Journal of Financial Economics 134 (2019) 48–69

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Size and value in China[☆]

Jianan Liu[a], Robert F. Stambaugh[a,d], Yu Yuan[b,c,*]

[a] The Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, United States
[b] Shanghai Mingshi Investment Co., Ltd., 2101 Taiping Finance Tower, No. 488 Middle Yincheng Road, Shanghai 200120, China
[c] Wharton Financial Institutions Center, University of Pennsylvania, United States
[d] NBER, United States

ARTICLE INFO

Article history:
Received 2 July 2018
Revised 27 August 2018
Accepted 28 August 2018
Available online 15 March 2019

JEL classification:
G12
G14
G15

Keywords:
China
Size
Value
Factors
Anomalies

ABSTRACT

We construct size and value factors in China. The size factor excludes the smallest 30% of firms, which are companies valued significantly as potential shells in reverse mergers that circumvent tight IPO constraints. The value factor is based on the earnings-price ratio, which subsumes the book-to-market ratio in capturing all Chinese value effects. Our three-factor model strongly dominates a model formed by just replicating the Fama and French (1993) procedure in China. Unlike that model, which leaves a 17% annual alpha on the earnings-price factor, our model explains most reported Chinese anomalies, including profitability and volatility anomalies.

1. Introduction

China has the world's second-largest stock market, helping to finance an economy that some predict will be the world's largest within a decade.[1] China also has political and economic environments quite different from those in the US and other developed economies. Moreover, China's market and investors are separated from the rest of the world. China largely prohibits participation by foreign investors in its domestic stock market as well as participation by its domestic investors in foreign markets.[2]

Factor models provide a cornerstone for investigating financial asset pricing and for developing investment strategies. Many studies of China's stock market use a three-factor model constructed by following the Fama and French (1993) procedure for US factors.[3] The advisability of simply

☆ We are grateful for comments from the referee (Eugene Fama), Jules van Binsbergen, Danting Chang, Zhi Da, Zhe Geng, Dong Lou, Jun Qian (discussant), Amir Yaron, Jianfeng Yu, seminar participants at the Wharton School, and conference participants at the 2018 Johns Hopkins Carey Finance Conference and the 2019 American Finance Association meetings. We also thank Wharton Research Data Services (WRDS) for providing the updated factors in our Chinese three-factor model on their website.

* Corresponding author at: Shanghai Mingshi Investment Co., Ltd., 2101 Taiping Finance Tower, No. 488 Middle Yincheng Road, Shanghai 200120, China.

E-mail addresses: jiananl@wharton.upenn.edu (J. Liu), stambaugh@wharton.upenn.edu (R.F. Stambaugh), yuanyu@mingshiim.com (Y. Yuan).

[1] According to the World Bank, the 2016 equity values of listed domestic companies, in trillions of US dollars, are 27.4 in the US and 7.3 in China, followed by 5.0 in Japan. For a forecast that China's gross domestic product will reach that of the US by 2028, see Bloomberg (https://www.bloomberg.com/graphics/2016-us-vs-china-economy/).

[2] At the end of 2016, 197 foreign institutions were authorized to invest in A-shares, China's domestically traded stocks, but with a quota of just 0.6% of total market value (and even less in earlier years). Chinese domestic investors can invest in international financial markets only through a limited authorized channel.

[3] Examples of such studies include include Yang and Chen (2003), Fan and Shan (2004), Wang and Chin (2004), Chen et al. (2010), Cheung et al. (2015), and Hu et al. (2019).

https://doi.org/10.1016/j.jfineco.2019.03.008
0304-405X/

replicating a US model in China is questionable, however, given China's separation and the many differences in economic and financial systems. We explore and develop factor models in China, allowing its unique environment to dictate alternative approaches.

We start by examining size and value effects in the Chinese market. These two effects have long been recognized elsewhere as important characteristics associated with expected return: Banz (1981) reports a firm-size effect, and Basu (1983) finds an effect for the earnings-price ratio, a popular value metric. Size and value are the most prominent characteristics used by many institutions to classify investment styles. The most widely used nonmarket factors in academic research are also size and value, following the influential study by Fama and French (1993). Our study reveals that size and value effects are important in China but with properties different from the US. We construct size and value factors for China.

The size factor is intended to capture size-related differences in stock risk and return that arise from size-related differences in the underlying businesses. In China, however, the stock of a small listed firm is typically priced to reflect a substantial component of value related not to the firm's underlying business but instead to the Chinese initial public offering (IPO) process. In China, the IPO market is strictly regulated, and a growing demand for public listing confronts the low processing capacity of the regulatory bureau to approve IPOs. As a consequence, private firms seek an alternative approach, a reverse merger, to become public in a timely manner. In a reverse merger, a private firm targets a publicly traded company, a so-called shell, and gains control rights by acquiring its shares. The shell then buys the private firm's assets in exchange for newly issued shares. While reverse mergers occur elsewhere, IPO constraints are sufficiently tight in China such that the smallest firms on the major exchanges become attractive shell targets, unlike in the US, for example.

The smallest listed firms are the most likely shells. In fact, 83% of the reverse mergers in China involve shells coming from the smallest 30% of stocks. For a typical stock in the bottom 30%, we estimate that roughly 30% of its market value reflects its potential shell value in a reverse merger. Our estimate combines the empirical probability of being targeted in a reverse merger with the average return accompanying that event. Consistent with the contamination of small-firm stock prices by shell value, we also find that when compared to other firms, the smallest 30% have returns less related to operating fundamentals, proxied by earnings surprises, but more related to IPO activity. Therefore, to avoid shell-value contamination when constructing any of our factors, we delete the bottom 30% of stocks, which account for 7% of the stock market's total capitalization.

The value effect in China is best captured by the earnings-price (*EP*) ratio, versus other valuation ratios. Following Fama and French (1992), we treat the choice among alternative valuation ratios as an empirical question, asking which variable best captures the cross-sectional variation in average stock returns. As in that study, we run a horse race among all candidate valuation ratios, including *EP*, book-to-market (*BM*), asset-to-market, and cash-flow-to-price ratios. In a Fama–MacBeth regression including those four ratios, *EP* dominates all others, just as Fama and French (1992) find *BM* dominates in the US market. Relying on the latter US result, Fama and French (1993) use *BM* to construct their value factor. Relying on our result for China, we use *EP* to construct our value factor.

Size and value are important factors in China, as revealed by their average premiums as well as their contributions to return variance. Our size and value factors both have average premiums exceeding 1% per month over our 2000–2016 sample period. For the typical stock in China, size and value jointly explain an additional 15% of monthly return variance beyond what the market factor explains. In comparison, size and value explain less than 10% of additional return variance for the typical US stock during the same period.

Our three-factor model, CH-3, includes the market factor as well as size and value factors incorporating the above China-specific elements. For comparison, we construct an alternative three-factor model, FF-3, by simply replicating the Fama and French (1993) procedure. We find that CH-3 strongly dominates FF-3. Specifically, FF-3 cannot price the CH-3 size and value factors, which have (significant) FF-3 annualized alphas of 5.6% and 16.7%. In contrast, CH-3 prices the FF-3 size and value factors, which have (insignificant) CH-3 annualized alphas of just −0.5% and 4.1%. A Gibbons et al. (1989) test of one model's ability to price the other's factors gives a p-value of 0.41 for CH-3's pricing ability but less than 10^{-12} for FF-3's ability.

We also investigate the ability of CH-3 to explain previously reported return anomalies in China. A survey of the literature reveals anomalies in nine categories: size, value, profitability, volatility, return reversal, turnover, investment, accruals, and illiquidity. We find each of the first six categories contains one or more anomalies that produce significant long-short alphas with respect to the single-factor capital asset pricing model (CAPM). CH-3 accommodates all anomalies in the first four of those six categories, including profitability and volatility, whose anomalies fail FF-3 explanations in the US. CH-3 fails only with some of the reversal and turnover anomalies. In contrast, FF-3 leaves significant anomalies in five of the six categories. A total of ten anomalies are unexplained by the CAPM; CH-3 explains eight of them, while FF-3 explains three. The average absolute CH-3 alpha for the ten anomalies is 5.4% annualized, compared to 10.8% for FF-3 (average absolute t-statistics: 1.12 versus 2.70).

Hou et al. (2015) and Fama and French (2015) add two factors based on investment and profitability measures in their recently proposed factor models, Q-4 and FF-5. Investment does not produce a significant CAPM alpha in China, and profitability is fully explained by CH-3. In an analysis reported in the Appendix, we find that a replication of FF-5 in China is dominated by CH-3.

Overall, CH-3 performs well as a factor model in China, and it captures most documented anomalies. In US studies, researchers often supplement the usual three factors (market, size, and value) with a fourth factor, such as the momentum factor of Carhart (1997) or the liquidity factor of Pástor and Stambaugh (2003). We also add a fourth factor, motivated by a phenomenon rather unique to China: a stock market dominated by individuals rather than institutions. Over 101 million individuals have stock trading accounts in China, and individuals own 88% of the market's free-floating shares. This heavy presence of individuals makes Chinese stocks especially susceptible to investor sentiment. To capture sentiment effects, we base our fourth factor on turnover, which previous research identifies as a gauge of both market-wide and stock-specific investor sentiment (e.g., Baker and Stein, 2004; Baker and Wurgler, 2006; Lee, 2013). The resulting four-factor model, CH-4, explains the turnover and reversal anomalies in addition to the anomalies explained by CH-3, thereby handling all of China's reported anomalies.

The remainder of the paper proceeds as follows. Section 2 discusses data sources and sample construction. Section 3 addresses the interplay between firm size and China's IPO constraints and explores the importance of shell-value distortions in small-stock returns. Section 4 investigates value effects in China. In Section 5, we construct CH-3 and FF-3 and compare their abilities to price each other's factors. In Section 6, we compare the abilities of those three-factor models to price anomalies. In Section 7, we construct CH-4 by including a turnover factor and then analyze the model's additional pricing abilities. Section 8 summarizes our conclusions.

2. Data source and samples

The data we use, which include data on returns, trading, financial statements, and mergers and acquisitions, are from Wind Information Inc. (WIND), the largest and most prominent financial data provider in China. WIND serves 90% of China's financial institutions and 70% of the Qualified Foreign Institutional Investors (QFII) operating in China.

The period for our main analysis is from January 1, 2000, through December 31, 2016. China's domestic stock market, the A-share market, began in 1990 with the establishment of the Shanghai and Shenzhen exchanges. We focus on the post-2000 period for two reasons. The first is to assure uniformity in accounting data. The implementation of rules and regulations governing various aspects of financial reporting in China did not largely take shape until about 1999. Although 1993 saw the origination of principles for fair trade and financial disclosure, firms received little guidance in meeting them. Companies took liberties and imposed their own standards, limiting the comparability of accounting data across firms. Not until 1998 and 1999 were laws and regulations governing trading and financial reporting more thoroughly designed and implemented. For example, detailed guidelines for corporate operating revenue disclosure were issued in December 1998 and implemented in January 1999. Securities laws were passed in December 1998 and implemented in July 1999. Only by 1999 did uniformity in accounting standards become widely accomplished. Because portfolios formed in 2000 use accounting data for 1999, our post-2000 sample for portfolio returns relies on accounting data more comparable across firms than in earlier years.

The second reason for beginning our sample in 2000 is to ensure sufficient numbers of observations. Portfolios are used in our study to construct factors and conduct many of the tests. To enable reasonable precision and power, we require at least 50 stocks in all portfolios after imposing our filters, which include eliminating stocks (i) in the bottom 30% of firm size, (ii) listed less than six months, and (iii) having less than 120 trading records in the past year or less than 15 trading records in the past month. This last pair of conditions is intended to prevent our results from being influenced by returns that follow long trading suspensions. Only by 1999 do the numbers of stocks in the market allow these criteria to be met.

WIND's data on reverse mergers begin in 2007, when the China Securities Regulatory Commission identified the criteria of a merger and acquisition (M&A) proposal that classify it as a reverse merger, making such deals easier to trace. In Section 3.2, we use reverse merger data to estimate shell values. Additional details about the data and the construction of empirical measures are provided in the Appendix.

3. Small stocks and IPO constraints

Numerous studies in finance address China's unique characteristics. For example, Allen et al. (2003, 2005) compare China to other developed countries along various political, economic, and financial dimensions. Brunnermeier et al. (2017) study China's government interventions in its trading environment. Bian et al. (2018) show the special nature of leveraged investors in China's stock market. Song and Xiong (2018) emphasize the necessity of accounting for the economy's uniqueness when analyzing risks in China's financial system. Allen et al. (2009) and Carpenter and Whitelaw (2017) provide broader overviews of China's financial environment.

One aspect of the Chinese market especially relevant for our study is the challenge faced by firms wishing to become publicly traded. As discussed earlier, market values of the smallest firms in China include a significant component reflecting the firms' potential to be shells in reverse mergers. Private firms often employ reverse mergers to become publicly traded rather than pursue the constrained IPO process. Section 3.1 describes that IPO process, while Section 3.2 describes reverse mergers and presents a notable example of one in China. In Section 3.3, we compute a simple estimate of the fraction of firm value associated with being a shell for a potential reverse merger, and we find the fraction to be substantial for the smallest stocks. Consistent with that result, we show in Section 3.4 that the returns on those stocks exhibit significantly less association with their underlying firms' fundamentals.

Our evidence demonstrating the importance of the shell component of small-firm values is buttressed by contemporaneous research on this topic, conducted independently from ours. In a study whose principal focus is the importance of shell values in China, Lee et al. (2017) also show that the shell component is a substantial fraction of small-firm values and that, as a result, the returns on small-firm stocks exhibit less sensitivity to fundamentals but greater sensitivity to IPO activity. Lee et al. (2017) explore models for pricing the shell-value firms, whereas we focus on models for pricing the "regular" stocks constituting the other 93% of the stock market's value.

3.1. The IPO process

In China, the IPO market is controlled by the China Securities Regulatory Commission (CSRC). As a central planner, the CSRC constrains the IPO process to macro-manage the total number of listed firms (e.g., Allen et al., 2014). Unlike the US, where an IPO can clear regulatory scrutiny in a matter of weeks, undertaking an IPO in China is long and tedious, easily taking three years and presenting an uncertain outcome. As detailed in the Appendix, the process involves seven administrative steps, three bureau departments, and a select 25-member committee that votes on each application. The committee meets for both an initial review and a final vote, with those meetings separated by years. As of November 2017, the CSRC reported 538 firms being processed, with just 31 having cleared the initial review. The IPOs approved in early 2017 all entered the process in 2015.

