Chapter 2

Random Matrix Theory Approach to Financial Markets

In the previous chapter we have applied the random matrix theory (RMT) approach to a random system. The RMT has been successfully used to investigate phenomena from different areas such as quantum field theory, quantum chaos, disordered systems, and recently to a large number of financial markets [9, 47, 48, 49, 50, 51, 78] to investigate the structure of cross-correlations in financial markets. Financial time series appears to be unpredictable but it does not mean that the time series of price of financial index does not reflect any important economic information. Instead the financial time series carries such a huge amount of information that it is difficult to extract the economic information from it. Thus in this chapter, we apply the RMT approach to a real world system i.e. the global financial markets to extract important information from the cross-correlation matrices constructed by using the global financial time series before and during the financial crisis of 2008.

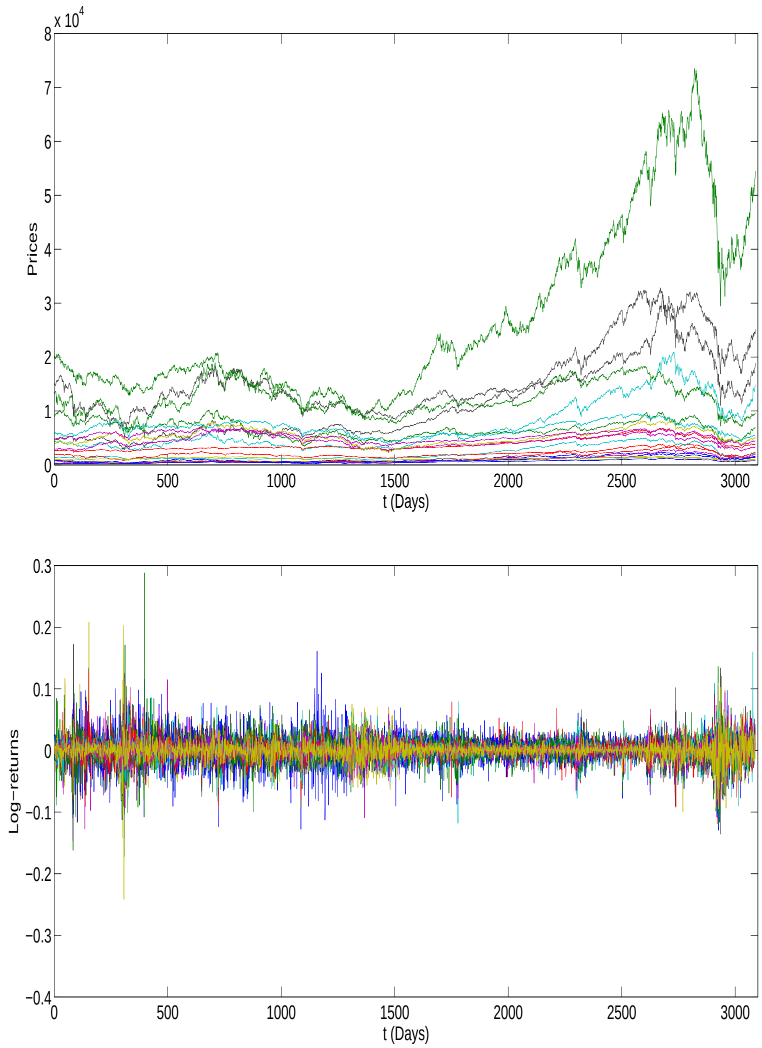


Figure 2.1: *(a) Daily closing prices of financial indices of 20 countries for the period July, 1997 to June, 2009 (b) Corresponding log-returns.*

2.1 Data Analyzed

We analyze the daily closing prices of 20 financial markets around the world traded from the period July 2, 1997 to June 1, 2009 as shown in Fig. 2.1. These indices are as follows: Argentina, MERV; Brazil, BVSP; Egypt, CCSI; India, BSESN; Indonesia, JKSE; Malaysia, KLSE; Mexico, MXX; South Korea, KS11; Taiwan, TWII; Australia, AORD; Austria, ATX; France, FCHI; Germany, GDAXI; Hong Kong, HSI; Israel, TA100; Japan, N225; Singapore, STI; Switzerland, SSMI; the

United Kingdom, FTSE; and the United States, GSPC. The data were obtained from [52]. There are differences in public holidays or weekends among countries so the data are shifted according to the rule that when more than 30% of markets did not open on a certain day, we remove that day from the data, and when it was fewer than 30%, we kept existing indices and inserted the last closing price for each of the remaining indices. Also these markets do not operate in the same time zones. It has been reported [50] that correlations of Asian with the U.S. indices increases when one considers the correlation of the U.S. indices with the next day indices of the Asian market. We did not consider weekly data to avoid the problem of different operating hours between international markets so that we do not miss major changes in markets which tend to occur during a small interval of days. Thus, we considered all indices taken on the same date and filtered the data as in [50].

