ELSEVIER

Power tails of index distributions in chinese stock market

J.W. Zhang[a,b,c,d,*], Y. Zhang[a], H. Kleinert[d]

[a]*Department of Physics, Teachers College, Shihezi University, Xinjiang 832003, PR China*
[b]*School of Physics and MOE Key Laboratory of Heavy Ion Physics, Peking University, Beijing 100871, PR China*
[c]*MOE Key Laboratory of Quantum Information and Measurements, Peking University, Beijing 100871, PR China*
[d]*Institut für Theotische Physik, Freie Universität, Arnimallee 14, 14195 Berlin, Germany*

Abstract

The power α of the Lévy tails of stock market fluctuations discovered in recent years are generally believed to be universal. We show that for the Chinese stock market this is not true, the powers depending strongly on anomalous daily index changes short before market closure, and weakly on the opening data.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Power-law; Lévy distribution; Stock market index

The study of price fluctuations of financial assets has long been an active field of finance [1–5]. In recent years, the possibility to access and analyze huge amounts of high-frequency financial data with the help of contemporary computer technology has attracted the interest of theoretical physicists [6–13], and led to several significant empirical findings [14,15,18–21]. One of these is that the distribution of stock price fluctuations [22] is characterized by a power-law decay with an exponent $\alpha \simeq 3$, which lies well outside the Lévy-stable range $0 < \alpha < 2$ [23–27]. This is the so-called *inverse cubic law* of returns. The observed tails are absent in the previous widely used distributions [12], most prominently in the normal distribution proposed by Bachelier [1], which forms the basis of the Black–Scholes theory, in the truncated Lévy distribution of Mantegna and Stanley [15], in the Meixner distributions [16], or in the generalized hyperbolic distributions [17]. Although the tails can be fitted with the pure Lévy distribution as proposed by Mandelbrot [5], the distribution of the central most probably events cannot.

Remarkably the empirical tail behavior appears to be *universal*, because it holds for stocks of different economies, such as German stocks [23], US stocks [24,26], as well as various market indices such as the S&P 500, the Dow Jones, the NIKKEI, the Hang Seng, the Milan, and the DAX index [23,25,27]. In addition, the scaling behavior of the distribution has been analyzed for US stocks and market indices [25,26], which indicates that the tails of the distribution are well described by a power-law decay for time scales Δt from 1 min up to a certain value $(\Delta t)_x$ [28]. By the central limiting theorem the distributions converge, of course, to

*Corresponding author. School of Physics and MOE Key Laboratory of Heavy Ion Physics, Peking University, Beijing 100871, PR China.

E-mail address: james@pku.edu.cn (J.W. Zhang).

Gaussian behavior for Δt longer than $(\Delta t)_x$, albeit quite slowly. All these empirical findings suggest that there might be some universal underlying regularities governing the complex financial systems [29–32].

In order to explain the mechanism of the empirical power-law distribution, several theoretical models have been proposed [33–35]. For instance, Solomon and Richmond [33] build a multiagent system by the use of a generalized Lotka-Volterra model. A theory of Gabaix et al. [34] find this behavior on the basis of the economic optimization by heterogeneous agents.

Such empirical results and corresponding theories have suggested the universality of the power-law distribution of stock price fluctuations. The origin of this, however, is far from understood. When Huang [36] analyzed the 1-min data of the Hang Seng index in the Hang Kong stock market from January 1994 to December 1997, he found that the tail properties of the probability distribution of index fluctuations depend on the opening effect of each trading day. When skipping the data in the first 20 min of each morning session, the asymptotic behavior of the probability distribution shows an exponential-type decay as $P(x) \sim \exp(-\alpha|x|)/|x|$, where the index move $x = x(t)$ over a time scale Δt satisfies $x(t) = \text{index}(t) - \text{index}(t - \Delta t)$. Moreover, he claims that this empirical result can be derived from a Langevin approach [37]. In contrast, the case without any skip of the data is characterized by a power-law decay with an exponent $\alpha \simeq 3$.

