Statistics of extreme events in Chinese stock markets*

Wu Gan-Hua(吴干华)[a)b)], Qiu Lu(邱 路)[a)], Mutua Stephen[a)c)], Li Xin-Li(李信利)[a)d)], Yang Yue(杨 悦)[a)], Yang Hui-Jie(杨会杰)[a)†], and Jiang Yan(蒋 艳)[a)]

[a)] *Business School, University of Shanghai for Science and Technology, Shanghai* 200093, *China*
[b)] *School of Software, South China Normal University, Foshan* 528225, *China*
[c)] *Computer Science Department, Masinde Muliro University of Science and Technology, P. O. Box* 190-50100, *Kakamega, Kenya*
[d)] *Logistic School, Linyi University, Linyi* 276000, *China*

(Received 10 April 2014; revised manuscript received 10 June 2014; published online 10 October 2014)

We investigate the impact of financial factors on daily volume recurrent time intervals in the developing Chinese stock markets. The tails of probability distribution functions (PDFs) of volume recurrent intervals behave as a power-law, and the scaling exponent decreases with the increase of stock lifetime, which are similar to those in the US stock markets, and they are typical representatives of developed markets. The difference is that the power-law exponent values remain almost the same with the changes of market capitalization, mean volume, and mean trading value, respectively. These findings enrich the results for event statistics for financial markets.

Keywords: extreme statistics, recurrent time interval, volume volatility

PACS: 89.65.Gh, 05.45.Tp, 89.75.Da **DOI:** 10.1088/1674-1056/23/12/128901

1. Introduction

Extreme events in the financial field may bring cascade damages to the whole of human society, and consequently attract persistent attention of researchers in diverse fields.[1–5] Event statistics is a powerful tool to understand extreme events, in which statistical behaviors of intervals between successive events are used to shed light on mechanisms of event occurrences. In a time series, the occurrence of an event is usually defined as the record overtakes a threshold. Detailed works show that the probability distribution function (PDF) of an event occurrence in volatility series of stock closing prices in the US stock market and the Chinese stock market behaves like a power-law,[6–9] instead of the expected normal distribution. That is, extreme events may occur with unexpected high probabilities, and there does not exist a characteristic interval between successive occurrent intervals.

In a very recent paper, Li *et al*.[10] investigated impacts of financial factors on PDF of event occurrence. It is found that the PDF of recurrent intervals for volume series has scaling behavior. The scaling exponents decrease with an increase of stock lifetime, market capitalization, mean volume, and mean trading value, respectively. The US stock markets are representative for developed markets, while those in China are yet immature markets. It is well known that immature markets have completely different behaviors compared with developed markets.[11–15] A natural question is what will happen on these issues in Chinese markets?

In this paper, we conduct analogous investigations on the records of stock markets in China. We investigate the relation between financial factors and daily volume recurrent time intervals in Chinese stock markets. It is found that the tail of PDF of volume recurrent intervals behaves like a power-law, and the scaling exponent decreases with the increase of stock lifetime, which are similar to those in the US stock markets. While the difference is that the power-law exponent values stay almost the same with the changes of market capitalization, mean volume, and mean trading value, respectively. These findings enrich the results for event statistics for financial markets.

2. Methods and materials

2.1. Data

We consider the records in the Shanghai Stock Exchange (SHSE) market and the Shenzhen Stock Exchange market (SZSE), which were established in 1990. Daily volume and closing price data are collected from the stock integration system of the China Merchants Securities Co. Ltd and stock lifetime, market capitalization, mean volume from the financial research database of the Beijing Gildata RESSET Data Tech Co. Ltd. The period covers more than twenty years from December 1990 to November 2011, in which great changes have taken place in Chinese society. The considered data includes a total of 2282 stocks (4.5 million records).

From a daily volume series V, one can obtain the return

*Project supported by the National Natural Science Foundation of China (Grant No. 10975099), the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning, the Innovation Program of Shanghai Municipal Education Commission (Grant No. 13YZ072), the Shanghai Leading Discipline Project (Grant No. XTKX2012), and the Innovation Fund Project for Graduate Students of Shanghai (Grant No. JWCXSL1302).