The long waiting time can impose significant costs. During the review process, firms are discouraged from any sort of expansion and must produce consistent quarterly earnings. Any change in operations can induce additional scrutiny and further delay. A firm undertaking an IPO may thus forgo substantial investment opportunities during the multi-year approval process. Moreover, policy changes can prolong the process even more. In 2013, the CSRC halted all reviews for nearly a year to cool down the secondary market.

3.2. Reverse mergers

Facing the lengthy IPO process, private firms wishing to become public often opt for an alternative: reverse merger. A reverse merger, which is regulated as an M&A, involves fewer administrative steps and is much faster. We illustrate the process via a real-life case involving the largest delivery company in China, SF Express (SF).

In 2016, SF decided to become public through a reverse merger. To be its shell firm, SF targeted the small public company, DT Material (DT), with market value of about \$380 million. SF and DT agreed on merger terms, and in May 2016, DT officially announced the deal to its shareholders. At the same time, DT submitted a detailed M&A proposal to the CSRC. The plan had DT issuing more than three billion shares to SF in exchange for all of SF's assets. The intent was clear: three billion shares would account for 97% of DT's stock upon the shares' issuance. With those shares, SF would effectively be the sole owner of DT, which would in turn be holding all of SF's assets. DT would become essentially the same old SF company but with publicly traded status. The M&A authorization went smoothly. By October 2016, five months after the application, the CSRC gave its conditional approval, and final authorization came two months later. The merged company was trading as SF on the Shenzhen Stock Exchange by February 2017. That same month, IPO applicants in the 2015 cohort had just begun their initial reviews.

The entire SF-DT process took less than a year, fairly typical for a reverse merger. The greater speed of a reverse merger comes with a price tag, however. In addition to regular investment banking and auditing fees, the private firm bears the cost of acquiring control of the public shell firm. In the SF-DT case, DT kept 3% of the new public SF's shares, worth about \$938 million. In the course of the deal, DT's original shareholders made about 150%.

Reverse mergers also occur in the US. As in China, they have long been recognized as an IPO alternative. From 2000 through 2008, the US averaged 148 reverse mergers annually (Floros and Sapp, 2011). There is, however, a fundamental difference between reverse mergers in the US versus China: because IPOs are less constrained in the US, the value of being a potential shell is much lower. In the US, the median shell's equity market value is only \$2 million (Floros and Sapp, 2011), versus an average of \$200 million in China. Nearly all shell companies in the US have minimal operations and few noncash assets. Their Chinese counterparts are typically much more expensive operating businesses. As a result, while small stocks on China's major exchanges are attractive shell targets, small stocks on the major US exchanges are not. Consistent with this difference, Floros and Sapp (2011) observe that their US reverse-merger sample includes almost no shell targets listed on the three major exchanges.

3.3. Small stocks with large shell values

A private firm's price tag for acquiring a reverse-merger shell depends essentially on the shell's market value. Not surprisingly, shells are most often small firms. Fig. 1 displays the size distribution of public shells in our sample of reverse mergers covering the 2007–2016 period. Of the 133 reverse mergers, 83% come from the bottom 30%, and more than half come from the bottom 10%. Given this evidence, we eliminate the bottom 30% when constructing factors to avoid much of the contamination of stock prices reflecting the potential to be targeted as shells. Although the 30% cutoff is somewhat arbitrary, our results are robust to using 25% and 35% as cutoffs.

What fraction of a firm's market value owes to the firm potentially becoming a reverse-merger shell? A back-of-the-envelope calculation suggests the fraction equals roughly 30% for the stocks we eliminate (the bottom 30%). Let p denote the probability of such a stock becoming a reverse-merger shell in any given period, and let G denote the stock's gain in value if it does become a shell. We can then compute the current value of this potential lottery-like payoff on the stock as

$$S = \frac{pG + (1-p)S}{1+r} = \frac{pG}{r+p}, \tag{1}$$

where r is the discount rate. We take p to be the annual rate at which stocks in the bottom 30% become reverse-merger shells, and we take G to be the average accompanying increase in stock value. Both quantities are estimated over a two-year rolling window. The annual discount rate, r, is set to 3%, the average one-year deposit rate from 2007 to 2016.

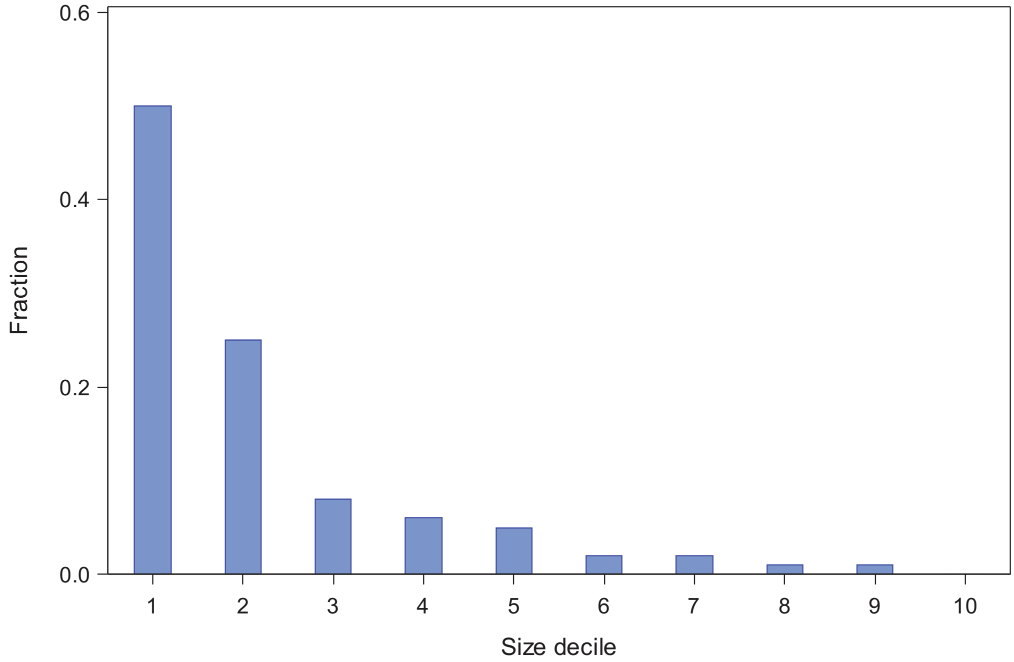

Fig. 1. Firm-size distribution of reverse-merger shells. The figure displays the size distribution of firms acquired in reverse-merger deals from January 2007 through December 2016. A total of 133 reverse-merger deals occurred, and the fraction of those deals falling into a given firm size decile is displayed in the bar chart. Size deciles reflect month-end market values three months before the deal month.

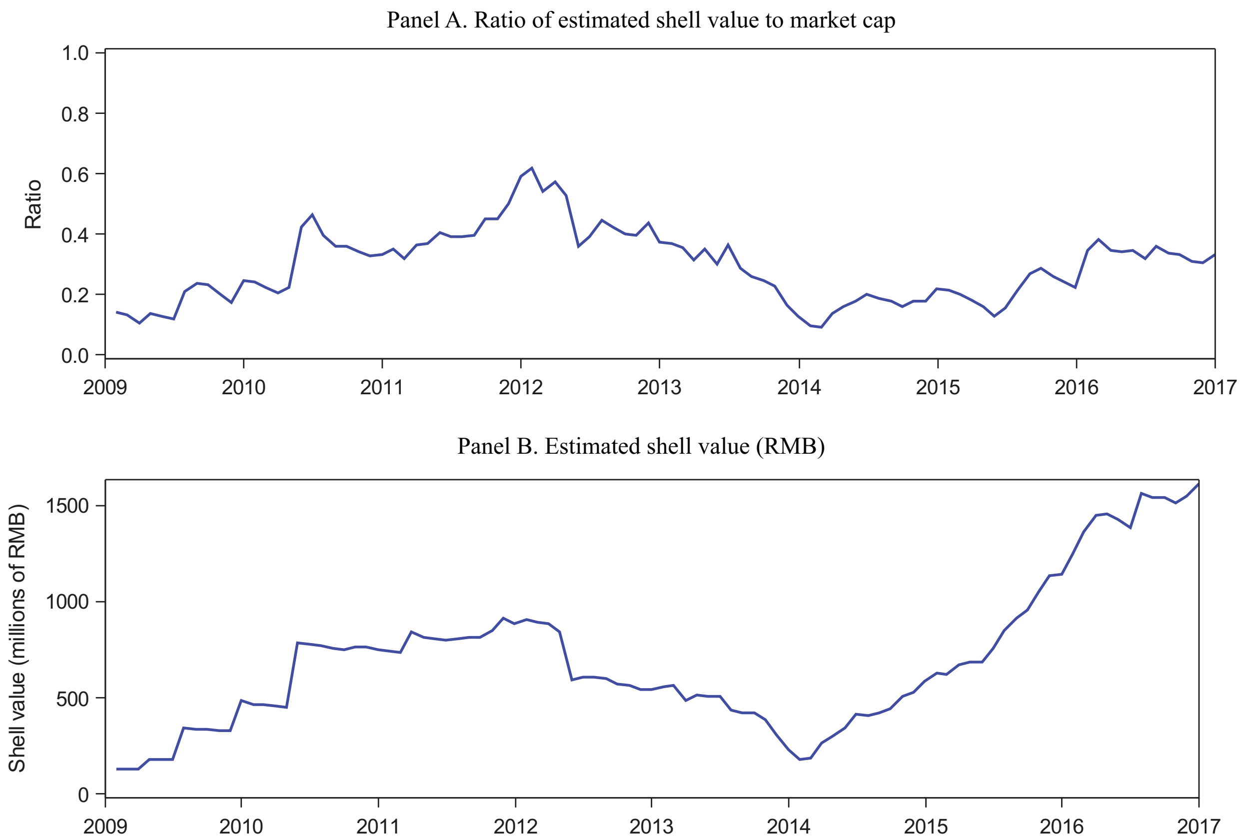

Fig. 2. Shell values over time. Panel A displays the time series of the ratio of estimated shell value to firm market capitalization. Panel B displays the time series of the estimated shell value (in RMB). The sample period is January 2009 through December 2016.

Panel A of Fig. 2 plots the estimated daily ratio of shell value to market value, S/V, with V equal to the median market value of stocks in the bottom 30%. Over the 2009–2016 sample period, the average value of S/V is 29.5%, while the series fluctuates between 10% and 60%. Eq. (1) implicitly assumes the stock remains a potential shell in perpetuity, until becoming a shell. In other words, the role of small stocks in reverse mergers is assumed to be rather permanent in China, as the IPO regulatory environment shows no overall trend toward loosening. Even if we reduce the horizon to 20 years, the average S/V remains about half as large as the series plotted.

Panel B of Fig. 2 plots the estimated shell value, S, expressed in renminbi (RMB). This value exhibits a fivefold increase over the eight-year sample period, in comparison to barely a twofold increase for the Shanghai–Shenzhen 300 index over the same period. The rise in S is consistent with the significant premium earned over the period by stocks in the bottom 30% of the size distribution. Recall, however, that these stocks account for just 7% of the stock market's total capitalization. As we demonstrate later, constructing factors that include these stocks, whose returns are distorted by the shell component, impairs the ability of those factors to price the regular stocks that constitute the other 93% of stock market value.

3.4. Return variation of small stocks

Given that the shell component contributes heavily to the market values and average returns of the smallest stocks, we ask whether this component also contributes to variation in their returns. If it does, then when compared to other stocks, returns on the smallest stocks should be explained less by shocks to underlying fundamentals but more by shocks to shell values. We explore both implications.

To compare responses to fundamentals, we analyze returns accompanying earnings announcements. We divide the entire stock universe into three groups, using the 30th and 70th size percentiles. Within each group, we estimate a panel regression of earnings-window abnormal return on standardized unexpected earnings (*SUE*),

$$R_{i,t-k,t+k} = a + b\,SUE_{i,t} + e_{i,t}, \tag{2}$$

in which earnings are announced on day t, and $R_{i,t-k,t+k}$ is the cumulative return on stock i, in excess of the market return, over the surrounding trading days from $t-k$ through $t+k$. We compute $SUE_{i,t}$ using a seasonal random walk, in which $SUE_{i,t} = \Delta_{i,t}/\sigma(\Delta_i)$, $\Delta_{i,t}$ equals the year-over-year change in stock i's quarterly earnings, and $\sigma(\Delta_i)$ is the standard deviation of $\Delta_{i,t}$ for the last eight quarters.

Under the hypothesis that the shell component is a significant source of return variation for the smallest stocks, we expect those stocks to have a lower b in Eq. (2) and a lower regression R^2 than the other groups. The first three columns of Table 1 report the regression results, which confirm our hypothesis. Panel A contains results for $k=0$ in Eq. (2), and Panel B has results for $k=3$. In both panels, the smallest stocks have the lowest values of b and R^2. For comparison, we conduct the same analysis in the US and report the results in the last three columns of Table 1. The US sample period is 1/1/1980–12/31/2016, before which the quality of quarterly data is lower. In contrast to the results for China, the smallest stocks in the US have the highest values of b and R^2.

We also compare stocks' return responses to shell-value shocks, using two proxies for such shocks. One is the average return that public stocks experience upon becoming shells in reverse mergers. Our rationale is that the higher the return, the greater is the potential value of becoming a shell. The other proxy is the log of the total number of IPOs, with the rationale that a greater frequency of IPOs could be interpreted by the market as a relaxing of IPO constraints. Consistent with the importance of the shell component for the smallest stocks, only that group's returns covary both positively with the reverse-merger premium and negatively with the log of the IPO number. The results are presented in the Appendix.

4. Value effects in China

A value effect is a relation between expected return and a valuation metric that scales the firm's equity price by an accounting-based fundamental. The long-standing intuition for value effects (e.g., Basu, 1983; Ball, 1992) is that a scaled price is essentially a catchall proxy for expected return: a higher (lower) expected return implies a lower (higher) current price, other things equal.

