One Country, Two Systems? The Heavy-Tailedness of Chinese A- and H- Share Markets [†]

Zhimin Chen

Imperial College Business School, Swiss Finance Institute and University of Lausanne

Rustam Ibragimov

Imperial College Business School

Abstract

Chinese A- and H- share markets operate in different institutional environments (emerging/developing v.s. developed) and thus may have different tail risk properties. This paper focuses on the analysis of heavy-tailedness properties of these two markets using recently developed robust inference methods. The equality of tail indices of returns for A and H dual-listed companies cannot be rejected, and some A- and H- share returns may have infinite second moments. Their heavy-tailedness properties did not change significantly with respect to the 2008 financial crisis and the date when the corresponding company starts to be dual-listed.

Keywords: Heavy-tailedness, crises, emerging markets, China, financial returns, A- and H-share markets

1. Introduction

1.1. Background

1.1.1. Heavy tails in financial markets

Since Mandelbrot (1963), who initially noticed that stable distribution could provide a good fit with the change of cotton price, heavy-tailed distributions and their applications in economics and finance have been studied intensively by generations of economists (see Ibragimov et al. (2015) and the references therein). Their aim is to built up an alternative framework to overcome the inability of classical normal or log-normal assumption in modeling the tail behavior of many economic and financial variables (Campbell et al., 1997).

As non-normal stable distributions are too heavy-tailed to be embraced by the empirical results that point out to finite second and higher moments of returns (Blattberg and Gonedes, 1974), economists turn to use power law distributions to model the tail behavior of financial

returns. The distribution of variable R of interest (daily return of stock or stock index in this paper) is usually assumed to have a power law tail[1], such that

$$P(R > x) \sim \frac{C_1}{x^{\zeta^+}} \quad , \quad P(R < -x) \sim \frac{C_2}{x^{\zeta^-}} \quad , \quad P(|R| > x) \sim \frac{C}{x^{\zeta}}$$

for sufficiently large x's. Here, C_1, C_2, C and ζ^+, ζ^-, ζ are some positive constants, and the parameters ζ^+, ζ^- and ζ are referred to as the right tail index, left tail index and tail index respectively. The tail index ζ reflects the degree of heavy-tailedness, with small values of them corresponding to heavier tail and thus larger likelihood of extreme observations (Ibragimov et al., 2015). The tail index ζ is also called the "maximal moment exponent" (Quintos et al., 2001, p. 635) since it determines the highest order of finite absolute moment (moment $E|R|^p$ is finite if and only if $p < \zeta$). Thus, as summarized by Ibragimov (2009), the tail index ζ is of great importance for the applicability and robustness of many economic and financial models (Ibragimov, 2007; Ibragimov et al., 2008, 2011; Ibragimov and Walden, 2010; Ibragimov and Prokhorov, 2017).

Thanks to the availability of more frequent data in recent years, power law distributions have become the workhorse in studying the tail of financial distributions. A host of empirical evidence has shown that financial variables in developed economies tend to have a tail index ζ lying between two and four (among others, see the reviews in Ibragimov et al. (2015)). Cont (2001) further argued that the left and right tail behavior for equity returns[2] tend to be asymmetric, with extreme negative returns more likely to be observed than extreme positive returns (gain/loss asymmetry[3]).

The rapid growth of empirical evidence inspires economists to build up the theoretical framework that can explain the endogeneity of these stylized facts. Gabaix et al. (2006) pioneered this work in the context of developed country financial market. In their structural model, the size distribution of large market participants (institutional investors) is assumed to follow Zipf's law[4] ($\zeta = 1$), and these institutional investors strategically choose the size and time for trading to moderate transaction cost and price impact, leading to a less heavy-tailed distribution of trading volume ($\zeta = 1.5$) and a square root relationship between trading volume and price change (Gabaix, 2016). Thus, in equilibrium, the tail index $\zeta = 3$ is obtained for returns from the transformation rules of power laws[5]. More recently, Kyle and Obizhaeva (2016) provided a microfoundation for Gabaix et al. (2006) model by suggesting that the stock markets are characterized by the so-called "market microstructure invariance".

Even though there is intensive research on the tail behavior of financial variables in developed financial markets, the research on their emerging/developing counterparts is relatively scarce. Ibragimov et al. (2013) argued that financial variables in developed and emerg-

---

[1]For a general introduction of the power law distributions and tail index, see Ibragimov et al. (2015).

[2]This is not true for exchange rates. See Cont (2001) and the evidence in Ibragimov et al. (2013).

[3]Heavy tails and gain/loss asymmetry are referred to as stylized facts in financial markets by Cont (2001).

[4]For Zipf's law, see Zipf (1949); Gabaix (1999). For Zipf's law for firm sizes, see Axtell (2001); Zhang et al. (2009).

[5]In general, $\zeta_{\alpha X} = \zeta_X$ and $\zeta_{X^\alpha} = \zeta_X / \alpha$ hold for a r.v. X and a positive constant α. Thus, a tail index $\zeta = 1.5$ for trading volume and a square root (1/2 power) price impact will lead to a tail index $\zeta = 3$ for returns. See the appendix in Gabaix et al. (2006) for more details.

ing/developing countries may have different tail behaviors, since government interventions may play an important role in emerging/developing markets so that some assumptions in Gabaix et al. (2006) should be modified. This argument is supported by following evidence found in Russian stock market (Ankudinov et al., 2017) and points to the extreme financial interdependences between emerging/developing and developed countries (Aloui et al., 2011; Mensi et al., 2014).

To further examine the tail behavior of emerging/developing country financial variables and their similarities and differences with those of developed markets, the Chinese A- share and H- share markets may be the natural and ideal experimental field.

1.1.2. Chinese A- and H- share markets

Broadly speaking, there are three stock exchanges in China: the Shanghai Stock Exchange (SSE), the Shenzhen Stock Exchange (SZSE) and the Hong Kong Stock Exchange (HKEx). The SSE and SZSE are operating in Chinese mainland, with the majority of Chinese stocks listed on them. They constitute the so-called A-share market. The HKEx is operating in Hong Kong, and shares of Chinese mainland-incorporated enterprises that are traded on it are generally known as H-shares. The overlaps of the issuers in A- and H- share markets are so-called "dual-listed" or "cross-listed" companies.

For these dual-listed companies, their A- and H- shares are traded at different prices even with same dividends and voting rights (Liu and Timmermann, 2013), which is a puzzle in Chinese stock market (Fernald and Rogers, 2002) and attracts a lot of research interest (among others, see the reviews in Zhang and Zhang (2017)). More recently, some studies (Su et al., 2007; Cai et al., 2011; Chen and Zhu, 2015) noted cointegration relationships between the prices of A- and H- shares for dual-listed companies. Liu and Timmermann (2013) further argued that a pairs trading strategy may exploit these cointegration relationships. In their empirical example, Liu and Timmermann (2013) showed that their strategy can be significantly profitable for some dual-listed Chinese banks. One should note, however, that the applicability of their econometric models depends on heavy-tailedness properties (e.g., finite variance and finite higher moments) of A- and H- share markets which, to the best of my knowledge, have not been systematically studied.

Under the famous "One country, two systems" principle, Chinese A- and H- share markets are subject to different institutional environment (emerging/developing v.s. developed). Firstly, Ng and Wu (2007) pointed out that A-share market is dominated by individual investors with limited trading experience, while H-share market behaves more like a developed country stock market where institutional investors act as the main market participants[6]. This suggests that the size distribution of market participants may be quite different for A- and H-share market. Secondly, inexperienced individual investors in A-share market are less likely to trade strategically compared to institutional investors in H-share markets (Hertz, 1998). Thus, the heavy-tailedness of trading volume and the volume-return relationship may be different

---

[6]This is because Chinese A- and H- share market are quite segmented. For a long time, A-share market opened exclusively to investors in Chinese mainland, and it was only available to international investors through Qualified Foreign Institutional Investor (QFII) scheme which is highly restricted by regulations and quota (Cai et al., 2011). This situation has somewhat changed since the Chinese government launched a series of market access schemes such as the Shanghai-Hong Kong Stock Connect program in 2014 and the Shenzhen-Hong Kong Stock Connect program in 2016.

for these two markets. Thirdly, A-share market is subject to active government interventions. Huang et al. (2016) showed that the Chinese government directly intervened in the stock market during the 2015 Chinese stock market crash through the trading activities of some quasi-government entities (so-called "national teams"). Brunnermeier et al. (2017) pointed out that these government interventions will mitigate price volatility, but may at the cost of reduced market information efficiency. The actions of government further distinguishes the A- and H- share markets with respect to the assumptions in Gabaix et al. (2006).

According to Gabaix et al. (2006), the institutional differences above will lead to divergent tail behavior for A- and H- share market. This relationship, however, has not been tested to the best of my knowledge.

1.2. Research Problems and Contributions

This paper provides a systematic approach to examining the tail behavior of A- and H-share markets. The main research problems and corresponding contributions of this paper are as follows.

Firstly, this paper analyzes whether the standard econometric models and investment strategies such as pairs trading (Liu and Timmermann, 2013) are applicable in A- and H-share markets. The analysis is based on the tail index estimation results obtained from inference method suggested by Gabaix and Ibragimov (2011). The theoretical and numerical results in the paper point out to robustness and favorable finite sample performance of the approaches. To the best of our knowledge, the estimation methods have not been applied to Chinese A- and H- share markets.

Secondly, this paper analyzes whether different institutional environment for A- and H-share market will lead to divergent tail behavior. The robust inference method proposed by Ibragimov and Müller (2016) is employed to test for the equality of tail indices of A- and H-shares. This is the first time that this robust inference method is applied to studies on heavy tails. In addition, the analysis is comprehensive, since all A and H dual-listed pairs (including delisted ones) with enough sample size are subject to the equality test in the paper. The empirical results may also inspire future research on the theoretical explanation of the similar or divergent tail behavior generated by different institutional environment.

Thirdly, this paper studies the (in)existence of gain/loss asymmetry for A- and H- share markets, by analyzing the estimation results for left and right tail indices. This analysis is of importance for investors whose portfolio has both long and short positions in these markets[7].

Fourthly, this paper examines the (in)stability of tail distributions with respect to some events of interest for A- and H- share markets. The (in)stability of heavy-tailedness is of importance for asset pricing and return predictability since increased evidence (Kelly and Jiang, 2014; Bollerslev et al., 2015; Acemoglu et al., 2017) has shown that tail risks are compensated in returns. The results can also be used by policy-makers for impact assessment and policy evaluation.

The conclusions of this paper may be of interest for investors, risk managers, financial regulators, and other professionals and policy-makers who are related to A- and H- share markets.

---

[7]Many investment strategies involve both long and short positions in these two markets. See Gatev et al. (2006) for pairs trading strategy and Liu and Timmermann (2013) for its performance in A- and H- share markets.

4

In addition, the comparisons of the heavy-tailedness properties of A- and H- share markets will contribute to the studies of the different tail behaviors of developed and emerging/developing country financial markets.

1.3. Outline of the Paper

The rest of the paper is divided into the following sections: Literature Review, Data and Methodology, Empirical Analysis, and Conclusions. Section Two reviews current literature and analyses the merits and demerits of each potential method that can be employed in the paper. Section Three provides a detailed description of the sample data and explains the methodology used to generate the results. Section Four presents the empirical results. Section Five discusses the implications of the results and areas in need of further research. Estimation results can be found in the Appendix.

2. Literature Review

2.1. Empirical evidence on heavy tails in finance

A number of articles have shown that power law distributions are a type of "universality" for stock market quantities since the distributions of equity returns, trading volume and other financial variables in different stock markets (both developed and emerging/developing markets) tend to behave in the same power-like way (Gabaix, 2016). Within the setting of power law distributions, Ibragimov et al. (2013) further advocated a dichotomy between developed and emerging/developing markets since the tail indices ζ may be pronouncedly different for these two kinds of markets.

In the context of developed economies, numerous studies have reported the estimates of tail index $\hat{\zeta}$ for returns on various stocks and stock indices (Jansen and De Vries, 1991; Loretan and Phillips, 1994), and most of these estimates are higher than two and less than four. Using larger data sets[8], some "econophysicists" (Gopikrishnan et al., 1999, 2000; Plerou et al., 2001) examined the "scaling behavior" in stock markets and noted that tail index ζ for short-term returns is around 3, which is referred to as the "inverse cubic law of returns" (Gabaix, 2009).

In terms of emerging/developing markets, robust empirical results are not well-established. Jondeau and Rockinger (2003) explored the tail behaviors of 20 stock indices covering both emerging/developing and developed markets, but their approach is based on the generalized extreme value (GEV) distribution which implicitly assumes that the return series are $i.i.d.$ Thus, the applicability of their approach is questionable for dependent financial time series. Assaf (2009) reported Hill estimates for daily returns of Middle East and North Africa equity markets, with confidence intervals calculated by 100 bootstrap simulations. His method, also, may not be appropriate under dependence. In terms of Chinese stock market, Yan et al. (2005) observed that daily returns for 104 Chinese stocks display a power-like tail but the tail indices were inconsistent with the inverse cubic law. Zhang et al. (2007) examined the tail behavior of one-minute returns for Chinese SSE index, finding that the inverse cubic law only holds when the near-closure returns are removed. Their conclusions are all obtained from Hill's estimation approach which may not be robust for emerging/developing country financial variables.

---

[8]These data sets include high-frequency data, with time interval ranging from several seconds to several days.

2.2. Robust inference for the heavy-tailedness

2.2.1. Inference on tail index

Numerous approaches have been proposed to inference on the tail index ζ. Hill et al. (1975), Pickands (1975) and their followers advocated the use of non(semi)-parametric estimation methods, while many statisticians argued that parametric techniques that are based on the Peak Over Thresholds (POT) theory or Block Maxima method have a more solid theoretical foundation (see the reviews in Embrechts et al. (1997)) . Among these available methods, Hill's estimation and log-log rank-size estimation are the most popular ones.

The main difficulty for the implementation of Hill's estimation comes from the choice of the truncation level, since its small-sample properties and performance under dependence are sensitive to the threshold (Embrechts et al., 1997). To overcome this difficulty, Embrechts et al. (1997) suggested the use of Hill's plot (see Drees et al. (2000)) in applications to demonstrate the robustness of the threshold selected. Theoretically, sophisticated methods using bootstrap procedure (Hall, 1990; Danielsson et al., 2001) or quantile driven technique (Danielsson et al., 2016) have been developed to select the optimal truncation level. In addition, Huisman et al. (2001) noted that the small sample bias can be reduced by using the weighted-average of Hill's estimators for different truncation levels.

Following Gabaix and Ibragimov (2011), the log-log rank-size estimation (with $\gamma = \frac{1}{2}$ in the regression $\log(\text{Rank} - \gamma) = a - b\log(\text{Size})$) has become popular in applications (Gabaix and Landier, 2008; Eaton et al., 2011; Acemoglu et al., 2012; Keller and Yeaple, 2013). According to the numerical results in Gabaix and Ibragimov (2011), the approach performs well under different dependence settings, including heavy-tailed dependent GARCH processes, and provides the correct standard errors (and thus the correct confidence intervals) for the tail index ζ. The approach has a favorable performance even when the heuristic rule of truncation level selection (for instance, 5% or 10% of the total sample size) is employed (Ibragimov et al., 2013). Except for Ankudinov et al. (2017) in the Russian case, this robust method has not been applied to emerging country stock market yet.

2.2.2. Testing for the equality of tail indices

To compare the degree of heavy-tailedness or examine the asymmetry of return distributions, numerous studies have tested the equality of tail indices, either between different assets or between left and right tails for the same asset. Koedijk et al. (1990) constructed a Q statistic by exploiting the asymptotic normality of Hill's estimator to test the equality of tail indices for several currency returns, with the same method applied to the asymmetric test of stock returns by Jansen and De Vries (1991). Jondeau and Rockinger (2003), by the likelihood-ratio test, found that the equality of the tail indices cannot be rejected either cross counties within a geographical group or between left and right tails for stock indices. In contrast, Hartmann et al. (2004) reported the asymmetry of left and right tail index in French and US stock markets by constructing a statistic that is based on bootstrap and the asymptotic normality proved by Hall (1982). More recently, Ibragimov et al. (2013) documented evidence of symmetry in foreign exchange markets by analyzing the heavy-tailedness properties of upward and downward moves for several developed and emerging country exchange rates.

Ibragimov and Müller (2010, 2016) have developed a series of inference methods that can be employed in the equality test of tail indices[9]. Ibragimov and Müller (2010) proposed a t-statistic approach to inference on an arbitrary scalar parameter[10], and Ibragimov and Müller (2016) extended the t-statistic approach to two-sample case by exploiting the Behrens–Fisher statistics that can be used to test the equality of means. Ibragimov and Müller (2010, 2016) further argued that the underlying assumptions of their methods are relatively moderate (asymptotic normality of group estimators) compared to other available methods.

2.2.3. Analysis of structural breaks in tail index

Since the distribution of economic and financial variables can be affected by political or macroeconomic shocks (Ibragimov and Ibragimov, 2017), the assumption of a constant tail index may be questionable especially for a long sample period. Motivated by the collapse of Bretton Woods, Koedijk et al. (1992) compared the tail index estimates of several Latin American currency returns for different exchange rate regimes and found that tail indices under flexible exchange rate regime are significantly higher than others. Similarly, Loretan and Phillips (1994) analyzed the constancy of tail index by artificially splitting the sample into two periods and estimating the tail index for each period respectively. More recently, Ibragimov et al. (2013) detected structural breaks of the heavy-tailedness for both developed and emerging markets by comparing the confidence intervals for tail index ζ constructed from pre- and post- 2008 financial crisis periods. These studies are all based on a single "known" breakpoint that is exogenously selected (Candelon and Straetmans, 2006).

To detect a structural break when the breakpoint is "unknown", Quintos et al. (2001) performed recursive, rolling and sequential tests to explore the tail behavior of stock market index for Thailand, Malaysia and Indonesia during the Asian financial crisis. They found that the structural breaks are detectable and the breakpoints picked by the tests have economic sense. Candelon and Straetmans (2006) extended Quintos et al. (2001) by allowing for multiple breakpoints. They identified multiple structural breaks in the tail index of six emerging Asian currencies, and these breaks are in line with the policy changes. With respect to the structural breaks caused by new trading rules or trading technology, Galbraith and Zernov (2004) noted significant changes in the tail behavior of U.S. stock market that are related to the introduction of circuit breakers or the prevelance of program trading.

Kelly and Jiang (2014), more recently, suggested a "dynamic power law" by assuming assets with different levels of firm-specific tail risk are governed by the same tail risk process. They utilized pooled cross-section data in each time interval to estimate the tail indices, which is a cause for concern since the dependence among returns (so-called "contagion effect") may lead to biased tail index estimates. The contagion issue should be dealt with care when implementing their method[11].

---

[9]The equality test is often conducted in the analysis of treatment effect. Ibragimov and Müller (2016) also advocated that their approach can be used in test for structural breaks.

[10]Their methods need to divide the samples into several groups (clusters, blocks). This blocking method coincides with the approach suggested by Brillinger (1973) in the homogeneous case.

[11]To mitigate the bias caused by contagion, Kelly and Jiang (2014) resorted to the Fama and French's three-factor model to remove the common return factors.

