Finding Anomalies in China

Kewei Hou, Fang Qiao, and Xiaoyan Zhang*

This Draft: February 2021

Abstract

Using data on stock trading and accounting information from 2000 to 2018, we construct 426 anomalies and propose the multiple hurdle of 2.85 in the Chinese A-share stock market. With single sort portfolio analysis on value-weighted returns, we find that 98 (27) anomalies have significant raw returns at the 5% level with absolute *t*-value larger than 1.96 (2.85). After risk adjustment using the Liu, Stambaugh and Yuan (2019) three-factor model, 16 (2) anomalies have significant alphas for single (multiple) tests, about half of which are based on liquidity information, while alphas for accounting anomalies are less significant. After regressing on the four-factor model with turnover, the liquidity anomalies become insignificant. We construct the composite anomalies, and find that the majority can pass the multiple test hurdle.

Keywords: China, A-share stocks, anomalies, liquidity.
JEL Classification: G12, G14.

* We thank Lu Zhang, Xintong Zhan, and participants at the 2019 Summer Institute of Finance conference for helpful discussions and comments. Hou is from Fisher college of Business, The Ohio State University, 820 Fisher Hall, 2100 Neil Avenue, Columbus, Ohio 43210; Email: hou.28@osu.edu; Qiao is from School of Banking and Finance at University of International Business and Economics, 10 Huixindongjie, Chaoyang District Beijing, 100029, P.R. China; Email: qiaofang@uibe.edu.cn. Zhang is from PBC School of Finance at Tsinghua University, 43 Chengfu Road, Haidian District Beijing, 100083, P.R. China; Email: zhangxiaoyan@pbcsf.tsinghua.edu.cn.

Finding Anomalies in China

This Draft: February 2021

Abstract

Using data on stock trading and accounting information from 2000 to 2018, we construct 426 anomalies and propose the multiple hurdle of 2.85 in the Chinese A-share stock market. With single sort portfolio analysis on value-weighted returns, we find that 98 (27) anomalies have significant raw returns at the 5% level with absolute *t*-value larger than 1.96 (2.85). After risk adjustment using the Liu, Stambaugh and Yuan (2019) three-factor model, 16 (2) anomalies have significant alphas for single (multiple) tests, about half of which are based on liquidity information, while alphas for accounting anomalies are less significant. After regressing on the four-factor model with turnover, the liquidity anomalies become insignificant. We construct the composite anomalies, and find that the majority can pass the multiple test hurdle.

Keywords: China, A-share stocks, anomalies, liquidity.
JEL Classification: G12, G14.

1. Introduction

The Chinese stock market is the second largest stock market in the world, with 3,666 listed stocks and total market capitalization of 43 trillion RMB in 2018. With the country's large population, the Chinese stock market is an important investment channel for domestic investors. Meanwhile, given its relatively low correlation with the rest of the global stock market, the Chinese stock market also attracts substantial attention of global investors for the purpose of international diversification. However, with its short history and unique growth path, it has been challenging for local investors and foreign investors to fully understand the opportunities and risks in Chinese stock market. In this study, we investigate anomalies in the Chinese stock market, in order to study the cross-sectional patterns among the stocks, which can help both global and local investors allocate their money efficiently in the Chinese stock market.

The classic asset pricing theory states that returns are compensations for risks, and many asset pricing models have been introduced in the previous literature to explain returns. For instance, the CAPM (Sharpe (1964); Lintner (1965)), the consumption CAPM (Breeden (1979)), and the intertemporal CAPM (Merton (1973)). In recent years, the Fama and French (1993) three-factor model has been widely adopted as an empirical benchmark model in the United States. A similar study by Liu, Stambaugh and Yuan (2019) investigates the size and value effects in China, and introduces a three-factor model for the Chinese stock market.

Unfortunately, many return patterns cannot be fully explained by asset pricing models, and these unexplained patterns are named "anomalies". Anomalies are an important channel to understand the cross-sectional stock return patterns. For instance, Hou, Xue and Zhang (2019) examine 452 stock return anomalies to understand the cross-sectional patterns in the United States. However, existing studies on Chinese stock anomalies mostly only focus on a small number of anomalies, and thus the understanding on anomalies in China has been fragmented. In this study, we collect 426 anomalies, to have a comprehensive understanding of the Chinese A-share stock market.

One might think that anomaly patterns in China would be similar to the rest of the world, including those in the United States. Nevertheless, previous literature shows that the reality can be much more complicated than that. For instance, Titman, Wei and Xie (2013) and Watanabe, Xu, Yao and Yu (2013) study investment anomalies in the United States and other markets, and find investment anomalies survive in more developed markets, while they do not in less developed markets. That is to say, the existence and patterns of anomalies can be heavily influenced by country-level environment. Given its short 30 years of history, the Chinese stock market is still under development, while the stock markets in other developed countries with long histories, such as the United States, are likely much more efficient. Therefore, the patterns of Chinese anomalies can be significantly different from these in other countries.

In our opinion, the Chinese stock market differs from the U.S. stock market in two important aspects: trading environment and information environment. The trading in China is mostly driven by retail investors, while the trading in the United States is mostly driven by institutional investors. According to the Shanghai Stock Exchange, which is the larger stock exchange of the two stock exchanges in China, more than 80% of daily volume comes from retail investors, while only 15% comes from institutional investors. For the United States, on the contrary, 90% of trading is from institutional investors, while only 10% is from retail investors. What is the difference between retail investors and institutional investors? A long literature studying retail investors' behaviors, such as Barber and Odean (2000), Barber and Odean (2008), Barber, Odean and Zhu (2009), shows that retail investors are less sophisticated, more overconfident, and can lose money by excessive trading. According to Jones, Shi, Zhang and Zhang (2020), the poorer Chinese retail investors normally buy and sell in the opposite directions of future price movements, and they on average trade too much. Meanwhile, retail investors' active trading provides significant liquidity to the market. So maybe it is not surprising to find that the Chinese stock market is more liquid than most of the other countries. However, too much of trading can also leads to price bubbles and excessive volatility. It would be interesting to see how the high liquidity affects the cross-sectional anomaly return patterns in China.

In terms of information environment, regulations promoting fair disclosure, law enforcements against insider trading, and sophisticated news releases system all support efficient information environment in the United States. In contrast, with a short history of barely 30 years, the rules and laws regarding capital market information environment are still incomplete in China. For instance, to integrate with the global capital market, China adopted the global accounting standards in 2007. However, the cost of manipulation of accounting statement is low, and therefore there may exist frauds or leakage in information disclosure. Meanwhile, financial service professionals, such as analysts, are an important part for the investment community. The analyst forecast industry only became standard in China in 2003, and has enjoyed tremendous growth over the past 17 years, which presumably positively affects overall market efficiency. However, the quality of analysts' forecasts is still low, and analysts' recommendations are often speculative. It is important to understand whether the information environment affects anomaly patterns, especially those based on information from financial statements.

Given the two major differences in trading and information environment, we divide the 426 anomalies into two groups: 192 are trading-related and 234 are accounting-related. We adopt the portfolio forming approach and sort all A-share stocks into ten groups based on each of the 426 anomaly variables identified in previous literature. Our calculation take into account many unique characteristics of the Chinese stock market, such as the shell-value contamination (see, Liu, Stambaugh and Yuan (2019)), the split-share structure reform, the daily price limit rule, shorting ban, and different financial reporting standards. We also exclude the smallest 30% of the stock universe to make sure that the anomalies are not driven by tiny stocks.

For long-short portfolio raw returns, we find that 98 out of 426 anomalies (about 23%) are significant at the 5% level for single test (the t-value≥1.96), with the monthly long-short returns ranging between 0.29% and 1.53%. Among these 98 significant anomalies, 31 anomalies use information related to trading, and 11 of them use firm-level liquidity measures, indicating that liquidity is an important driver of cross-sectional returns in China. Meanwhile, the rest 67 significant anomalies are based on accounting information, and 27 of these use firm-level profitability measures. That is, firm profitability also seems to be an important driver of cross-sectional return patterns.

It is possible that the significant anomaly raw returns are compensations for exposures to systematic risk factors. Therefore, we compute risk-adjusted returns, or alphas, using both the CAPM and the Chinese Fama-French 3-factor model (CH3 model) and 4-factor model with turnover (CH4) from Liu, Stambaugh and Yuan (2019). After risk adjustment, we find that 97 CAPM alphas are significant, 16 CH3 alphas are significant, and 15 CH4 alphas are significant. The CAPM alphas are very similar to the raw returns, indicating that the market risk fails to explain these anomaly returns. In contrast, with embedded size, value and turnover effects, the CH3 and CH4 models are able to explain away a significant part of anomaly raw returns, and thus we focus our discussion on alphas using this model. Among the 16 significant alphas relative to the CH3 factors, 7 anomalies are constructed using liquidity-related information, and

4 are related to accounting-related information. Among the 15 significant alphas relative to the CH4 factors, half of anomalies are from trading-related information and half of them are from accounting-related information. It explains away the liquidity anomalies.

Based on the 426 anomalies, we propose that the multiple test hurdle should be 2.85 at the 5% level in Chinese stock market using the double bootstrapping method of Harvey and Liu (2020). We find that 27 anomalies are significant. 8 are constructed with trading-related information, and 19 are constructed with accounting-related information. With risk adjustment with the CH3 factors, only 2 anomalies are significant, which are from the liquidity group. After risk adjustment with the CH4 factors, no anomalies are significant.

Then we construct the composite anomalies with score, Fama-MacBeth regression, Lasso, and Random Forest methods by combining many signals. In total, we construct 10 composite anomalies with all characteristics in each category, sub-category, and in total. We find that the majority of the composite anomalies can pass the multiple hurdle of 2.85. After risk adjustment with the CH3 factors, almost half of the composite anomalies are still significant, with the majority of them from trading-related anomalies. There are still significant CH4 alphas using random forecast method, which are from the trading-related group. It indicates that the popular machine learning methods of Lasso and Random Forest outperform the traditional methods of score and Fama-MacBeth regression. The composite anomalies confirm that trading-related anomalies are very important in China.

We compare the anomaly results in China with those in the United States, as in Hou, Xue and Zhang (2019), and identify three major differences. First, out of the 426 anomalies, 23% of anomalies are significant in China, which is lower than the replication rate (35%) in the United States, with the same hurdle being the absolute *t*-value above 1.96. It is likely that the return patterns in China are fundamentally different from the United States, and it is also possible that stock returns in China are more volatile than those in the United States, and the t-statistics are in general lower.

Second, most of the significant anomalies in China are trading-related anomalies, using characteristics such as volatility, turnover, and dollar trading volume, especially liquidity measures; while almost all trading frictions anomalies are insignificant in the United States as in Hou, Xue and Zhang (2019). This difference is likely driven by the fact that daily trading volume in China mostly comes from individual investors, who chase past good performance and drive up trading liquidity, while this excessive trading volume leads to future low returns. The magnitudes of trading frictions in the United States are much smaller, possibly because the U.S. market is far more efficient than the Chinese market, with its longer history and dominance of institutional investors.

Third, most of the significant anomalies in the United States are driven by firm fundamental information in accounting statements, such as value, investment, and profitability; while in China, these variables mostly cannot provide statistically significant information about future stock returns. These results indicate that the accounting information is not as informative in China about future stock returns, possibly due to the immature accounting information infrastructure, frauds or leakage in information disclosure.

Our study is related to three different areas of previous literature. First, our work naturally connects to previous studies on Chinese stock return patterns, such as Hsu, Viswanathan, Wang and Wool (2017), Liu, Stambaugh and Yuan (2019), Carpenter, Lu and Whitelaw (2018), and Chen, Kim, Yao and Yu (2010). Second, we are also connected to the anomaly literature. Our collection and construction of all anomalies in China follow the study of Hou, Xue and Zhang

(2019) in the U.S. market. We construct the multiple test hurdle following Harvey and Liu (2020). Finally, our comprehensive study of Chinese stock return anomalies is particularly related to the previous anomaly papers for the global capital market. For instance, Titman, Wei and Xie (2013) and Watanabe, Xu, Yao and Yu (2013) study investment in more than 40 countries, and Chui, Titman and Wei (2010) studies momentum in 41 countries. All papers find that some anomalies working in the United States may not be replicated in other countries. Jacobs and Muller (2019) study how 241 anomalies decay after publication across 39 international markets, and find that the United States is the only country with a reliable post-publication decline in long-short returns.

Compared to the previous literature, our study contributes to the literature in three ways. First, we provide the first comprehensive study on 426 anomalies in the Chinese A-share stock market, and complement prior studies on Chinese stock return patterns. Second, we first propose the multiple hurdle should be 2.85 in China. Third, we document significant differences between the anomalies in China and in the United States. Trading-related anomalies, especially those related to liquidity, are more significant in China, while the accounting information based anomalies are less significant in China. Our finding complements the literature on anomalies in international financial markets.

The remainder of this paper is organized as follows. Section 2 introduces the data and methodology used in this paper. We provide our main empirical results on anomalies in Section 3. Section 4 concludes.

2. Data and Methodology

2.1 Data

There are two stock exchanges in China: the Shanghai Stock Exchange and the Shenzhen

Stock Exchange. We obtain daily stock trading data and quarterly accounting data for stocks traded on these two exchanges from Wind Information Inc., which is the largest financial data provider in China.

The Chinese data are available from 1990 to 2018, and we choose the sample period for our main analysis to be January 2000 to December 2018 for three reasons. First, to build anomalies in the cross section of stocks, we need a large number of stocks, and the stock sample before 2000 cannot ensure sufficient number of stocks in some years. Second, there are several changes in accounting standards in China. The laws and regulations on financial reporting were implemented around 2000, and the accounting reporting becomes more comparable across firms afterwards. Third, regular trading rules, such as the daily price limit rule, are put in place before 2000, which greatly reduces noises and extreme moves in stock prices.

The trading data include stock code, trading date, close price, trading volume, trading volume in RMB, turnover, return with dividends, free float A shares, total A shares, and total shares. Three different types of shares coexist in the Chinese stock market, which are A shares, B shares, and H shares. A shares are the shares of Chinese firms that listed on Shanghai and Shenzhen exchanges, and are open to domestic investors to trade. These stocks are our main research subject in this study. More detailed discussion on the share classes are provided in Online Appendix A. By the end of 2018, there are 3,666 A-share stocks, including both actively traded and delisted stocks.

For the accounting information, the data on the balance sheet and income statement start in 1990, and the data on the cash flow statement start from 1997. Before 2002, financial statements are reported annually or semiannually, and after 2002, they are reported quarterly. To keep data structure consistent throughout the paper, we transform the annual (in calendar December) and semiannual data (in calendar June) to quarterly data as follows. Before 2002, for annual income and cash flow statements released in December, we compute the quarterly data as half the net change between the calendar June and December data. For the semiannual data in calendar June, we compute the quarterly data as half of the semi-annual data. For the data on balance sheet, income statement and cash flow statements, we fill in the missing quarterly data for calendar March and September with the most recent quarterly data.

We obtain the analyst earnings forecasts data from Go-goal.com, which is the most comprehensive, accurate analysts' forecasts database, and widely used in fund industry in China. The analyst sample period is from 2005 to 2018. It includes 1.6 million reports produced by more than 22,000 analysts (90% of analysts in the market) from 383 organizations. The analyst dataset contains information regarding analysts' name, analysts' organization, forecast periods, analysts' forecast release date, forecasted earnings per share, and earnings announcement date. Finally, we obtain quarterly institutional ownership from CSMAR, covering a sample period of 2000 to 2018. It reports the proportion of shares held by funds, brokers, insurance, security funds, insurance companies, and trust companies etc.

We merge the above data based on stock code, which is unique for each firm. Following previous literature, we impose several filters on our datasets. First, we drop the firms that have become public within the past six months to avoid extreme volatility and illiquidity in stock prices right after IPOs. Second, since it is hard to apply for IPO, small firms are likely targets as shells for firms which want to go public. To avoid this shell-value contamination in small firms, we drop 30% of microcap firms in the sample following Liu, Stambaugh and Yuan (2019). Third, in order to guarantee stock liquidity and data quality, we only keep firms that have a minimum of 75% of non-zero-volume trading days with trading records during our recent certain sample periods. This filter also helps to prevent influences from firms with long trading suspension.

Table 1 provides simple summary statistics of overall Chinese A-share stock market by year from 2000 to 2018. It includes the number of firms (including actively traded and delisted firms), total A-share market capitalization (sum of all A-share stocks' capitalization), average firm size (product of number of total A shares outstanding and close price), annualized average stock returns, annualized average volatility (computed as annualized standard deviation of daily stock returns over one year), average stock share turnover (discussed below), and average institutional ownership (discussed below). From columns (1)-(5), it is clear to see that the Chinese stock market grows rapidly. The number of firms rises from 681 in 2000 to 2,465 in 2018, and the total A-share market capitalization rises from 1,255 billion RMB in 2000 to 40,533 billion RMB in 2018. In comparison, for the same time span, the total market capitalization of the United States rises from 16,094 billion dollar in 2000 to 33,979 billion dollar in 2018, which is decreasing from around 90 times larger to 5.5 times larger than the Chinese market. The average Chinese firm size also grows from 1,843 million RMB in 2000 to 16,444 million RMB in 2018. Columns (6)-(9) show the firm average return and volatility patterns in China and the United States. The annual stock market returns take plenty of extreme positive and negative values. For instance, for a total of 19 years, the annual returns are above 20% in seven years, and they are below -20% for four years. For average stock volatility, 11 years have annual volatility above 40%. In comparison, the annual stock market returns are mostly between -20% and 20%, and the annual volatilities are mostly below 45% in the United States. That is, the Chinese stock returns are much more volatile than those in the U.S. market.

As mentioned in the introduction, the Chinese stock market is significantly different from the U.S. stock market in two aspects: trading and information environment. Columns (10)-(11)

show the average firm-level turnover. For each stock, we measure stock share turnover as the sum of trading volume in one year divided by free float A shares outstanding at the year end. Then for each year, we average across all firms to obtain average firm-level turnover ratio. The overall average of turnover is around 4.73 in China, significantly higher than that in the U.S. market, which is 2.33. We also observe that turnover exhibits substantial fluctuations over time in China, and there are large spikes in 2007, 2009, and 2015, with turnover exceeding 8. The large spikes are likely driven by excessive trading around large movements in stock prices, especially during market boom periods.

Columns (12) and (13) present the annual average institutional ownership across stocks for both China and the United States. For each stock and year, we first sum the proportion of shares held by funds, brokers, entrust etc., and then calculate the average of institutional ownership across all firms. The overall average institutional ownership is around 7% in China, while for the same period, the number is 47% for institutional holdings in the United States using the Thomson Reuters 13F database. All institutional investment managers, including mutual funds, hedge funds, trust companies, pension funds, insurance companies, etc., with at least $100 million in assets under management are required to disclose to the SEC their end-of-quarter equity holdings on Form 13F. Over time, we observe that the Chinese market institutional ownership increases from 2% in 2000, arriving at a peak above 11% in 2008. But afterwards, over 2008 to 2018, it decreases to a low level of 6% in 2018. This is because the issue of funds was suspended in the last quarter of 2007, and then it follows market downturn afterwards.

2.2 Anomaly Construction

Hou, Xue and Zhang (2019) provide one of the most comprehensive anomaly studies in the United States, with 452 anomalies. In this study, we follow their construction of anomalies.

Altogether, we are able to collect 426 anomaly variables in China due to data availability and differences in accounting standards. To better organize this large number of anomalies, we divide the 426 anomalies into two larger categories: trading-related anomalies, and accounting information-related anomalies.

For trading-related anomalies, we further separate them into three sub-categories: 94 anomalies using liquidity proxies, such as volume and turnover; 46 anomalies using risk proxies, such as betas and volatilities; and 52 anomalies using past return information. For accounting information related anomalies, we also further separate them into four sub-categories: 67 anomalies using profitability proxies, such as return on assets and return on equity; 44 anomalies using value proxies, such as book-to-market ratios and earnings-to-price ratios; 51 anomalies using investment proxies, such as asset growth and investment growth; and 72 anomalies using other accounting information variables, such as R&D expense ratio, leverage, and analysts' forecasts etc. The categories and subcategories are all reported in Table 2. More details on the sort variables are provided in Online Appendix B.

[Insert Table 2 here]

We implement single portfolio analysis to test whether the anomalies can predict stock returns. For each anomaly, we follow the [k, m, n] approach as in Jegadeesh and Titman (1993). Suppose the current month is t. We first collect firm-level information to be used as anomaly sorting variables over [t-k+1, t]. At time t, we sort A-share stocks into decile groups based on the sorting variable. Decile 1 contains the 10% of stocks with the lowest sorting variable values, while decile 10 contains the 10% of stocks with the largest sorting variable values. Then we wait a period of m months over [t, t+m]. If investors prefer immediate investment, we set m to be zero. We hold the decile portfolios, determined at time t, over the next [t+m+1, t+m+n] months. The value-weighted decile portfolio raw returns are computed using previous month-end firm market capitalization as weights. We also provide results using equal weights. If the anomaly sorting variable can predict future return, we expect to see a significant difference in raw returns between decile 10 and decile 1, over the holding period of [t+m+1, t+m+n].

The choice of the parameters k, m, n depends on the specific anomalies. For instance, the waiting period m is zero except for momentum anomalies, and the holding period n is generally set to be 1-, 6-, or 12-month. The anomaly sorting variables, if coming from accounting information, might be constructed over the previous year or quarter, while if the sorting variables are from trading, they can be constructed over the previous month. The choice of k, m, n for each anomaly is clearly marked for each anomaly in Online Appendix B.

Since Chinese trading rules and accounting standards are different from the United States, we made several adjustments to the original sorting procedures by Hou, Xue, and Zhang (2019) for the U.S. data. For example, due to daily price limit of 10%, we use the average of the five highest daily returns in one month to construct the maximum daily return, rather than using one highest daily return in one month as in the original procedure. We also exclude the stock in the portfolio if it is suspended with zero trading volume during the portfolio sorting date.

2.3 Multiple Tests

Multiple testing gains attention in finance literature in recent years for data snooping biases. This can help solve the concern that many discoveries may be false. For example, Harvey, Liu, and Zhu (2016) propose that the multiple test hurdle controlling for false discovery rates in the U.S. market should be above 3 based on more than 300 anomalies. Since 426 anomalies are constructed in China, what should be the multiple test hurdle based on Chinese data property?

To construct the multiple test hurdle in China, we first consider three methods that control for false discovery rates in financial economics, including the methods of Benjamini and Hochberg (1995) (BH), Benjamini, Hochberg, and Yekutieli (2001) (BHY), and Barras, Scaillet, and Wermers (2010) (BSW), which is based on the method of Storey (2002). The detail of these four methods is shown in Internet Appendix C. We find that the BHY method is too conservative to produce the multiple t-cutoff in China. The BH and BSW methods can produce a t-cutoff 3.16 and 2.55, respectively, but they do not allow dependence and cannot solve short sample.

In order to allow dependence and solve short sample for 426 anomalies in China, we consider the double bootstrapping method of Harvey and Liu (2020) controlling for both Type I error (false discovery rates among all discoveries) and Type II error (missing discovery rates among all non discoveries) as an optimal choice to find out the multiple test hurdle in China. The advantage of controlling for both Type I and II errors is as follows. After controlling for Type I error rates, the t-cutoff becomes higher. If the t-cutoff is too high (too conservative), it may lead to high Type II error rates (cannot detect true anomalies).

Suppose we have N anomalies and D time periods. We arrange the data into a D*N data matrix X_0. We believe that a fraction, p_0 of the N anomalies are true to control for Type II error.

Step I: We bootstrap the time periods and create an alternative panel of long-short returns, X_i. For X_i, we calculate the corresponding *1*N* vector of t-statistics, t_i.