2.1.1 Global Financial Crisis of 2008

The global financial crisis of 2007-2009 is known to be the worst financial crisis since the Great Depression of the 1930s and had its origins in the United States then spread to the world [18]. The triggering cause of this crisis was the collapse of the housing bubble in the U.S. in 2007 and increase in interest rates [53]. This has caused severe losses that destabilized the financial strength of banking institutions. Several major institutions like Lehman Brothers were failed. The interconnections of the global financial systems has played a major role in spreading of this financial crisis from one institution to the rest of the financial system. In order to investigate the effect of the global financial crisis of 2008 on the correlations in global financial indices, we have considered two periods for the study i.e. the period before the crisis (June 7, 2006 to November 30, 2007) and during the crisis (December, 2007 to June, 2009). The periods before (calm) and during

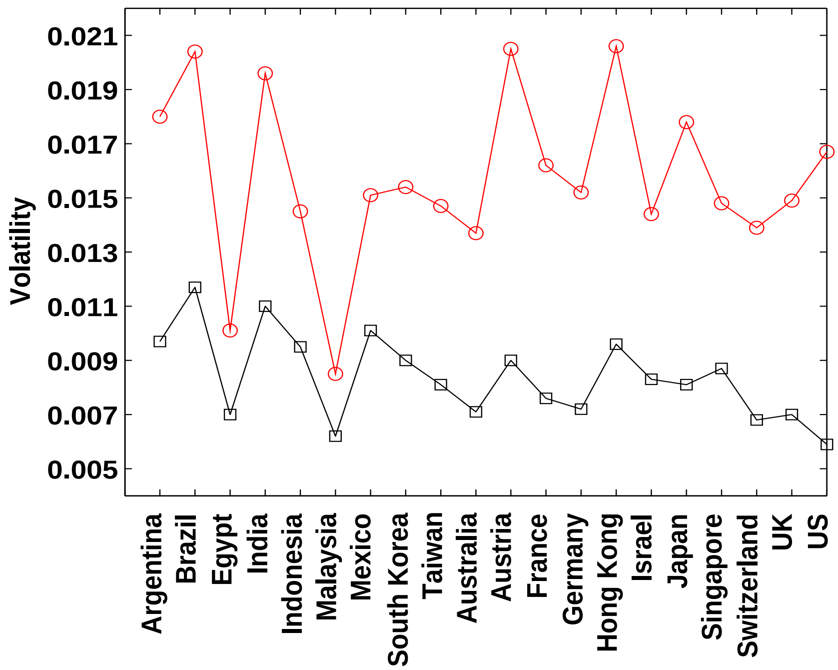


Figure 2.2: *Volatility as a measure of fluctuation in financial indices before (△) and during (○) the crisis.*

the crisis are chosen by observing the volatility. The **volatility** gives a measure of fluctuations in the market. We quantify the volatility, as the local average of the absolute value of daily returns of indices in an appropriate time window of T days, as an estimate of volatility in that period $v(t) = \frac{\sum_{t=1}^{T-1} |R(t)|}{T-1}$. The volatility for these two periods for individual countries before and during the crisis is shown in Fig. 2.2. We find that there is an increase in the value of volatility of each index during the crisis period.

2.2 Random Matrix Theory in Finance

The Random Matrix Theory (RMT) is a quantitative technique which has been used here to extract important information contained in the cross-correlation matrix of the global financial time series.

2.2.1 Correlations in Global Financial Indices

Let $S_i(t)$ and $R_i(t)$ denote the daily closing prices and returns of financial indices i at time t ($i = 1, 2, ..., N; t = 1, 2, ..., T$), respectively. The logarithmic returns is defined as,

$$R_i(t) \equiv \ln(S_i(t + \Delta t)) - \ln(S_i(t)), \tag{2.1}$$

where $\Delta t = 1$ day is the time lag. The normalized return for index i is defined as,

$$r_i(t) \equiv (R_i(t) - \langle R_i \rangle)/\sigma_i, \tag{2.2}$$

where $\sigma_i \equiv \sqrt{\langle R_i^2 \rangle - \langle R_i \rangle^2}$ is the standard deviation of R_i and $\langle \cdots \rangle$ denotes the time average over the period studied. The equal-time cross-correlation matrix is computed with elements

