i$, where f_k is the kth factor. Since the eigenvalues embedded in the predicted bounds represent noises, it is natural to choose the top k largest eigenvalues $\lambda_{max-k}, \ldots, \lambda_{max-1}$; thus, we get k corresponding eigenvectors $v_{max-k}, \ldots, v_{max-1}$. In other words, the k principle components in the PCA. To simplify the model, it is reasonable to consider the sector information revealed in the eigenvectors; in other words, the corresponding eigenvector components belonging the the sector are reserved.

Last but not least, the present work still has several limitations that should not be neglected and are worth further efforts in future works. First, this work only considered two major markets in an outdated time period. More global markets and updated periods can be considered in future work. Second, this work provides findings largely through empirical analysis rather than rigorous statistical approaches. In order to further validate the findings, statistical testing should be considered. Third, this work only reports observations, and no practical applications are developed to further evaluate the values of the methodology and findings. In the future, applications like quantitative trading strategies, portfolio management, and risk management can be developed around the findings to demonstrate the values in financial practices.

Author Contributions: Conceptualization, Y.T.; methodology, Y.T. and Y.Z. (Yicheng Zhang); formal analysis, Y.T., J.X. (Jason Xiong), Z.C., Y.Z. (Yan Zhuang), K.L., J.X. (Jingcong Xie) and Y.Z. (Yicheng Zhang); data curation, Y.T., writing—original draft preparation, Y.T., writing—review and editing, Y.T., J.X. (Jason Xiong), Z.C., Y.Z. (Yan Zhuang), K.L., J.X. (Jingcong Xie) and Y.Z. (Yicheng Zhang), visualization, Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data used in this study are publicly available from financial websites.

Acknowledgments: The authors would like to thank editors and anonymous reviewers for their constructive discussions, comments, and patience.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Markowitz, H. Portfolio Selection. *J. Financ.* **1952**, *7*, 77–91. [CrossRef]
2. Sharpe, W.F. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *J. Financ.* **1964**, *19*, 425–442. [CrossRef]
3. Livan, G.; Rebecchi, L. Asymmetric correlation matrices: An analysis of financial data. *Eur. Phys. J. B—Condens. Matter Complex Syst.* **2012**, *85*, 213. [CrossRef]
4. Wang, D.; Podobnik, B.; Horvatić, D.; Stanley, H.E. Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices. *Phys. Rev. E* **2011**, *83*, 046121. [CrossRef] [PubMed]
5. Fenn, D.J.; Porter, M.A.; Williams, S.; McDonald, M.; Johnson, N.F.; Jones, N.S. Temporal evolution of financial-market correlations. *Phys. Rev. E* **2011**, *84*, 026109. [CrossRef]
6. Di Matteo, T.; Pozzi, F.; Aste, T. The use of dynamical networks to detect the hierarchical organization of financial market sectors. *Eur. Phys. J. B—Condens. Matter Complex Syst.* **2010**, *73*, 3–11. [CrossRef]
7. Tilak, G.; Széll, T.; Chicheportiche, R.; Chakraborti, A. Study of Statistical Correlations in Intraday and Daily Financial Return Time Series. In *Econophysics of Systemic Risk and Network Dynamics, New Economic Windows*; Abergel, F., Chakrabarti, B.K., Chakraborti, A., Ghosh, A., Eds.; Springer: Milan, Italy, 2013; pp. 77–104.
8. Mantegna, R.N. Hierarchical structure in financial markets. *Eur. Phys. J. B* **1999**, *11*, 193–197. [CrossRef]
9. Nobi, A.; Maeng, S.E.; Ha, G.G.; Lee, J.W. Network Topologies of Financial Market During the Global Financial Crisis. *arXiv* **2013**, arXiv:1307.6974.