†Corresponding author. E-mail: hjyang@ustc.edu.cn

© 2014 Chinese Physical Society and IOP Publishing Ltd

series R, an element of which reads,

$$R_t = \ln \frac{V_{t+1}}{V_t}. \tag{1}$$

In order to compare different stocks, we use the relative volatility series, an element of which is defined as,

$$\upsilon_t = \frac{|R(t)|}{\sqrt{\langle |R(t)|^2 \rangle - \langle |R(t)| \rangle^2}}, \tag{2}$$

where $\langle \cdot \rangle$ denotes the average over the series.

2.2. Recurrent interval and power-law distributions

A recurrent interval is the interval between successive events defined as the record overtakes a given threshold q, as illustrated in Fig. 1. If we set $q = 1.5$, the set for event occurrent time includes time 10, 17, and 37. The corresponding set of recurrent interval is $\tau(q = 1.5) = \{7,20\}$. When q is large enough, the PDF for the recurrent interval can be used to understand the statistical characteristics of extreme events.

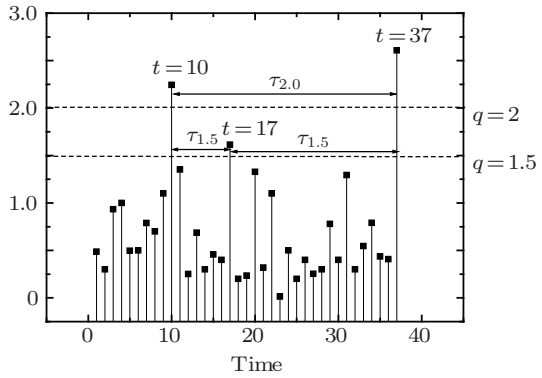


Fig. 1. Illustration of recurrent time series. Sets of recurrent intervals for two thresholds q=1.5 and 2.0, $\tau(1.5)$ and $\tau(2.0)$, are displayed.

For a given threshold, q, one can obtain the recurrent interval set for all the stocks, $\tau(q) \equiv \{\tau_1(q), \tau_2(q), \ldots, \tau_N(q)\}$. $\tau_i(q) \equiv \{\tau_i^1(q), \tau_i^2(q), \ldots, \tau_i^{L(i)}(q)\}$ is the recurrent interval set for the i-th stock. N is the total number of stocks and $L(i)$ is the number of elements in $\tau_i(q)$. Statistics of $\tau(q)$ represent the ensemble feature of the considered markets. In a large amount of complicated systems, the PDF for recurrent intervals is normally distributed with a characteristic interval, but in some records, the PDF turns out to behave like a power-law, which reads,

$$P_q(\tau) \sim \tau^{-\gamma}, \tag{3}$$

where γ represents the power-law exponent. In this kind of system, extreme events can occur with high probability and a characteristic interval between events does not exist. Numerically, the maximum likelihood-based estimator (MLE) proposed in the novel paper[16] is used to obtain reliable values of γ from the finite sets of empirical data.

In calculations, to reserve information in the original time series as far as possible, the dividing procedure about the recurrent interval is conducted also by starting from the end of series. The variance is obtained from a total of $2 \cdot [T/s]$ segments.

3. Results

3.1. Distribution of volume recurrent intervals

Firstly, we investigate the behavior of the PDF for recurrent intervals at different q. Figure 2(a) presents the relations between the PDF of τ versus q. The tails of the PDFs behave like a power-law. With the increase of q from 2 to 4, the scaling exponent decreases monotonically from 4.5 to 2.5. Figure 2(b) provides the re-scaled PDFs, $P_q(\tau)\langle\tau\rangle$, for $\tau/\langle\tau\rangle$, namely standardized recurrent interval, for different values of q. All the five curves collapse into a single curve. Hence, a general scaling relation is obeyed, the general scaling exponent of which is 2.6.

