A Market Crash or Tail Risk? Heavy Tails and Asymmetry of Returns in the Chinese Stock Market*

Zeyu Xing

Imperial College Business School

Rustam Ibragimov†

Imperial College Business School

Centre for Econometrics and Business Analytics (CEBA), St. Petersburg University

Abstract

Rapid stock market growth without real economic back-up has led to the 2015 Chinese Stock Market Crash with thousands of stocks hitting the down limit simultaneously multiple times. We provide a detailed analysis of structural breaks in heavy-tailedness and asymmetry properties of returns in Chinese A-share markets due to the crash using recently proposed robust approaches to tail index inference. The empirical analysis points out to heavy-tailedness properties often implying possibly infinite second moments and gain/loss asymmetry for daily returns on individual stocks. We further present an analysis of the main determinants of heavy-tailedness in Chinese financial markets. It points out to liquidity and company size as being the most important factors affecting the returns' heavy-tailedness properties. At the same time, we do not observe statistically significant differences in tail indices of the returns on A-shares and the coefficients on factors affecting them in the pre-crisis and post-crisis periods.

Keywords: Market crases, crises, tail risk, heavy tails, gain/loss asymmetry, structural breaks, Chinese stock market

1 Introduction

1.1 Heavy-tailedness in financial markets

Numerous studies in finance and economics indicate many key variables in these fields, including financial returns and foreign exchange rates, are heavy-tailed and have thick-tailed power law distributions (see, among others, Loretan & Phillips (1994), Embrechts et al., 1997, Cont, 2001, Park, 2002, Gabaix, 2009, 2016, Ibragimov et al., 2015, Ch. 3 in McNeil et al. (2015), Ibragimov & Prokhorov, 2017, and references therein). For a heavy-tailed power law distributed random variable (r.v.) X (e.g., representing a risk, financial return or exchange rate), one has

*We thank the Editors, an anonymous referee and the participants at the seminar series at the Centre for Econometrics and Business Analytics (CEBA), St. Petersburg University, for helpful comments and suggestions. Rustam Ibragimov's research for this paper was supported in part by a grant from the Russian Science Foundation (Project No. 22-18-00588)

†Corresponding author. E-mail addresses: zeyu.xing18@imperial.ac.uk (Z. Xing), irustam@imperial.ac.uk (R. Ibragimov)

$$P(X > x) \sim \frac{c_1}{x^{\xi^+}} \tag{1}$$

$$P(X < -x) \sim \frac{c_2}{x^{\xi^-}} \tag{2}$$

$$P(|X| > x) \sim \frac{c}{x^{\xi}}, \tag{3}$$

for large x's, where ξ^-, $\xi^-, \xi = min(\xi^-,\ \xi^-)$, c, c_1, c_2 are positive constants. The parameters ξ^+, ξ^-, ξ in (1)-(3) are referred to as, respectively, the right tail index, left tail index and tail index (or tail exponent) of the variable X. The parameters ξ^+, ξ^- and ξ characterize the degree of heavy-tailedness (the rate of decay of the tails) of power law distributions. Heavy-tailed distributions provide a convenient framework for modelling and quantifying (by their key parameters - the tail index values ξ_+, ξ_- and ξ) the likelihood and the magnitude of large downfalls, large fluctuations and crises in financial and economic markets. In particular, in (1)-(3), the *smaller* values of the tail index parametes ξ_+, ξ_- and ξ correspond to a *higher* degree of heavy-tailedness and thus to a larger likelihood of crises, large downfalls and large fluctuations in (e.g., financial returns) time series of observations on the variable X, and vice versa (see the discussion in Chs. 1 in Ibragimov et al., 2015, and Ibragimov & Prokhorov, 2017). The tail index parameters are further important as they govern existence of moments of X with, for instance, the variance of X being defined and finite: $Var(X) < \infty$ if and only if $\xi > 2$, and, more generally, the pth moment $E|X|^p$, $p > 0$, being finite: $E|X|^p < \infty$ if and only if $\xi > p$.[1] The most of the empirical literature on heavy-tailed distributions agrees that, in the case of developed financial markets, the returns and foreign exchange rates' tail indices ξ belong to the interval $(2, 4)$, thus implying finite variances and infinite fourth moments (op. cit.).[2] Importantly, the stylized facts of heavy-tailedness and volatility clustering are exhibited by widely used GARCH-type models for financial returns and foreign exchange rates (see the results and discussion in Section 8.4 in Embrechts et al. (1997), Davis et al. (1998), Mikosch & Starica (2000*b*), Park (2002), Han & Park (2008), Ch. 4 in McNeil et al. (2015), and Ibragimov et al. (2020)). One should note that finiteness of variances and higher moments for economic and financial indicators like financial returns and exchange rates is crucial in the analysis of many models in economics and finance and also for applicability of classical statistical and econometric approaches, including regression and least squares methods. In a similar fashion, the problem of heavy-tailedness with potentially infinite fourth moments, nonlinear dependence and potential nonstationarity of economic and financial time series needs to be taken into account in applications of regression and autocorrelation-based methods, and related inference procedures in their analysis (see, among others, the discussion in Granger & Orr (1972), and in a number of more recent studies, e.g., Loretan & Phillips (1994); Embrechts et al. (1997); Davis et al. (1998); Mikosch & Starica (2000*b*); Section 7 in Cont (2001); Chung & Park (2007); Miller & Park (2010); Ch. 1 in Ibragimov et al. (2015), Anatolyev (2019), Ibragimov et al. (2020), and the references therein).

Gabaix et al. (2003, 2006) review the empirical results that imply that, in developed financial markets, the tail index of stock returns is very close to 3 and develop a theoretical model explaining this empirical regularity that the authors call the "Cubic Law of the Stock Returns". In the model, heavy-tailedness of financial returns is implied by trading actions of largest market participants (mutual funds and other institutional investors) that have a size distribution with the tail index $\xi = 1$ (Zipf's law). The empirical results in the literature imply that the "Cubic Law of the Stock Returns" does not hold in the case of emerging and developing country financial returns and foreign exchange rates as many of them, including those in Chinese markets, have tail

[1] Naturally, therefore, the standard OLS regression methods and autocorrelation-based time series analysis methods are in principle inapplicable directly and need to be modified in the case of heavy-tailed time series with tail indices ξ smaller than two and infinite (or undefined) variances.

[2] The property that the financial returns' tail indices are smaller than 4 and their fourth moments are infinite implies that the use of the common measure of heavy-tailedness, the kurtosis, is inappropriate: E.g., under $\xi \in (2, 4)$, its estimate given by the sample kurtosis diverges to infinity in probability as the sample size grows. Thus, the sample kurtosis of financial returns is expected to take on increasingly larger values as the sample size increases (see the results and the discussion in Park (2002) and Han & Park (2008)).

indices smaller than 3, and even tail indices smaller than 2 and infinite variances are not uncommon (see Ibragimov et al., 2013, Gu & Ibragimov, 2018, Chen & Ibragimov, 2019, Ch. 3 in Ibragimov et al., 2015 and references therein). This implies that the model in Gabaix et al. (2003, 2006) may need to be modified in the case of emerging and developing markets, e.g., with possible deviations of the distribution of sizes of market participants from Zipf's law due to the governments' regulatory interventions. The analysis in Quintos et al. (2001) and Candelon & Straetmans (2006) point to important structural breaks in heavy-tailedness properties of developed and emerging country stock index returns and foreign exchange rates, in particular, those due to the Asian financial crisis.[3]. Ankudinov et al. (2017) provides the analysis of tail index regressions for financial returns in Russia that indicates importance of stock liquidity and company size as determinants of the returns' heavy-tailedness.

1.2 Chinese equity market and the 2015 crash

In terms of scale, China's economy is currently the second largest one in the world and is on the track of exceeding that of the United States. However, institutions and financial management in Chinese stock markets work both similar and different from the western world (Jiang et al. 2017).

The Shanghai Stock Exchange (SSE) and Shenzhen Stock Exchange (SZSE) date back to the 1990s. Both the exchanges are order-driven without specialists or market makers with both upward cap and downward floor of 10% daily price fluctuation limit from the closing price of the prior day. In March 2010, China introduced a pilot scheme and eased the ban on short-selling and margin-trading for a certain list of stocks. Chang et al. (2014) find that this scheme has improved Chinese stock market efficiency and decreased its volatility.

A-shares are regular domestic stocks settled in renminbi (RMB). There are also stocks settled in foreign currencies other than RMB. B-shares are denominated in U.S. or Hong Kong dollars with the same cash flow rights as A-shares. Cross-listed shares are stocks listed both in Chinese mainland and Hong Kong or foreign markets, such as H-shares (listed in Hong Kong), N-shares (listed in New York), S-shares (listed in Singapore) and L-shares (listed in London). The Shanghai-Hong Kong Stock Connect and Shenzhen-Hong Kong Stock Connect were introduced in 2014 and 2016 respectively, since when qualified investors are able to directly invest in Hong Kong stock markets (see Jiang et al. (2017) and also the review in Chen & Ibragimov (2019)).

Before having peaked on June 12, 2015, stock market in China had increased around 150% in one single year, while the nominal GDP growth rate had been on the downwards track from 14.23% in 2007 to 6.9% in 2015. However, it took far less time for the market to drop, when it went down 40% from around 5166 to 3052 points in less than four months.

Zeng et al. (2016) summarize the causes of China's 2015 stock market crash as the inconsistency of economic fundamentals and market performance. Besides the slowing GDP growth rate, stock prices skyrocketed in companies with meager earnings or even losses. On the demand side, a huge amount of inexperienced investors flooded into the market with their own savings and the number of investors added up to 7% of the population then (Sornette et al. 2015). On the supply side, short selling and trading on the margin have been approved during that period by authorities, contributing to the market leverage.

More than 40 measures had been enacted to rescue the market by the government and authorities. However, by the end of August, as the SSE Composite Index fell down below 4500 points, it became clear that the government has failed to boost the stock market as was initially planned. Among so many actions deployed, only a few have worked except one, which was the establishment of of the so-called "National Team" that directly purchased shares from the market (see Zeng et al. (2016)).