$$C_{ij} \equiv \langle r_i(t) r_j(t) \rangle \tag{2.3}$$

which are limited to the domain $[-1, 1]$. For the global financial indices C_{ij} = 1 (-1) corresponds to perfect correlation (anti-correlation) in indices and C_{ij} = 0 corresponds to no correlation. The financial data of N = 20 indices for T = 387 days have been used to analyze the crisis. The value of the average correlation coefficient $\langle C_{ij} \rangle$ increases from 0.4353 (before the crisis) to 0.4634 (during the crisis) in response to the financial crisis. The statistical properties of a Wishart matrix (a correlation matrix of uncorrelated time series with finite length) are known [6]. In the limit $N \to \infty, T \to \infty$ with $Q \equiv T/N$ (≥ 1), the probability distribution of the eigenvalue λ_{rm} is given by equation 1.1 within the bounds $\lambda_{min} \leq \lambda_i \leq \lambda_{max}$ and 0 otherwise [9]. The smallest (largest) eigenvalue of the random correlation matrix is given by equation 1.2.

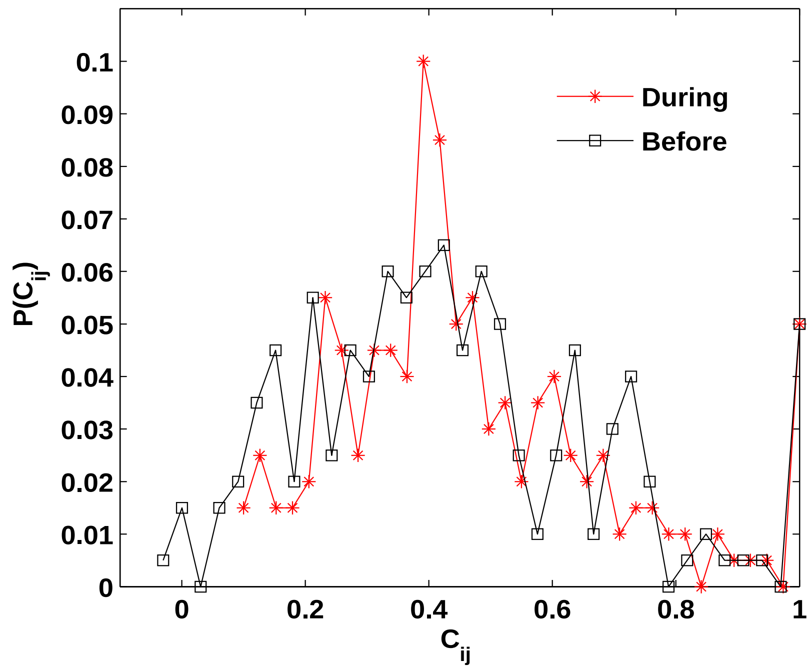


Figure 2.3: *Plot of the probability density of elements of correlation matrix C calculated using daily returns of 20 indices before and during the crisis. We find the average magnitude of correlation* $\langle |C| \rangle = 0.435$ *before and* $\langle |C| \rangle = 0.463$ *during the crisis respectively.*

2.2.2 Statistics of Correlation Coefficients

We calculate the cross-correlation matrix (C_{ij}) from the daily normalized returns of $N = 20$ indices before and during crisis periods. The probability densities ($P(C_{ij})$) of C_{ij} for both periods are compared in Fig. 2.3. We find that $P(C_{ij})$ is asymmetric and centred around a positive mean value which indicate that positive correlation in empirical correlation matrix is larger than the negative correlation. We observe that the value of the average correlation coefficient ($\langle C_{ij} \rangle$) increases from 0.4353 (before the crisis) to 0.4634 (during the crisis) in response to the financial crisis.

2.2.3 Largest Eigenvalue and Eigenvalue Distribution

The largest eigenvalue deviating from RMT prediction reflects that it contain important information hidden in the empirical correlation matrix (C_{ij}). However