10. Tumminello, M.; Lillo, F.; Mantegna, R.N. Correlation, hierarchies, and networks in financial markets. *J. Econ. Behav. Organ.* **2010**, *75*, 40–58. [CrossRef]
11. Dimov, I.I.; Kolm, P.N.; Maclin, L.; Shiber, D.Y.C. Hidden noise structure and random matrix models of stock correlations. *Quant. Financ.* **2012**, *12*, 567–572. [CrossRef]
12. Bouchaud, J.P.; Laloux, L.; Miceli, M.A.; Potters, M. Large dimension forecasting models and random singular value spectra. *Eur. Phys. J. B—Condens. Matter Complex Syst.* **2007**, *55*, 201–207. [CrossRef]
13. Dyson, F.J. Statistical Theory of the Energy Levels of Complex Systems. I. *J. Math. Phys.* **1962**, *3*, 140–156. [CrossRef]
14. Wigner, E.P. Characteristic Vectors of Bordered Matrices with Infinite Dimensions I. In *The Collected Works of Eugene Paul Wigner: Part A: The Scientific Papers*; Wightman, A.S., Ed.; Springer: Berlin/Heidelberg, Germany, 1993; pp. 524–540. ._35. [CrossRef]
15. Potters, M.; Bouchaud, J.P. *A First Course in Random Matrix Theory: For Physicists, Engineers and Data Scientists*; Cambridge University Press: New York, NY, USA, 2020.
16. Tao, T. *Topics in Random Matrix Theory*; American Mathematical Society: Providence, RI, USA, 2023; Volume 132.
17. Laloux, L.; Cizeau, P.; Bouchaud, J.P.; Potters, M. Noise Dressing of Financial Correlation Matrices. *Phys. Rev. Lett.* **1999**, *83*, 1467–1470. [CrossRef]
18. Chen, H.; Mai, Y.; Li, S.P. Analysis of network clustering behavior of the Chinese stock market. *Phys. A Stat. Mech. Appl.* **2014**, *414*, 360–367. [CrossRef]
19. Jiang, X.F.; Chen, T.T.; Zheng, B. Structure of local interactions in complex financial dynamics. *Sci. Rep.* **2014**, *4*, 5321. [CrossRef]

20. Jamali, T.; Jafari, G.R. Spectra of empirical autocorrelation matrices: A random-matrix-theory-inspired perspective. *EPL (Europhys. Lett.)* **2015**, *111*, 10001. [CrossRef]
21. Kumar, S.; Kumar, S.; Kumar, P. Diffusion entropy analysis and random matrix analysis of the Indian stock market. *Phys. A Stat. Mech. Appl.* **2020**, *560*, 125122. [CrossRef]
22. Saeedian, M.; Jamali, T.; Kamali, M.; Bayani, H.; Yasseri, T.; Jafari, G.R. Emergence of world-stock-market network. *Phys. A Stat. Mech. Appl.* **2019**, *526*, 120792. [CrossRef]
23. Raei, R.; Namaki, A.; Vahabi, H. Analysis of collective behavior of Iran banking sector by random matrix theory. *Iran. J. Financ.* **2019**, *3*, 60–75. [CrossRef]
24. Vahabi, H.; Namaki, A.; Raei, R. Comparing the collective behavior of banking industry in emerging markets versus mature ones by random matrix approach. *Front. Phys.* **2022**, *10*, 896303. [CrossRef]
25. Taştan, B.; Imamoglu, H. The analysis of cross-correlation between Istanbul Stock Exchange and major stock markets and indices: An empirical analysis using Random Matrix Theory. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e7113. . [CrossRef]
26. Bun, J.; Bouchaud, J.P.; Potters, M. Cleaning large correlation matrices: Tools from random matrix theory. *Phys. Rep.* **2017**, *666*, 1–109. [CrossRef]
27. Heidari Haratemeh, M. Portfolio Optimization and Random Matrix Theory in Stock Exchange. *Innov. Manag. Oper. Strateg.* **2021**, *2*, 257–267. [CrossRef]
28. Namaki, A.; Shirazi, A.H.; Raei, R.; Jafari, G. Network analysis of a financial market based on genuine correlation and threshold method. *Phys. A Stat. Mech. Appl.* **2011**, *390*, 3835–3841. [CrossRef]
29. Tang, Y.; Xiong, J.J.; Jia, Z.Y.; Zhang, Y.C. Complexities in Financial Network Topological Dynamics: Modeling of Emerging and Developed Stock Markets. *Complexity* **2018**, *2018*, 4680140. [CrossRef]
30. Tang, Y.; Xiong, J.J.; Luo, Y.; Zhang, Y.C. How Do the Global Stock Markets Influence One Another? Evidence from Finance Big Data and Granger Causality Directed Network. *Int. J. Electron. Commer.* **2019**, *23*, 85–109. [CrossRef]
31. Pafka, S.; Kondor, I. Noisy covariance matrices and portfolio optimization II. *Phys. A Stat. Mech. Appl.* **2003**, *319*, 487–494. [CrossRef]
32. Laloux, L.; Cizeau, P.; Potters, M.; Bouchaud, J.P. Random matrix theory and financial correlations. *Int. J. Theor. Appl. Financ.* **2000**, *3*, 391–397. [CrossRef]
33. Mantegna, R.N.; Stanley, H.E. *An Introduction to Econophysics: Correlations and Complexity in Finance*; Cambridge University Press: New York, NY, USA, 2000.
34. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L.A.N.; Stanley, H.E. A random matrix theory approach to financial cross-correlations. *Phys. A Stat. Mech. Appl.* **2000**, *287*, 374–382. [CrossRef]
35. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L.A.N.; Stanley, H.E. Econophysics: Financial time series from a statistical physics point of view. *Phys. A Stat. Mech. Appl.* **2000**, *279*, 443–456. [CrossRef]
36. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L.A.N.; Stanley, H.E. Collective behavior of stock price movements—A random matrix theory approach. *Phys. A Stat. Mech. Appl.* **2001**, *299*, 175–180. [CrossRef]
37. Tola, V.; Lillo, F.; Gallegati, M.; Mantegna, R.N. Cluster analysis for portfolio optimization. *J. Econ. Dyn. Control* **2008**, *32*, 235–258. [CrossRef]
38. Rosenow, B.; Plerou, V.; Gopikrishnan, P.; Amaral, L.A.N.; Stanley, H.E. Application of random matrix theory to study cross-correlations of stock prices. *Int. J. Theor. Appl. Financ.* **2000**, *03*, 399–403. [CrossRef]
39. Burda, Z.; Jarosz, A.; Nowak, M.A.; Jurkiewicz, J.; Papp, G.; Zahed, I. Applying free random variables to random matrix analysis of financial data. Part I: The Gaussian case. *Quant. Financ.* **2011**, *11*, 1103–1124. [CrossRef]
40. Bai, Z.; Liu, H.; Wong, W.K. Enhancement of the applicability of Markowitz's portfolio optimization by utilizing random matrix theory. *Math. Financ.* **2009**, *19*, 639–667. [CrossRef]
41. Biely, C.; Thurner, S. Random matrix ensembles of time-lagged correlation matrices: Derivation of eigenvalue spectra and analysis of financial time-series. *Quant. Financ.* **2008**, *8*, 705–722. [CrossRef]
42. Luo, Y.; Xiong, J.; Dong, L.G.; Tang, Y. Statistical correlation properties of the SHIBOR interbank lending market. *China Financ. Rev. Int.* **2015**, *5*, 91–102. [CrossRef]
43. Pafka, S.; Kondor, I. Estimated correlation matrices and portfolio optimization. *Phys. A Stat. Mech. Appl.* **2004**, *343*, 623–634. [CrossRef]
44. Jiang, X.F.; Zheng, B. Anti-correlation and subsector structure in financial systems. *EPL (Europhys. Lett.)* **2012**, *97*, 48006. [CrossRef]
45. Ouyang, F.; Zheng, B.; Jiang, X. Spatial and temporal structures of four financial markets in Greater China. *Phys. A Stat. Mech. Appl.* **2014**, *402*, 236–244. [CrossRef]
46. Lim, K.; Kim, M.J.; Kim, S.; Kim, S.Y. Statistical properties of the stock and credit market: RMT and network topology. *Phys. A Stat. Mech. Appl.* **2014**, *407*, 66–75. [CrossRef]