Further doubts on the universality were raised by K. Matia et al. [38], who tested the daily returns from November 1994 to June 2002 for the 49 largest stocks of the National Stock Exchange, which has the highest volume of trade in India. These authors found an exponential probability density function of normalized returns g, to be defined below, as $P(g) \sim \exp(-\beta g)$, with the characteristic decay scales $\beta = 1.51 \pm 0.05$ for the negative tail and $\beta = 1.34 \pm 0.04$ for the positive tail. This led them to suggest that the power-law behavior merely holds for highly developed economies while the less highly developed ones follow a scale-dependent behavior.

On the other hand, most financial markets exhibit rich patterns caused by periodic market closures. For example, Cajueiro et al. [39] found that the intensity of the long-range dependence phenomena presented in this market depends on the time of the day that the phenomena is measured. This kind of pattern seems to be related to trading performed by different types of investors and the flow of information over the day.

The purpose of this paper is to test this hypothesis by analyzing the distributions of the Chinese market, which should be comparable to the Indian one. In particular, we consider the influence of the opening effect, and of the near-closure returns [40]. In addition, we study the scaling behavior of the distribution for time scales from 1 min up to 64 min.

We analyze the Shanghai Stock Exchange (SSE) index [41], which is a market-value weighted index. The data contain minute-by-minute records of every trading day from November 2000 to June 2004 [42], with the total number of data exceeding 2×10^5.

Define $S(t)$ as the value of SSE index and the return $G(t)$ over a time scale Δt as the forward change in the logarithm of $S(t)$, namely, $G(t) = \ln S(t + \Delta t) - S(t)$. Then we define the normalized return $g(t)$ as,

$$g(t) = \frac{G(t) - \langle G(t) \rangle_T}{\sigma}. \tag{1}$$

Here σ is the standard deviation of $G(t)$, and $\langle \cdots \rangle_T$ denotes an average over the entire length of the time series.

Fig. 1(a) and (b) display the cumulative distributions of 1-min returns of the SSE index for the positive and the negative tails, respectively. When the distributions are calculated with all data, which include near-closure returns, both positive and negative tails cannot be fitted well by the regression fits, especially for large values of g [43]. Consequently, in this case we obtain $\alpha \simeq 2.46$ in the region $1 \leqslant g \leqslant 8$, and $\alpha \simeq 1.52$ in the region $8 \leqslant g \leqslant 20$ for the positive tail. Likewise, we obtain $\alpha \simeq 2.64$ in the region $1 \leqslant g \leqslant 4$, and $\alpha \simeq 1.72$ in the region $4 \leqslant g \leqslant 20$ for the negative tail (see Eq. (2) for the definition of α). According to these results, there is definite evidence for a power-law asymptotic behavior described by Lévy distributions with $\alpha \approx 3$.

When the irregular near-closure returns are removed we still observe, for both positive and negative tails, a power-law asymptotic behavior