Our approach to creating a value factor in China follows the same path established by the two-study sequence of Fama and French (1992, 1993). Following Fama and French (1992), the first step is to select the valuation ratio exhibiting the strongest value effect among a set of candidate ratios. The valuation ratios Fama and French (1992) consider include *EP*, *BM*, and assets-to-market (*AM*). The authors find *BM* exhibits the strongest value effect, subsuming the other candidates. Based on that result, the subsequent study by Fama and French (1993) uses *BM* to construct the value factor (*HML*).

In this section, we conduct the same horse race among valuation ratios. Our entrants are the same as in Fama and French (1992), plus cash-flow-to-price (*CP*). As in that study, we estimate cross-sectional Fama and MacBeth (1973) regressions of individual monthly stock returns on the valuation ratios, with a stock's market capitalization and estimated CAPM beta (β) included in the regression. For the latter variable we use the beta estimated from the past year's daily returns, applying a five-lag Dimson (1979) correction. Following Fama and French (1992), we use *EP* to construct both EP^+ and a dummy variable, with EP^+ equal to *EP* when *EP* is positive, and zero otherwise, and with the dummy variable, $D(EP<0)$, equal to one when *EP* is negative, and zero otherwise. In the same manner, we construct CP^+ and $D(CP<0)$ from *CP*. Due to the shell-value contamination of returns discussed earlier, we exclude the smallest 30% of stocks.

Table 2 reports average slopes from the month-by-month Fama–MacBeth regressions. Similar to results in the US market, we see from column (1) that β does not enter significantly. Also as in the US, the size variable, $\log ME$, enters with a significantly negative coefficient that is insensitive to including β: in columns (2) and (3), without

Table 1
Return reactions to earnings surprises across different size groups in China and the US.

The table reports slope estimates and R-squares in a panel regression of earnings-window returns on earnings surprises,

$$R_{i,t-k,t+k} = a + b\,SUE_{i,t} + e_{i,t},$$

in which earnings are announced on day t; $R_{i,t-k,t+k}$ is the cumulative return on stock i, in excess of the market return, over the surrounding trading days from $t-k$ through $t+k$; $SUE_{i,t} = \Delta_{i,t}/\sigma(\Delta_i)$; $\Delta_{i,t}$ equals the year-over-year change in stock i's quarterly earnings; and $\sigma(\Delta_i)$ is the standard deviation of $\Delta_{i,t}$ for the last eight quarters. Panel A contains results for $k=0$; Panel B contains results for $k=3$. The regression is estimated within each of three size groups in both the China and US markets. The groups are formed based on the top 30%, middle 40%, and bottom 30% of the previous month's market capitalizations. The sample periods are January 2000 through December 2016 for China and January 1980 through December 2016 for the US. The US returns data are from the Center for Research in Security Prices (CRSP) and the earnings data are from Compustat. White (1980) heteroskedasticity-consistent t-statistics are reported in parentheses. The estimates of b are multiplied by 100.

	China			US		
Quantity	Smallest	Middle	Largest	Smallest	Middle	Largest
Panel A: $k=0$						
b	0.14	0.17	0.24	0.19	0.07	0.05
	(6.34)	(12.42)	(17.28)	(7.51)	(6.99)	(9.90)
R^2	0.003	0.010	0.017	0.005	0.003	0.002
Panel B: $k=3$						
b	0.43	0.58	0.59	0.52	0.20	0.13
	(9.74)	(17.91)	(17.60)	(7.84)	(6.03)	(10.68)
R^2	0.006	0.016	0.021	0.012	0.005	0.003

Table 2
Fama–MacBeth regressions of stock returns on beta, size, and valuation ratios.

The table reports average slope coefficients from month-by-month Fama–MacBeth regressions. Individual stock returns are regressed cross-sectionally on stock characteristics as of the previous month. The columns correspond to different regression specifications, with nonempty rows indicating the included regressors. The regressors include preranking CAPM β_t estimated using the past 12 months of daily returns with a five-lag Dimson (1979) correction; the log of month-end market cap (log$M$); the log of book-to-market (logBM); the log of assets-to-market (logAM); EP^+, which equals the positive values of earnings-to-price, and zero otherwise; $D(EP<0)$, which equals one if earnings are negative, and zero otherwise; CP^+; and $D(CP<0)$ (with the last two similarly defined). The last row reports the average adjusted R-squared for each specification. The sample period is January 2000 through December 2016. The t-statistics based on Newey and West (1987) standard errors with four lags are reported in parentheses.

Quantity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.0149	0.0581	0.0571	0.0659	0.0629	0.0690	0.0564	0.0716	0.0728
	(1.94)	(3.32)	(3.19)	(3.90)	(3.74)	(4.03)	(3.19)	(4.40)	(4.39)
β	−0.0002		−0.0010	−0.0018	−0.0017	0.0002	−0.0010	−0.0002	−0.0004
	(−0.09)		(−0.37)	(−0.71)	(−0.67)	(0.07)	(−0.37)	(−0.06)	(−0.15)
logME		−0.0049	−0.0046	−0.0046	−0.0048	−0.0068	−0.0047	−0.0066	−0.0064
		(−2.91)	(−2.69)	(−2.73)	(−3.00)	(−4.34)	(−2.80)	(−4.49)	(−4.40)
logBM				0.0057				0.0022	0.0035
				(3.21)				(1.31)	(1.76)
logAM					0.0045			0.0014	
					(3.03)			(0.99)	
EP^+						0.9503		0.7825	0.7960
						(4.88)		(4.38)	(5.06)
$D(EP<0)$						0.0006		−0.0005	−0.0001
						(0.31)		(−0.29)	(−0.04)
CP^+							0.0546	0.0181	
							(3.41)	(1.35)	
$D(CP<0)$							0.0019	0.0016	
							(3.11)	(2.37)	
R^2	0.0196	0.0277	0.0441	0.0652	0.0677	0.0615	0.0454	0.0832	0.0776

and with β beta included, the size slopes are −0.0049 and −0.0046 with t-statistics of −2.91 and −2.69. These results confirm a significant size effect in China.

Columns (4) through (7) of Table 2 report results when each valuation ratio is included individually in its own regression. All four valuation ratios exhibit significant explanatory power for returns. When the four valuation ratios are included in the regression simultaneously, as reported in column (8), EP dominates the others. The t-statistic for the coefficient on EP^+ is 4.38, while the t-statistics for logBM, logAM, and CP^+ are just 1.31, 0.99, and 1.35. In fact, the coefficient and t-statistic for EP^+ in column (8) are very similar to those in column (6), in which EP is the only valuation ratio in the regression. The estimated EP effect in column (8) is also economically significant. A one standard-deviation difference in EP^+ implies a difference in expected monthly return of 0.52%.

Because BM likely enters the horse race as a favorite, we also report in column (9) the results when BM and EP are the only valuation ratios included. The results are very similar, with the coefficient and t-statistic for EP^+ quite close to those in column (8) and with the coefficient on $\log BM$ only marginally significant.

Fama and French (1992) exclude financial firms, whereas we include them in Table 2. We do so because we also include financial firms when constructing our factors, as do Fama and French (1993) when constructing their factors. If we instead omit financial firms (including real estate firms) when constructing Table 2, the results (reported in the Appendix) are virtually unchanged.

In sum, we see that *EP* emerges as the most effective valuation ratio, subsuming the other candidates in a head-to-head contest. Therefore, in the next section, we construct our value factor for China using *EP*. The dominance of *EP* over *BM* is further demonstrated in the next section, where we show that our CH-3 model with the *EP*-based value factor prices a *BM*-based value factor, whereas the *BM*-based model, FF-3, cannot price the *EP*-based value factor.

5. A three-factor model in China

In this section, we present our three-factor model, CH-3, with factors for size, value, and the market. Our approach incorporates the features of size and value in China discussed in the previous sections. Section 5.1 provides details of the factor construction. We then compare our approach to one that ignores the China-specific insights. Section 5.2 illustrates the problems with including the smallest 30% of stocks, while Section 5.3 shows that using *EP* to construct the value factor dominates using *BM*.

5.1. Size and value factors

Our model has two distinct features tailored to China. First, we eliminate the smallest 30% of stocks, to avoid their shell-value contamination, and we use the remaining stocks to form factors. Second, we construct our value factor based on *EP*. Otherwise, we follow the procedure used by Fama and French (1993). Specifically, each month we separate the remaining 70% of stocks into two size groups, small (*S*) and big (*B*), split at the median market value of that universe. We also break that universe into three *EP* groups: top 30% (value, *V*), middle 40% (middle, *M*), and bottom 30% (growth, *G*). We then use the intersections of those groups to form value-weighted portfolios for the six resulting size-*EP* combinations: *S*/*V*, *S*/*M*, *S*/*G*, *B*/*V*, *B*/*M*, and *B*/*G*. When forming value-weighted portfolios, here and throughout the study, we weight each stock by the market capitalization of all its outstanding A shares, including nontradable shares. Our size and value factors, denoted as *SMB* (small-minus-big) and *VMG* (value-minus-growth), combine the returns on these six portfolios as follows:

$$SMB = \frac{1}{3}(S/V + S/M + S/G) - \frac{1}{3}(B/V + B/M + B/G),$$

$$VMG = \frac{1}{2}(S/V + B/V) - \frac{1}{2}(S/G + B/G).$$

The market factor, *MKT*, is the return on the value-weighted portfolio of our universe, the top 70% of stocks, in excess of the one-year deposit interest rate.

Table 3 reports summary statistics for the three factors in our 204-month sample period. The monthly standard deviations of *SMB* and *VMG* are 4.52% and 3.75%, each roughly half of the market's standard deviation of 8.09%. The averages of *SMB* and *VMG* are 1.03% and 1.14% per month, with *t*-statistics of 3.25 and 4.34. In contrast, the market factor has a 0.66% mean with a *t*-statistic of just 1.16. Clearly, size and value command substantial premiums in China over our sample period. All three factors are important for pricing, however, in that each factor has a significantly positive alpha with respect to the other two factors. Specifically, those two-factor monthly alphas for *MKT*, *SMB*, and *VMG* are 1.57%, 1.91%, and 1.71%, with *t*-statistics of 2.30, 6.92, and 7.94. Each factor's two-factor alpha exceeds its corresponding simple average essentially due to the negative correlations of *VMG* with both *MKT* and *SMB* (−0.27 and −0.62). In China, smaller stocks tend to be growth stocks, making the negative correlation between size and value stronger than it is in the US. Fama–Macbeth regressions also reveal a substantial negative correlation between China's size and value premiums. For example, the correlation between the coefficients on log*ME* and EP^+ underlying the results reported in column (6) of Table 2 equals 0.42. Note that a positive correlation there is consistent with a negative correlation between the premiums on (small) size and value.

Table 3
Summary statistics for the CH-3 factors.
This table reports the means, standard deviations, *t*-statistics, and pairwise correlations for the three factors in the CH-3 model. The means and standard deviations are expressed in percent per month. The sample period is January 2000 through December 2016 (204 months).

				Correlations		
Factor	Mean	Std. dev.	*t*-stat.	*MKT*	*SMB*	*VMG*
MKT	0.66	8.09	1.16	1.00	0.12	−0.27
SMB	1.03	4.52	3.25	0.12	1.00	−0.62
VMG	1.14	3.75	4.34	−0.27	−0.62	1.00

As Ross (2017) argues, explaining average return is one of two desiderata for a parsimonious factor model. Explaining return variance is the other. Table 4 reports the average *R*-squared values in regressions of individual stock returns on one or more of the CH-3 factors. Panel A includes all listed stocks in China, while Panel B omits the smallest 30%. For comparison with the US, over the same period from January 2000 through December 2016, Panel C reports results when regressing NYSE/Amex/Nasdaq stocks on one or more of the three factors of Fama and French (1993). All regressions are run over rolling three-year windows, and the *R*-squared values are then averaged over time and across stocks.

We see from Table 4 that our size and value factors explain substantial fractions of return variance beyond what the market factor explains. Across all Chinese stocks, for example, the three CH-3 factors jointly explain 53.6% of the typical stock's return variance, versus 38.5% explained by just the market factor. The difference between these values, 15.1%, is actually higher than the corresponding 9.6% difference for the US (27.3% minus 17.7%). Size and value individually explain substantial additional variance, again with each adding more *R*-squared in China than in the US. We also see that the explanatory power of the CH-3 factors, which are constructed using the largest 70% of stocks, improves when averaging just over that universe (Panel B versus Panel A). The improvement is rather modest,

Table 4
Average R-squares for individual stocks in China and the US.

The table compares the average R-squares in regressions of monthly individual stocks' returns on factors in China's and the US stock markets. Regressions are estimated for four models: one with just the excess market return (*MKT*); one with *MKT* plus the size factor; one with *MKT* plus the value factor; and the three-factor model with *MKT* plus the size and value factors. In China's stock market, we use our CH-3 model's market (*MKT*), size (*SMB*), and value (*VMG*) factors, while in the US market, we use FF-3's three factors: market, *SMB* and *BM*-based *HML*. For each stock, we run rolling-window regressions of each stock's monthly returns on factors over the past three years (36 months). We average the R-square across time for each stock and then compute the mean of these averages across all stocks. Panel A reports average R-squares across all individual stocks on China's main boards and the Growth Enterprise Market (GEM), including the bottom 30% of stocks. Panel B reports average R-squares of all but the smallest 30% of stocks. Panel C reports average R-squares of all common stocks from the NYSE, Amex, and Nasdaq for the US. The sample periods for both China and the US are from January 2000 through December 2016 (204 months).

Factors	Avg. R-square
Panel A: All individual stocks in China	
MKT	0.385
MKT, SMB	0.507
MKT, VMG	0.471
MKT, SMB, VMG	0.536
Panel B: All but the smallest 30% of stocks in China	
MKT	0.417
MKT, SMB	0.528
MKT, VMG	0.501
MKT, SMB, VMG	0.562
Panel C: All individual stocks in the US	
MKT	0.177
MKT, SMB	0.231
MKT, HML	0.226
MKT, SMB, HML	0.273

however, indicating that our factors explain substantial variance even for the shell stocks.

A striking China–US difference is that the market factor in China explains more than twice as large a fraction of the typical stock's variance than the market factor explains in the US: 38.5% versus 17.7%. The high average R-squared in China is more typical of earlier decades in US history. For example, Campbell et al. (2001) report average R-squared values exceeding 30% in the US during the 1960s. Exploring potential sources of the higher explanatory power of the market factor in China seems an interesting direction for future research.