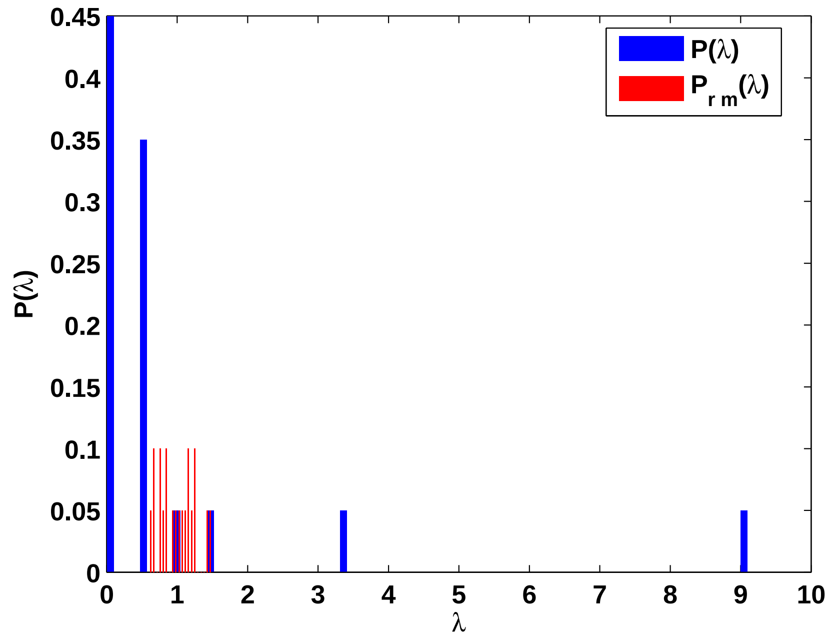


Figure 2.4: *Comparison of probability density function of 20 financial indices before the crisis. For N=20 indices, T=387 days and Q=19.35, $\lambda^{rand}_{min} = 0.597$ and $\lambda^{rand}_{max} = 1.506$ and $\lambda^{real}_{min} = 0.0527$ and $\lambda^{real}_{max} = 9.045$.*

the eigenvalues that are in the RMT bounds, do not reflect any information. If there is no correlation between financial time series then the smallest (largest) eigenvalue should be bounded between numerical RMT predictions, i.e., $\lambda_{min(max)}$ = $0.597(1.5063)$ (for Wishart matrix in appendix A). But for the global financial indices we find that the smallest (largest) eigenvalue is $\lambda^{real}_{min(max)}$ = $0.0527(9.0454)$ before the financial crisis and $\lambda^{real}_{min(max)}$ = $0.0388(9.5282)$ during the financial crisis. Thus, the largest eigenvalues deviate significantly from the RMT prediction in both periods, which shows that there is a strong correlation in the financial indices. The eigenvalue distribution for the empirical correlation matrix before and during the crisis are shown in Figs. 2.4 and 2.5. Here a bulk of eigenvalues lies in the RMT bounds which shows that contents of C_{ij} are mostly random excepting the deviating largest eigenvalues. We find that a few largest eigenvalues lies outside the RMT bound.

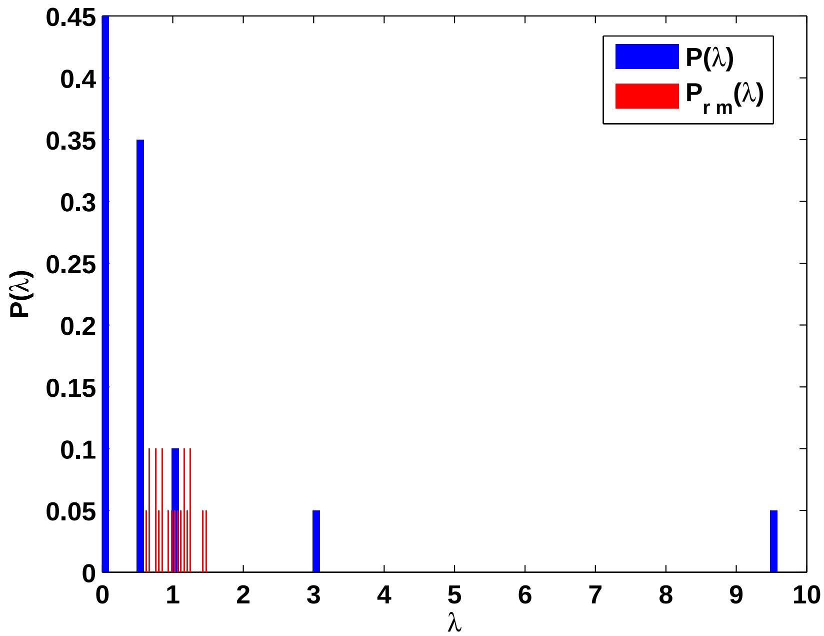


Figure 2.5: *Comparison of probability density function of 20 financial indices during the crisis. For N=20 indices, T=387 days and Q=19.35,* $\lambda_{min}^{rand} = 0.597$ *and* $\lambda_{max}^{rand} = 1.506$ *and* $\lambda_{min}^{real} = 0.0388$ *and* $\lambda_{max}^{real} = 9.528$.

