Collective behavior of stock price movements in an emerging market

Raj Kumar Pan and Sitabhra Sinha

CONTENT

- INTRODUCTION
- DATA ANALYSIS FROM NSE
- ANALYSIS USING CORRELATION MATRIX
- EIGEN VALUE SPECTRUM AND ITS PROPERTIES
- RELATION CORRELATION WITH MARKET EVOLUTION
- MODEL OF MARKET DYNAMICS
- CONCLUSIONS

INTRODUCTION

- Fundamental characteristics of financial markets as complex systems

  Financial markets have many interacting elements which exhibit large fluctuations in properties such as stock price or market index. The interactions within the market is governed by either traders or stocks and are also influenced significantly by the arrival of external information.

- Analysis of statistical properties of stock price fluctuations and correlations between different stocks

  Inverse Cubic Law: Physicists have observed that the fluctuation distribution of stock prices follows a power law with an exponent α 3, known as the "inverse cubic law."

- Spectral Properties of the Correlation Matrix

The properties of the empirical correlation matrix differed from those of a random matrix. Deviations from random matrix theory (RMT) in the correlation matrix can offer clues about these interactions. For example, the largest eigenvalue of the correlation matrix represents the influence of the entire market, common for all stocks, while the remaining large eigenvalues are associated with different business sectors.

- Correlation Analysis in Emerging Markets

It is generally believed that stock prices in emerging markets tend to be relatively more correlated than the developed ones. The Indian market shows significant deviations from developed markets in terms of the properties of its collective modes with a higher degree of correlation attributed to a dominant market mode affecting all stocks

- Analysis in Indian Market

A higher degree of correlation in the Indian market compared to developed markets is found to be the result of a dominant market mode affecting all the stocks according to the study, which is further accentuated by the relative absence of clusters of mutually interacting stocks.

- Methodology employed to analyze the network of interactions within the Indian financial market

The methodology involved here is reconstructing the network of interactions within the market, using a filtered correlation matrix from which the common market influence and random noise were removed.This method gives a more accurate representation of the intra-market structure than the commonly used method of constructing minimal spanning tree from the unfiltered correlation matrix

- Two factor model of market dynamics

In this model, it is assumed that the market consists of several correlated groups of stocks which are also influenced by a common external signal, i.e., market mode. It demonstrates that decreasing the intra-group interactions result in spectral distribution properties similar to that seen for the Indian market. It also implies that the appearance and consolidation of distinct group identities signify the transition of a market from emerging to developed status.

DATA ANALYSIS FROM NSE

- Why National Stock Exchange(NSE)?

The National Stock Exchange (NSE) is the largest stock market in India and also world's third largest stock exchange.

- Analysis of missing data points

It was assumed that no trading took place in case of missing data points, and thus the price was considered to remain the same as the preceding day. Multiple random samples of 201 stocks each, from the set of 434 NYSE stocks, were used for comparison with the NSE data, and it was verified that the results obtained were independent of the particular sample of 201 stocks chosen.

Analysis Using Correlation Matrix

- To analyse the price movements of different stocks we calculate the logarithmic price return of a stock over a time delta t and divide it by the volatility to get the normalized return.
- Then a correlation matrix C is made containing C(I,j) as the correlation between the normalized returns i,jth stocks.

$R_i(t, \Delta t) \equiv \ln P_i(t + \Delta t) - \ln P_i(t).$

$r_i(t,\Delta t) \equiv (R_i - \langle R_i \rangle)/\sigma_i$

$\sigma_i \equiv \sqrt{\langle R_i^2 \rangle - \langle R_i \rangle^2}$, is the standard deviation of R_i

$C_{ij} \equiv \langle r_i r_j \rangle,$

Results

- By construction C is symmetric with C(I,i)=1 and C(I,j) between [-1,1]
- Upon Plotting the probability density function of the elements in correlation matrix C we observe that the correlation among stocks in NSE is larger on the average compared to that amongst the stocks in NYSE.
- This study supports the belief that developing markets tend to be more correlated than developed ones.