[3]See also Koedijk et al., 1992, for the analysis of changes in the tail behavior of Latin American foreign exchange rates in response to changes in foreign exchange rate regimes

1.3 Research problems and contributions

Despite the existence of a few studies on heavy-tailedness properties of developed markets reviewed in Section 1.1, the research on heavy-tailedness of emerging markets, in particular, those in China, and especially their determinants remains limited.

This paper attempts to partially fill this gap in the literature by providing a detailed econometric analysis of several important research questions on heavy-tailedness characteristics of financial markets in China.

i. Using recently proposed robust approaches to tail index inference, we provide a rigorous analysis of heavy-tailedness and asymmetry properties of returns in Chinese A-share markets.

ii. We further provide a detailed analysis of structural changes in these properties due to the 2015 crash.

iii. The paper also presents a study of the main determinants of heavy-tailedness in Chinese financial markets, including stock market factors, stock liquidity and firm-specific variables and attributes such as company size, ownership structure and sector affiliation.

iv. In addition, we provide the results on the analysis of potential structural breaks in the relationship between the degree of heavy-tailedness in A-share returns and the explanatory factors considered.

Among other results, the paper provides the estimates of the degree of heavy-tailedness of daily returns on individual stocks that point out to their fat-tailedness properties with possibly infinite variances and second moments. The estimates further indicate gain/loss asymmetry in the tails of the returns' distributions.

The results the paper point out to liquidity and company size as being the most important factors affecting A-share returns' heavy-tailedness properties. At the same time, we do not observe statistically significant differences in tail indices of the returns on A-shares and the coefficients on factors affecting them in the pre-crisis and post-crisis periods.

1.4 Organization of the paper

The paper is organized as follows. Section 2 reviews tail index inference approaches employed in the analysis. Section 3 describes the data and variables used in the analysis. Section 4 provides the results of the empirical analysis and their discussion. In Section 6 we make some concluding remarks.

2 Inference on Heavy-Tailedness

Several approaches to inference about the tail index of heavy-tailed distributions (1)-(3) are available in the literature (see, among others, the reviews in Embrechts et al., 1997, Beirlant et al., 2004, and Ch. 3 in Ibragimov et al., 2015). The two most commonly used are Hill's tail index estimates and the OLS log–log rank-size regression approach to tail index estimation.

Below, we describe Hill's and log-log rank-size regression estimates $\hat{\xi}_{Hill}$ and $\hat{\xi}_{RS}$ of the tail index ξ in power law distributions (3). Hill's and log-log rank-size regression estimates $\hat{\xi}^+_{Hill}$, $\hat{\xi}^+_{RS}$ and $\hat{\xi}^-_{Hill}$, $\hat{\xi}^-_{RS}$ of the right and left tail indices ξ^+ and ξ^- (1)-(2) are defined in a similar way, with the use of the largest positive returns and largest negative returns instead of the returns' largest absolute values.

Let $r_1, r_2, ..., r_N$ be a sample of returns that have power law distribution (3). Further, let, for $n < N$;

$$|r|_{(1)} \geq |r|_{(2)} \geq ... \geq |r|_{(n)} \tag{4}$$

be decreasingly ordered largest absolute values of observations in the sample (in practical applications, one usually takes the number n of observations in the tails of power law distribution used for estimation of the tail index ξ to be equal to some small fraction, e.g., 5% or 10%, of the total sample size $N : n \approx 0.05N, 0.1N$). Hill's estimator $\hat{\xi}_{Hill}$ of the tail index ξ in (3) is given by

$$\hat{\xi}_{Hill} = \frac{n}{\Sigma_{t=1}^{n}(log|r|_{(t)} - log|r|_{(n+1)})} \tag{5}$$

with the standard error $S.E._{Hill} = \frac{1}{\sqrt{n}}\hat{\xi}_{Hill}$. The corresponding 95% confidence interval for the true value of the tail index ξ is given by

$$\left(\hat{\xi}_{Hill} - 1.96 \times \frac{1}{\sqrt{n}}\hat{\xi}_{Hill},\ \hat{\xi}_{Hill} + 1.96 \times \frac{1}{\sqrt{n}}\hat{\xi}_{Hill}\right).^{4} \tag{6}$$

It was reported in a number of studies that inference on the tail index using Hill's estimator suffers from several problems, including sensitivity to dependence, deviation from power laws in the tails and small sample sizes (see, among others, the discussion in Ch. 6 in Embrechts et al., 1997, Gabaix & Ibragimov, 2011, and Ch. 3 in Ibragimov et al., 2015). Motivated by these problems, several studies have focused on alternative robust approaches to the tail index estimation, including small-sample weighted analogues of Hill's estimator (see Huisman et al., 2001) and nonlinear tail index estimation approaches (see Embrechts et al., 1997). A popular simple approach to tail index estimation that appears to be more robust than inference procedures based on Hill's estimates is that based on log-log rank-size regressions.[5] The log-log rank-size regression tail index estimation approach is motivated by the linear relationships like $\log(P(|X| > x)) \sim c - \xi \log(x)$ for large $x's$ implied by (3). Gabaix & Ibragimov (2011) show that the empirical analogues of such relationships that correct the bias in tail index estimation to the first-order can be based on regressions of shifted ranks of power law distributed observations on their sizes, that is, the regressions $\log(Rank - 1/2) = a - b\log(Size)$ with the optimal shift of 1/2 in ranks. More precisely, following the approach, one runs the following OLS regression:

$$\log(t - 1/2) = a - b\log|r|_{(t)}, \quad t = 1, 2, ..., n, \tag{7}$$

and takes b as the log-log rank-size regression estimate $\hat{\xi}_{RS}$ of the tail index ξ in (3). Gabaix & Ibragimov (2011) show that the standard error of the log-log rank-size regression tail index estimator $\hat{\xi}_{RS}$ is different from the OLS standard error and is given by $S.E. = \sqrt{\frac{2}{n}}\hat{\xi}_{RS}$. The corresponding 95% confidence interval for the true value of the tail index ξ is given by

$$\left(\hat{\xi}_{RS} - 1.96 \times \sqrt{\frac{2}{n}}\hat{\xi}_{RS},\ \hat{\xi}_{RS} + 1.96 \times \sqrt{\frac{2}{n}}\hat{\xi}_{RS}\right). \tag{8}$$

The analysis of (a)symmetry in heavy-tailedness properties of upward and downward fluctuations, that is large upward moves and large downfalls, in financial returns r_t can be based on estimates of the right and left tail indices ξ^+, ξ^- in power laws (1)-(2) and their standard errors and confidence intervals constructed as described in this section (see also Ch. 3 in Ibragimov et al., 2015, and Section 3.2 in Ibragimov et al. (2015) for examples of the analysis of asymmetry in right and left tails of the distribution of emerging country exchanges using their right and left tail indices).

In a similar way, the analysis of structural breaks in heavy-tailedness properties of financial returns due

[4]The standard errors and confidence intervals for the tail index estimators dealt with are asymptotically valid under the assumption that r_t are i.i.d. observations from power laws (3)-(2). As discussed below, the log-log rank-size regression approach to tail estimation appears to be more robust to dependence in observations r_t, including GARCH-type dependence typical for financial returns, as compared to that based on Hill's tail index estimates (see also the discussion in Gabaix & Ibragimov (2011) and the online appendix to that paper).

[5]In particular, the log-log rank-size regression tail index estimation approaches appear to be more robust than those based on Hill's estimates to nonlinear dependence in the form of GARCH-type volatility dynamics as in the case of real-world economic and financial and to deviations from power laws (1)-(3) in the form of slowly varying functions (see the discussion in Gabaix & Ibragimov, 2011 and the numerical results in the online appendix to that paper,
$https://scholar.harvard.edu/files/xgabaix/files/rank\_1\_2\_appendix.pdf$

to the 2015 crash can be conducted using the estimates $\hat{\xi}_{pre}$ and $\hat{\xi}_{post}$ of the tail indices in the pre-crash and post-crash periods, the standard errors of the estimates described in this section and the corresponding confidence intervals.[6]

3 Data

The analysis is based on the data for the period from 13/04/2012 to 04/07/2019 collected from Bloomberg. The data includes the time series of daily closing prices, daily volume, daily bid-ask spread, quarterly market capital, daily P/B ratio, yearly return on equity (ROE) and total investment to total assets ratio.

The analysis of structural breaks due to the 2015 crash in the degree of heavy-tailedness of financial returns and in its dependence on explanatory factors considered is conducted using the data for the pre-crash period from 13/04/2012 to 12/06/2015 and the post-crash break period from 13/06/2015 to 04/07/2019.

Many companies in the sample have a record of hundreds of days on suspension. To deal with such missing data problems, the analysis focuses on shares that have at least 1600 of non-zero returns in the 1885 trading days considered. The sample considered consists of 1038 public companies.

4 Empirical analysis

Due to the large volume of stock indices, we present and describe in detail the results of tail index estimation for 4 companies in the sample. The results of the analysis for the full sample of companies are presented in the Appendix.

[6]See Ibragimov et al. (2013), Chen & Ibragimov (2019) for applications of the approach in the analysis of structural breaks due to the beginning of the 2008 global financial crisis in the degree of heavy-tailedness of developed and emerging country foreign exchange rates and the returns on several A- and H-shares in China.