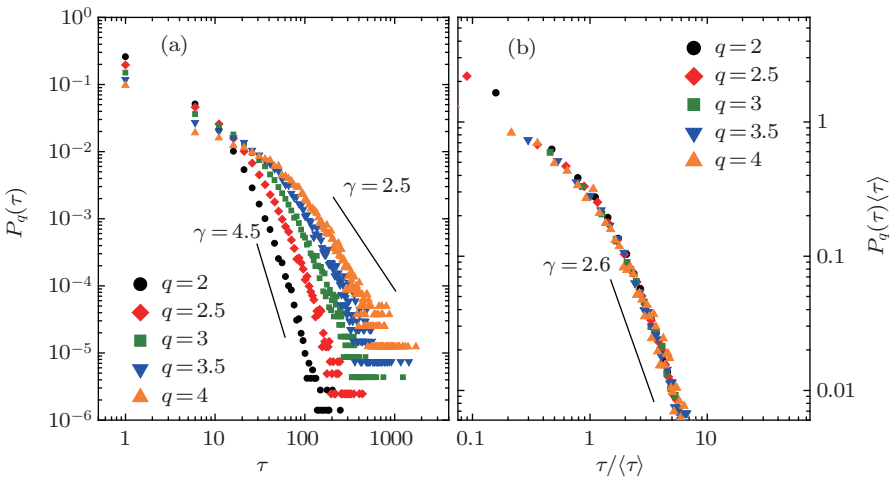


Fig. 2. (color online) Probability distribution of recurrent intervals for volume volatility series. A total of 2282 stocks in Chinese markets are considered. (a) $P_q(\tau)$ versus τ; (b) re-scaled PDF, namely, $P_q(\tau)\langle\tau\rangle$ versus $\tau/\langle\tau\rangle$. The curves collapse into a single curve, indicating that the scaling-law is q-independent.

Li *et al.*[10] also found a general scaling relation in US markets, but the scaling exponent of which is 3.2, a significantly large value compared with that for Chinese markets.

3.2. Impacts of financial factors

The key financial factors affecting stock volatilities include stock lifetime, market capitalization, mean volume, and mean trading value. To investigate impacts of the key factors on re-scaled PDFs, one can divide the total of 2282 stocks into a certain number of subsets according to the four factors, and calculate the re-scaled PDFs for the subsets, respectively.

Figures 3(a)–3(d) show the results for lifetime, capitalization, mean volume, and mean trading value, respectively. For each factor, the stocks are divided into four subsets. To cite an example, the lifetime of the stocks reaches 20 years, which is divided into four intervals, namely, $[0,5]$, $[5,10]$, $[10,15]$, and $[15,20]$ years. The stocks whose lifetimes are in the intervals are catalogued into the corresponding subsets, respectively. The tails of the re-scaled PDFs behave like power-laws. The curves for the market capitalization, mean volume, and mean trading value collapse into single curves with scaling exponent values 3.4, 3.0, and 3.2, respectively. On the contrary, the curves for stock lifetime subsets do not collapse, i.e., with the increase of lifetime, the scaling exponent decreases from 4.7 to 3.7.