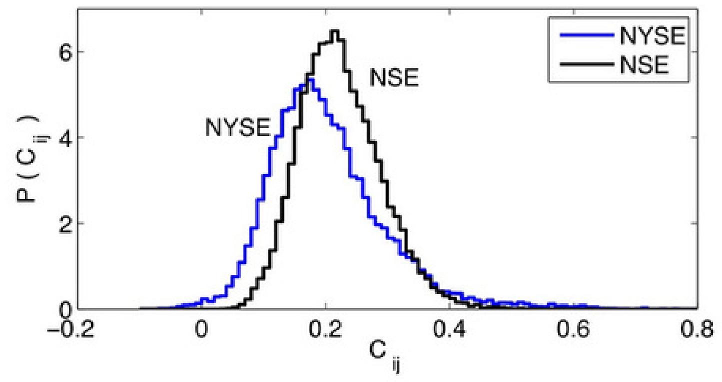
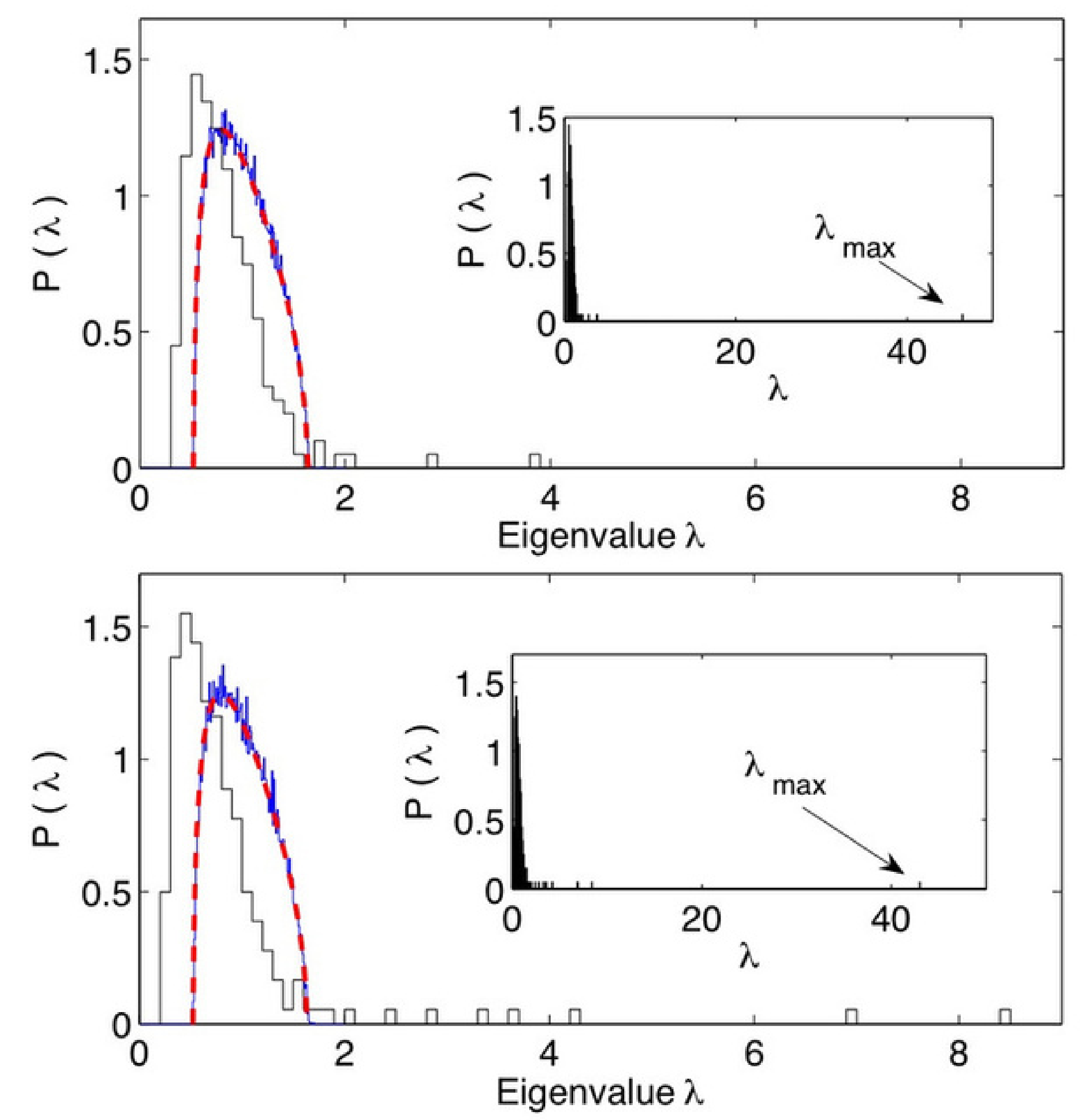


FIG. 1: The probability density function of the elements of the correlation matrix **C** for 201 stocks in the NSE of India and NYSE for the period Jan 1996-May 2006. The mean value of elements of **C** for NSE and NYSE, $\langle C_{ij} \rangle$, are 0.22 and 0.20 respectively.