3. Data and Methodology

3.1. Data Description

Daily logarithmic return series for 2 pairs of stock indices and 93 pairs of stocks are analyzed in this paper, with the maximum[12] sample period from 3 January 2000 to 31 July 2017. Trading suspension dates are excluded. The daily logarithmic return series r_t is calculated from the adjusted close price series P_t with

$$r_t = \ln P_t - \ln P_{t-1}. \tag{1}$$

This sample contains over 500000 data points, including almost all A and H dual-listed companies (even including the delisted ones to avoid the "survivorship bias"[13]) and important periods such as the 2008 financial crisis and the 2015 Chinese stock market crash[14]. All data used is from Wind China[15].

3.1.1. Data of stock indices

Two pairs of stock indices are investigated to get a general view of the heavy-tailedness properties for A- and H- markets. The first pair of stock indices is Shanghai SE Composite Index (SSEC)[16] and Hang Seng Index (HSI) which are generally considered as the most important stock market index for Chinese mainland and Hong Kong respectively. This pair of stock indices reflects the general market environment where A- and H- shares are listed[17]. The sample period for this pair of stock indices is from 3 January 2000 to 31 July 2017.

The second pair of stock indices is Hang Seng China AH (A) Index (AHXA) and Hang Seng China AH (H) Index (AHXH)[18]. They include the most liquid A and H dual-listed companies and tracks their performance in A- and H- share markets respectively based on the free-float methodology[19]. The sample period for this pair of stock indices is from 3 January 2006 (the date when they start to be compiled) to 31 July 2017.

3.1.2. Data of A and H dual-listed pairs

According to the Wind China, there are 95 A and H dual-listed companies by 31 July 2017. Three of them are not included into the sample due to small sample sizes. To avoid

---

[12]For those stocks and stock indices that are listed or compiled later than 3 January 2000, the start point of the sample period is the date when they started to be listed or compiled, see Table A6 in the Appendix.

[13]The survivorship bias suggests that overlooking failed companies may lead to a biased conclusion. See Brown et al. (1992); Carpenter and Lynch (1999).

[14]For more details about the Chinese stock market crash, see Huang et al. (2016).

[15]Wind China is commonly regarded as a Chinese equivalent to Bloomberg or Thomson Reuters but focusing on Chinese markets. See http://www.wind.com.cn/en/product/product.html.

[16]With just few exceptions, the majority of A-shares of dual-listed companies are listed on the Shanghai Stock Exchange (SSE) so that Shanghai SE Composite Index (SSEC) is a desirable indicator.

[17]The SSEC that is based on the full-market capitalization methodology (e.g., including both tradable and non-tradable shares), is often condemned for not truly reflecting the overall stock market condition. To address this concern, we check the robustness of the conclusions by including in the study another pair of stock indices based on the free-float methodology (and calculated using tradable shares only).

[18]It is worth mentioning that the Hang Seng Indexes Company changed the name of the two indices to "Hang Seng Stock Connect China AH (A) Index " and "Hang Seng Stock Connect China AH (H) Index" recently, but since our sample set comes from the pre-change period, we still use the pre-change name here.

[19]For the index methodology, see HSIndexes (2017).

the "survivorship bias", HKEx Fact Books[20] from 2000 to 2016 are checked to investigate the delisted companies that were once dual-listed. Three such companies are found and one of them (Jingwei Textile) is added back to the sample[21].

The start point of the sample period for a stock of interest is the later of 3 January 2000 and the date when this stock starts to be listed[22]. The end point of the sample period is 31 July 2017, except for the delisted one which sample ends at the date when it was delisted (28 December 2015).

3.2. Methodology

3.2.1. Inference on tail index

As per Gabaix and Ibragimov (2011), OLS log-log rank-size method with bias reduction is employed to estimate the (left and right) tail index estimates $\hat{\zeta}^-$, $\hat{\zeta}^+$ and $\hat{\zeta}$ for each A and H dual-listed pairs.

Let $r_1, r_2, r_3, ..., r_N$ be the daily log return series for the stock or stock index of interest. And, let $r_1^-, r_2^-, r_3^-, ..., r_{N_1}^-$ and $r_1^+, r_2^+, r_3^+, ..., r_{N_2}^+$ be the series of negative and positive values of that return series respectively.

Further, let, for $n_1 < N_1$, $n_2 < N_2$ and $n < N$,

$$r_{(1)}^- \leq r_{(2)}^- \leq r_{(3)}^- \leq ... \leq r_{(n_1)}^- < 0 \tag{2}$$

be increasingly ordered smallest returns, and

$$r_{(1)}^+ \geq r_{(2)}^+ \geq r_{(3)}^+ \geq ... \geq r_{(n_2)}^+ > 0 \tag{3}$$

be decreasingly ordered largest returns, and

$$|r|_{(1)} \geq |r|_{(2)} \geq |r|_{(3)} \geq ... \geq |r|_{(n)} \tag{4}$$

be decreasing ordered largest absolute returns.

The log-log rank-size regression estimates $\hat{\zeta}^-$, $\hat{\zeta}^+$ and $\hat{\zeta}$ of tail index ζ^-, ζ^+ and ζ are the OLS estimates $\hat{b}^-$, $\hat{b}^+$ and $\hat{b}$ from the regressions

$$\log(t - \frac{1}{2}) = a^- - b^- \log|r_{(t)}^-|, \qquad t = 1, 2, 3, ..., n_1, \tag{5}$$

$$\log(t - \frac{1}{2}) = a^+ - b^+ \log|r_{(t)}^+|, \qquad t = 1, 2, 3, ..., n_2, \tag{6}$$

and

$$\log(t - \frac{1}{2}) = a - b \log|r|_{(t)}, \qquad t = 1, 2, 3, ..., n. \tag{7}$$

In addition to bias-reduction properties of the rank shift of 1/2, Gabaix and Ibragimov (2011) also proved that the correct standard error of the estimator $\hat{\zeta}$ is given by $S.E. = \sqrt{\frac{2}{n}}\hat{\zeta}$. Thus

---

[20]HKEx Fact Books can be downloaded from the website of HKEx, see HKEX (2017).

[21]The other two companies are excluded due to small sample size or unavailable data.

[22]See Table A6 in the Appendix for their date of listing.

the 95% correct confidence intervals for the true tail indices ζ^-, ζ^+ and ζ are

$$\left(\hat{\zeta}^- - 1.96 \times \sqrt{\frac{2}{n_1}} \hat{\zeta}^- \quad , \quad \hat{\zeta}^- + 1.96 \times \sqrt{\frac{2}{n_1}} \hat{\zeta}^- \right), \tag{8}$$

$$\left(\hat{\zeta}^+ - 1.96 \times \sqrt{\frac{2}{n_2}} \hat{\zeta}^+ \quad , \quad \hat{\zeta}^+ + 1.96 \times \sqrt{\frac{2}{n_2}} \hat{\zeta}^+ \right), \tag{9}$$

and

$$\left(\hat{\zeta} - 1.96 \times \sqrt{\frac{2}{n}} \hat{\zeta} \quad , \quad \hat{\zeta} + 1.96 \times \sqrt{\frac{2}{n}} \hat{\zeta} \right). \tag{10}$$

It is worth highlighting that the difference between N_1, N_2 and N (thus the difference between n_1, n_2 and n) should be dealt with care. Further, to ensure the regression is based on a relatively large sample[23], only result with N (N_1, N_2) larger than 1000 is reported.

Throughout this paper, $n \approx 5\%N$ and $n \approx 10\%N$ ($n_1 \approx 5\%N_1$ and $n_1 \approx 10\%N_1$, $n_2 \approx 5\%N_2$ and $n_2 \approx 10\%N_2$) are used for the truncation levels. This heuristic rule is widely used in applications and seems to have a robust performance (among others, see Ibragimov et al. (2013)). To check the sensitivity of tail index estimate to different truncation levels, a log-log rank-size version of Hill's plot (the plot of the tail index estimates $\hat{\zeta}$ and corresponding confidence intervals CI against different truncation levels) is provided for some cases.

Further, although this paper mainly focuses on the applications of log-log rank-size estimation, a number of results are compared to those obtained using Hill's estimation procedure for illustration. Hill's estimator $\hat{\zeta}_{Hill}$ is given by[24]

$$\hat{\zeta}_{Hill} = \frac{n}{\sum_{t=1}^{n} (\log |r|_{(t)} - \log |r|_{(n+1)})}, \tag{11}$$

with the standard error $S.E._{Hill} = \frac{1}{\sqrt{n}} \hat{\zeta}_{Hill}$ and the 95% corresponding confidence interval given by

$$\left(\hat{\zeta}_{Hill} - 1.96 \times \sqrt{\frac{1}{n}} \hat{\zeta}_{Hill} \quad , \quad \hat{\zeta}_{Hill} + 1.96 \times \sqrt{\frac{1}{n}} \hat{\zeta}_{Hill} \right). \tag{12}$$

3.2.2. Testing for the equality of tail indices of A and H dual-listed pairs

The robust inference method suggested by Ibragimov and Müller (2016) is employed to conduct the equality test. Since this method needs to partition the whole sample into several groups, only the equality of tail indices ζ_A and ζ_H of A and H dual-listed pairs is tested to ensure there are enough data points in each groups. Further, the the start point of the sample period for a dual-listed pair is set to be the later of the time when this company starts to be listed as A- and H- share. This is to make sure that the equality is tested over the same period.

The empirical strategy is as follows[25]:

---

[23]For a small sample, otherwise, Huisman et al. (2001) method should be employed.
[24]See Embrechts et al. (1997); Ibragimov et al. (2015) for a review of Hill's estimator.
[25]See the empirical strategy summarized in Ibragimov and Müller (2016) for a general description.

(i) Partition the samples of A- and H- share returns into q_A and q_H isometric groups chronologically[26], with $q_A = q_H = 2$ or 4.[27]

(ii) Conduct log-log rank-size estimation using group j of i-share returns only to obtain $\hat{\zeta}_{i,j}$, $j = 1,..,q_i$, $i = A, H$.

(iii) Compute $\bar{\hat{\zeta}}_i = q_i^{-1} \sum_{j=1}^{q_i} \hat{\zeta}_{i,j}$ and $S_i^2 = (q_i - 1)^{-1} \sum_{j=1}^{q_i} \left(\hat{\zeta}_{i,j} - \bar{\hat{\zeta}}_i \right)^2$, $i = A, H$.

(iv) Compute Behrens-Fisher statistic

$$BF = \frac{\bar{\hat{\zeta}}_A - \bar{\hat{\zeta}}_H}{\sqrt{\frac{S_A^2}{q_A} + \frac{S_H^2}{q_H}}}. \tag{13}$$

(v) Reject $H_0 : \zeta_A = \zeta_H$ at level α if $|BF| > cv(\alpha, min(q_1, q_2) - 1)$, where $cv(\alpha, min(q_1, q_2) - 1)$ is the two-sided critical value of the Student-t distribution with $min(q_1, q_2) - 1$ degrees of freedom of level $\alpha = 5\%$.[28]

3.2.3. Analysis of the asymmetry of left and right tails

The analysis of the asymmetry of left and right tails is based on comparing the confidence intervals CI^- and CI^+ for the true left and right tail indices ζ^- and ζ^+. If these confidence intervals do not intersect, the tail indices ζ^- and ζ^+ are statistically distinguishable, implying that the likelihood of large upward and downward fluctuations are pronouncedly different. In addition, some other evidence can be found from the inequality of the point estimates $\hat{\zeta}^-$ and $\hat{\zeta}^+$.

3.2.4. Analysis of structural breaks in tail index

The analysis of structural break with a "known" breakpoint is conducted in this paper. Since this analysis needs to split the whole sample period, only the structural break of tail index ζ is tested to ensure there are enough data points in each sub-period.

For a "known" breakpoint of interest (the date of financial crisis for stock indices or the date of being dual-listed for stocks in this paper), the log-log rank-size estimation in section 3.2.1 is employed for both the pre- and post-break period, with the resulting tail indices denoted as $\hat{\zeta}_{pre}$ and $\hat{\zeta}_{post}$ and corresponding 95% confidence intervals denoted as CI_{pre} and CI_{post}.

The confidence intervals CI_{pre} and CI_{post} are compared to detect the structural break. If these confidence intervals do not intersect, the tail indices ζ_{pre} and ζ_{post} are statistically distinguishable, implying that the degree of heavy-tailedness and therefore the likelihood of large fluctuations change pronouncedly after the breakpoint. In addition, some other evidence can be found from the variation of the point estimates $\hat{\zeta}$.

---

[26]Here implicitly assumes that the q_A and q_H groups provide approximately independent and Gaussian information about ζ_A and ζ_H respectively. This assumption may be suspectable, but it is still weaker than assumptions of other available methods (Ibragimov and Müller, 2016).

[27]Indeed, q_A and q_H do not need to be identical. The identity assumed here is just for simplicity. q_A and q_H can also be larger than 4 if there are enough data.

[28]As per Ibragimov and Müller (2016), 10% is also valid since $2 \leq q_A, q_H \leq 14$.

[Ibragimov and Müller](2016) method in section 3.2.2 can also be employed in tests for structural breaks. This method is not used here and left for future research since it needs to further partition the sub-periods.

4. Empirical Analysis

Throughout this Section, for brevity, the superscript A is used for the stock or stock index that is related to A-share market, with the superscript H used for its counterpart related to H-share market within the stock or stock index pair. Estimation results obtained using log-log rank-size regression are denoted with subscript RS, and results from Hill's estimation method are denoted with subscript $Hill$. Superscript $-$ and $+$ are used for left and right tails respectively.

Failure to reject the null hypothesis $H_0 : \zeta = \zeta_0$ refers to the 5% significance level and the two-sided alternative $H_a : \zeta \neq \zeta_0$. Rejection of H_0 refers to the 2.5% significance level and the one-sided alternatives $H_a : \zeta > \zeta_0$ or $H_a : \zeta < \zeta_0$. The confidence intervals discussed in this Section are all 95% confidence intervals.

4.1. Analysis of pairs of stock indices
4.1.1. Inference on the heavy tails

The estimation results for the tail indices ζ for the two pairs of stock indices are presented in Table 1. The tail index estimates $\hat{\zeta}_{RS}$ are obtained from log-log rank-size regressions (7) with the correct standard errors $S.E._{RS} = \sqrt{\frac{2}{n}}\hat{\zeta}_{RS}$ and 95% corresponding confidence intervals as shown in (3.10). For comparison, the last three columns in Table 1 provide results obtained using Hill's estimation approach (11) with standard errors $S.E._{Hill} = \sqrt{\frac{1}{n}}\hat{\zeta}_{Hill}$ and 95% corresponding confidence intervals as shown in (12). These estimates are related to truncation levels $n \approx 5\%N$ and $n \approx 10\%N$, where N is the total sample size for the daily return series.

Table 1: Estimation results for the tail indices ζ for the two pairs of stock indices

Index	N	Truncation(%)	$\hat{\zeta}_{RS}$	$S.E._{RS}$	$95\%CI_{RS}$	$\hat{\zeta}_{Hill}$	$S.E._{Hill}$	$95\%CI_{Hill}$
SSEC	4255	5	**3.50**	0.34	(2.83 , 4.16)	**2.99**	0.20	(2.59 , 3.39)
		10	**3.10**	0.21	(2.68 , 3.52)	**2.76**	0.13	(2.50 , 3.02)
HSI	4334	5	**3.33**	0.32	(2.70 , 3.95)	**3.16**	0.21	(2.74 , 3.59)
		10	**3.13**	0.21	(2.72 , 3.55)	**2.83**	0.14	(2.56 , 3.10)
AHXA	2803	5	**4.25**	0.51	(3.26 , 5.25)	**3.54**	0.30	(2.95 , 4.12)
		10	**3.35**	0.28	(2.79 , 3.90)	**2.66**	0.16	(2.35 , 2.98)
AHXH	2845	5	**3.42**	0.41	(2.63 , 4.22)	**3.18**	0.27	(2.66 , 3.71)
		10	**3.05**	0.26	(2.55 , 3.55)	**2.55**	0.15	(2.26 , 2.85)

The results in Table 1 can be summarized as follows. For stock indices that are related to A-share market, the point estimates $\hat{\zeta}_{RS}$ lie between 3.10 and 4.30, and the the point

estimates $\hat{\zeta}_{Hill}$ lie between 2.60 and 3.60. By contrast, for stock indices that are related to H-share market, the point estimates $\hat{\zeta}_{RS}$ lie between 3.00 and 3.50, and the the point estimates $\hat{\zeta}_{Hill}$ lie between 2.50 and 3.20. All these estimates are in the interval $(2, 5)$ which is in line with the stylized fact in developed country financial markets, even through A-share market is a emerging country stock market.

The null hypothesis $\zeta = 2$ is rejected in favor of $\zeta > 2$ for all these stock indices by both the log-log rank-size regression and Hill's estimation approach. Both approaches do not reject the null hypothesis $\zeta = 3$ for SSEC, HSI and AHXH. The log-log rank-size method rejects the null hypothesis $\zeta = 4$ in favor of $\zeta < 4$ for all these stock indices at 10% truncation level, but fails to reject it for SSEC, AHXA and AHXH at 5% truncation level. The Hill's estimation approach rejects the the null hypothesis $\zeta = 4$ in favor of $\zeta < 4$ for SSEC, HSI and AHXH, but fails to reject it for AHXA at 5% truncation level. The results above imply that all these stock indices have finite first and second moments but the third and fourth moments may be infinite. In addition, according to both methods, the fourth moments of HSI are infinite.

Surprisingly, the confidence interval CI_{RS} with 5% truncation level and the confidence interval CI_{Hill} with 10% truncation level do not intersect, which points to the fact that the results obtained from log-log rank-size method and Hill's approach may be significantly different. The difference may arise from the sensitivity of the Hill's estimation procedure to dependence and sample size (Embrechts et al., 1997), but the truncation level selected is also suspectable. In order to eliminate this concern, a log-log rank-size version of Hill's plot following Ibragimov et al. (2013) is provided below.
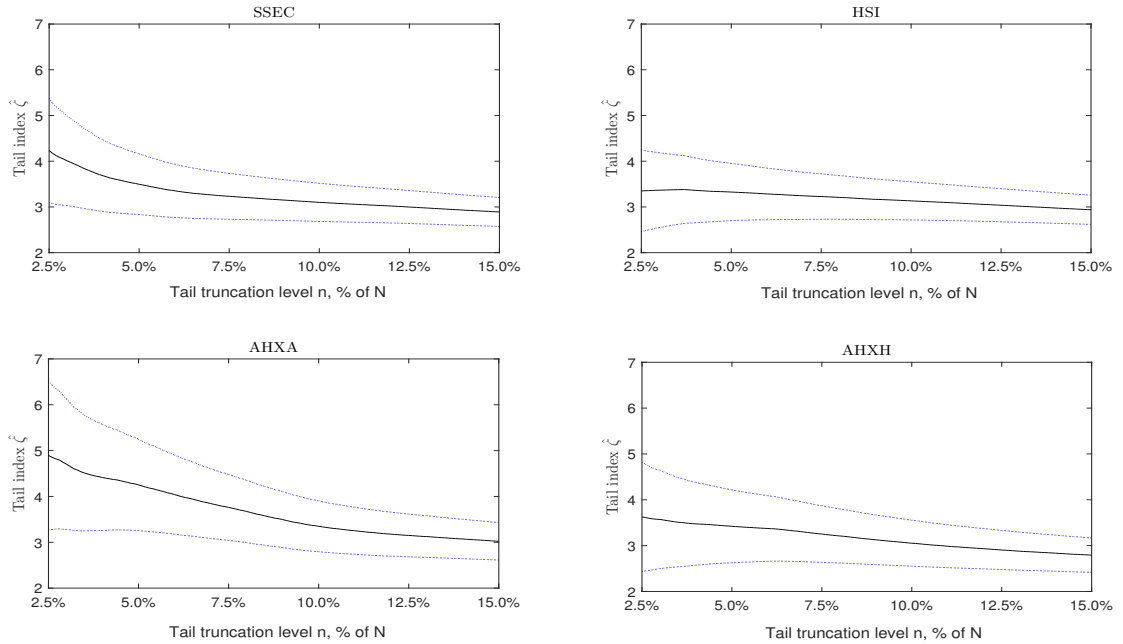
Figure 1: Log-log rank-size version of Hill's plot for tail indices ζ of stock indices

As shown in Figure 1, the point estimates $\hat{\zeta}$ is relatively stable, with the 95% confidence intervals CI_{RS} constructed for different truncation levels intersecting. This points to the statistical indistinguishability of tail indices for different truncation levels (Ibragimov et al., 2013). Moreover, the qualitative conclusions on the (in)finiteness of the first, second, third and

fourth moments generally hold except for **HSI** whose fourth moments may be finite according to the estimation with small truncation levels.