2.2.4 Dynamical Behavior of Largest Eigenvalue

The largest eigenvalue represents the collective information about the correlation between different indices therefore its trend is expected to be dependent on the market conditions. We find that there is an increase in the largest eigenvalue during the financial crisis of 2008. We calculated the first, second, and third largest eigenvalues over sliding windows of 25 days. The trend of these eigenvalues are shown in Fig. 2.6 where the first largest eigenvalues increases during the financial crisis of 2008 while the second and third largest eigenvalues do not show significant change.

2.2.5 Largest Eigenvalue and Corresponding Eigenvector

The components of the eigenvector corresponding to the first largest eigenvalue, before and during the crisis, are shown in Fig. 2.7. We find that all components

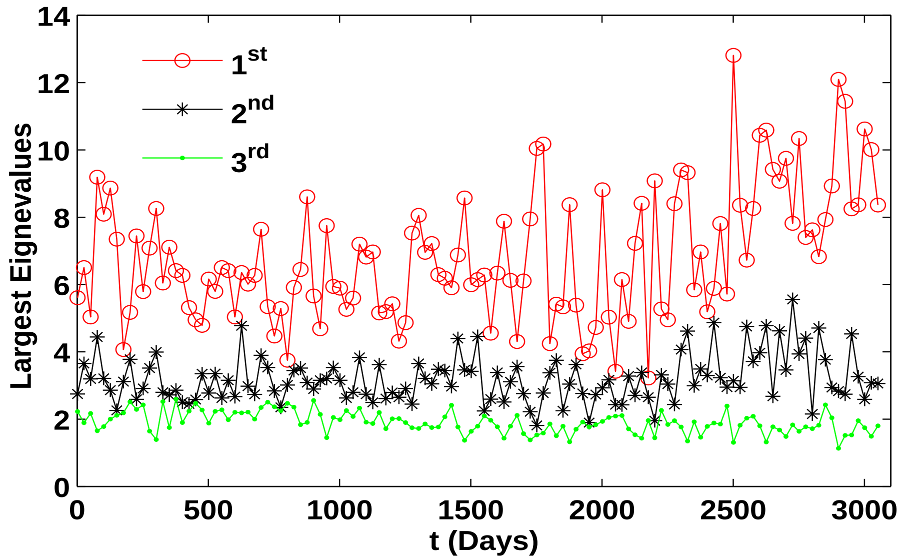


Figure 2.6: *First three largest eigenvalues of the correlation matrices of financial indices using sliding time windows of 25 days.*

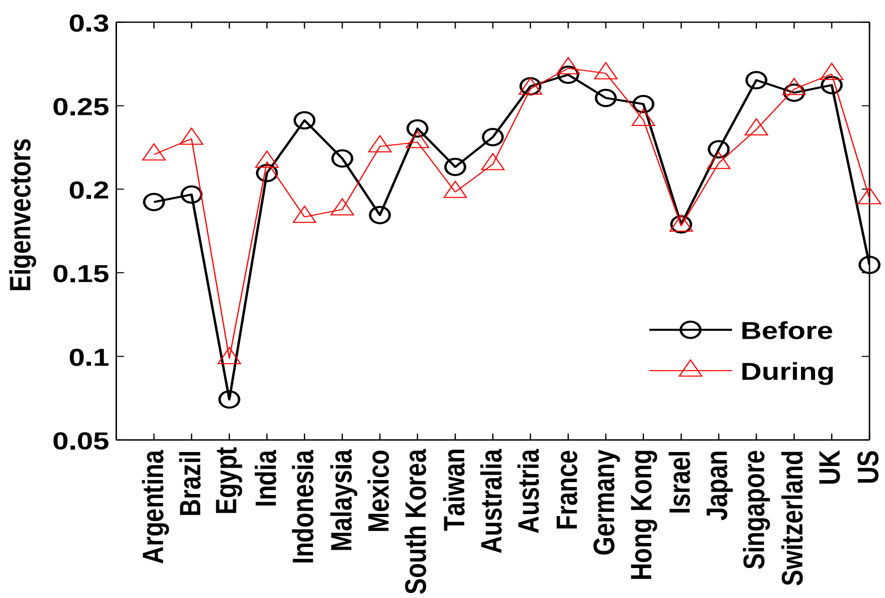


Figure 2.7: *Components of eigenvectors corresponding to first largest eigenvalue. All positive components reflects a common global financial market mode.*