4.1 Tail index estimates

Table 1: Tail index estimates for absolute returns

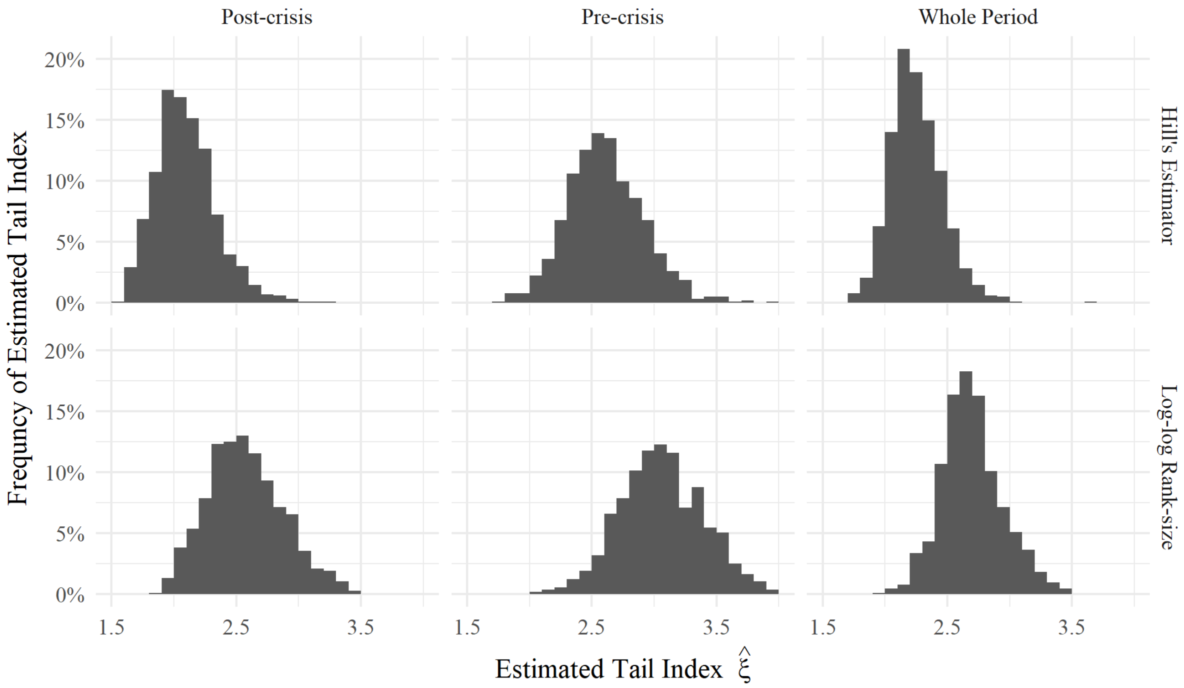
Stock	Truncation(%)	N	Log-log Rank-size $\hat{\xi}_{RS}$	$CI_{95\%,RS}$	Hill's Estimate $\hat{\xi}_{Hill}$	$CI_{95\%,Hill}$
After the 2015 China's Stock Market Crash						
000963 CH Equity	7.5	951	3.54	(2.38 , 4.7)	3.17	(2.44 , 3.9)
	15.0	951	3.28	(2.52 , 4.04)	2.92	(2.44 , 3.4)
600021 CH Equity	7.5	887	2.84	(1.89 , 3.79)	2.19	(1.67 , 2.72)
	15.0	887	2.08	(1.58 , 2.58)	1.67	(1.39 , 1.95)
600770 CH Equity	7.5	884	4.98	(3.29 , 6.67)	3.21	(2.44 , 3.98)
	15.0	884	2.55	(1.94 , 3.16)	2.02	(1.68 , 2.36)
601010 CH Equity	7.5	910	2.97	(1.98 , 3.96)	2.20	(1.68 , 2.73)
	15.0	910	2.08	(1.59 , 2.57)	1.67	(1.39 , 1.95)
Before the 2015 China's Stock Market Crash						
000963 CH Equity	7.5	743	4.36	(2.76 , 5.96)	3.90	(2.89 , 4.91)
	15.0	743	3.67	(2.71 , 4.63)	2.90	(2.36 , 3.44)
600021 CH Equity	7.5	717	2.61	(1.63 , 3.59)	1.98	(1.46 , 2.51)
	15.0	717	2.09	(1.54 , 2.64)	1.88	(1.53 , 2.23)
600770 CH Equity	7.5	746	3.87	(2.45 , 5.29)	3.11	(2.3 , 3.91)
	15.0	746	3.24	(2.4 , 4.08)	2.92	(2.38 , 3.46)
601010 CH Equity	7.5	745	2.74	(1.73 , 3.75)	2.25	(1.67 , 2.84)
	15.0	745	2.11	(1.56 , 2.66)	1.87	(1.53 , 2.22)
Through the Whole Period of 2012 to 2019						
000963 CH Equity	7.5	1694	3.89	(2.94 , 4.84)	3.45	(2.85 , 4.05)
	15.0	1694	3.47	(2.87 , 4.07)	2.92	(2.56 , 3.28)
600021 CH Equity	7.5	1604	2.81	(2.1 , 3.52)	2.11	(1.73 , 2.48)
	15.0	1604	2.09	(1.72 , 2.46)	1.75	(1.53 , 1.97)
600770 CH Equity	7.5	1630	3.91	(2.93 , 4.89)	2.80	(2.31 , 3.3)
	15.0	1630	2.74	(2.25 , 3.23)	2.27	(1.99 , 2.55)
601010 CH Equity	7.5	1655	2.90	(2.18 , 3.62)	2.22	(1.83 , 2.61)
	15.0	1655	2.10	(1.73 , 2.47)	1.76	(1.54 , 1.98)

Table 1 provides Hill's and the log-log rank-size regression estimates $\hat{\xi}_{Hill}$ and $\hat{\xi}_{RS}$ of the tail index ξ of the returns on the shares of companies considered, together with their standard errors and the corresponding 95% confidence intervals (see Section 2 for details). Due to the moderate size of financial returns time series used in the analysis, tail index estimation uses the truncation levels n equal to 7.5% and 15% of the total sample size N of time series observations: $n \approx 0.075N, 0.15N$. The analysis below using Hill's plots shows that the tail index estimates are generally robust to the truncation level choice.

The results in Table 1 suggest that the obtained Hill's estimates of tail indices of absolute returns are in general smaller than their log-log rank-size regression estimates. This is further confirmed by Figure 1 that provides the histograms of Hill's and log-log rank-size regression tail index estimates of returns with 15% truncation across the companies considered. In both the pre-crisis and post-crisis periods and over the whole period considered, Hill's estimates are distributed leftward on the x axis as compared to the log-log rank-size regression estimates. Both of the histograms have a long tail on the right hand side indicating less pronounced heavy-tailedness for many companies. The estimates are distributed nearly symmetrically around the center.

The average differences of Hill's estimate and the log-log rank-size estimate are 0.4787, 0.4617 and 0.4431 post-crisis, pre-crisis and in the whole period respectively.

Figure 1: Histogram of tail index estimates with 15% Truncation



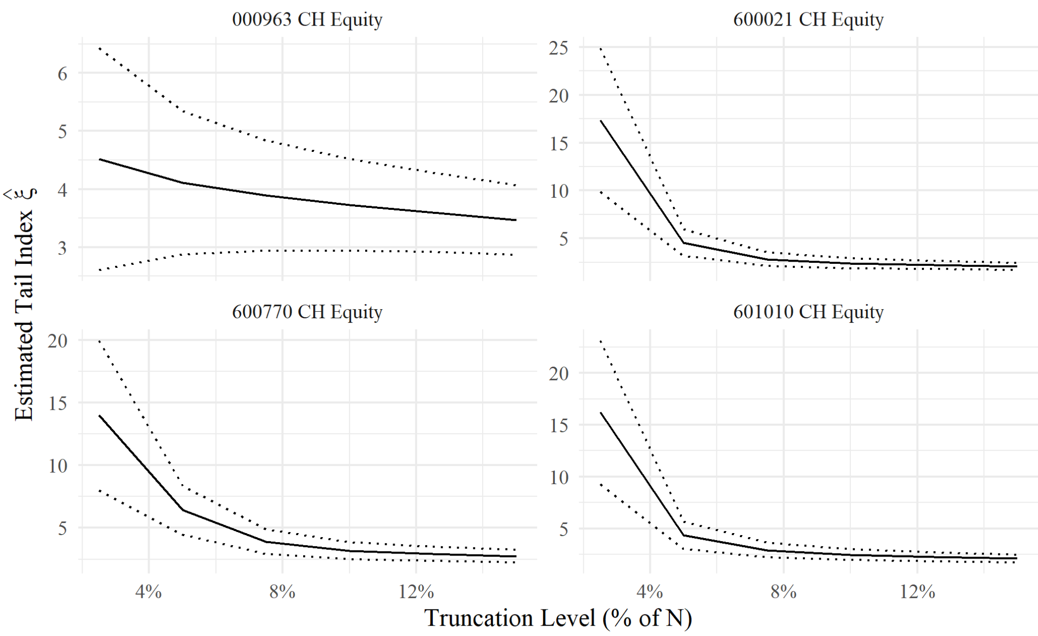
Importantly, with the 15% truncation level, there are only seven companies for which the 95% confidence intervals for the tail indices of their shares constructed using the log-log rank-size regression and Hill's estimates do not intersect. For these companies, this is the case for confidence intervals for the tail index of absolute returns over the whole period. However, for the 7.5% truncation level, there are 148 out of $9,342$ cases where the above confidence intervals do not intersect with each other. The difference may be due to strong sentivity of Hill's estimates to truncation levels used (Embrechts et al. 1997). Overall, the above results point out to similarity of Hill's and log-log rank-size estimates of the tail indices of financial returns in the Chinese stock market in the period from April 2012 to July 2019, even though Hill's point estimates are in general smaller.

Due to the disadvantages of Hill's estimators discussed in Section 2 and robustness of log-log rank-size regression approach, the analysis and the discussion below will be mostly based on log-log rank-size regression approach to tail index estimation with the optimal shift of $1/2$ in ranks (see the discussion in Gabaix & Ibragimov (2011) and Section 2).