4.1.2. Testing for the equality of tail indices

As shown in Table 1, the point estimates $\hat{\zeta}^A$ and $\hat{\zeta}^H$ within stock index pairs seem to be similar and the confidence intervals CI^A and CI^H intersect. To achieve a statistically robust conclusion, a formal equality test with the null hypothesis $H_0 : \zeta^A = \zeta^H$ is conducted using Ibragimov and Müller (2016) method.

Table 2: The results of equality test for pairs of stock indices

Panel A. $q = 2$, $cv(5\%, 1) = 12.71$

Index	Truncation(%)	RS : $H_0 : \zeta^A = \zeta^H$		Hill : $H_0 : \zeta^A = \zeta^H$	
		BF statistic	p-value[1]	BF statistic	p-value
SSEC $v.s$ HSI	5	0.82	0.56	1.42	0.39
	10	0.18	0.89	0.55	0.68
AHXA $v.s$ AHXH	5	0.23	0.86	0.19	0.88
	10	0.02	0.99	0.20	0.88

Panel B. $q = 4$, $cv(5\%, 3) = 3.18$

Index	Truncation(%)	RS : $H_0 : \zeta^A = \zeta^H$		Hill : $H_0 : \zeta^A = \zeta^H$	
		BF statistic	p-value	BF statistic	p-value
SSEC $v.s$ HSI	5	0.41	0.71	0.41	0.71
	10	0.21	0.85	0.35	0.75
AHXA $v.s$ AHXH	5	1.10	0.35	0.39	0.72
	10	0.53	0.63	0.36	0.74

[1] The Student-t p-values with m degrees of freedom ($m = 1$ in Panel A and $m = 3$ in Panel B) presented here are not valid. They are just for illustration. Ibragimov and Müller (2016) have shown that the p-value rounded up to multiples of 0.001 is only valid if $BF > cv(8.3\%, m)$ for $2 \leq q_1, q_2 \leq 50$, and if $BF > cv(10\%, m)$ for $2 \leq q_1, q_2 \leq 14$.

The test results are presented in Table 2, where Panel A and Panel B shows the results for group number $q = 2$ and $q = 4$ respectively. The BF statistics are obtained from (13). The Student-t p-values are just for illustration and are not valid[29]. Since all the BF statistics are smaller than the two-sided Student-t critical values ($cv(5\%, q - 1) = 12.71$ for group number $q = 2$ and $cv(5\%, q - 1) = 3.18$ for group number $q = 4$), the null hypothesis $H_0 : \zeta^A = \zeta^H$ is not rejected using either log-log rank-size regression or Hill's estimation approach with either 5% or 10% truncation level. These results confirm the intersecting confidence intervals,

---

[29]See the footnote 1 in Table 2.

implying that the A- and H- share markets have similar tail behavior, with the probabilities of having extreme observations statistically indistinguishable.

4.1.3. Analysis of the asymmetry of left and right tails

The results of left and right tail index estimates $\hat{\zeta}^-$ and $\hat{\zeta}^+$ for the two pairs of stock indices using log-log rank-size regression are presented in Table 3. Since the estimation of left (right) tail index can only use negative (positive) values of the return series, the sample size is not as large as that in the estimation of tail index which could use both positive and negative values. To avoid small-sample bias, the results obtained form Hill's estimation are not reported.

Table 3: The left and right tail index estimates $\hat{\zeta}^-$ and $\hat{\zeta}^+$ for stock indices

Index	Truncation (%)	N_1	$\hat{\zeta}_{RS}^-$	$95\%CI_{RS}^-$	N_2	$\hat{\zeta}_{RS}^+$	$95\%CI_{RS}^+$
SSEC	5	1992	**3.77**	(2.72 , 4.81)	2263	**3.46**	(2.56 , 4.36)
	10		**2.99**	(2.41 , 3.58)		**3.34**	(2.72 , 3.95)
HSI	5	2112	**3.42**	(2.50 , 4.34)	2222	**3.21**	(2.37 , 4.06)
	10		**3.13**	(2.53 , 3.72)		**3.10**	(2.53 , 3.68)
AHXA	5	1354	**4.35**	(2.89 , 5.81)	1449	**4.25**	(2.86 , 5.63)
	10		**3.24**	(2.47 , 4.01)		**3.62**	(2.79 , 4.46)
AHXH	5	1403	**3.67**	(2.45 , 4.89)	1442	**3.17**	(2.13 , 4.21)
	10		**3.13**	(2.40 , 3.87)		**2.94**	(2.26 , 3.62)

The results in Table 3 shows that the confidence intervals of left tail and right tail intersects, implying that the left tail index and right tail index are statistically indistinguishable. These results point to the symmetry between left and right tail behavior, inconsistent with the gain/loss asymmetry that is one of the stylized facts in developed financial stock markets.

In the perspective of point estimates, all the point estimates of left tail index $\hat{\zeta}^-$ are larger than the point estimates of right tail index $\hat{\zeta}^+$ for stock indices related to H-share market. This indicates that extreme upward moves are somewhat more likely to be observed than extreme downside moves, which is the opposite of the gain/loss asymmetry in developed financial markets. For stock indices related to A-share market, the symmetry between left and right tail seems to be higher, which may indicate the government intervention[30] during stock market crash[31] .

---

[30]See Huang et al. (2016) for the intervention of Chinese government during 2015 Chinese stock market crash.

[31]Besides the government intervention, the explanations for the differences in the statistical properties of A- and H-share, including their heavy-tailedness properties, may also relate to other factors such as investor composition and the foreign exchange rate risk. The detailed investigation and quantification of the effects of all the above factors is beyond the scope of the paper and is left for future research.

4.1.4. Analysis of structural breaks in tail index

Table 4 presents the tail index estimation results with respect to the effect of 2008 financial crisis on the heavy-tailedness of stock indices. Since AHXA and AHXH have a relatively short sample when splitting the whole sample period, only results of SSEC and HSI are reported. The results in Panel A and Panel B are obtained from log-log rank-size regressions and Hill's estimation approach respectively. The "known" breakpoint of interest is 15 September 2008 which is generally seen as the start point of recent global financial crisis[32].

Table 4: The tail index estimates $\hat{\zeta}$ for pre- and post- 2008 financial crisis periods

Panel A. Log-log rank-size regression

Index	Truncation (%)	$Pre-$crisis			$Post-$crisis		
		N	$\hat{\zeta}_{RS}$	CI_{RS}	N	$\hat{\zeta}_{RS}$	CI_{RS}
SSEC	5	2099	**3.53**	(2.58 , 4.49)	2155	**3.35**	(2.46 , 4.25)
	10		**3.10**	(2.51 , 3.69)		**3.06**	(2.48 , 3.63)
HSI	5	2145	**3.85**	(2.82 , 4.88)	2188	**2.95**	(2.17 , 3.73)
	10		**3.52**	(2.86 , 4.19)		**2.83**	(2.30 , 3.36)

Panel B. Hill's estimation

Index	Truncation (%)	$Pre-$crisis			$Post-$crisis		
		N	$\hat{\zeta}_{Hill}$	CI_{Hill}	N	$\hat{\zeta}_{Hill}$	CI_{Hill}
SSEC	5	2099	**3.13**	(2.53 , 3.73)	2155	**2.86**	(2.32 , 3.39)
	10		**2.80**	(2.42 , 3.18)		**2.61**	(2.26 , 2.96)
HSI	5	2145	**3.40**	(2.76 , 4.05)	2188	**2.95**	(2.40 , 3.50)
	10		**3.17**	(2.75 , 3.59)		**2.56**	(2.22 , 2.90)

The results in Table 4 can be summarized as follows. The pre- and post- crisis confidence intervals intersect for both stock index using both estimation approach. This implies that the tail indices before and after the beginning of the 2008 financial crisis are statistically indistinguishable from each other, which is somewhat consistent with the findings of Ibragimov et al. (2013) for currency markets. But from the view of point estimates, the post-crisis point estimates $\hat{\zeta}_{post}$ are seen to be smaller than pre-crisis point estimates $\hat{\zeta}_{pre}$ for both indices no matter what estimation methods are employed. This may imply that it is more likely to observe large fluctuations after the beginning of financial crisis, which conforms to the recent

---

[32]In order to address robustness of the conclusions on absence of statistical significance of changes in the tail index following the beginning of the 2008 crisis, we have also conducted a similar analysis using other breakpoint dates like September 15, 2007. The results on tail index estimation pre- and post-that date are similar to those in Table 4 for the breakpoint date at 15 September 15, 2008 thus providing further support for the conclusions.

history of international financial market. In addition, the decrease in point estimates for SSEC is not as pronounced as that for HSI. These results may arise from the capital control and other kinds of government interventions in China. It is a general view that the 2008 financial crisis has a moderate impact on the Chinese financial market compared with that on the developed financial markets thanks to China's relatively closed and regulated financial market environment.

The effect of the 2015 Chinese stock market crash is left for further research since the samples for post-crash period are not enough at this moment.

4.2. Analysis of A and H dual-listed pairs

This subsection investigates the heavy-tailedness properties of A and H dual-listed pairs. All the results, for brevity, are placed in the Appendix. The tail index estimates $\hat{\zeta}$ in this section are all obtained from log-log rank-size regressions (7) with correct 95% corresponding confidence intervals as shown in (10). Only results with N (N_1, N_2) larger than 1000 are reported.

4.2.1. Inference on the heavy tails

The estimation results for the tail indices ζ for each A and H dual-listed pair are presented in Table A1. The results for A-shares are shown in the left panel, and the results for their H-share counterparts are shown in the right panel and are in the same row.

As shown in the left panel of Table A1, there are 29 A-shares whose confidence intervals CI constructed for the true tail index ζ using 5% and 10% truncation levels do not intersect, even though the sample sizes for them are quite large (some sample sizes are larger than 4000). For these A-shares, the confidence intervals constructed using 5% truncation levels all lie to the right of the confidence intervals constructed using 10% truncation levels. In addition, using 5% truncation level, some of them even have a point estimate $\hat{\zeta}$ bigger than 10. These abnormal results may indicate the inappropriateness of the tail truncation levels used.

To further address this problem, we use the log-log rank-size version of Hill's plot and the histograms of the daily return for the above abnormal A-shares are resorted to. For illustration, we provide the figures of four of them here (LUOYANG GLASS, COMEC, HAITONG SEC, and TIANJIN CAPITAL). These four stocks are selected from different sectors for robustness, and figures for other abnormal stocks are similar[33]. The plots and histograms are presented in Figure 2 and Figure 3 respectively.

As shown in Figure 2, these plots all have a hump shape for truncation levels smaller than 5%, and the confidence intervals constructed using truncation levels smaller than 5% are seen to lie to the right of the confidence intervals constructed using truncation levels bigger than 5%. This implies that the point estimates $\hat{\zeta}$ for small and relatively big truncation levels are statistically distinguishable. The point estimates $\hat{\zeta}$ seem to be stable for truncation levels bigger than 7.5%, with the confidence intervals intersecting with each others. Thus, the commonly-used 10% truncation level seems to provide a robust result.

Apparently, all these stocks have clustered observations in the intervals (9%, 10%] and

---

[33]The figures are available upon request.
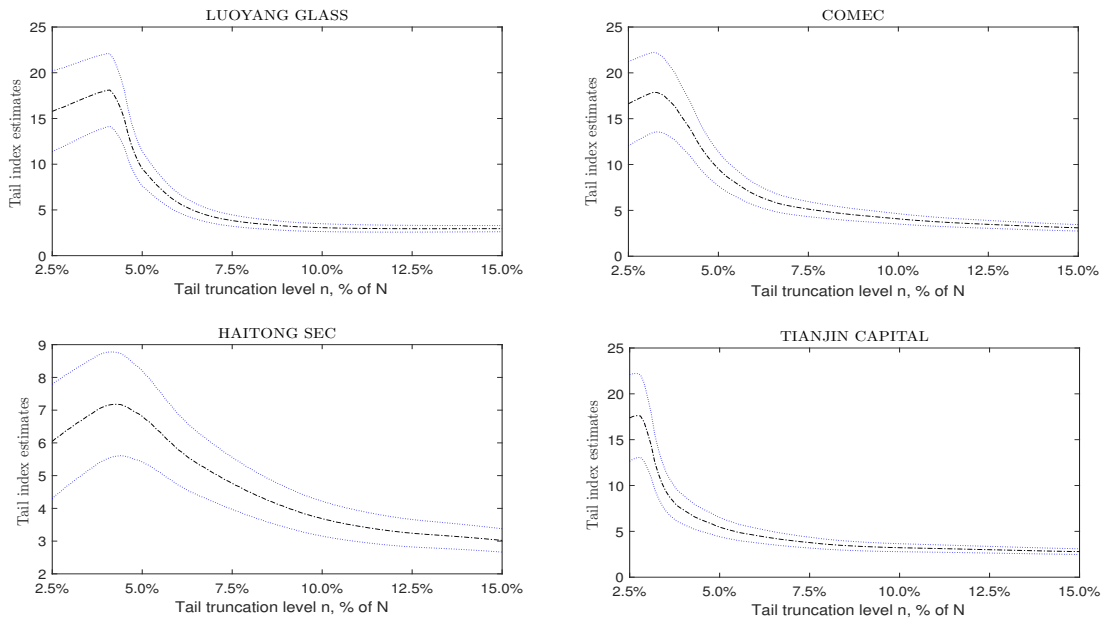
Figure 2: Log-log rank-size version of Hill's plot for tail indices ζ of four A-share stocks
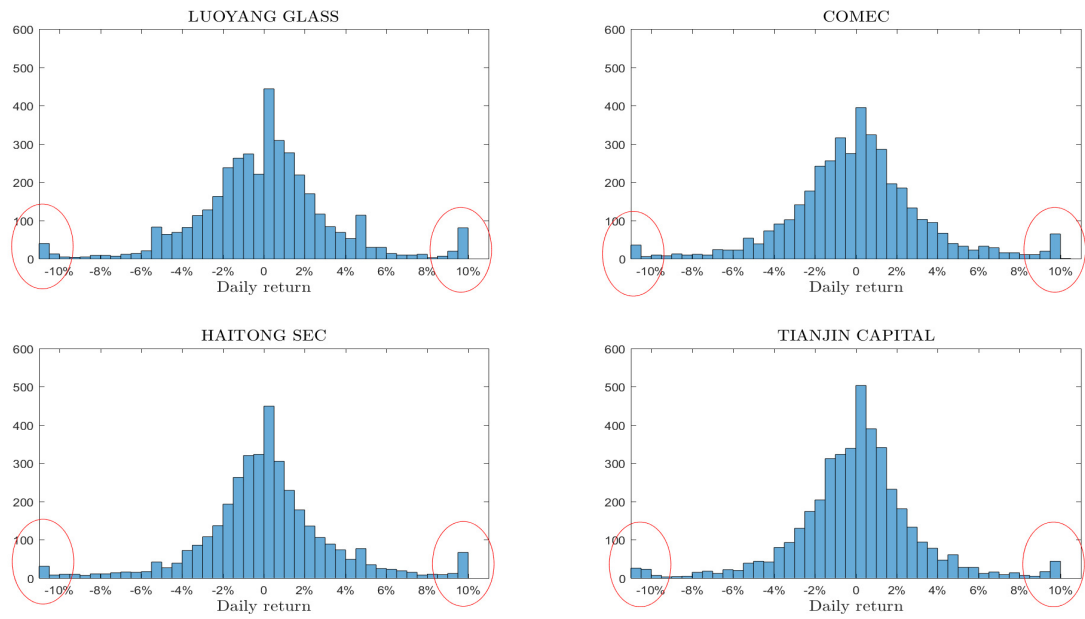
Figure 3: The histograms of the return series for four A-share stocks

18

[-10%, -9%). This may arise from the ±10% daily price limits[34] in Chinese stock market and the so-called magnet effect related to it. The magnet effect indicates that illiquidity risk and the fear of behavioral investors tend to accelerate the near limit price towards the limit (Cho et al., 2003), leading to clusters at the tails of the limited return distribution. The clusters in the tails, which seem to be deviations from power law distribution, will give rise to a quite flat truncated tail and thus a large point estimates $\hat{\zeta}$ if the truncation level selected is too small.

In contrast, as shown in Table A1, the confidence intervals constructed using 5% and 10% truncation levels intersect for H-shares, implying that the estimation results for them are stable regardless of the truncation level selected. This is consistent with similar studies for developed financial markets where the authors advocate the appropriateness of 5% and 10% truncation level used in log-log rank-size regressions[35].
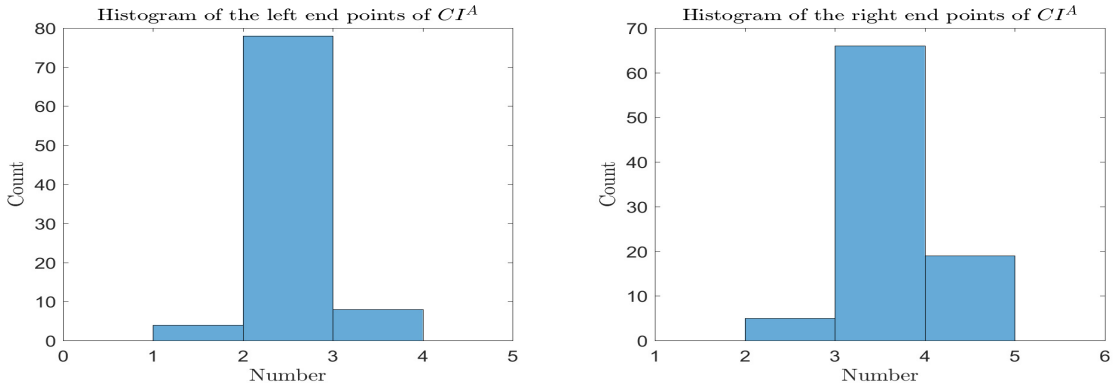
Figure 4: Histograms of left and right end points of CI^A for A shares

For illustration, Figure 4 and Figure 5 below show the histograms of the end points of the confidence intervals CI_A and CI_H respectively. The confidence intervals CI_A for A-shares in Figure 4 are constructed using the relatively appropriate 10% truncation level. In order to get a robust result, the confidence intervals CI_H for H-shares in Figure 5 are the unions of the confidence intervals obtained from 5% and 10% truncation levels. The lengths of these confidence intervals are all larger than one unit so that the conclusion can be drawn from analyzing the left and right end points respectively.