Step II: We rank the anomalies based on their t-statistics, t_i. For the top $\mathrm{p}_0 * N$ anomalies with the highest t-statistics, we find the corresponding anomalies in X_0. We adjust these anomalies in X_0 so that their means are the same means for the top $\mathrm{p}_0 * N$ in X_i. We denote the data matrix of these adjusted strategies by $X_{0,1}^{(i)}$. For the remaining anomalies in X_0, we adjust them so they have a zero in-sample mean. We denote the data matrix for these adjusted anomalies by $X_{0,0}^{(i)}$. We arrange $X_{0,1}^{(i)}$ and $X_{0,0}^{(i)}$ into a new data matrix Y_i by concatenating the two data matrices.

Step III: We bootstrap Y_i the time periods J times. For each bootstrapped sample, we calculate the error rates for Y_i for a statistical procedure, denoted by $f_{i,j}$.

Step IV: Repeat Steps I to III I times. We calculate the final bootstrapped error rate as $\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} f_{i,j}$.

We suppose that 15% of anomalies are believed to be true. The reasons for the choice of $\mathrm{p}_0 = 15\%$ are as follows. First, 22% of 426 anomalies have t-statistics greater than 2. The choice of p_0 should below 22%. Second, we balance Type I and Type II errors and achieve an odds ratio of around 1/5, that is, on average five misses for each false discovery. Given the significance level (Type I error), the Type II error rate and odds ratio should be at a reasonable level. When the Type I error rate is 5%, the Type II error rate is 6%, and the odds ratio is 15%. The Type II error rate and odds ratio are at a reasonable level, compared with the other two values of p_0 (p_0=10% and 20%).

Figure 1 plots the error rates across a range of t-statistics for 426 anomalies. When the threshold t-statistic increases, the Type I error rate declines while the Type II error rate increases. In addition, the odds ratio, which is the ratio of false discoveries to misses, also decreases as the threshold t-statistic increases. We highlight the threshold t-statistic that achieves 5% significance level (Type I error rate). Therefore, when we choose p_0=15%, the multiple t-cutoff should be 2.85 at the 5% significant level.

[Insert Figure 1 here]

In order to show the t-cutoff is robust and stable to the choice of p_0, we plot the t-cutoff against p_0 in Figure 2. In general, the cutoff t-statistic declines as p_0 becomes larger, since a discovery is less likely to be false when a larger fraction of anomalies are true. The t-cutoff is relatively stable around 2.85.

[Insert Figure 2 here]

3. Empirical Results

This section contains our main empirical results. We report the significant individual anomalies in Section 3.1. Result for composite anomalies is provided in Section 3.2.

3.1 Individual Anomalies

3.1.1 Summary of Anomaly Portfolio Returns

Following the portfolio sorting procedure in Section 2.2, we compute the time-series average of monthly decile portfolio returns over the whole sample period. If there are significant return differences between decile 1 and decile 10, then the sorting variable can generate cross-sectional return differences. Other than documenting cross-sectional return patterns, anomalies have been routinely used as trading strategies. For instance, investors would long the portfolio with high return and short the portfolio with low return, which forms a long-short strategy. Following the previous literature, our results are based on the returns on the long-short strategy.

In addition to the raw returns, we also compute risk-adjusted returns using the CAPM and the three factor model, CH3, and the four-factor model, CH4, from Liu, Stambaugh and Yuan (2019). Under the CAPM assumption, the only relevant pricing factor for returns is the market factor, measured by market returns minus risk-free interest rate. We compute market return as the return on the value-weighted portfolio of the top 70% of A-share stocks, and we use the one-year deposit rate as the risk-free rate. Both data items are obtained from Liu, Stambaugh and Yuan (2019) through WRDS. For the long-short strategy, we estimate the CAPM-adjusted return, α_j^{CAPM}, for anomaly j as:

$$R_{jt}^{long} - R_{jt}^{short} = \alpha_j^{CAPM} + \beta_{MKT,j}^{CAPM} MKT_t + e_{jt}. \quad (1)$$

Here variable R_{jt}^{long} represents the long-leg return, R_{jt}^{short} represents the short-leg return, variable is MKT_t is the excess return on the market portfolio, and $\beta_{MKT,j}^{CAPM}$ is the anomaly j's exposure to the market risk. The alternative model, CH3, has three factors. The first factor, the market factor, is the same as in the CAPM. To construct the size and value factors, we first independently sort stocks into two size groups and three value groups. The breakpoints for the size groups is the median market capitalization of the top 70% of A-share stocks, and the breakpoints for the value groups are the 30th and 70th percentiles of the earnings-to-price ratio (E/P) for the top 70% of A-share stocks. We then construct six portfolios by the intersection of the size and value groups. The size factor, SIZE, and is computed as the difference between the simple average of the returns on three portfolios of small stocks (under the median market capitalization) and the simple average of the returns on three portfolios of big stocks (above the median market capitalization). The value factor, VALUE, is computed as the difference between the simple average of the returns on two portfolios of value stocks (above the 70th percentile of E/P) and the simple average of the returns on two portfolios of growth stocks (below the 30th percentile of E/P). The CH4 model has four factors. Except for the market, size, and value factors, the fourth factor is the turnover factor, 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 the 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 estimate the CH3 and CH4 risk-adjusted returns similarly to those in equation (1) by replacing the market factor by the three and four factors.

To have an overview of all 426 anomalies in Chinese stock market, we summarize in Table 3 the number of significant anomalies after excluding the microcap stocks in the bottom 30% of the market. The statistical significance is based on the t-statistics, which is computed using the Newey and West (1987) standard errors for the time series of monthly returns with 4 lags. Following the previous literature, we choose to report anomalies with absolute t-statistics larger than 1.96 for single test and larger than 2.85 for multiple test, which corresponds to a 5% significance level.

[Insert Table 3 here]

Table 3 Panels A and B present significant anomalies using the long-short strategies for value-weighted and equally-weighted portfolio returns with single test and multiple test, respectively. In total, 98 out of 426 anomalies (23%) have significant long-short raw returns for single test. The 98 significant anomalies include 11, 11, 9, 27, 14, 2, and 24 from the liquidity, risk, past returns, profitability, value, investment, and other subcategories, respectively. Among the six sub-categories, liquidity, risk, profitability, value, and "other" sorting variables seem to generate the majority of the significant anomalies in the cross section of stocks.

Given the possibility that the anomaly raw returns could be compensations for exposures to systematic risk factors, we present the risk-adjusted returns using the CAPM, the CH3 factors, and CH4 factors in the second, third, and fourth rows, respectively. For the CAPM, risk-adjusted returns are significant in 97 cases, which is similar to the number of significant anomalies using raw returns. This indicates that significant returns are not compensations for exposures to the market risk. When we use the CH3 model, with size and value factors, 16 anomalies have significant long-short risk-adjusted returns, indicating that most of the anomaly raw returns can be attributed to size or value effects. Out of the 16 significant CH3 alphas, 12 of them are related to trading information, while only 4 are related to accounting information. In particular, 7 of the significant alphas are in the liquidity sub-category, and 3 belong to past returns category. That is, liquidity seems to be an important main driver of the cross-sectional alpha patterns. But the accounting information does not create significant CH3 alphas. When regressing on the CH4 factors, we find that 15 anomalies have significant alphas with 7 from trading-related anomalies and 8 from accounting-related anomalies. It indicates that liquidity anomalies can be explained by the CH4 factors with turnover.

When it comes to multiple test, we find that only 27 anomalies can pass the hurdle, which contain 4, 4, 10, 4, 5 from liquidity, risk, profitability, value, and other anomalies. The CAPM alphas of these 27 anomalies are still significant for multiple test. When regressing on the CH3 factors, only 2 anomalies can pass the hurdle, which come from the liquidity group. When we use the CH4 factors as risk adjustment, no anomalies have significant alphas.

For the equally-weighted portfolios in Panel B, we find that 175 out of 426 anomalies (41%) have significant long-short raw returns, which is much larger than the number of significant cases in Panel A. This might not be surprising, because value-weighted returns are tilted towards large firms, and equal-weighted returns are more balanced for small firms. Previous anomaly literature shows that anomalies are more likely to be significant for smaller firms due to difficulty of arbitrage on these firms. Among the 175 significant anomalies, 42, 26, 14, 37, 21, 7, and 28 are from the liquidity, risk, past returns, profitability, value, investment, and other subcategories, respectively. The result for equally-weighted portfolios further confirms that liquidity, profitability, and “other” variables are important drivers of cross-sectional return patterns. After we use the CAPM, the CH3 and CH4 factors as risk adjustment, the result is similar to the value-weighted result. When using the CAPM as risk adjustment, we find that 175 anomalies have significant alphas. Again, this finding indicates that the market risk premium cannot attribute to these significant returns. For the CH3 factor model, 59 risk-adjusted returns are significant with 30 from liquidity and 12 from accounting-related anomalies. For the CH4 factor model, 54 alphas are significant with 18 from liquidity and 16 from accounting-related anomalies. It means that liquidity is very important for the alpha returns, while accounting-related information matters less.

For multiple test, we find that 94 equally-weighted anomalies have significant long-short returns. 25, 13, 6, 25, 7, 2, 16 anomalies come from liquidity, risk, return, profitability, value, investment, and “other” groups. When we regress on the CAPM, CH3, and CH4 factor models, the result is quite similar to the single test.

To understand how large the returns are, we plot in Figure 1 the magnitudes of monthly value-weighted long-short raw and risk-adjusted returns relative to the CAPM, the CH3, and CH4 model. In Panel A, we present the 98 portfolios with t-statistics for raw returns to be higher than 1.96, ranked by the magnitude of raw returns in Panel A. These long-short portfolios have returns ranging between 0.29% and 1.53%. Consistent with Panel A of Table 3, the CAPM alphas are similar in the magnitude of the raw returns, indicating that the market risk exposure is not the reason for the significant long-short returns. The CH3 alphas are substantially smaller than the raw returns. In many cases, the alphas reduce to close to zero, and that is why the CH3 alphas are only significant in 16 cases. For the 16 significant CH3 alphas, 4 of them are still highly significant with monthly risk-adjusted returns larger than 1%, such as returns on the strategy “vturn1-1” (discussed in Section 3.2), indicating that the CH3 factors cannot explain away all significant long-short strategy returns.[1] The CH4 alphas are similar to the CH3 alphas for the

[1] Given the shorting constraints in Chinese stock market, the short side of the strategy can be difficult to implement. Here our purpose of this study is to identify cross-sectional return patterns rather than trading on them, so we still rely on long-short strategy for our empirical analysis. For interested readers, we report long-leg returns for references in Online Appendix D.

majority of anomalies. The CH4 alphas are below the CH3 alphas for the first several anomalies, which belong to the liquidity group.

[Insert Figure 1 here]

Panel B plots the 27 anomalies that pass the multiple test hurdle of 2.85. These long-short portfolio returns range between 0.54% and 1.53% per month. The CAPM alphas have similar magnitude with the raw returns. The CH3 and CH4 alphas are similar.

3.1.2 Individual Anomaly Portfolio Returns

In this subsection, we provide details on 98 anomalies with significant value-weighted long-short raw returns, with the cutoff t-statistic being 1.96. We report the long-short strategy's value-weighted raw returns, CAPM alpha, CH3 alpha, CH4 alpha, and their t-statistics in Table 4. The result for insignificant anomalies is reported in Online Appendix D.

[Insert Table 4 here]

Trading-Related Anomalies

Liquidity

Liquidity is generally described as the ability to trade large quantities of assets quickly at low cost with little price impact. It is now well known and accepted that measures of liquidity can influence asset prices (see Amihud and Mendelson (1986)). For more than twenty years, researchers have examined the importance of liquidity in cross-sectional return patterns, and found various ways to measure liquidity. In these papers, they document a negative and strong relation between average returns and liquidity. In this study, we examine 94 liquidity anomalies, which are measured using information regarding share turnover, dollar trading volume, Amihud illiquidity, etc.

Table 4 Panel A shows the raw returns and abnormal returns of 11 anomalies with significant raw returns for single test. The first liquidity measure is average of share turnover, as introduced in Datar, Naik and Radcliffe (1998). Share turnover is computed as the one-month average of trading volume divided by float A shares. For this strategy, we choose the forming period to be 1 month (k=1), without waiting (m=0), and holding period to be 1 month (n=1). The results are reported in the row of "turn1-1". The relationship between share turnover and stock returns is negative, indicating that stocks with low liquidity perform better in the next month. If we long the stocks in decile with the lowest share turnover, and short the stocks in the decile with the highest share turnover, the raw long-short monthly return is 1.43%, with a t-statistic of 3.17. Using the CAPM as a risk adjustment model, the abnormal return (alpha) becomes 1.54% per month with a t-statistic of 3.39. When we use the CH3 model for risk adjustment, the abnormal return stays at 1.14% with a significant t-statistic of 2.63. When we use the CH4 factor model for risk adjustment, the alpha becomes 0.37% with the t-statistic of 0.92. All the statistics show that share turnover is a statistically significant anomaly. The higher return on firms with lower liquidity is consistent with the illiquidity premium hypothesis in the previous literature, such as Liu, Stambaugh and Yuan (2019).

Similar patterns exist when we use one-month average of dollar trading volume. The strategy that buys the stocks with the lowest dollar volume and shorts the stocks with the highest dollar volume can produce significant monthly raw returns of 1.07%. The dollar trading volume based liquidity anomalies cannot be fully explained by the CAPM or the CH3 model, but can be explained by the CH4 model. The CAPM alpha is 1.08% with a t-statistic of 2.51, the CH3 alpha is 0.49% with a t-statistic of 2.24, and the CH4 alpha is 0.17% with a t-statistic of 0.89.

When we use variations and coefficient of variations of turnover and dollar volume to measure liquidity, the anomaly returns are also significant. For instance, for anomaly "vturn1-1", we long the stocks in the decile with the lowest variation of share turnover, and short the stocks in the decile with the highest variation of share turnover, the monthly long-short raw return is 1.58%, with a significant t-statistic of 3.45. After the CH3 model for risk adjustment, we find that the monthly abnormal return is 1.38%, with a significant t-statistic of 3.38. After regressing on CH4 factors, the alpha is 0.71% with the t-statistic of 1.82.

Overall, we find that liquidity anomalies are quite significant in China. It is likely driven by the fact that daily trading volume in China mostly comes from retail investors. Retail investors usually trade excessively to chase past good performance. Their excessive purchase towards past winners can lead to price bubbles in these winner stocks, which lead to future bad returns. Retail investors in the United States, on the contrary, mostly pursue contrarian strategy, as documented in Boehmer, Jones, Zhang and Zhang (2020). Meanwhile, the majority of trading in the U.S. stock market is driven by institutional investors, who are likely better informed than retail investors. As a result, the relation between liquidity and future stock returns is less significant.

Risk

The classic capital asset pricing theory of Sharpe (1964) and Lintner (1965) claims that a security's risk premium should depend on the security's market beta or other measure(s) of systematic risk. High risk should be related to high returns. Due to its importance in asset pricing, researchers have found several different methods to measure risk, and examined the relationship between risk and returns in the empirical data in the past 50 years. However, many empirical papers document the opposite findings to the theory, which means that high risk is associated with lower returns.

Our study includes 46 risk anomalies, which are constructed on volatility, skewness, beta, and tail risk measures. Altogether, 11 of them have significant long-short raw returns with t-statistics above 1.96, and their monthly raw returns and abnormal returns are reported in Table 4 Panel B.

Our first measure of risk is idiosyncratic volatility computed as the standard deviation of residuals from regressing a stock's daily excess returns on the CH3 factors over one month, or total volatility computed the standard deviation of a stock's daily returns over one month. According to the study of Ang, Hodrick, Xing and Zhang (2006) on the U.S. stocks, the stocks with high idiosyncratic volatility have lower returns. In the case of China, the long-short strategies for idiosyncratic volatility have raw returns ranging between 0.74% and 1.19% per month, and the t-statistic is always above 2. In another word, Chinese firms with higher volatility also underperform firms with lower volatility, which confirms the finding in Ang, Hodrick, Xing and Zhang (2006). However, when we use risk adjustment with the CAPM, the CH3 and CH4 models, the alphas are between -0.79% and 1.31%. The CAPM alpha is statistically significant, but the CH3 and CH4 alpha is not.

We also measure a stock's risk by using their skewness, as in Boyer, Mitton and Vorkink (2009) and Amaya, Christoffersen, Jacobs and Vasquez (2015). For each stock, we calculate its total skewness as the skewness of daily stock returns over one month and its idiosyncratic skewness as the skewness of the residuals from regressing a stock's daily excess returns on the CH3 factors over one month. In terms of raw returns, we find that only total skewness and idiosyncratic skewness have significant long-short returns (0.65% and 0.83% per month) at the 1-month horizon. The stocks with lowest skewness have higher returns, which is consistent with the finding of Boyer, Mitton and Vorkink (2009) in the United States. When we use the CAPM as risk adjustment, the alphas of total skewness and idiosyncratic skewness are 0.81% and 0.65% per month with the t-statistics above 2.7. When using the CH3 model, we find that the alphas of total skewness and idiosyncratic skewness are 0.67% with the t-statistic of 2.02 and 0.92% with the t-statistic of 2.81, respectively. After regressing on the CH4 model, the alpha of total skewness is 0.51% per month with the t-statistic of 1.30 and the alpha of idiosyncratic skewness is 0.88% with the t-statistic of 2.16.

Another set of risk measures are beta measures. We first compute two market beta measures based on rolling-window regressions, including the Frazzini and Pedersen (2014) method with daily stock returns over a one-year rolling window, and the Dimson (1979) method with the daily returns over a one-month rolling window. The long-short returns of the two beta measures are between 0.93% and 1.10 % per month, which are significant at the 1-month horizon. We notice that stocks with higher market beta have lower returns, except for the Dimson beta method. Similar to volatility patterns, the CH3 factors can explain these significant long-short returns of beta anomalies, while the CAPM cannot. The CH4 factor can explain the Frazzini and Pedersen (2014) beta but cannot explain the Dimson (1979) beta.

To summarize, 11 out of 46 anomalies have significant raw returns based on risk measures, but most of them are insignificant after the CH3 and CH4 model risk adjustment. This finding is consistent with the findings in Hou, Xue and Zhang (2019) for risk anomalies in the United States.

Past Returns

Past returns have been used widely in the previous literature as sorting variables for trading strategies. The most popular ones are the momentum strategy, first documented by Jegadeesh and Titman (1993), and contrarian strategy. We construct 52 past return anomaly variables using momentum strategy, contrarian strategy, seasonality etc., and 9 of them have significant long-short raw returns for single test with the t-statistic greater than 1.96. We report details on raw returns and abnormal returns of these 9 anomalies in Panel C.

Momentum strategy means buying past winners and selling past losers. Only a small number of momentum anomalies, including momentum change, residual momentum, and industry lead-lag momentum, have significant long-short portfolio raw returns. For momentum change, which is computed as cumulative daily returns from the most recent 6 months minus cumulative daily returns from the first 6 months in one year, the long-short return is 0.57% per month with the t-statistic of 2.79 at the 6-month horizon. The residual momentum, computed as the k-month (k=11) cumulative residuals of regressing monthly stock returns on the CH3 factors over the prior 36 months scaled by their standard deviation over the same period, has the long-short raw returns of 0.68% per month with the t-statistic above 2.70. The industry momentum, constructed as 1-month value-weighted returns of the portfolio consisting of the 30% biggest (total A share market capitalization) firms within an industry, has the long-short spread of 0.63% with the t-statistic of 2.00.

Altogether, we construct 30 momentum strategies, and 4 of them are statistically significant, corresponding to a 13.3% significance rate. In contrast, the significance rate of momentum anomalies in the United States by Hou, Xue and Zhang (2019) is 63.2%. That is, the momentum anomalies are relatively weak in China. Daniel and Moskowitz (2016) show that momentum crashes in panic states, following market declines and high market volatility. The majority of momentum does not work in China possibly due to a long period of bear market conditions since 2000. The weak momentum in China finding is consistent with the international evidence for momentum. Cakici, Chan and Kudret (2015) also fail to find that momentum is a useful predictor in the Chinese and other emerging stock markets.

The contrarian strategy means buying past loser stocks and selling past winner stocks, and is normally treated as the opposite of the momentum strategy. For China, the long-term reversal (constructed as the cumulative daily returns over four years) and the maximum daily return (constructed as the average of the five highest daily returns in one month) have significant long-short raw portfolio returns ranging between 0.81% to 0.89% per month with significant t-statistics above 2 at the 5% level. However, the abnormal returns become insignificant after we use the CH3 model for risk adjustment. The alpha of long-term reversal cannot be explained by the CH4 factors, but the alpha of maximum daily return can.

We also examine seasonality in China. The long-short raw returns of the average of 2-5 years' lagged monthly returns in the same months (*Ra25*) and the average of 2-5 years' lagged monthly returns in other months (*Rn25*) are 0.65% and 0.93% per month with the t-statistics of above 2.30 at the 1-month horizon, respectively. The CAPM cannot explain the raw returns of *Ra25* and *Rn25*, but the CH3 and CH4 factors can explain.

Overall, we find that 9 out of 52 return anomalies are significant, and 5 of them are significant after risk adjustment with the CH3 model. The percentage of significant momentum anomalies is quite low. Though there exist some contrarian strategies and seasonality effects in raw returns, they disappear after regressing on the CH3 factors.

Accounting-Related Anomalies

Profitability

Profitability is an important fundamental accounting variable, and is examined in many studies, including Fama and French (2006), Fama and French (2008), Fama and French (2015), and Novy-Marx (2013). For instance, Novy-Marx (2013) finds that gross profitability is strongly related to average future returns in the sense that stocks with high profitability should earn higher expected returns than stocks with low profitability. We measure profitability using proxies such as return on equity or assets, changes in return on equity or assets, return on operating assets, gross profits, and operating profits etc. Out of the 67 profitability-sorted long-short returns, 27 have significant long-short raw returns that pass the hurdle of 1.96, and are reported in Panel D of Table 4.

We first measure profitability with return on equity and return assets from Hou, Xue and Zhang (2015). Return on equity and return on assets have significant long-short portfolio spreads at the 1- and 6-month horizons. The raw returns are ranging from 0.83% to 1.14% per month with the t-statistics above 2.4. The CAPM alphas have similar magnitudes and significances. However, the CH3 and CH4 alphas are much smaller, ranging between -0.12% and 0.20% with the t-statistics below 0.. Similar patterns can be found for all other profitability proxies, such as changes in return on equity and assets, gross profits, operating profits, and F score, etc.

Overall, we find about 39% of profitability anomalies are significant for long-short raw portfolio returns in China, which is close to the 44.3% significance rate using the U.S. data, as documented in Hou, Xue and Zhang (2019). However, most of the alphas lose significances after we use the CH3 model as risk adjustment.

Value

One of the most popular studies on capital market phenomena is the relation between an asset's return and the ratio of its "long-run" value (or book value) relative to its current market value, or the value effect. Previous studies, such as Fama and French (1992), find that value stocks consistently outperform growth stocks on average.

We construct 44 value anomalies with the book-to-market ratio, earnings-to-price ratio, sales-to-price ratios, intangible return, enterprise book-to-price etc., and 14 of them have significant long-short raw returns, as reported in Panel E of Table 4.

We first follow Fama and French (1992) and measure value using the book-to-market ratio. For annual portfolio sorting, we long the stocks with the highest book-to-market ratio and short the stocks with the lowest book-to-market ratio at the end of each June, and hold the portfolios for one year. This strategy generates 0.88% per month with a t-statistic of 2.15. If we conduct monthly sorting using the latest information on the book-to-market ratio and hold it for 1-, 6- or 12-month horizons, the long-short raw returns range 0.81-1.18% per month with the t-statistics above 2, respectively. All of the CAPM alphas are significant at the 5% level, but all of the alphas relative to the CH3 and CH4 factors are not significant. The results are quite similar when we use the earnings-to-price ratio as the value measure, as in Liu, Stambaugh and Yuan (2019), or the sales to price ratio, as in Hsu, Viswanathan, Wang and Wool (2017).

