The Power (Law) of Indian Markets: Analysing NSE and BSE Trading Statistics

Sitabhra Sinha and Raj Kumar Pan

The Institute of Mathematical Sciences, C. I. T. Campus, Taramani, Chennai - 600 113, India. sitabhra@imsc.res.in

The nature of fluctuations in the Indian financial market is analyzed in this paper. We have looked at the price returns of individual stocks, with tick-by-tick data from the National Stock Exchange (NSE) and daily closing price data from both NSE and the Bombay Stock Exchange (BSE), the two largest exchanges in India. We find that the price returns in Indian markets follow a fat-tailed cumulative distribution, consistent with a power law having exponent $\alpha \sim 3$, similar to that observed in developed markets. However, the distributions of trading volume and the number of trades have a different nature than that seen in the New York Stock Exchange (NYSE). Further, the price movement of different stocks are highly correlated in Indian markets.

1 Introduction

Over the past decade, a growing number of physicists have got involved in searching for statistical regularities in the behavior of financial markets. A key motivation for such "econophysicists" is the prospect of discovering *universal* features in financial data [1], i.e., statistical properties that are invariant with respect to stocks, markets, the time interval over which the data is collected, etc. The most prominent candidate for such universality is the distribution of fluctuations in the price of individual stocks [2, 3], as well as, market indices [4] which reflect the composite value of many such stocks. Studies in various markets have reported evidence for the cumulative distribution of price fluctuations having positive and negative tails that obey a power law decay, i.e., $P_c(x) \sim x^{-\alpha}$. It has also been claimed that the exponent for this power law, α, is around 3 for most markets (the "inverse cubic law") [5]. It may be useful here to distinguish between the power law reported for individual stock price fluctuations and that for market index fluctuations, as the former is more fundamental and implies the latter, provided most of the stocks comprising the index have significant cross-correlation in their price movement.

We will, therefore, focus on the behavior of individual stocks, although we will also mention in brief our study of a particular Indian market index.

The prime motivation for our study of the Indian financial market is to check recent claims that emerging markets (including those in India) have behavior that departs significantly from the previously mentioned "universal" behavior for developed markets. Although a recent paper [6] reported heavy tailed behavior of the fluctuation distribution for an Indian market index between Nov 1994 and Oct 2004, the generalized Pareto distribution fit to the data did not suggest a power law decay of the tails. Moreover, an almost contemporaneous study [7] of the fluctuations in the price of 49 largest stocks in the NSE between Nov 1994 and Jun 2002, has claimed that the distribution has exponentially decaying tails. This implies the presence of a characteristic scale, and therefore, the breakdown of universality of the power law tail for the price fluctuation distribution. The contradiction between the results of the two groups indicates that a careful analysis of the Indian market is necessary to come to a conclusive decision. Note that, both of the above-mentioned studies looked at low-resolution data, namely, the daily closing time series.

In this study, we have looked at the high-frequency transaction by transaction stock price data, as well as taken a fresh look at the low-frequency daily data. We conclude that, far from being different, the distribution of price fluctuations in Indian markets is *quantitatively* almost identical to that of developed markets. However, the distributions for trading volume and number of trades seem to be market-specific, with the Indian data being consistent with a log-normal distribution for both of these quantities. Next, we look at the distribution of fluctuations in the NSE market index and find it to also follow the "inverse cubic law". Given the result for the price fluctuation distribution of individual stocks, this is expected if the price movements of the various stocks are highly correlated. Therefore, we also study the cross-correlations among the price fluctuations of most of the stocks comprising the index. We find that, on the whole, stock price movements in the Indian market are remarkably correlated.

2 The Indian financial market

There are 23 different stock markets in India. The two largest are the National Stock Exchange (NSE) and the Bombay Stock Exchange (BSE) which together accounted for more than 98% of the total turnover for all markets in 2003-04 [8]. Of these, the NSE is the larger one, with a turnover that is slightly more than double that of BSE, although their market capitalizations are comparable. BSE was founded in 1875, and is the oldest stock market in Asia. It has the largest number of companies listed and traded, among all the exchanges in India. The market indices associated with it, namely BSE 30, BSE 100 and BSE 500, are closely followed indicators of the health of the Indian financial market. The stocks belonging to BSE 500 represent nearly

93% of the total market capitalisation in that exchange, and therefore in this study we have confined ourselves to these stocks.