As shown in Figure 4 and Table A1, the null hypothesis $\zeta^A = 1$ is rejected in favor of $\zeta^A > 1$ for all these A-shares. The null hypothesis $\zeta^A = 2$ is not rejected for 4 of them. For most of them, the null hypothesis $\zeta^A = 3$ cannot be rejected, while the null hypothesis $\zeta^A = 4$ is rejected in favor of $\zeta^A < 4$. The null hypothesis $\zeta^A = 5$ is rejected in favor of $\zeta^A < 5$ for all of them. Thus, the conclusions on the (in)existence of first, second, third and fourth moments can be drawn as follows. All these A-shares have finite first moments, but some of them may have infinite second moments. For most of them, the third moments may be infinite and the fourth moments are infinite.

The results in Table A1 for H-shares are illustrated in Figure 5. The null hypothesis $\zeta^H = 1$ is rejected in favor of $\zeta^H > 1$ for all these H-shares. The null hypothesis $\zeta^H = 2$ is not rejected

---

[34]There are some observations beyond the -10% limit in Figure 3. This is because they are logarithmic returns calculated from adjusted close prices as shown in (3.1).

[35]See the Section 2 for the review of these studies.
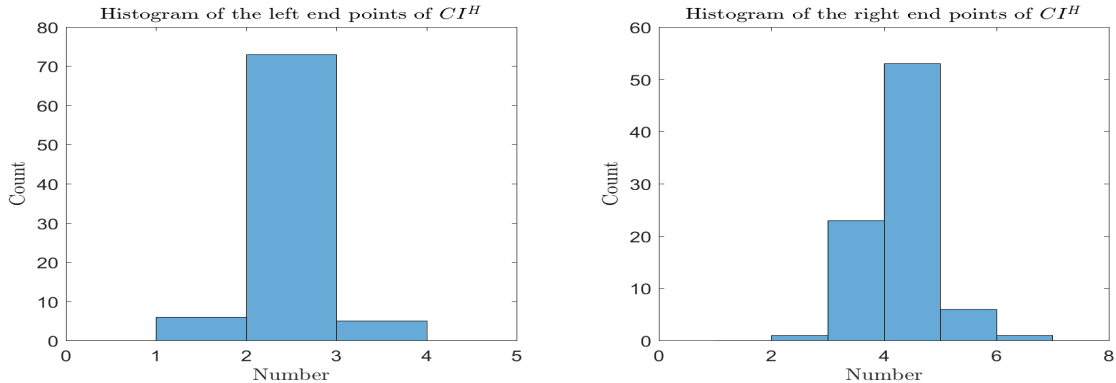
Figure 5: Histograms of left and right end points of CI^H for H shares

for 6 of them. For most of them, the null hypotheses $\zeta^H = 3$ and $\zeta^H = 4$ cannot be rejected. This implies that all these H-shares have finite first moments but some of them may have infinite second moments. In addition, for most of them, the third and fourth moments may be infinite.

4.2.2. Testing for the equality of tail indices

As in line with the analysis for pairs of stock indices in section 4.1.2, a equality test with the null hypothesis $H_0 : \zeta^A = \zeta^H$ is conducted for each A and H dual-listed pair using Ibragimov and Müller (2016) method. Within the dual-listed pairs, the longer sample period is truncated[36] to ensure that sample periods are same for both shares in the equality test. In order to mitigate the abnormal effect of 5% truncation level on the tail index estimates $\hat{\zeta}^A$, only 10% truncation level is used in the tests. In addition, only results with N^A and N^H larger than 2000 are reported[37].

The results of the equality tests are presented in Table A2. The third and fourth columns show the results for the number of groups $q = 2$, with the results for $q = 4$ shown in the same rows in the last two columns. N^A and N^H are the sample sizes used in the test for A- and H-shares respectively. Again, the invalid[38] Student-t p-values are listed for illustration. As shown in Table A2, all the BF statistics obtained from (13) are smaller than the corresponding two-sided Student-t critical values ($cv(5\%, q-1) = 12.71$ for group number $q = 2$ and $cv(5\%, q-1) = 3.18$ for group number $q = 4$), implying that the null hypothesis $H_0 : \zeta^A = \zeta^H$ is not rejected for all these A and H dual-listed pairs. There results are consistent with that obtained in section 4.1.2 where the equality is not rejected for stock index pairs.

The equality of tail indices, again, implies that the A- and H- share markets have similar tail behavior, with the likelihood of large fluctuations statistically indistinguishable. This may indicate the arbitrage between these two markets and the market access schemes launched by

---

[36]For stocks which have longer sample period compared to their counterpart within the pair, the tail index estimation results for the truncated period can be found in the right panel of Table A5, since they are also used for the analysis of structural breaks in section 4.2.4.

[37]Since Ibragimov and Müller (2016) method need to split the samples into groups, the criterion for sample size here accordingly increases.

[38]See the footnote 1 in Table 2, the Student-t p-value is only valid if the test is statistically significant with respect to the conditions described in Ibragimov and Müller (2016).

Chinese government such as the Shanghai-Hong Kong Stock Connect program and Shenzhen-Hong Kong Stock Connect program.

4.2.3. Analysis of the asymmetry of left and right tails

The estimation results for the left and right tail indices ζ^- and ζ^+ for each A and H dual-listed pair are presented in Table A3 and Table A4. The results for A-shares are shown in Table A3, while the Table A4 is for results of H-shares.

The abnormal phenomenon in section 4.2.1 can also be found in Table A3, in which some point estimates for A-shares are abnormally large with 5% truncation level and the the corresponding confidence intervals lie to the right of the confidence intervals constructed with 10% truncation level. To make a robust analysis, for A-shares, only results obtained from 10% truncation level are discussed.

As shown in Table A3 and Table A4, the confidence intervals for the left and right tail indices CI^- and CI^+ intersect for each stock regardless of A- or H- shares. This indicates that the left and right tail indices ζ^- and ζ^+ are statistically indistinguishable. These results are in line with the findings for pairs of stock indices in section 4.1.3, implying that the stylized fact of gain/loss asymmetry does not hold for either A- or H- share market.
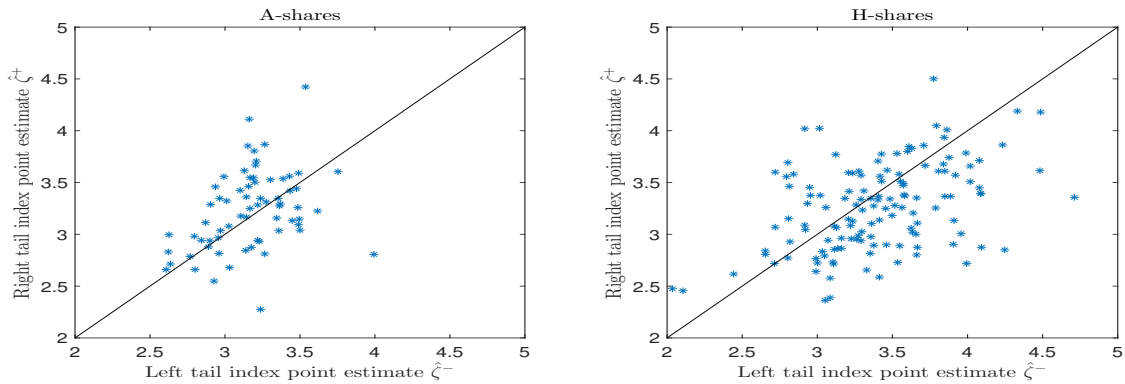
Figure 6: The relationship between the left and right tail index point estimates $\hat{\zeta}^-$ and $\hat{\zeta}^+$

For illustration, the left and right tail index point estimates $\hat{\zeta}^-$ and $\hat{\zeta}^+$ for each stocks are plotted in Figure 6. As shown in the right panel of Figure 6, a large part of the points $(\hat{\zeta}^-, \hat{\zeta}^+)$ for H-shares are seen to lie below the diagonal, which indicates that H-shares may have somewhat more pronounced heavy-tailedness in the right tail than in the left tail. This result points to the opposite of the gain/loss asymmetry. In contrast, as per the left panel of Figure 6, the points $(\hat{\zeta}^-, \hat{\zeta}^+)$ for A-shares seem to be more evenly distributed around the diagonal, which indicates a higher degree of symmetry. These conclusions, again, are in line with that in section 4.1.3 for pairs of stock indices. The higher degree of symmetry for A-shares may arise from the 10% daily price limits set for both negative and positive change, as seen in Figure 3, or the government interventions during crisis.

4.2.4. Analysis of structural breaks in tail index

For a A and H dual-listed company, the date when it starts to be dual-listed is of great importance since the interaction between A- and H- shares after that date, as seen in section

4.2.3, may have an effect on its tail behavior. In order to detect this effect, for each dual-listed company, the first dual-listed date is used as the "known" breakpoint to split the whole sample period, with the single-listed period denoted by superscript S and the dual-listed period denoted by superscript D. Table A5 presents the tail index estimation results with respect to the effect of being dual-listed on the heavy-tailedness of once single-listed shares. Again, for robustness, only 10% truncation level is used for A-shares, and only results with N^S and N^D larger than 1000 are reported.
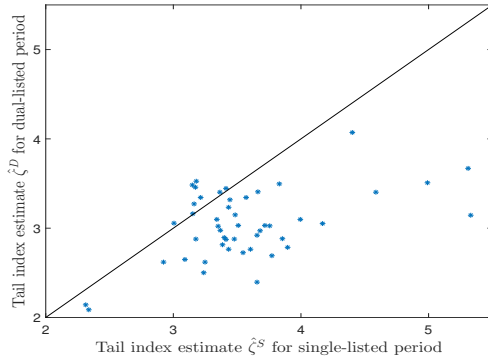
Figure 7: The relationship between tail index estimates for single- and dual- listed period

The results in Table A5 can be summarized as follows. The confidence intervals for single- and dual- listed period intersect for all these stocks, which indicates that the tail indices before and after being dual-listed are statistically indistinguishable from each other. But with respect to point estimates, the point estimates $\hat{\zeta}^S$ for single-listed period appear to be larger than the point estimates $\hat{\zeta}^D$ for dual-listed period. As shown in Figure 7, most of the points $(\hat{\zeta}^S, \hat{\zeta}^D)$ are seen to lie below the diagonal, implying that the heavy-tailedness is somewhat more pronounced after being dual-listed. This may indicate increased vulnerability caused by the interaction and arbitrage between these two markets.

5. Conclusion

5.1. Implications of the results

The estimation results indicate that some A- and H- shares may have infinite second moments, which indicates the inapplicability of many classical econometric models such as OLS regressions and autoregressive models (Ibragimov et al., 2013). Apparently, the cointegration relationship (Su et al., 2007; Cai et al., 2011; Chen and Zhu, 2015) and thus the convergence trading strategy (Liu and Timmermann, 2013) based on it may not be valid under this circumstance. In addition, for most of them, the third and fourth moments may be (or are) infinite. These results pose a challenge to the reliability of ACF inference, since sample ACFs of returns and absolute returns are poor estimators (slow rate of convergence and thus wide confidence interval) and sample ACFs of squared returns are inconsistent (non–degenerate limit distributions) if the fourth moments do not exist (Davis and Mikosch, 1998; Mikosch and Starica, 2000; Basrak et al., 2002). The invalidity of traditional approaches, in turn, highlights the importance of the robust inference methods under heavy-tails such as Ibragimov and Müller (2010, 2016).

The results of equality tests imply that different institutional environment generate similar tail behavior for A- and H- share market[39], which is a puzzle as it opposes to the implication of Gabaix et al. (2006). This puzzle may indicate the necessity of modification in models explaining heavy-tailed distributions of equity returns in developing/emerging markets. In terms of the size distribution of market participants, on the one hand, the lack of institutional investors implies a less heavy-tailed size distribution ($\zeta > 1$), on the other hand, the participation of quasi-government entities during the market crash makes the size distribution much more heavy-tailed ($\zeta < 1$). The counteraction of these two factors results in an unclear conclusion for the tail index of size distribution. In terms of the trading activities of market participants, the "herding effect"(Barber and Odean, 2000; Cont and Bouchaud, 2000) and "disposition effect" (Odean, 1998) are more pronounced for inexperienced individual investors, leading to suboptimal trading activities. In addition, the trading activities of quasi-government entities may not be economically optimal since the government's cost and benefit analysis may be beyond the scope of stock market (Huang et al., 2016). Thus, the tail index of trading volume and the volume-return relationship are also unclear. These unclear issues are of great interest for future research.

The results obtained in this paper also have economic implications. The symmetry between left and right tails highlights the impact of the symmetric price limits in A-share market and the arbitrage between these two markets. The estimation results can also be used by investors to adjust their portfolios with respect to long and short positions. In addition, the stability of tail indices implies that the 2008 financial crisis and being dual-listed do not have significant impact on the tail behaviors of A- and H- shares. Thus, no extra tail risk premium should be demanded with respect to these events.

5.2. Suggestions for future work

In addition to the research suggestions mentioned in section 5.1, some other topics may also be of interest for future research.

(i) Inference on and comparisons of heavy-tailedness in other volatile markets. The analysis and methods in this paper can be used in different markets. For example, due to the recent surge in interest in academia, industry and among the general public, the inference on and comparison of the degrees of heavy-tailedness and the likelihood of large fluctuations, downfalls and crises in different cryptocurrency markets as well as their comparison with stock markets seems to be of particular interest. Importantly, the institutional differences (e.g., investor composition and exchange rate risk) of these markets may be considered and investigated.

(ii) Modeling the volatility dynamics and inference on GARCH-type models for A- and H-stock market returns. As is well known, the nonlinear dynamics of volatility in GARCH-type models generate power-law distributions of the time series even in the case of light-tailed normal innovations; the effect is further pronounced under heavy-tailed (e.g., Student-t distributed innovations). The future research may focus on the analysis and

---

[39]Despite of similar tail behavior, the information efficiency may be quite different for these two markets, since the government intervention itself will serve as an extra pricing factor. See Brunnermeier et al. (2017).

comparison of such volatility dynamics models for the markets considered and other settings and their implications for heavy-tailedness, crises and their propagation.

(iii) Robust inference method under deviations from power laws. As is noted in the paper, price limits which are common in developing country financial markets tend to generate return distributions that have clusters in the tail and deviate from power laws. Under this circumstance, the heuristic rule of truncation level selection seems to be too subjective, since even log-log rank-size estimation method will lead to unstable tail index estimates $\hat{\zeta}$ for small truncation levels. Thus, future research that focuses on the refinement of the estimation method would be helpful. One direction is to endogenously detect the true tail threshold for these return distributions. Another direction may come from reshaping the clusters based on the volatility spillover hypothesis (Kim and Rhee, 1997; Kim, 2001). It is of importance to generate a robust indicator that could accurately reflect the degree of heavy-tailedness for these return series.

(iv) Testing for structural breaks in tail index with "unknown breakpoints". The analysis of structural breaks in this paper is based on a single "known breakpoint" which is exogenously selected. Further research could employ the methods developed by Quintos et al. (2001) and Candelon and Straetmans (2006) to detect (single or multiple) "unknown" structural breaks. In addition, to embed Ibragimov and Müller (2016) method into the tests for structural breaks is of much interest since this method has a better performance under heavy-tails.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A. and Tahbaz-Salehi, A. (2012), 'The network origins of aggregate fluctuations', *Econometrica* **80**(5), 1977–2016.

Acemoglu, D., Ozdaglar, A. and Tahbaz-Salehi, A. (2017), 'Microeconomic origins of macroeconomic tail risks', *The American Economic Review* **107**(1), 54–108.

Aloui, R., Aïssa, M. S. B. and Nguyen, D. K. (2011), 'Global financial crisis, extreme interdependences, and contagion effects: The role of economic structure?', *Journal of Banking & Finance* **35**(1), 130–141.

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

Assaf, A. (2009), 'Extreme observations and risk assessment in the equity markets of MENA region: Tail measures and Value-at-Risk', *International Review of Financial Analysis* **18**(3), 109–116.

Axtell, R. L. (2001), 'Zipf distribution of US firm sizes', *Science* **293**(5536), 1818–1820.

Barber, B. M. and Odean, T. (2000), 'Trading is hazardous to your wealth: The common stock investment performance of individual investors', *The Journal of Finance* **55**(2), 773–806.

Basrak, B., Davis, R. A. and Mikosch, T. (2002), 'Regular variation of GARCH processes', *Stochastic Processes and Their Applications* **99**(1), 95–115.

Blattberg, R. C. and Gonedes, N. J. (1974), 'A comparison of the stable and Student distributions as statistical models for stock prices', *The Journal of Business* **47**(2), 244–280.

Bollerslev, T., Todorov, V. and Xu, L. (2015), 'Tail risk premia and return predictability', *Journal of Financial Economics* **118**(1), 113–134.

Brillinger, D. R. (1973), 'Estimation of the mean of a stationary time series by sampling', *Journal of Applied Probability* **10**(2), 419–431.

Brown, S. J., Goetzmann, W., Ibbotson, R. G. and Ross, S. A. (1992), 'Survivorship bias in performance studies', *The Review of Financial Studies* **5**(4), 553–580.

Brunnermeier, M., Sockin, M. and Xiong, W. (2017), 'China's model of managing the financial system', Available from: http://www.princeton.edu/~wxiong/papers/ChinaTrading.pdf. [Accessed: 14th October 2017].

Cai, C. X., McGuinness, P. B. and Zhang, Q. (2011), 'The pricing dynamics of cross-listed securities: The case of Chinese A-and H-shares', *Journal of Banking & Finance* **35**(8), 2123–2136.

Campbell, J. Y., Lo, A. W.-C. and MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Candelon, B. and Straetmans, S. (2006), 'Testing for multiple regimes in the tail behavior of emerging currency returns', *Journal of International Money and Finance* **25**(7), 1187–1205.

Carpenter, J. N. and Lynch, A. W. (1999), 'Survivorship bias and attrition effects in measures of performance persistence', *Journal of financial economics* **54**(3), 337–374.

Chen, H. and Zhu, Y. (2015), 'An empirical study on the threshold cointegration of Chinese A and H cross-listed shares', *Journal of Applied Statistics* **42**(11), 2406–2419.

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

Cont, R. (2001), 'Empirical properties of asset returns: Stylized facts and statistical issues', *Quantitative Finance* **1**(2), 223–236.

Cont, R. and Bouchaud, J.-P. (2000), 'Herd behavior and aggregate fluctuations in financial markets', *Macroeconomic dynamics* **4**(2), 170–196.

Danielsson, J., de Haan, L., Peng, L. and de Vries, C. G. (2001), 'Using a bootstrap method to choose the sample fraction in tail index estimation', *Journal of Multivariate Analysis* **76**(2), 226–248.