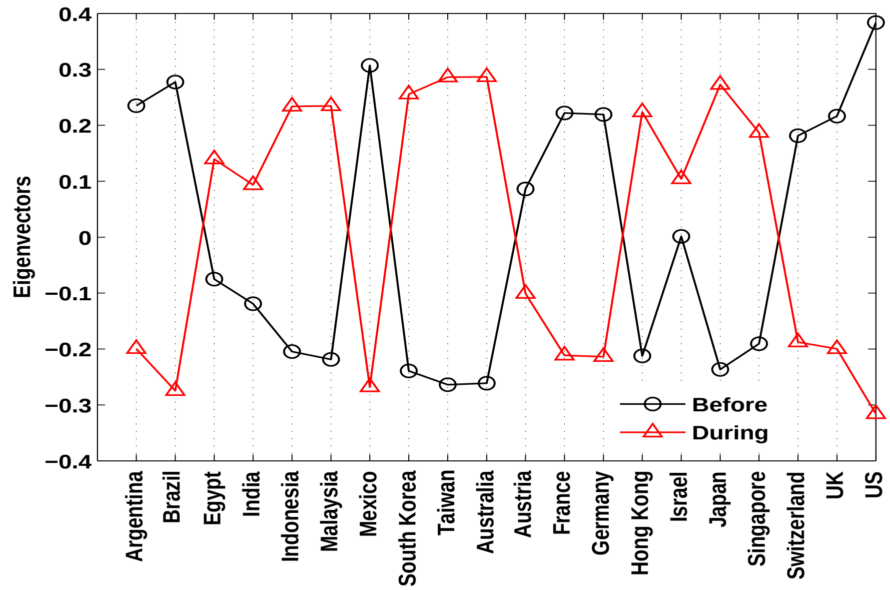


Figure 2.8: *Components of eigenvectors corresponding to second largest eigenvalue. The financial indices form two clusters in the positive and negative directions, respectively. The positive significant contributions of the components are associated with the cluster of American (Argentina, Brazil, Mexico, United States) and European (Austria, France, Germany, Switzerland) indices. The negative significant contributions of the components are associated with the cluster of indices corresponding to Asia-Pacific (Egypt, India, Indonesia, Malaysia, South Korea, Taiwan, Australia, Hong Kong, Japan, Singapore) indices (except Israel). The components of the two clusters switch in opposite directions during the crisis.*

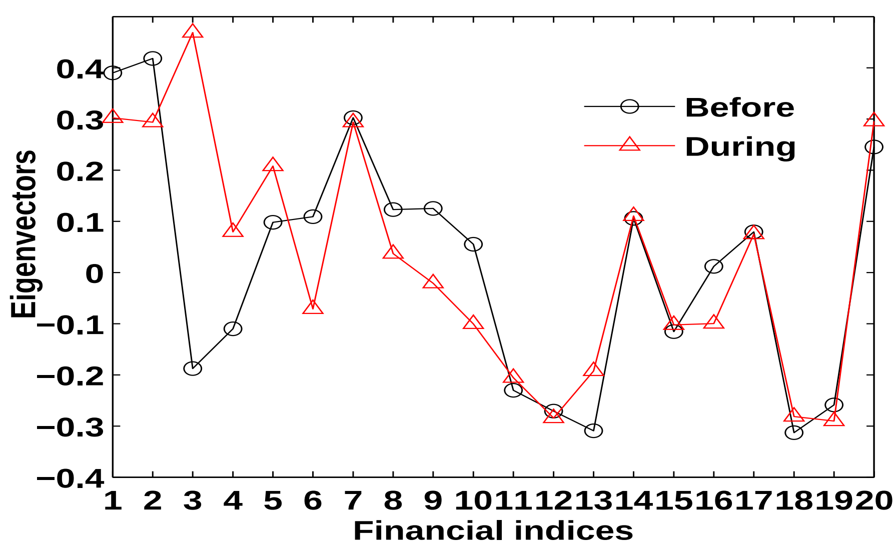
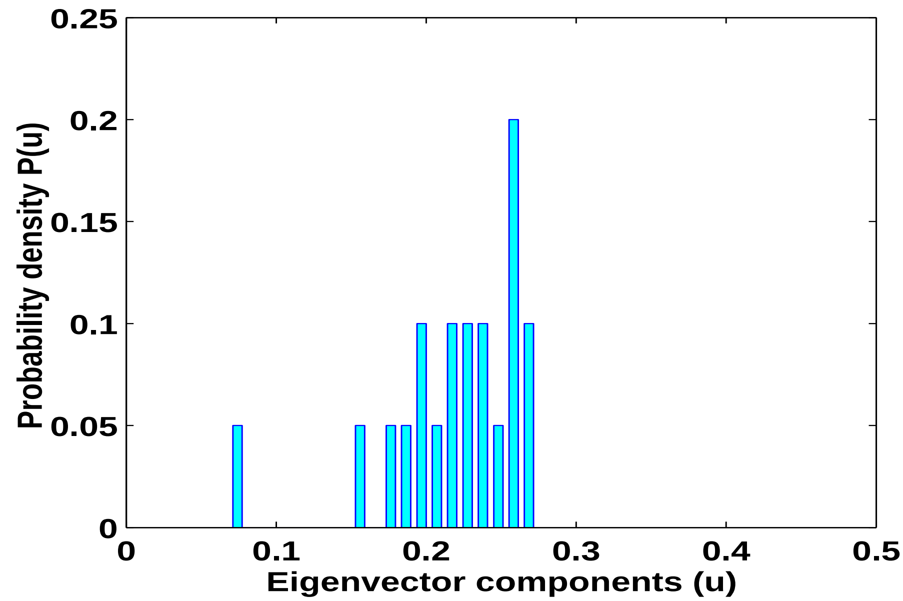