Similar to Ibragimov et al. (2013), in order to illustrate the appropriateness of the tail truncation levels (7.5% and 15%) used in this section, we follow the analysis and suggestions in Embrechts et al. (1997) and Mikosch & Starica (2000*b*) and present the analogues of Hill's plots for the log-log rank-size regression tail index estimates for four companies in the sample. These are graphs of the log–log rank-size regression point estimates $\hat{\xi}_{RS}$ of the tail indices of the returns on the companies shares, together with the corresponding 95%-confidence intervals in (8) for the true tail index values computed using log–log rank-size regressions,

Similar to the four companies dealt with in the figure, the log-log rank-size estimates for returns on shares of companies in the sample, the log-log rank-size tail index estimates start to stabilize above the truncation level roughly 6% while the length of the 95% confidence intervals shrinks down evidently as well. The conclusions indicate that the use of the truncation levels of 7.5% and 15% in tail index estimation is reasonable. In general, the log-log rank-size regression tail index estimates are robust to the choice the truncation level choice greater than 6%, as the corresponding confidence intervals for the tail index intersect. At the same time, in contrast to the analysis for emerging country foreign exchange rates in Ibragimov et al. (2013), due to smaller sample sizes in this paper, the confidence intervals constructed using small truncation levels are rather wide and do not intersect with those constructed using the above larger truncation levels.

Figure 2: Hill's plots for log-log rank-size regression tail index estimates $\hat{\xi}_{RS}$



The plots in Figure 3 illustrate the dynamics of daily returns from four sampled public companies with different tail indices and the degrees of heavy-tailedness. The blank parts of the diagrams for the returns on 600021 CH and 600770 CH equities correspond to their long periods of trading suspensions. It was more than normal for companies in China to suspend their trading due to high market volatility or other major issues during and after the 2015 stock market crash.

The histogram in Figure 4 further illustrates the differences in heavy-tailedness of the distribution of the returns plotted in Figure 3. As is seen from the histogram, heavy-tailedness of the returns' distribution and presence of outlier becomes more pronounced as the tail indices decrease. In addition, as is seen from Figure 3, the decrease in tail indices implying more pronounced heavy-tailedness is accompanied by the increase in the likelihood and the magnitude of large downfalls and large fluctuations in the returns' time series.

In the plots in Figure 4, one observes clusters of returns in the intervals $[-10\%, -9\%)$, $(9\%, 10\%]$. This is due to the cap (floor) of 10% on the daily price variation in the Chinese market. The clusters at the boundary of 10% indicate a so-called magnet effect of daily price limits. Liquidity risk and behavioral investors catalyze the process to touch and trigger the limit (Cho et al. 2003). Due to the presence of clusters at the boundary of 10%, (Hill's and log-log rank-size regression) tail index estimates with small truncation levels used in estimation are expected to be large (Chen & Ibragimov 2019), which is why, in particular, the tail index estimates at the starting points with small truncation levels in the graphs in the Figure 2 are so high as compared to the estimates with larger truncation levels.

Figure 3: The dynamics of daily returns

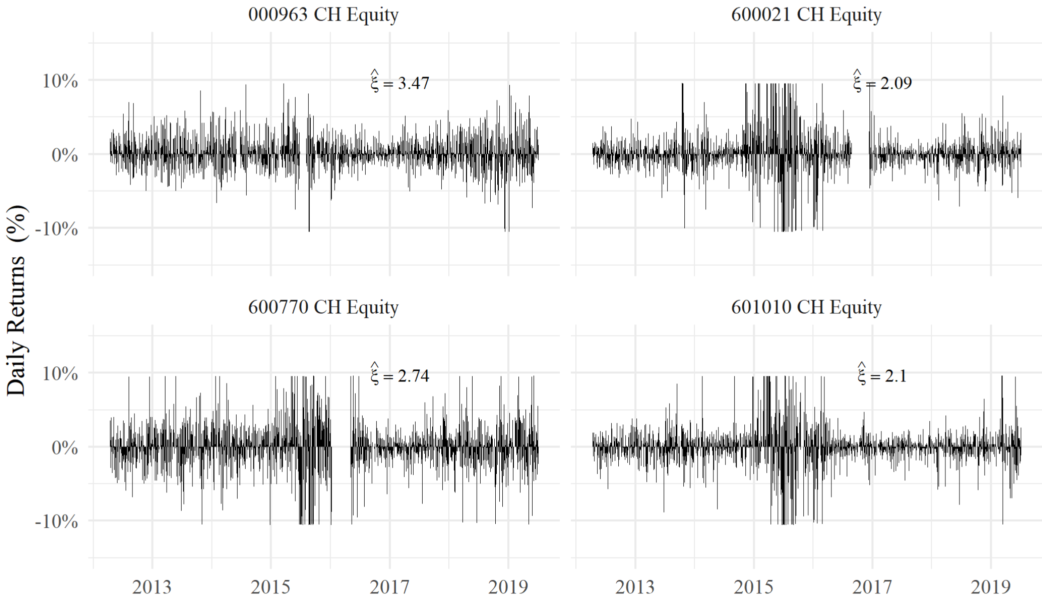
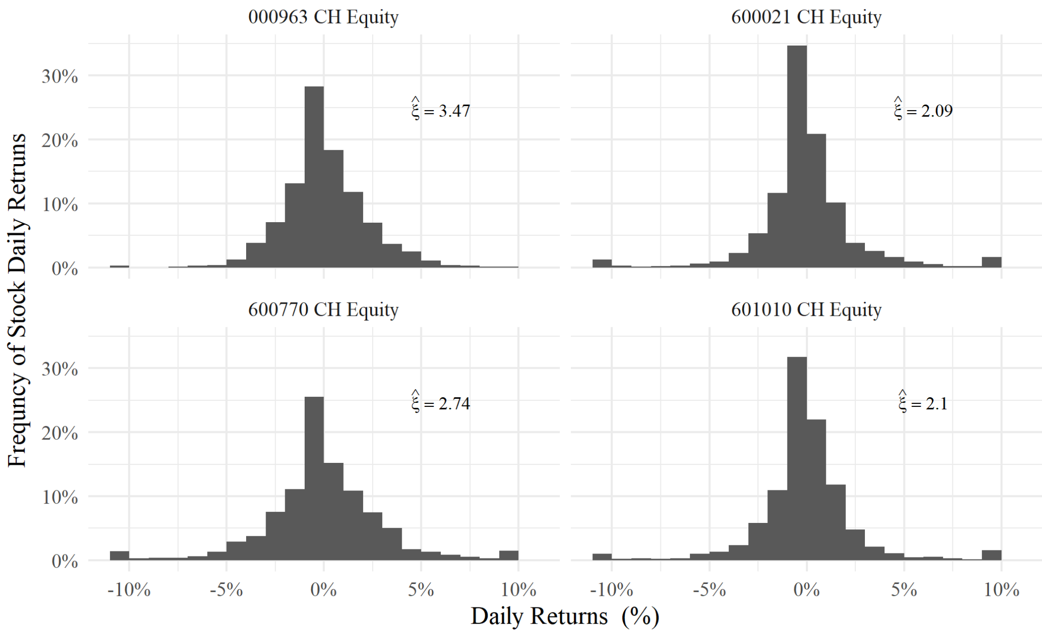


Figure 4: Histograms of daily returns



4.2 (A)symmetry in the left and right heavy-tailedness

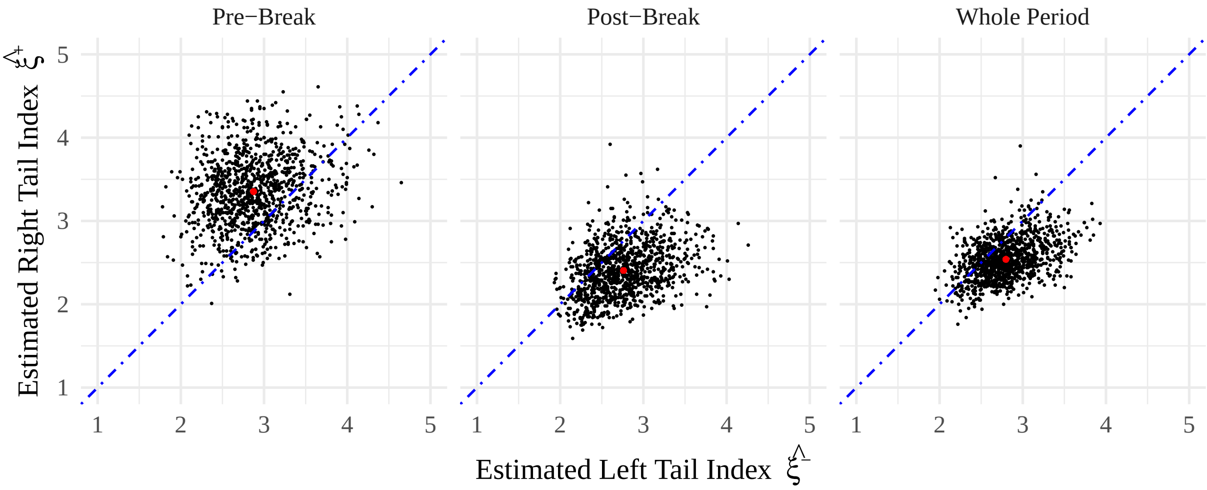
This section provides the analysis of asymmetry in heavy-tailedness in the right and left tails of the distribution of the returns considered. It also analyses structural changes in the asymmetry due to the 2015 market crash.

Figure 5 provides scatterplots of the log-log rank-size regression estimates $\hat{\xi}^+_{RS}$ and $\hat{\xi}^-_{RS}$ of the right and left tail indices ξ^+ and ξ^- in power laws (1) and (2) for financial returns considered. One can observe, due to smaller sample sizes used in estimation, a larger variance in estimates of right tail and left tail indices ξ^+ and ξ^- in (1) and (2) as compared to the estimates of the tail index ξ for the absolute values of the returns in (3).

In Figure 5, if a point is located downwards to the diagonal line, then this suggests that the right tail index ξ^+ is smaller than the left tail index $\xi^- : \xi^+ < \xi^-$. That is, in this case, heavy-tailedness is more pronounced in the right tail of the distribution of the return on the corresponding stock. The conclusions are reversed for a point located above the diagonal line. The latter suggests that the left tail index is smaller than the right tail index: $\xi^- < \xi^+$ and thus heavy-tailedness is more pronounced in the left tail of the return's distribution.