Danielsson, J., Ergun, L. M., de Haan, L. and de Vries, C. G. (2016), 'Tail index estimation: Quantile driven threshold selection', Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2717478. [Accessed: 10th August 2017].

Davis, R. A. and Mikosch, T. (1998), 'Limit theory for the sample ACF of stationary process with heavy tails with applications to ARCH', *Annals of Statistics* **26**, 2049–2080.

Drees, H., De Haan, L. and Resnick, S. (2000), 'How to make a Hill plot', *Annals of Statistics* pp. 254–274.

Eaton, J., Kortum, S. and Kramarz, F. (2011), 'An anatomy of international trade: Evidence from French firms', *Econometrica* **79**(5), 1453–1498.

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

Fernald, J. and Rogers, J. H. (2002), 'Puzzles in the Chinese stock market', *The Review of Economics and Statistics* **84**(3), 416–432.

Gabaix, X. (1999), 'Zipf's law for cities: An explanation', *The Quarterly Journal of Economics* **114**(3), 739–767.

Gabaix, X. (2009), 'Power laws in economics and finance', *Annu. Rev. Econ.* **1**(1), 255–294.

Gabaix, X. (2016), 'Power laws in economics: An introduction', *The Journal of Economic Perspectives* **30**(1), 185–205.

Gabaix, X., Gopikrishnan, P., Plerou, V. and Stanley, H. E. (2006), 'Institutional investors and stock market volatility', *The Quarterly Journal of Economics* **121**(2), 461–504.

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

Gabaix, X. and Landier, A. (2008), 'Why has CEO pay increased so much?', *The Quarterly Journal of Economics* **123**(1), 49–100.

Galbraith, J. W. and Zernov, S. (2004), 'Circuit breakers and the tail index of equity returns', *Journal of Financial Econometrics* **2**(1), 109–129.

Gatev, E., Goetzmann, W. N. and Rouwenhorst, K. G. (2006), 'Pairs trading: Performance of a relative-value arbitrage rule', *The Review of Financial Studies* **19**(3), 797–827.

Gopikrishnan, P., Plerou, V., Amaral, L. A. N., Meyer, M. and Stanley, H. E. (1999), 'Scaling of the distribution of fluctuations of financial market indices', *Physical Review E* **60**(5), 5305.

Gopikrishnan, P., Plerou, V., Gabaix, X. and Stanley, H. E. (2000), 'Statistical properties of share volume traded in financial markets', *Physical Review E* **62**(4), R4493.

Hall, P. (1982), 'On some simple estimates of an exponent of regular variation', *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 37–42.

Hall, P. (1990), 'Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems', *Journal of Multivariate Analysis* **32**(2), 177–203.

Hartmann, P., Straetmans, S. and De Vries, C. G. (2004), 'Asset market linkages in crisis periods', *The Review of Economics and Statistics* **86**(1), 313–326.

Hertz, E. (1998), *The Trading Crowd: An Ethnography of the Shanghai Stock Market*, Vol. 108, Cambridge University Press.

Hill, B. M. et al. (1975), 'A simple general approach to inference about the tail of a distribution', *The Annals of Statistics* **3**(5), 1163–1174.

HKEX (2017), 'HKEX fact book', Available from: https://www.hkex.com.hk/eng/stat/statrpt/factbook/factbook.htm. [Accessed: 10th August 2017].

HSIndexes (2017), 'Index methodology: For managing the Hang Seng China AH index series', Available from: https://www.hsi.com.hk/static/uploads/contents/zh_hk/dl_centre/methodologies/IM_chinaahc.pdf. [Accessed: 10th August 2017].

Huang, Y., Miao, J. and Wang, P. (2016), 'Saving China's stock market', Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2821664. [Accessed: 10th August 2017].

Huisman, R., Koedijk, K. G., Kool, C. J. M. and Palm, F. (2001), 'Tail-index estimates in small samples', *Journal of Business & Economic Statistics* **19**(2), 208–216.

Ibragimov, M. and Ibragimov, R. (2017), 'Heavy tails and upper-tail inequality: The case of Russia', *Empirical Economics* pp. 1–15.

Ibragimov, M., Ibragimov, R. and Kattuman, P. (2013), 'Emerging markets and heavy tails', *Journal of Banking & Finance* **37**(7), 2546–2559.

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

Ibragimov, R. (2007), 'Efficiency of linear estimators under heavy-tailedness: convolutions of α-symmetric distributions', *Econometric Theory* **23**(3), 501–517.

Ibragimov, R. (2009), Heavy-tailed densities, *in* S. N. Durlauf and L. E. Blume, eds, '*The New Palgrave Dictionary of Economics*', Palgrave Macmillan.

Ibragimov, R., Jaffee, D. and Walden, J. (2008), 'Nondiversification traps in catastrophe insurance markets', *The Review of Financial Studies* **22**(3), 959–993.

Ibragimov, R., Jaffee, D. and Walden, J. (2011), 'Diversification disasters', *Journal of Financial Economics* **99**(2), 333–348.

Ibragimov, R. and Müller, U. K. (2010), 't-statistic based correlation and heterogeneity robust inference', *Journal of Business & Economic Statistics* **28**(4), 453–468.

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

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

Ibragimov, R. and Walden, J. (2010), 'Optimal bundling strategies under heavy-tailed valuations', *Management Science* **56**(11), 1963–1976.

Jansen, D. W. and De Vries, C. G. (1991), 'On the frequency of large stock returns: Putting booms and busts into perspective', *The Review of Economics and Statistics* pp. 18–24.

Jondeau, E. and Rockinger, M. (2003), 'Testing for differences in the tails of stock-market returns', *Journal of Empirical Finance* **10**(5), 559–581.

Keller, W. and Yeaple, S. R. (2013), 'The gravity of knowledge', *The American Economic Review* **103**(4), 1414–1444.

Kelly, B. and Jiang, H. (2014), 'Tail risk and asset prices', *The Review of Financial Studies* **27**(10), 2841–2871.

Kim, K. A. (2001), 'Price limits and stock market volatility', *Economics Letters* **71**(1), 131–136.

Kim, K. A. and Rhee, S. (1997), 'Price limit performance: evidence from the Tokyo Stock Exchange', *the Journal of Finance* **52**(2), 885–901.

Koedijk, K. G., Schafgans, M. M. and De Vries, C. G. (1990), 'The tail index of exchange rate returns', *Journal of International Economics* **29**(1-2), 93–108.

Koedijk, K. G., Stork, P. A. and de Vries, C. G. (1992), 'Differences between foreign exchange rate regimes: The view from the tails', *Journal of International Money and Finance* **11**(5), 462–473.

Kyle, A. S. and Obizhaeva, A. A. (2016), 'Market microstructure invariance: Empirical hypotheses', *Econometrica* **84**(4), 1345–1404.

Liu, J. and Timmermann, A. (2013), 'Optimal convergence trade strategies', *The Review of Financial Studies* **26**(4), 1048–1086.

Loretan, M. and Phillips, P. C. (1994), 'Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets', *Journal of Empirical Finance* **1**(2), 211–248.

Mandelbrot, B. (1963), 'The variation of certain speculative prices', *The Journal of Business* **36**(4), 394–419.

Mensi, W., Hammoudeh, S., Reboredo, J. C. and Nguyen, D. K. (2014), 'Do global factors impact BRICS stock markets? A quantile regression approach', *Emerging Markets Review* **19**, 1–17.

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

Ng, L. and Wu, F. (2007), 'The trading behavior of institutions and individuals in Chinese equity markets', *Journal of Banking & Finance* **31**(9), 2695–2710.

Odean, T. (1998), 'Are investors reluctant to realize their losses?', *The Journal of Finance* **53**(5), 1775–1798.

Pickands, J. (1975), 'Statistical inference using extreme order statistics', *the Annals of Statistics* pp. 119–131.

Plerou, V., Gopikrishnan, P., Gabaix, X., Amaral, L. N. and Stanley, H. E. (2001), 'Price fluctuations, market activity and trading volume', *Quantitative Finance* **1**(2), 262–269.

Quintos, C., Fan, Z. and Phillips, P. C. (2001), 'Structural change tests in tail behaviour and the Asian crisis', *The Review of Economic Studies* **68**(3), 633–663.

Su, Q., Chong, T. T.-L. and Yan, I. K.-M. (2007), 'On the convergence of the Chinese and Hong Kong stock markets: A cointegration analysis of the A and H shares', *Applied Financial Economics* **17**(16), 1349–1357.

Yan, C., Zhang, J., Zhang, Y. and Tang, Y. (2005), 'Power–law properties of Chinese stock market', *Physica A: Statistical Mechanics and its Applications* **353**, 425–432.

Zhang, J., Chen, Q. and Wang, Y. (2009), 'Zipf distribution in top Chinese firms and an economic explanation', *Physica A: statistical Mechanics and its applications* **388**(10), 2020–2024.

Zhang, J., Zhang, Y. and Kleinert, H. (2007), 'Power tails of index distributions in Chinese stock market', *Physica A: Statistical Mechanics and its Applications* **377**(1), 166–172.

Zhang, R. and Zhang, T. (2017), 'Understanding AH premium in Chinese stock market', Available from: http://tongbinzhang.weebly.com/uploads/9/1/6/6/91669244/draft.pdf. [Accessed: 10th August 2017].

Zipf, G. K. (1949), *Human Behavior and the Principle of Least Effort*, Cambridge: Addison-Wesley.

Appendix A. Tables

Table A1: Estimation results for the tail indices ζ for A and H dual-listed pairs

| Company | Truncation (%) | A share | | | H share | | |
		N	$\hat{\zeta}$	$95\%CI$	N	$\hat{\zeta}$	$95\%CI$
BANK OF CHINA	5	2678	3.07	(2.34 , 3.81)	2754	2.98	(2.27 , 3.68)
	10		2.69	(2.24 , 3.15)		2.86	(2.38 , 3.34)
CRRC	5	2076	3.86	(2.81 , 4.91)	2129	2.64	(1.93 , 3.36)
	10		2.50	(2.02 , 2.98)		2.63	(2.13 , 3.13)
MCC	5	1828	3.02	(2.14 , 3.90)	1856	3.30	(2.35 , 4.25)
	10		2.25	(1.79 , 2.71)		3.06	(2.44 , 3.68)
ICBC	5	2595	3.08	(2.33 , 3.83)	2649	3.22	(2.44 , 3.99)
	10		2.77	(2.29 , 3.24)		2.82	(2.34 , 3.30)
NCI	5	1337	4.42	(2.92 , 5.91)	1353	3.85	(2.55 , 5.14)
	10		3.36	(2.56 , 4.16)		3.25	(2.47 , 4.02)
ABC	5	1707	2.67	(1.87 , 3.47)	1735	3.35	(2.35 , 4.34)
	10		2.40	(1.89 , 2.91)		3.12	(2.47 , 3.78)
CHINA RAIL CONS	5	2265	3.87	(2.86 , 4.88)	2298	3.42	(2.54 , 4.31)
	10		2.66	(2.17 , 3.15)		3.16	(2.59 , 3.74)
YANZHOU COAL	5	4180	5.41	(4.37 , 6.44)	4322	3.76	(3.05 , 4.46)
	10		3.50	(3.02 , 3.97)		3.37	(2.92 , 3.82)
LUOYANG GLASS	5	3910	9.49	(7.61 , 11.37)	3483	2.96	(2.34 , 3.59)
	10		3.06	(2.63 , 3.49)		2.64	(2.25 , 3.03)
DONGFANG ELEC	5	4090	4.77	(3.84 , 5.69)	4223	3.44	(2.79 , 4.10)
	10		3.23	(2.79 , 3.67)		3.16	(2.73 , 3.58)
TIANJIN CAPITAL	5	4175	5.47	(4.42 , 6.51)	4328	3.53	(2.86 , 4.19)
	10		3.22	(2.78 , 3.65)		3.15	(2.73 , 3.56)
SINOPEC SSC	5	4052	3.27	(2.64 , 3.91)	4252	3.16	(2.56 , 3.76)
	10		3.38	(2.91 , 3.84)		3.04	(2.63 , 3.45)
CITIC BANK	5	2473	3.72	(2.80 , 4.65)	2526	3.43	(2.58 , 4.28)
	10		3.03	(2.50 , 3.57)		3.24	(2.68 , 3.81)
HISENSE KELON	5	3846	2.87	(2.30 , 3.44)	3325	2.98	(2.34 , 3.62)
	10		3.36	(2.89 , 3.84)		2.95	(2.50 , 3.40)
SHANDONG XINHUA	5	4124	4.20	(3.39 , 5.01)	4200	3.48	(2.82 , 4.15)
	10		2.94	(2.54 , 3.34)		3.06	(2.65 , 3.47)
CHINA EAST AIR	5	4056	4.14	(3.34 , 4.95)	4197	3.19	(2.58 , 3.80)
	10		3.06	(2.64 , 3.48)		3.05	(2.63 , 3.46)
NANJING PANDA	5	4085	7.01	(5.65 , 8.37)	4114	3.50	(2.82 , 4.17)
	10		3.45	(2.97 , 3.92)		3.06	(2.64 , 3.48)
CHINA RAILWAY	5	2327	3.97	(2.95 , 4.99)	2361	3.28	(2.44 , 4.11)

Company	Truncation (%)	A share			H share		
		N	$\hat{\zeta}$	95%CI	N	$\hat{\zeta}$	95%CI
	10		2.60	(2.13 , 3.07)		3.06	(2.51 , 3.61)
SINOPEC CORP	5	3822	3.40	(2.71 , 4.08)	4135	3.71	(2.99 , 4.42)
	10		2.88	(2.47 , 3.29)		3.38	(2.92 , 3.84)
JINGWEI TEXTILE	5	4051	5.07	(4.08 , 6.05)	3800	3.15	(2.52 , 3.79)
	10		3.45	(2.98 , 3.93)		2.95	(2.53 , 3.36)
ANGANG STEEL	5	4186	4.77	(3.85 , 5.68)	4327	3.72	(3.02 , 4.42)
	10		3.33	(2.87 , 3.78)		3.49	(3.02 , 3.95)
SHANGHAI PECHEM	5	4137	3.31	(2.68 , 3.95)	4329	3.84	(3.12 , 4.57)
	10		3.55	(3.07 , 4.04)		3.41	(2.95 , 3.86)
MAANSHAN IRON	5	4187	4.18	(3.38 , 4.98)	4334	3.64	(2.95 , 4.32)
	10		3.04	(2.63 , 3.46)		3.30	(2.86 , 3.74)
COMEC	5	4023	9.53	(7.66 , 11.39)	4079	3.43	(2.76 , 4.09)
	10		4.07	(3.51 , 4.64)		2.96	(2.56 , 3.37)
KUNMING MACHINE	5	3918	3.97	(3.18 , 4.76)	3788	3.55	(2.83 , 4.27)
	10		3.38	(2.91 , 3.86)		3.15	(2.70 , 3.60)
JINGCHENG MAC	5	3988	6.01	(4.83 , 7.19)	3939	3.51	(2.81 , 4.20)
	10		3.39	(2.92 , 3.86)		2.99	(2.57 , 3.41)
TSINGTAO BREW	5	4153	3.19	(2.57 , 3.80)	4322	3.39	(2.75 , 4.03)
	10		2.87	(2.48 , 3.26)		2.94	(2.55 , 3.33)
NE ELECTRIC	5	3940	3.80	(3.05 , 4.55)	3945	2.73	(2.19 , 3.26)
	10		3.10	(2.67 , 3.53)		2.49	(2.14 , 2.83)
HTSC	5	1794	4.78	(3.38 , 6.18)	534	——	——
	10		3.03	(2.40 , 3.66)		——	——
HAITONG SEC	5	3679	6.81	(5.42 , 8.20)	1293	2.82	(1.85 , 3.79)
	10		3.69	(3.15 , 4.22)		2.65	(2.00 , 3.29)
CEB BANK	5	1668	3.41	(2.37 , 4.44)	875	——	——
	10		2.63	(2.07 , 3.20)		——	——
EB SECURITIES	5	1909	4.23	(3.03 , 5.43)	233	——	——
	10		3.05	(2.44 , 3.66)		——	——
CMSC	5	1858	4.02	(2.87 , 5.18)	198	——	——
	10		2.91	(2.32 , 3.50)		——	——
CITIC SEC	5	3455	6.24	(4.93 , 7.56)	1431	3.57	(2.40 , 4.73)
	10		3.44	(2.92 , 3.95)		3.08	(2.37 , 3.80)
CM BANK	5	3644	4.03	(3.2 , 4.86)	2668	3.48	(2.64 , 4.31)
	10		3.22	(2.75 , 3.68)		2.92	(2.42 , 3.41)
FUYAO GLASS	5	4190	4.33	(3.50 , 5.16)	573	——	——
	10		3.36	(2.90 , 3.81)		——	——
SH PHARMA	5	4077	4.43	(3.57 , 5.29)	1525	3.00	(2.05 , 3.95)

Company	Truncation (%)	A share			H share		
		N	$\hat{\zeta}$	95%CI	N	$\hat{\zeta}$	95%CI
	10		3.27	(2.82 , 3.72)		3.03	(2.35 , 3.71)
CPIC	5	2325	4.74	(3.52 , 5.96)	1872	4.04	(2.88 , 5.19)
	10		3.70	(3.03 , 4.38)		3.61	(2.87 , 4.34)
GOLDWIND	5	2307	4.99	(3.70 , 6.27)	1670	4.03	(2.81 , 5.26)
	10		3.62	(2.96 , 4.28)		3.52	(2.77 , 4.28)
FOSUN PHARMA	5	4172	4.26	(3.44 , 5.07)	1152	3.12	(1.98 , 4.25)
	10		3.32	(2.87 , 3.77)		3.06	(2.27 , 3.85)
CIMC	5	4151	4.54	(3.67 , 5.41)	1129	2.80	(1.76 , 3.84)
	10		3.52	(3.04 , 4.00)		2.78	(2.05 , 3.50)
MINSHENG BANK	5	3961	3.84	(3.08 , 4.60)	1888	4.11	(2.94 , 5.29)
	10		3.27	(2.82 , 3.73)		3.49	(2.78 , 4.19)
CHENMING PAPER	5	3952	4.41	(3.54 , 5.28)	2224	3.41	(2.51 , 4.30)
	10		3.23	(2.78 , 3.69)		3.21	(2.61 , 3.81)
GF SEC	5	3321	2.62	(2.05 , 3.18)	568	——	——
	10		2.69	(2.28 , 3.10)		——	——
DZUG	5	4154	5.82	(4.70 , 6.94)	154	——	——
	10		3.32	(2.87 , 3.77)		——	——
LIVZON PHARMA	5	3945	4.32	(3.46 , 5.17)	792	——	——
	10		3.40	(2.93 , 3.88)		——	——
ZOOMLION	5	3935	4.86	(3.90 , 5.82)	1616	4.01	(2.78 , 5.25)
	10		3.24	(2.79 , 3.69)		3.54	(2.77 , 4.31)
ZTE	5	4138	4.12	(3.33 , 4.91)	3086	3.14	(2.44 , 3.84)
	10		3.33	(2.88 , 3.79)		3.15	(2.66 , 3.65)
ZMJ	5	1543	3.80	(2.60 , 4.99)	1117	3.57	(2.24 , 4.89)
	10		3.10	(2.41 , 3.79)		3.33	(2.46 , 4.20)
CGS	5	125	——	——	1032	2.83	(1.74 , 3.92)
	10		——	——		2.74	(1.99 , 3.49)
CMOC	5	1148	18.27	(11.56 , 24.98)	2513	3.36	(2.53 , 4.19)
	10		3.72	(2.76 , 4.68)		3.16	(2.61 , 3.71)
BANKCOMM	5	2471	3.34	(2.51 , 4.17)	2984	3.37	(2.61 , 4.14)
	10		2.96	(2.44 , 3.48)		2.99	(2.51 , 3.47)
ZIJIN MINING	5	2213	4.62	(3.41 , 5.84)	3320	3.97	(3.12 , 4.83)
	10		2.91	(2.37 , 3.46)		3.41	(2.89 , 3.93)
CHINA OILFIELD	5	2379	4.84	(3.61 , 6.07)	3624	3.60	(2.86 , 4.34)
	10		3.33	(2.73 , 3.93)		3.15	(2.69 , 3.61)
DALIAN PORT	5	1611	6.21	(4.30 , 8.13)	2774	2.94	(2.25 , 3.64)
	10		2.28	(1.78 , 2.78)		2.80	(2.33 , 3.27)
COSCO SHIP DEV	5	2234	5.34	(3.94 , 6.74)	3149	3.82	(2.97 , 4.66)