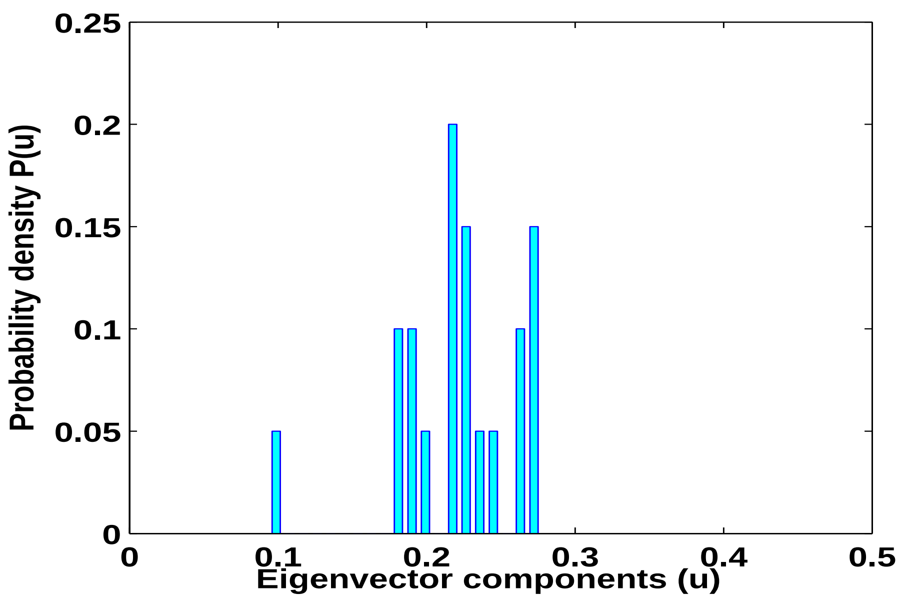
Figure 2.9: *Comparison of eigenvectors corresponding to third largest eigenvalue before and during the financial crisis of 2008 respectively.*

are positive which reflects a common global financial market mode. These components do not show appreciable change due to the financial crisis of 2008. Fig. 2.8 shows the components of the eigenvectors corresponding to the second largest value. Here we find that the positive components of the second eigenvector (Argentina, Brazil, Mexico, Austria, France, Germany, Switzerland, the UK, and the US) switch to negative values during the crisis while the negative components (Egypt, India, Indonesia, Malaysia, South Korea, Taiwan, Australia, Hong Kong, Japan, and Singapore) switch to large positive values during the crisis. However, the components of eigenvector corresponding to third largest eigenvalue does not carry much information as the third largest eigenvalue is near the random matrix bound.

We analyze the distribution of components of eigenvector corresponding to the largest eigenvalue of empirical correlation matrix of global financial indices. The distribution $P_{rm}(u)$ of components $u_l^k; l = 1, ..., N$ of eigenvector u^k of a random correlation matrix follow the Gaussian distribution having zero mean and unit variance i.e. $P_{rm}(u) = \frac{1}{\sqrt{2\pi}} exp(\frac{-u^2}{2})$. The probability distribution of components of eigenvector u^{20} corresponding to largest eigenvalue of empirical correlation matrix are shown in Figs. 2.10 (a) and (b), before and during the crisis. We find that all components of eigenvector are distributed in the positive direction only. Therefore this distribution of components of eigenvector u^{20} deviate significantly from the RMT prediction (see appendix A). This suggests that all components (global financial indices) participate in the eigenvector corresponding to the largest eigenvalue. Therefore the largest eigenvalue and its eigenvector is represented as the collective response of the global financial indices. We observe that the probability distribution of square of eigenvector components of the empirical correlation matrix does not follow the Porter-Thomas distribution as discussed in chapter 1.