Naturally, diversification allows the CH-3 factors to explain larger fractions of return variance for portfolios than for individual stocks. For example, we form value-weighted portfolios within each of 37 industries, using classifications provided by Shenyin–Wanguo Security Co., the leading source of industry classifications in China. On average across industries, the CH-3 factors explain 82% of the variance of an industry's return, versus 72% explained by the market factor. For the anomalies we analyze later, the CH-3 factors typically explain 90% of the return variance for a portfolio formed within a decile of an anomaly ranking variable, versus 85% explained by the market.

We keep negative-*EP* stocks in our sample and categorize them as growth stocks, observing that negative-*EP* stocks comove with growth stocks. Returns on the negative-*EP* stocks load negatively on a value factor constructed using just the positive-*EP* sample, with a slope coefficient of −0.28 and a t-statistic of −3.31. As a robustness check, we exclude negative-*EP* stocks and find all our results hold. On average across months, negative-*EP* stocks account for 15% of the stocks in our universe.

In sum, size and value, as captured by our model's *SMB* and *VMG*, are important factors in China. This conclusion is supported by the factors' average premiums as well as their ability to explain return variances.

5.2. Including shell stocks

If we construct our three factors without eliminating the smallest 30% of stocks, the monthly size premium increases to 1.36%, while the value premium shrinks to 0.87%. As observed earlier, the value of being a potential reverse-merger shell has grown significantly over time, creating a shell premium that accounts for a substantial portion of the smallest stocks' average returns. Consequently, a size premium that includes shell stocks is distorted upward by the shell premium. At the same time, the shell premium distorts the value premium downward. Market values of small firms with persistently poor or negative earnings nevertheless include significant shell value, so those firms' resulting low *EP* ratios classify them as growth firms. Misidentifying shell firms as growth firms then understates the value premium due to the shell premium in returns on those "growth" firms.

High realized returns on shell stocks during our sample period should not necessarily be interpreted as evidence of high expected returns. The high returns could reflect unanticipated increases in rationally priced shells, or they could reflect overpricing of shells in the later years (implying low expected subsequent returns). With rational pricing, an increase in shell value could either raise or lower expected return on the shell firms' stocks, depending on the extent to which shell values contain systematic risks. We do not attempt to explain expected returns on shell stocks. Lee et al. (2017) link expected returns on these stocks to systematic risk related to regulatory shocks.

Including shell stocks also impairs the resulting factor model's explanatory power. When the three factors include the bottom 30% of stocks, they fail to price *SMB* and *VMG* from CH-3, which excludes shells: shell-free *SMB* produces an alpha of −23 basis points (bps) per month (t-statistic: -3.30), and *VMG* produces an alpha of 27 bps (t-statistic: 3.32). These results further confirm that the smallest 30% of stocks are rather different animals. Although they account for just 7% of the market's total capitalization, including them significantly distorts the size and value premiums and impairs the resulting model's explanatory ability. Therefore, excluding shells is important if the goal is to build a model that prices regular stocks.

5.3. Comparing size and value factors

The obvious contender to CH-3 is FF-3, which follows Fama and French (1993) in using *BM* instead of *EP* as the value metric. In this section, we compare CH-3 to FF-3, asking whether one model's factors can explain the other's.

Table 5
Abilities of models CH-3 and FF-3 to explain each other's size and value factors.

Panel A reports a factor's estimated monthly alpha (in percent) with respect to the other model (with White, 1980, heteroskedasticity-consistent *t*-statistics in parentheses). Panel B computes the Gibbons–Ross–Shanken (1989) *F*-test of whether a given model produces zero alphas for the factors of the other model (*p*-value in parentheses). The sample period is January 2000 through December 2016.

	Alphas with respect to:	
Factors	CH-3	FF-3
Panel A: Alpha (t-statistic)		
FFSMB	−0.04	–
	(−0.66)	–
FFHML	0.34	–
	(0.97)	–
SMB	–	0.47
	–	(7.03)
VMG	–	1.39
	–	(7.93)
Panel B: GRS F-statistics (p-value)		
FFSMB, FFHML	0.88	–
	(0.41)	–
SMB, VMG	–	33.90
	–	(2.14×10^{-13})

Using the same stock universe as CH-3, we construct the FF-3 model's size and value factors, combining the six size-*BM* value-weighted portfolios (*S*/*H*, *S*/*M*, *S*/*L*, *B*/*H*, *B*/*M*, *B*/*L*). The size groups are again split at the median market value, and the three *BM* groups are the top 30% (*H*), middle 40% (*M*), and bottom 30% (*L*). The returns on the resulting six portfolios are combined to form the FF-3 size and value factors as follows:

$$FFSMB = \frac{1}{3}(S/H + S/M + S/L) - \frac{1}{3}(B/H + B/M + B/L),$$

$$FFHML = \frac{1}{2}(S/H + B/H) - \frac{1}{2}(S/L + B/L).$$

The market factor is the same as in the CH-3 model.

Our CH-3 model outperforms FF-3 in China by a large margin. Panel A of Table 5 reports the alphas and corresponding *t*-statistics of each model's size and value factors with respect to the other model. CH-3 prices the FF-3 size and value factors quite well. The CH-3 alpha of *FFSMB* is just −4 bps per month, with a *t*-statistic of −0.66, while the alpha of *FFHML* is 34 bps, with a *t*-statistic of 0.97. In contrast, FF-3 prices neither the size nor the value factor of CH-3. FF-3 removes less than half of our model's 103 bps size premium, leaving an *SMB* alpha of 47 bps with a *t*-statistic of 7.03. Most strikingly, the alpha of our value factor, *VMG*, is 139 bps per month (16.68% annually), with a *t*-statistic of 7.93.

Panel B of Table 5 reports Gibbons–Ross–Shanken (GRS) tests of whether both of a model's size and value factors jointly have zero alphas with respect to the other model. The results tell a similar story as above. The test of zero CH-3 alphas for both *FFSMB* and *FFHML* fails to reject that null, with a *p*-value of 0.41. In contrast, the test strongly rejects jointly zero FF-3 alphas for *SMB* and *VMG*, with a *p*-value less than 10^{-12}. The Appendix reports additional details of the regressions underlying the results in Table 5.

The above analysis takes a frequentist approach in comparing the abilities of models to explain each other's factors. Another approach to making this model comparison is Bayesian, proposed by Barillas and Shanken (2018) and also applied by Stambaugh and Yuan (2017). This approach compares factor models in terms of posterior model probabilities across a range of prior distributions. Consistent with the above results, this Bayesian comparison of FF-3 to CH-3 also heavily favors the latter. Details of the analysis are presented in the Appendix.

In the US, two additional factors, profitability and investment, appear in recently proposed models by Hou et al. (2015) and Fama and French (2015). Guo et al. (2017) construct the Fama–French five-factor model in China (FF-5) and find that, when benchmarked against the CAPM, the investment factor is very weak, while the profitability factor is significant. We also find that the investment effect is weak in China, yielding no significant excess return spread or CAPM alpha. A profitability spread has a significant CAPM alpha but does not survive CH-3. Accordingly, in the same tests as above, CH-3 again dominates. The CH-3 alphas for the nonmarket factors in FF-5 produce a GRS *p*-value of 0.88, whereas the FF-5 alphas for the *SMB* and *VMG* factors of CH-3 produce a GRS *p*-value of 0.0003. Details are presented in the Appendix.

6. Anomalies and factors

A factor model is often judged by its ability not only to price another model's factors but also to explain return anomalies. In this section, we explore the latter ability for CH-3 versus FF-3. We start by compiling a set of anomalies in China that are reported in the literature. For each of those anomalies, we compute a long-short return spread in our sample, and we find ten anomalies that produce significant alphas with respect to a CAPM benchmark. Our CH-3 model explains eight of the ten, while FF-3 explains three.

6.1. Anomalies in China

Our survey of the literature reveals 14 anomalies reported for China. The anomalies fall into nine categories: size, value, profitability, volatility, reversal, turnover, investment, accruals, and illiquidity. The literature documenting Chinese anomalies is rather heterogeneous with respect to sample periods, data sources, and choice of benchmarking model (e.g., one factor, three factors, or no factors). Our first step is to reexamine all of the anomalies using our data and sample period. As discussed earlier, our reliance on post-2000 data and our choice of WIND as the data provider offer the most reliable inferences. We also use one model, the CAPM, to classify all the anomalies as being significant or not. Unlike the previous literature, we also evaluate the anomalies within our stock universe that eliminates the smallest 30% so that shell values do not contaminate anomaly effects. For the 14 anomalies we find in the literature, the Appendix reports their CAPM alphas as well as the conclusions of each previous study examining one or more of the anomalies.

For our later analysis of the pricing abilities of the three-factor models, we retain only the anomalies that generate significant CAPM alphas for long-short spreads between portfolios of stocks in the extreme deciles. This nonparametric approach of comparing the extreme deciles, as is common in the anomalies literature, is robust to any monotonic relation but relies on having a sufficiently large sample to achieve power. After imposing our filters, the number of stocks grows from 610 in 2000 to 1872 in 2016, so each portfolio contains at least 60 stocks even early in the sample period. Nevertheless, our 17-year period is somewhat shorter than is typical of US studies, so any of our statements about statistical insignificance of an anomaly must be tempered by this power consideration.

We compute alphas for both unconditional and size-neutral sorts. We conduct the latter sort because correlation between an anomaly variable and size could obscure an anomaly's effect in an unconditional sort, given China's large size premium of 12.36% annually. For each of the 14 anomalies, the two sorting methods are implemented as follows. The unconditional sort forms deciles by sorting on the anomaly variable. (For EP and CP, we sort only the positive values.) We then construct a long-short strategy using deciles one and ten, forming value-weighted portfolios within each decile. The long leg is the higher-performing one, as reported by previous studies and confirmed in our sample. For the size-neutral version, we first form size deciles by sorting on the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly decile portfolios used in our tests. We pool all stocks that fall within a given anomaly decile for any size decile. The returns on those stocks are then value-weighted, using the individual stocks' market capitalizations, to form the portfolio return for that anomaly decile. As with the unconditional sort, the long-short strategy again uses deciles one and ten.

Our procedure reveals significant anomalies in six categories: size, value, profitability, volatility, reversal, and turnover. Almost all of the anomalies in these categories produce significant CAPM-adjusted return spreads from both unconditional and size-neutral sorts. Although the investment, accrual, and illiquidity anomalies produce significant CAPM alphas in the US, they do not in China, for either unconditional or size-neutral sorts. The estimated monthly alphas for investment are small, at 0.22% or less per month, and the accrual alphas are fairly modest as well, at 0.42% or less. The estimated illiquidity alphas, while not quite significant at conventional levels, are nevertheless economically substantial, as high as 0.83% per month. This latter result raises the power issue mentioned earlier. Also unlike the US, there is no momentum effect in China. There is, however, a reversal effect, as past losers significantly outperform past winners.

Reversal effects in China are especially strong. Past performance over any length window tends to reverse in the future. In contrast, past returns in the US correlate in different directions with future returns, depending on the length of the past-return window. That is, past one-month returns correlate negatively with future returns, past two-to-twelve-month returns correlate positively (the well-documented momentum effect), and past three-to-five-year returns correlate negatively. In China, past returns over various windows all predict future reversals. In untabulated results, we find that past returns over windows of one, three, six, and twelve months, as well as five years, all negatively predict future returns, in monotonically weakening magnitudes. For a one-month window of past return, the decile of biggest losers outperforms the biggest winners with a CAPM alpha of 18% annually (t-statistic: 2.96). The alpha drops to 6% and becomes insignificant (t-statistic: 0.90) when sorting by past one-year return.

We choose one-month reversal for the anomaly in the reversal category. One potential source of short-run reversals that does not appear to be related to this anomaly is bid-ask bounce, e.g., Niederhoffer and Osborne (1966). The WIND data beginning in 2012 allow us to average each stock's best bid and ask prices at the day's close of trading. Using the resulting mid-price returns to compute the one-month reversal anomaly gives a result virtually identical to (even slightly higher than) that obtained using closing price returns: 2.21% versus 2.15% for the average long-short monthly return over the 2012–2016 subperiod.

Altogether we find ten significant anomalies. Table 6 reports their average excess returns along with their CAPM alphas and betas. The results for the unconditional sorts appear in Panel A. The monthly CAPM alphas range from 0.53%, for 12-month turnover, to 1.49%, for one-month reversal, and most display significant t-statistics. The average alpha for the ten anomalies is 1.02%, and the average t-statistic is 2.21.

Panel B of Table 6 reports the corresponding results for the size-neutral sorts. Two differences from Panel A emerge. First, size-neutralization substantially increases the alphas of several anomalies. For example, the ROE monthly alpha increases by 0.57%, the EP alpha increases by 0.52%, and the alpha for 12 month turnover increases by 0.21%. Second, for almost all of the long-short spreads, standard deviations decrease and thus t-statistics increase. The decrease in standard deviations confirms that size is an important risk factor. The size-neutral sorting essentially gives the long-short spreads a zero *SMB* loading and thus smaller residual variance in the single-factor CAPM regression. Panel B conveys a similar message as Panel A, just more strongly: all ten anomalies generate significant CAPM-adjusted return spreads. The average monthly CAPM alpha for the size-neutral sorts is 1.17%, and the average t-statistic is 2.91.

6.2. Factor model explanations of anomalies

Table 7 reports CH-3 alphas and factor loadings for the ten anomalies that survive the CAPM, the same anomalies as in Table 7. For the most part, our CH-3 model explains the anomalies well. Panel A of Table 7 reports results for the unconditional sorts. Not surprisingly, CH-3 explains the size anomaly. More noteworthy is that the model explains all the value anomalies (EP, BM, and CP), each of which loads positively on our value factor. The monthly CH-3 alphas of the three value anomalies are 0.64% or less, and the highest t-statistic is just 1.02. These findings echo the earlier Fama–MacBeth regression results, in which *EP* subsumes both *BM* and *CP* in terms of cross-sectional abilities to explain average returns.