$$P(g > x) \sim x^{-\alpha}. \tag{2}$$

The powers of these distributions are completely consistent with the previous empirical power-law distributions found for different stock markets [23–27].
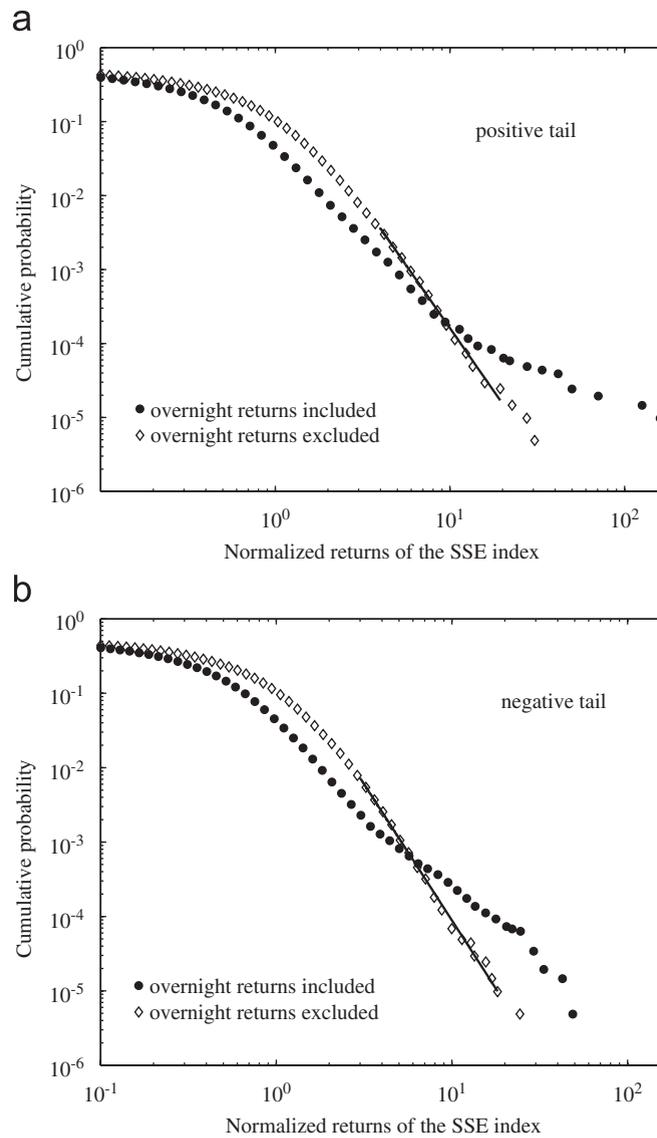
Fig. 1. Cumulative distribution of the 1-min returns of the SSE index for (a) the positive tail and (b) the negative tail. Each figure displays the distribution with (solid symbols) and without (open symbols) near-closure returns, respectively. Solid lines are the power-law regression fits in the region $3 \leqslant g \leqslant 20$ for the distributions without near-closure returns, which yield estimates $\alpha = 3.19 \pm 0.01$ for the positive tail and $\alpha = 3.67 \pm 0.01$ for the negative tail.

The power-law fit method [24–26] is applied to obtain the initial values of the exponents of power-law decay for the positive and negative returns, respectively. Similar to the way in Refs. [44,45], the Hill estimator [46] and its varieties [47,48] are further adapted for evaluating their standard errors. In fact, we use over 1% of the total data to fit the tail distributions. Specifically, in the appropriate region of g for the positive and the negative tails, the fits yield

$$\alpha = \begin{cases} 3.19 \pm 0.01, & 3 \leqslant g \leqslant 20 \quad \text{(positive tail)}, \\ 3.67 \pm 0.01, & 3 \leqslant g \leqslant 20 \quad \text{(negaive tail)}, \end{cases} \tag{3}$$

well outside the Lévy-stable range, $0 < \alpha < 2$. We observe a slight asymmetry between the 1-min positive and negative returns from the SSE index. Notwithstanding, it behaves inconsistently and indicates a significant

asymmetry in the distribution for the daily positive and negative returns of individual stocks from the "SSE" and "Shanghai Stock Exchange" [44]. It is found that the values of the exponents of the negative tails are larger than those of the positive tails for individual stocks.

The notable difference between the cases with and without near-closure returns shows that the tail parameter of the distribution is greatly affected by these irregular events. This is remarkable since the number of these events is relatively small (847) compared to the total number of returns (205215). In order to study their influence, as well as the data of returns for other certain time periods which might have a similar effect, we calculate the daily pattern $A(t_{day})$ [27]. This quantity is defined as the average absolute value of index returns:

$$A(t_{day}) = \frac{1}{N} \sum_{j=1}^{N} |G^j(t_{day})|. \tag{4}$$

The symbol t_{day} indicates the trading time within a day [49], and the index j runs over all trading days. We further find it useful to renormalize $A(t_{day})$ to a new quantity $a(t_{day})$ which has unit average. Generally speaking, if the value of $a(t_{day})$ for the time t_{day} is much larger than the values for other time periods, it may imply that the data of g for the time t_{day} contain more extreme values of the distribution, which gives them greater influence on the tail properties. Fig. 2 depicts the value of $a(t_{day})$ for each trading time within a day. Obviously, the near-closure value $a(t = 242)$ is much larger than the others (more than 10 times the average level). This is why the near-closure data have severe impact upon the tail parameter.

Apart from this, there are other values of $a(t_{day})$ which exceed the average level: those in the first, roughly 10 min within a trading day, near the end of the morning session, and those at the beginning of the afternoon session. For the tails of the g-distributions these are not significant, however. In fact, removing them causes only a small change of the tail parameter— the most significant change coming from the near-closure returns [50].