Company	Truncation (%)	A share			H share		
		N	$\hat{\zeta}$	95%CI	N	$\hat{\zeta}$	95%CI
	10		2.98	(2.43 , 3.54)		3.15	(2.66 , 3.64)
SH ELECTRIC	5	1865	5.05	(3.60 , 6.50)	2937	3.08	(2.38 , 3.79)
	10		2.66	(2.12 , 3.20)		2.89	(2.43 , 3.36)
CHINA LIFE	5	2558	4.48	(3.38 , 5.57)	3357	3.69	(2.90 , 4.47)
	10		3.21	(2.66 , 3.77)		3.29	(2.79 , 3.78)
CHALCO	5	2458	7.94	(5.96 , 9.92)	3841	3.97	(3.18 , 4.77)
	10		3.45	(2.84 , 4.07)		3.68	(3.16 , 4.20)
WEICHAI POWER	5	2432	4.71	(3.53 , 5.89)	3273	3.95	(3.10 , 4.81)
	10		3.51	(2.88 , 4.13)		3.46	(2.93 , 3.99)
GREATWALL MOTOR	5	1398	5.14	(3.44 , 6.85)	3343	3.39	(2.66 , 4.11)
	10		3.54	(2.71 , 4.37)		3.19	(2.71 , 3.68)
PING AN	5	2469	4.40	(3.30 , 5.50)	3179	3.62	(2.82 , 4.42)
	10		3.40	(2.80 , 4.00)		3.10	(2.62 , 3.59)
GAC GROUP	5	1285	4.25	(2.78 , 5.72)	1704	4.74	(3.31 , 6.16)
	10		2.99	(2.26 , 3.72)		4.41	(3.47 , 5.35)
BBMG	5	1533	5.24	(3.58 , 6.89)	1958	3.52	(2.54 , 4.51)
	10		3.27	(2.54 , 4.01)		3.31	(2.66 , 3.97)
COSCO SHIP HOLD	5	2272	7.23	(5.35 , 9.11)	2891	3.53	(2.72 , 4.35)
	10		3.41	(2.78 , 4.04)		3.04	(2.55 , 3.54)
CHINA COAL	5	2296	4.41	(3.27 , 5.55)	2612	3.37	(2.55 , 4.19)
	10		3.12	(2.55 , 3.69)		3.17	(2.63 , 3.72)
CHINA COMM CONS	5	1310	6.84	(4.51 , 9.18)	2615	3.43	(2.60 , 4.26)
	10		2.78	(2.10 , 3.45)		3.16	(2.62 , 3.70)
BYD COMPANY	5	1467	5.80	(3.92 , 7.68)	3693	3.51	(2.79 , 4.22)
	10		3.59	(2.77 , 4.41)		3.23	(2.76 , 3.69)
COSCO SHIP ENGY	5	3535	5.03	(3.98 , 6.08)	4244	3.44	(2.79 , 4.10)
	10		3.36	(2.86 , 3.85)		3.28	(2.84 , 3.72)
CHINA SHENHUA	5	2342	3.89	(2.89 , 4.89)	2991	2.96	(2.29 , 3.63)
	10		3.12	(2.56 , 3.69)		3.03	(2.55 , 3.52)
HUADIAN POWER	5	2975	4.35	(3.36 , 5.34)	4330	3.66	(2.97 , 4.35)
	10		3.04	(2.55 , 3.53)		3.31	(2.87 , 3.75)
ZHEJIANG SHIBAO	5	994	——	——	2420	2.30	(1.72 , 2.87)
	10		——	——		2.29	(1.88 , 2.69)
CHINA SOUTH AIR	5	3303	5.89	(4.62 , 7.17)	4297	3.81	(3.09 , 4.54)
	10		3.33	(2.82 , 3.84)		3.51	(3.04 , 3.98)
CHONGQING IRON	5	2189	7.50	(5.51 , 9.49)	4236	3.10	(2.51 , 3.69)
	10		3.42	(2.78 , 4.06)		2.86	(2.48 , 3.25)
ANHUIEXPRESSWAY	5	3480	3.42	(2.70 , 4.14)	4333	4.23	(3.43 , 5.02)

Table A1 – continued from previous page

Company	Truncation (%)	A share			H share		
		N	$\hat{\zeta}$	95%CI	N	$\hat{\zeta}$	95%CI
	10		2.75	(2.34 , 3.16)		3.69	(3.20 , 4.18)
DATANG POWER	5	2539	5.46	(4.12 , 6.80)	4329	3.29	(2.67 , 3.91)
	10		3.13	(2.59 , 3.68)		3.05	(2.65 , 3.46)
CCB	5	2378	3.45	(2.57 , 4.32)	2898	3.33	(2.57 , 4.10)
	10		2.75	(2.25 , 3.24)		2.95	(2.47 , 3.43)
ANHUI CONCH	5	3695	4.41	(3.51 , 5.30)	4274	3.96	(3.21 , 4.71)
	10		3.49	(2.99 , 4.00)		3.36	(2.91 , 3.81)
HUANENG POWER	5	3723	3.93	(3.13 , 4.73)	4324	3.79	(3.08 , 4.51)
	10		3.14	(2.69 , 3.59)		3.42	(2.97 , 3.88)
DONGJIANG ENV	5	1178	4.53	(2.90 , 6.17)	3469	3.17	(2.51 , 3.84)
	10		3.18	(2.36 , 3.99)		2.95	(2.51 , 3.39)
BAIYUNSHAN PH	5	3809	5.35	(4.27 , 6.42)	4183	4.12	(3.33 , 4.91)
	10		3.45	(2.96 , 3.95)		3.41	(2.95 , 3.87)
PETROCHINA	5	2360	3.32	(2.47 , 4.16)	4266	3.51	(2.84 , 4.18)
	10		2.70	(2.21 , 3.19)		3.19	(2.77 , 3.62)
XINHUA WINSHARE	5	237	——	——	2507	2.82	(2.12 , 3.52)
	10		——	——		2.75	(2.27 , 3.23)
AIR CHINA	5	2594	6.21	(4.70 , 7.72)	3061	3.57	(2.77 , 4.37)
	10		3.64	(3.02 , 4.27)		3.27	(2.75 , 3.79)
BEIJING N STAR	5	2601	8.04	(6.08 , 9.99)	4320	3.26	(2.65 , 3.88)
	10		3.22	(2.67 , 3.77)		3.04	(2.64 , 3.45)
SHANDONG MOLONG	5	1644	7.75	(5.37 , 10.12)	3208	3.39	(2.65 , 4.13)
	10		3.97	(3.11 , 4.83)		3.04	(2.57 , 3.51)
SHENZHENEXPRESS	5	3720	3.64	(2.90 , 4.38)	4319	3.92	(3.18 , 4.65)
	10		2.91	(2.49 , 3.33)		3.42	(2.96 , 3.88)
GUANGSHEN RAIL	5	2567	3.53	(2.66 , 4.39)	4334	4.19	(3.40 , 4.98)
	10		2.86	(2.36 , 3.35)		3.44	(2.98 , 3.90)
JIANGXI COPPER	5	3692	4.69	(3.73 , 5.64)	4318	3.66	(2.97 , 4.35)
	10		3.50	(2.99 , 4.00)		3.23	(2.80 , 3.66)
JIANGSU EXPRESS	5	3930	3.31	(2.66 , 3.96)	4318	3.73	(3.03 , 4.44)
	10		2.88	(2.48 , 3.29)		3.38	(2.93 , 3.83)
SICHUAN EXPRESS	5	1929	3.50	(2.51 , 4.49)	4315	3.34	(2.71 , 3.97)
	10		2.40	(1.92 , 2.88)		2.96	(2.57 , 3.36)
FIRST TRACTOR	5	1209	15.06	(9.67 , 20.45)	4328	3.58	(2.91 , 4.26)
	10		3.93	(2.94 , 4.92)		3.25	(2.81 , 3.68)
CHINA VANKE	5	4041	4.29	(3.45 , 5.13)	749	——	——
	10		3.34	(2.88 , 3.80)		——	——

Table A2: The results for equality test for dual-listed pairs

Company	N^A	N^H	$H_0 : \zeta^A = \zeta^H$			
			$q = 2$		$q = 4$	
			BF Statistic	p-value	BF Statistic	p-value
BANK OF CHINA	2678	2730	1.23	0.43	1.23	0.31
CRRC	2073	2129	0.24	0.85	0.50	0.65
ICBC	2595	2649	1.64	0.35	1.88	0.16
CHINA RAIL CONS	2262	2298	0.83	0.56	0.69	0.54
YANZHOU COAL	4180	4321	0.35	0.78	0.92	0.42
LUOYANG GLASS	3910	3482	2.44	0.25	1.29	0.29
DONGFANG ELEC	4090	4222	0.54	0.68	0.20	0.85
TIANJIN CAPITAL	4175	4327	0.03	0.98	2.27	0.11
SINOPEC SSC	4052	4251	0.54	0.68	1.04	0.37
CITIC BANK	2473	2526	1.18	0.45	1.28	0.29
HISENSE KELON	3846	3324	0.35	0.79	0.93	0.42
SHANDONG XINHUA	4124	4199	0.36	0.78	0.42	0.70
CHINA EAST AIR	4056	4196	0.02	0.99	1.82	0.17
NANJING PANDA	4085	4113	4.45	0.14	2.49	0.09
CHINA RAILWAY	2323	2361	1.28	0.42	0.42	0.70
SINOPEC CORP	3822	3941	2.93	0.21	0.45	0.68
JINGWEI TEXTILE	4051	3799	1.25	0.43	1.91	0.15
ANGANG STEEL	4186	4326	0.31	0.81	0.14	0.90
SHANGHAI PECHEM	4137	4328	0.44	0.74	1.03	0.38
MAANSHAN IRON	4187	4333	0.79	0.58	0.19	0.86
COMEC	4023	4078	2.84	0.22	2.33	0.10
KUNMING MACHINE	3918	3788	1.13	0.46	0.70	0.53
JINGCHENG MAC	3988	3938	4.52	0.14	1.74	0.18
TSINGTAO BREW	4153	4321	1.32	0.41	0.40	0.71
NE ELECTRIC	3940	3944	3.23	0.19	1.41	0.25
CM BANK	2608	2667	0.22	0.86	0.44	0.69
CHENMING PAPER	2182	2223	2.70	0.23	2.58	0.08
ZTE	2975	3085	0.60	0.66	0.32	0.77
BANKCOMM	2516	2471	0.96	0.51	1.43	0.25
ZIJIN MINING	2251	2213	0.78	0.58	0.09	0.93
CHINA OILFIELD	2422	2379	0.07	0.95	0.08	0.94
COSCO SHIP DEV	2286	2234	0.27	0.83	0.30	0.79
SH ELECTRIC	2076	1865	0.02	0.98	0.98	0.40
CHINA LIFE	2601	2558	0.58	0.67	0.74	0.51
CHALCO	2524	2458	0.22	0.86	0.64	0.57

| Company | N^A | N^H | $H_0 : \zeta^A = \zeta^H$ | | | |
| | | | $q = 2$ | | $q = 4$ | |
			BF Statistic	p-value	BF Statistic	p-value
WEICHAI POWER	2520	2432	0.16	0.90	0.83	0.47
PING AN	2513	2469	0.65	0.63	0.38	0.73
COSCO SHIP HOLD	2400	2272	0.52	0.69	1.52	0.23
CHINA COAL	2337	2296	0.24	0.85	0.17	0.87
COSCO SHIP ENGY	3661	3535	0.31	0.81	0.21	0.85
CHINA SHENHUA	2417	2342	0.22	0.86	0.70	0.53
HUADIAN POWER	3073	2975	0.57	0.67	0.07	0.95
CHINA SOUTH AIR	3427	3303	0.02	0.98	0.97	0.40
CHONGQING IRON	2469	2189	1.91	0.31	1.20	0.32
ANHUIEXPRESSWAY	3593	3480	1.11	0.47	0.70	0.53
DATANG POWER	2608	2539	0.27	0.83	1.04	0.37
CCB	2425	2378	1.27	0.43	1.50	0.23
ANHUI CONCH	3793	3695	1.38	0.40	0.48	0.66
HUANENG POWER	3851	3723	1.04	0.49	0.71	0.53
BAIYUNSHAN PH	3915	3809	1.46	0.38	0.70	0.53
PETROCHINA	2398	2360	1.17	0.45	1.93	0.15
AIR CHINA	2651	2594	0.15	0.90	0.98	0.40
BEIJING N STAR	2644	2601	0.60	0.66	0.90	0.44
SHENZHENEXPRESS	3832	3720	1.69	0.34	1.30	0.29
GUANGSHEN RAIL	2610	2567	0.77	0.58	0.30	0.79
JIANGXI COPPER	3821	3692	0.81	0.56	0.99	0.39
JIANGSU EXPRESS	4070	3930	1.74	0.33	2.15	0.12

Table A3: The left and right tail index estimates $\hat{\zeta}^-$ and $\hat{\zeta}^+$ for A-shares

Company	Truncation (%)	N_1	$\hat{\zeta}^-$	$95\%CI^-$	N_2	$\hat{\zeta}^+$	$95\%CI^+$
BANK OF CHINA	5	1191	3.24	(2.08 , 4.40)	1163	3.12	(1.98 , 4.25)
	10		2.80	(2.09 , 3.51)		2.66	(1.98 , 3.34)
ICBC	5	1157	2.60	(1.65 , 3.55)	1180	3.55	(2.27 , 4.83)
	10		2.63	(1.95 , 3.30)		3.00	(2.23 , 3.76)
CHINA RAIL CONS	5	1114	3.28	(2.07 , 4.50)	1066	4.94	(3.06 , 6.82)
	10		2.64	(1.94 , 3.33)		2.71	(1.98 , 3.44)
YANZHOU COAL	5	2018	4.61	(3.34 , 5.88)	2088	5.40	(3.93 , 6.87)
	10		3.44	(2.77 , 4.11)		3.42	(2.76 , 4.07)
LUOYANG GLASS	5	1837	4.79	(3.41 , 6.18)	1947	217.54	(156.32 , 278.77)
	10		2.87	(2.28 , 3.46)		3.11	(2.49 , 3.73)
DONGFANG ELEC	5	1948	4.35	(3.13 , 5.58)	2079	4.59	(3.34 , 5.83)
	10		3.01	(2.41 , 3.61)		3.32	(2.68 , 3.96)
TIANJIN CAPITAL	5	1935	4.76	(3.42 , 6.09)	2106	5.29	(3.86 , 6.72)
	10		3.14	(2.52 , 3.77)		3.17	(2.56 , 3.77)
SINOPEC SSC	5	1901	3.28	(2.35 , 4.21)	1959	3.01	(2.17 , 3.86)
	10		3.35	(2.67 , 4.02)		3.15	(2.53 , 3.78)
CITIC BANK	5	1205	3.97	(2.55 , 5.39)	1136	3.46	(2.19 , 4.74)
	10		3.27	(2.44 , 4.09)		2.81	(2.08 , 3.54)
HISENSE KELON	5	1802	2.97	(2.10 , 3.83)	1905	2.52	(1.80 , 3.23)
	10		3.50	(2.78 , 4.22)		3.04	(2.43 , 3.65)
SHANDONG XINHUA	5	1898	4.03	(2.89 , 5.18)	2114	3.83	(2.80 , 4.86)
	10		2.96	(2.36 , 3.55)		2.82	(2.28 , 3.35)
CHINA EAST AIR	5	1945	3.88	(2.79 , 4.97)	1896	3.62	(2.59 , 4.65)
	10		2.97	(2.38 , 3.56)		3.04	(2.43 , 3.65)
NANJING PANDA	5	1906	8.66	(6.20 , 11.12)	2080	5.44	(3.96 , 6.92)
	10		3.21	(2.56 , 3.85)		3.67	(2.96 , 4.37)
CHINA RAILWAY	5	1132	3.25	(2.06 , 4.44)	1043	5.55	(3.42 , 7.68)
	10		2.61	(1.93 , 3.29)		2.66	(1.94 , 3.38)
SINOPEC CORP	5	1834	3.36	(2.39 , 4.33)	1764	3.39	(2.39 , 4.40)
	10		2.89	(2.30 , 3.48)		2.88	(2.28 , 3.48)
JINGWEI TEXTILE	5	1875	4.81	(3.43 , 6.18)	2046	5.28	(3.83 , 6.72)
	10		3.21	(2.56 , 3.86)		3.71	(2.99 , 4.43)
ANGANG STEEL	5	1966	4.04	(2.91 , 5.17)	1997	5.61	(4.06 , 7.17)
	10		3.16	(2.54 , 3.79)		3.55	(2.85 , 4.24)
SHANGHAI PECHEM	5	1949	3.12	(2.24 , 4.00)	2009	3.29	(2.38 , 4.20)
	10		3.48	(2.79 , 4.17)		3.44	(2.77 , 4.11)
MAANSHAN IRON	5	1911	4.30	(3.09 , 5.52)	1915	4.12	(2.96 , 5.29)