According to tail index estimation results, before the market crash on 12^{th} June 2015, 832 of 1038 companies had a smaller estimate of the right tail indices ξ^+ as compared to the estimates of the left tail indices ξ^-. This number declines to 191 post-crisis. These results are in accordance with the behavior observed in the market, with more large upward moves of stock price during the bull market and large downfalls during the bear market. This effect appears to be more pronounced for large downfalls in the post-crisis period, with the points $(\hat{\xi}^-_{RS}, \hat{\xi}^+_{RS})$ for 815 of 1038 stocks lying below the diagonal line (although very close to it).

Figure 5: Scatterplots of the left and right tail index estimates $(\hat{\xi}^-,\ \hat{\xi}^+)$



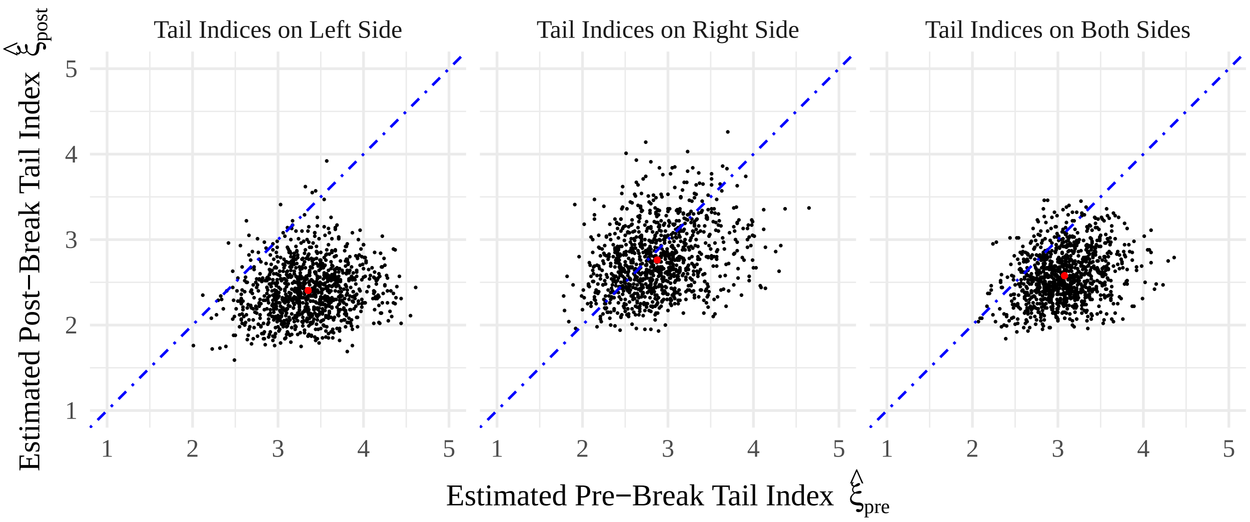
One should note that the 95% confidence intervals for the right and left tail indices ξ^+ and ξ^- for all the stocks considered that are constructed using their estimates in Figure 5 intersect. This implies that the null hypothesis $H_0:\ \xi^- = \xi^+$ of symmetry in the degree of heavy-tailedness properties of the right and left tails of the returns' distributions cannot be rejected in favor of the two-sided alternative $H_a:\ \xi^- \neq \xi^+$ at 5% level. That is, the right and left tail indices ξ^+ and ξ^- of the returns' distributions are statistically indistinguishable (see also Ibragimov et al. (2013) for the analysis and conclusions on symmetry in the degree of heavy-tailedness in right and left tails of distributions of emerging country foreign exchange rates).

4.3 Structural breaks in heavy-tailedness

Similar to the analysis of heavy-tailedness (a)symmetry in the previous section, this section of the paper provides the analysis of structural breaks in heavy-tailedness properties of returns' distribution due to the 2015 crash using the estimates and confidence intervals for tail indices in the pre-crash and post-crash periods (see Ibragimov et al. (2013) for the related analysis of structural breaks in heavy-tailedness properties of distributions of emerging country foreign exchange rates due to the beginning of the 2008 financial crisis). Figure 6 provides the scatterplots of the pre-crash and post-crash estimates $(\hat{\xi}^-_{pre},\ \hat{\xi}^-_{post})$, $(\hat{\xi}^+_{pre},\ \hat{\xi}^+_{post})$ and $(\hat{\xi}_{pre},\ \hat{\xi}_{post})$ of tail indices of the returns in (1)-(3).

In the scatterplots in Figure 6, if a point is located below the diagonal line, then this suggests a larger value of the corresponding tail index in the pre-crash period as compared to its post-crash value. In this case, therefore the corresponding return has a thinner tail before 12^{th} June 2015 as compared to the period after that date. In contrast, if a point in the scatterplots is located above the diagonal line, then this suggest that the corresponding pre-crash tail index value is smaller than its post-crash value, and thus the corresponding public company's return has a heavier tail in the period before 12^{th} June 2015 as compared to the period after that date.

Figure 6: Scatterplots of the pre-crash & post-crash tail index estimates ($\hat{\xi}_{pre}$, $\hat{\xi}_{post}$)



As is seen from the scatterplots for the pre-crash and post-crash left tail indices of large stock price downfalls in Figure 6 there appears to be a significant change in the tail indices at the time of the crash. The pre-crash left tail index value is smaller than its post-crash value only in the case of 25 of 1038 return time series. This indicates that, for nearly all of the returns in the sample, the left tails become heavier after the stock market crash. This conclusion may be due to the bear market with large stock price downfalls replacing the bull market before the crash.

In the corresponding scatterplot for the right tail indices of large stock price upward movements in Figure 6, one observes less deviations of the points ($\hat{\xi}^+_{pre}$, $\hat{\xi}^+_{post}$) from the diagonal line. Specifically, the points lie above the diagonal for 603 out of 1038 companies: ($\hat{\xi}^+_{pre} < \hat{\xi}^+_{post}$), while the points for the other 425 companies lie below the diagonal with ($\hat{\xi}^+_{pre} > \hat{\xi}^+_{post}$). Although more positive large upward movements are expected during the bull market suggesting that pre-crises right tail indices are expected to be smaller than their post-crisis values (with more pronounced heavy-tailedness in the right tails of returns' distributions in the pre-crisis period), the relative symmetry of the scatterplot of the pre- and post-crash estimates of the right tail indices may be explained by a positive and active market environment in the post-crisis period. In particular, the post-crash period includes the first half of 2019. During the first half of 2019, the SSE Composite opened on 2^{nd} January at 2497.88 points and closed on 4^{th} July at 3005.25 points, while SSEC hit 3288.45 as its highest price on 8^{th} April with the maximum price growth of 31.73% during the period. Such a positive and active market environment brings positive returns, which could neutralize the effect of the bear market. As this period is a part of the post-crisis period, the estimates $\hat{\xi}^+_{pre}$ and $\hat{\xi}^+_{post}$ are closer to each other.

As is seen from Figure 6, due to larger sample sizes used in estimation, the tail index estimates for absolute returns (3) have smaller variance than the estimates of right and left tail indices in (1)-(2). The points ($\hat{\xi}_{pre}, \hat{\xi}_{post}$) with pre- and post-crash estimates of the tail index $\hat{\xi}$ are located above the diagonal for returns of only 105 of 1038 companies. The conclusion that, in general, the returns' tail indices become smaller post-crash thus implying more pronounced heavy-tailedness is in accordance with the property that more large stock price downfalls is observed during the bear market.

Similar to the heavy-tailedness (a)symmetry analysis in Section 4.2, for most of the returns in the sample, the 95% confidence intervals with 15% truncation levels for the pre-break and post-break tail indices intersect. There are only 3 and 13 cases for negative and absolute returns, respectively, where the intersections of the confidence intervals are empty. Even with the lower truncation level of 7.5%, the total number of stocks with empty intersections of the confidence intervals adds up to 96 out of 3114 cases (1038 companies $\times$ 3 types of tail indices).

Figure 7: Daily Returns Variation Between Pre-/Post-crisis with $\hat{\xi}_{RS,15\%}$