Table 6
CAPM alphas and betas for anomalies.
For each of ten anomalies, the table reports the monthly long-short return spread's, average ($\bar{R}$), CAPM alpha (α), and CAPM beta (β). In Panel A, for the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. In Panel B, long/short legs are neutralized with respect to size. That is, we first form size deciles by sorting on the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size groups, again with value weighting. Panel B omits the size anomaly, whose alpha equals zero by construction with size-neutral sorts. Our sample period is January 2000 through December 2016 (204 months). All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Category	Anomaly	$\bar{R}$	α	β	$t(\bar{R})$	$t(\alpha)$	$t(\beta)$
Panel A: Unconditional sorts							
Size	Market cap	1.09	0.97	0.18	1.92	1.81	1.90
Value	EP	1.27	1.37	−0.16	2.58	2.93	−2.15
Value	BM	1.14	1.14	0.01	2.08	2.13	0.06
Value	CP	0.73	0.70	0.04	1.67	1.69	0.55
Profitability	ROE	0.83	0.93	−0.15	1.77	2.11	−2.09
Volatility	1-Month vol.	0.81	1.03	−0.34	1.64	2.31	−5.55
Volatility	MAX	0.57	0.81	−0.36	1.26	2.02	−6.39
Reversal	1-Month return	1.47	1.49	−0.02	2.96	3.06	−0.22
Reversal	12-Month turn.	0.33	0.53	−0.29	0.63	1.09	−3.46
Turnover	1-Mo. abn. turn.	1.14	1.27	−0.19	2.44	2.92	−2.66
Panel B: Size-neutral sorts							
Value	EP	1.80	1.89	−0.14	4.32	4.72	−2.25
Value	BM	1.14	1.10	0.05	2.23	2.22	0.64
Value	CP	0.78	0.76	0.04	2.27	2.25	0.71
Profitability	ROE	1.45	1.50	−0.07	3.90	4.11	−1.30
Volatility	1-Month Vol.	0.66	0.90	−0.37	1.41	2.19	−6.20
Volatility	MAX	0.39	0.60	−0.32	0.93	1.61	−6.14
Reversal	1-Month return	1.67	1.65	0.02	3.65	3.68	0.32
Reversal	12-Month turn.	0.51	0.74	−0.34	1.06	1.74	−4.94
Turnover	1-Mo. abn. turn.	1.29	1.39	−0.15	3.19	3.68	−2.56

Table 7
CH-3 alphas and factor loadings for anomalies.
For each of ten anomalies, the table reports the monthly long-short return spread's CH-3 alpha and factor loadings. For each anomaly, the regression estimated is

$$R_t = \alpha + \beta_{MKT} MKT_t + \beta_{SMB} SMB_t + \beta_{VMG} VMG_t + \epsilon_t,$$

where R_t is the anomaly's long-short return spread in month t, MKT_t is the excess market return, SMB_t is CH-3's size factor, and VMG_t is the EP-based value factor. In Panel A, for the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. In Panel B, long/short legs are neutralized with respect to size. That is, we first form size deciles by sorting on the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size groups, again with value weighting. Panel B omits the size anomaly, whose alpha equals zero by construction with size-neutral sorts. Our sample period is January 2000 through December 2016. All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Category	Anomaly	α	β_{MKT}	β_{SMB}	β_{VMG}	$t(\alpha)$	$t(\beta_{MKT})$	$t(\beta_{SMB})$	$t(\beta_{VMG})$
Panel A: Unconditional sorts									
Size	Market cap	0.21	0.01	1.45	−0.54	1.71	0.77	40.88	−11.70
Value	EP	0.04	0.04	−0.38	1.40	0.16	1.31	−4.73	14.75
Value	BM	0.64	0.06	−0.03	0.43	1.02	0.65	−0.14	1.64
Value	CP	0.20	0.14	−0.28	0.64	0.45	2.10	−1.95	4.00
Profitability	ROE	−0.36	0.03	−0.29	1.28	−0.88	0.70	−2.35	9.43
Volatility	1-Month vol.	0.23	−0.23	−0.12	0.75	0.44	−3.81	−0.67	3.86
Volatility	MAX	0.27	−0.30	−0.05	0.48	0.65	−4.57	−0.30	2.55
Reversal	1-Month return	0.93	−0.06	0.56	0.01	1.70	−0.69	3.15	0.03
Turnover	12-Month turn.	0.42	−0.14	−0.85	0.77	1.30	−3.69	−9.33	7.90
Turnover	1-Mo. abn. turn.	1.28	−0.22	0.18	−0.16	2.86	−2.78	0.93	−0.76
Panel B: Size-neutral sorts									
Value	EP	0.23	0.02	0.05	1.32	0.82	0.47	0.57	11.97
Value	BM	0.61	0.13	−0.15	0.53	0.98	1.60	−0.80	2.19
Value	CP	0.18	0.11	−0.06	0.52	0.54	1.98	−0.50	3.84
Profitability	ROE	−0.37	0.05	0.41	1.20	−1.04	1.23	4.14	9.66
Volatility	1-Month vol.	0.20	−0.28	−0.08	0.64	0.42	−4.96	−0.49	3.34
Volatility	MAX	0.00	−0.26	0.05	0.45	0.00	−4.38	0.30	2.50
Reversal	1-Month return	1.13	0.01	0.41	0.10	2.12	0.11	2.59	0.55
Turnover	12-Month turn.	0.25	−0.22	−0.43	0.74	0.69	−4.94	−3.91	5.74
Turnover	1-Mo. abn. turn.	1.24	−0.18	0.25	−0.08	3.04	−2.79	1.55	−0.43

Perhaps unexpectedly, given the US evidence, CH-3 fully explains the profitability anomaly, return on equity (ROE). In the US, profitability's strong positive relation to average return earns it a position as a factor in the models recently advanced by Hou et al. (2015) and Fama and French (2015). In China, however, profitability is captured by our three-factor model. The ROE spread loads heavily on the value factor (t-statistic: 9.43), and the CH-3 monthly alpha is −0.36%, with a t-statistic of just −0.88.

CH-3 also performs well on the volatility anomalies. It produces insignificant alphas for return spreads based on the past month's daily volatility and the past month's maximum daily return (MAX). The CH-3 monthly alphas for both anomalies are 0.27% or less, with t-statistics no higher than 0.65. We also see that both of the anomalies load significantly on the value factor. That is, low (high) volatility stocks behave similarly to value (growth) stocks.

Recall from the previous section that the estimated CAPM alpha for the illiquidity anomaly, while not quite clearing the statistical-significance hurdle, is as high as 0.83% per month. In contrast, we find that the corresponding CH-3 alpha is just 0.23%, with a t-statistic of 1.14. That is, if we were to add the illiquidity anomaly to our set of ten, given its substantial estimated CAPM alpha, we see that illiquidity would also be included in the list of anomalies that CH-3 explains.

To say for short that our CH-3 model "explains" an anomaly, as in several instances above, must prompt a nod to the power issue mentioned earlier. Of course, more accurate would be to say that the test presented by the anomaly merely fails to reject the model. In general, however, the anomalies for which we can make this statement produce not only insignificant t-statistics but also fairly small estimated CH-3 alphas. Across the eight anomalies that the CH-3 model explains, the average absolute estimated monthly alpha is 0.30% in the unconditional sorts and 0.26% in size-neutral sorts. In contrast, the same anomalies produce average absolute FF-3 alphas of 0.84% and 0.90% in the unconditional and size-neutral sorts.

CH-3 encounters its limitations with anomalies in the reversal and turnover categories. While the reversal spread loads significantly on *SMB*, its monthly alpha is nevertheless 0.93% (t-statistic:1.70). In the turnover category, CH-3 accommodates 12 month turnover well but has no success with abnormal 1 month turnover. The latter anomaly's return spread has small and insignificant loadings on *SMB* and *VMG*, and its CH-3 monthly alpha is 1.28%, nearly identical to its CAPM alpha (t-statistic: 2.86).

The size-neutral sorts, reported in Panel B of Table 7, deliver the same conclusions as the unconditional sorts in Panel A. CH-3 again explains all anomalies in the value, profitability, and volatility categories. The monthly alphas for those anomalies have absolute values of 0.61% or less, with t-statistics less than 0.98 in magnitude. For the reversal and turnover categories, CH-3 displays the same limitations as in Panel A. The CH-3 monthly alpha for reversal is 1.13%, with a t-statistic of 2.12. Abnormal turnover has an alpha of 1.24%, with a t-statistic of 3.04.

In the same format as Table 7, Table 8 reports the corresponding results for the FF-3 model. These results clearly demonstrate that FF-3 performs substantially worse than CH-3, leaving significant anomalies in five of the six categories—all categories except size. Consider the results in Panel A, for example. Similar to FF-3's inability to price our EP-based value factor, FF-3 fails miserably with the EP anomaly, leaving a monthly alpha of 1.54% (t-statistic: 5.57). Moreover, as in the US, FF-3 cannot accommodate profitability. The ROE anomaly leaves a monthly alpha of 1.75% (t-statistic: 5.67). Finally, for all anomalies in the volatility, reversal, and turnover categories, FF-3 leaves both economically and statistically significant alphas.

Table 9 compares the abilities of models to explain anomalies by reporting the average absolute alphas for the anomaly long-short spreads, the corresponding average absolute t-statistics, and GRS tests of whether a given model produces jointly zero alphas across anomalies. The competing models include unconditional means (i.e., zero factors), the single-factor CAPM, and both of the three-factor models, CH-3 and FF-3. As in Tables 7 and 8, Panel A reports results for the unconditional sorts, and Panel B reports the size-neutral sorts. First, in both panels, observe that CH-3 produces much smaller absolute alphas than do the other models: 0.45% for CH-3 versus at least 0.9% for the other models. In Panel A, for the unconditional sorts, the GRS p-value of 0.15 for CH-3 fails to reject the joint hypothesis that all ten anomalies produce zero CH-3 alphas. In contrast, the corresponding p-values for the other models are all less than 10^{-4}. For the size-neutral sorts (Panel B), a similar disparity occurs for a test of jointly zero alphas on nine anomalies (size is omitted). The CH-3 p-value is 0.05 versus p-values less than 10^{-4} for the other models. Because size, *EP*, and *BM* are used to construct factors, we also eliminate those three anomalies and conduct the GRS test using the remaining seven. As shown in the last two rows of each panel, the results barely change—CH-3 again dominates.

7. A four-factor model in China

Notwithstanding the impressive performance of CH-3, the model does leave significant alphas for reversal and turnover anomalies, as noted earlier. Of course, we see above that these anomalies are not troublesome enough to cause the larger set that includes them to reject CH-3 when accounting for the multiple comparisons inherent in the GRS test. At the same time, however, the latter test confronts the same power issue discussed earlier. Moreover, the reversal and turnover anomalies both produce alpha estimates that are not only statistically significant but also economically large, over 1% per month in the size-neutral sorts reported in Panel B of Table 7. We therefore explore the addition of a fourth factor based on turnover. In Section 7.1, we discuss this turnover factor's sentiment-based motivation, describe the factor's construction, and explain how we also modify the size factor when building the four-factor model, CH-4. Section 7.2 then documents CH-4's ability to explain all of China's reported anomalies.

7.1. A turnover factor

A potential source of high trading intensity in a stock is heightened optimism toward the stock by

Table 8
FF-3 alphas and factor loadings for anomalies.
For each of ten anomalies, the table reports the monthly long-short return spread's CH-3 alpha and factor loadings. For each anomaly, the regression estimated is

$$R_t = \alpha + \beta_{MKT} MKT_t + \beta_{SMB} FFSMB_t + \beta_{HML} FFHML_t + \epsilon_t,$$

where R_t is the anomaly's long-short return spread in month t, MKT_t is the excess market return, SMB_t is FF-3's size factor, and $FFHML_t$ is the BM-based value factor. In Panel A, for the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. In Panel B, long/short legs are neutralized with respect to size. That is, we first form size deciles by sorting on the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size groups, again with value weighting. Panel B omits the size anomaly, whose alpha equals zero by construction with size-neutral sorts. Our sample period is January 2000 through December 2016. All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Category	Anomaly	α	β_{MKT}	β_{SMB}	β_{HML}	$t(\alpha)$	$t(\beta_{MKT})$	$t(\beta_{SMB})$	$t(\beta_{HML})$
Panel A: Unconditional sorts									
Size	Market cap	0.16	0.04	1.55	−0.11	1.36	1.86	34.29	−2.43
Value	EP	1.54	−0.07	−0.98	0.48	5.57	−1.63	−13.29	6.54
Value	BM	−0.28	0.00	0.12	1.60	−1.25	0.09	1.48	30.61
Value	CP	0.63	0.08	−0.49	0.43	1.40	1.31	−3.84	2.31
Profitability	ROE	1.75	−0.06	−1.01	−0.28	5.67	−1.36	−12.74	−3.08
Volatility	1-Month vol.	0.83	−0.30	−0.40	0.52	2.11	−5.41	−3.07	4.18
Volatility	MAX	0.74	−0.33	−0.28	0.28	1.85	−5.70	−1.79	1.96
Reversal	1-Month return	0.94	−0.06	0.53	0.28	1.97	−0.83	3.50	1.58
Turnover	12-Month turn.	0.83	−0.19	−1.08	0.38	2.96	−4.97	−13.10	3.72
Turnover	1-Mo. abn. turn.	1.34	−0.21	0.16	−0.20	2.86	−2.76	0.95	−0.99
Panel B: Size-neutral sorts									
Value	EP	1.76	−0.09	−0.53	0.52	5.49	−1.79	−6.52	6.17
Value	BM	−0.01	0.07	−0.10	1.39	−0.04	1.75	−1.36	20.12
Value	CP	0.52	0.06	−0.23	0.44	1.73	1.33	−2.33	4.05
Profitability	ROE	2.01	−0.04	−0.36	−0.35	5.72	−0.71	−3.75	−3.38
Volatility	1-Month vol.	0.76	−0.33	−0.33	0.40	2.06	−6.04	−2.64	3.47
Volatility	MAX	0.43	−0.30	−0.16	0.31	1.14	−5.67	−1.10	2.60
Reversal	1-Month return	1.21	−0.01	0.35	0.29	2.55	−0.09	2.63	1.78
Turnover	12-Month turn.	0.80	−0.28	−0.68	0.39	2.46	−5.89	−6.37	3.74
Turnover	1-Mo. abn. turn.	1.37	−0.17	0.21	−0.12	3.26	−2.82	1.43	−0.68

Table 9
Comparing the abilities of models to explain anomalies.
The table reports measures summarizing the degree to which anomalies produce alphas under three different factor models: CAPM, FF-3, and CH-3. Also reported are measures for "unadjusted" return spreads (i.e., for a model with no factors). For each model, the table reports the average absolute monthly alpha (in percent), average absolute t-statistic, the Gibbons–Ross–Shanken (1989) "GRS" F-statistic with associated p-value, and the number of anomalies for which the model produces the smallest absolute alpha among the four models. In Panel A, for the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. In Panel B, long/short legs are neutralized with respect to size. That is, we first form size deciles by sorting on the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size groups, again with value weighting. Two versions of the GRS test are reported. In Panel A, GRS_{10} uses all ten anomalies, while GRS_7 excludes the anomalies for size, BM, and EP, which are variables used to construct factors. Panel B omits the size anomaly, whose alpha equals zero by construction with size-neutral sorts. All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980). The sample period is from January 2000 through December 2016 (204 months).