EigenValue Spectrum



$$P_{\rm rm}(\lambda) = \frac{Q}{2\pi}\frac{\sqrt{(\lambda_{max}-\lambda)(\lambda-\lambda_{min})}}{\lambda},$$

- Now to find reason behind this extra correlation,we do an eigenvalue spectrum analysis of the correlation matrix . We use eigenvalues as a measure because they help reduce the dimensionality of the data while retaining the important properties.
- So we make a random Correlation matrix(Wishart Matrix) based on statistical properties .and compare it with the correlation matrix(NSE Data) we generated.
- The NSE matrix we generate has N=201 stocks, T=2606 returns so Q=12.97 and the range of values without correlation are λ_{max} = 1.63, λ_{min} =0.5
- Upon observing the correlation matrices we see that bulk of the eigen value spectrum for NSE matrix is in agreement with the random matrix except for certain large eigenvalues.
- To be sure that these outliers are not noise we randomly shift the return series and then again calculated the eigenvalue distribution.
- We observe that the results are exactly same indicating these values are not noise but genuine indicators among the stocks. So by understanding these indicators we can get to know the striucture of the interactions between the stocks in the market.

Properties of Deviating EigenValues

- The largest eigenvalue in NSE cross-correlation matrix is 28 times and NYSE is 26 times greater than random matrix upper bound.These values also remain almost same for synthetic data conatining same no of missing points.

- Also another interesting feature of this of this eigenvaue is that the corresponding eigenvector shows a relatively uniform composition with almost all stocks contributing to it and with same sign. This shows there is a common factor which affects all the stocks with the same bias and this is called the market mode (collective response of market to external info)

- To understand more about the market we see intermediate eigenvalues and analyse them. The eigenvectors of these corresponding eigenvalues are localised in NYSE ie. only a small no. of stocks contribute to them. Whereas in NSE such a straighforward division into groups is possible.

- To do such analysis on a more deeper level we use IPR(Inverse Participation Ratio) which is defined as summation [u(k,i)]^4 where u(k,i)are components of eigenvector k. IPR gives us the reciprocal of the of the no. of eigenvector components with significant contribution.

- The values of IPR at lower and higher ends show deviation for both NSE,NYSE indicating the presence of localized modes. However these deviations are much less significant in NSE as compared to NYSE.