Overall, there exists value effect in China, which is consistent with the findings of Hsu, Viswanathan, Wang and Wool (2017), Liu, Stambaugh and Yuan (2019), and Cakici, Tang and Yan (2016). When we compare the result with the U.S. market, the percentage of significant value anomalies using long-short returns in China is 27%, which is lower than 42% in the United States, as documented in Hou, Xue and Zhang (2019). But the value effect can be easily accounted for using the CH3 model, which might not be surprising, given that the CH3 model contains a value factor.

Investment

One of the primary functions of capital markets is to raise capital or make investments. Many papers, such as Cooper, Gulen and Schill (2008), show that corporate investment can affect asset prices. Due to its importance in asset pricing, Hou, Xue and Zhang (2015) propose the q-factor model by including investment growth as an important market factor based on the q theory of investment. All of these papers document a negative relationship between various forms of corporate investment and the cross-section of returns. That is, high investment firms (or high asset growth firms) tend to earn lower returns in the future. We construct 51 investment anomalies with investment, asset growth, accruals, etc., but only three have significant long-short raw returns. The detail for these three significant anomalies is reported in Table 4 Panel F.

For the composite equity issuance portfolios, we sort the portfolio at the end of June of each year, we find that the long-short portfolio return is 0.66% per month with the t-statistic of 2.15. The CAPM alpha is large and significant, but the CH3 and CH4 alphas are not. When we use change in noncurrent operating assets as sorting variables, we find significant long-short raw returns around 0.38% per month. The CAPM alpha is still significant, but neither the CH3 and CH4 alphas are significant.

Overall, there is almost no investment effect in China. In contrast, the replication rate of investment anomalies in the United States is as high as 73.7% in Hou, Xue and Zhang (2019).

The missing investment effect in China is consistent with the previous literature in China, such as Chen, Kim, Yao and Yu (2010), Hsu, Viswanathan, Wang and Wool (2017), Guo, Zhang, Zhang and Zhang (2017), Lin (2017), and Liu, Stambaugh and Yuan (2019). It is also consistent with the investment or asset growth effect in international markets by Titman, Wei and Xie (2013), and Watanabe, Xu, Yao and Yu (2013). Specifically, the latter studies find that the investment effect is stronger in countries with more developed or efficient markets than in countries with less developed or less efficient markets. According to earlier studies of Bekaert and Harvey (2002), Bhattacharya, Daouk, Jorgenson and Kehr (2000), and Griffin, Kelly and Nardari (2010), emerging markets tend to be inefficient, possibly due to higher transaction costs and information costs. In the case of China, with less than 30 years' history, higher transaction costs and information costs might lead to lower efficiencies in Chinese corporate managers' investment decisions, which leads to the insignificance of investment anomalies in China.

Others

Our 72 other accounting-related anomalies are mainly constructed using advertising expense, R&D, cash related, and analysts' forecasts. In Table 4 Panel G, we report the 24 anomalies with significant long-short returns at the 5% level.

We first construct anomalies using accounting information, such as advertising expense, the R&D expense-to-market ratio by Chan, Lakonishok and Sougiannis (2011), tangibility, cash to assets by Palazzo (2012), value relevance of earnings, asset liquidity, standardized unexpected earnings by Foster, Olsen and Shevlin (1984), and revenue surprise by Jegadeesh and Livnat (2006). The long-short portfolio returns range between 0.45% and 1.26% per month with t-statistics above 1.96. After risk adjustment with the CAPM, all of these anomalies still have significant CAPM alphas. However, the CH3 alphas only remain significant and large for anomalies related to the R&D expense-to-market ratio, tangibility, and asset liquidity, with magnitudes ranging from 0.65% to 0.82% per month. The CH4 alphas only remain significant for anomalies related to the R&D expense-to-market ratio and asset liquidity.

Next, we use analysts' earnings forecasts for forming anomalies. According to Elgers, Lo and Pfeiffer (2001), stocks with high analysts' earnings forecasts tend to have higher returns than the stocks with low earnings forecast. The long-short raw return of analysts' earnings forecasts-to-price is 1.39% per month with the t-statistic of 2.47 at the 1-month horizon. The long-short raw return of change in analysts' earnings forecasts is 0.32% per month with the t-statistic of 2.32 at the 12-month horizon. These raw returns cannot be explained by the CAPM, but they can be explained by the CH3 and CH4 factors. The insignificance of analyst-related anomalies in China is more likely due to low information quality of analysts' forecasting.

Overall, 30% of other accounting-related anomalies have significant long-short portfolio raw returns. It is almost consistent with the findings by Hou, Xue and Zhang (2019), who find that 25.2% of intangible anomalies are significant in the United States.

In total, we find that 98 of 426 (23%) anomalies are significant at the 5% level in the Chinese A-share stock market from 2000 to 2018. There are significant differences between our findings on Chinese anomalies and findings in Hou, Xue and Zhang (2019) for the anomalies in the United States. First, about one third of these significant anomalies are from the trading-related anomalies, especially from trading liquidity anomalies, while Hou, Xue and Zhang (2019) find that most (96%) of the trading frictions anomaly variables are insignificant at the 5% level in the United States. This is possibly because the daily trading volume in China mostly comes from individual investors, who chase past good performance and drive up trading liquidity, while this excessive trading volume leads to future low returns. The magnitudes of trading frictions in the United States are much smaller, possibly because the U.S. market is far more efficient than the Chinese market, with its longer history and dominance of institutional investors.

Second, for accounting information-based anomalies, many do not have significance in China, such as the investment anomalies. However, most of the significant anomalies in the United States are driven by firm fundamental information in accounting statements, such as value, investment, and profitability. This indicates that accounting information is not informative enough to reflect future price movement, possibly due to frauds or leakage in information disclosure or poor market efficiency. These results are, in general, consistent with the international evidence on anomalies, in the sense that anomaly patterns heavily depend on country-level environment.

3.1.3 Robustness

We provide two robustness checks in this section. First, we include the bottom 30% of stocks, in terms of market capitalization, in the sample, and check whether patterns we document earlier still hold. Previous studies exclude the small stocks due to concerns of illiquidity, while in China the small stocks also enjoy adequate liquidity, and are important components of the market. Second, we examine the subperiod between 2007 and 2018, after the split-share structure reform and the adoption of the global accounting standards, to examine whether our results are robust for this more recent period, while the information environment is presumably improved.

Including Smaller Stocks

Our earlier results exclude the smallest 30% of all A-share stocks, which account for 7% of the stock market's total capitalization, to minimize the impact of shell contamination. In previous studies, such as Hou, Xue and Zhang (2019), the microcap stocks, with only 3% of the aggregate market capitalization of the NYSE-Amex-NASDAQ universe but account for 60% of the total number of stocks (Fama and French, 2008), are excluded because they do not have enough liquidity for trading. Unlike the United States, the smallest 30% of stocks in Chinese market actually have adequate liquidity, due to the bounty trading of retail investors. For instance, the average annual turnover is 5 for these microcap stocks, while it is only about 2.3 for the U.S. microcap (NASDQ) stocks. When we include the smaller stocks in the sample, we expect the number of significant anomalies to be higher.

Results on the anomalies using all A-share stocks are presented in Table 5 Panel A. Out of the 426 anomalies, 95 anomalies (22.3%) are significant with t-statistics higher than 1.96, which is similar to what we find in Table 3 Panel A. Among these significant anomalies, about 17% is related to liquidity proxies, 22% is related to other accounting information, and 11 are profitability anomalies. Using the CAPM as risk adjustment does not reduce the number of significant long-short returns. In fact, the number of significant returns is 94, indicating that these long-short returns cannot be explained away by market risk exposures. When we move on to the CH3 model as risk adjustment, the number of significant long-short returns become 33, with 24 of them using liquidity proxies, again indicating the importance of trading environment. That is, when including the small firms in the sample, the liquidity anomalies are more significant for the cross-sectional patterns of stock returns. With the CH4 model as risk adjustment, there are 38 significant alphas with a half of them from liquidity. For accounting information-related anomalies, the number of significant cases using CH3 actually reduces from 5 in Table 3 to 3 in Table 5. The number of significant CH4 alphas is 10. One potential explanation for these findings is that the retail investors might be more speculative towards smaller stocks, and thus the bubble from excessive trading could be larger for these small stocks. Meanwhile, the information environment is normally less transparent and less efficient for these smaller stocks, and accounting information might contain more errors to be predictive for future stock returns.

[Insert Table 5 here]

For multiple tests, we find that there are 26 anomalies passing the hurdle of 2.85. Among 26 significant anomalies, 6, 5, 1, 5, 2, 7 are from liquidity, risk, past returns, profitability, value, and other anomalies. The CAPM cannot explain any anomalies. After regressing on the CH3 factors, there are 5 significant anomalies with 4 from the liquidity group. Similarly, with the CH4 model as risk adjustment, there are only 3 significant anomalies with all of them from liquidity.

Subsample between 2007 and 2018

The Chinese stock market experiences many changes in information structure and regulations over the whole sample period 2000-2018. For instance, the accounting standards changed a few times over the past 30 years. After 2007, the Chinese accounting system is set to the global accounting standards, which potentially improve the quality of accounting data. Another important regulation change in the Chinese stock market is the split share reform, which started in 2005 and was mostly finished by the end of 2006.[2] The split-share structure reform solves the conflicts between different shareholders, and likely improves the market quality and efficiency. Combining the changes of the accounting standards and the split-share structure reform, we examine whether our earlier results hold in this subsample between 2007 and 2018, when the information environment gradually improves. We expect that the anomalies based on trading information to be weaker, while anomalies based on accounting information to be more significant than the results over 2000-2018. But one potential drawback of this approach is that with only 12 years of monthly data, the estimation of the standard errors can be noisy and may reduce significances of these anomalies.

In Table 5 Panel B, we report the number of significant anomalies for the A-share sample without the 30% of small firms. For long-short portfolios, 68 out of 426 anomalies (16%) are significant with t-statistics higher than 1.96. The percentage of significant anomalies is significantly lower than those (98) in the full sample period from 2000 to 2018. It is possible that the shorter sample period has less precise estimate of volatility, which reduces the t-statistics, and it is also possible that both trading environment and information environment improve and thus change the patterns of anomalies. Among these 68 significant anomalies during 2007-2018, 19, 8,

[2] Chinese stocks have a unique split-share structure, and with free floating and non-floating shares. The Chinese government and semi-government entities hold equity shares of the state-owned enterprises (SOEs), which are in the form of non-floating shares, while the floating shares are issued to the general public, and are listed and traded on exchanges. In order to meet the needs of state-owned enterprises for funds, liquidity, better governance, and re-organization, the China Securities Regulatory Commission (CSRC) launched the split-share structure reform on April 29, 2005. Through terms negotiated with the owners of floating shares, the non-floating shares were gradually converted into floating shares.

6, 15, 1, 6, and 13 anomalies are from the liquidity, risk, past returns, profitability, value, investment, and other accounting-related groups, respectively. For trading-related anomalies, more liquidity anomalies are significant than the results in Table 3, while accounting-related anomalies seem less significant in this subsample.

The alphas from the CH3 model tell a different story. Altogether, there are 12 anomalies with significant CH3 alphas, with 8 of them related to trading frictions and 4 related to accounting information. The change in composition is quite interesting. In Table 3, there are 12 trading anomalies and four accounting anomalies with significant CH3 alphas, over 2000 to 2018. For the abnormal return relative to the CH4 model, there are 7 significant anomalies with 4 from trading-related anomalies and 3 from accounting-related anomalies. That is, the trading anomalies are slightly less significant over the later sample period of 2007-2018, possibly because trading environment improves over the recent period. Meanwhile, the accounting information-related anomalies become more significant over this recent period, implying that the information environment might improve to the point that anomalies constructed on accounting information become more relevant.

When we adopt the multiple test hurdle of 2.85, there are only 3 significant anomalies with 2 of them from liquidity. When using the CH3 and CH4 models as risk adjustment, there are 1 and 0 significant abnormal returns, which is from liquidity.

3.2 Composite Anomalies

In Section 3.1, we find that only 27 anomalies can pass the multiple test hurdle of 2.85. After risk adjustment with the CH4 model, almost no anomalies can pass the multiple hurdle. We then adopt four methods to construct composite anomalies, which include the score, Fama-MacBeth regression, Lasso and Random Forest methods. The first two methods are traditional ones, while the last two belongs to the popular machine learning. We construct composite anomalies by combining characteristics in each of 2 categories (trading-related and accounting-related anomalies), 7 subcategories (liquidity, risk, return, profitability, value, investment, and others) and among all anomalies. In total, we have 10 composite anomalies.

3.2.1 Score

The score method is a traditional method to construct composite anomalies, which is used in Hou, Xue and Zhang (2019). For a given set of anomalies, we construct its composite score for a stock by equal-weighting the stock's percentile rankings for the significant anomalies in each group at the end of each month. If the stock falls in the long portfolio, the ranking is 10, and in the short portfolio, the ranking is 1. If it falls in other groups, we have the ranking of the group (e.g. 2, 3, … 9). We then take a simple average of these ranking scores across all anomalies in each group to obtain the composite score. At the end of each month, we sort all A-share sample but 30% microcap firms into 10 groups based on the composite score, we then calculate value-weighted portfolio returns in the next month.

Panel A of Table 6 shows the long-short returns and the alphas relative to the CAPM, CH3 factors, and CH4 factors. In each category, all long-short returns are significant at the 5% level with single test. 9 of 10 composite anomalies can pass the multiple test hurdle of 2.85. The long-short returns range between 0.83% and 2.48% per month with the t-statistics ranging between 2.05 and 6.69. The composite anomaly with all characteristics ("total") provides the highest long-short returns, while the composite anomaly with all characteristics in the investment sub-category provides the lowest long-short return. All of the composite anomalies cannot be explained by the CAPM. With the CH3 model as risk adjustment, 5 composite anomalies are significant at the 5% level for single test, and 4 can pass the multiple hurdle of 2.85, These significant anomalies are mainly from trading-related anomalies. When regressing on the CH4 factors, the alphas range from 0.20% to 1.12% per month. We find that 3 composite anomalies have significant alphas for single test, and no one has significant alphas for multiple test.

[Insert Table 6 here]

3.2.2 Fama-MacBeth regressions

Lewellen (2015) use Fama-MacBeth (FM) regressions to combine many firm characteristics to obtain a composite estimate of a stock's expected return. There are four steps. First, we estimate the cross-sectional regression of monthly stock returns on firm characteristics for each period t.

$$R_t^i = a + b * X_{t-1}^i + \varepsilon_t^i \tag{2}$$

where X_{t-1}^i denotes a set of firm characteristics in month t-1 for firm *i*, R_t^i is stock return in month t for firm *i*. Second, we calculate 3-year (36-month) rolling-window average of parameters ($\hat{a}$ and $\hat{b}$) in the regression, starting from 1997. Third, we estimate expected stock returns in the next month.

$$\hat{R}_{t+1}^i = \hat{a} + \hat{b} * X_t^i \tag{3}$$

Fourth, we do portfolio sorting on predicted returns $\hat{R}_{t+1}^i$ at the end of month t, and calculate value-weighted portfolio returns in the next month t+1.

Panel B of Table 6 reports the composite anomalies with the Lewellen (2015)' FM method. We find that 7 composite anomalies have significant long-short returns for single test, and 6 can pass the multiple hurdle of 2.85. The long-short returns range 0.11-1.51% per month with the highest return in trading-related anomalies. The CAPM still cannot explain these anomalies. Only 2 composite anomalies with all trading-related anomalies and other anomalies have significant CH4 alphas at the 5% level for single test. Compared with the score method, the FM method performs poorly.

3.2.3 Lasso

Machine learning is current popular technique used in finance area. Gu, Kelly, and Xiu (2020) use machine learning methods to predict the cross-sectional returns in the U.S stock market. We employ the stable method of Lasso. The expected stock return is:

$$\mathrm{E_t}\left(\mathrm{r_{i,t+1}}\right) = g^{*}(z_{i,t}) \tag{4}$$

The least squares (or loss function) of stock returns is:

$$\mathrm{L}(\theta) = \frac{1}{\mathrm{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\mathrm{r_{i,t+1}} - g(z_{i,t};\theta)\right)^{2} \tag{5}$$

The Lasso loss function is

$$\mathrm{L}(\theta;.) = \mathrm{L}(\theta) + \lambda\sum_{j=1}^{P}|\theta_j| \tag{6}$$

We use both least squares and Huber objective function to minimize equation (6).

Panel C reports the composite anomalies with the Lasso method with the long-short returns ranging between 0.49% and 2.74% per month. We find that 9 composite anomalies have significant long-short return at the 5% level for single test and 7 of them can pass the multiple test hurdle of 2.85. The CAPM still cannot explain these anomalies. With the risk adjustment of CH3 model, the CH3 alphas of 5 composite anomalies are significant for both single test and multiple test, which are trading-related, liquidity, accounting-related, others, and total anomalies. With risk adjustment relative to the CH4 model, these 5 anomalies still have significant alphas for single test and 4 anomalies have significant alphas for multiple test.

3.2.4 Random Forest

Random forest is a supervised learning algorithm which is used for both classification as well as regression. But however, it is mainly used for classification problems. As we know that a forest is made up of trees and more trees means more robust forest. Similarly, random forest algorithm creates decision trees on data samples and then gets the prediction from each of them and finally selects the best solution by means of voting. It is an ensemble method which is better than a single decision tree because it reduces the over-fitting by averaging the result. First, it starts with the selection of random samples from a given dataset. Second, this algorithm will construct a decision tree for every sample. Then it will get the prediction result from every decision tree. Third, voting will be performed for every predicted result. Finally, it selects the most voted prediction result as the final prediction result.

Panel D presents the composite anomalies with random forest. The long-short raw returns range between 0.57% and 3.32% per month. The t-statistics range between 1.94 and 8.58. All composite anomalies except for investment have significant long-short returns at the 5% level for multiple test. After risk adjustment, all CAPM alphas are still significant for single test. The CH3 alphas of 6 anomalies are significant for single test. After regressing on CH4 factors, 6 composite anomalies have significant alphas for single test, and 3 have significant alphas for multiple test.

Overall, we find that the composite anomalies perform better than the individual anomalies. The composite anomalies can provide higher long-short returns and t-statistics. The majority of the most prominent anomalies are related to trading information. Among these 4 composite methods, the random forest performs better than the other three methods.

4. Conclusion

In this study, we document the returns patterns of 426 anomalies in China. We find that 98 (23%) are significant at the 5% level in all A-share sample without the 30% of small firms from 2000 to 2018. Among these, 11 liquidity anomalies and 27 profitability anomalies have significant long-short returns. After risk adjustment using the CAPM or the CH3 model, most of these liquidity anomalies still have significant alphas, but the profitability anomalies become insignificant. With risk adjustment with the CH4 model, the liquidity anomalies become insignificant. The liquidity anomalies are robust to small stocks and subperiods, while the profitability anomalies are not. Over the whole sample period, liquidity is important in Chinese stock market, possibly due to the dominance of retail investors in trading. The accounting anomalies are less significant, possibly because of low quality of accounting information.

We also propose that the multiple test hurdle should be 2.85 in China. We find that only 27 anomalies can pass the hurdle. After risk adjustment with the CH3 or CH4 model, almost no anomalies can pass the hurdle. We construct the composite anomalies, and some can pass the hurdle, even after risk adjustment with the CH3 and CH4 factors.

References

Amaya, Diego, Peter Christoffersen, Kris Jacobs, and Aurelio Vasquez, 2015, Does realized skewness predict the cross-section of equity returns?, *Journal of Financial Economics* 118, 135-167.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223-249.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773-806.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785-818.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009, Do retail trades move markets?, *Review of Financial Studies* 22, 151-186.

Bekaert, Geert, and Campbell R. Harvey, 2002, Research in emerging markets finance: Looking to the future, *Emerging Makrets Review* 3, 429-448.

Bhattacharya, Uptal, Hazem Daouk, Brian Jorgenson, and Carl-Heinrich Kehr, 2000, When an event is not an event: The curious case of an emerging market, *Journal of Financial Economics* 55, 69-101.

Boyer, Brian, Todd Mitton, and Keith Vorkink, 2009, Expected idiosyncratic skewness, *Review of Financial Studies* 23, 169-202.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265-296.

Cakici, Nusret, Kalok Chan, and Topyan. Kudret, 2015, Cross-sectional stock return predictability in china, *working paper*.

Cakici, Nusret, Yi Tang, and An Yan, 2016, Do the size, value, and momentum factors drive stock returns in emerging markets?, *Journal of International Money and Finance* 69, 179-204.

Carpenter, Jennifer N., Fangzhou Lu, and Robert F. Whitelaw, 2018, The real value of china's stock market, *working paper*.

Chan, Louis K. C., Josef Lakonishok, and Theodore Sougiannis, 2011, The stock market valuation of research and development expenditures, *Journal of Finance* 56, 2431-2456.

Chen, Xuanjuan, Kenneth A. Kim, Tong Yao, and Tong Yu, 2010, On the predictability of chinese stock returns, *Pacific-Basin Finance Journal* 18, 403-425.

Cohen, Randolph B., Paul A. Gompers, and Tuomo Vuolteenaho, 2002, Who underreacts to cashflow news? Evidence from trading between individuals and institutions, *Journal of Financial Economics* 66, 409-462.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609-1651.

Daniel, Kent, and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221-247.

Datar, Vinay T., Narayan Y. Naik, and Robert Radcliffe, 1998, Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203-219.

Dimson, Elroy, 1979, Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7, 197-226.

Elgers, Pieter T., May H. Lo, and Ray J. Pfeiffer, 2001, Delayed security price adjustments to financial analysts' forecasts of

annual earnings, *Accounting Review* 76, 613-32.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, Eugene F., and Kenneth R. French, 2006, Profitability, investment, and average returns, *Journal of Financial Economics* 82, 491-518.

Fama, Eugene F., and Kenneth R. French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653-1678.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Foster, George, Chris Olsen, and Terry Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, *Accounting Review* 59, 574-603.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1-25.

Griffin, John M., Patrick J. Kelly, and Federico Nardari, 2010, Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets, *Review of Financial Studies* 23, 3225-3277.

Guo, Bin, Wei Zhang, Yongjie Zhang, and Han Zhang, 2017, The five-actor asset pricing model tests for the chinese stock market, *Pacific-Basin Finance Journal* 43, 84-106.

Hou, Kewei, Chen Xue, and Lu Zhang, 2019, Replicating anomalies, *Review of Financial Studies* forthcoming.

Hsu, Jason, Vivek Viswanathan, Michael Wang, and Phillip Wool, 2017, Anomalies in chinese a-shares, *working paper*.

Jacobs, Heiko, and Sebastian Muller, 2019, Anomalies across the globe: Once public, no longer existent?, *Journal of Financial Economics* forthcoming.

Jegadeesh, Narasimhan, and Joshua Livnat, 2006, Revenue surprises and stock returns, *Journal of Accounting and Economics* 41, 147-171.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jones, Charles M., Donghui Shi, Xiaoyan Zhang, and Xinran Zhang, 2020, The behavior and performance of chinese retail investors, *working paper*.

Lin, Qi, 2017, Noisy prices and the fama-french five-factor asset pricing model in china, *Emerging Makrets Review* 31, 141-163.

Lintner, John, 1965, The valuation of risk assets and the selection of risky
investments in stock portfolios and capital budgets, *The Review of Economics and Statistics* 47, 13-37.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan, 2019, Size and value in china, *Journal of Financial Economics* 134, 48-69.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887.

Newey, Whitney K., and Kenneth D. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1-28.

Palazzo, Berardino, 2012, Cash holdings, risk, and expected returns, *Journal of Financial Economics* 104, 162-185.

Sharpe, William F. , 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Titman, Sheridan, John K. C. Wei, and Feixue Xie, 2013, Market development and the asset growth effect: International evidence, *Journal of Financial and Quantitative Analysis* 48, 140-1432.

Watanabe, Akiko, Yan Xu, Tong Yao, and Tong Yu, 2013, The asset growth effect: Insights from international equity markets, *Journal of Financial Economics* 108, 529-563.