Measure	Unadjusted	CAPM	FF-3	CH-3
Panel A: Unconditional sorts				
Average $\|\alpha\|$	0.94	1.02	0.90	0.45
Average $\|t\|$	1.89	2.21	2.70	1.12
GRS_{10}	7.30	7.31	6.00	1.49
p_{10}	< 0.0001	< 0.0001	< 0.0001	0.15
GRS_7	4.40	4.45	6.86	1.74
p_7	0.0002	0.0001	0.0001	0.10
Panel B: Size-neutral sorts				
Average $\|\alpha\|$	1.08	1.17	0.99	0.47
Average $\|t\|$	2.55	2.91	2.72	1.07
GRS_9	8.24	8.08	7.97	1.97
p_9	< 0.0001	< 0.0001	< 0.0001	0.05
GRS_7	8.15	8.10	9.11	2.33
p_7	< 0.0001	< 0.0001	< 0.0001	0.03

Table 10
Anomaly alphas under a four-factor model.
For each of ten anomalies, the table reports the monthly long-short return spread's CH-4 alpha and factor loadings. For each anomaly, the regression estimated is

$$R_t = \alpha + \beta_{MKT} MKT_t + \beta_{SMB} SMB_t + \beta_{VMG} VMG_t + \beta_{PMO} PMO_t + \epsilon_t,$$

in which R_t is the anomaly's long-short return spread in month t, MKT_t is the excess market return, SMB_t is CH-3's size factor, VMG_t is the *EP*-based value factor, and PMO_t (pessimistic minus optimistic) is the sentiment factor based on abnormal turnover. In Panel A, for the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. In Panel B, long/short legs are neutralized with respect to size. That is, we first form size deciles by sorting on the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size groups, again with value weighting. Panel B omits the size anomaly, whose alpha equals zero by construction with size-neutral sorts. Our sample period is January 2000 through December 2016. All *t*-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Category	Anomaly	α	β_{MKT}	β_{SMB}	β_{VMG}	β_{PMO}	$t(\alpha)$	$t(\beta_{MKT})$	$t(\beta_{SMB})$	$t(\beta_{VMG})$	$t(\beta_{PMO})$
Panel A: Unconditional sorts											
Size	Market cap	0.23	0.05	1.50	−0.42	0.01	1.41	3.05	39.38	−6.88	0.17
Value	EP	0.02	0.00	−0.46	1.34	0.04	0.08	0.07	−5.38	13.15	0.33
Value	BM	0.75	0.03	−0.02	0.43	−0.13	1.04	0.32	−0.09	1.55	−0.45
Value	CP	0.31	0.09	−0.26	0.65	−0.20	0.57	1.43	−1.81	3.60	−1.10
Profitability	ROE	−0.29	−0.03	−0.36	1.28	−0.10	−0.68	−0.70	−3.21	8.74	−0.96
Volatility	1-Month vol.	−0.27	−0.16	−0.27	0.59	0.72	−0.51	−2.71	−1.85	3.27	5.13
Volatility	MAX	−0.59	−0.18	−0.13	0.44	0.88	−1.64	−3.07	−0.90	2.91	7.91
Reversal	1-Month return	0.49	0.02	0.54	0.04	0.46	0.87	0.29	3.19	0.18	2.48
Turnover	12-Month turn.	0.04	−0.11	−0.94	0.64	0.43	0.11	−3.36	−10.60	5.09	3.69
Turnover	1-Mo. abn. turn.	−0.00	−0.01	0.07	−0.27	1.44	−0.01	−0.32	0.68	−2.43	16.47
Panel B: Size-neutral sorts											
Value	EP	0.43	−0.04	−0.03	1.28	−0.12	1.42	−0.79	−0.35	11.74	−1.09
Value	BM	0.57	0.12	−0.18	0.46	0.09	0.82	1.58	−1.04	1.83	0.39
Value	CP	0.19	0.08	−0.03	0.57	−0.12	0.49	1.56	−0.26	3.67	−0.92
Profitability	ROE	−0.30	0.02	0.35	1.23	−0.05	−0.76	0.51	3.67	8.99	−0.40
Volatility	1-Month vol.	−0.27	−0.21	−0.20	0.51	0.63	−0.59	−3.95	−1.37	2.87	4.90
Volatility	MAX	−0.77	−0.17	0.00	0.45	0.74	−2.05	−2.86	0.01	2.95	5.81
Reversal	1-Month return	0.71	0.07	0.40	0.12	0.42	1.28	1.14	2.62	0.68	2.60
Turnover	12-Month turn.	−0.07	−0.19	−0.53	0.61	0.44	−0.19	−4.16	−4.38	3.77	3.39
Turnover	1-Mo. abn. turn.	0.17	−0.00	0.17	−0.17	1.21	0.67	−0.05	1.83	−1.90	15.23

sentiment-driven investors. This argument is advanced by Baker and Stein (2004), for example, and Lee (2013) uses turnover empirically as a sentiment measure at the individual stock level. High sentiment toward a stock can affect its price, driving it higher than justified by fundamentals and thereby lowering its expected future return. Two assumptions underly such a scenario. One is a substantial presence in the market of irrational, sentiment-driven traders. The other is the presence of short-sale impediments.

China's stock market is especially suited to both assumptions. First, individual retail investors are the most likely sentiment traders, and individual investors are the major participants in China's stock market. As of year-end 2015, over 101 million individuals had trading accounts, and individuals held 88% of all free-floating shares (Jiang et al., 2016). Second, shorting is extremely costly in China.[4]

Shorting constraints not only impede the correction of overpricing. They also sign the likely relation between sentiment and turnover. As Baker and Stein (2004) argue, when pessimism about a stock prevails among sentiment-driven investors, those who do not already own the stock simply do not participate in the market, as short-sale constraints prevent them from acting on their pessimistic views. In contrast, when optimism prevails, sentiment-driven investors can participate broadly in buying the stock. Thus, shorting constraints make high turnover (greater liquidity) more likely to accompany strong optimism as opposed to strong pessimism.

Given this sentiment-based motivation, to construct our fourth factor we use abnormal turnover, which is the past month's share turnover divided by the past year's turnover. We construct this turnover factor in precisely the same manner as our value factor, again neutralizing with respect to size. That is, abnormal turnover simply replaces *EP*, except the factor goes long the low-turnover stocks, about which investors are relatively pessimistic, and goes short the high-turnover stocks, for which greater optimism prevails. We denote the resulting factor *PMO* (pessimistic minus optimistic). We also construct a new *SMB*, taking a simple average of the *EP*-neutralized version of *SMB* from CH-3 and the corresponding turnover-neutralized version. The latter procedure for modifying *SMB* when adding additional factors essentially follows Fama and French (2015). The new size and turnover factors have annualized averages of 11% and 12%. The market and value factors in CH-4 are the same as in CH-3.

[4] Costs of short selling in China are discussed, for example, in the CSRC publication, Chinese Capital Market Development Report (translated from Mandarin).

7.2. Explaining all anomalies with four factors

For model CH-4, Table 10 reports results of the same analyses conducted for models CH-3 and FF-3 and reported in Tables 7 and 8. Adding the fourth factor produces insignificant alphas not just for the abnormal turnover anomaly but also for reversal. In Panel A, for the unconditional sorts, the CH-4 monthly alphas for those anomalies are 0.00% and 0.49%, with t-statistics of −0.01 and 0.87. The size-neutral sorts in Panel B produce similar results. Adding the turnover factor essentially halves the reversal anomaly's unconditional alpha relative to its CH-3 value in Table 7, even though the rank correlation across stocks between the sorting variables for the turnover and reversal anomalies is just 0.3, on average.

CH-4 accommodates the above two anomalies, thus now explaining all ten, while also lowering the average magnitude of all the alphas. For the unconditional sorts, the average absolute alpha drops to 0.30%, versus 0.45% for CH-3, and the average absolute t-statistic drops to 0.69, versus 1.12 for CH-3. The GRS test of jointly zero alphas for all ten anomalies produces a p-value of 0.41, versus 0.15 for CH-3, thereby moving even farther from rejecting the null. Similar improvements occur for the size-neutral sorts.

8. Conclusion

Size and value are important factors in the Chinese stock market, with both having average premiums exceeding 12% per year. Capturing these factors well, however, requires that one not simply replicate the Fama and French (1993) procedure developed for the US.

Unlike small listed stocks in the US, China's tight IPO constraints cause returns on the smallest stocks in China to be significantly contaminated by fluctuations in the value of becoming corporate shells in reverse mergers. To avoid this contamination, before constructing factors we eliminate the smallest 30% of stocks, which account for just 7% of the market's total capitalization. Eliminating these stocks yields factors that perform substantially better than using all listed stocks to construct factors, whereas the Fama and French (1993) procedure essentially does the latter in the US.

Value effects in China are captured much better by *EP* than by *BM*, used in the US by Fama and French (1993). The superiority of *EP* in China is demonstrated at least two ways. First, in an investigation paralleling Fama and French (1992), cross-sectional regressions reveal that *EP* subsumes other valuation ratios, including *BM*, in explaining average stock returns. Second, our three-factor model, CH-3, with its *EP*-based value factor, dominates the alternative FF-3 model, with its *BM*-based value factor. In a head-to-head model comparison, CH-3 prices both the size and value factors in FF-3, whereas FF-3 prices neither of the size and value factors in CH-3. In particular, FF-3 leaves a 17% annual alpha for our value factor.

We also survey the literature that documents return anomalies in China, and we find ten anomalies with significant CAPM alphas in our sample. Our CH-3 model explains eight of the anomalies, including not just all value anomalies but also profitability and volatility anomalies not explained in the US by the three-factor Fama–French model. In contrast, the only two anomalies in China that FF-3 explains are size and BM. The two anomalies for which CH-3 fails, return reversal and abnormal turnover, are both explained by a four-factor model that adds a sentiment-motivated turnover factor.

Appendix A

Section A.1 provides details of the data sources and the filters we apply. Section A.2 details the anomalies and their construction. Section A.3 provides further details about China's IPO review process. Section A.4 explains the reverse-merger data and the shell-value estimation, while section A.5 examines the hypothesis that the smallest 30% of stocks covary more with shell-value proxies. Section A.6 reports the results of recomputing the regressions reported in Table 2 with financial firms excluded. Section A.7 presents additional comparisons of model CH-3 to models FF-3 and FF-5.

A.1. Data sources and filters

Our stock trading data and firm financial data come from WIND. The Internet Appendix lists the WIND data items we use, including both the English and Chinese item codes. Our sample includes all A-share stocks from the main boards of the Shanghai and Shenzhen exchanges as well as the board of the GEM, essentially the Chinese counterpart of Nasdaq. In China, stock tickers for listed firms are nonreusable unique identifiers. The ticker contains six digits, of which the first two indicate the exchange and the security type. We include stocks whose first two digits are 60, 30, and 00. Our series for the riskfree rate, the one-year deposit rate, is obtained from the China Stock Market and Accounting Research (CSMAR) database on Wharton Research Data Services (WRDS). Our sample period is January 2000 through December 2016.

We also impose several filters. First, we exclude stocks that have become public within the past six months. Second, we exclude stocks having (i) less than 120 days of trading records during the past 12 months or (ii) less than 15 days of trading records during the most recent month. The above filters are intended to prevent our results from being influenced by returns that follow long trading suspensions. Third, for the reason explained in Section 3, we eliminate the bottom 30% of stocks ranked by market capitalization at the end of the previous month. Market capitalization is calculated as the closing price times total shares outstanding, including nontradable shares.

When we use financial statement information to sort stocks, in constructing either factors or anomaly portfolios, the sort at the end of a given month uses the information in a firm's financial report having the most recent public release date prior to that month's end. (The WIND data include release dates.) Firms' financial statement data are from quarterly reports beginning January 1, 2002, when public firms were required to report quarterly. Prior to that date, our financial statement data are from semi-annual reports.

A.2. Firm characteristics and anomaly portfolios

We survey the literature documenting anomalies in China, and we compile here, to our knowledge, an exhaustive list of stock characteristics identified as cross-sectional predictors of future returns. The list comprises nine categories: size, value, profitability, volatility, investment, accruals, illiquidity, reversal, and turnover. Within each category, one or more firm-level characteristics are identified as return predictors. The anomalies, by category, are as follows:

1. Size. The stock's market capitalization is used in this category. It is computed as the previous month's closing price times total A shares outstanding, including nontradable shares.[5]
2. Value. Three variables are used.
   - Earnings-price ratio (*EP*). Earnings equals the most recently reported net profit excluding nonrecurrent gains/losses. A stock's *EP* is the ratio of earnings to the product of last month-end's close price and total shares.
   - Book-to-market ratio (*BM*). Book equity equals total shareholder equity minus the book value of preferred stocks. A stock's *BM* is the ratio of book equity to the product of last month-end's close price and total shares.
   - Cash-flow-to-price (*CP*). Cash flow equals the net change in cash or cash equivalents between the two most recent cash flow statements.[6] A stock's *CP* is the ratio of cash flow to the product of last month-end's close price and total shares.
3. Profitability. Firm-level *ROE* at the quarterly frequency is used. The value of *ROE* equals the ratio of a firm's earnings to book equity, with earnings and book equity defined above.
4. Volatility. Two variables are used.
   - One-month volatility. A firm's one-month volatility is calculated as the standard deviation of daily returns over the past 20 trading days.
   - *MAX*. *MAX* equals the highest daily return over the past 20 trading days.
5. Investment. As in Fama and French (2015), a firm's investment is measured by its annual asset growth rate. Specifically, a firm's asset growth equals total assets in the most recent annual report divided by total assets in the previous annual report.
6. Accruals. Two variables are used.
   - Accruals. We construct firm-level accruals following Sloan (1996). Specifically, a firm's accruals in year t can be expressed as

     $$Accrual = (\Delta CA - \Delta Cash) - (\Delta CL - \Delta STD - \Delta TP) - Dep,$$

     in which ΔCA equals the most recent year-to-year change in current assets, $\Delta Cash$ equals the change in cash or cash equivalents, ΔCL equals the change in current liabilities, ΔSTD equals the change in debt included in current liabilities, ΔTP equals the change in income taxes payable, and *Dep* equals the most recent year's depreciation and amortization expenses.
   - Net-operating-assets (*NOA*). We construct firm-level *NOA*, following Hirshleifer et al. (2004). Specifically, *NOA* is calculated as

     $$NOA_t = (Operating\ asset_t - Operating\ liability_t)/Total\ asset_{t-1},$$

     in which $Operating\,asset_t$ equals total assets minus cash and short-term investment, and $Operating\,liability_t$ equals total assets minus short-term debt, long-term debt, minority interest, book preferred stock, and book common equity.
7. Illiquidity. We compute a stock's average daily illiquidity over the past 20 trading days. Following Amihud (2002), a stock's illiquidity measure for day t is calculated as

   $$Illiq_t = |ret_t|/volume_t,$$

   in which $|ret_t|$ is the stock's absolute return on day t, and $volume_t$ is the stock's dollar trading volume on day t.
8. Turnover. Two variables are used:
   - Twelve-month turnover. We measure 12-month turnover as the average daily share turnover over the past 250 days. A firm's daily turnover is calculated as its share trading volume divided by its total shares outstanding.
   - One-month abnormal turnover. A firm's abnormal turnover is calculated as the ratio of its average daily turnover over the past 20 days to its average daily turnover over the past 250 days.
9. Reversal. The sorting measure used is the stock's one-month return, computed as the cumulative return over the past 20 trading days.