Furthermore, we study the distribution of the normalized returns (near-closure returns are removed) for longer time scales. In the region $3 \leqslant g \leqslant 20$, for positive tails, the power-law regression fits yield $\alpha = 3.63 \pm 0.14$, and $\alpha = 3.08 \pm 0.10$ for $\Delta t = 16$ and 64 min, respectively. Likewise, the corresponding values of α for negative tails are $\alpha = 3.71 \pm 0.08$ and $\alpha = 3.29 \pm 0.11$, respectively. It seems that the distribution maintains the power-law functional form for all these time scales. Note that here the values of exponents of both the positive and the negative tails are almost balanced, even if it appears a little bit larger bias for the values of the positive tails.
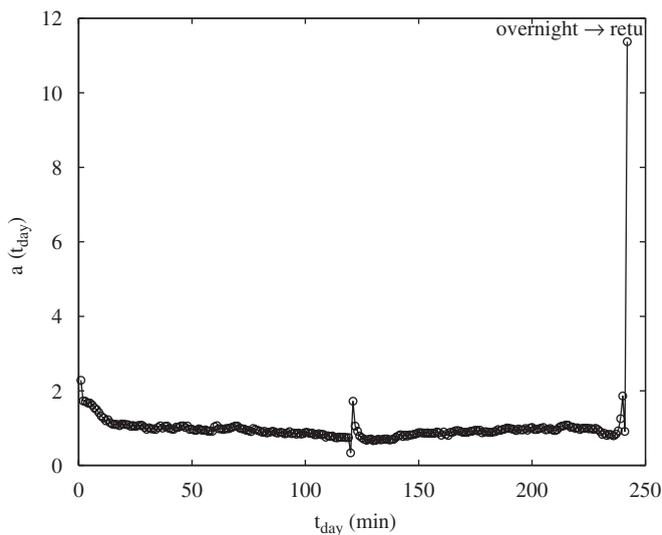

Fig. 2. Normalized 1-min interval daily pattern for the return of the SSE index, where t_{day} is the ordinal number of the trading time within a day, $a(t_{day})$ is the normalized daily pattern defined as the average of the absolute value of returns, $G^j(t_{day})$. Note that $G^j(242)$ which is calculated by near-closure returns largely exceed the average level.
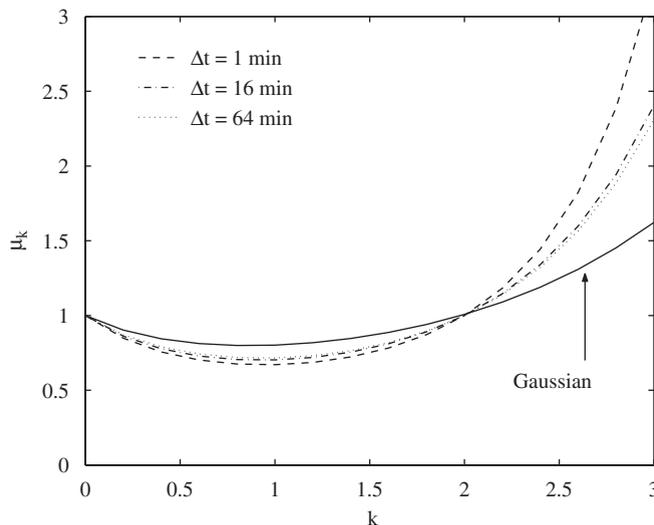
Fig. 3. The slow convergence to Gaussian behavior is shown by the moment μ_k with increasing $\Delta t = 1$, 16, and 64 min, respectively.

In the earlier study of the stock price returns of the US [26], near-closure returns were not removed, since these would merely bring small changes to the tail parameters. Thus the question arises why their influence is so much more relevant to the SSE index. This must be due to some specific characteristics of the Chinese stock market. One reason may be that compared to mature stock markets such as those in the US, the Chinese market depends more on government policies, and cases of insider trading are frequent. When some traders acquire the information of certain forthcoming government policies that might affect the stock markets, they act in last few minutes of the trading day, when it is too late for the others to react. Also, the supervision of excessive prices of market orders may not be as strict as in the mature markets. The smaller fluctuations in the morning are the overnight spillover of these activities.