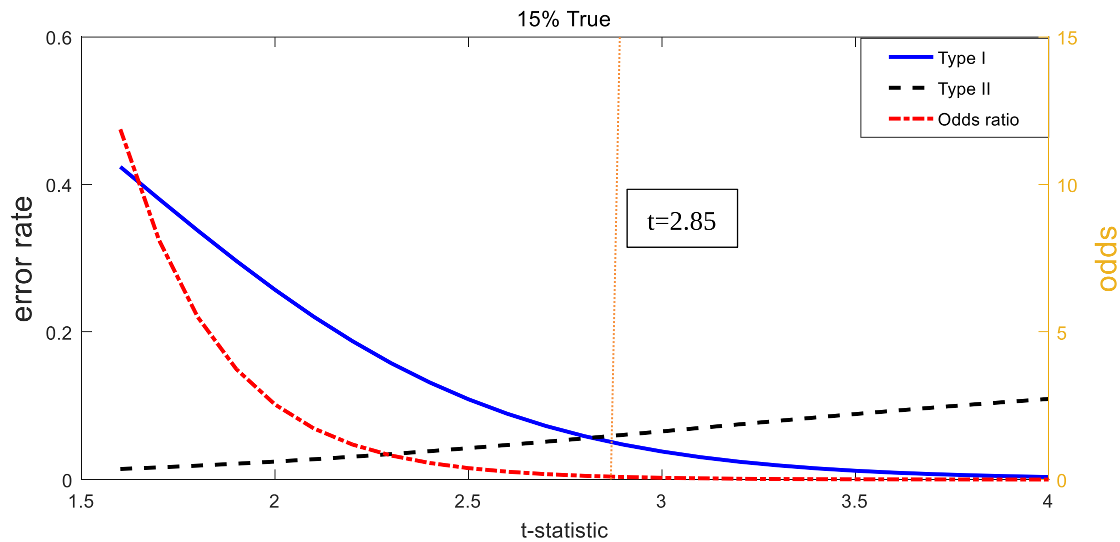


Figure 1. Error rates for fixed t-statistics threshold: 426 anomalies

This figure shows simulated Type I and Type II error rates for 426 anomalies. For 15% of strategies that are believed to be true, we follow the method of Harvey and Liu (2020) and set I=100 (for each i, we bootstrap to obtain the ranking of strategies and set the top p_0 as true) and J=1, 000 (conditional on i, for each j, we bootstrap the time periods) to run 100,000 (=100×1,000) bootstrapped simulations to calculate the empirical Type I (fraction of false discoveries) and Type II (fraction of missed discoveries among all non discoveries) error rates. On the graph, the (blue) solid line is the Type I error rate, the (black) dashed line is the Type II error rate, and the (red) dotted line is the odds ratio. The left axis is the error rate that applies to the Type I and Type II error rates, whereas the right axis is the odds ratio. The vertical (orange) line marks the t-statistic cutoff that corresponds to a Type I error rate of 5%.

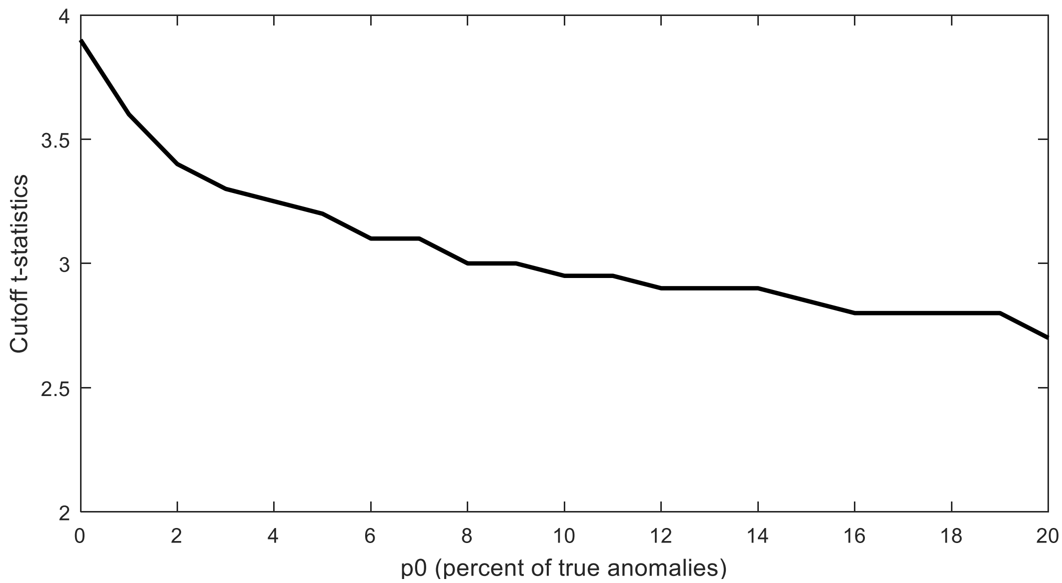
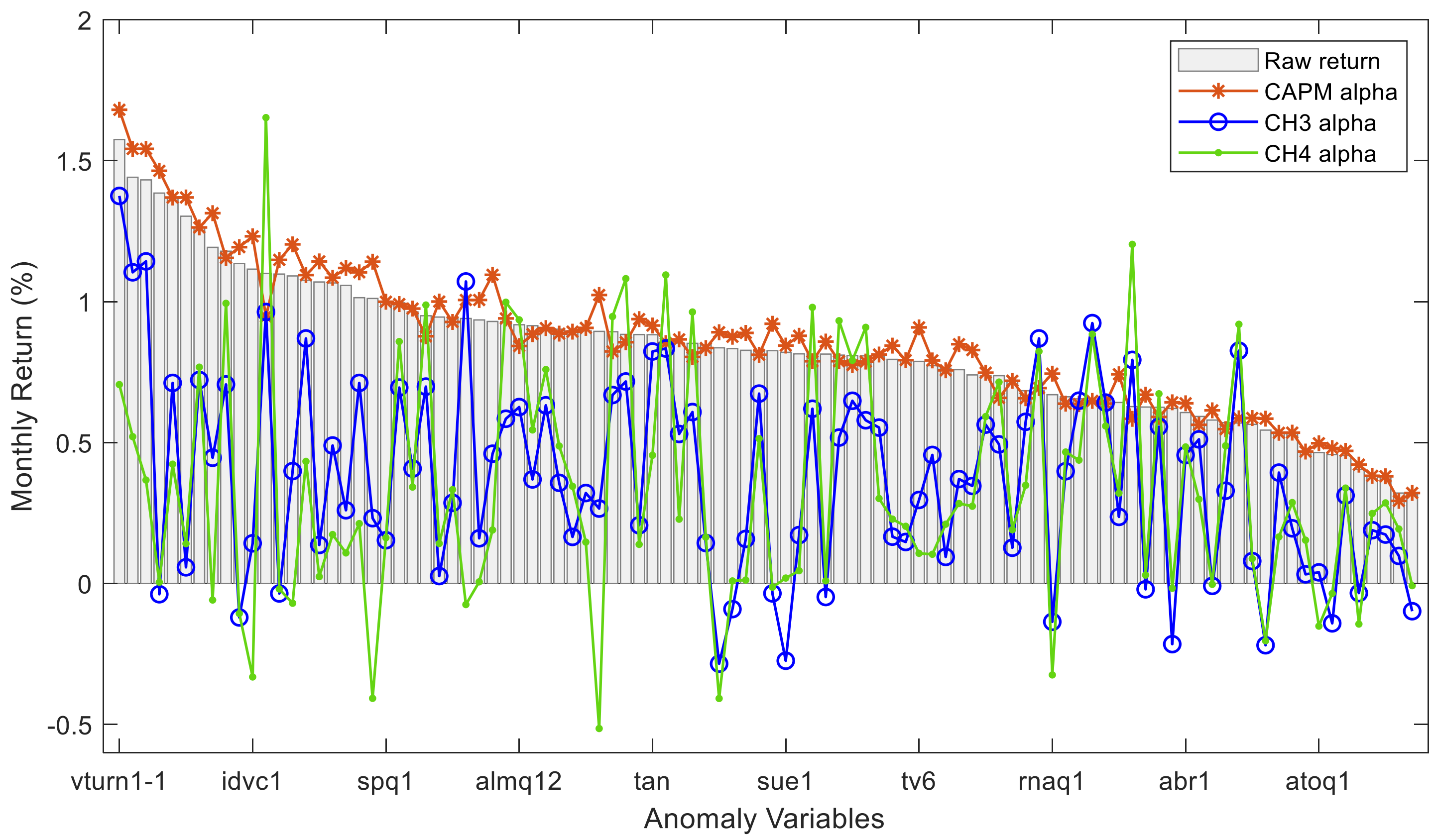


Figure 2. Cutoff t-statistics as a function of p_0

This figure shows cutoff t-statistics as a function of p_0 for 426 anomalies. For each hypothetical level of p_0 between 0% and 20% (with 1% increments), we search for the optimal cutoff t-statistic between 2 and 4 (with 0.1 increments) that corresponds to a Type I error rate of 5%. The kinks in the graph are caused by the discrete increments in both p_0 and the range of t-statistics from which we search.

Panel A. The Ranked Returns for 98 Significant Anomalies With Single Test



Panel B. The Ranked Returns for 27 Significant Anomalies With Multiple Test

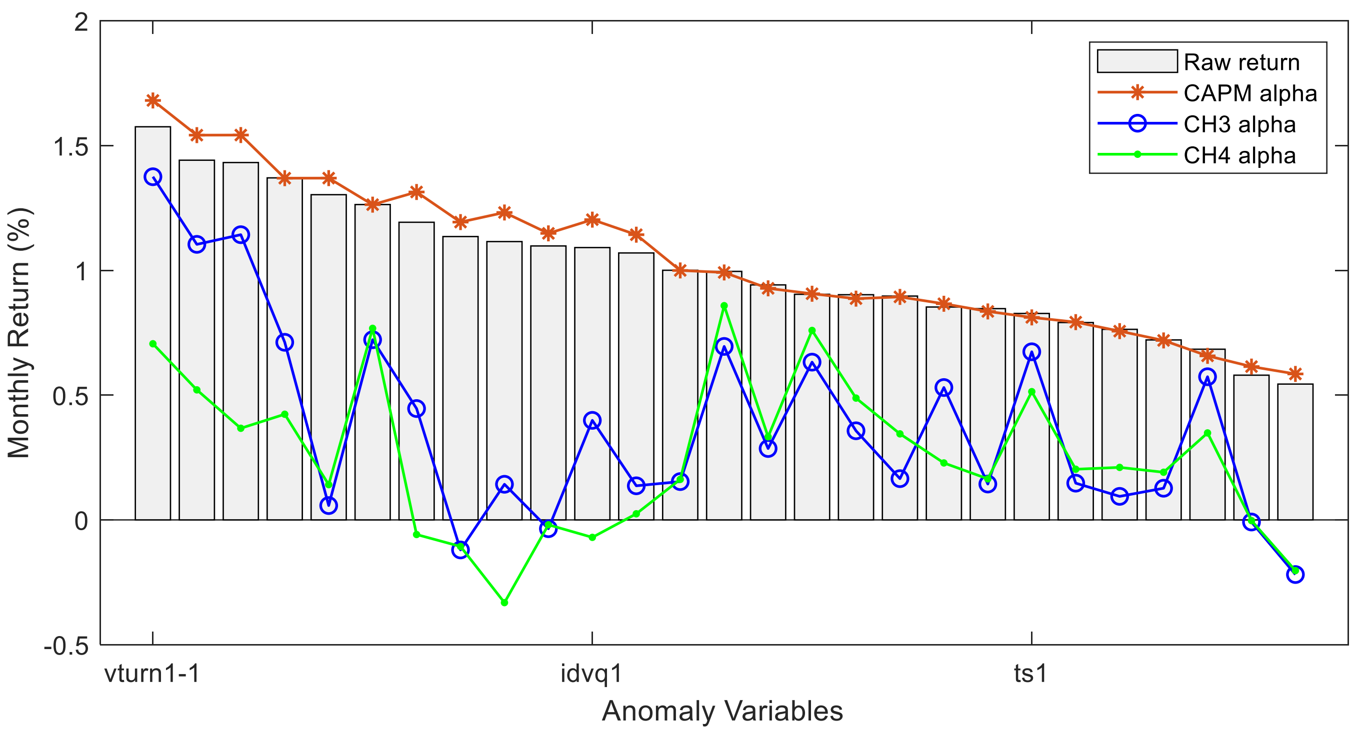


Figure 3. The Ranked Returns and Alphas for Long-Short Portfolios of Significant Anomalies

This figure plots the monthly value-weighted long-short raw returns, alphas relative to the CAPM, the CH3, and CH4 factors for significant anomalies with t-statistics for raw returns greater than 1.96 and 2.85. The long-short raw returns are ranked in a descending order and denoted in monthly percentage. The portfolio analysis is among all A-share sample but 30% microcap firms from July 2000 to December 2018.

Table 1. Summary on Chinese A-Share Stock Market

This table provides the summary of overall Chinese A-share stock market and compares with the U.S. market. It reports the number of firms, total A-share market capitalization, average firm size, annualized average stock returns, annualized average stock volatility, average stock share turnover, and average institutional ownership (IO) after data filters by year from 2000 to 2018.

Year	NO. of Firms	Total Market Cap (Billion RMB)	US Total Cap (Billion Dollar)	Average Firm size (million RMB)	US Average Firm size (Million Dollar)	Average Stock Return (Annual)	US Stock Return (Annual)	Average Stock Volatility (Annual)	US Stock Volatility (Annual)	Average Stock Turnover (Annual)	US Stock Turnover (Annual)	Average Stock IO	US Stock IO
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
2000	681	1,255	16,094	1,843	1,979	0.62	-0.08	0.42	0.81	4.73	1.49	0.02	0.30
2001	767	1,119	14,205	1,459	1,910	-0.22	0.14	0.32	0.72	2.04	1.23	0.01	0.32
2002	808	997	11,285	1,234	1,608	-0.17	-0.11	0.35	0.65	1.90	1.29	0.02	0.35
2003	855	1,079	14,978	1,262	2,239	-0.04	0.72	0.30	0.49	2.07	1.46	0.04	0.38
2004	917	996	16,924	1,086	2,514	-0.12	0.19	0.36	0.41	2.88	1.63	0.04	0.44
2005	947	1,326	17,976	1,400	2,664	-0.14	0.05	0.41	0.38	3.44	1.62	0.05	0.45
2006	931	5,379	20,388	5,777	2,976	0.76	0.17	0.48	0.37	5.62	1.78	0.08	0.46
2007	995	23,734	21,135	23,853	3,019	2.04	-0.02	0.63	0.40	9.34	2.26	0.10	0.48
2008	1,088	11,501	12,639	10,571	1,862	-0.58	-0.42	0.69	0.79	5.04	2.86	0.11	0.47
2009	1,100	22,132	16,598	20,120	2,532	1.37	0.56	0.52	0.69	8.66	2.88	0.09	0.48
2010	1,286	22,932	19,336	17,832	2,931	0.12	0.23	0.44	0.42	5.83	3.02	0.09	0.47
2011	1,527	19,816	18,672	12,977	2,793	-0.29	-0.09	0.38	0.45	3.71	3.22	0.08	0.49
2012	1,683	21,596	21,259	12,832	3,200	0.04	0.15	0.38	0.36	3.69	2.70	0.07	0.50
2013	1,723	22,266	27,371	12,923	4,032	0.24	0.32	0.41	0.33	4.91	2.89	0.08	0.52
2014	1,764	34,731	30,033	19,689	4,246	0.45	0.02	0.39	0.33	4.91	2.73	0.07	0.54
2015	1,943	48,214	28,702	24,814	3,952	0.65	-0.07	0.68	0.38	8.17	2.80	0.08	0.56
2016	2,005	44,573	31,221	22,231	4,347	-0.12	0.14	0.45	0.40	5.37	2.89	0.08	0.56
2017	2,290	51,930	37,341	22,677	5,118	-0.05	0.16	0.31	0.32	3.74	2.54	0.07	0.55
2018	2,465	40,534	33,979	16,444	4,532	-0.30	-0.13	0.40	0.37	3.80	2.88	0.06	0.51
Average	1,357	19,795	21,586	12,159	3,077	0.22	0.10	0.44	0.48	4.73	2.33	0.07	0.47

Table 2. Categories of Anomalies

This table lists the categories of 426 anomalies. The anomalies are categorized into two groups, which are (1) trading-related anomalies, (2) accounting-related anomalies. Online Appendix B shows the detail of defining these anomaly variables.

Trading-related anomalies	Liquidity	Size, turnover, dollar trading volume, Amihud illiquidity, turnover-adjusted liquidity, liquidity betas
	Risk	Market beta, downside beta, Frazzini-Pedersen beta, Dimson beta, idiosyncratic volatility, total volatility, idiosyncratic skewness, total skewness, co-skewness, tail risk
	Past Returns	Momentum, industry momentum, maximum daily return, short/long-term reversal, seasonality, 52-week high, share price, cumulative abnormal returns
Accounting-related anomalies	Profitability	Return on equity/assets/operating assets, change in return on equity/assets, profit margin, assets/capital turnover, operating/gross profits to book equity (assets), taxable income to book income, book leverage, sales growth, fundamental score, Ohlson's O-score, Altman's Z-score
	Value	Book/assets/debt-to-market-ratio, earnings/cash flow/sales-to-price ratio, liabilities-to-book equity ratio, sales growth (rank), enterprise multiple, intangible return
	Investment	Corporate investment, asset/inventory/investment/other asset growth, net operating assets, stock/debt issuance, accruals, changes in inventory/PPE and inventory-to-assets/net noncash working capital/different operating assets (liabilities)/net financial assets/investments
	Others	Advertising expense ratio, R&D expense ratio, operating leverage, firm age, effective tax rate, labor force efficiency, tangibility, secured debt-to-total debt, cash flow volatility, cash to assets, earnings-related, asset liquidity, standard unexpected earnings, revenue/tax surprise, % change in sales (gross margin) minus % change in inventory (accounts receivables, SG&A growth, sales), analyst earnings forecasts, analyst coverage, dispersion in analyst forecasts

Table 3. The Number of Significant Anomalies

The table reports the number of significant anomalies for long-short raw returns and returns relative to different factor models in each category and in total at the 5% level. The anomalies are categorized into 2 groups, which are trading-related anomalies and accounting-related anomalies. We adopt single portfolio analysis to test whether these anomalies are significant with all A-share but 30% of microcap firms from July 2000 to December 2018. The table reports the number of anomalies that are significant with the *t*-cutoff of 1.96 and 2.85 for raw returns, alphas relative to the CAPM, the CH3, and CH4 factors. Panel A shows the result for value-weighted portfolios, and Panel B provides the result for equally-weighted portfolios.

Panel A. Value-Weighted Portfolios

	Trading-Related Anomalies			Accounting-Related Anomalies				
	liquidity	risk	past returns	profitability	value	investment	Others	Total
	94	46	52	67	44	51	72	426
Single Test								
Raw returns	11	11	9	27	14	2	24	98
CAPM alphas	11	11	8	27	14	2	24	97
CH3 alphas	7	2	3	0	0	0	4	16
CH4 alphas	1	2	4	0	1	0	7	15
Multiple Test								
Raw return	4	4	0	10	4	0	5	27
CAPM alpha	4	4	0	10	4	0	5	27
CH3 alpha	2	0	0	0	0	0	0	2
CH4 alpha	0	0	0	0	0	0	0	0

Panel B. Equally-Weighted Portfolios

	Trading-Related Anomalies			Accounting-Related Anomalies				
	liquidity	risk	past returns	profitability	value	investment	others	Total
	94	46	52	67	44	51	72	426
Single Test								
Raw returns	42	26	14	37	21	7	28	175
CAPM alphas	42	26	14	37	21	7	28	175
CH3 alphas	30	11	6	4	1	0	7	59
CH4 alphas	18	12	8	5	3	0	8	54
Multiple Test								
Raw return	25	13	6	25	7	2	16	94
CAPM alpha	25	13	6	25	7	2	16	94
CH3 alpha	12	8	1	1	0	0	2	24
CH4 alpha	5	6	1	0	0	0	3	15

Table 4. Significant Anomalies Among All A-Share but 30% Microcap Firms

This table provides monthly value-weighted long-short raw returns, alphas in percentages relative to the CAPM and the CH3 factors for 98 significant anomalies with single test. These significant anomalies are with the t-statistics for long-short raw returns greater than 1.96. Single portfolio analysis is among all A-share sample but 30% microcap firms from July 2000 to December 2018. It also provides the Newey-West *t*-statistics with 4 lags adjusted for heteroscedasticity and autocorrelation. The symbols and definitions of these anomalies are described in Online Appendix B. * denotes that it passes the multiple test hurdle 2.85.