For every anomaly except reversal, we sort the stock universe each month using the most recent month-end measures and then hold the resulting portfolios for one month. Because one-month return reversal is a short-term anomaly, we sort the stock universe each day based on the most recently available 20-day cumulative return. Using this sort, we rebalance a one-fifth "slice" of the total portfolio that is then held for five trading days. Each day we average the returns across the five slices. Those resulting daily returns are then compounded across days to compute the reversal anomaly's monthly return. For all anomalies, value-weighted portfolios of stocks within the top and bottom deciles are formed using the most recent month-end market capitalizations as weights.

A.3. The IPO review process in China

The process of IPO review by the CSRC involves seven steps:

[5] In China's stock market, a firm can issue three types of share classes: A, B, and H shares. Domestic investors can trade only A shares, while foreign investors can trade only B shares. H shares are issued by domestic companies but traded on Hong Kong exchanges. We measure size based on total A shares but compute valuation ratios by scaling with total shares, including B and H shares (treating earnings and book-values of a firm as applying to all shareholders, including foreign and Hong Kong investors).

[6] Prior to January 2002, cash flow equals half the net change between the two most recent semi-annual statements.

Table A1
Chinese anomalies in the literature.
The table provides a compilation of the anomalies reported by one or more studies as being significant in China. The studies analyzing each anomaly are listed, including those reporting an anomaly to be insignificant (identified as such in the table). For each anomaly variable, we also report CAPM alphas for both unconditional and conditional sorts. For the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. For the conditional sorts, long/short legs are neutralized with respect to size. That is, we first form size deciles by sorting on the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size groups, again with value weighting. Our sample period is January 2000 through December 2016 (204 months). All t-statistics (in parentheses) are based on the heteroskedasticity-consistent standard errors of White (1980).

			CAPM alpha (monthly %)	
Category	Anomaly	References	Unconditional	Size-neutral
Size	Market cap	Wang and Xu (2004), Eun and Huang (2007), Cheung et al. (2015), Cakici et al. (2017), Hsu et al. (2018), Carpenter et al. (2018), and Hu et al. (2019). Reported insignificant: Chen et al. (2010) and Cheung et al. (2015).	0.97 (1.81)	–
Value	EP	Cakici et al. (2017) and Hsu et al. (2018). Reported insignificant: Chen et al. (2010) and Hu et al. (2019).	1.37 (2.93)	1.89 (4.72)
Value	BM	Wang and Xu (2004), Eun and Huang (2007), Chen et al. (2010), Cheung et al. (2015), Cakici et al. (2017), Hsu et al. (2018), and Carpenter et al. (2018). Reported insignificant: Hu et al. (2019).	1.14 (2.13)	1.10 (2.22)
Value	CP	Cakici et al. (2017). Reported insignificant: Wang and Di Iorio (2007) and Chen et al. (2010).	0.70 (1.69)	0.76 (2.25)
Profitability	ROE	Guo et al. (2017). Reported insignificant: Li et al. (2007).	0.93 (2.11)	1.50 (4.11)
Volatility	1-Mo. vol.	Cheung et al. (2015), Cakici et al. (2017), and Hsu et al. (2018). Reported insignificant: Chen et al. (2010).	1.03 (2.31)	0.90 (2.19)
Volatility	MAX	Carpenter et al. (2018).	0.81 (2.02)	0.60 (1.61)
Reversal	1-Month return	Cakici et al. (2017), Hsu et al. (2018), and Carpenter et al. (2018). Reported insignificant: Cheung et al. (2015).	1.49 (3.06)	1.65 (3.68)
Turnover	12-Month turn.	Zhang and Liu (2006) and Eun and Huang (2007). Reported insignificant: Chen et al. (2010).	0.53 (1.09)	0.74 (1.74)
Turnover	1-Mo. abn. turn.	Li and Wu (2003) and Zhang and Liu (2006).	1.27 (2.92)	1.39 (3.68)
Investment	Asset growth	Chen et al. (2010). Reported insignificant: Hsu et al. (2018), Guo et al. (2017), and Lin (2017).	0.22 (0.72)	−0.05 (−0.20)
Accruals	Accruals	Li et al. (2011) and Hsu et al. (2018). Reported insignificant: Chen et al. (2010).	0.08 (0.39)	−0.15 (−0.70)
Accruals	NOA	Chen et al. (2010) and Hsu et al. (2018).	0.38 (1.03)	0.42 (1.22)
Illiquidity	Amihud illiq.	Carpenter et al. (2018) and Chen et al. (2010).	0.83 (1.62)	0.63 (1.55)

1. Confirmation of application receipt. The reception department in the bureau organizes all of the application packages, confirms with each applicant firm the receipt of all materials, and makes the offering proposals public.
2. Application material precheck (feedback provided). In this step, the offering administration department assigns a team to precheck all application materials and prepare a written report indicating whether more material/information is needed and suggesting the potential concerns or issues with the offering for further reviews. After receiving the written report, the applicant firm can work with investment bankers to revise the application package.
3. Communication meeting. The applicant's IPO team meets with the offering administration department, but no material issues are discussed.
4. Update disclosed offering proposal. After revising the application package and upon receiving the offering administration team's approval, the applicant's team can revise the initially disclosed offering proposal.
5. Initial review meeting. An offering committee team and the offering administrative department attend the meeting to review thoroughly whether the applicant firm's current condition and growth prospects satisfy IPO criteria.
6. Final offering review meeting. A different offering committee team votes on whether or not to approve the firm's IPO application based on the results of the initial review meeting discussion.
7. Offering. The applicant firm may still need to revise its application package based on the decision in the final offering review meeting. Following that, the firm can prepare the offering.

A.4. Reverse-merger data and shell-value estimation

Our reverse merger data are from WIND and cover the 2007–2016 period. In July 2007, the CSRC issued a ruling that formed a special committee in its Public Offering Department to strengthen the review process for M&A applications. Subsequently, the CSRC took a more active role in the M&A review process, increasing reporting transparency and quality. In the same ruling, the CSRC identified several characteristics of M&A cases that classify them as reverse mergers, to be handled with more scrutiny. The improved reporting transparency and CSRC's specification of reverse mergers made it possible to trace reverse-merger cases.

To construct the daily series of estimated shell values, we first estimate p and G in Eq. (1) as follows:

1. Probability of becoming a shell (p). Each day, we calculate the fraction of stocks, among those in the bottom 30%, involved in reverse merger deals during the past 730 calendar days. We use that fraction as the empirical probability estimate of becoming a shell within in a two-year window, converting the value to a one-year probability (essentially by halving it).
2. Value appreciation upon becoming a shell (G). For each reverse merger case among the bottom 30%, the value appreciation equals the change in market value over a time window beginning 60 days before the board's announcement of a deal proposal and ending 60 days after the deal's CSRC approval. On each day, the estimated value of G is the average value appreciation in all reverse-merger deals whose windows end during the past 730 calendar days.

A.5. Return variation related to shell value

Here we investigate the sensitivity of the returns on the smallest stocks to two variables that proxy for fluctuations in shell values. The first variable reflects variation in the reverse-merger premium. For each reverse-merger event, we define an event window beginning 60 days before the shell firm's board meeting announcing the proposal and ending 60 days after CSRC approval. Each day we compute the average return of all stocks that are within their event window, and then we compound those daily average returns to form a reverse-merger return for month t, RM_t. We associate a relatively high value of RM_t with an increase in shell value during that month. Our hypothesis is that returns on stocks in the bottom 30% of the size distribution are more sensitive to RM_t than are other stocks. For each of three size groups, formed by dividing the universe of all listed stocks at the 30th and 70th percentiles, we estimate the regression,

$$R_t = a + bRM_t + \gamma F_t + \epsilon_t, \qquad (3)$$

in which R_t is the size group's value-weighted monthly return, and the vector F_t contains the other two size groups' returns. For example, when the regression has the return on the smallest stocks on the left-hand side, F_t includes the returns on the middle- and large-cap groups. We expect b to be positive for the smallest stocks but not for the other two size groups. Panel A of Table A2 reports the regression results. We see that indeed the estimate of b is significantly positive for the smallest stocks but is negative for the other two size groups.

The second proxy for capturing fluctuations in shell value is the log of the total number of IPOs in month t, with the rationale that a greater frequency of IPOs could be interpreted by the market as a relaxing of IPO constraints and thus a reduction in reverse-merger shell value. Under our shell-value hypothesis, we expect stock returns to covary negatively with IPO numbers for the smallest stocks but not for the other two size groups. The regression is the same as that in Eq. (3) except that RM_t is replaced by the log of the IPO number, $\log(NIPO_t)$. Panel B of Table A2 reports the results. Consistent with our hypothesis, the estimate of b is negative for the smallest stocks, although the t-statistic of -1.53 falls somewhat short of significance. In contrast, the estimate of b is insignificant and positive (t-statistic: 1.18) for the middle group, while it has a t-statistic of just -0.50 for the large-cap group.

Table A2
Sensitivities of size group returns to proxies for fluctuations in shell values.
For each of three size groups, formed by dividing the universe of all listed stocks at the 30th and 70th percentiles, the table reports the estimate of b in the regression,

$$R_t = a + bShell_t + \gamma F_t + \epsilon_t,$$

in which R_t is the value-weighted return for the size group, $Shell_t$ is a proxy for fluctuations in the value of potentially becoming a reverse-merger shell, and the vector F_t contains the returns on the other two size groups. In Panel A, $Shell_t$ is the average return of shell stocks experiencing reverse mergers, RM_t. In Panel B, $Shell_t$ is the log of the number of IPOs, $\log(NIPO_t)$. The sample period is January 2007 through December 2016. All t-statistics (in parentheses) are based on the heteroskedasticity-consistent standard errors of White (1980).

Smallest stocks	Middle	Largest stocks
Panel A. $Shell_t = RM_t$		
0.963	−1.267	−0.330
(2.19)	(−3.02)	(−4.65)
Panel B. $Shell_t = \log(NIPO_t)$		
−0.023	0.013	−0.001
(−1.53)	(1.18)	(−0.50)

The results from both sets of regressions are consistent with our hypothesis that the returns on stocks in the bottom 30% reflect, in substantial part, variation in the value of being a potential reverse-merger shell. In other words, the values of these "small" stocks reflect more than the businesses of the underlying small firms.

A.6. Excluding financial firms from Table 2

Table A3 reports the results of recomputing the same regressions reported in Table 2 but with financial firms excluded. On average, the financial firms include 79 real estate firms and 20 other firms in banking and financial services. The results in Table A3 are very similar to those in Table 2.

A.7. Additional model-comparison results

Table A4 reports details of the regressions underlying the results reported in Table 5. Each model's size factor loads strongly on the other model's size factor, producing t-statistics of about 39 in each case. The models' value factors also load significantly on each other, but less strongly, with t-statistics of 2.83 and 3.49. The size factor of FF-3 loads significantly negatively on the value factor of CH-3, with a t-statistic of -8.3, but CH-3's size factor does not load significantly on FF-3's value factor. Similarly, CH-3's value factor loads significantly negatively on FF-3's size factor (t-statistic: -11.15) but not vice versa.

Model CH-3 dominates FF-3 not only in the frequentist comparisons reported in Table 5 but also in a

Table A3
Fama–MacBeth regressions of stock returns on beta, size, and valuation ratios; nonfinancial firms.
The table reports average slope coefficients from month-by-month Fama–MacBeth regressions using nonfinancial firms. Individual stock returns are regressed cross-sectionally on stock characteristics as of the previous month. The columns correspond to different regression specifications, with nonempty rows indicating the included regressors. The regressors include preranking CAPM β_t estimated using the past 12 months of daily returns with a five-lag Dimson (1979) correction; the log of month-end market cap (log*ME*); the log of book-to-market (log*BM*); the log of assets-to-market (log*AM*); EP^+, which equals the positive values of earnings-to-price, and zero otherwise; $D(EP<0)$, which equals one if earnings are negative, and zero otherwise; CP^+; and $D(CP<0)$ (with the last two similarly defined). The last row reports the average adjusted R-squared for each specification. The sample period is January 2000 through December 2016. The t-statistics based on Newey and West (1987) standard errors with four lags are reported in parentheses.