Company	Truncation (%)	N_1	$\hat{\zeta}^-$	$95\%CI^-$	N_2	$\hat{\zeta}^+$	$95\%CI^+$
	10		3.23	(2.58 , 3.87)		2.93	(2.34 , 3.52)
COMEC	5	1941	6.04	(4.34 , 7.74)	2006	16.68	(12.06 , 21.31)
	10		3.54	(2.83 , 4.24)		4.42	(3.56 , 5.29)
KUNMING MACHINE	5	1852	5.40	(3.85 , 6.95)	1976	2.73	(1.97 , 3.49)
	10		3.99	(3.18 , 4.8)		2.80	(2.25 , 3.36)
JINGCHENG MAC	5	1875	5.44	(3.88 , 7.00)	2007	5.23	(3.78 , 6.68)
	10		3.49	(2.79 , 4.20)		3.10	(2.49 , 3.70)
TSINGTAO BREW	5	2013	3.71	(2.69 , 4.73)	2068	2.74	(1.99 , 3.49)
	10		3.03	(2.44 , 3.62)		2.68	(2.16 , 3.20)
NE ELECTRIC	5	1784	4.02	(2.84 , 5.20)	1955	3.33	(2.40 , 4.26)
	10		3.14	(2.49 , 3.79)		2.84	(2.28 , 3.41)
HAITONG SEC	5	1764	4.97	(3.50 , 6.44)	1822	13.62	(9.66 , 17.58)
	10		3.27	(2.58 , 3.95)		3.87	(3.07 , 4.66)
CITIC SEC	5	1697	4.70	(3.29 , 6.12)	1691	7.68	(5.37 , 9.99)
	10		3.15	(2.48 , 3.82)		3.85	(3.03 , 4.68)
CM BANK	5	1801	3.57	(2.53 , 4.62)	1759	4.48	(3.16 , 5.81)
	10		2.93	(2.33 , 3.54)		3.46	(2.74 , 4.18)
FUYAO GLASS	5	2027	3.90	(2.83 , 4.98)	2046	4.74	(3.44 , 6.05)
	10		3.20	(2.58 , 3.82)		3.50	(2.82 , 4.18)
SH PHARMA	5	1978	3.95	(2.85 , 5.04)	1983	4.75	(3.42 , 6.07)
	10		3.10	(2.49 , 3.71)		3.42	(2.75 , 4.10)
CPIC	5	1110	4.40	(2.77 , 6.03)	1184	4.87	(3.11 , 6.62)
	10		3.75	(2.77 , 4.74)		3.60	(2.68 , 4.52)
GOLDWIND	5	1112	4.41	(2.78 , 6.05)	1162	4.49	(2.85 , 6.12)
	10		3.49	(2.57 , 4.41)		3.59	(2.67 , 4.51)
FOSUN PHARMA	5	1960	4.22	(3.04 , 5.40)	2123	4.16	(3.04 , 5.27)
	10		3.24	(2.60 , 3.88)		3.35	(2.71 , 3.98)
CIMC	5	2006	3.98	(2.87 , 5.08)	2094	4.82	(3.52 , 6.12)
	10		3.19	(2.57 , 3.82)		3.80	(3.07 , 4.53)
MINSHENG BANK	5	1893	3.76	(2.69 , 4.83)	1893	3.91	(2.80 , 5.02)
	10		3.22	(2.57 , 3.87)		3.29	(2.62 , 3.95)
CHENMING PAPER	5	1832	4.15	(2.95 , 5.35)	1967	4.78	(3.44 , 6.12)
	10		2.99	(2.38 , 3.60)		3.56	(2.85 , 4.26)
GF SEC	5	1653	3.84	(2.67 , 5.01)	1588	1.84	(1.27 , 2.41)
	10		3.24	(2.54 , 3.94)		2.28	(1.78 , 2.78)
DZUG	5	1977	5.05	(3.65 , 6.46)	2036	5.58	(4.05 , 7.11)
	10		3.14	(2.52 , 3.76)		3.36	(2.71 , 4.02)
LIVZON PHARMA	5	1884	3.60	(2.57 , 4.63)	1988	4.80	(3.46 , 6.14)
	10		3.13	(2.50 , 3.76)		3.62	(2.91 , 4.33)

Company	Truncation (%)	N_1	$\hat{\zeta}^-$	95%CI^-	N_2	$\hat{\zeta}^+$	95%CI^+
ZOOMLION	5	1886	4.70	(3.36 , 6.05)	1935	4.86	(3.50 , 6.23)
	10		3.17	(2.53 , 3.80)		3.25	(2.60 , 3.90)
ZTE	5	2017	3.87	(2.81 , 4.94)	2055	4.08	(2.97 , 5.20)
	10		3.27	(2.63 , 3.91)		3.31	(2.67 , 3.95)
BANKCOMM	5	1161	3.19	(2.03 , 4.35)	1151	3.40	(2.16 , 4.64)
	10		2.96	(2.20 , 3.72)		2.97	(2.20 , 3.74)
CHINA OILFIELD	5	1152	4.73	(3.01 , 6.46)	1192	4.23	(2.71 , 5.74)
	10		3.37	(2.50 , 4.24)		3.27	(2.44 , 4.10)
COSCO SHIP DEV	5	1060	4.47	(2.77 , 6.18)	1028	12.54	(7.67 , 17.41)
	10		3.18	(2.32 , 4.03)		2.88	(2.09 , 3.66)
CHINA LIFE	5	1300	4.58	(3.00 , 6.15)	1231	4.27	(2.76 , 5.77)
	10		3.36	(2.54 , 4.17)		3.04	(2.28 , 3.80)
CHALCO	5	1176	6.28	(4.02 , 8.55)	1188	9.62	(6.15 , 13.09)
	10		3.62	(2.70 , 4.54)		3.22	(2.41 , 4.04)
WEICHAI POWER	5	1193	4.75	(3.05 , 6.45)	1204	4.44	(2.85 , 6.03)
	10		3.43	(2.56 , 4.30)		3.56	(2.66 , 4.46)
PING AN	5	1205	3.81	(2.44 , 5.17)	1239	4.95	(3.21 , 6.70)
	10		3.19	(2.39 , 3.99)		3.55	(2.66 , 4.43)
COSCO SHIP HOLD	5	1123	5.27	(3.32 , 7.22)	1065	52.94	(32.79 , 73.1)
	10		3.50	(2.58 , 4.41)		3.15	(2.30 , 3.99)
CHINA COAL	5	1110	3.90	(2.45 , 5.34)	1107	4.81	(3.01 , 6.61)
	10		2.91	(2.14 , 3.67)		3.29	(2.42 , 4.15)
COSCO SHIP ENGY	5	1659	4.69	(3.26 , 6.12)	1759	4.90	(3.45 , 6.35)
	10		3.35	(2.63 , 4.07)		3.35	(2.65 , 4.04)
CHINA SHENHUA	5	1153	3.87	(2.46 , 5.27)	1154	3.59	(2.28 , 4.90)
	10		3.22	(2.39 , 4.05)		2.94	(2.18 , 3.70)
HUADIAN POWER	5	1386	4.19	(2.79 , 5.59)	1416	4.26	(2.86 , 5.66)
	10		3.03	(2.32 , 3.74)		3.08	(2.36 , 3.79)
CHINA SOUTH AIR	5	1557	4.77	(3.27 , 6.27)	1601	6.51	(4.49 , 8.52)
	10		3.16	(2.46 , 3.86)		3.46	(2.70 , 4.22)
ANHUIEXPRESSWAY	5	1609	3.14	(2.17 , 4.12)	1708	3.45	(2.41 , 4.49)
	10		2.62	(2.05 , 3.19)		2.83	(2.23 , 3.43)
DATANG POWER	5	1179	4.71	(3.01 , 6.41)	1231	5.49	(3.55 , 7.42)
	10		3.11	(2.31 , 3.90)		3.17	(2.38 , 3.97)
CCB	5	1076	3.21	(2.00 , 4.42)	1094	3.91	(2.45 , 5.37)
	10		2.77	(2.03 , 3.51)		2.78	(2.05 , 3.52)
ANHUI CONCH	5	1770	4.00	(2.82 , 5.17)	1859	4.48	(3.19 , 5.77)
	10		3.39	(2.68 , 4.10)		3.54	(2.82 , 4.26)
HUANENG POWER	5	1779	3.72	(2.62 , 4.81)	1804	4.06	(2.88 , 5.25)

Table A3 – continued from previous page

Company	Truncation (%)	N_1	$\hat{\zeta}^-$	$95\%CI^-$	N_2	$\hat{\zeta}^+$	$95\%CI^+$
	10		2.97	(2.35 , 3.58)		3.34	(2.65 , 4.04)
BAIYUNSHAN PH	5	1846	5.21	(3.71 , 6.72)	1881	4.91	(3.51 , 6.32)
	10		3.30	(2.63 , 3.97)		3.53	(2.81 , 4.24)
PETROCHINA	5	1148	3.86	(2.44 , 5.28)	1090	2.82	(1.77 , 3.87)
	10		2.93	(2.17 , 3.68)		2.55	(1.87 , 3.22)
AIR CHINA	5	1237	5.50	(3.56 , 7.43)	1271	6.08	(3.97 , 8.19)
	10		3.16	(2.37 , 3.95)		4.11	(3.10 , 5.12)
BEIJING N STAR	5	1175	8.54	(5.46 , 11.62)	1277	6.45	(4.22 , 8.69)
	10		3.45	(2.57 , 4.33)		3.13	(2.37 , 3.90)
SHENZHENEXPRESS	5	1713	3.51	(2.46 , 4.56)	1798	3.63	(2.57 , 4.69)
	10		2.90	(2.29 , 3.52)		2.94	(2.33 , 3.54)
GUANGSHEN RAIL	5	1147	3.31	(2.09 , 4.52)	1244	3.68	(2.39 , 4.98)
	10		2.80	(2.07 , 3.52)		2.98	(2.24 , 3.72)
JIANGXI COPPER	5	1813	4.64	(3.29 , 5.99)	1811	3.32	(2.35 , 4.28)
	10		3.49	(2.77 , 4.21)		3.26	(2.59 , 3.93)
JIANGSU EXPRESS	5	1818	3.30	(2.34 , 4.26)	1924	3.32	(2.38 , 4.26)
	10		2.84	(2.26 , 3.43)		2.94	(2.35 , 3.53)
CHINA VANKE	5	1943	3.93	(2.82 , 5.03)	1990	4.71	(3.40 , 6.01)
	10		3.37	(2.70 , 4.04)		3.29	(2.65 , 3.94)

Table A4: The left and right tail index estimates $\hat{\zeta}^-$ and $\hat{\zeta}^+$ for H-shares

Company	Truncation (%)	N_1	$\hat{\zeta}^-$	$95\%CI^-$	N_2	$\hat{\zeta}^+$	$95\%CI^+$
BANK OF CHINA	5	1260	3.33	(2.17 , 4.49)	1244	2.65	(1.72 , 3.59)
	10		3.00	(2.26 , 3.74)		2.72	(2.04 , 3.40)
CRRC	5	1048	3.09	(1.90 , 4.27)	1005	2.39	(1.45 , 3.32)
	10		3.05	(2.23 , 3.88)		2.36	(1.71 , 3.01)
ICBC	5	1258	3.99	(2.60 , 5.39)	1257	2.72	(1.77 , 3.67)
	10		3.09	(2.32 , 3.85)		2.58	(1.94 , 3.21)
CHINA RAIL CONS	5	1131	3.28	(2.07 , 4.48)	1095	3.61	(2.26 , 4.96)
	10		3.30	(2.44 , 4.16)		3.03	(2.23 , 3.82)
YANZHOU COAL	5	2027	4.02	(2.91 , 5.12)	2011	3.51	(2.54 , 4.48)
	10		3.52	(2.83 , 4.20)		3.28	(2.64 , 3.92)
LUOYANG GLASS	5	1632	2.99	(2.07 , 3.91)	1411	2.77	(1.86 , 3.68)
	10		2.72	(2.13 , 3.31)		2.72	(2.09 , 3.36)
DONGFANG ELEC	5	1987	3.24	(2.33 , 4.14)	1928	3.59	(2.58 , 4.61)
	10		2.96	(2.37 , 3.54)		3.38	(2.70 , 4.05)
TIANJIN CAPITAL	5	1983	3.84	(2.77 , 4.91)	1868	3.68	(2.62 , 4.73)
	10		3.66	(2.94 , 4.38)		3.00	(2.39 , 3.61)
SINOPEC SSC	5	2003	3.66	(2.65 , 4.68)	1788	2.80	(1.98 , 3.63)
	10		3.35	(2.70 , 4.01)		2.82	(2.23 , 3.40)
CITIC BANK	5	1199	4.25	(2.73 , 5.77)	1190	2.85	(1.83 , 3.87)
	10		3.67	(2.74 , 4.59)		2.87	(2.14 , 3.61)
HISENSE KELON	5	1574	2.82	(1.94 , 3.70)	1469	2.93	(1.98 , 3.88)
	10		2.72	(2.12 , 3.32)		3.07	(2.37 , 3.77)
SHANDONG XINHUA	5	1832	3.89	(2.76 , 5.01)	1843	3.37	(2.39 , 4.34)
	10		3.27	(2.60 , 3.94)		2.99	(2.38 , 3.60)
CHINA EAST AIR	5	1982	3.36	(2.42 , 4.29)	1833	2.98	(2.12 , 3.84)
	10		3.16	(2.54 , 3.78)		2.96	(2.36 , 3.57)
NANJING PANDA	5	1986	3.01	(2.17 , 3.85)	1779	4.02	(2.84 , 5.21)
	10		2.95	(2.37 , 3.53)		3.46	(2.74 , 4.17)
CHINA RAILWAY	5	1169	3.21	(2.04 , 4.38)	1099	3.42	(2.14 , 4.69)
	10		3.26	(2.43 , 4.10)		2.95	(2.17 , 3.73)
SINOPEC CORP	5	1902	4.09	(2.93 , 5.25)	1938	3.40	(2.44 , 4.35)
	10		3.56	(2.85 , 4.28)		3.26	(2.61 , 3.91)
JINGWEI TEXTILE	5	1710	2.84	(1.99 , 3.69)	1637	3.58	(2.49 , 4.68)
	10		2.81	(2.21 , 3.41)		3.15	(2.47 , 3.84)
ANGANG STEEL	5	2085	3.81	(2.77 , 4.84)	1960	3.61	(2.60 , 4.62)
	10		3.57	(2.89 , 4.26)		3.48	(2.79 , 4.16)
SHANGHAI PECHEM	5	1954	4.08	(2.94 , 5.22)	1958	3.71	(2.67 , 4.75)

Company	Truncation (%)	N_1	$\hat{\zeta}^-$	$95\%CI^-$	N_2	$\hat{\zeta}^+$	$95\%CI^+$
	10		3.66	(2.93 , 4.39)		3.35	(2.69 , 4.01)
MAANSHAN IRON	5	2017	4.71	(3.41 , 6.01)	1849	3.36	(2.39 , 4.33)
	10		3.96	(3.18 , 4.73)		3.00	(2.39 , 3.62)
COMEC	5	1960	3.31	(2.38 , 4.24)	1811	3.42	(2.43 , 4.42)
	10		2.92	(2.34 , 3.49)		3.09	(2.45 , 3.72)
KUNMING MACHINE	5	1751	4.02	(2.83 , 5.21)	1600	3.66	(2.53 , 4.79)
	10		3.63	(2.87 , 4.39)		3.02	(2.36 , 3.68)
JINGCHENG MAC	5	1779	3.53	(2.49 , 4.56)	1632	3.78	(2.62 , 4.94)
	10		3.13	(2.48 , 3.77)		3.06	(2.40 , 3.73)
TSINGTAO BREW	5	1999	3.79	(2.74 , 4.84)	1997	3.25	(2.35 , 4.16)
	10		3.14	(2.52 , 3.75)		2.86	(2.30 , 3.43)
NE ELECTRIC	5	1786	2.65	(1.87 , 3.43)	1552	2.84	(1.95 , 3.73)
	10		2.44	(1.94 , 2.95)		2.61	(2.03 , 3.20)
CM BANK	5	1268	3.91	(2.54 , 5.28)	1304	3.13	(2.06 , 4.21)
	10		3.05	(2.30 , 3.80)		2.79	(2.11 , 3.47)
ZTE	5	1530	2.72	(1.86 , 3.58)	1472	3.60	(2.44 , 4.76)
	10		2.79	(2.17 , 3.42)		3.55	(2.74 , 4.37)
CMOC	5	1229	4.09	(2.64 , 5.55)	1170	2.88	(1.84 , 3.91)
	10		3.46	(2.60 , 4.33)		2.90	(2.15 , 3.64)
BANKCOMM	5	1407	3.67	(2.45 , 4.88)	1422	3.11	(2.09 , 4.13)
	10		3.16	(2.42 , 3.90)		2.86	(2.20 , 3.53)
ZIJIN MINING	5	1570	3.79	(2.61 , 4.98)	1507	4.05	(2.75 , 5.34)
	10		3.21	(2.50 , 3.92)		3.59	(2.78 , 4.41)
CHINA OILFIELD	5	1749	3.56	(2.50 , 4.61)	1676	3.51	(2.45 , 4.58)
	10		3.06	(2.42 , 3.70)		3.26	(2.56 , 3.96)
DALIAN PORT	5	1261	3.41	(2.22 , 4.61)	1211	2.59	(1.67 , 3.50)
	10		2.99	(2.25 , 3.73)		2.64	(1.98 , 3.31)
COSCO SHIP DEV	5	1507	3.86	(2.63 , 5.10)	1358	4.01	(2.66 , 5.36)
	10		3.30	(2.56 , 4.05)		3.24	(2.47 , 4.01)
SH ELECTRIC	5	1386	3.54	(2.36 , 4.71)	1356	2.73	(1.81 , 3.65)
	10		3.11	(2.38 , 3.84)		2.71	(2.07 , 3.36)
CHINA LIFE	5	1540	3.60	(2.46 , 4.74)	1579	3.80	(2.61 , 4.98)
	10		3.29	(2.56 , 4.03)		3.34	(2.60 , 4.07)
CHALCO	5	1842	4.23	(3.01 , 5.46)	1754	3.86	(2.72 , 5.00)
	10		3.92	(3.12 , 4.72)		3.57	(2.82 , 4.32)
WEICHAI POWER	5	1599	3.85	(2.65 , 5.04)	1563	3.94	(2.70 , 5.17)
	10		3.42	(2.67 , 4.17)		3.56	(2.77 , 4.35)
GREATWALL MOTOR	5	1620	3.50	(2.42 , 4.58)	1532	3.18	(2.18 , 4.18)
	10		3.21	(2.51 , 3.91)		3.15	(2.44 , 3.85)
							Continued on next page