Panel A. Liquidity anomalies

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
turn: share turnover, k=1, m=0, n=1, Datar, et al. (1998)								
turn1-1	1.43	3.17*	1.54	3.39*	1.14	2.63	0.37	0.92
vturn: variation of share turnover, k=1, 6, m=0, n=1, Chordia, et al. (2001)								
vturn1-1	1.58	3.45*	1.68	3.77*	1.38	3.38*	0.71	1.82
vturn6-1	1.01	2.20	1.10	2.31	0.71	1.75	0.21	0.52
abturn: abnormal turnover, k=12, m=0, n=1, Liu, et al. (2019)								
abturn1	0.94	2.42	1.01	2.78	1.07	2.49	-0.07	-0.26
dtv: dollar trading volume, k=1,m=0, n=1, Brennan, et al. (1998)								
dtv1-1	1.07	2.41	1.08	2.51	0.49	2.24	0.17	0.89
vdtv: variation of share dollar trading volume, k=1, 6, 12, m=0, n=1, 6, Chordia, et al. (2001)								
vdtv1-1	1.37	3.16*	1.37	3.28*	0.71	3.39*	0.42	1.85
vdtv6-1	0.90	2.05	0.91	2.14	0.32	1.57	0.15	0.76
vdtv12-1	0.96	2.30	0.98	2.41	0.41	2.16	0.34	1.97
cvdtv: coefficient of variation of dollar trading volume, k=6, 12, m=0, n=1, 6, Chordia, et al. (2001)								
cvdtv6-6	0.54	2.01	0.53	2.02	0.39	1.45	0.17	0.65
cvdtv12-1	1.08	2.75	1.09	2.79	0.87	1.80	0.43	1.04
Lm: turnover-adjusted number of zero daily trading volume, k=1, m=0, n=1, Liu (2006)								
Lm1-1	1.44	3.51*	1.54	3.75*	1.10	2.45	0.52	1.19

Panel B. Risk anomalies

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
idvc: idiosyncratic volatility per the CAPM, k=1, m=0, n=1, Ali, et al. (2003)								
idvc1	1.12	2.85*	1.23	3.07*	0.14	0.33	-0.33	-0.75
idvff: idiosyncratic volatility per the CH3 factor model, k=1, m=0, n=1, 6, Ang, et al. (2006)								
idvff1	1.19	3.07*	1.31	3.37*	0.45	1.03	-0.06	-0.14
idvff6	0.74	2.07	0.83	2.37	0.35	1.18	0.27	0.67
idvq: idiosyncratic volatility per the q-factor model, k=1, m=0, n=1, 6								
idvq1	1.09	2.99*	1.20	3.41*	0.40	0.94	-0.07	-0.16
idvq6	0.76	2.12	0.85	2.43	0.37	1.31	0.28	0.73
tv: total volatility, k=1, m=0, n=1, 6, 12, Ang, et al. (2006)								
tv1	1.01	2.35	1.14	2.57	0.23	0.41	-0.41	-0.79
tv6	0.79	2.15	0.91	2.50	0.30	1.04	0.11	0.26
idsc: idiosyncratic skewness per the CAPM, k=1, m=0, n=1								
idsc1	0.65	2.71	0.65	2.75	0.92	2.81	0.88	2.16
ts: total skewness, k=1, m=0, n=1, Amaya, et al. (2015)								

ts1	0.83	2.86*	0.81	2.91*	0.67	2.02	0.51	1.30
betaFP: the Frazzini-Pedersen beta, k=12, m=0, n=1, Frazzini and Pedersen (2013)								
betaFP1	0.93	2.14	1.10	2.77	0.46	1.05	0.19	0.43
betaDM: the Dimson beta, k=1, m=0, n=1, Dimson (1979)								
betaDM1	1.10	2.59	0.95	2.46	0.96	1.92	1.65	3.05*

Panel C. Past Return Anomalies

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
mchg: momentum change, k=12, m=1, n=6, Gettleman and Marks (2006)								
mchg6	0.57	2.79	0.59	2.94*	0.83	3.29*	0.92	3.55*
im: residual momentum, prior 11-month returns, k=11, 6, m=0, n=1, 6, Blitz, et al. (2011)								
im11-1	0.68	2.70	0.69	2.71	0.87	2.52	0.82	2.16
lrev: long-term reversal, k=48, m=12, n=1, 6, 12, De Bondt and Thaler (1985)								
lrev1	0.81	2.02	0.79	2.01	0.52	1.35	0.93	2.04
52wh: 52-week high, k=12, m=0, n=12, George and Hwang (2004)								
52wh12	0.64	2.02	0.74	2.43	0.24	0.53	0.32	0.77
mdr: maximum daily return, k=1, m=0, n=1, Bali, et al. (2011)								
mdr1	0.89	2.29	1.02	2.70	0.27	0.54	-0.51	-1.08
ilr: industry lead-lag effect in prior returns, k=1, m=0, n=1								
ilr1	0.63	2.00	0.58	1.94	0.79	2.22	1.20	3.54*
abr: cumulative abnormal returns around earnings announcement dates, n=1, Chan, et al. (1996)								
abr1	0.61	2.09	0.64	2.29	0.46	1.33	0.48	1.47
Ra25: Years 2-5 lagged return, annual, Heston and Sadka (2008)								
Ra25	0.65	2.36	0.64	2.43	0.64	1.87	0.56	1.37
Rn25: Year 2-5-lagged return, nonannual, Heston and Sadka (2008)								
Rn25	0.93	2.30	0.94	2.41	0.58	1.37	1.00	1.85

Panel D. Profitability anomalies

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
roe: return on equity, n=1, 6, Hou, Xue, and Zhang (2015)								
roe1	1.14	3.16*	1.19	3.35*	-0.12	-0.37	-0.11	-0.32
roe6	0.83	2.47	0.87	2.63	-0.10	-0.39	0.00	0.00
droe: 4-quarter change in return on equity, n=1, 6, 12								
droe1	0.90	3.44*	0.89	3.47*	0.17	0.45	0.34	0.99
droe6	0.76	3.79*	0.75	3.78*	0.09	0.35	0.21	0.80
droe12	0.47	2.75	0.47	2.75	0.03	0.18	0.15	0.85
roa: return on assets, n=1, 6, Balakrishnan, Bartov, and Faurel (2010)								
roa1	1.07	2.90*	1.14	3.15*	0.14	0.42	0.02	0.07
roa6	0.88	2.62	0.93	2.86*	0.20	0.70	0.13	0.42
droa: 4-quarter change in return on assets, n=1, 6, 12								
droa1	0.79	3.63*	0.79	3.65*	0.15	0.51	0.20	0.69
droa6	0.72	3.99*	0.72	4.00*	0.12	0.59	0.19	0.88
droa12	0.32	2.01	0.32	2.04	-0.10	-0.55	-0.01	-0.04
rna: return on net operating assets, n=1, Soliman (2008)								

rnaq1	0.67	2.35	0.74	2.69	-0.14	-0.40	-0.32	-0.87
pmq: quarterly profit margin, n=1								
pmq1	0.83	2.16	0.92	2.58	-0.04	-0.16	-0.01	-0.06
atoq: asset turnover, monthly sort, n=1, Soliman (2008)								
ato	0.40	1.98	0.42	2.04	-0.03	-0.12	-0.14	-0.46
atoq1	0.46	2.22	0.50	2.36	0.04	0.13	-0.15	-0.47
gpla: gross profits-to-lagged assets, n=1, 6								
gplaq1	1.06	2.68	1.12	2.89*	0.26	0.79	0.11	0.30
gplaq6	0.83	2.31	0.89	2.63	0.16	0.54	0.02	0.05
ople: operating profits-to-lagged book equity, n=1, 6, 12								
opleq1	1.10	3.18*	1.15	3.37*	-0.04	-0.11	-0.02	-0.06
opleq6	0.81	2.51	0.85	2.69	-0.06	-0.23	0.00	0.00
opleq12	0.63	2.04	0.67	2.22	-0.02	-0.10	0.03	0.14
opla: operating profits-to-lagged assets, n=1, 6								
oplaq1	0.94	2.40	1.01	2.66	0.16	0.50	0.01	0.02
oplaq6	0.81	2.36	0.87	2.68	0.16	0.61	0.04	0.12
sg: annual sales growth, n=1, 6, Lakonishok, Shleifer, and Vishny (1994)								
sgq1	0.85	2.99*	0.84	3.03*	0.14	0.43	0.16	0.44
sgq6	0.55	2.70	0.55	2.74	0.21	0.85	0.30	1.21
sgq12	0.38	2.16	0.38	2.19	0.17	0.77	0.29	1.33
Fq: Fundamental score, n=1, 6, 12, Piotroski (2000)								
Fq1	0.84	2.81	0.89	3.34*	-0.28	-0.83	-0.41	-1.14
Fq6	0.54	3.02*	0.58	3.57*	-0.22	-1.25	-0.20	-1.08
Fq12	0.58	3.34*	0.61	3.88*	-0.01	-0.06	0.00	-0.02

Panel E. Value anomalies

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
bm: book-to-market equity								
bm	0.88	2.15	0.86	2.14	0.72	1.32	1.08	1.78
bmj: book-to-June-end market equity, annual sort, n=12, Asness and Frazzini (2013)								
bmj	0.87	2.45	0.85	2.43	0.83	1.76	1.09	2.14
bmq: quarterly book-to-market equity, n=1, 6, 12								
bmq1	1.18	2.81	1.16	2.85*	0.71	1.21	0.99	1.59
bmq6	0.81	2.00	0.79	2.00	0.62	1.12	0.98	1.56
bmq12	0.81	2.12	0.79	2.10	0.58	1.11	0.91	1.59
ep: earnings-to-price ratio, n=1, 6, 12, Basu (1983)								
epq1	1.30	3.45*	1.37	3.71*	0.06	0.25	0.14	0.58
epq6	0.95	2.74	1.00	2.99*	0.03	0.14	0.14	0.69
epq12	0.80	2.24	0.84	2.45	0.17	0.87	0.23	1.13
sp: sales-to-price ratio, n=1, 6, 12, Barbee, Mukherji, and Raines (1996)								
sp	0.91	2.61	0.88	2.68	0.37	0.87	0.54	1.13
spq1	1.00	3.23*	1.00	3.29*	0.15	0.41	0.16	0.40
spq6	0.94	3.02*	0.93	3.11*	0.29	0.76	0.33	0.80

spq12	0.90	3.08*	0.89	3.17*	0.36	0.99	0.49	1.23
ir: intangible return, annual sort, n=12, Daniel and Titman (2006)								
ir	0.85	2.19	0.80	2.14	0.61	1.40	0.96	1.91
ebp: enterprise book-to-price, n=1, Penman, Richardson, and Tuna (2007)								
ebpq1	0.81	2.24	0.77	2.23	0.65	1.42	0.79	1.47

Panel F. Investment anomalies

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
cei: composite equity issuance, annual sort, n=12, Daniel and Titman (2006)								
cei	0.66	2.15	0.64	2.14	0.40	1.11	0.47	1.20
dnca: change in noncurrent operating assets								
dnca	0.38	2.10	0.38	2.15	0.19	0.79	0.25	1.18

Panel G. Other anomalies

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
adm: advertising expense-to-market ratio, annual sort, n=12								
adm	0.56	2.80	0.59	2.87*	0.08	0.40	0.09	0.42
rdm: R&D expense-to-market, n=1, 6, 12, Chan, Lakonishok, and Sougiannis (2001)								
rdm	0.74	2.82	0.75	2.91*	0.56	1.82	0.59	1.81
rdmq1	1.26	4.28*	1.26	4.38*	0.72	1.95	0.77	2.05
rdmq6	1.00	3.55*	0.99	3.68*	0.70	2.15	0.86	2.45
rdmq12	0.90	3.57*	0.91	3.69*	0.63	1.97	0.76	2.33
tan: tangibility, n=1, 6, 12								
tan	0.88	2.63	0.92	2.96*	0.82	2.06	0.45	1.00
tanq1	0.81	2.80	0.81	2.94*	0.55	1.55	0.30	0.74
tanq6	0.85	2.95*	0.87	3.16*	0.53	1.53	0.23	0.57
tanq12	0.77	2.68	0.79	2.92*	0.46	1.31	0.10	0.26
cta: cash-to-assets, n=1, Palazzo (2012)								
cta1	0.57	2.05	0.55	2.03	0.33	1.06	0.49	1.47
evr: value relevance of earnings, annual sort, n=12, Francis, et al. (2004)								
evr	0.63	2.03	0.59	1.97	0.56	1.49	0.67	1.79
ala: asset liquidity, n=1, 6, 12								
alaq1	0.66	2.79	0.64	2.81	0.65	2.14	0.44	1.21
alaq6	0.68	2.87*	0.66	2.94*	0.57	1.92	0.35	1.06
alaq12	0.59	2.55	0.56	2.59	0.51	1.83	0.30	0.95
alm	0.74	2.07	0.66	2.21	0.49	1.66	0.71	2.11
almq1	0.89	2.38	0.82	2.42	0.67	1.63	0.95	2.02
almq6	0.95	2.58	0.88	2.71	0.70	1.84	0.99	2.24
almq12	0.92	2.48	0.84	2.63	0.63	1.74	0.94	2.22
sue: earnings surprise, n=1, 6, 12, Foster, Olsen, and Shevlin (1984)								
sue1	0.82	2.60	0.84	2.60	-0.27	-0.74	0.02	0.06
sue6	0.62	2.42	0.64	2.44	-0.22	-0.67	-0.02	-0.06
sue12	0.46	1.97	0.48	2.03	-0.14	-0.57	-0.04	-0.14

rs: revenue surprise, n=12, Jegadeesh and Livnat (2006)								
rs12	0.45	2.11	0.47	2.23	0.31	1.46	0.34	1.57
def: change in analysts' earnings forecasts, n=12, Hawkins, Chamberlin, and Daniel (1984)								
def12	0.32	2.36	0.29	2.42	0.10	0.75	0.19	1.71
efp: analysts' earnings forecasts-to-price, n=1, Elgers, Lo, and Pfeiffer (2001)								
efp1	1.39	2.47	1.46	2.65	-0.04	-0.08	0.00	0.01

Table 5. The Number of Significant Anomalies from Robust Checks

The table reports the number of significant anomalies with value-weighted long-short returns for different samples, different sample periods, and returns relative to the CAPM, CH3, and CH4 factor models in each category and in total at the 5% level with the hurdle of the *t*-cutoffs of 1.96 and 2.85, respectively. The anomalies are categorized into 2 groups, which are trading-related anomalies and accounting-related anomalies. We adopt single portfolio analysis to test whether these anomalies are significant with different samples (all A-share sample) in Panel A, and in different sample periods (2007-2018) in Panel B. Panel C shows the number of significant anomalies among the components of HS300 and ZZ500.

Panel A. All A-share sample, 2000-2018

	Trading-Related Anomalies			Accounting-Related Anomalies				
	liquidity	risk	past returns	profitability	value	investment	Others	Total
	94	46	52	67	44	51	72	426
Single test								
Raw return	26	13	10	11	11	3	21	95
CAPM alpha	26	13	10	11	11	3	20	94
CH3 alpha	24	2	4	0	0	0	3	33
CH4 alpha	19	3	6	0	1	0	9	38
Multiple test								
Raw return	6	5	1	5	2	0	7	26
CAPM alpha	6	5	1	5	2	0	7	26
CH3 alpha	4	1	0	0	0	0	0	5
CH4 alpha	3	0	0	0	0	0	0	3

Panel B. All A-share sample but 30% microcap firms, 2007-2018

	Trading-Related Anomalies			Accounting-Related Anomalies				
	liquidity	risk	past returns	profitability	value	investment	Others	Total
	94	46	52	67	44	51	72	426
Single test								
Raw return	19	8	6	15	1	6	13	68
CAPM alpha	19	8	6	15	1	5	13	67
CH3 alpha	8	0	0	0	0	0	4	12
CH4 alpha	2	1	1	0	0	0	3	7
Multiple test								
Raw return	2	0	0	1	0	0	0	3
CAPM alpha	2	0	0	1	0	0	0	3
CH3 alpha	1	0	0	0	0	0	0	1
CH4 alpha	0	0	0	0	0	0	0	0

Table 6. Composite Anomalies

This table shows the average raw returns, alphas relative to the CAPM, CH3 factors, and CH4 factors for value-weighted long-short portfolios of composite anomalies. We adopt four methods, which include score, Lewellen Fama-MacBeth regression, Lasso, and Random Forest methods. The composite anomalies are constructed with all firm characteristics in each category, subcategory, and in total. We construct 10 composite anomalies. The sample period is from July 2000 to December 2018. Note, the anomalies that pass the multiple test hurdle are denoted with *.

Panel A: Score

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
Trading-related	2.27	6.65*	2.39	7.20*	1.43	4.23*	0.85	2.01
liquidity	1.84	5.28*	1.92	5.67*	1.06	3.39*	0.59	1.42
risk	1.39	3.91*	1.50	4.41*	0.76	2.39	0.41	0.97
return	1.16	3.12*	1.26	3.66*	1.00	2.95*	0.73	1.87
Accounting-related	1.61	4.31*	1.62	4.40*	0.58	1.51	0.87	1.86
profitability	0.89	2.05	0.94	2.26	0.47	1.80	0.20	0.65
value	1.17	2.85*	1.19	2.95*	0.57	1.37	0.78	1.60
investment	0.83	3.59*	0.85	3.68*	0.51	1.95	0.60	2.24
others	0.98	3.13*	0.94	3.37*	0.20	0.62	0.49	1.30
Total	2.48	6.69*	2.55	6.73*	1.36	3.71*	1.12	2.01

Panel B. Lewellen FM

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
Trading-related	1.51	3.60*	1.53	3.78*	1.66	3.96*	1.22	2.71
liquidity	1.36	3.14*	1.37	3.29*	1.40	3.33*	0.90	1.90
risk	1.36	3.80*	1.40	4.11*	0.75	2.24	0.59	1.44
return	0.84	2.46	0.93	2.83	0.35	0.95	0.02	0.05
Accounting-related	0.88	2.94*	0.88	3.12*	0.05	0.15	0.08	0.23
profitability	0.14	0.37	0.12	0.31	0.06	0.15	-0.07	-0.17
value	0.37	1.06	0.40	1.18	-0.31	-1.01	-0.20	-0.57
investment	0.12	0.40	0.12	0.44	-0.43	-1.35	-0.28	-0.90
others	1.11	4.11*	1.09	4.21*	0.79	3.00*	1.02	3.67*
Total	1.30	3.29*	1.31	3.41*	1.07	2.80	0.62	1.80

Panel C. Lasso

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
Trading-related	2.34	5.51*	2.39	5.92*	1.55	3.91*	0.73	2.10
liquidity	2.23	5.74*	2.26	5.97*	1.79	5.57*	1.12	2.98*
risk	1.32	3.55*	1.42	3.85*	0.56	1.60	0.32	0.67
return	0.80	2.28	0.90	2.96*	0.25	0.62	-0.51	-1.58
Accounting-related	1.76	4.82*	1.75	4.88*	0.90	2.97*	0.99	3.13*
profitability	0.86	2.82	0.85	2.82	0.29	1.04	0.36	1.26
value	1.42	3.32*	1.42	3.41*	0.38	1.11	0.40	1.23
investment	0.49	1.53	0.50	1.62	0.44	1.19	0.23	0.58
others	1.35	5.02*	1.34	5.17*	0.85	3.01*	0.95	3.27*
Total	2.74	7.26*	2.77	7.51*	1.74	5.57*	1.17	3.27*

Panel D. Random Forest

	Raw return	t	CAPM alpha	t	CH3 alpha	t	CH4 alpha	t
Trading-related	3.11	7.60*	3.12	7.96*	2.35	5.85*	2.01	4.37*
liquidity	3.27	8.58*	3.28	9.00*	2.65	6.83*	2.08	5.29*
risk	1.70	5.88*	1.71	6.20*	1.11	3.03*	1.13	2.65
return	1.83	5.30*	1.86	5.67*	0.82	1.85	0.53	1.18
Accounting-related	1.80	5.13*	1.79	5.19*	0.74	2.64	0.77	2.62
profitability	1.05	3.41*	1.05	3.44*	0.46	1.31	0.28	0.74
value	1.37	3.69*	1.36	3.71*	0.54	1.47	0.46	1.41
investment	0.57	1.94	0.58	2.02	0.45	1.38	0.31	0.84
others	1.26	4.82*	1.27	4.96*	0.63	2.27	0.76	2.66
Total	3.32	8.15*	3.33	8.40*	2.09	5.42*	1.59	4.11*

Finding Anomalies in China
---Online Appendix

A. Different Types of Shares

There exist three different types of shares in Chinese stock market. Based on different investors and exchanges, the types of shares include A-, B-, and H-shares. A Shares are the shares of Chinese firms that listed on Shanghai and Shenzhen exchanges. They are denominated in Renminbi (RMB). A shares are open to domestic investors to trade. B Shares were established in 1992 on both the Shanghai and Shenzhen exchanges. Initially, the participants were exclusively foreign investors. After February 19, 2001, however, this market was opened to domestic individual investors. On the Shanghai Stock Exchange, prices are denominated in U.S. dollars, while on the Shenzhen Stock Exchange, prices are denominated in Hong Kong dollars. H shares refer to shares of companies registered in mainland China but listed and traded on the HongKong Exchange. Many companies issue their shares simultaneously on the HongKong Exchange and one of the two stock exchanges in mainland China. A shares are generally traded at a premium to H shares. Domestic investors are restricted from investing abroad, and foreign investors are also restricted from investing in the A-share market in mainland China. In this article, we use A shares, which comprise approximately 96% of all shares traded.

From the perspective of holdings and trading, there exist two types of shares, which are free floating and non-floating shares. The Chinese government and semi-government entities hold equity shares of the state-owned enterprises (SOEs), which are in the form of non-floating or non-tradable shares. The non-floating shares were traded between government and semi-government entities and later other legal entities through negotiations, typically at book value. They cannot be traded on exchanges. The free floating shares are issued to the general public, which are listed and traded on exchanges. Total shares are the sum of free floating and non-floating shares.

B. Anomaly Construction

B.1 Trading-Related Anomalies

B.1.1 Liquidity

B.1.1.1 Firm Size (size)

Banz (1981) presents evidence of a size effect in U.S. stock returns; that is, the negative relation between firm size and stock returns. Fama and French (1992) confirm that firm size has a negative relationship with stock returns. Firm size is equal to close price (unadjusted) multiplying by total A shares following Liu, Stambaugh and Yuan (2019). We sort stocks based on firm size at the end of June of each year t, and hold the portfolio over the next [t+1, t+n] months (n=12, size). We also sort stocks based on firm size at the end of each month t, wait for 0 month (m=0), and hold the portfolio over the next [t+m+1, t+m+n] months (n=1, 6, 12, size1, size6, size12).

B.1.1.2 Share Turnover (turn1, turn6, and turn12)

Following Datar, Naik and Radcliffe (1998), we calculate a stock's share turnover, *turn*, as the average of daily share turnover over the prior k months (k=1, 6 and 12). Daily turnover is calculated as the trading volume on a given day divided by the number of shares outstanding on that day. In China, shares outstanding are float A shares outstanding. We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on *turn* over [t-k+1, t] at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.3 Variation of Share Turnover (vturn1, vturn6, and vturn12)

Chordia, Subrahmanyam and Anshuman (2001) show a negative cross-sectional relationship between stock returns and the variability of dollar trading volume and share turnover, and the level of dollar volume or share turnover. Following Chordia, Subrahmanyam and Anshuman (2001), we measure the variation of share turnover, *vturn*, as the standard deviation of daily share turnover in prior k months (k=1, 6, and 12). Daily turnover is calculated as the trading volume on a given day divided by float A shares on that day. We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on *vt* over [t-k+1, t] at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.4 Coefficient of Variation of Share Turnover (cvturn1, cvturn6, and cvturn12)

Following Chordia, Subrahmanyam and Anshuman (2001), we calculate a stock's 1-month coefficient of variation of share turnover, *cvturn*, as the ratio of the standard deviation to the mean of daily share turnover in prior k months (k=1, 6, and 12). We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on c*vturn* over [t-k+1, t] at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.5 One-month Abnormal Turnover (abturn)

Following Liu, Stambaugh and Yuan (2019), we estimate abnormal turnover, *abturn*, as the ratio of average daily turnover in month t to its average daily turnover in prior one year over t-11 to t. We require a minimum of 75% of non-zero-volume trading days with trading records in month t and in prior 12 months. We sort stocks based on *abturn* over [t-k+1, t] months (k=12) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.6 Dollar Trading Volume (dtv1, dtv6, and dtv12)

Following Brennan, Chordia and Subrahmanyam (1998), we calculate a stock's dollar trading volume, *dtv1*, at the end of each month t, as the average of daily dollar trading volume in prior k months (k=1, 6, and 12). Daily dollar trading volume is downloaded from Wind. We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on *dtv* over [t-k+1, t] at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.7 Variation of Dollar Trading Volume (vdtv1, vdtv6, and vdtv12)

Following Chordia, Subrahmanyam and Anshuman (2001), we measure the variation of share turnover, *vdtv*, as the standard deviation of daily share turnover in prior k months (k=1, 6, and 12). Daily dollar trading volume is downloaded from Wind. We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on v*dtv* over [t-k+1, t] at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.8 Coefficient of Variation of Dollar Trading Volume (cvd1, cvd6, and cvd12)

Following Chordia, Subrahmanyam and Anshuman (2001), we measure a stock's coefficient of variation of dollar trading volume, *cvd*, as the ratio of the standard deviation to the mean of daily dollar trading volume in prior k months (k=1, 6, and 12). We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on *cvd* over [t-k+1, t] at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.9 Amihud Illiquidity (Absolute Return-to-Volume) (Ami1, Ami6, and Ami12)

Amihud (2002) shows that expected market illiquidity positively affects ex ante stock excess returns. Following Amihud (2002), we calculate the measure of Amihud illiquidity, *Ami1*, as the average of the ratio of the absolute daily stock returns to its daily dollar trading volume in prior k months (k=1, 6, 12). Daily dollar trading volume is downloaded from Wind. We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on *Ami* over [t-k+1, t] at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.10 Turnover-adjusted Number of Zero Daily Volume (Lm1, Lm6, and Lm12)

Liu (2006) documents that turnover-adjusted number of zero daily volume, *Lm,* has a positive relationship with stock returns. Following Liu (2006), the standardized turnover adjusted number of zero daily trading volume over the prior *k* months, *Lm*, is calculated as follows,

$$\mathrm{Lm}^{\mathrm{x}}=\left[\text{Number of volumes}<150000\text{ in prior x months}+\frac{\frac{1}{\text{k-month turnover}}}{\text{Deflator}}\right]\frac{21\mathrm{x}}{\mathrm{NoTD}} \tag{B.1.1}$$

where $\mathrm{turover}_{\mathrm{t}}$ is the sum of daily turnover over the prior *k* months (*k*=1, 6 and 12). Daily share turnover is calculated as the number of shares traded on a given day divided by the number of total A shares. *NoTD* is the total number of trading days over the prior *k* months. W set the deflator to $\max\left\{\frac{1}{\text{k}-\text{month Turnover}}\right\}+1$, in which the maximization is taken across all sample stocks each month. The choice of the deflator ensures that (1/(*k*-month turnover))/Deflator is between zero and one for all stocks. Since there are many suspension days for some stocks in China, we replace Number of zero volumes in prior x months by number of volumes less than 150000 in prior months. We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on *Lm* over [t-k+1, t] at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.1.11 Liquidity Betas (return-return, Bret, illiquidity-illiquidity, Blcc, return-illiquidity, Blrc, liquidity-return, Blcr, and net, Bnet)