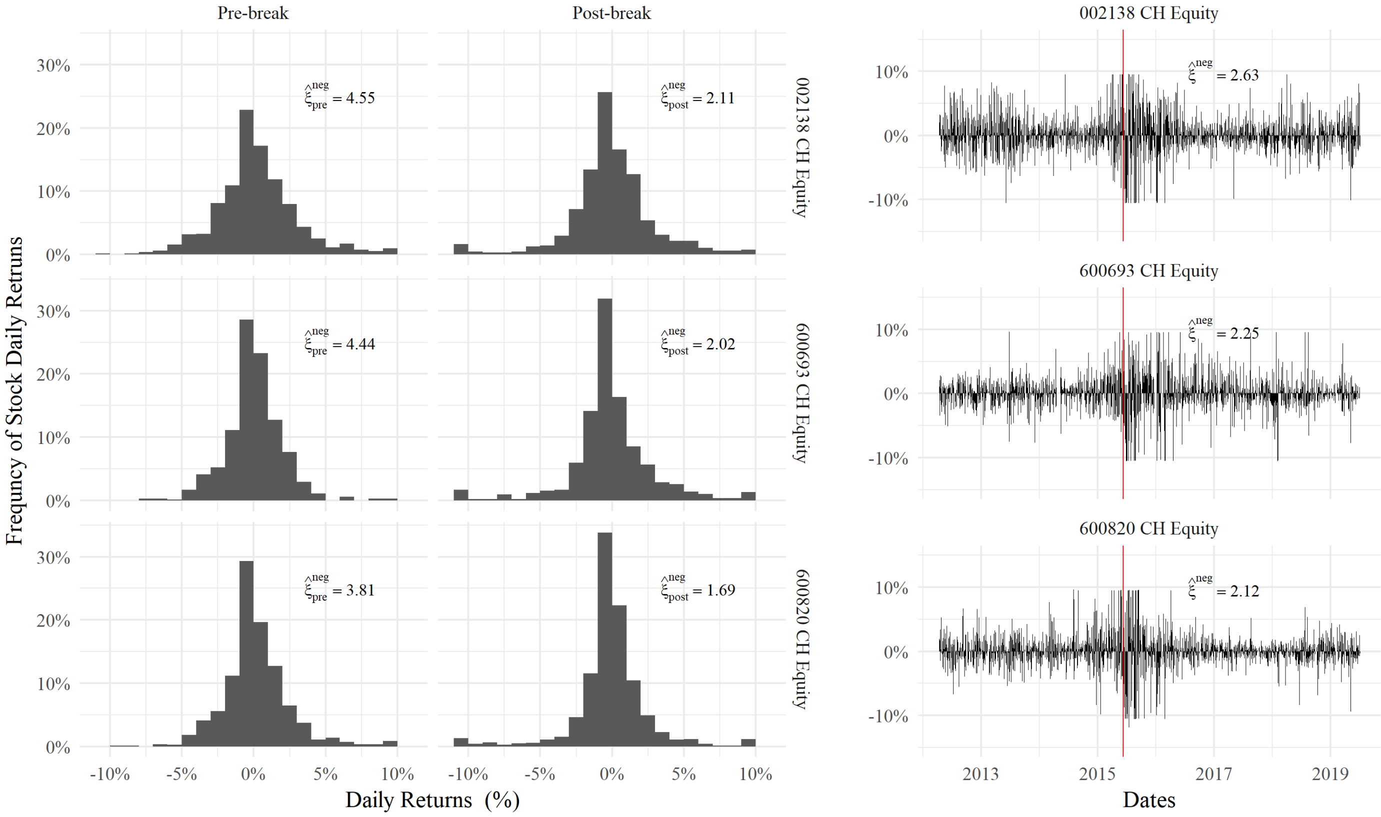


Figure 7 illustrates the above 3 cases where the confidence intervals for the left tail indices in the pre-crash and post-crash periods do not intersect and thus a structural break in the left tail indices at the 2015 crash date is observed. The change in the distribution of negative returns of these stocks is visible from the returns' histograms. Taking 002138 CH Equity as an example, the pre-break left tail index estimate is $\hat{\xi}^{-}_{pre,RS} = 4.55$ with the 95% confidence interval $CI_{95\%,RS} = (2.83,\ 6.27)$. In the post-crisis period, the left tail index estimate becomes $\hat{\xi}^{-}_{post,RS} = 2.11$ with the 95% confidence intervals $CI_{95\%,RS} = (1.43,\ 2.79)$, implying that the tails of the negative returns' distribution become heavier post-crisis. This, in turn, explains the fact that more frequent and larger in magnitude downfalls are observed in the returns' time series considered in the post-crisis period..

Similar conclusions are also implied by the corresponding diagrams for the 13 returns' time series with statistically significant changes in the tail indices of their absolute values. Table 2 provides the estimates and the corresponding 95% confidence intervals for the tail indices of these returns in the pre- and post-crisis periods. The fact that the confidence intervals do not intersect implies that the null hypothesis that the tail indices are the same in the pre- and post-crisis periods: $H_0 : \xi_{pre} = \xi_{post}$ is rejected in favor of the alternative that the post-break tail index is smaller: $H_a : \xi_{pre} > \xi_{post}$ (at 2.5% level of significance) and thus heavy-tailedness of these returns becomes more pronounced in the post-crash period.

As indicated above, except for the above individual cases, the 95% confidence intervals for the tail indices in (1)-(3) do intersect. Thus, the hypothesis on equality of the tail indices in the pre-crisis and post-crisis periods: $H_0 : \xi_{pre} = \xi_{post}$ is not rejected in favor of the alternative $H_a : \xi_{pre} \neq \xi_{post}$ at 5% signficance level. Thus, the returns' tail indices in the pre-break and post-break periods appear to be statistically indistinguishable, and there is no evidence of statistically significant change in the tail indices due to the 2015 crash for most of the returns.

Table 2: Stocks with Significant Variation of Tail Index Estimator on Absolute Returns

		Pre-Crash			Post-Crash		
Stock	Truncation(%)	N_{Pre}	$\hat{\xi}_{Pre}$	$CI_{95\%,Pre}$	N_{Post}	$\hat{\xi}_{Post}$	$CI_{95\%,Post}$
000726 CH Equity	0.15	721	3.65	(2.68 , 4.62)	951	2.04	(1.57 , 2.51)
002073 CH Equity	0.15	729	4.13	(3.04 , 5.22)	941	2.42	(1.86 , 2.98)
002433 CH Equity	0.15	746	3.89	(2.88 , 4.90)	943	2.22	(1.70 , 2.74)
002459 CH Equity	0.15	733	3.99	(2.94 , 5.04)	892	2.31	(1.76 , 2.86)
002521 CH Equity	0.15	748	3.76	(2.78 , 4.74)	944	2.07	(1.59 , 2.55)
300206 CH Equity	0.15	755	4.23	(3.13 , 5.33)	947	2.47	(1.90 , 3.04)
300258 CH Equity	0.15	762	4.15	(3.08 , 5.22)	965	2.48	(1.91 , 3.05)
600252 CH Equity	0.15	749	3.35	(2.48 , 4.22)	876	1.96	(1.49 , 2.43)
600509 CH Equity	0.15	750	3.63	(2.68 , 4.58)	923	2.07	(1.58 , 2.56)
600674 CH Equity	0.15	747	3.54	(2.62 , 4.46)	938	2.06	(1.58 , 2.54)
600757 CH Equity	0.15	658	3.53	(2.55 , 4.51)	952	2.02	(1.55 , 2.49)
600969 CH Equity	0.15	751	3.67	(2.72 , 4.62)	962	2.14	(1.65 , 2.63)
601058 CH Equity	0.15	728	3.87	(2.85 , 4.89)	888	2.22	(1.69 , 2.75)

The results in Figure 7 and Table 2 and the corresponding results for other companies in the sample emphasize the importance of the structural break in the tail indices of the returns considered. In particular, the 95% confidence intervals for the tail indices in the pre-crash period $CI_{95\%,Pre}$ in Table 2 are located on the right of 2, thus implying the tail indices greater than 2 and finite variances for the returns' time series. At the same time, the confidence intervals for the post-crash period contain the value of 2 and thus imply possibly infinite variances for the returns. In the full sample, there are 571 more stocks for which the confidence intervals for the returns' tail indices contain the value of two thus implying possibly infinite second moments and variances for the returns. As discussed in the introduction, finiteness of variances and higher moments for financial returns and exchange rates is crucial for many models in finance and economics also for applicability of widely used standard statistical and econometric approaches, including the regression and least squares methods.

5 Empirical Analysis: Heavy-Tailedness Determinants

5.1 Tail index regressions: Factors Affecting Heavy-Tailedness

The analysis of the determinants of heavy-tailedness of financial returns in this section is based on tail index regressions that relate the (estimates of) tail indices of A-share returns obtained in the previous section to several firm-specific factors. The estimated tail index regressions evaluate the effects of specific characteristics of market participants - Chinese companies - on the degree of heavy-tailedness of the companies' stock returns.

7

More precisely, we provide estimates of dependence of tail indices of A-share returns on the factors in the following model:

$$\xi_i = f(SOE_i,\ FinSect_i,\ SSECList_i,\ ForList_i,\ Liq_i,\ Size_i,\ Value_i,\ Profit_i,\ Invest_i) \qquad (9)$$

where

ξ_i is the tail index of daily returns on shares of company i,

SOE_i is a dummy that equals one if the government is one of stakeholders of company i, and zero otherwise,

$FinSect_i$ is a dummy variable that equals one if company i belongs to the financial sector,

$SSECList_i$ is a dummy variable that equals one if company i is listed on the SSE Composite Index,

$ForList_{i,t}$ is a dummy variable that equals one if company i is listed in foreign exchanges,

$Liq_{i,t}$ is a dummy variable that equals one if the volume weighted average bid-ask spread (in percentage) of company i is in the first decile $D_1(Spreadt^{VWA})$,[8]

$Size_i$ is the natural logarithm of the quarterly average market capitalization of company i,

$Value_i$ is a dummy variable that equals one if the price-to-book (P/B) ratio characterizing the value of company i is in the ninth decile $D_9(P/Bratio)$,

$Profit_i$ is a dummy variable that equals one if the return on equity (ROE) of company i is in the ninth decile $D_9(ROE)$,[9]

and $Invest_i$ is a dummy variable that equals one if the total investment to total assets (I/A) ratio of company i is in the ninth decile $D_9(I/A\ Ratio)$.[10]

The analysis is based on OLS regressions of tail index estimates $\hat{\xi}$ on company-specific factors on the right-hand size of (9).[11]

5.2 Tail index regressions: Estimation

5.2.1 Absolute returns

Table 3 presents the estimation results for tail index regressions (9) for tail index estimates $\hat{\xi}_{RS}$ for absolute returns obtained using log-log rank-size regression approach with 15% truncation level, for the whole sample, the pre-crash period and the post-crash period, respectively. The overall $F-$statistics on the nine firm-specific factors indicate their joint significance at 1% level.

We note that, in the tail regression, a positive value of the estimate of the coefficient on a particular factor implies that an increase in the factor leads, on average, to an *increase* in the *tail index* and thus *decreases* the degree of heavy-tailedness (see the discussion in Section 1.1).

[7]See Ankudinov et al., 2017, for estimates of related tail index regressions for financial markets in Russia.

[8]The volume weighted average daily bid-ask spread percentage used in the analysis characterizes how liquid an underlying asset is. The definition of the liquidity dummy Liq_i using the 10^{th} percentile of the volume weighted bid-ask spread percentage is motivated by Fama & French (2015). Ankudinov et al. (2017) regards a stock as liquid if its weighted average bid-ask spread percentage is less than 0.04 with an average number of transactions of minimum 1000 per month.

[9]The literature has also used other characteristics of a company's profitability, including the earnings yield, expected profitability, and others.

[10]The value one for $Invest_i$ thus suggests that the underlying company invests in an aggressive style.

[11]As indicated above, due to it is robustness as compared to Hill's estimates, tail index regressions (9) in the next section are based on log-log rank-size regression tail index estimates.