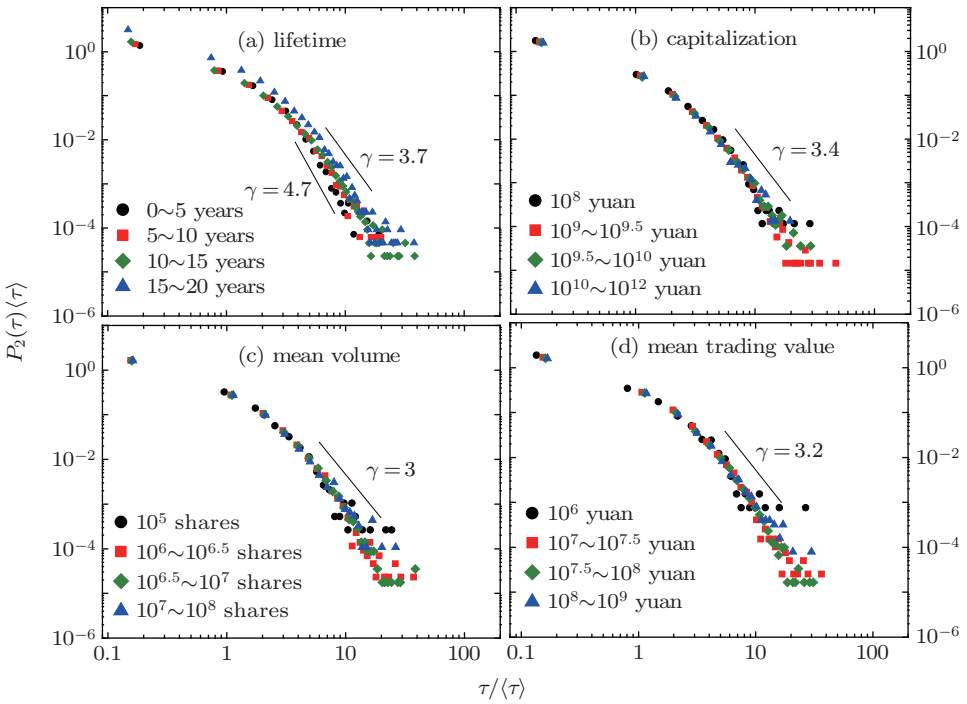


Fig. 3. (color online) Impacts of financial factors on re-scaled PDFs. The threshold is set to be q=2.0. Stocks are sorted into four subsets for each financial factor. (a) Impact of lifetime. γ varies from 3.7 to 4.7; (b) impact of market capitalization. γ stays almost the same ($\gamma = 3.4$) for different subsets; (c) impact of mean volume. γ stays almost the same ($\gamma = 3$) for different subsets; (d) impact of mean trading value. γ stays almost the same ($\gamma = 3.2$) for different subsets.

To confirm the findings, all the stocks are catalogued further into 10 subsets according to lifetime. The ten subsets contain the stocks whose lifetimes are in the intervals of $[0,500],[500,1000],\ldots,[4500,5000]$ days, respectively. The scaling exponent varies between 2.9 and 3.7, which covers a significantly sharper interval compared with that for the four subsets, as shown in Fig. 4(a). Hence, the scaling behavior is lifetime-dependent. From the inset, one can find that the exponent decreases and tends to a constant with the increase of lifetime, which is similar to that for US markets (starts from 4.2 and saturates to 2.8).

Analogous calculations are conducted for the capitalization, mean volume, and mean trading value also, and the resulting re-scaled PDFs are presented in Figs. 4(b)–4(d). The curves collapse into single curves with scaling exponents 3.2, 3.4, and 3.0, respectively. The behavior of re-scaled PDF is independent of the three factors. However, for USA stocks, the scaling exponents decrease from 4.1 and saturate to 3.0 for the subsets with respect to capitalization, mean volume, and mean trading value, respectively.

To understand the different behaviors of re-scaled PDFs for the four factors, one should answer if there exists any dependent relation between these factors. Figure 5 depicts the scattering plots between each pair of the four factors. The absolute values of cross-correlation coefficients,[17,18] denoted with *ccc*, between lifetime and the other three factors are all less than 0.2, which indicates very weak correlations between the lifetime and the other three factors. The absolute values of correlation coefficients between trading value, capitalization, and volume are all larger than or equal to 0.48, which reveals the strong correlations between them. The strong correlations maybe understood by the simple fact that the three factors are all strongly affected by firm size.