Compared to BSE, NSE is considerably younger, having commenced operations in the capital (equities) market from Nov 1994. However, as of 2004, it is already the world's third largest stock exchange (after NASDAQ and NYSE) in terms of transactions [8]. The most important market index associated with the NSE is the Nifty. The 50 stocks comprising the Nifty index represent about 58% of the total market capitalization and 47% of the traded value of all stocks in the NSE (as of Dec 2005).

Description of the data set. The low-frequency data that we analyze consists of the daily closing price, volume and number of trades for individual stocks from BSE (starting from as early as 1991) and NSE (starting from as early as 1994). This data is available from the web-sites of the corresponding exchanges [9]. The high-frequency tick-by-tick data contains information of all transactions carried out in the NSE between Jan 1, 2003 and Mar 31, 2004. This information includes the date and time of trade, the price of the stock during transaction and the volume of shares traded. This database is available in the form of CDs published by NSE. For calculating the price return, we have focused on 479 stocks, which were all used to calculate the BSE 500 index during this period. To calculate the distribution of index fluctuations, we have looked at the daily closing value of Nifty between Jan 1, 1995 and Dec 31, 2005. For cross-correlation analysis, we have focused on daily closing price data of 45 NSE stocks (all belonging to the Nifty index) from Jan 1, 1997 to Dec 31, 2005.

3 Price return distribution of individual stocks

To measure the price fluctuations (or the fluctuations in the market index) such that the result is independent of the scale of measurement, we calculate the logarithmic return of price (or index). If $P_i(t)$ is the stock price of the ith stock at time t, then the (logarithmic) price return is defined as

$$R_i(t, \Delta t) \equiv \ln P_i(t + \Delta t) - \ln P_i(t). \tag{1}$$

However, the distribution of price returns of different stocks may have different widths, owing to differences in their volatility. To be able to compare the distribution of various stocks, we must normalize the returns by dividing them with their standard deviation (which is a measure of the volatility), $\sigma_i = \sqrt{\langle R_i^2 \rangle - \langle R_i \rangle^2}$. The normalized price return is, therefore, given by

$$r_i(t, \Delta t) \equiv \frac{R_i - \langle R_i \rangle}{\sigma_i}, \tag{2}$$

where $\langle \dots \rangle$ represents time average.

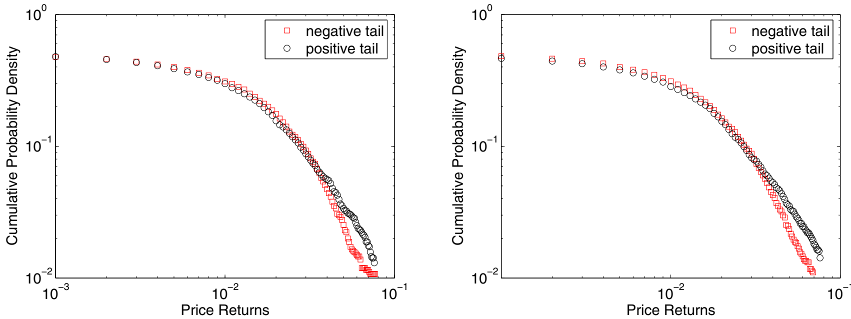

Fig. 1. Cumulative distribution of daily price returns for a particular stock (Reliance) at BSE (left) between July 12, 1995 and Jan 31, 2006, and at NSE (right) between Nov 3, 1994 and Jan 30, 2006.

3.1 Daily price returns in BSE and NSE

We start by focussing on the daily price variation of individual stocks, i.e., $\Delta t = 1$ day. By using the time series of daily closing price of a particular stock (e.g., Reliance) we can obtain the corresponding daily returns. Binning this data appropriately we can obtain the probability density function, and by integrating it over a suitable range, the cumulative distribution function (CDF), which is essentially the probability that a return is larger than a given value. Fig. 1 shows the CDF for daily price returns for the same stock in BSE (left) and NSE (right). Note that, we have shown the tails for the positive and negative returns in the same figure. The distribution for the two exchanges are almost identical, and both show long tails consistent with a power law decay.