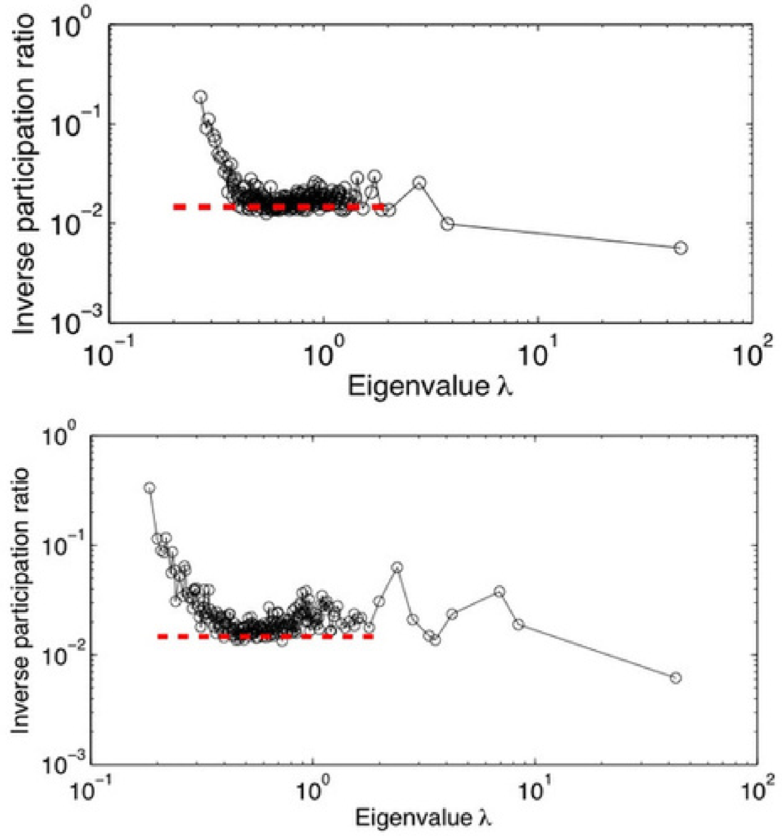


FIG. 4: Inverse participation ratio as a function of eigenvalue for the correlation matrix **C** of NSE (top) and NYSE (bottom). The broken line indicates the average value of $\langle I \rangle = 1.49 \times 10^{-2}$ for the eigenvectors of a matrix constructed by randomly shuffling each of the N time series.

This implies that distinct groups whose members are mutually correlated in their price movement do exist in NYSE, but their existence is far less in NSE.

Filtering the Correlation Matrix

- The previous analysis shows that there is a market factor which influences all stocks and makes it difficult to analyse individual correlations.So we use a filtering method to seperate out the noise and the market mode.
- We write the entries of the correlation matrix as the product of eigenvalues ,eigenvectors and their transposes. Also the correlation matrix is written as a combination of three factors market,group and random
- The results of this decomposition is generated and studied for both Indian and US markets and it is observed that the group component shows a significant truncated tail in former as compared to latter.
- This indicates that intra group correlations are not prominent in NSE while they are comparable with market correlations in NYSE.
- So collective behaviour in Indian stocks is dominated by external information and hence,correlaton between sectors is difficult to observe.

$$\mathbf{C} = \sum_{i=0}^{N-1} \lambda_i \mathbf{u}_i \mathbf{u}_i^T,$$

$$\begin{aligned}\mathbf{C} &= \mathbf{C}^{market} + \mathbf{C}^{group} + \mathbf{C}^{random} \\ &= \lambda_0 \mathbf{u}_0 \mathbf{u}_0^T + \sum_{i=1}^{N_g} \lambda_i \mathbf{u}_i \mathbf{u}_i^T + \sum_{i=N_g+1}^{N-1} \lambda_i \mathbf{u}_i \mathbf{u}_i^T, \quad (6)\end{aligned}$$

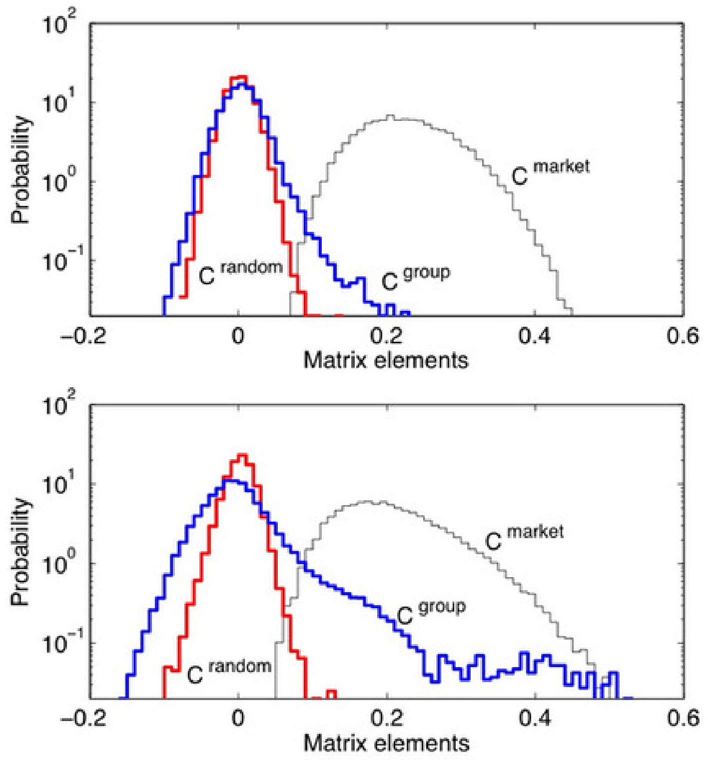


FIG. 5: The distribution of elements of correlation matrix corresponding to the market, $\mathbf{C}^{market}$, the group, $\mathbf{C}^{group}$, and the random interaction, $\mathbf{C}^{random}$. For NSE (top) $N_g = 5$ whereas for NYSE (bottom) $N_g = 10$. The short tail for the distribution of the $\mathbf{C}^{group}$ elements in NSE indicates that the correlation generated by mutual interaction among stocks is relatively weak.