Following Acharya and Pedersen (2005), we measure illiquidity using the Amihud (2002) measure, *Ami*. For stock i in month t, Ami_t^i is the average ratio of absolute daily return to daily dollar trading volumes. We require a minimum of 75% of non-zero-volume trading days with trading records. Daily dollar trading volume is downloaded from Wind. The market illiquidity, Ami_t^M, is the value-weighted average of $\min(\mathrm{Ami}_t^i, (30 - 0.25)/(0.30P_{t-1}^M))$, in which P_{t-1}^M is the ratio of the total market capitalization of Wind A index at the end of month t-1 to its value at the beginning of 1992. We measure market illiquidity innovations, $\epsilon_{\mathrm{M}t}^c$, as the residual from the regression below:

$$(0.25 + 0.3\mathrm{Ami}_t^M P_{t-1}^M) = a_0 + a_1(0.25 + 0.30Ami_{t-1}^M P_{t-1}^M) + a_2(0.25 + 0.3Ami_{t-2}^M P_{t-1}^M) + \epsilon_{Mt}^c \quad \text{(B.1.2)}$$

Similarly, we measure innovations to individual stock's illiquidity, ϵ_{it}^c, by replacing $\mathrm{Ami}^{\mathrm{M}}$ with $\min(\mathrm{Ami}_t^i,\ (30 - 0.25)/(0.30P_{t-1}^M))$ in equation (B.1.2). Finally, we measure innovations to the market return ϵ_{Mt}^r as the residual from the second-order autoregression of the market return. We define five measures of liquidity betas:

$$\text{Return-Return: } \mathrm{Bret_i} = \frac{Cov(r_{it}, \epsilon_{Mt}^r)}{var(\epsilon_{Mt}^r - \epsilon_{Mt}^c)} \quad \text{(B.1.3)}$$

$$\text{Illiquidity-illiquidity: } \mathrm{Blcc_i} = \frac{Cov(\epsilon_{it}^c,\ \epsilon_{Mt}^C)}{var(\epsilon_{Mt}^r - \epsilon_{Mt}^c)} \quad \text{(B.1.4)}$$

$$\text{Return-illiquidity: } \mathrm{Blrc_i} = \frac{Cov(r_{it}, \epsilon_{Mt}^C)}{var(\epsilon_{Mt}^r - \epsilon_{Mt}^c)} \quad \text{(B.1.5)}$$

$$\text{Illiquidity-Return: } \mathrm{Blcr_i} = \frac{Cov(\epsilon_{it}^c, \epsilon_{Mt}^r)}{var(\epsilon_{Mt}^r - \epsilon_{Mt}^c)} \quad \text{(B.1.6)}$$

$$\text{Net: } \mathrm{Bnet_i} = \mathrm{Bret_i} + \mathrm{Blcc_i} - \mathrm{Blrc_i} - \mathrm{Blcr_i} \quad \text{(B.1.7)}$$

We estimate $\mathrm{Bret_i}, \mathrm{Blcc_i}, \mathrm{Blrc_i}, \mathrm{Blcr_i}$, and $\mathrm{Bnet_i}$ with the prior 60 months from t-59 to t. We require a minimum of 75% of non-zero-volume trading days with trading records in prior 6 years. We sort stocks based on liquidity betas over [t-k+1, t] months (k=60) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2 Risk

B.1.2.1 Idiosyncratic Volatility (idv)

Based on Ali, Hwang and Trombley (2003), we calculate idiosyncratic volatility per the CAPM, *idvc*, as the standard deviation of residuals from regressing a stock's daily excess returns over one-year deposit rate on the daily value-weighted excess market returns (Wind A-share index returns minus one-year deposit rate) in the prior one year (k=12). We require a minimum of 75% of non-zero-volume trading days with trading records in prior k months. We sort stocks based on *idv* over [t-k+1, t] months (k=12) at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.1.2.2 Idiosyncratic Volatility per the CH3 Factor Model (idvff)

Ang, Hodrick, Xing and Zhang (2006) find that there is a negative relationship between idiosyncratic volatility and stock returns. Following Ang, Hodrick, Xing and Zhang (2006), we calculate idiosyncratic volatility relative to the Chinese three-factor model (CH3), *idvff*, is calculated as the standard deviation of residuals from regressing a stock's daily excess returns over one-year deposit rate on the CH3 factors during month t (k=1). We require a minimum of 75% of non-zero-volume trading days with trading records in month t. The CH3 factors are obtained from Liu, Stambaugh and Yuan (2019). We sort stocks based on *idvff* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.3 Idiosyncratic Volatility per the q-factor Model (idvq)
We calculate idiosyncratic volatility relative to the q-factor model, *idvq*, is calculated as the standard deviation of residuals from regressing a stock's daily excess returns over one-year deposit rate on the q-factor during month t. We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *idvq* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.4 Idiosyncratic Volatility per the CAPM (idvc)
We calculate idiosyncratic volatility per the CAPM, *idvc*, as the standard deviation of residuals from regressing a stock's daily excess returns over one-year deposit rate on value-weighted excess market returns (Wind A-share index returns minus one-year deposit rate) during month t. We require a minimum of 75% of non-zero-volume trading days with trading records in month *t*. We sort stocks based on *idvc* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.5 Total Volatility (tv)
Following Ang, Hodrick, Xing and Zhang (2006), we estimate total volatility, *tv*, as the standard deviation of a stock's daily returns in month *t* (k=1). We require a minimum of 75% of non-zero-volume trading days with trading records in month *t*. We sort stocks based on *tv* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.6 Idiosyncratic Skewness per the CH3 Factor Model (idsff)
Boyer, Mitton and Vorkink (2009) find that expected idiosyncratic skewness and returns are negatively correlated. We estimate idiosyncratic skewness relative to the CH3 factor model, *idsff*, as the skewness of the residuals from regressing a stock's daily excess returns on the CH3 factors in month *t*. We require a minimum of 75% of non-zero-volume trading days with trading records in month *t*. The CH3 factors are obtained from Liu, Stambaugh, and Yuan (2019). We sort stocks based on *idsff* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.7 Idiosyncratic Skewness per the q-Factor Model (idsq)
We estimate idiosyncratic skewness relative to the q-factor model, *idsq*, as the skewness of the residuals from regressing a stock's daily excess returns on the q-factor in month *t*. We require a minimum of 75% of non-zero-volume trading days with trading records in month *t*. We sort stocks based on *idsq* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.8 Idiosyncratic Skewness per the CAPM (idsc)

We calculate idiosyncratic skewness per the CAPM, *idsc*, as the skewness of residuals from regressing a stock's daily excess returns on the daily value-weighted excess market returns (Wind A-share index returns minus the one-year deposit rate) in month t. We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *idsc* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.9 Total Skewness (ts)

Amaya, Christoffersen, Jacobs and Vasquez (2015) find that the relation between realized skewness and next week's stock returns is negative. They use intraday high-frequency returns to calculate realized skewness. Following Hou, Xue and Zhang (2019), we calculate total skewness, *ts*, as the skewness of a stock's daily returns from month *t*. We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *ts* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.10 Co-skewness (cs)

Harvey and Siddique (2000) find that systematic skewness commands a risk premium. They suggest that the momentum effect is related to systematic skewness. The low expected return momentum portfolios have higher skewness than the high expected return momentum portfolios. Following Harvey and Siddique (2000), we calculate co-skewness, *cs*, as:

$$cs = \frac{E[\epsilon_i, \epsilon_m^2]}{\sqrt{E[\epsilon_i^2]}E[\epsilon_m^2]} \tag{B.1.8}$$

In which ϵ_i is the residual from regressing a stock's excess returns on excess market returns, and ϵ_m is the demeaned market returns. At the end of each month *t*, we estimated co-skewness, *cs*, with daily stock returns from month *t*. We require a minimum of 75% of non-zero-volume trading days with trading records in month *t*. We sort stocks based on *cs* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.11 Market Beta Using Monthly Returns (betam)

The CAPM indicates that market beta should have a positive relationship with stock returns. Market beta is the sensitivity of stock returns to market returns, $beta_t = \frac{cov(r_i, r_m)}{var(r_m)}$, where r_m is market index excess returns, which is value-weighted A-share stock excess returns across Chinese A-share stocks. At the end of each month *t*, we estimate market beta, *betam*, with monthly returns from month *t*-59 to month *t* following Fama and MacBeth (1973). We require a minimum of 75% of months with monthly returns in prior 6 years. We sort stocks based on *betam* over [t-k+1, t] months (k=60) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.12 Market Beta Using Daily Returns (beta)

At the end of each month *t*, we estimate market beta, *beta*, with daily returns from the prior 12 months from month t-11 to month t. We require a minimum of 75% of non-zero-volume trading days with trading records from month t-11 to month t. We sort stocks based on *beta* over [t-k+1, t] months (k=12) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.13 Downside Beta (dbeta)

Ang, Chen and Xing (2006) find that the downside beta can improve the positive beta-return relationship. Based on Ang, Chen and Xing (2006), we calculate the downside beta, *dbeta*, as:

$$\beta^- = \frac{Cov(r_i, r_m | r_m < \mu_m)}{Var(r_m | r_m < \mu_m)} \qquad \text{(B.1.9)}$$

where r_i and r_m are stock and market excess returns, respectively, and μ_m is the average of excess market returns. At the end of each month t, we estimate *dbeta*, with daily returns from the prior 12 months from month t-11 to month t. We only use daily observations conditional on $r_m < \mu_m$. We require a minimum of 75% of non-zero-volume trading days with trading records from month t-11 to month t. We sort stocks based on *dbeta* over [t-k+1, t] months (k=12) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.14 The Frazzini-Pedersen Beta (betaFP)

Following Frazzini and Pedersen (2014), we estimate the market beta for stock i, $betaFP = \hat{\rho}\frac{\hat{\sigma}_i}{\hat{\sigma}_m}$, in which $\hat{\sigma}_i$ and $\hat{\sigma}_m$ are the estimated stock and market return volatilities, respectively, and $\hat{\rho}$ is the stock and market return correlation. To estimate return volatilities, we compute the standard deviation of daily log returns over a one-year rolling window (a minimum of 75% of non-zero-volume trading days with trading records in one year). To estimate return correlations, we use overlapping three-day log returns, $r_{i,t}^{3d} = \sum_{k=0}^{2} log\left(1 + R_{t+k}^{i}\right)$, over a five-year rolling window (with at least 75% of trading days in 5 years). We sort stocks based on *betaFP* over [t-k+1, t] months (k=120) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.15 The Dimson Beta (betaDM)

Following Dimson (1979), we use the lead and lag of market returns along with the contemporaneous market returns, when estimating the market beta, *betaDM*,

$$r_{id} - r_{fd} = \alpha_i + \beta_{i1}\left(r_{md-1} - r_{fd-1}\right) + \beta_{i2}\left(r_{md} - r_{fd}\right) + \beta_{i3}\left(r_{md+1} - r_{fd}\right) + \epsilon_{id} \qquad \text{(B.1.10)}$$

in which r_{id} is stock i's returns on day d, r_{md} is value-weighted market returns on day d, and r_{fd} is the risk-free rate (one-year deposit rate). The Dimson beta for stock i, $betaDM = \hat{\beta}_{i1} + \hat{\beta}_{i2} + \hat{\beta}_{i3}$. At the end of each month t, we estimate *betaDM* with the daily returns from month t. We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *betaDM* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.2.16 Tail Risk (tail)

Following Kelly and Jiang (2014), we estimate common tail risk,

$$\lambda_t = \frac{1}{K_t}\sum_{k=1}^{K_t}\log\left(\frac{R_{kt}}{\mu_t}\right) \qquad \text{(B.1.11)}$$

where μ_t is the fifth percentile of all daily returns in month t, R_{kt} is the *k*th daily return that is below μ_t, and K_t is the total number of daily returns that are below μ_t. Tail risk, *tail*, is the slope from regressing a stock's monthly excess returns on 1-month-lagged common tail risk over the most recent 120 months from t-120 to t. We require a minimum of 75% of months with monthly returns in the most recent one year. We sort stocks based on *tail* over [t-k+1, t] months (k=120) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3 Past Returns

B.1.3.1 Prior k-month Momentum (m11, m9, m6, m3)

Following Jegadeesh and Titman (1993), we estimate the x-month momentum, *m*, in month *t* as the x-month cumulative daily returns from month *t-k+1* to month *t* (k=3, 6, 9, and 11). We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *m11, m9, m6, or m3* over [t-k+1, t] months (k=3, 6, 9, 11) at the end of each month t, wait for 1 month (m=1), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.2 Industry Momentum (inm)

We start with Shenwanhongyuan 28 industry classification. We have 28 industries after excluding financial firms. At the end of each month *t*, we sort industries based on their prior 6-month value-weighted returns from *t-5* to *t*. We do not skip month *t*. We form nine portfolios (9*3), each of which contains three different industries. We define the return of a given portfolio as the simple average of the three industry returns within the portfolio. We sort stocks based on *inm* over [t-k+1, t] months (k=6) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.3 Prior 24-month Momentum (m24)

Following Jegadeesh and Titman (1993), we estimate the 24-month momentum, *m24*, in month *t*, as the 24-month cumulative daily returns from month *t*-23 to month *t*. We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *m24* over [t-k+1, t] months (k=24) at the end of each month t, wait for 12 months (m=12), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.4 Momentum Change (mchg)

Following Gettleman and Marks (2006), we estimate momentum change, *mchg*, in month *t* as calculated as cumulative daily returns from month *t-5* to month *t* minus cumulative daily returns from month *t-11* to month *t-6*. We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *mchg* over [t-k+1, t] months (k=12) at the end of each month t, wait for 1 month (m=1), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.5 Short-Term Reversal (srev)

Jegadeesh (1990) finds that the monthly return in the previous month is negatively related to monthly stock returns in the next month. Short-term reversal is the monthly return in month t. We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *srev* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.6 Long-Term Reversal (lrev)

Following De Bondt and Thaler (1985), the long-term reversal, *lrev*, in month *t* is calculated as the cumulative daily returns from month *t*-47 to month *t* (k=48). We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *lrev* over [t-k+1, t] months (k=48) at the end of each month t, wait for 0 month (m=12), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.7 k-month Residual Momentum (im11, im6)

Blitz, Huij and Martens (2011) find that residual momentum earns risk-adjusted profits that are about twice as large as those associated with total return momentum. According to Blitz, Huij and Martens (2011), the k-month residual momentum, *im*, is the k-month cumulative residuals of regressing stock returns on the CH3 factors over the prior 36 months from month *t*-35 to month *t* scaled by their standard deviation over the same period (k=6 and 11). To reduce the estimation noisiness, we require monthly returns to be available for all prior 36 months. We sort stocks based on *im11, im6* over [t-k+1, t] months (k=11, 6) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.8 52-Week High (52w)

George and Hwang (2004) show that firms with stock prices nearest to their 52-week highs earn higher factor-adjusted returns on average than firms whose stock prices are farthest from their 52-week highs. At the end of each month *t*, we estimate 52-week high, *52w*, as the ratio of its split-adjusted price per share at the end of month *t* to its highest (daily) split-adjusted price per share during the prior one-year period from month *t*-11 to month *t*. We sort stocks based on *52w* over [t-k+1, t] months (k=12) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.9 Maximum Daily Return (mdr)

Bali, Cakici and Whitelaw (2011) report that portfolio-level analysis and firm-level cross-sectional regressions indicate a negative and significant relation between the maximum daily return over the past one month and expected stock returns. Because of the daily 10% price limit rule after 1996 in China, we define the maximum daily return, *mdr*, as the average of the 5 highest daily returns of the given stock in a given month following Bali, Brown, Murray and Tang (2017). We require a minimum of 75% of non-zero-volume trading days with trading records in month t. We sort stocks based on *mdr* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.10 Share Price (pr)

Following Miller and Scholes (1982), share price, *pr*, is observed at the end of month *t*. Share price is adjusted for splitting and delisting. We sort stocks based on *pr* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.11 Industry Lead-lag Effect in Prior Returns (ilr)

We start with Shenwanhongyuan 28 industry classification. We have 27 industries after excluding financial firms. At the end of each month t, we sort industries based on their prior 1-month value-weighted returns of the portfolio consisting of the 30% biggest (total A-share market capitalization) firms within an industry in month t. We form nine portfolios (9*3), each of which contains three different industries. We define the return of a given portfolio as the simple average of the three industry returns within the portfolio. We sort stocks based on *ilr* over [t-k+1, t] months (k=1) at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.12 Cumulative Abnormal Returns around Earnings Announcement Dates (abr)

We calculate cumulative abnormal stock return (*abr*) around the latest quarterly earnings announcement date (Chan, Jegadeesh, and Lakonishok, 1996):

$$\mathrm{abr_i} = \sum_{d=-2}^{+1}(r_{id} - r_{md}) \qquad \text{(B.1.12)}$$

where r_{id} is stock i's return in day d (with earnings announcement date on day 0) and r_{md} is the value-weighted market index return. We calculate returns until one (trading) day after the announcement date to account for the 1-day-delayed reaction to earnings news. At the end of each month, we sort stocks into deciles based on their most recent past *abr*. For a firm to enter our portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent *abr* to be within 12 months prior to the portfolio formation. The reason is that there are only semiannual data released in China prior 2002. Note, there is no earnings announcement date in China. Earnings are announced with the accounting statements. Wind provides the accounting announcement dates. We sort stocks based on *abr* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.1.3.13 Seasonality (Ra1, Rn1, Ra25, Rn25)

Following Heston and Sadka (2008), at the end of each month t, we estimate various measures of past performance. It includes returns in month t-11(*Ra1*), average returns from month t-10 to month t (*Rn1*), average returns across month t-23, t-35, t-47, and t-59 (*Ra25*), and average returns from month t-59 to month t-12 except for returns in month t-23, t-35, t-47, and t-59 (*Rn25*). We sort stocks based on these seasonality variables at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1).

B.2. Accounting-Related Anomalies

B.2.1 Profitability

B.2.1.1 Return on Equity (roe)

According to Hou, Xue and Zhang (2015), return on equity, *roe*, is equal to earnings before extraordinary items divided by lagged common shareholders' equity. At the end of each month t, we measure *roe*, as the quarterly net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses ("PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP") for the latest fiscal quarter after its announcement date divided by one-quarter lagged book value of equity. The book value of equity is total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus preferred shares (book value is 1 in China, zero if missing). We sort stocks based on *roe* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.2 4-quarter Changes in Return on Equity (droe)

At the end of each month *t*, we estimate change in return on equity, *droe*, as *roe* (in 2.1.1) for the latest fiscal quarter after its announcement date minus its value from four quarters ago. We sort stocks based on *droe* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.3 Return on Assets (roa)

Following Balakrishnan, Bartov and Faurel (2010), at the end of each month *t*, we measure return on assets, *roa*, as quarterly net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses (Wind income statement item "PLUS_NON_OPER_REV" minus Wind income statement item "LESS_NON_OPER_EXP") for the latest fiscal quarter after announcement divided by one-quarter lagged total assets (Wind balance sheet item "TOT_ASSETS"). We sort stocks based on *roa* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.4 4-quarter Changes in Return on Assets (droa)

At the end of each month *t*, we estimate change in return on assets, *droa*, as *roa* (in 2.1.3) for the latest fiscal quarter after its announcement date minus its value from four quarters ago. We sort stocks based on *droa* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.5 Return on Net Operating Assets, Profit Margin, Assets Turnover (rna, pm, ato)

Soliman (2008) decomposes $roe = rna + flev * spread$, in which roe is return on equity, *rna* is return on net operating assets, *flev* is financial leverage, and spread is the difference between return net operating assets and borrowing costs. We further decompose *rna* as $pm * ato$, in which pm is profit margin, and *ato* is assets turnover.

At the end of June of year *t*, we measure *rna* as operating income (Wind income statement item "OPER_PROFIT") for the fiscal year ending in calendar year *t*-1 divided by net operating assets (*noa*) for the fiscal year ending in calendar year *t*-2. Net operating assets (*noa*) are operating assets minus operating liabilities. Operating assets are total assets (Wind balance sheet item "TOT_ASSETS") minus cash and short-term investment (Wind balance sheet item "MONETARY_CAP", 0 if missing). Operating liabilities are total assets (Wind balance sheet item "TOT_ASSETS") minus debt in current liabilities (Wind balance sheet item "ST_BORROW", zero if missing) minus long-term debt (Wind balance sheet item "LT_BORROW", zero if missing) minus minority interests (Wind balance sheet item "MINORITY_INT", zero if missing) minus preferred stock shares (zero if missing) minus common equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT"). *pm* is operating income (Wind income statement item "OPER_PROFIT") divided by sales (Wind income statement item "OPER_REV") for the fiscal year ending in calendar year t-1. *ato* is sales

(Wind income statement item "OPER_REV") for the fiscal year ending in calendar year t-1 divided by *noa* for the fiscal year ending in year t-2. We sort stocks based on these anomaly variables at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.6 Quarterly Return on Net Operating Assets, Profit Margin, Assets Turnover (rnaq, pmq, atoq)

At the end of each month of month *t*, we measure quarterly return on net operating assets, *rnaq*, as quarterly operating income (Wind income statement item "OPER_PROFIT") for the latest fiscal quarter after its announcement date divided by 1-quarter lagged net operating assets. Net operating assets (*noa*) is operating assets minus operating liabilities. Operating assets are total assets (Wind balance sheet item "TOT_ASSETS") minus cash and short-term investment (Wind balance sheet item "MONETARY_CAP", zero if missing). Operating liabilities are total assets (Wind balance sheet item "TOT_ASSETS") minus debt in current liabilities (Wind balance sheet item "ST_BORROW", zero if missing) minus long-term debt (Wind balance sheet item "LT_BORROW", zero if missing) minus minority interests (Wind balance sheet item "MINORITY_INT", zero if missing) minus preferred stock shares (zero if missing) minus common equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT"). *pm* is operating income (Wind income statement item "OPER_PROFIT") divided by sales (Wind income statement item "OPER_REV") the latest fiscal quarter after its announcement date. *ato* is sales (Wind income statement item "OPER_REV") the latest fiscal quarter after its announcement date divided by one-quarter lagged *noa*. We sort stocks based on these anomaly variables at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.7 Capital Turnover (ct)

According to Haugen and Baker (1996), at the end of June of each year *t*, we measure capital turnover, *ct*, as sales/operating revenue (Wind income statement item "OPER_REV") for the fiscal year ending in calendar year *t*-1 divided by one-year lagged total assets (Wind income statement item "OPER_REV") for the fiscal year ending in year *t*-2. We sort stocks based on *ct* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.8 Quarterly Capital Turnover (ctq)

We measure quarterly capital turnover, *ctq*, as quarterly sales divided by operating revenue (Wind income statement item "OPER_REV") for the latest fiscal quarter after its announcement divided by one-quarter-lagged total assets (Wind balance sheet item "TOT_ASSETS"). We sort stocks based on *ctq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.9 Gross Profits to Assets (gpa)

Novy-Marx (2013) discovers that sorting on gross profit-to-assets creates abnormal benchmark-adjusted returns, with more profitable firms having higher returns than less profitable ones. Following Novy-Marx (2013), we measure gross-profits-to-assets, *gpa*, as total revenue minus cost of goods sold (gross profit, Wind income statement item "TOT_PROFIT") for the fiscal year ending in calendar year *t*-1 divided by total assets (Wind balance sheet item

"TOT_ASSETS") for the fiscal year ending in calendar year *t*-1. We sort stocks based on *gpa* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.10 Gross Profits to Lagged Assets (gpla)