To confirm that this is a general property, and not unique to the particular stock that is being analysed, we next perform the same analysis for other stocks. To be able to compare between stocks, we normalize the returns for each stock by their standard deviation. Fig. 2 (left) shows that four stocks chosen from different sectors have very similar normalized cumulative distributions. Moreover, the tail of each of these distributions approximately follow a power law with exponent $\alpha \simeq 3$. However, the daily closing price data set for any particular stock that we have analyzed is not large enough for an unambiguous determination of the nature of the tail. For this, we aggregate the data for 43 frequently traded stocks, all of which are used for calculating the Nifty index, over 3 years, and obtain the corresponding CDF (Fig. 2, right). Putting together the time series of different stocks to form a single large time series is justified because, after normalization, the different stocks have almost identical distributions [3]. From this figure we confirm that the distribution does indeed follow a power law decay, albeit with different exponents for the positive and negative return tails. The different exponents of the positive and negative tails have also been observed in the case of stocks listed in the New

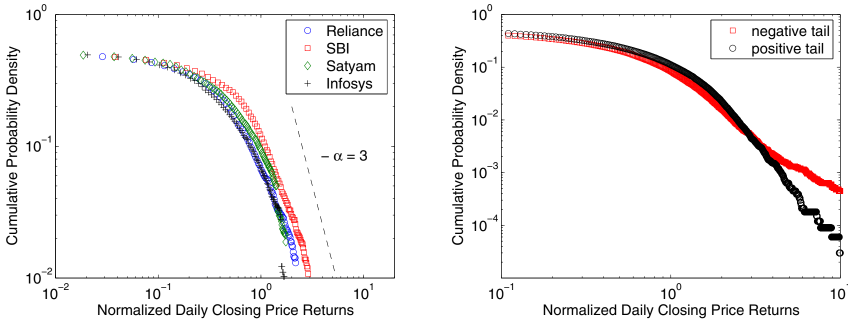

Fig. 2. (Left) Cumulative distribution of the normalized daily closing price returns for four stocks in BSE between July 12, 1995 and Jan 31, 2006. (Right) Cumulative distribution of the aggregated normalized daily closing price returns for 43 stocks (included in the Nifty index) at NSE between Jan 1, 2003 and Jan 30, 2006.

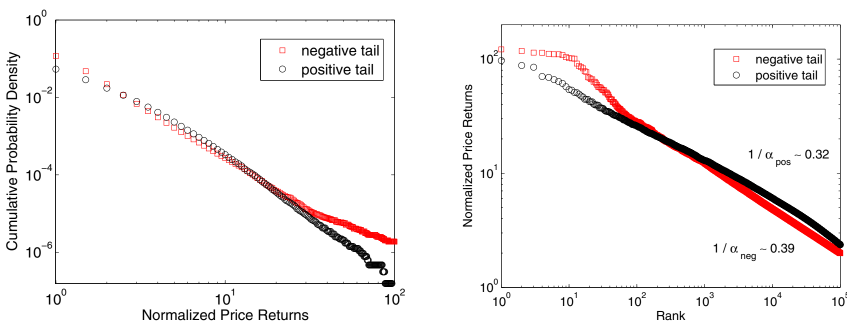

Fig. 3. Cumulative distribution (left) and rank-ordered plot (right) for the 5-minute interval price returns aggregated for 479 stocks at NSE between Jan 1, 2003 to Mar 31, 2004.

York Stock Exchange (NYSE) [3]. For comparison, we carried out a similar study with the daily closing price series of several NYSE stocks from Yahoo! Finance [10], and obtained distributions qualitatively similar to that shown here for the Indian market.

3.2 Price return for tick-by-tick data in NSE

The daily data is strongly suggestive of a power law tail for the price return distribution, but for conclusive evidence we next turn to the tick-by-tick data for stocks listed in the NSE. Choosing an appropriate Δt, we obtain the corresponding return by taking the log ratio of consecutive average prices, averaged over a time window of length Δt; for the results reported here $\Delta t = 5$ minutes. We have verified that the nature of the distribution is not sensitive to the exact value of this parameter. For individual stocks, the cumulative distribution of returns again show power law decay, but as the data set for each stock is not large enough, we carry out an aggregation procedure similar to that outlined above. Picking 479 frequently traded stocks from NSE, we put together their normalized returns to form a single large data set. The corresponding CDF is shown in Fig. 3 (left), with the exponents for the positive and negative tails estimated to be $\alpha \sim 3.2$ and 2.7, respectively. To check the accuracy of these exponents, obtained using linear least square fitting on a doubly logarithmic plot, we next plot the return data in descending order. This *rank-ordered plot* is an alternative visualization of the CDF, interchanging the ordinate and abscissae. It is easy to show that if the CDF has a power-law form, so does the rank-ordered plot, and the two exponents are the inverses of each other [11]. Exponents obtained by least square fitting on this graph produces similar values of α, namely, 3.1 and 2.6 for the positive and negative tails, respectively.