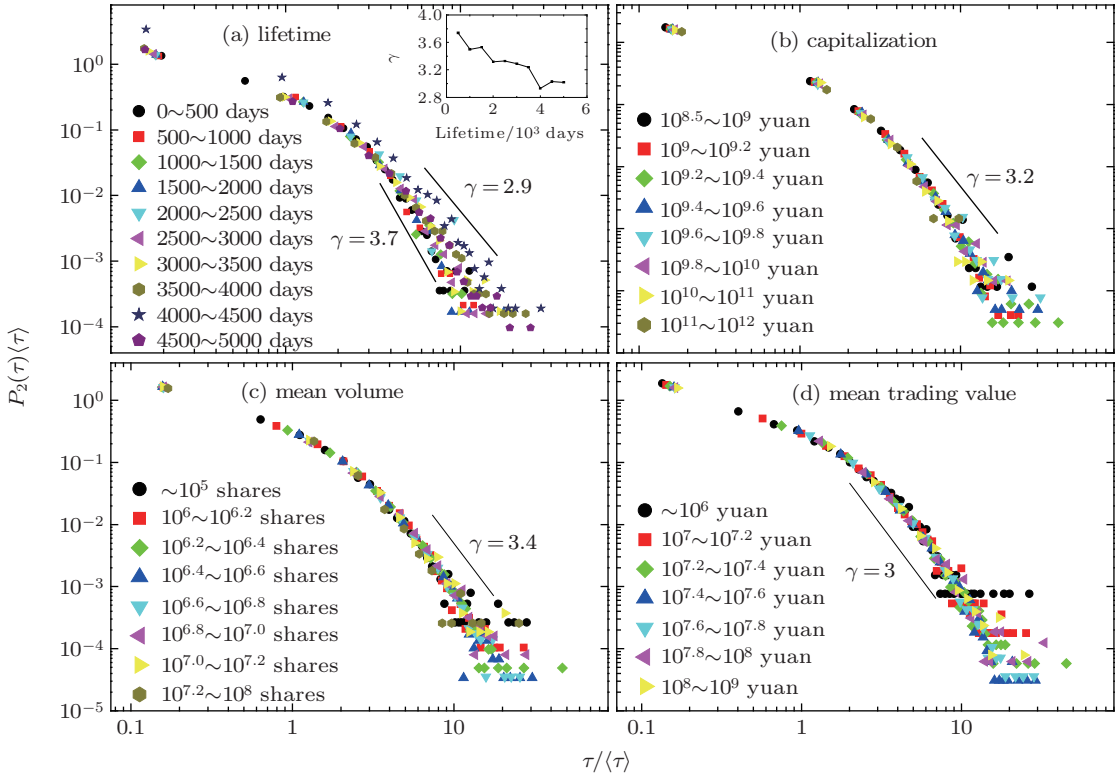


Fig. 4. (color online) Impacts of financial factors on re-scaled PDFs. $q = 2$. (a) Impact of lifetime, the stocks are divided into ten subsets. γ distributes in the interval from 3.7 to 2.9. Inset: γ decreases with the increase of lifetime; (b) impact of capitalizations, stocks are sorted into eight subsets. γ stays almost the same; (c) impact of mean volume, stocks are separated into eight subsets. γ stays almost the same; (d) impact of mean trading value, stocks are sorted into seven subsets. γ stays almost the same.