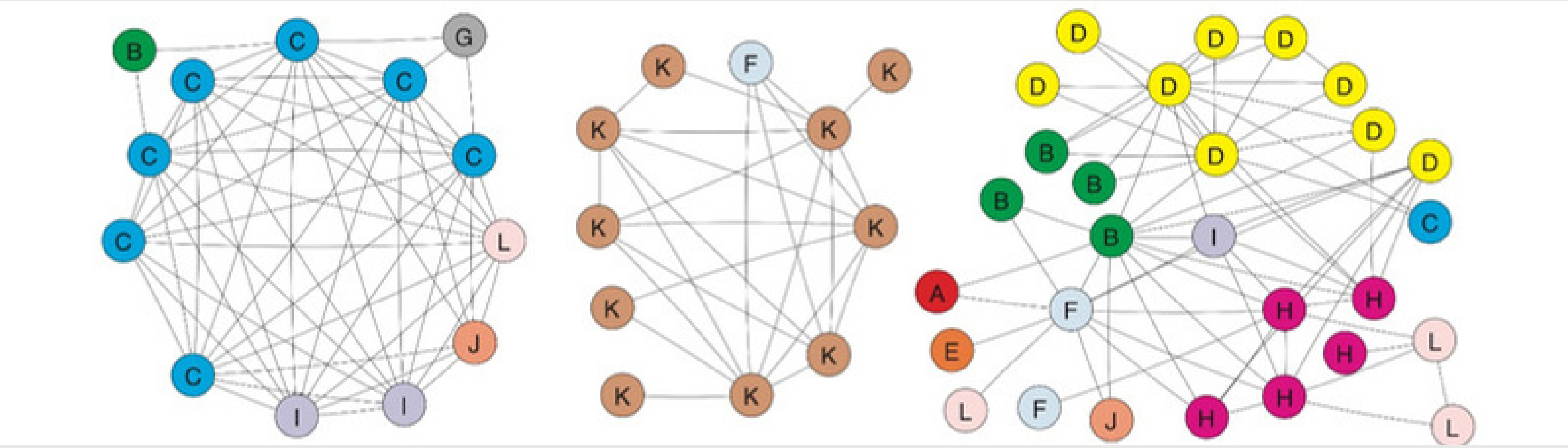


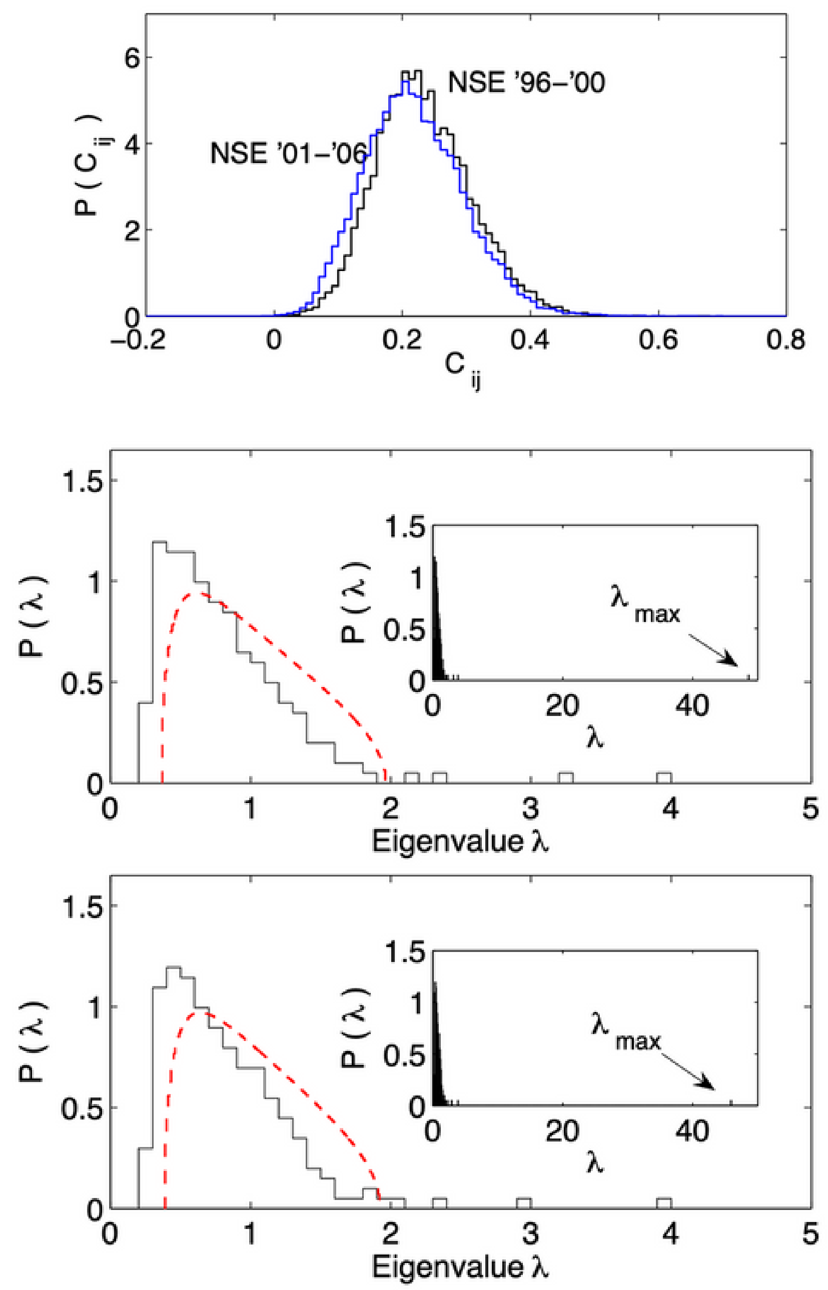
FIG. 7: The structure of interaction network in the Indian financial market at threshold $c^* = 0.09$. The left cluster comprises of mostly Technology stocks, while the middle cluster is composed almost entirely of Healthcare & Pharmaceutical stocks. By contrast, the cluster on the right is not dominated by any particular sector. The node labels indicate the business sector to which a stock belongs and are as specified in the caption to Fig 3.

- We now use the information present in the correlation matrix to create the networks graphs shown above.
- To do this we create an adjacency matrix based on a threshold that A(I,j)=1 if C(I,j) group >cth and A(I,j)=0 otherwise. We choose different values of cth and analyse the network graph for the same.
- We find that for all values of cth the no of clusters is less in NSE as compared to NYSE
- So the fact that majority of the NSE cannot be organised into well segregated groups illustrates that **intra-group interaction is much weaker than market-wide correlation in India.**

In the above example the network has 52 nodes,298 links partitioned into 3 isolated clusters, but only two business sectors can be propery identified ie.(Technology and Pharmaceutical sector

Relating correlation with market evolution