3.3 The "inverse cubic law" for price and index fluctuations

The results reported above provide conclusive evidence that the Indian financial market follows a price fluctuation distribution with long tails described by a power law. Moreover, the exponent characterizing this power law is close to 3, as has been observed for several financial markets of developed economies, most notably the NYSE, where the "inverse cubic law" has been found to be valid from $\Delta t = 1$ day to 1 month.

Most observations of this "law" have been in the context of market indices, rather than the price of individual stocks. We have, therefore, carried out a similar analysis for the Nifty index of NSE during the period Jan 1, 1995 to Dec 31, 2005. Fig. 4 (left) shows that the distribution of index returns also shows a power law decay, with an exponent very close to 3. As the index is a

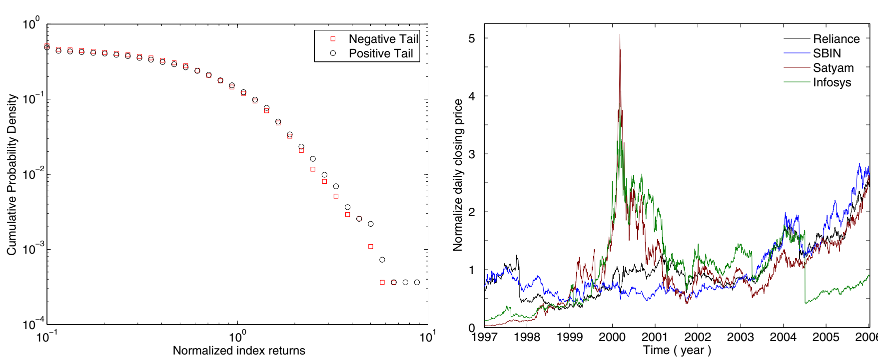

Fig. 4. (Left) Cumulative distribution of daily returns for the Nifty index between Jan 1, 1995 and Dec 31, 2005. (Right) Comparison of the daily closing price for four stocks in NSE from Jan 1, 1997 to Dec 31, 2005, showing the high degree of correlation among the stocks.

composite of several stocks, this behavior can be understood as a consequence of the power law decay for the tails of individual stock price returns, provided the movement of these stocks are correlated. As is evident from Fig 4 (right), this condition is indeed satisfied in the Indian market. In a later section we provide a more detailed look into the cross-correlation structure of these price fluctuations.

These findings assume importance in view of the recent claims that emerging markets behave very differently from developed markets, in particular, exhibiting an exponentially decaying return distribution [7]. India is one of the largest emerging markets, and our analysis of the price fluctuations in the major Indian stock exchanges challenges these claims, while at the same time, providing strong support to the universality for the "inverse cubic law" which had previously only been seen in developed markets.

4 Distribution of trading volume and number of trades

Besides the price of stocks, one can also measure market activity by looking at the trading volume (the number of shares traded), $V(t)$, and the number of trades, $N(t)$. To obtain the corresponding distributions, we normalize these variables by subtracting the mean and dividing by their standard deviation, such that, $v = \frac{V - \langle V \rangle}{\sqrt{\langle V^2 \rangle - \langle V \rangle^2}}$ and $n = \frac{N - \langle N \rangle}{\sqrt{\langle N^2 \rangle - \langle N \rangle^2}}$. Fig. 5 shows the distribution of these two quantities for several stocks, based on daily data for BSE. As is evident, the distribution is very similar for the different stocks, and the nature of the decay is significantly different from a power law. To better characterize the distribution, we have also looked at the intra-day distributions for volume and number of trades, based on high-frequency data from NSE. Fig. 6 shows the distributions of the two quantities for trading conducted on a particular stock in 5 minute intervals. Analysis of data for other stocks show qualitatively

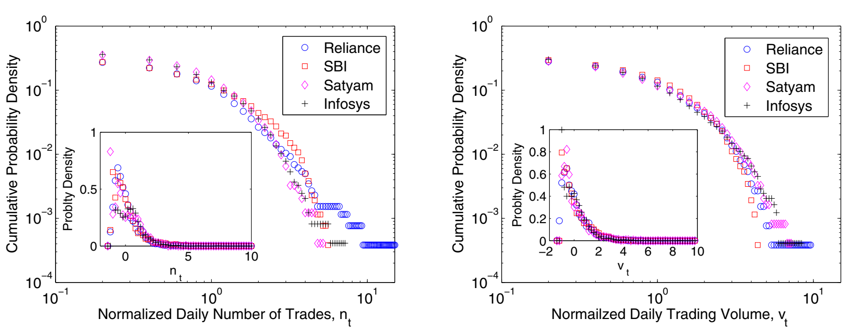

Fig. 5. Cumulative distribution of the number of trades (top left) and the volume of shares traded (top right) in a day for four stocks at BSE between July 12, 1995 and Jan 31, 2006.