Table 3: Regression Analysis of Heavy-tailedness of Absolute Stock Returns on Factors

	Dependent variable:		
	Tail Index on Absolute Returns		
	Whole Period	Pre-break	Post-break
Constant	2.500***	3.065***	2.615***
	(0.083)	(0.117)	(0.107)
SOE	−0.112***	−0.119***	−0.104***
	(0.015)	(0.022)	(0.019)
FinSect	−0.099**	−0.235***	−0.021
	(0.050)	(0.075)	(0.065)
SSECList	−0.088***	−0.076***	−0.077***
	(0.015)	(0.023)	(0.019)
ForList	−0.059**	−0.153***	−0.023
	(0.027)	(0.041)	(0.034)
Liquid	0.111***	0.110***	0.175***
	(0.027)	(0.040)	(0.034)
Size	0.030***	0.011	0.002
	(0.009)	(0.014)	(0.012)
Value	0.125***	0.107***	0.162***
	(0.024)	(0.036)	(0.030)
Profit	0.083***	0.065*	0.065**
	(0.025)	(0.038)	(0.032)
Invest	−0.014	0.002	−0.049
	(0.027)	(0.040)	(0.034)
Observations	1,038	1,038	1,038
R^2	0.221	0.121	0.152
Adjusted R^2	0.214	0.114	0.144
F Statistic$_{(9,1028)}$	32.429***	15.763***	20.425***
Note:			*p<0.1; **p<0.05; ***p<0.01

The intercept appears to be highly statistically significant and rather large as compared to the coefficients on the factors in the regressions for all three periods considered.[12] This suggests that heavy-tailedness of Chinese stock returns is more influenced by general market trends rather than the individual company characteristics. This conclusion is similar to heavy-tailedness properties of the Russian stock market (see Ankudinov et al. (2017)) and to the fact that, according to the analysis in Morck et al. (2000) and Jin & Myers (2006), in emerging economies, the country-level volatility of returns turns out to be significantly higher compared to the firm-specific volatility.

[12]One should that the reported standard errors of the coefficients of the tail index regressions and the comparisons and the analysis of their significance should be considered as only indicative as the standard errors do not account for uncertainty in (the first-stage) estimation of tail indices that are used as dependent variables in the regressions. A more detailed analysis of the effects of different factors on the degree of heavy-tailedness and tail indices of financial retruns would require extensions of recently developed maximum likelihood-type methods for power laws with factor-dependent tail indices (see Wang & Tsai (2009), Ma et al. (2019), and references therein) to the case of time series. The analysis may also use robust $t-$statistic inference approaches developed in Ibragimov & Müller (2010, 2016) that do not require consistent estimation of limiting variances of estimators dealt with (e.g., those of tail index regression coefficients). Applications of these approaches are left for further research.

The regression estimates in Table 3 point to the higher level of heavy-tailedness for financial sector companies as compared to those compared to those representing the real sector of economy, with statisically significant negative values of the coeffient at the variable FinSect in regressions for the whole sample and the pre-crash period. The effect is expected as higher degree of financial leverage results in higher returns volatility of financial intermediaries. The difference between the degree of heavy-tailedness of the returns of financial sector companies as compared to those in the real sector appears to be particularly large and significant in the pre-crisis period, while it is not statistically significant in the post-crash period. According to the estimates, during the pre-crash period, the financial sector companies had on average tail indices smaller by 0.235 as compared to the real sector companies. These conclusions are apparently due to the fact that the bull market before the crash was led by financial companies and banks. We also note that statistical significance of the coefficient is in contrast to the conclusions for tail index regressions in the case of Russian stock market Ankudinov et al. (2017) where it is in general absent.

Stock liquidity appears to reduce the degree of returns' heavy-tailedness. The effect is expected since, as discussed, for example, in Ankudinov et al. (2017), all other factors held equal, actively traded stocks are expected to have lower volatility due to the regular and prompt revaluation by the market as any significant difference between a stock's fundamental and market values emerges. In contrast, for illiquid shares, rare and limited in value transactions may significantly affect their market prices. The effect of (il)liquidity on the degree of returns' heavy-tailedness is statistically significant in all the periods considered. The tail indices of the returns on stocks of the most liquid companies (with the volume-weighted bid-ask spread percentage in the 10^{th} percentile of the sample) are, on average, greater than those for less liquid companies by about 0.11 in the pre-crash period and over the whole time period considered. Illiquidity amplifies heavy-tailedness in the post-crash bear market with a larger average difference (of 0.175) between the returns on stocks of the most liquid and less companies.

The returns on stocks of larger companies have, on average, lower degree of heavy-tailedness as compared to those for smaller one. The effect of company size on the degree of heavy-tailedness is statistically significant in regressions over the whole time period considered. The above negative relation between the degree of heavy-tailedness and companies size is expected since, as discussed in Ankudinov et al. (2017), large firms are generally more stable financially due to government support for very large "too big to fail" companies, better access to capital markets and diversified sources of funding. These effects are stronger in emerging markets under financial constraints resulting from, among other factors, the lesser development of financial intermediation. Higher stability of large firms tends to reduce volatility of their stock prices.

According to the tail index regressions, state owned companies tend to have, on average, higher degree of heavy-tailedness in their stock returns as compared to other companies. On average, higher degree of heavy-tailedness is also observed for companies listed on the SSE Composite Index SSE and on foreign exchanges. These conclusions are surprising since state-owned enterprises are expected to be more stable, especially in emerging markets like China and Russia (see the discussion in Ankudinov et al. (2017)) due to government's preferential policies, low-cost funding, etc. and thus to have less volatile stock prices. In addition, listings on the SSE or foreign exchanges would be expected to reduce the degree of heavy-tailedness due to their regulatory requirements on corporate governance, financing, transparency, and other factors that favor more stable companies. The analysis of the above conclusions from the empirical results is an important problem that is left for further research.

Higher value and profitability indicated by a higher price-to-book (P/B) ratio and higher return on equity (ROE) tend to be associated with lower degrees of heavy-tailedness, with statistically significant coefficients on these regressors over all the periods considered. The analysis of the effects of these variables on the degree of heavy-tailedness and explanations for them merits a further investigation.

The investment factor does not appear to be statistically sigificant in any of the regression models considered.

5.2.2 One-side returns

Table 4: Tail index estimates for positive and negative returns

	Dependent variable:					
	Tail Index on One-side Returns					
	Left Tail			Right Tail		
	Whole Period	Pre-break	Post-break	Whole Period	Pre-break	Post-break
Constant	2.663***	4.175***	3.089***	2.189***	2.139***	2.116***
	(0.087)	(0.141)	(0.116)	(0.114)	(0.147)	(0.139)
SOE	−0.085***	−0.081***	−0.053**	−0.133***	−0.141***	−0.145***
	(0.015)	(0.027)	(0.020)	(0.020)	(0.028)	(0.024)
FinSect	0.015	−0.296***	0.114	−0.211***	−0.174*	−0.241***
	(0.052)	(0.090)	(0.070)	(0.069)	(0.094)	(0.084)
SSECList	−0.073***	−0.055**	−0.061***	−0.095***	−0.097***	−0.093***
	(0.015)	(0.027)	(0.020)	(0.020)	(0.029)	(0.025)
ForList	−0.012	−0.098**	0.014	−0.106***	−0.197***	−0.079*
	(0.028)	(0.049)	(0.037)	(0.037)	(0.051)	(0.044)
Liquid	0.108***	0.077	0.217***	0.118***	0.150***	0.131***
	(0.028)	(0.048)	(0.037)	(0.037)	(0.050)	(0.044)
Size	−0.008	−0.085***	−0.072***	0.076***	0.094***	0.080***
	(0.010)	(0.017)	(0.013)	(0.013)	(0.018)	(0.015)
Value	0.119***	0.074*	0.115***	0.079**	0.118***	0.126***
	(0.025)	(0.043)	(0.033)	(0.033)	(0.045)	(0.039)
Profit	0.044*	−0.022	0.062*	0.168***	0.099**	0.104**
	(0.027)	(0.045)	(0.035)	(0.035)	(0.047)	(0.042)
Invest	−0.036	−0.058	−0.046	−0.002	0.038	−0.060
	(0.028)	(0.048)	(0.037)	(0.037)	(0.050)	(0.044)
Observations	1,038	1,038	1,038	1,038	1,038	1,038
R^2	0.133	0.129	0.097	0.207	0.132	0.161
Adjusted R^2	0.125	0.122	0.089	0.200	0.124	0.153
F Statistic (df = 9; 1028)	17.512***	16.952***	12.250***	29.821***	17.340***	21.852***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4 presents the estimation results for analogues of tail regressions dealt with in the previous section for the right and left tail indices ξ^+ and ξ^- in power law models (1) and (2). Again, the analysis is based on log-log rank-size regression estimates $\hat{\xi}^+_{RS}$ and $\hat{\xi}^-_{RS}$ with 15% truncation regressed on firm-specific factors in (9), As in the previous section, the tail index regression estimates are provided for the whole time period considered as well as for the pre-crash and post-crash periods.

As in the case of the analysis of tail index regressions for absolute returns, the overall $F-$statistics appear to be significant in all the periods considered. In addition, similar to the previous section, high significance is observed for the intercepts in all the regressions, pointing out to the apparently greater importance of general market trends rather than the individual company characteristics for heavy-tailedness properties (and thus the likelihood and the magnitude) of both the large downfalls and large upward movements in the returns and price time series in the three periods. Importantly, a statistically significant change is observed in the intercept of the regression for left tail indices, with its point estimates declining from a rather high value of 4.2 indicating rather low degree of heavy-tailedness pre-crash to the value of 3.1.

As expected, the magnitude of essentially all the coefficients in tail index regressions estimated separately for the returns' right and left tail indices in Table 4 are smaller than those for the regressions for tail indices of absolute returns in the previous section.

The analysis of confidence intervals for the coefficients in regressions in Table 4 does not indicate their statistically significant difference for the right and left tails.

In addition, importantly, with the exception of the intercept, the confidence intervals for all the coefficients in the regressions in the pre-crash and post-crash periods intersect. This points out to absence of statistically significant changes in the regressions' coefficients due to the 2015 crash, and in the effects of firm-specific factors on heavy-tailedness properties of large downfalls and large upward movements in the return and price time series.