(a)



(b)

Figure 2.10: *Distribution of components of eigenvector u^{20} corresponding to the largest eigenvalue of empirical correlation matrix of global financial indices (a) before the crisis (b) during the crisis.*

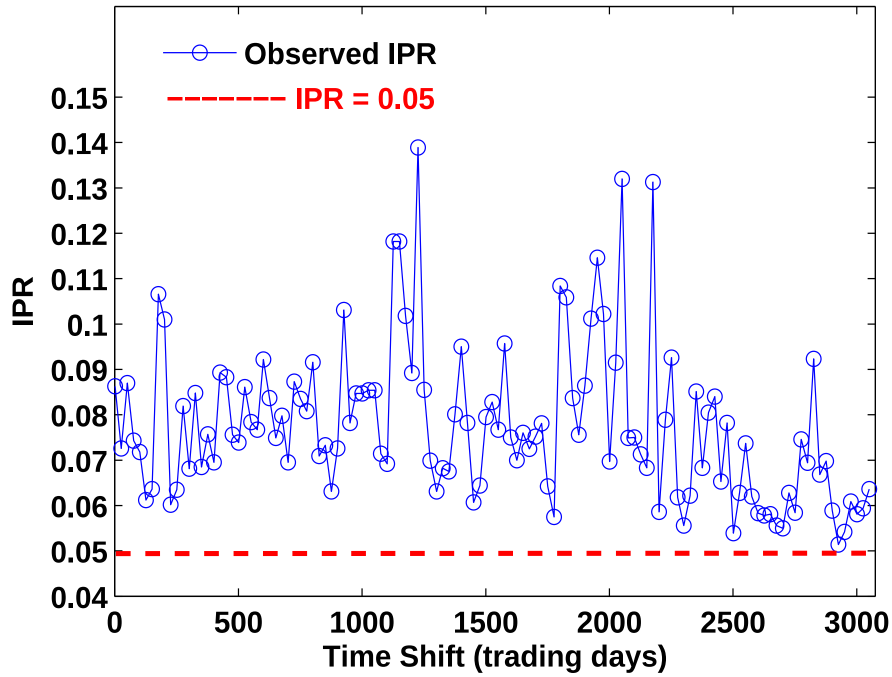


Figure 2.11: *IPR for the eigenvector U^{20} as a function of time, obtained from correlation matrices of financial indices using windows of 25 days. The dashed line marks the value 0.05 of the IPR when all components contribute equally.*

2.2.6 Inverse Participation Ratio

Next we compute the inverse participation ratio (IPR) which allows the quantification of the number of components that participate significantly in each eigenvector and tells us more about the level and nature of the deviation from the RMT. The IPR of the eigenvector u^k is defined by $I^k \equiv \sum_{l=1}^{N} [u_l^k]^4$, where u_l^k, $l = 1, \ldots, N$, are the components of eigenvector u^k. The IPR allows us to compute the inverse of the number of eigenvector components that contribute significantly to each eigenvector. Fig. 2.11 shows the IPR for the eigenvector U^{20}, calculated from the empirical correlation matrices of global financial indices over 25 days time windows. The value of IPR are found close to 0.05 $(= 1/20)$ which is the value of IPR when all components contribute equally. The value of IPR for the empirical correlation matrix before (during) the crisis is found to be 0.056 (0.055).

2.3 Discussion

In this chapter we have applied the random matrix theory approach to investigate the structure of cross-correlations in global financial markets from the period July 2, 1997 to June 1, 2009. The financial data as well as the filtration of data is discussed. We apply RMT to study the global financial crisis of 2008 and investigate the structure of organization of global financial indices. In the random matrix theory approach the first few largest eigenvalues deviate significantly from the RMT prediction and these deviation changes during the crisis. The largest eigenvalue represents the collective information about the correlation between different indices and its trend depends on the market conditions. We have shown that the RMT analysis of correlation matrices provides information about the formation of clusters of indices. We find that the components of the eigenvectors corresponding to the second largest eigenvalue form two clusters of indices in the positive and negative directions. The components of these two clusters switch in opposite directions during the financial crisis of 2008. We also computed the IPR that allows the quantification of the number of components that participate significantly in each eigenvector and tells us more about the level and nature of the deviation from the RMT.

In the next chapter we will apply the complex network techniques to the filtered empirical correlation matrix to capture the structure of organization of global financial indices clearly.