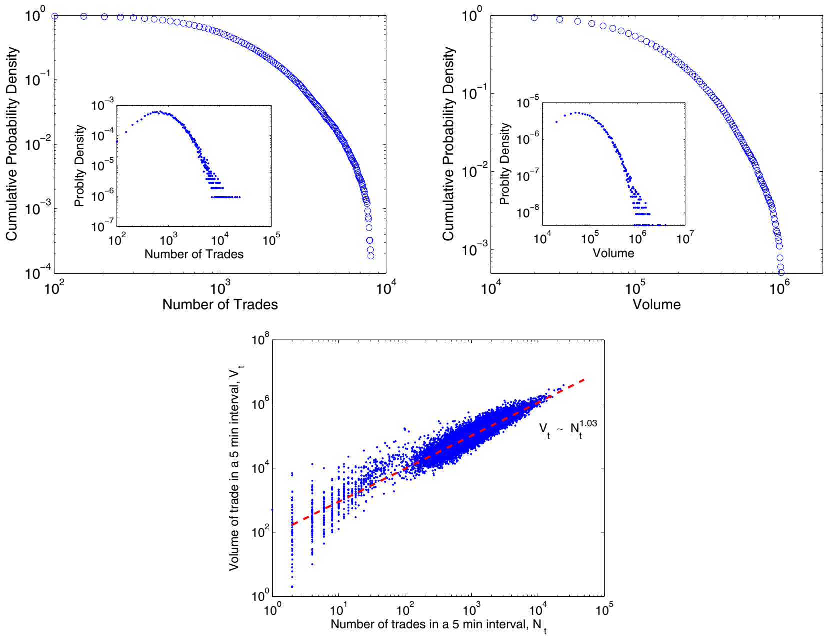

Fig. 6. Cumulative distribution of the number of trades (top left) and the volume of shares traded (top right) for a particular stock (Reliance) in 5-minute intervals at NSE between Jan 1, 2003 to March 31, 2004. The bottom figure shows an almost linear relation between the number of trades in a 5-minute interval and the corresponding trading volume. The broken line indicates the best fit on a doubly logarithmic scale.

similar results. As is clear, both of these distributions are non-monotonic, and are suggestive of a log-normal form. The fact that these distributions are very similar to each other is not surprising in view of the almost linear relationship between the two (Fig. 6, bottom). This supports previous observation in major US stock markets that statistical properties of the number of shares traded and the number of trades in a given time interval are closely related [13].

For US markets, power law tails have been reported for the distribution of both the number of trades [12] and the volume [13]. It has also been claimed that these features are observed on the Paris Bourse, and therefore, these features are as universal as the "inverse cubic law" for price returns distribution [14]. However, analysis of other markets, e.g., the London Stock Exchange [15] have failed to see any evidence of power law behavior. Our results confirm the latter assertion that the power law behavior in this case may not be universal, and the particular form of the distribution of these quantities may be market specific.

5 Correlated stock movement in the Indian market

As indicated in a previous section, we now return to look at the cross-correlation among price movements. The data that we analyze for this purpose consists of 2255 daily returns each for 45 stocks. We divide this data into M overlapping windows of width T, i.e., each window contains T daily returns. The displacement between two consecutive windows is given by the window step length parameter δt. In our study, T is taken as six months (125 trading days), while δt is taken to be one month (21 trading days). The correlation between returns for stocks i and j is calculated as

$$C_{ij} = \langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle, \tag{3}$$

where $\langle \ldots \rangle$ represents the time average within a window. The resulting correlation matrices, C, can be analysed to get further understanding of the relations between movements of the different stocks.

We now look at the eigenvalues of C which contain significant information about the cross-correlation structure [16]. Fig. 7 (left) shows the eigenvalues of C as a function of time. It is clear that the majority of these are very close to zero at all times. The largest eigenvalues contain almost all information about the market, which is evident from Fig. 7 (right). This shows the variation of the average correlation coefficient, as well as the largest eigenvalue λ_{max}, with time. The two are strongly correlated, indicating that λ_{max} captures the behavior of the entire market. Our results indicate that the Indian market is highly correlated, as indicated by the strong cross-correlations among the most traded stocks.