To further verify the scaling behavior we analyze the moments of the distribution [25],

$$\mu_k = \langle |g|^k \rangle, \tag{5}$$

where $\langle \cdots \rangle$ denotes an average over all normalized returns, and μ_k is calculated for $0 \leqslant k \leqslant 3$. The moments demonstrate a slow convergence to Gaussian behavior with increasing Δt, as shown in Fig. 3.

In summary, we have analyzed the asymptotic behavior of the distribution of normalized 1-min returns for the SSE index. We find that the tail properties of the distribution are caused by the near-closure returns, the relevant fall-off powers being $\alpha \simeq 3$. Moreover, for time scales Δt from 1 min up to 64 min, the distribution shows a slow convergence to Gaussian behavior as required by the central limiting theorem. It is argued that to remove opening and closure returns in the study of high-frequency changes of the emerging stock markets qualifies as a standard empirical methodology and should be taken into account as a standard caveat in empirical research.

References

[1] L. Bachelier, Ann. Sci. École Norm. Suppl. 3 (1900) 21.
[2] M.J. Kendall, J. Roy. Stat. Soc. 96 (1953) 11.

[3] H.V. Roberts, J. Finance 14 (1959) 1.
[4] M.M. Osborne, Oper. Res. 7 (1959) 145.
[5] B.B. Mandelbrot, J. Business 36 (1963) 294.
[6] B.B. Mandelbrot, Fractals and Scaling in Finance, Springer, Berlin, 1997.
[7] J.P. Bouchaud, M. Potters, Theory of Financial Risks: From Statistical Physics to Risk Management, Cambridge University Press, Cambridge, 2000, 2003, The first and the second extension editions.
[8] I. Kondor, J. Kértesz, (Eds.), Econophysics: An Emerging Science, Kluwer, Dordrecht, 1999.
[9] R.N. Mantegna, H.E. Stanley, An Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, Cambridge, 1999.
[10] R.N. Mantegna, (Ed.), Proceedings of the International Workshop on Econophysics and Statistical Finance, Physica A 269 (1999), Special issue.
[11] D. Sornette, Why Stock Markets Crash: Critical Events in Complex Financial Systems, Princeton University Press, Princeton, 2003.
[12] H. Kleinert, Path integrals in quantum mechanics, statistics, polymer physics, and financial markets, World Scientific, Singapore, 2004, third extended edition. ⟨http://www.physik.fu-berlin.de/~kleinert/b5⟩.
[13] J.W. Zhang, C.S. Wang, A Glimpse of Econophysics, in: K.H. Zhao, K.C. Qin (Eds.), Physics Lightens the World, Peking University Press, Beijing, 2005.
[14] M.M. Dacorogna, U.A.M. Zuller, R.J. Nagler, R.B. Oslen, O.V. Pictet, J. Int. Money Finance 12 (1993) 413.
[15] R.N. Mantegna, H.E. Stanley, Nature 376 (1995) 46.
[16] B. Grigelionis, Lith. Math. J. 39 (1999) 33. See also www.maf.vu.lt/mif/mokslas/MAF00pre.html for a technical report of the author;
W. Schoutens, Meixner Processes in Finance, Report 2001-002, EURANDOM, Eindhoven. (www.eurandom.tue.nl/reports/2001/002wsreport.ps).
[17] J.P. Fouqué, G. Papanicolaou, K.R. Sircar, Derivatives in Financial Markets with Stochastic Volatility, Cambridge University Press, Cambridge, 2000;
J.P. Fouqué, G. Papanicolaou, K.R. Sircar, Int. J. Theor. Appl. Finance 3 (2000) 101.
[18] M.H.R. Stanley, L.A. Nunes Amaral, S.V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M.A. Salinger, H.E. Stanley, Nature 379 (1996) 804.
[19] D. Sornette, A. Johansen, J.P. Bouchaud, J. Phys. I (France) 6 (1996) 167.
[20] V. Plerou, L.A. Nunes Amaral, P. Gopikrishnan, M. Meyer, H.E. Stanley, Nature 400 (1999) 433.