- On comparing 2 different time periods, Jan 1996-Dec 2000 and between Jan 2001-May 2006, we see that the average value for the elements of the correlation matrix is slightly lower for the later period, suggesting a greater homogeneity between the stocks at the earlier period.
- Looking at the eigenvalue distribution for C in the 2 periods, the Q value being aprroximately equal, the largest deviating eigenvalue for Period I is higher than Period II, again suggesting that over time market has become less homogeneous.
- When the interaction networks between stocks are gen- erated for the two periods, they show less distinction into clearly defined sectors than was obtained with the data for the entire period.
- On further analysis, it was revealed that the Indian market is evolving, the inter- actions between stocks are tending to get arranged into clearly identifiable groups.
- Thus, we need a structural rearrangement or new market model.



MODEL OF MARKET DYNAMICS

$$r_i^k(t) = \beta_i r_m(t) + \gamma_i^k r_g^k(t) + \sigma_i \eta_i(t)$$

We split the normalized returns if i-th stock in k-th sector into ‘market factor’, ‘sector factor’ and ‘idiosyncratic term’ (ie. random variations unique for that stock) with β, γ and σ being the strengths of each term.

Assuming unit variance for the returns, we get $(\beta_i)^2+(\gamma_i^k)^2+(\sigma_i)^2 = 1$. As a result we can get beta values by fixing the other two.