6 Conclusion

Many studies in the literature have focused on the analysis of stylized facts of developed financial markets, including heavy-tailedness properties of financial returns and foreign exchange rates (see, among others, the review in Cont (2001)). The empirical research mostly agrees, in particular, that, in the case of developed markets, tail indices of financial returns and foreign exchange rates lie in the interval $(2, 4)$, thus implying finite variances and infinite fourth moments. At the same time, the research on heavy-tailedness properties of emerging financial markets still remains limited.

This paper partially fills this gap in the literature by providing a detailed analysis of heavy-tailedness properties of financial returns on more than a thousand A-shares in China using the recently developed robust tail index inference approaches. Among other results, the analysis points to the tail index estimate smaller than two for several stocks considered, thus implying possibly infinite second moments and variances. These results are particularly important as infinite second moments and variances lead to the inapplicability of many classical econometric and statistic models including least-squares analysis and (auto-)correlation based approaches, while standard auto-correlation based methods require infinite fourth moments (see, among others, the discussion in Davis et al., 1998, Mikosch & Starica, 2000*a*, Cont, 2001, and Ibragimov et al., 2015.).

We further focus on the analysis of (a)symmetry and structural breaks in heavy-tailedness properties of Chinese financial markets due to the 2015 crash. The returns' distributions display some (gain/loss) asymmetry in their right and left heavy-tailedness properties that varies before and after the crash, with a relatively fatter right tail observed before the 2015 market crash (see Figure 5). However, none of the stocks shows a statistically significant difference between the degree of heavy-tailedness in the left and right tails.

In terms of the structural breaks in the tail indices, differences in the degree of heavy-tailedness in the pre-crash and post-crash periods generally exist, and we find 3 cases where left tails become heavier statistically significantly and 13 cases where the tails of the returns' large fluctuations of either sign get fatter after the break (see Figure 7 and Table 2). Importantly, the break causes fluctuations in the tail indices for most of the stocks that imply infinite second moments.

Motivated, in part, by the analysis in Fama & French (2015) and Ankudinov et al. (2017), the paper also provides a detailed regression analysis of firm-specific characteristics and attributes affecting heavy-tailedness properties of returns on their A-shares before and after the break. Important effects on the returns' heavy-tailedness properties are observed, in particular, for the momentum, liquidity, and company size. The analysis also points out to the importance of government ownership, sector affiliation, and dual listing as determinants of returns' heavy-tailedness properties.

Further research may focus on applications of $t-$statistic approaches in Ibragimov & Müller (2010, 2016) to robust inference on coefficients in tail index regressions (see Section 5.2). It may also focus on applications of the above approaches and QLR-type procedures in robust tests for structural breaks in tail indices and the coefficients at their determinants, possibly with unknown date. The research in these directions is currently under way by the authors, and will be presented elsewhere.

References

Anatolyev, S. (2019), 'Volatility filtering in estimation of kurtosis (and variance)', *Dependence Modeling* **7**, 1–23.

Ankudinov, A., Ibragimov, R. & Lebedev, O. (2017), 'Heavy tails and asymmetry of returns in the Russian stock market', *Emerging Markets Review* **32**, 200–219.

Beirlant, J., Goegebeur, Y., Teugels, J. & Segers, J. (2004), *Statistics of Extremes: Theory and Applications*, Wiley Series in Probability and Statistics, Chichester: John Wiley & Sons Ltd.

Candelon, B. & Straetmans, S. (2006), 'Testing for multiple regimes in the tail behavior of emerging currency returns', *Journal of International Money and Finance* **25**, 1187–1205.

Chang, E. C., Luo, Y. & Ren, J. (2014), 'Short-selling, margin-trading, and price efficiency: Evidence from the Chinese market', *Journal of Banking & Finance* **48**, 411–424.

Chen, Z. & Ibragimov, R. (2019), 'One country, two systems? The heavy-tailedness of Chinese A-and H-share markets', *Emerging Markets Review* **38**, 115–141.

Cho, D. D., Russell, J., Tiao, G. C. & Tsay, R. (2003), 'The magnet effect of price limits: evidence from high-frequency data on Taiwan stock exchange', *Journal of Empirical Finance* **10**(1-2), 133–168.

Chung, H. & Park, J. Y. (2007), 'Nonstationary nonlinear heteroskedasticity in regression', *Journal of Econometrics* **137**, 230–259.

Cont, R. (2001), 'Empirical properties of asset returns: Stylized facts and statistical issues', *Quantitative Finance* **1**, 223–236.

Davis, R. A., Mikosch, T. et al. (1998), 'The sample autocorrelations of heavy-tailed processes with applications to ARCH', *The Annals of Statistics* **26**, 2049–2080.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997), *Modelling Extremal Events: For Insurance and Finance*, Vol. 33 of *Applications of Mathematics*, Springer.

Fama, E. F. & French, K. R. (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**, 1–22.

Gabaix, X. (2009), 'Power laws in economics and finance', *Annual Review of Economics* **1**, 255– 293.

Gabaix, X. (2016), 'Power laws in economics: An introduction', *Journal of Economic Perspectives* **30**, 185–206.

Gabaix, X., Gopikrishnan, P. & Plerou, V. and1 Stanley, H. E. (2006), 'Institutional investors and stock market volatility', *Quarterly Journal of Economics* **121**, 461–504.

Gabaix, X., Gopikrishnan, P., Plerou, V. & Stanley, H. E. (2003), 'A theory of power-law distributions in financial market fluctuations', *Nature* **423**, 267–270.

Gabaix, X. & Ibragimov, R. (2011), 'Rank-1/2: a simple way to improve the OLS estimation of tail exponents', *Journal of Business & Economic Statistics* **29**, 24–39.

Granger, C. W. J. & Orr, D. (1972), 'Infinite variance and research strategy in time series analysis', *Journal of the American Statistical Association* **67**, 275–28.

Gu, Z. & Ibragimov, R. (2018), 'The "Cubic law of the stock returns" in emerging markets', *Journal of Empirical Finance* **46**, 182–190.

Han, H. & Park, J. Y. (2008), 'Times series properties of arch processes with persistent covariates', *Journal of Econometrics* **146**, 275–292.

Huisman, R., Koedijk, K. G., Kool, C. J. M. & Palm, F. (2001), 'Tail-index estimates in small samples', *Journal of Business & Economic Statistics* **19**, 208–216.

Ibragimov, M., Ibragimov, R. & Kattuman, P. (2013), 'Emerging markets and heavy tails', *Journal of Banking & Finance* **37**, 2546–2559.

Ibragimov, M., Ibragimov, R. & Walden, J. (2015), *Heavy-Tailed Distributions and Robustness in Economics and Finance*, Vol. 214 of *Lecture Notes in Statistics*, Springer.

Ibragimov, R. & Müller, U. K. (2010), '$t-$statistic based correlation and heterogeneity robust inference', *Journal of Business and Economic Statistics* **28**, 453–468.

Ibragimov, R. & Müller, U. K. (2016), 'Inference with few heterogeneous clusters', *Review of Economics and Statistics* **98**(1), 83–96.

Ibragimov, R., Pedersen, R. S. & Skrobotov, A. (2020), 'New approaches to robust inference on market (non-)efficiency, volatility clustering and nonlinear dependence', *Working paper, Imperial College Business School and Department of Economics, University of Copenhagen* .
URL: *https://arxiv.org/abs/2006.01212*

Ibragimov, R. & Prokhorov, A. (2017), *Heavy Tails and Copulas: Topics in Dependence Modelling in Economics and Finance*, World Scientific.

Jiang, F., Jiang, Z. & Kim, K. A. (2017), 'Capital markets, financial institutions, and corporate finance in China', *Journal of Corporate Finance* .

Jin, L. & Myers, S. (2006), 'R^2 around the world: New theory and new tests', *Journal of Financial Economics* **79**, 257–292.

Koedijk, K., Stork, P. & de Vries, C. (1992), 'Differences between foreign exchange rate regimes: The view from the tails', *Journal of International Money and Finance* **11**, 462–473.

Loretan, M. & Phillips, P. C. B. (1994), 'Testing the covariance stationarity of heavy-tailed time series', *Journal of Empirical Finance* **1**, 211–248.

Ma, Y., Jiang, Y. & Huang, W. (2019), 'Tail index varying coefficient model', *Communications in Statistics - Theory and Methods* **48**, 235–256.

McNeil, A. J., Frey, R. & Embrechts, P. (2015), *Quantitative Risk Management: Concepts, Techniques and Tools*, revised edn, Princeton University Press.

Mikosch, T. & Starica, C. (2000*a*), 'Is it really long memory we see in financial returns', *Extremes and Integrated Risk Management* **12**, 149–168.

Mikosch, T. & Starica, C. (2000*b*), 'Limit theory for the sample autocorrelations and extremes of a GARCH(1, 1) process', *Annals of Statistics* **28**, 1427–1451.

Miller, J. I. & Park, J. (2010), 'Nonlinearity, nonstationarity, and thicktails: How they interact to generate persistence in memory', *Journal of Econometrics* **155**, 83–89.

Morck, R., Yeung, B. & Yu, W. (2000), 'The information content of stock markets: Why do emerging markets have synchronous stock price movements?', *Journal of Financial Economics* **58**, 215–260.

Park, J. Y. (2002), 'Nonstationary nonlinear heteroskedasticity', *Journal of Econometrics* **110**, 383 – 415.

Quintos, C., Fan, Z. & Phillips, P. C. (2001), 'Structural change tests in tail behaviour and the Asian crisis', *Review of Economic Studies* **68**, 633–663.

Sornette, D., Demos, G., Zhang, Q., Cauwels, P., Filimonov, V. & Zhang, Q. (2015), 'Real-time prediction and post-mortem analysis of the Shanghai 2015 stock market bubble and crash', *Swiss Finance Institute Research Paper No. 15-31* .

Wang, H. & Tsai, C.-L. (2009), 'Tail index regression', *Journal of the American Statistical Association* **104**, 104–487.

Zeng, F., Huang, W.-C. & Hueng, J. (2016), 'On Chinese government's stock market rescue efforts in 2015', *Modern Economy* **7**, 411–418.