[21] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A. Nunes Amaral, H.E. Stanley, Phys. Rev. Lett. 83 (1999) 1471.
[22] The price fluctuations are usually described by the variable "return", which is defined as the forward change in the logarithm of price. Some studies use the difference in the price instead, such as in Refs. [15,36]. However, both have the same statistics for events that are not too large [25].
[23] T. Lux, Appl. Fin. Econ. 6 (1996) 463.
[24] P. Gopikrishnan, M. Meyer, L.A. Nunes Amaral, H.E. Stanley, Eur. Phys. J. B 3 (1998) 139.
[25] P. Gopikrishnan, V. Plerou, L.A. Nunes Amaral, M. Meyer, H.E. Stanley, Phys. Rev. E 60 (1999) 5305.
[26] V. Plerou, P. Gopikrishnan, L.A. Nunes Amaral, M. Meyer, H.E. Stanley, Phys. Rev. E 60 (1999) 6519.
[27] B.H. Wang, P.M. Hui, Eur. Phys. J. B 20 (2001) 573.
[28] Here are two examples: $(\Delta t)_x \simeq 4$ days for returns of the S&P 500 index [25] and $(\Delta t)_x \simeq 16$ days for returns of stock prices of US individual companies [26].
[29] An empirical study (Ref. [32]) suggests that stock prices are in a different universality class other than spot prices for commodities.
[30] P. Gopikrishnan, V. Plerou, X. Gabaix, H.E. Stanley, Phys. Rev. E 62 (2000) R4493.
[31] V. Plerou, P. Gopikrishnan, L.A. Nunes Amaral, X. Gabaix, H.E. Stanley, Phys. Rev. E 62 (2000) R3023.
[32] K. Matia, L.A. Nunes Amaral, S.P. Goodwin, H.E. Stanley, Phys. Rev. E 66 (2002) 045103(R).
[33] S. Solomon, P. Richmond, Physica A 299 (2001) 188.
[34] X. Gabaix, P. Gopikrishnan, V. Plerou, H.E. Stanley, Nature 423 (2003) 267.
[35] M.G. Daniels, J.D. Farmer, L. Gillemot, G. Iori, E. Smith, Phys. Rev. Lett. 90 (2003) 108102.
[36] Z.F. Huang, Physica A 287 (2000) 405.
[37] L.H. Tang, Z.F. Huang, Physica A 288 (2000) 444.
[38] K. Matia, M. Pal, H.E. Stanley, H. Salunkay, cond-mat/0308013.
[39] D.O. Cajueiro, B.M. Tabak, N.A. Souza, Physica A 351 (2005) 512.
[40] On the American market, these are usually referred as overnight returns.
[41] The Shanghai Stock Exchange (SSE) is the first market of government-approved securities which was founded in Shanghai on November 26, 1990 and started operating on December 19 of the same year. The SSE index is usually considered as a benchmark of the Chinese equity performance.
[42] The data are provided by the Stock Star Corporation, ⟨http://www.stockstar.com⟩.
[43] The values of α are sensitive to the bounds of the region used for fitting [25]. The error caused by the missing digits that might affect the values of α has been checked and removed [26].
[44] C. Yan, J.W. Zhang, Y. Zhang, Y.N. Tang, Physica A 353 (2005) 425.
[45] Y. Zhang, M.Sc. Thesis, Peking University, 2005.
[46] B.M. Hill, Ann. Stat. 3 (1975) 1163.
[47] P. Hall, A.H. Welsh, Ann. Stat. 13 (1985) 331.

[48] P.C.B. Phillips, J. Appl. Econ. 11 (1996) 1.
[49] There are two time periods within a trading day in Chinese stock market. The morning session runs from 9:30 to 11:30, and the afternoon session runs from 13:00 to 15:00. Therefore, the length of a trading day is 242 minutes.
[50] For instance, the total number of t_{day}^i satisfying $a(t_{\text{day}}^i) > 1.2$ is 17. They occur during the first 11 minutes of a trading day. When all data of corresponding $g(t_{\text{day}}^i)$ are removed, the distribution of returns retains its power-law form. The tail exponents ($\alpha \simeq 3.52$ for the positive tail and $\alpha \simeq 4.34$ for the negative tail) are only a little larger than those caused by the overnight returns.