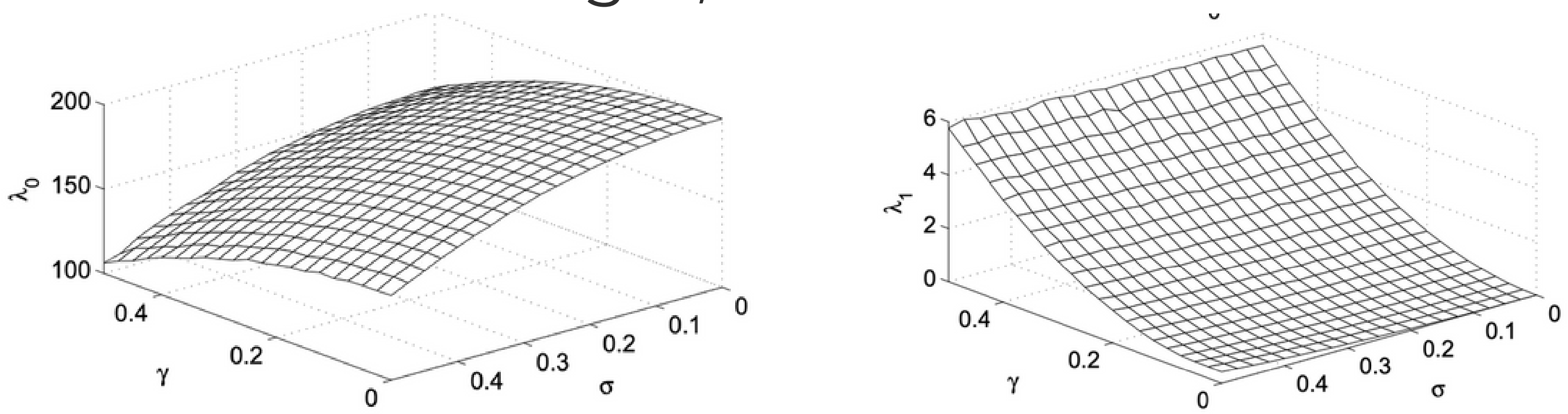
We choose σ_i and γ_i from a uniform distribution having width δ and centered about the mean values σ and γ, respectively.

Artificial Market Simulation

- The artificial market simulation involves N stocks belonging to K sectors. Sectors are composed of different numbers of stocks (n1, n2, n3, ..., nk) such that n1+n2+...+nK=N.
- We analyse the collective behaviour constructing the correlation matrix C and its eigenvalues. For simplicity, we neglect *βi*'s and take sector factor strength equal for all stocks within a sector, the spectrum of the correlation matrix is composed of K large eigenvalues and N-K small eigenvalues.
- When a nonzero market factor (*βi*=*β*) is considered, along with equal sector factor strength for all stocks (*γik*=*γ*), similar patterns are observed in terms of large and small eigenvalues.

$$\begin{aligned}\lambda_0 &\sim N\beta^2,\\ \lambda_1 &\sim n_l(1-\beta^2)\end{aligned}$$

We now choose the strength γ k i and σi from a uniform distribution with mean γ and σ respectively and with width δ = 0.05. We get,



What does this artificial market model tell us?

A decrease in the strength of the sector factor relative to the market factor results in a decrease in the second-largest eigenvalue ($\lambda 1$).

As $Q = T/N$ is fixed, the RMT bounds for the bulk of the eigenvalue distribution, [λmin, λmax], remain unchanged.

A reduction in $\lambda 1$ indicates that intermediate eigenvalues align more closely with the spectrum predicted by RMT, as observed in the NSE case.

The model analysis substantiates our hypothesis: the correlation matrix's spectral properties in the NSE reflect a market where the dominance of shared information (the market mode) leads to a notable correlation among all stocks.

CONCLUSION

Our study reveals heightened stock correlations in emerging markets compared to developed ones, with a similar eigenvalue spectrum. Fewer deviations from RMT upper bound occurs, indicating the dominance of common market effects over sector interactions. This dominance of the market mode relative to modes arising through interactions between stocks makes an emerging market appear more correlated than developed markets.

A two-factor model confirms a robust market factor's alignment with empirical observations in the Indian market. This insight contributes to understanding market evolution, suggesting strong interactions may emerge over time within stock groups.

The study challenges assumptions about intra-market interactions, emphasizing the ‘simple’ mean field-like description for certain properties in complex financial systems. This prompts further exploration into the relationship between self-organization, market development, and external signals in econophysics research.

Thank You