Quantity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.0134	0.0646	0.0610	0.0693	0.0642	0.0714	0.0602	0.0722	0.0750
	(1.78)	(3.69)	(3.40)	(4.08)	(3.79)	(4.16)	(3.39)	(4.41)	(4.51)
β	0.0010		0.0003	−0.0006	−0.0004	0.0014	0.0002	0.0007	0.0008
	(0.34)		(0.09)	(−0.23)	(−0.17)	(0.49)	(0.09)	(0.26)	(0.31)
log*ME*		−0.0056	−0.0052	−0.0051	−0.0051	−0.0072	−0.0053	−0.0068	−0.0069
		(−3.37)	(−3.04)	(−3.07)	(−3.14)	(−4.64)	(−3.09)	(−4.48)	(−4.44)
log*BM*				0.0053				0.0017	0.0032
				(2.79)				(0.96)	(1.76)
log*AM*					0.0041			0.0015	
					(2.55)			(0.94)	
EP^+						1.0123		0.8555	0.8572
						(4.78)		(4.51)	(4.51)
$D(EP<0)$						0.0001		−0.0012	−0.0007
						(0.03)		(−0.68)	(−0.36)
CP^+							0.0391	0.0122	
							(2.26)	(0.79)	
$D(CP<0)$							0.0009	0.0009	
							(1.31)	(1.20)	
R^2	0.0193	0.0264	0.0429	0.0643	0.0660	0.0615	0.0434	0.0825	0.0776

Table A4
Regressions of CH-3 and FF-3 size and value factors on the other model's factors.
Each row reports a factor's mean, its estimated monthly alpha (in percent), and the factor's loadings with respect to the other model. White (1980) heteroskedasticity-consistent t-statistics are in parentheses. The sample is from 2000 to 2016.

			Factors				
Dep. variable	Mean (%)	Alpha (%)	*MKT*	*SMB*	*VMG*	*FFSMB*	*FFHML*
FFSMB	0.64	−0.04	−0.01	0.95	−0.25		
	(1.84)	(−0.66)	(−0.69)	(38.79)	(−8.33)		
FFHML	0.84	0.34	0.05	−0.00	0.41		
	(2.60)	(0.97)	(0.96)	(−0.04)	(2.83)		
SMB	1.03	0.47	−0.01			0.89	0.00
	(3.26)	(7.03)	(−1.24)			(39.28)	(0.07)
VMG	1.14	1.39	−0.08			−0.50	0.14
	(4.35)	(7.93)	(−3.15)			(−11.15)	(3.49)

Bayesian comparison that follows the Stambaugh and Yuan (2017) procedure in applying the analysis of Barillas and Shanken (2018). Suppose we compare two models, M_1 and M_2, and before observing the data, we assign probabilities $p(M_1)$ and $p(M_2)$ to each model being the right one, with $p(M_1)+p(M_2)=1$. After observing the data, D, the posterior probability of model i is given by

$$p(M_i|D) = \frac{p(M_i)\cdot p(D|M_i)}{p(M_1)\cdot p(D|M_1) + p(M_2)\cdot p(D|M_2)}, \quad (4)$$

with model i's marginal likelihood given by

$$p(D|M_i) = \int_{\theta_i} p(\theta_i)p(D|\theta_i)d\theta_i, \quad (5)$$

in which $p(\theta_i)$ is the prior distribution for model i's parameters, and $p(D|\theta_i)$ is the likelihood function for model i. As shown by Barillas and Shanken (2018), when D includes observations of the factors in both models (including the market) as well as a common set of "test" assets, the latter drop out of the computation in Eq. (4). Moreover, that study also shows that when $p(\theta_i)$ follows a form as in Pástor and Stambaugh (2000), then $p(D|M_i)$ can be computed analytically. The key feature of the prior, $p(\theta_i)$, is that it is informative about how large a Sharpe ratio can be produced by combining a given set of assets; in this case, the assets represented by the model's factors. Specifically, the prior implies a value for the expected maximum squared Sharpe ratio, relative to the (observed) Sharpe ratio of the market. We use the Barillas and Shanken (2018) analytical results here to compute posterior model probabilities in the above two-way model comparison. Prior model probabilities of each model are set to one-half.

Fig. A1 displays posterior model probabilities in the comparison of CH-3 to FF-3. The value on the horizontal axis is the square root of the prior expected maximum

Table A5
Abilities of models CH-3 and FF-5 to explain each other's factors.
Panel A reports a factor's estimated monthly alpha (in percent) with respect to the other model (with White, 1980, heteroskedasticity-consistent t-statistics in parentheses). Panel B computes the Gibbons et al. (1989) F-test of whether a given model produces zero alphas for the factors of the other model (p-value in parentheses). The sample period is January 2000 through December 2016.

	Alphas with respect to:	
Factors	CH-3	FF-5
Panel A: Alpha (t-statistics)		
SMB	–	0.14
	–	(2.41)
VMG	–	0.43
	–	(4.39)
FFSMB	0.01	–
	(0.18)	–
FFHML	0.34	–
	(0.96)	–
FFRMW	−0.10	–
	(−0.86)	–
FFCMA	−0.08	–
	(−0.51)	–
Panel B: GRS F-statistics (p-value)		
SMB, VMG	–	8.43
	–	(0.0003)
FFSMB, FFHML,	0.29	–
FFRMW, FFCMA	(0.88)	–

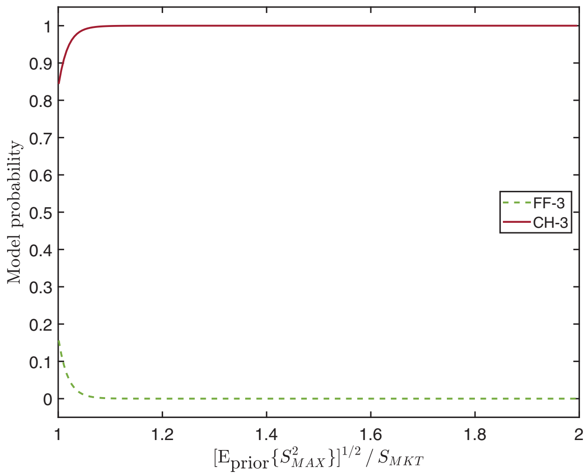

Fig. A1. Model probabilities comparing model CH-3 to model FF-3. The figure displays Bayesian posterior model probabilities for the two-way model comparison. The value on the horizontal axis is the square root of the prior expected maximum Sharpe ratio achievable by combining the model's factors, divided by the observed Sharpe ratio of the market. Prior model probabilities are equal. The sample period is from January 2000 through December 2016 (204 months).

squared Sharpe ratio achievable by combining the model's factors, $[\mathrm{E}_{\text{prior}}\{S^2_{MAX}\}]^{1/2}$, divided by the observed Sharpe ratio of the market, S_{MKT}. We see that the data strongly favor CH-3 over FF-3. In fact, the posterior probability of CH-3 essentially equals one if the Sharpe ratio multiplier on the horizontal axis exceeds only 1.05 or so, corresponding to a prior expectation that the market's Sharpe ratio can be improved only very modestly.

We also find that CH-3 dominates the frequentist comparisons reported in Table 5 when FF-3 is replaced by FF-5, the replication in China of the five-factor model in Fama and French (2015). Table A5 reports those results. As with FF-3, FF-5 fails to explain either *SMB* or *VMG* from model CH-3, leaving those factors' FF-5 alphas with t-statistics of 2.41 and 4.39 and producing a GRS p-value of just 0.0003. In contrast, the CH-3 alphas for the four nonmarket factors in FF-5 are all insignificant, having t-statistics of 0.87 or less in magnitude and producing a large GRS p-value of 0.88.

References

Allen, F., Qian, J., Qian, M., 2003. Comparing China' s financial system. China J. Financ. 1, 1–28.
Allen, F., Qian, J., Qian, M., 2005. Law, finance, and economic growth in China. J. Financ. Econ. 77 (1), 57–116.
Allen, F., Qian, J., Qian, M., Zhao, M., 2009. A review of China's financial system and initiatives for the future. In: Barth, J.R., Tatom, J.A., Yago, G. (Eds.), China's Emerging Financial Markets: Challenges and Opportunities. Springer, Boston, MA, pp. 3–72.
Allen, F., Qian, J., Shan, C., Zhao, M., 2014. The IPO of industrial and commercial bank of China and the "Chinese Model" of privatizing large financial institutions. Eur. J. Financ. 20 (7–9), 599–624.
Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. J. Financ. Mark. 5 (1), 31–56.
Baker, M., Stein, J.C., 2004. Market liquidity as a sentiment indicator. J. Financ. Mark. 7 (3), 271–299.
Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. J. Financ. 61 (4), 1645–1680.
Ball, R., 1992. The earnings-price anomaly. J. Account. Econ. 15 (2–3), 319–345.
Banz, R.W., 1981. The relationship between return and market value of common stocks. J. Financ. Econ. 9 (1), 3–18.
Barillas, F., Shanken, J., 2018. Comparing asset pricing models. J. Financ. 73, 715–753.
Basu, S., 1983. The relationship between earnings' yield, market value and return for NYSE common stocks: further evidence. J. Financ. Econ. 12 (1), 129–156.
Bian, J., Da, Z., Lou, D., Zhou, H., 2018. Leverage Network and Market Contagion. University of International Business and Economics, University of Notre Dame, London School of Economics, and Tsinghua University.
Brunnermeier, M.K., Sockin, M., Xiong, W., 2017. China's Model of Managing the Financial System. Princeton University and University of Texas at Austin.
Cakici, N., Chan, K., Topyan, K., 2017. Cross-sectional stock return predictability in China. Eur. J. Financ. 23 (7–9), 581–605.
Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. J. Financ. 56, 1–43.
Carhart, M.M., 1997. On persistence in mutual fund performance. J. Financ. 52 (1), 57.
Carpenter, J.N., Lu, F., Whitelaw, R.F., 2018. The real value of China's stock market. New York University and Massachusetts Institute of Technology. Unpublished working paper.
Carpenter, J.N., Whitelaw, R.F., 2017. The development of China's stock market and stakes for the global economy. Annu. Rev. Financ. Econ. 9, 233–257.
Chen, X., Kim, K.A., Yao, T., Yu, T., 2010. On the predictability of Chinese stock returns. Pac.-Basin Financ. J. 18 (4), 403–425.
Cheung, C., Hoguet, G., Ng, S., 2015. Value, size, momentum, dividend yield, and volatility in China's A-share market. J. Portf. Manag. 41 (5), 57–70.
Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. J. Financ. Econ. 7 (2), 197–226.
Eun, C.S., Huang, W., 2007. Asset pricing in China's domestic stock markets: is there a logic? Pac.-Basin Financ. J. 15 (5), 452–480.
Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. J. Financ. 47 (2), 427–465.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33 (1), 3–56.
Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116 (1), 1–22.
Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. J. Polit. Econ. 81 (3), 607–636.

Fan, L., Shan, Y., 2004. Trading volume, valuation ratio, momentum, and a three-factor model (Translated from Mandarin). J. Manag. Sci. 7 (3), 13–22.
Floros, I.V., Sapp, T.R.A., 2011. Shell games: on the value of shell companies. J. Corp. Financ. 17 (4), 850–867.
Gibbons, M.R., Ross, S.A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica 57 (5), 1121–1152.
Guo, B., Zhang, W., Zhang, Y., Zhang, H., 2017. The five-factor asset pricing model tests for the Chinese stock market. Pac.-Basin Financ. J. 43, 84–106.
Hirshleifer, D., Hou, K., Teoh, S., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? J. Account. Econ. 38, 297–331.
Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: an investment approach. Rev. Financ. Stud. 28 (3), 650–705.
Hsu, J., Viswanathan, V., Wang, M., Wool, P., 2018. Anomalies in Chinese A-shares. J. Portfolio Manag. 44, 108–123.
Hu, G.X., Chen, C., Shao, Y., Wang, J., 2019. Fama–French in China: size and value factors in Chinese stock returns. Int. Rev. Financ. 19, 3–44.
Jiang, J., Qian, K., Gong, F., 2016. Institutional Holdings' Structure Analysis in China (Translated from Mandarin). Shenwan Hongyuan Research, Co.
Lee, C., Qu, Y., Shen, T., 2017. Reverse Mergers, Shell Value, and Regulation Risk in Chinese Equity Markets. Stanford University and Tsinghua University.
Lee, E., 2013. Individual Stock Investor Sentiment, Stock Issuance, and Financial Market Anomalies. University of Pennsylvania.
Li, Y., Niu, J., Zhang, R., Largay, J.A., 2011. Earnings management and the accrual anomaly: evidence from China. J. Int. Financ. Manag. Account. 22 (3), 205–245.
Li, Y., Wu, S., 2003. An empirical analysis of the liquidity premium in China (Translated from Mandarin). Manag. Rev. 11, 34–42.
Li, Z., Yao, Z., Pu, J., 2007. What are the economics of ROE in China's stock market? (Translated from Mandarin). Chin. Account. Rev. 3, 305–314.
Lin, Q., 2017. Noisy prices and the Fama–French five-factor asset pricing model in China. Emerg. Mark. Rev. 31, 141–163.
Newey, W.K., West, K.D., 1987. A simple positive-definite heteroskedasticity and autocorrelation-consistent covariance matrix. Econometrica 55, 703–708.
Niederhoffer, V., Osborne, M.F.M., 1966. Market making and reversal on the stock exchange. J. Am. Stat. Assoc. 61, 897–916.
Pástor, L., Stambaugh, R.F., 2000. Comparing asset pricing models: an investment perspective. J. Financ. Econ. 56 (3), 335–381.
Pástor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. J. Polit. Econ. 111 (3), 642–685.
Ross, S.A., 2017. Factors—theory, statistics, and practice. J. Portf. Manag. 43 (5), 1–5.
Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? Account. Rev. 71 (3), 289–315.
Song, Z., Xiong, W., 2018. Risks in China's financial system. Annu. Rev. Financ. Econ. 10, 261–286.
Stambaugh, R.F., Yuan, Y., 2017. Mispricing factors. Rev. Financ. Stud. 30 (4), 1270–1315.
Wang, C., Chin, S., 2004. Profitability of return and volume-based investment strategies in China's stock market. Pac.-Basin Financ. J. 12 (5), 541–564.
Wang, F., Xu, Y., 2004. What determines Chinese stock returns? Financ. Anal. J. 60 (6), 65–77.
Wang, Y., Di Iorio, A., 2007. The cross-section of expected stock returns in the Chinese A-share market. Glob. Financ. J. 17 (3), 335–349.
White, H., 1980. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. Econometrica 48, 817–838.
Yang, X., Chen, Z., 2003. An empirical analysis of the three-factor model in China (Translated from Mandarin). Quant. Econ. Rev. 12, 137–141.
Zhang, Z., Liu, L., 2006. The turnover premium: is it a liquidity premium or a bubble? (Translated from Mandarin). Q. Econ. J. 3 (5), 871–892.