The gross-profits-to-lagged-assets ratio, *gpla*, is calculated as total revenue minus cost of goods sold (gross profit, Wind income statement item "TOT_PROFIT") for the fiscal year ending in calendar year *t*-1 divided by one-year-lagged total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-2. We sort stocks based on *gpla* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.11 Quarterly Gross Profits to Lagged Assets (gplaq)

We measure the quarterly gross profits to lagged assets ratio, *gplaq*, as quarterly total revenue minus cost of goods sold (Wind income statement item "TOT_PROFIT") for the latest fiscal quarter after its announcement date divided by one-quarter-lagged total assets (Wind balance sheet item "TOT_ASSETS"). We sort stocks based on *gplaq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.12 Operating Profits to Equity (ope)

Following Fama and French (2015), we measure the operating profits to equity ratio, *ope*, as total revenue minus cost of goods sold minus selling, general, and administrative expense and minus interest expense (operating profits, Wind income statement item "OPER_PROFIT") for the fiscal year ending in calendar year *t*-1 scaled by book equity for the fiscal year ending in calendar year *t*-1. Book equity is total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus preferred stock shares (zero if missing). We sort stocks based on *ope* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.13 Operating Profits to Lagged Equity (ople)

We estimate operating profitability to lagged equity, *ople*, as total revenue minus cost of goods sold minus selling, general, and administrative expense and minus interest expense (operating profit, Wind income statement item "OPER_PROFIT") for the fiscal year ending in calendar year *t*-1 scaled by one-year-lagged book equity for the fiscal year ending in calendar year *t*-2. Book equity is total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus preferred stock shares (zero if missing). We sort stocks based on *ople* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.14 Quarterly Operating Profits to Lagged Equity (opleq)

Quarterly operating profits to lagged equity, *opleq*, is quarterly operating profits (Wind income statement item "OPER_PROFIT") for the latest fiscal quarter after its announcement divided by one-quarter-lagged book equity (total shareholders' equity). We sort stocks based on *opleq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.15 Operating Profits to Assets (opa)

Following Ball, Gerakos, Linnaimma and Nikolaev (2015), we measure the operating profits to total assets ratio, *opa*, as total revenue minus cost of goods sold minus selling, general, and administrative expense and minus interest expense (operating profits, Wind income statement item "OPER_PROFIT") for the fiscal year ending in calendar year *t*-1 scaled by total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-1. We sort stocks based on *opa* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.16 Operating Profits to Lagged Assets (opla)

The operating profits to lagged equity ratio, *opla*, is calculated as total revenue minus cost of goods sold minus selling, general, and administrative expense and minus interest expense (operating profits, Wind income statement item "OPER_PROFIT") for the fiscal year ending in calendar year *t*-1 scaled by one-year-lagged total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-2. We sort stocks based on *opla* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.17 Quarterly Operating Profits to Lagged Assets (oplaq)

The quarterly operating profits to lagged assets ratio, *oplaq*, is quarterly operating profits (Wind income statement item "OPER_PROFIT") for the fiscal quarter after its announcement date divided by one-quarter-lagged total assets (Wind balance sheet item "TOT_ASSETS"). We sort stocks based on *oplaq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.18 Taxable Income to Book Income (tbi)

Following Lev and Nissim (2004), the taxable income to book income ratio, *tbi*, is measured as pretax income (Wind income statement "EBIT" minus "LESS_INT_EXP") for the fiscal year ending in calendar year *t*-1 divided by net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") for the fiscal year ending in calendar year *t*-1. We sort stocks based on *tbi* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.19 Quarterly Taxable Income to book Income (tbiq)

At the end of each month *t*, we measure the quarterly taxable income to book income ratio *tbiq*, as quarterly pretax income (Wind income statement "EBIT" minus "LESS_INT_EXP") divided by net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") from the latest fiscal quarter after announcement date. We sort stocks based on *tbiq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.20 Book Leverage (bl)

According to Fama and French (1992), the book leverage, *bl*, is total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-1 divided by book equity for the fiscal year ending in calendar year *t*-1. Book equity is total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus preferred stock shares (zero if missing). We sort stocks based on *bl* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.21 Quarterly Book Leverage (blq)
At the end of each month *t*, we measure the quarterly book leverage, *blq*, as total assets (Wind balance sheet item "TOT_ASSETS") divided by book equity for the latest fiscal quarter after announcement date. Book equity is total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus preferred stock shares (zero if missing). We sort stocks based on *blq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.22 Annual Sales Growth (sg)
Lakonishok, Shleifer and Vishny (1994) find growth in sales has a negative relation to stock returns. At the end of June of year *t*, we estimate sales growth, *sg*, as the annual growth in sales (operating revenue, Wind income statement item "OPER_REV") from the fiscal year ending in calendar year *t*-2 to the sales for the fiscal year ending in calendar year *t*-1. We exclude firms with non-positive sales in the sample. We sort stocks based on *sg* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.23 Quarterly Sales Growth (sgq)
The quarterly sales growth, *sgq*, is quarterly sales (operating revenue, Wind income statement item "OPER_REV") divided by its value four quarters ago. At the end of each month *t*, we measure *sgq* for the latest fiscal quarter after its announcement date. We sort stocks based on *sgq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.24 Fundamental Score (Fscore)
Piotroski (2000) classifies each fundamental signal as either good or bad depending on the signal's implication for future stock prices and profitability. An indicator variable for a particular signal is one if its realization is good and zero if it is bad. The aggregate signal, denoted Fscore, is the sum of the nine binary signals. F is designed to measure the overall quality, or strength, of the firm's financial position. The nine fundamental signals are chosen to measure three areas of a firm's financial condition, profitability, liquidity, and operating efficiency.

We use four variables to measure profitability: (1) Roa is net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses ("PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP") scaled by 1-year-lagged total assets (Wind balance sheet item "TOT_ASSETS"). If the firm's Roa is positive, the indicator variable F_{roa} equals one and zero otherwise. (2) Cf/A is cash flow from operation (Wind cash flow statement item "NET_CASH_FLOWS_OPER_ACT") scaled by 1-year-lagged total assets (Wind balance sheet item "TOT_ASSETS"). If the firm's Cf/A is positive, the indicator variable $F_{Cf/A}$ equals one and zero otherwise. (3) dRoa is the current year's Roa less the prior year's Roa. If *dRoa* is positive, the indicator variable F_{dROA} is one and zero otherwise. Finally, (4) the indicator F_{Acc} equals one if Cf/A > Roa and zero otherwise.

We use three variables to measure changes in capital structure and a firm's ability to meet debt obligations. Piotroski (2000) assumes that an increase in leverage, a deterioration of liquidity, or the use of external financing is a bad signal about financial risk. (1) dLever is the change in the ratio of total long-term debt (long-term borrow, Wind balance sheet item "LT_BORROW") to the average of current and 1-year-lagged total assets. F_{dLever} is one if the firm's leverage ratio falls, that is, dLever<0, and zero otherwise. (2) *dLiquid* measures the change in a firm's current ratio from the prior year, in which the current ratio is the ratio of current assets (Wind balance sheet item "TOT_CUR_ASSETS") to current liabilities (Wind balance sheet item "TOT_CUR_LIAB"). An improvement in liquidity (dLiquid>0) is a good signal about the firm's ability to service debt obligations. The indicator $F_{dLiquid}$ equals one if the firm's liquidity improves and zero otherwise. (3) The indicator, Eq, equals one if the firm does not issue common equity during the current year and zero otherwise. The issuance of common equity is sales of common and preferred stock shares minus any increase in preferred stock shares. Issuing equity is interpreted as a bad signal (inability) to generate sufficient internal funds. Since there is no data to indicate the issuance of common equity, we do not include this indicator.

The remaining two signals are designed to measure changes in the efficiency of the firm's operations that reflect two key constructs underlying the decomposition of return on assets. (1) *dMargin* is the firm's current gross margin ratio, measured as gross margin (operating profit, Wind income statement item "OPER_PROFIT") scaled by sales (operating revenue, Wind income statement item "OPER_REV"), less the prior year's gross margin ratio. An improvement in margins signifies a potential improvement in factor costs, a reduction in inventory costs, or a rise in the price of the firm's product. The indictor $F_{dMargin}$ equals one if dMargin>0 and zero otherwise. (2) *dTurn* is the firm's current year asset turnover ratio, measured as total sales (operating revenue, Wind income statement item "OPER_REV") scaled by 1-year-lagged total assets (Wind balance sheet item "TOT_ASSETS"), minus the prior year's asset turnover ratio. An improvement in asset turnover ratio signifies greater productivity from the asset base. The indicator, F_{dTurn}, equals one if *dTurn*>0 and zero otherwise. Piotroski (2000) forms a composite score, F, as the sum of the individual binary signals:

$$\text{Fscore} = F_{roa} + F_{droa} + F_{cf/A} + F_{Acc} + F_{dMargin} + F_{dTurn} + F_{dLever} + F_{dLiquid} \quad \text{(B.2.1)}$$

At the end of June of each year t, we measure Fscore for the fiscal year ending in calendar year t−1 to form seven portfolios: low (F = 0, 1, 2), 3, 4, 5, 6, and high (F = 7, 8). Because extreme F scores are rare, we combine scores 0, 1, and 2 into the low portfolio and scores 7 and 8 into the high portfolio. We sort stocks based on Fscore at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.25 Quarterly Fundamental score (fq)

We use quarterly data to measure quarterly F-score. Four variables to measure profitability are: (1) *ROA* is quarterly net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses ("PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP") scaled by 1-quarter-lagged total assets (Wind balance sheet item "TOT_ASSETS"). If the firm's *ROA* is positive, the indicator variable F_{roa} equals one and zero otherwise. (2) CF/A is cash flow from operation (Wind cash flow statement item "NET_CASH_FLOWS_OPER_ACT") scaled by 1-quarter-lagged total assets (Wind balance sheet item "TOT_ASSETS"). If the firm's CF/A is positive, the indicator variable $F_{CF/A}$ equals one and zero otherwise. (3) *dROA* is the current year's *ROA* less the prior year's *ROA*. If dRoa is positive, the indicator variable F_{dROA} is one and zero otherwise. Finally, (4) the indicator F_{Acc} equals one if CF/A > *ROA* and zero otherwise.

We use three variables to measure changes in capital structure and a firm's ability to meet debt obligations. (1) *dLever* is the change in the ratio of total long-term debt (long-term borrow, Wind balance sheet item "LT_BORROW") to the average of current and 1-quarter-lagged total assets. F_{dLever} is one if the firm's leverage ratio falls, that is, *dLever*<0, and zero otherwise. (2) *dLiquid* measures the change in a firm's current ratio from the prior year, in which the current ratio is the ratio of current assets (Wind balance sheet item "TOT_CUR_ASSETS") to current liabilities (Wind balance sheet item "TOT_CUR_LIAB"). An improvement in liquidity (*dLiquid*>0) is a good signal about the firm's ability to service debt obligations. The indicator $F_{dLiquid}$ equals one if the firm's liquidity improves and zero otherwise.

The remaining two signals are designed to measure changes in the efficiency of the firm's operations that reflect two key constructs underlying the decomposition of return on assets. (1) *dMargin* is the firm's current gross margin ratio, measured as gross margin (operating profit, Wind income statement item "OPER_PROFIT") scaled by sales (operating revenue, Wind income statement item "OPER_REV"), less gross margin ratio four quarters ago. An improvement in margins signifies a potential improvement in factor costs, a reduction in inventory costs, or a rise in the price of the firm's product. The indictor $F_{dMargin}$ equals one if dMargin>0 and zero otherwise. (2) *dTurn* is the firm's current year asset turnover ratio, measured as total sales (operating revenue, Wind income statement item "OPER_REV") scaled by 1-quarter-lagged total assets (Wind balance sheet item "TOT_ASSETS"), minus the asset turnover ratio four quarters ago. An improvement in asset turnover ratio signifies greater productivity from the asset base. The indicator, F_{dTurn}, equals one if dTurn>0 and zero otherwise. The composite score, F, as the sum of the individual binary signals:

$$\text{Fscore} = \mathrm{F_{roa}} + F_{droa} + F_{cf/A} + F_{Acc} + F_{dMargin} + F_{dTurn} + F_{dLever} + F_{dLiquid} \quad \text{(B.2.2)}$$

At the end of each month t, we measure Fscore, *fq*, for the latest fiscal after its announcement to form seven portfolios: low (F = 0, 1, 2), 3, 4, 5, 6, and high (F = 7, 8). Because extreme F scores are rare, we combine scores 0, 1, and 2 into the low portfolio and scores 7 and 8 into the high portfolio. We sort stocks based on *fq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.26 Ohlson O-score (Oscore)

We follow Ohlson (1980) to construct O-score as

$$\begin{aligned} \mathrm{O} = {} & -1.32 - 0.407\log(\mathrm{TA}) + 6.03\mathrm{TLTA} - 1.43\mathrm{WCTA} + 0.076\mathrm{CLCA} - 1.72\mathrm{OENEG} - \\ & 2.37\mathrm{NITA} - 1.83\mathrm{FUTL} + 0.285\mathrm{IN2} - 0.521\mathrm{CHIN} \qquad \text{(B.2.3)} \end{aligned}$$

in which TA is total assets (Wind balance sheet item "TOT_ASSETS"), TLTA denotes the leverage ratio defined as total debt (Wind balance sheet item "TOT_LIAB") divided by total assets, WCTA is working capital, which is measured as current assets minus current liabilities divided by total assets. CLCA is current liability (Wind balance sheet item "TOT_CUR_LIAB") divided by current assets (Wind balance sheet item "TOT_CUR_ASSETS"). OENEG is one if total liabilities (Wind balance sheet item "TOT_LIAB") exceeds total assets and zero otherwise.

NITA is net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") divided by total assets. FUTL is the fund provided by operations (Wind income statement item "EBITDA" minus interest expense, Wind income statement item "LESS_INT_EXP") divided by total liabilities. IN_2 is equal to one if net income is negative for the last two years and zero otherwise. CHIN is $(NI_s - NI_s - 1)/(|NI_s| + |NI_s - 1|)$, in which NI_s and $NI_s - 1$ are the net income for the current and prior years. We winsorize all nondummy variables on the right-hand side of equation (B.2.3) at the 1st and 99th percentiles of their distributions each year. At the end of June of each year *t*, we measure O-score for the fiscal year ending in calendar year t−1. We sort stocks based on Oscore at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.27 Quarterly O-score (oq)

We use quarterly accounting data to construct O-score as

$$\begin{aligned} O = {} & -1.32 - 0.407\log(\mathrm{TA}) + 6.03\mathrm{TLTA} - 1.43\mathrm{WCTA} + 0.076\mathrm{CLCA} - 1.72\mathrm{OENEG} - \\ & 2.37\mathrm{NITA} - 1.83\mathrm{FUTL} + 0.285\mathrm{IN2} - 0.521\mathrm{CHIN} \end{aligned} \qquad \text{(B.2.4)}$$

in which TA is total assets (Wind balance sheet item "TOT_ASSETS"), TLTA denotes the leverage ratio defined as total debt (Wind balance sheet item "TOT_LIAB") divided by total assets, WCTA is working capital, which is measured as current assets minus current liabilities divided by total assets. CLCA is current liability (Wind balance sheet item "TOT_CUR_LIAB") divided by current assets (Wind balance sheet item "TOT_CUR_ASSETS"). OENEG is one if total liabilities (Wind balance sheet item "TOT_LIAB") exceeds total assets and zero otherwise. NITA is net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") divided by total assets. FUTL is the fund provided by operations (Wind income statement item "EBITDA" minus interest expense, Wind income statement item "LESS_INT_EXP") divided by total liabilities. IN_2 is equal to one if net income is negative for the current quarter and four quarters ago, and zero otherwise. CHIN is $(NI_s - NI_s - 1)/(|NI_s| + |NI_s - 1|)$, in which NI_s and $NI_s - 1$ are the net income for the current quarter and four quarters ago. We winsorize all nondummy variables on the right-hand side of equation (B.2.4) at the 1st and 99th percentiles of their distributions each year. At the end of each month t, we measure O-score for the latest fiscal after its announcement. We sort stocks based on *oq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.1.28 Altman's Z-score (Zscore)

We follow Altman (1968) to construct the Z-score as:

$$Z = 1.2\mathrm{WCTA} + 1.4\mathrm{RETA} + 3.3\mathrm{EBITTA} + 0.6\mathrm{METL} + \mathrm{SALETA} \qquad \text{(B.2.5)}$$

in which WCTA is working capital (current assets, Wind balance sheet item "TOT_CUR_ASSETS" minus current liabilities, Wind balance sheet item "TOT_CUR_LIAB") divided by total assets (Wind balance sheet item "TOT_ASSETS"), RETA is retained earnings divided by total assets, EBITTA is earnings before interest and taxes divided by total assets, METL is the market equity (at fiscal year end) divided by total liabilities (Wind balance sheet item "TOT_LIAB"), and SALETA is sales (operating revenue, Wind income statement item "OPER_REV") divided by total assets. We winsorize all nondummy variables above at the 1^{st} and 99^{th} percentiles of their distributions each year. At the end of June of each year t, we measure Z-score for the fiscal year ending in calendar year t−1. We sort stocks based on Zscore at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.1.29 Quarterly Z-score (zq)
We use quarterly accounting data to construct the Z-score as:

$$Z = 1.2\text{WCTA} + 1.4\text{RETA} + 3.3\text{EBITTA} + 0.6\text{METL} + \text{SALETA} \quad \text{(B.2.6)}$$

in which WCTA is working capital (current assets, Wind balance sheet item "TOT_CUR_ASSETS" minus current liabilities, Wind balance sheet item "TOT_CUR_LIAB") divided by total assets (Wind balance sheet item "TOT_ASSETS"), RETA is retained earnings divided by total assets, EBITTA is earnings before interest and taxes divided by total assets, METL is the market equity (at the end of each month) divided by total liabilities (Wind balance sheet item "TOT_LIAB"), and SALETA is sales (operating revenue, Wind income statement item "OPER_REV") divided by total assets. We winsorize all nondummy variables above at the 1st and 99th percentiles of their distributions each year. At the end of each month t, we measure Z-score, zq, for the latest fiscal quarter after its announcement. We sort stocks based on *zq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2 Value

B.2.2.1 Book-to-Market Equity (bm)
Basu (1983) finds the book-to-market effect; the book-to-market ratio has a positive relation with stock returns. Fama and French (1992) confirm that the book-to-market ratio is positively related to stock returns. Following Fama and French (1992), the book-to-market equity ratio is measured as the book value of equity, plus balance sheet deferred taxes if available, minus the book value of preferred stock divided by fiscal-year-end market capitalization based on total shares (It is also used for other ratios below). At the end of June of year *t*, we estimate the book-to-market ratio, *bm*, as total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus preferred shares (book value is 1 in China, zero if missing) for the fiscal year ending in calendar year *t*-1 divided by market equity at the end of December in year *t*-1. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *bm* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.2 Book-to-June-end Market Equity (bmj)
Following Asness and Frazzini (2013), at the end of June of each year *t*, we estimate the book-to-June-end market equity ratio, *bmj*, as total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus the preferred stock shares (book value is 1 in China, zero if missing) for the fiscal year ending in calendar year *t*-1 divided by the market equity at the end of June in year *t*. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *bmj* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.3 Quarterly Book-to-Market Equity (bmq)

At the end of each month *t*, we sort stocks into decile portfolios based on the quarterly book-to-market equity ratio, *bmq*, which is total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_INCL_MIN_INT") minus the preferred stock shares (book value is 1 in China, zero if missing) for the most recent quarter after its announcement date divided by the market equity at the end of month *t*. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *bmq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2.4 Liabilities-to-Market Equity (dm)
According to Bhandari (1988), at the end of June of year *t*, we estimate the debt-to-market equity ratio, *dm*, as total liabilities (Wind balance sheet item "TOT_LIAB") for the fiscal year ending in calendar year *t*-1 divided by the market equity at the end of December in year *t*-1. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *dm* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.5 Quarterly Liabilities-to-Market Equity (dmq)
At the end of each month *t*, we sort stocks into decile portfolios based on the quarterly debt-to-market equity ratio, *dmq*, which is total liabilities (Wind balance sheet item "TOT_LIAB") for the most recent quarter after its announcement date divided by the market equity at the end of month *t*. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *dmq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2.6 Assets-to-Market Equity (am)
As specified in Fama and French (1992), the assets-to-market ratio is equal to total assets divided by market capitalization. At the end of June of year *t*, we estimate the assets-to-market ratio, *am*, as total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-1 divided by the market equity at the end of December in year *t*-1. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *am* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.7 Quarterly Assets-to-Market Equity (amq)
At the end of each month *t*, we sort stocks into decile portfolios based on the quarterly assets-to-market equity ratio, *amq*, which is total assets (Wind balance sheet item "TOT_ASSETS") for the most recent quarter after its announcement date divided by the market equity at the end of month *t*. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *amq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2.8 Earnings-to-Price Ratio (ep)
Basu (1977) finds that the P/E ratio is negatively related to stock returns at short legs. Lakonishok, Shleifer and Vishny (1994) find that the earnings-to-price ratio (E/P) has a positive relation to stock returns. Following Basu (1977), the earnings-to-price ratio, *ep*, is income before extraordinary items divided by market equity. At the end of June of year *t*, we estimate *ep* as net profit (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus

nonrecurrent gains/losses ("PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP") for the fiscal year ending in calendar year t-1 divided by the market equity at the end of December in year t-1. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *ep* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.9 Quarterly Earnings-to-Price Ratio (epq)
At the end of each month t, we sort stocks into decile portfolios based on the quarterly earnings-to-price ratio, epq, which is net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses ("PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP") for the most recent quarter after its announcement date divided by the market capitalization at the end of month t. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on epq at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2.10 Cash Flow to Price (cfp)
At the end of June of each year t, we estimate the cash flow to price ratio, *cfp*, as the net change in cash or cash equivalents between two most recent cash flow statement (WIND cash flow statement item "NET_INCR_CASH_CASH_EQU") for the fiscal year ending in calendar year t-1 divided by the market equity at the end of December in year t-1. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *cfp* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.11 Quarterly Cash Flow to Price (cfpq)
At the end of June of each year t, we the cash flow to price ratio, *cfpq*, as the net change in cash or cash equivalents between two most recent cash flow statement (WIND cash flow statement item "NET_INCR_CASH_CASH_EQU") for the most recent quarter after its announcement date divided by the market capitalization at the end of month t. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *cfpq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2.12 5-year Sales Growth Rank (sr)
Following Lakonishok, Shleifer, and Vishny (1994), we measure the 5-year growth rank, *sr*, at the end of June in year t as the weighted average of the annual sales growth ranks for the prior 5 years: $\sum_{j=1}^{5}(6-j)*Rank(t-j)$. Only firms with data for all five prior years are used to determine the annual growth ranks, and we exclude firms with nonpositive sales. We sort stocks based on *sr* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.13 Enterprise Multiple (em)
Following Loughran and Wellman (2011), we estimate the enterprise multiple, *em*, as the enterprise value divided by operating income before depreciation (WIND income statement "OPER_PROFIT") . Enterprise value is estimated as market equity plus total debt (WIND balance sheet item "TOT_LIAB") plus the preferred stock shares (0, if missing) minus cash and short-term investments (WIND balance sheet "MONETARY_CAP") for the fiscal year ending in

calendar year t-1. The market equity is unadjusted close price multiplying by total shares at the end of December of year t-1. We sort stocks based on *em* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.14 Quarterly Enterprise Multiple (emq)

The quarterly enterprise multiple (*emq*) is estimated as enterprise value divided by operating income before depreciation (WIND income statement "OPER_PROFIT") for the most recent quarter after its announcement date. Enterprise value is market equity plus total debt (WIND balance sheet item "TOT_LIAB") plus the preferred stock shares minus cash and short-term investments (WIND balance sheet "MONETARY_CAP"). The market equity is unadjusted close price multiplying by total shares at the end of each month t. We sort stocks based on *emq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2.15 Sales-to-Price Ratio (sp)