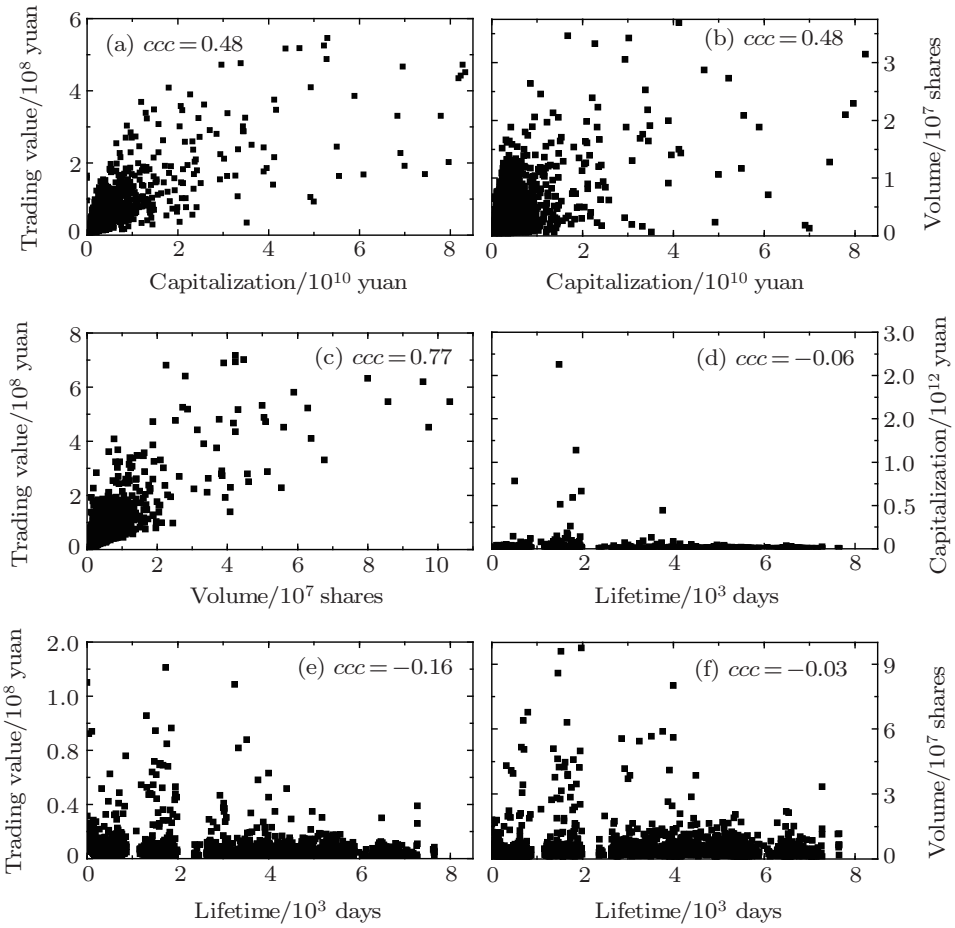


Fig. 5. Scattering plots. (a) Trading value versus capitalization; (b) mean volume versus capitalization; (c) trading value versus mean volume; (d) capitalization versus lifetime; (e) trading value versus lifetime; (f) volume versus lifetime.

4. Conclusions

In the present paper, we investigate the impacts of financial factors on recurrent volume intervals in Chinese markets, a representative of developing stock markets. The results are compared with those for USA's stock markets, a typical developed market. The developed and developing markets share two facts, namely, the tail of the probability distribution function for recurrent volume interval behaves like a power-law, and there does not exist a general scaling behavior in the relation of re-scaled PDF versus lifetime (i.e., lifetime-dependent). However, for Chinese markets, the scaling behaviors in re-scaled PDFs are independent of the capitalization, mean volume, and trading value, while those for USA markets start from larger values and saturate to constants with the increase of the three factors, respectively.

We hope our findings can enrich the statistical investigations on recurrent interval series, and provide a certain reference for risk management and portfolio optimization in the Chinese stock market.

References

[1] Malevergne Y and Sornette D 2006 *Extreme Financial Risks: From Dependence to Risk Management* (Berlin: Springer)

[2] Santhanam M S and Kantz H 2008 *Phys. Rev. E* **78** 051113

[3] Bunde A, Eichner J F, Havlin S and Kantel-hardt J W 2004 *Physica A* **342** 308

[4] McNeil A J and Frey R 2000 *Journal of Empirical Finance* **7** 271

[5] Longin F and Solnik B 2001 *The Journal of Finance* **56** 649

[6] Wood R A, McInish T H and Ord K 1985 *Journal of Finance* **40** 723

[7] Yamasaki K, Muchnik L, Havlin S, Bunde A and Stanley H E 2005 *Proc. Natl. Acad. Sci. USA* **102** 9424

[8] Ren F and Zhou W X 2010 *Phys. Rev. E* **81** 066107

[9] Qiu T, Guo L and Chen G 2008 *Physica A* **387** 6812

[10] Li W, Wang F, Havlin S and Stanley H E 2011 *Phys. Rev. E* **84** 046112

[11] Zhou W X and Sornette D 2004 *Physica A* **337** 243

[12] Qiu T, Zhong L, Chen G and Wu X 2009 *Physica A* **388** 2427

[13] Liang Y, Yang G and Huang J P 2013 *Frontiers of Physics* **8** 438

[14] Zhou W C, Xu H C, Cai Z Y, Wei J R, Zhu X Y, Wang W, Zhao L and Huang J P 2009 *Physica A* **388** 891

[15] Wang Y G and Zhang J 2011 *Chin. Phys. Lett.* **28** 038901

[16] Clauset A, Shalizi C R and Newman M E J 2009 *SIAM Review* **51** 661

[17] Rodgers J L and Nicewander W A 1988 *The American Statistician* **42** 59

[18] Roumen T 2013 *Chin. Phys. Lett.* **30** 088901