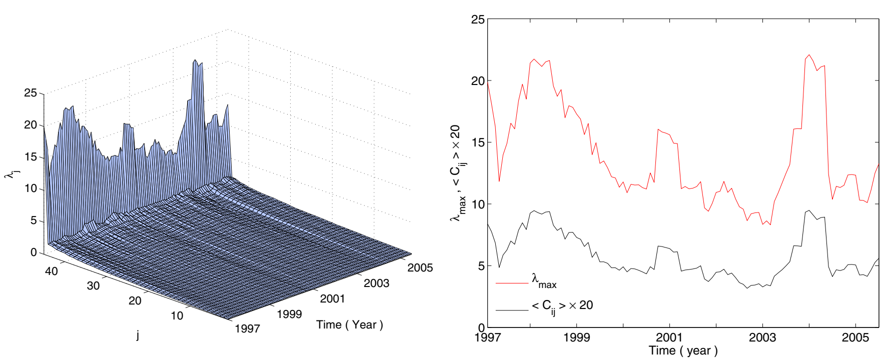

Fig. 7. (Left) The eigenvalues, sorted in descending order, for the correlation matrices of daily price returns for 45 stocks in NSE, across time. (Right) The variation across time of the largest eigenvalue λ_{max} of the correlation matrices and the average correlation $\langle C \rangle$. The window size chosen for calculating correlation is 6 months and the window is shifted in steps of 21 days. The total period is from Jan 1997 to Dec 2005.

6 Conclusions

In this paper, we have examined the statistical properties of trading in the two largest Indian financial markets, BSE and NSE. Using both low-frequency (daily) and high-frequency (tick-by-tick), we demonstrate that the price return cumulative distribution has long tails, consistent with a power law having exponent close to 3. This lends strong support to the claim that the price return distribution has an universal form across different markets, namely, the "inverse cubic law". On the other hand, the distributions for volume and number of trades appear to be log-normal, the two quantities being almost linearly related. We also find the market index fluctuation distribution to have the same form as the distribution of individual stock price returns. This implies that stocks in the Indian market are highly correlated. We verify that this is indeed the case with a cross-correlation analysis of most of the frequently traded stocks in the Indian market.

Acknowledgements

We are grateful to M. Krishna for invaluable assistance in obtaining and analyzing the high-frequency NSE data. We thank S. Sridhar and N. Vishwanathan for technical assistance in arranging the data, and J.-P. Onnela for helpful discussions.

References

1. Farmer JD, Shubik M, Smith E (2005) Physics Today 58(9): 37–42
2. Lux T (1996) Applied Financial Economics 6: 463–475
3. Plerou V, Gopikrishnan P, Amaral LAN, Meyer M, Stanley HE (1999) Phys. Rev. E 60:6519–6529
4. Gopikrishnan P, Plerou V, Amaral LAN, Meyer M, Stanley HE (1999) Phys. Rev. E 60: 5305–5316
5. Gopikrishnan P, Meyer M, Amaral LAN, Stanley HE (1998) Eur. Phys. J. B 3: 139–140
6. Sarma M (2005) EURANDOM Report 2005-003 (http://www.eurandom.tue.nl/reports/2005/003MSreport.pdf)
7. Matia K, Pal M, Salunkay H, Stanley HE (2004) Europhys. Lett. 66: 909–914
8. National Stock Exchange (2004) Indian securities market: A review. (http://www.nseindia.com/content/us/ismr2005.zip)
9. BSE: http://www.bseindia.com/, NSE: http://www.nseindia.com/
10. http://finance.yahoo.com/
11. Adamic LA, Huberman BA (2002) Glottometrics 3:143–150
12. Plerou V, Gopikrishnan P, Amaral LAN, Gabaix X, Stanley HE (2000) Phys. Rev. E 62: 3023–3026
13. Gopikrishnan P, Plerou V, Gabaix X, Stanley HE (2000) Phys. Rev. E 62: 4493–4496

14. Gabaix X, Gopikrishnan P, Plerou V, Stanley HE (2003) Nature 423: 267–270
15. Farmer JD, Lillo F (2004) Quantitative Finance 4: C7–C11
16. Plerou V, Gopikrishnan P, Rosenow B, Amaral LAN, Guhr T, Stanley HE (2002) Phys. Rev. E 65: 066126