Company	Truncation (%)	N_1	$\hat{\zeta}^-$	$95\%CI^-$	N_2	$\hat{\zeta}^+$	$95\%CI^+$
PING AN	5	1539	3.43	(2.34 , 4.51)	1528	3.77	(2.57 , 4.97)
	10		2.94	(2.28 , 3.59)		3.30	(2.56 , 4.04)
COSCO SHIP HOLD	5	1418	3.12	(2.10 , 4.15)	1323	3.77	(2.48 , 5.06)
	10		2.81	(2.16 , 3.47)		3.47	(2.63 , 4.30)
CHINA COAL	5	1298	2.92	(1.91 , 3.92)	1209	4.02	(2.58 , 5.45)
	10		2.80	(2.12 , 3.49)		3.69	(2.76 , 4.62)
CHINA COMM CONS	5	1319	3.41	(2.24 , 4.57)	1221	3.37	(2.17 , 4.56)
	10		3.22	(2.44 , 4.00)		3.08	(2.31 , 3.86)
BYD COMPANY	5	1778	3.29	(2.33 , 4.26)	1778	3.57	(2.52 , 4.62)
	10		3.02	(2.39 , 3.65)		3.38	(2.67 , 4.08)
COSCO SHIP ENGY	5	2046	3.36	(2.44 , 4.28)	1929	3.35	(2.40 , 4.30)
	10		3.18	(2.57 , 3.80)		3.35	(2.68 , 4.01)
CHINA SHENHUA	5	1417	3.11	(2.08 , 4.13)	1433	2.73	(1.84 , 3.63)
	10		3.07	(2.36 , 3.79)		2.94	(2.26 , 3.63)
HUADIAN POWER	5	1972	3.61	(2.60 , 4.62)	1864	3.85	(2.74 , 4.95)
	10		3.46	(2.77 , 4.14)		3.25	(2.59 , 3.91)
ZHEJIANG SHIBAO	5	1082	2.04	(1.27 , 2.80)	1053	2.47	(1.53 , 3.42)
	10		2.11	(1.54 , 2.67)		2.45	(1.79 , 3.12)
CHINA SOUTH AIR	5	1995	3.71	(2.68 , 4.74)	1934	3.86	(2.77 , 4.94)
	10		3.46	(2.79 , 4.14)		3.62	(2.90 , 4.34)
CHONGQING IRON	5	1922	3.24	(2.32 , 4.15)	1753	3.13	(2.20 , 4.05)
	10		3.11	(2.49 , 3.74)		2.86	(2.26 , 3.46)
ANHUIEXPRESSWAY	5	1949	4.33	(3.11 , 5.55)	1924	4.19	(3.00 , 5.38)
	10		3.63	(2.91 , 4.35)		3.83	(3.06 , 4.60)
DATANG POWER	5	1953	3.40	(2.45 , 4.36)	1906	3.33	(2.38 , 4.28)
	10		3.29	(2.64 , 3.95)		2.94	(2.35 , 3.53)
CCB	5	1349	3.91	(2.58 , 5.23)	1376	2.90	(1.93 , 3.87)
	10		3.03	(2.31 , 3.75)		2.84	(2.17 , 3.51)
ANHUI CONCH	5	1974	3.99	(2.88 , 5.10)	1977	3.78	(2.73 , 4.84)
	10		3.43	(2.75 , 4.10)		3.51	(2.82 , 4.21)
HUANENG POWER	5	1934	3.88	(2.79 , 4.97)	2005	3.74	(2.70 , 4.78)
	10		3.58	(2.87 , 4.30)		3.37	(2.71 , 4.03)
DONGJIANG ENV	5	1528	3.11	(2.12 , 4.10)	1403	3.08	(2.06 , 4.10)
	10		2.92	(2.27 , 3.58)		3.04	(2.33 , 3.76)
BAIYUNSHAN PH	5	1955	3.77	(2.72 , 4.83)	1864	4.50	(3.21 , 5.79)
	10		3.40	(2.73 , 4.08)		3.71	(2.95 , 4.46)
PETROCHINA	5	1913	3.57	(2.56 , 4.58)	2001	3.49	(2.53 , 4.46)
	10		3.36	(2.69 , 4.03)		3.10	(2.49 , 3.71)
XINHUA WINSHARE	5	1179	2.80	(1.79 , 3.82)	1141	2.77	(1.75 , 3.79)

44

Table A4 – continued from previous page

Company	Truncation (%)	N_1	$\hat{\zeta}^-$	95%CI^-	N_2	$\hat{\zeta}^+$	95%CI^+
	10		2.65	(1.98 , 3.33)		2.80	(2.08 , 3.53)
AIR CHINA	5	1479	3.55	(2.40 , 4.69)	1412	3.58	(2.40 , 4.76)
	10		3.26	(2.52 , 4.00)		3.28	(2.52 , 4.05)
BEIJING N STAR	5	1968	3.63	(2.61 , 4.64)	1854	3.06	(2.18 , 3.94)
	10		3.37	(2.71 , 4.04)		2.89	(2.30 , 3.48)
SHANDONG MOLONG	5	1512	3.85	(2.63 , 5.08)	1379	3.61	(2.40 , 4.81)
	10		3.55	(2.75 , 4.35)		2.89	(2.21 , 3.57)
SHENZHENEXPRESS	5	1897	4.48	(3.21 , 5.76)	1924	3.62	(2.59 , 4.64)
	10		3.58	(2.86 , 4.30)		3.38	(2.70 , 4.06)
GUANGSHEN RAIL	5	1932	4.49	(3.23 , 5.75)	1898	4.18	(2.99 , 5.37)
	10		3.85	(3.08 , 4.62)		3.37	(2.69 , 4.05)
JIANGXI COPPER	5	2047	4.09	(2.97 , 5.21)	2000	3.39	(2.45 , 4.33)
	10		3.41	(2.75 , 4.08)		3.14	(2.52 , 3.75)
JIANGSU EXPRESS	5	1937	4.08	(2.93 , 5.23)	2007	3.45	(2.49 , 4.40)
	10		3.64	(2.91 , 4.36)		3.20	(2.58 , 3.83)
SICHUAN EXPRESS	5	1871	3.49	(2.49 , 4.48)	1812	3.34	(2.37 , 4.31)
	10		3.23	(2.57 , 3.88)		2.95	(2.35 , 3.56)
FIRST TRACTOR	5	2095	3.72	(2.71 , 4.72)	1878	3.66	(2.61 , 4.71)
	10		3.37	(2.73 , 4.02)		3.27	(2.61 , 3.94)

Table A5: Tail index estimates $\hat{\zeta}^S$ and $\hat{\zeta}^D$ for single- and dual- listed period

Panel A. A-shares

Company	Truncation (%)	Single-listed period			Dual-listed period		
		N^S	$\hat{\zeta}^S$	$95\%CI^S$	N^D	$\hat{\zeta}^D$	$95\%CI^D$
HAITONG SEC	10	2402	4.17	(3.42 , 4.91)	1276	3.05	(2.31 , 3.80)
CITIC SEC	10	2042	3.90	(3.14 , 4.65)	1411	2.79	(2.14 , 3.44)
CM BANK	10	1034	3.15	(2.29 , 4.01)	2608	3.49	(2.89 , 4.08)
SH PHARMA	10	2571	3.51	(2.90 , 4.12)	1505	3.03	(2.35 , 3.71)
FOSUN PHARMA	10	3033	3.48	(2.92 , 4.03)	1138	2.88	(2.13 , 3.62)
CIMC	10	3033	3.85	(3.24 , 4.47)	1117	2.88	(2.13 , 3.64)
MINSHENG BANK	10	2103	3.61	(2.92 , 4.30)	1856	2.77	(2.20 , 3.33)
CHENMING PAPER	10	1768	3.16	(2.50 , 3.82)	2182	3.27	(2.66 , 3.88)
ZOOMLION	10	2339	3.55	(2.90 , 4.19)	1594	2.72	(2.13 , 3.32)
ZTE	10	1162	3.17	(2.36 , 3.99)	2975	3.46	(2.90 , 4.01)

Panel B. H-shares

Company	Truncation (%)	Single-listed period			Dual-listed period		
		N^S	$\hat{\zeta}^S$	$95\%CI^S$	N^D	$\hat{\zeta}^D$	$95\%CI^D$
CMOC	5	1340	3.36	(2.23 , 4.50)	1172	3.40	(2.18 , 4.63)
	10		3.21	(2.44 , 3.98)		3.34	(2.49 , 4.20)
ZIJIN MINING	5	1068	4.99	(3.09 , 6.89)	2251	3.51	(2.60 , 4.43)
	10		3.99	(2.92 , 5.07)		3.10	(2.53 , 3.67)
CHINA OILFIELD	5	1201	5.33	(3.42 , 7.24)	2422	3.15	(2.35 , 3.94)
	10		3.65	(2.73 , 4.58)		2.92	(2.40 , 3.44)
DALIAN PORT	5	1138	3.09	(1.96 , 4.22)	1635	2.65	(1.84 , 3.46)
	10		2.92	(2.17 , 3.68)		2.62	(2.05 , 3.19)
CHALCO	5	1316	5.31	(3.50 , 7.12)	2524	3.67	(2.76 , 4.58)
	10		4.59	(3.48 , 5.69)		3.40	(2.81 , 4.00)
GREATWALL MOTOR	5	1923	3.40	(2.44 , 4.36)	1419	2.89	(1.94 , 3.85)
	10		3.35	(2.68 , 4.02)		3.02	(2.32 , 3.73)
CHINA COMM CONS	5	1288	3.18	(2.08 , 4.28)	1326	3.52	(2.32 , 4.73)
	10		3.15	(2.38 , 3.92)		3.16	(2.40 , 3.92)
BYD COMPANY	5	2204	3.57	(2.63 , 4.51)	1488	3.34	(2.27 , 4.42)
	10		3.37	(2.74 , 4.00)		2.97	(2.30 , 3.65)
HUADIAN POWER	5	1256	3.83	(2.49 , 5.17)	3073	3.50	(2.72 , 4.28)
	10		3.43	(2.59 , 4.28)		3.23	(2.72 , 3.74)
ZHEJIANG SHIBAO	5	1328	2.31	(1.52 , 3.10)	1091	2.14	(1.34 , 2.94)
	10		2.34	(1.77 , 2.90)		2.09	(1.53 , 2.64)
CHONGQING IRON	5	1766	3.77	(2.66 , 4.89)	2469	2.69	(2.02 , 3.36)
	10		3.25	(2.57 , 3.92)		2.62	(2.16 , 3.08)

Table A5 – continued from previous page

Company	Truncation (%)	N^S	$\hat{\zeta}^S$	$95\%CI^S$	N^D	$\hat{\zeta}^D$	$95\%CI^D$
		Single-listed period			Dual-listed period		
DATANG POWER	5	1720	3.48	(2.44 , 4.52)	2608	3.15	(2.38 , 3.92)
	10		3.39	(2.67 , 4.10)		2.81	(2.33 , 3.30)
DONGJIANG ENV	5	2237	3.66	(2.70 , 4.61)	1231	2.40	(1.55 , 3.24)
	10		3.24	(2.64 , 3.84)		2.50	(1.88 , 3.13)
PETROCHINA	5	1867	3.66	(2.61 , 4.71)	2398	3.41	(2.55 , 4.27)
	10		3.34	(2.66 , 4.02)		3.10	(2.54 , 3.65)
BEIJING N STAR	5	1675	3.68	(2.57 , 4.79)	2644	2.97	(2.26 , 3.69)
	10		3.43	(2.70 , 4.17)		2.76	(2.29 , 3.24)
SHANDONG MOLONG	5	1540	3.76	(2.57 , 4.94)	1667	3.02	(2.10 , 3.94)
	10		3.18	(2.47 , 3.89)		2.88	(2.26 , 3.49)
GUANGSHEN RAIL	5	1723	4.40	(3.09 , 5.72)	2610	4.07	(3.09 , 5.06)
	10		3.41	(2.69 , 4.13)		3.45	(2.85 , 4.04)
SICHUAN EXPRESS	5	2337	3.44	(2.56 , 4.32)	1977	3.32	(2.40 , 4.24)
	10		3.01	(2.46 , 3.55)		3.06	(2.45 , 3.66)
FIRST TRACTOR	5	3104	3.72	(2.89 , 4.55)	1223	3.03	(1.96 , 4.11)
	10		3.41	(2.87 , 3.95)		2.87	(2.15 , 3.59)

Table A6: Information of companies in the sample

Company Name*	A-Share		H-share	
	Code	Date of listing	Code	Date of listing
BANK OF CHINA	601988.SH	2006-07-05	3988.HK	2006-06-01
CRRC	601766.SH	2008-08-18	1766.HK	2008-08-21
MCC	601618.SH	2009-09-21	1618.HK	2009-09-24
ICBC	601398.SH	2006-10-27	1398.HK	2006-10-27
NCI	601336.SH	2011-12-16	1336.HK	2011-12-15
ABC	601288.SH	2010-07-15	1288.HK	2010-07-16
CHINA RAIL CONS	601186.SH	2008-03-10	1186.HK	2008-03-13
YANZHOU COAL	600188.SH	1998-07-01	1171.HK	1998-04-01
LUOYANG GLASS	600876.SH	1995-10-31	1108.HK	1994-07-08
DONGFANG ELEC	600875.SH	1995-10-10	1072.HK	1994-06-06
TIANJIN CAPITAL	600874.SH	1995-06-30	1065.HK	1994-05-17
SINOPEC SSC	600871.SH	1995-04-11	1033.HK	1994-03-29
CITIC BANK	601998.SH	2007-04-27	0998.HK	2007-04-27
HISENSE KELON	000921.SZ	1999-07-13	0921.HK	1996-07-23
SHANDONG XINHUA	000756.SZ	1997-08-06	0719.HK	1996-12-31
CHINA EAST AIR	600115.SH	1997-11-05	0670.HK	1997-02-05
NANJING PANDA	600775.SH	1996-11-18	0553.HK	1996-05-02
CHINA RAILWAY	601390.SH	2007-12-03	0390.HK	2007-12-07
SINOPEC CORP	600028.SH	2001-08-08	0386.HK	2000-10-19
ANGANG STEEL	000898.SZ	1997-12-25	0347.HK	1997-07-24
SHANGHAI PECHEM	600688.SH	1993-11-08	0338.HK	1993-07-26
MAANSHAN IRON	600808.SH	1994-01-06	0323.HK	1993-11-03
COMEC	600685.SH	1993-10-28	0317.HK	1993-08-06
KUNMING MACHINE	600806.SH	1994-01-03	0300.HK	1993-12-07
JINGCHENG MAC	600860.SH	1994-05-06	0187.HK	1993-08-06
TSINGTAO BREW	600600.SH	1993-08-27	0168.HK	1993-07-15
NE ELECTRIC	000585.SZ	1995-12-13	0042.HK	1995-07-06
HTSC	601688.SH	2010-02-26	6886.HK	2015-06-01
HAITONG SEC	600837.SH	1994-02-24	6837.HK	2012-04-27
CEB BANK	601818.SH	2010-08-18	6818.HK	2013-12-20
EB SECURITIES	601788.SH	2009-08-18	6178.HK	2016-08-18
CMSC	600999.SH	2009-11-17	6099.HK	2016-10-07
CITIC SEC	600030.SH	2003-01-06	6030.HK	2011-10-06
CM BANK	600036.SH	2002-04-09	3968.HK	2006-09-22
FUYAO GLASS	600660.SH	1993-06-10	3606.HK	2015-03-31
SH PHARMA	601607.SH	1994-03-24	2607.HK	2011-05-20

Company Name	A-Share		H-share	
	Code	Date of listing	Code	Date of listing
CPIC	601601.SH	2007-12-25	2601.HK	2009-12-23
GOLDWIND	002202.SZ	2007-12-26	2208.HK	2010-10-08
CHINA VANKE	00002.SZ	1991-01-29	2202.HK	2014-06-25
FOSUN PHARMA	600196.SH	1998-08-07	2196.HK	2012-10-30
CIMC	000039.SZ	1994-04-08	2039.HK	2012-12-19
MINSHENG BANK	600016.SH	2000-12-19	1988.HK	2009-11-26
CHENMING PAPER	000488.SZ	2000-11-20	1812.HK	2008-06-18
GF SEC	000776.SZ	1997-06-11	1776.HK	2015-04-10
DZUG	600635.SH	1993-03-04	1635.HK	2016-12-05
LIVZON PHARMA	000513.SZ	1993-10-28	1513.HK	2014-01-16
ZOOMLION	000157.SZ	2000-10-12	1157.HK	2010-12-23
ZTE	000063.SZ	1997-11-18	0763.HK	2004-12-09
ZMJ	601717.SH	2010-08-03	0564.HK	2012-12-05
CGS	601881.SH	2017-01-23	6881.HK	2013-05-22
CMOC	603993.SH	2012-10-09	3993.HK	2007-04-26
BANKCOMM	601328.SH	2007-05-15	3328.HK	2005-06-23
ZIJIN MINING	601899.SH	2008-04-25	2899.HK	2003-12-23
CHINA OILFIELD	601808.SH	2007-09-28	2883.HK	2002-11-20
DALIAN PORT	601880.SH	2010-12-06	2880.HK	2006-04-28
COSCO SHIP DEV	601866.SH	2007-12-12	2866.HK	2004-06-16
SH ELECTRIC	601727.SH	2008-12-05	2727.HK	2005-04-28
CHINA LIFE	601628.SH	2007-01-09	2628.HK	2003-12-18
CHALCO	601600.SH	2007-04-30	2600.HK	2001-12-12
WEICHAI POWER	000338.SZ	2007-04-30	2338.HK	2004-03-11
GREATWALL MOTOR	601633.SH	2011-09-28	2333.HK	2003-12-15
PING AN	601318.SH	2007-03-01	2318.HK	2004-06-24
GAC GROUP	601238.SH	2012-03-29	2238.HK	2010-08-30
BBMG	601992.SH	2011-03-01	2009.HK	2009-07-29
COSCO SHIP HOLD	601919.SH	2007-06-26	1919.HK	2005-06-30
CHINA COAL	601898.SH	2008-02-01	1898.HK	2006-12-19
CHINA COMM CONS	601800.SH	2012-03-09	1800.HK	2006-12-15
BYD COMPANY	002594.SZ	2011-06-30	1211.HK	2002-07-31
COSCO SHIP ENGY	600026.SH	2002-05-23	1138.HK	1994-11-11
CHINA SHENHUA	601088.SH	2007-10-09	1088.HK	2005-06-15
HUADIAN POWER	600027.SH	2005-02-03	1071.HK	1999-06-30
ZHEJIANG SHIBAO	002703.SZ	2012-11-02	1057.HK	2006-05-16
CHINA SOUTH AIR	600029.SH	2003-07-25	1055.HK	1997-07-31
CHONGQING IRON	601005.SH	2007-02-28	1053.HK	1997-10-17

Company Name	A-Share		H-share	
	Code	Date of listing	Code	Date of listing
ANHUIEXPRESSWAY	600012.SH	2003-01-07	0995.HK	1996-11-13
DATANG POWER	601991.SH	2006-12-20	0991.HK	1997-03-21
CCB	601939.SH	2007-09-25	0939.HK	2005-10-27
ANHUI CONCH	600585.SH	2002-02-07	0914.HK	1997-10-21
HUANENG POWER	600011.SH	2001-12-06	0902.HK	1998-01-21
DONGJIANG ENV	002672.SZ	2012-04-26	0895.HK	2003-01-29
BAIYUNSHAN PH	600332.SH	2001-02-06	0874.HK	1997-10-30
PETROCHINA	601857.SH	2007-11-05	0857.HK	2000-04-07
XINHUA WINSHARE	601811.SH	2016-08-08	0811.HK	2007-05-30
AIR CHINA	601111.SH	2006-08-18	0753.HK	2004-12-15
BEIJING N STAR	601588.SH	2006-10-16	0588.HK	1997-05-14
SHANDONG MOLONG	002490.SZ	2010-10-21	0568.HK	2004-04-15
SHENZHENEXPRESS	600548.SH	2001-12-25	0548.HK	1997-03-12
GUANGSHEN RAIL	601333.SH	2006-12-22	0525.HK	1996-05-14
JIANGXI COPPER	600362.SH	2002-01-11	0358.HK	1997-06-12
JIANGSU EXPRESS	600377.SH	2001-01-16	0177.HK	1997-06-27
SICHUAN EXPRESS	601107.SH	2009-07-27	0107.HK	1997-10-07
FIRST TRACTOR	601038.SH	2012-08-08	0038.HK	1997-06-23
JINGWEI TEXTILE**	000666.SZ	1996-12-10	0350.HK	1996-02-02

*Company names in this table are their short names in HKEx.

**JINGWEI TEXTILE was delisted from HKEx on 28 December 2015.