47. Namaki, A.; Raei, R.; Ardalankia, J.; Hedayatifar, L.; Hosseiny, A.; Haven, E.; Jafari, G.R. Analysis of the Global Banking Network by Random Matrix Theory. *Front. Phys.* **2021**, *8*, 586561. [CrossRef]
48. Glasserman, P.; Young, H.P. Contagion in Financial Networks. *J. Econ. Lit.* **2016**, *54*, 779–831. [CrossRef]
49. Elliott, M.; Golub, B.; Jackson, M.O. Financial Networks and Contagion. *Am. Econ. Rev.* **2014**, *104*, 3115–3153. [CrossRef]

50. Li, F.; Kang, H.; Xu, J. Financial stability and network complexity: A random matrix approach. *Int. Rev. Econ. Financ.* **2022**, *80*, 177–185. [CrossRef]
51. Amini, H.; Cont, R.; Minca, A. RESILIENCE TO CONTAGION IN FINANCIAL NETWORKS. *Math. Financ.* **2016**, *26*, 329–365. [CrossRef]
52. Plerou, V.; Gopikrishnan, P.; Rosenow, B.; Amaral, L.A.N.; Guhr, T.; Stanley, H.E. Random matrix approach to cross correlations in financial data. *Phys. Rev. E* **2002**, *65*, 066126. [CrossRef]
53. Alaoui, M.E. Random matrix theory and portfolio optimization in Moroccan stock exchange. *Phys. A Stat. Mech. Appl.* **2015**, *433*, 92–99. [CrossRef]
54. Sharifi, S.; Crane, M.; Shamaie, A.; Ruskin, H. Random matrix theory for portfolio optimization: A stability approach. *Phys. A Stat. Mech. Appl.* **2004**, *335*, 629–643. [CrossRef]
55. Kwapień, J.; Drożdż, S.; Oświe¸cimka, P. The bulk of the stock market correlation matrix is not pure noise. *Phys. A Stat. Mech. Appl.* **2006**, *359*, 589–606. [CrossRef]
56. Nie, C.X. Analyzing financial correlation matrix based on the eigenvector–eigenvalue identity. *Phys. A Stat. Mech. Appl.* **2021**, *567*, 125713. [CrossRef]
57. Pharasi, H.K.; Sharma, K.; Chakraborti, A.; Seligman, T.H., Complex Market Dynamics in the Light of Random Matrix Theory. In *New Perspectives and Challenges in Econophysics and Sociophysics*; Abergel, F., Chakrabarti, B.K., Chakraborti, A., Deo, N., Sharma, K., Eds.; Springer International Publishing: Cham, Switaerland, 2019; pp. 13–34. [CrossRef]
58. Bun, J.; Bouchaud, J.P.; Potters, M. Overlaps between eigenvectors of correlated random matrices. *Phys. Rev. E* **2018**, *98*, 052145. [CrossRef]
59. García-Medina, A.; Sandoval, L.; Bañuelos, E.U.; Martínez-Argüello, A. Correlations and flow of information between the New York Times and stock markets. *Phys. A Stat. Mech. Appl.* **2018**, *502*, 403–415. [CrossRef]
60. Ji, J.; Huang, C.; Cao, Y.; Hu, S. The network structure of Chinese finance market through the method of complex network and random matrix theory. *Concurr. Comput. Pract. Exp.* **2019**, *31*, e4877.
61. Zitelli, G.L. Random matrix models for datasets with fixed time horizons. *Quant. Financ.* **2020**, *20*, 769–781. [CrossRef]
62. Baruccaand, P.; Kieburg, M.; Ossipov, A. Eigenvalue and eigenvector statistics in time series analysis. *EPL (Europhys. Lett.)* **2020**, *129*, 60003. [CrossRef]
63. Han, R.Q.; Xie, W.J.; Xiong, X.; Zhang, W.; Zhou, W.X. Market Correlation Structure Changes Around the Great Crash: A Random Matrix Theory Analysis of the Chinese Stock Market. *Fluct. Noise Lett.* **2017**, *16*, 1750018. [CrossRef]
64. Yang, Y. Comparison and Analysis of Chinese and United States Stock Market. *J. Financ. Risk Manag.* **2020**, *9*, 44. [CrossRef]
65. Zhang, Y.; Mao, J. COVID-19's impact on the spillover effect across the Chinese and U.S. stock markets. *Financ. Res. Lett.* **2022**, *47*, 102684. [CrossRef]
66. Jin, Z.; Guo, K. The Dynamic Relationship between Stock Market and Macroeconomy at Sectoral Level: Evidence from Chinese and US Stock Market. *Complexity* **2021**, *2021*, 6645570. [CrossRef]
67. Chen, Y.; Pantelous, A.A. The U.S.-China trade conflict impacts on the Chinese and U.S. stock markets: A network-based approach. *Financ. Res. Lett.* **2022**, *46*, 102486. [CrossRef]
68. Urama, T.C.; Ezepue, P.O.; Nnanwa, C.P. Analysis of Cross-Correlations in Emerging Markets Using Random Matrix Theory. *J. Math. Financ.* **2017**, *7*, 18. [CrossRef]
69. Pharasi, H.K.; Sharma, K.; Chatterjee, R.; Chakraborti, A.; Leyvraz, F.; Seligman, T.H. Identifying long-term precursors of financial market crashes using correlation patterns. *New J. Phys.* **2018**, *20*, 103041. [CrossRef]
70. Batondo, M.; Uwilingiye, J. Comovement across BRICS and the US Stock Markets: A Multitime Scale Wavelet Analysis. *Int. J. Financ. Stud.* **2022**, *10*, 27. [CrossRef]
71. Vuong, G.T.H.; Nguyen, M.H.; Huynh, A.N.Q. Volatility spillovers from the Chinese stock market to the U.S. stock market: The role of the COVID-19 pandemic. *J. Econ. Asymmetries* **2022**, *26*, e00276. [CrossRef]
72. Pan, Q.; Mei, X.; Gao, T. Modeling dynamic conditional correlations with leverage effects and volatility spillover effects: Evidence from the Chinese and US stock markets affected by the recent trade friction. *N. Am. J. Econ. Financ.* **2022**, *59*, 101591. [CrossRef]
73. Ren, F.; Zhou, W.X. Dynamic Evolution of Cross-Correlations in the Chinese Stock Market. *PLoS ONE* **2014**, *9*, e97711. [CrossRef]
74. Said, S.; Heuveline, S.; Mostajeran, C. Riemannian statistics meets random matrix theory: Toward learning from high-dimensional covariance matrices. *IEEE Trans. Inf. Theory* **2022**, *69*, 472–481. [CrossRef]
75. Zhu, W.; Ma, X.; Zhu, X.H.; Ugurbil, K.; Chen, W.; Wu, X. Denoise Functional Magnetic Resonance Imaging With Random Matrix Theory Based Principal Component Analysis. *IEEE Trans. Biomed. Eng.* **2022**, *69*, 3377–3388. [CrossRef]
76. Bouchaud, J.P.; Potters, M. *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*; Cambridge University Press: New York, NY, USA, 2003.
77. Rosenow, B.; Gopikrishnan, P.; Plerou, V.; Stanley, H.E. Dynamics of cross-correlations in the stock market. *Phys. A Stat. Mech. Appl.* **2003**, *324*, 241–246. [CrossRef]
78. Giardina, I.; Bouchaud, J.P.; Mézard, M. Microscopic models for long ranged volatility correlations. *Phys. A Stat. Mech. Appl.* **2001**, *299*, 28–39. [CrossRef]
79. Kenett, D.Y.; Huang, X.; Vodenska, I.; Havlin, S.; Stanley, H.E. Partial correlation analysis: Applications for financial markets. *Quant. Financ.* **2015**, *15*, 569–578. [CrossRef]

80. Iori, G.; Masi, G.D.; Precup, O.V.; Gabbi, G.; Caldarelli, G. A network analysis of the Italian overnight money market. *J. Econ. Dyn. Control* **2008**, *32*, 259–278. [CrossRef]
81. Tanaka-Yamawaki, M.; Yang, X.; Kido, T.; Yamamoto, A. Extracting Market Trends from the Cross Correlation between Stock Time Series. In *Advanced Techniques for Knowledge Engineering and Innovative Applications: Proceedings of the 16th International Conference, KES 2012, San Sebastian, Spain, 10–12 September 2012*; Revised Selected Papers; Tweedale, J.W., Jain, L.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 25–38. [CrossRef]