Barbee, Jr. and Raines (1996) find that the sales-to-price ratio has a positive relationship with stock returns. At the end of June of year *t*, we estimate the sales to price ratio, *sp*, estimated as operating revenue (Wind income statement item "OPER_REV") for the fiscal year ending in calendar year *t*-1 divided by the market equity at the end of December in year *t*-1. The market equity is unadjusted close price multiplying by total shares at the end of December of year t-1. We sort stocks based on *sp* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.16 Quarterly Sales-to-Price Ratio (spq)

At the end of each month *t*, we estimate the quarterly sales-to-price ratio, *spq*, as quarterly operating revenue (Wind income statement item "OPER_REV") for the latest fiscal quarter after its announcement date divided by the market capitalization at the end of month *t*. The market equity is unadjusted close price multiplying by total shares. We sort stocks based on *spq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2.17 Operating Cash Flow to Price Ratio (ocfp)

Desai, Rajgopal and Venkatachalam (2004) show that value (glamour) stocks, characterized by low (high) past sales growth, high (low) book-to-market (B/M), high (low) earnings-to-price (E/P), and high (low) cash flow-to-price (C/P), are known to earn positive (negative) future abnormal returns. Following Desai, Rajgopal and Venkatachalam (2004), at the end of June of year *t*, we estimate the operating cash flow to price ratio, *ocfp*, as operating cash flows (Wind cash flow statement item "NET_CASH_FLOWS_OPER_ACT") for the fiscal year ending in calendar year *t*-1 divided by the market equity at the end of December in year *t*-1. We sort stocks based on *ocfp* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.18 Quarterly Operating Cash Flow to Price Ratio (ocfpq)

At the end of each month *t*, we estimate the quarterly operating cash flow to price ratio, *ocfpq*, as operating cash flows (Wind cash flow statement item "NET_CASH_FLOWS_OPER_ACT") for the latest fiscal quarter after its announcement date divided by the market capitalization at the end of month *t*. We sort stocks based on *ocfpq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.2.19 Liabilities-to-Book Equity (de)

Bhandari (1988) find that expected common stock returns are positively related to the ratio of debt (noncommon equity liabilities) to equity. At the end of June of year *t*, we estimate the debt-to-equity ratio, *de*, as total liabilities (Wind balance sheet item "TOT_LIAB") for the fiscal year ending in calendar year *t*-1 divided by total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus the preferred stock shares (0 if missing) for the fiscal year ending in calendar year *t*-1. We sort stocks based on *de* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.20 Intangible Return (ir)

At the end of June of each year t, we compute the intangible return, *ir*, as the residual from running the cross-sectional regression of each firm's past 5-year stock return on its 5-year lagged log book-to-market and 5-year log book return:

$$\mathrm{r(t-5,t)} = \gamma_0 + \gamma_1 \mathrm{bm_{t-5}} + \gamma_2 \mathrm{r^B(t-5,t)} + \mu_t \qquad \text{(B.2.7)}$$

in which r(t-5,t) is the past 5-year log stock return from the end of year t-6 to the end of t-1, $\mathrm{bm_{t-5}}$ is the 5-year lagged log book-to-market, and $\mathrm{r^B(t-5,t)}$ is the 5-year log book return. The 5-year lagged log book-to-market $\mathrm{bm_{t-5}} = \log\left(\frac{B_{t-5}}{M_{t-5}}\right)$, in which B_{t-5} is thB book equity for the fiscal year ending in calendar year t-6 and M_{t-5} is the market equity at the end of December of year t-6. We compute the 5-year log book return as $\mathrm{r^B(t-5,t)} = \log\left(\frac{B_t}{B_{t-5}}\right) + \sum_{s=t-5}^{t-1}(r_s - \log\left(\frac{P_s}{P_{s-1}}\right))$, in which B_t is the book equity for the fiscal year ending in calendar year t-1, r_s is the stock return from the end of year s-1 to the end of year s, and P_s is the stock price per share at the end of year s. Book equity is total shareholders' equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") minus the preferred stock shares (book value is 1 in China, zero if missing). We sort stocks based on *ir* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.21 Enterprise Book-to-Price (ebp) and Net Debt-to-Price (ndp)

Following Penman, Richardson and Tuna (2007), we measure enterprise book-to-price, *ebp*, as the ratio of the book value of net operating assets (net debt plus book equity) to the market value of net operating assets (net debt plus market equity) for the fiscal year ending in calendar year *t*-1. The net debt-to-price, *ndp*, is the ratio of net debt for the fiscal year ending in calendar year *t*-1 to the market equity at the end of December in year *t*-1. Net debt is financial liabilities minus financial assets. We measure financial liabilities as the sum of long-term debt (long-term borrow, Wind balance sheet item "LT_BORROW"), debt in current liabilities (short-term borrow, Wind balance sheet item "ST_BORROW"), carrying value of preferred stock (shares of preferred stocks, 0 if missing). We measure financial assets as cash and short-term investments (Wind balance sheet item "MONETARY_CAP"). We sort stocks based on *ebp or ndp* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.2.22 Quarterly Enterprise Book-to-Price (ebpq) and Quarterly Net Debt-to-Price

(ndpq)
We measure the quarterly enterprise book-to-price ratio, *ebpq*, as the ratio of the book value of net operating assets (net debt plus book equity) to the market value of net operating assets (net debt plus market equity) for the latest fiscal quarter after its announcement date. The net debt-to-price, *ndp*, is the ratio of net debt for the latest fiscal quarter after its announcement date to the market equity at the end of each month t. Net debt is financial liabilities minus financial assets. We measure financial liabilities as the sum of long-term debt (long-term borrow, Wind balance sheet item "LT_BORROW"), debt in current liabilities (short-term borrow, Wind balance sheet item "ST_BORROW"), carrying value of preferred stock (shares of preferred stocks, 0 if missing). We measure financial assets as cash and short-term investments (Wind balance sheet item "MONETARY_CAP"). We sort stocks based on *ebpq or ndpq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.3 Investment

B.2.3.1 Abnormal Corporate Investment (aci)
At the end of June of year *t,* we measure abnormal corporate investment, *aci*, as $\frac{\mathrm{Ce}_{t-1}}{\left[\frac{Ce_{t-2}+Ce_{t-3}+Ce_{t-4}}{3}\right]}-1$, in which Ce_{t-j} is capital expenditure (Wind cash flow statement item "CASH_PAY_ACQ_CONST_FIOLTA") scaled by sales (Wind income statement item "OPER_REV") for the fiscal year ending in calendar year *t-j.* We exclude firms with negative sales. Note, there is no exact capital expenditure in Chinese financial statement. It is in the fixed asset item of balance sheet. We use the cash flows paid out for purchasing fixed assets, intangible assets, and other long-term investment. We sort stocks based on *aci* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.2 Investment-to-Assets (ag)
Cooper, Gulen and Schill (2008) find companies that grow their total asset more earn lower subsequent returns. At the end of June of year *t*, we investment-to-assets, *ag*, as the annual growth in total assets (Wind balance sheet item "TOT_ASSETS") from the fiscal year ending in calendar year *t*-2 to the fiscal year ending in calendar year *t*-1. We sort stocks based on *ag* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.3 Quarterly Investment-to-Assets (agq)
We measure quarterly investment-to-assets, *agq*, as quarterly total assets (Wind balance sheet item "TOT_ASSETS") divided by four-quarter-lagged total assets minus one for the latest fiscal quarter after its announcement date. We sort stocks based on *agq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.3.4 Changes in PPE and inventory-to-assets (dpia)
Changes in PPE and Inventory-to-assets, dpia, is defined as the annual change in gross property, plant, and equipment (fixed assets, Wind balance sheet item "FIX_ASSETS") plus the annual change in inventory (Wind balance sheet item "INVENTORIES") scaled by 1-year-lagged total assets (Wind balance sheet item "TOT_ASSETS"). We sort stocks based on *dpia* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months

(n=12).

B.2.3.5 Net Operating Assets, Changes in Net Operating Assets (noa, dnoa)
We measure net operating assets as operating assets minus operating liabilities. Operating assets are total assets (Wind balance sheet item "TOT_ASSETS") minus cash and short-term investment (Wind balance sheet item "MONETARY_CAP"). Operating liabilities are total assets (Wind balance sheet item "TOT_ASSETS") minus debt included in current liabilities (Wind balance sheet item "ST_BORROW", zero if missing), minus long-term debt (Wind balance sheet item "LT_BORROW", zero if missing), minus minority interests (Wind balance sheet item "MINORITY_INT"), minus shares of preferred stocks (zero if missing), and minus common equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT"). *noa* is net operating assets for the fiscal year ending in calendar year *t*-1 scaled by 1-year lagged total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-2. Changes in net operating assets, *dnoa*, is the annual change in net operating assets from the fiscal year ending in calendar year *t*-2 to the fiscal year ending in calendar year *t*-1 scaled by 1-year lagged total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-2. We sort stocks based on *noa, dnoa* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.6 x-year Investment Growth (ig, 2ig, 3ig)
At the end of June of each year t, we measure investment growth, *ig*, as the growth rate in capital expenditure (Wind cash flow statement item "CASH_PAY_ACQ_CONST_FIOLTA") from the fiscal year ending in calendar year t-x-1 to the fiscal year ending in year t-1 (x=1, 2, and 3). We sort stocks based on *ig, 2ig, 3ig* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.7 Net Share Issues (nsi)
Pontiff and Woodgate (2008) show that post-1970, share issuance exhibits a strong and negative cross-sectional ability to predict stock returns. We calculate share change in China as the annual percentage change in total A shares, including non-tradable shares. The number of shares outstanding adjusted for splits and other events are:

$$AdjustedShare_t = SharesOutstanding_t / TotalFactor_t \tag{B.2.8}$$

At the end of June of each year *t*, we measure net stock issues, *nsi*, as the natural logarithm of the ratio of split-adjusted total shares for the fiscal year ending in calendar year *t*-1 to split-adjusted total shares for the fiscal year ending in calendar year *t*-2. Split-adjusted total shares are calculated as total shares multiplying by the adjustment factor. We sort stocks based on *nsi* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.8 Composite Equity Issuance (cei)
Following Daniel and Titman (2006), at the end of June of year *t*, we measure composite equity issuance, *cei*, as measured as the log growth rate in market equity not attributing to stock returns from year *t*-5 to year *t*, $cei = \log\left(\frac{\mathrm{ME_t}}{ME_{t-5}}\right) - r(t-5, t)$, where $r(t-5, t)$ is the cumulative log stock returns from the last trading day of June in year *t*-5 to the last trading day of June in year *t*, and ME is the market equity on the last trading day of June in year *t*. We sort stocks based on *cei*

at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.9 Composite Debt Issuance (cdi)

Following Lyandres, Sun and Zhang (2007), at the end of June of each year *t*, we measure composite debt issuance, *cdi*, as the log growth rate of total liabilities (Wind balance sheet item "TOT_LIAB") from the fiscal year ending in calendar year *t*-6 to the fiscal year ending in year *t*-1. We sort stocks based on *cdi* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.10 Inventory Growth (ivg)

Following Belo and Lin (2011), at the end of June of each year *t*, we measure inventory growth, *ivg*, as the annual growth rate in inventory (Wind balance sheet item "INVENTORIES") from the fiscal year ending in year *t*-2 to the fiscal year ending in year *t*-1. We sort stocks based on *ivg* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.11 Inventory Change (ivchg)

Following Thomas and Zhang (2002), at the end of June of each year *t*, we measure inventory change *ivchg*, as the annual change in inventory (Wind balance sheet item "INVENTORIES") from the fiscal year ending in year *t*-2 to the fiscal year ending in year *t*-1 scaled by average total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-2 and t-1 . We sort stocks based on *ivchg* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.12 Operating Accruals (oacc)

We measure operating accruals, *oacc*, at the end of June of each year *t* as net profit (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") for the fiscal year ending in calendar year *t-1* minus operating cash flow (Wind cash flow statement item "NET_CASH_FLOWS_OPER_ACT") for the fiscal year ending in calendar year *t-1* scaled by one-year lagged total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-2. We sort stocks based on *oacc* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.13 Total Accruals (tacc)

We measure total accruals, *tacc*, as net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus cash flows (Wind cash flow statement item "NET_INCR_CASH_CASH_EQU") for the fiscal year ending in calendar year *t*-1 scaled by total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year t-2. We sort stocks based on *tacc* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.14 Changes in Net Noncash Working Capital, Current Operating Assets, and Current Operating Liabilities (dwc, dcoa, dcol)

Net noncash working capital (wc) is equal to current operating assets (coa) minus current operating liabilities (col). Current operating assets is equal to current assets (Wind balance sheet item “TOT_CUR_ASSETS”) minus cash and short-term investments (Wind balance sheet item “MONETARY_CAP”). Current operating liabilities are equal to current liabilities (Wind balance sheet item “TOT_CUR_LIAB”) minus debt in current liabilities (Wind balance sheet item “ST_BORROW”, zero if missing). At the end of June of each year t, we measure *dwc*, *dcoa*, and *dcol* as annual changes in net noncash working capital, current operating assets, and current operating liabilities from the fiscal year ending in calendar year t-2 to the fiscal year ending in calendar year t-1 scaled by 1-year-lagged total assets (Wind balance sheet item “TOT_ASSETS”) for the fiscal year ending in calendar year t-2. We sort stocks based on *dwc, dcoa, dcol* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.15 Changes in Net Noncurrent Operating Assets, in Noncurrent Operating Assets, in Noncurrent Operating Liabilities (dnco, dnca, dncl)

Net noncurrent operating assets (*nco*) is equal to noncurrent operating assets minus noncurrent operating liabilities. Noncurrent operating assets (*nca*) is total assets (Wind balance sheet item “TOT_ASSETS”) minus current assets (Wind balance sheet item “TOT_CUR_ASSETS”) minus long-term investments (zero if missing). In China, there is long-term investments item in financial statement. We use the sum of held-to-maturity investment (Wind balance sheet item “HELD_TO_MTY_INVEST”), long-term equity investment (Wind balance sheet item “LONG_TERM_EQY_INVEST”), investment in real estate (Wind balance sheet item “INVEST_REAL_ESTATE”), and fixed deposit (Wind balance sheet item “TIME_DEPOSITS”). Noncurrent operating liabilities (ncl) is total liabilities (Wind balance sheet item “TOT_LIAB”) minus current liabilities (Wind balance sheet item “TOT_CUR_LIAB”) minus long-term debt (Wind balance sheet item “LT_BORROW”, zero if missing). At the end of June of each year t, we measure *dnco*, *dnca*, and *dncl* are annual changes in net noncurrent operating assets, noncurrent operating assets, and noncurrent operating liabilities from the fiscal year ending in calendar year t-2 to the fiscal year ending in calendar year t-1 scaled by 1-year lagged total assets (Wind balance sheet item “TOT_ASSETS”) for the fiscal year ending in calendar year t-2. We sort stocks based on *dnco, dnca, dncl* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.16 Changes in Net Financial Assets, in Short-Term Investments, in Long-Term Investments, in Financial Liabilities, and in Book Equity (dfin, dsti, dlti, dfnl, dbe)

Net financial assets are the difference between financial assets and financial liabilities. We measure financial assets as short-term investments (Wind balance sheet item “MONETARY_CAP”) plus long-term investment (zero if missing). In China, there is long-term investments item in financial statement. We use the sum of held-to-maturity investment (Wind balance sheet item “HELD_TO_MTY_INVEST”), long-term equity investment (Wind balance sheet item “LONG_TERM_EQY_INVEST”), investment in real estate (Wind balance sheet item “INVEST_REAL_ESTATE”), and fixed deposit (Wind balance sheet item “TIME_DEPOSITS”). Financial liabilities is the sum of long-term debt (Wind balance sheet item “LT_BORROW”, zero if missing), debt in current liabilities (Wind balance sheet item “ST_BORROW”, zero if missing), and preferred stock shares. At the end of June of each year t, we measure *dfin* is the annual change in net financial assets, short-term investments, long-term investments, financial liabilities from the fiscal year ending in calendar year t-2 to the fiscal year ending in calendar year t-1 scaled by 1-year lagged total assets for the fiscal year ending in calendar year t-2. We measure *dbe* as the change in book equity (Wind balance sheet item "TOT_SHRHLDR_EQY_EXCL_MIN_INT") for the fiscal year ending in calendar year t-1 scaled by 1-year lagged total assets for the fiscal year ending in calendar year t-2. We sort stocks based on *dfin, dsti, dlti, dfnl, dbe* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.17 Discretionary Accruals (dacc)
We measure discretionary accruals, dacc, from

$$\frac{Oa_{i,t}}{A_{i,t-1}} = \alpha_1 \frac{1}{A_{it-1}} + \alpha_2 \frac{dSALE_{it} - dREC_{it}}{A_{it-1}} + \alpha_3 \frac{PPE_{i,t}}{A_{it-1}} + e_{it} \qquad \text{(B.2.9)}$$

in which $\mathrm{Oa}_{i,t}$ is operating accruals for firm i (see Appendix 2.3.11), A_{it-1} is total assets (Wind balance sheet item "TOT_ASSETS") at the end of year t-1, $dSALE_{it}$ is the annual change in sales (operating revenue, Wind income statement item "OPER_REV") from year t-1 to t, $dREC_{it}$ is the annual change in net account receivables (WIND balance sheet item "ACCT_RCV") from year t-1 to t, and $PPE_{i,t}$ is gross property, plant, and equipment (fixed assets, Wind balance sheet item "FIX_ASSETS") at the end of year t. We winsorize the variables at the right hand side of equation (B.2.9) at the 1^{st} and 99^{th} percentiles of their distributions each year.

At the end of June of each year t, we run a cross-sectional regression (B.2.9) using the firms in each Shenwanhongyuan industry for each fiscal year ending in t-1. We require at least 6 firms for each regression. The discretionary accrual for stock *i* is defined as the residual from the regression, e_{it}. We sort stocks based on *dacc* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.18 Percent Operating Accruals (poacc)
At the end of June of each year t, we measure percent operating accruals, *poacc*, as operating accruals in (2.3.11) or the fiscal year ending in calendar year t-1 divided by the absolute value of net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") for the fiscal year ending in calendar year t-1. We sort stocks based on *poacc* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.19 Percent Total Accruals (ptacc)
At the end of June of each year t, we measure percent total accruals, *ptacc*, as total accruals in (2.3.12) for the fiscal year ending in calendar year t-1 divided by the absolute value of net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") for the fiscal year ending in calendar year t-1. We sort stocks based on *ptacc* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.20 Percent Discretionary Accruals (pdacc)
At the end of June of each year t, we calculate percent discretionary accruals, *pdacc*, as the discretionary accruals (see Appendix 2.3.16) for the fiscal year ending in calendar year t-1 multiplied with total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in t-2 scaled by the absolute value of net income (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") for the fiscal year ending in t-1. We sort stocks based

on *pdacc* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.3.21 Quarterly Current Asset Growth, Non-Current Asset Growth (cagq, ncagq)
Total assets are equal to current assets plus non-current assets. The total asset growth can be decomposed into current asset growth and non-current asset growth. We measure quarterly current asset growth, *cagq*, as quarterly current assets (Wind balance sheet item "TOT_CUR_ASSETS") minus four-quarter-lagged current assets divided by four-quarter-lagged total assets. We measure quarterly asset growth, *ncagq*, as quarterly non-current assets (Wind balance sheet item "TOT_NON_CUR_ASSETS") minus four-quarter-lagged non-current assets, and then divided by four-quarter-lagged total assets. At the end of each month *t*, we measure quarterly *cagq* and *ncagq* for the latest fiscal quarter after its announcement date, sort stocks based on *ebpq or ndpq*, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.3.22 Quarterly Cash Growth, Fixed Asset Growth, Non-Cash Current Assets Growth, Other Asset Growth (cashgq, fagq, nccgq, oagq)
According to Cooper, Gulen and Schill (2008), the asset growth decomposition is as follows:
Total asset growth = cash growth + noncash current asset growth + property, plant, and equipment growth + other assets growth. We measure quarterly cash growth, c*ashgq*, as quarterly cash (Wind balance sheet item "MONETARY_CAP") minus four-quarter-lagged cash and then divided total assets four quarters ago. We measure quarterly fixed asset growth, *fagq*, as quarterly fixed assets (Wind balance sheet item "FIX_ASSETS") minus four-quarter-lagged fixed assets and then divided total assets four quarters ago. We measure quarterly non-cash current asset growth, *nccagq*, as annual change in quarterly current assets (Wind balance sheet item "TOT_CUR_ASSETS") minus cash (Wind balance sheet item "MONETARY_CAP") divided by four-quarter-lagged total assets. Other assets are equal to non-current assets (Wind balance sheet item "TOT_NON_CUR_ASSETS") minus fixed assets (Wind balance sheet item "FIX_ASSETS"). We measure quarterly other asset growth, *oagq*, as annual change in quarterly other assets divided by four-quarter-lagged total assets. We sort stocks based on *cashq, fagq, nccgq, oagq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4 Other Anomalies

B.2.4.1Advertising Expense-to-Market (adm)
We measure advertising expense-to-market, *adm*, as advertising expenses (Wind income statement item "LESS_SELLING_DIST_EXP") for the fiscal year ending in calendar year t-1 divided by the market equity at the end of December of t-1. There is no exact item recording advertising expense in China, but it belongs to selling expense in income statement. We sort stocks based on *adm* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.2 Growth in Advertising Expense (gad)
At the end of June of each year t, we measure growth in advertising expenses, *gad*, as the growth rate of advertising expenses (Wind income statement item "LESS_SELLING_DIST_EXP") from the fiscal year ending in calendar year t-2 to the fiscal year ending in calendar year t-1. We sort

stocks based on *gad* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.3 R&D Expense to Market Equity (rdm)
We follow Chan, Lakonishok and Sougiannis (2011) to calculate the R&D expense to market equity, *rdm*, as R&D expense for the fiscal year ending in calendar year t-1 to market equity at the end of December of t-1. Since R&D expense is not available in the financial statement in China, we use administrative expenses (Wind income statement item "LESS GERL ADMIN_EXP") as a proxy for R&D expense following Chen, Kim, Yao and Yu (2010). We only keep firms with positive R&D expense. We sort stocks based on *rdm* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.4 Quarterly R&D Expense to Market Equity (rdmq)
At the end of each month *t*, we measure the quarterly R&D expense to market ratio, *rdmq*, as quarterly R&D expense (administrative expenses, Wind income statement item "LESS GERL ADMIN_EXP") for the latest fiscal quarter after its announcement divided by the market equity at the end of month *t*. We only keep firms with positive R&D expense. We sort stocks based on *rdmq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.5 R&D Expense to Sales Ratio (rds)
Following Chan, Lakonishok and Sougiannis (2011), at the end of June of each year *t*, we measure the R&D expense-to-sales ratio, *rds*, as R&D expense (administrative expenses, Wind income statement item "LESS GERL ADMIN_EXP") for the fiscal year ending in calendar year *t*-1 divided by sales (operating revenue, Wind income statement item "OPER_REV") for the fiscal year ending in calendar year *t*-1. We only keep firms with positive R&D expense. We sort stocks based on *rds* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.6 Quarterly R&D Expense to Sales (rdsq)
At the end of each month *t*, we measure the quarterly R&D expense to sales ratio, *rdsq*, as quarterly R&D expense (administrative expenses, Wind income statement item "LESS GERL ADMIN_EXP") divided by quarterly sales (operating revenue, Wind income statement item "OPER_REV") for the latest fiscal quarter after its announcement date. We only keep firms with positive R&D expense. We sort stocks based on *ebpq* or *ndpq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.7 Operating Leverage (ol)
Following Novy-Marx (2011), we estimate the operating leverage, *ol*, as operating costs (Wind income statement item "LESS_OPER_COST") for the fiscal year ending in calendar year *t*-1 scaled by total assets (Wind balance sheet item "TOT_ASSETS") for the fiscal year ending in calendar year *t*-1. We sort stocks based on *ol* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.8 Quarterly Operating Leverage (olq)
At the end of each month *t*, we measure the quarterly operating leverage, *olq*, as quarterly

operating costs (Wind income statement item "LESS_OPER_COST") divided by total assets (Wind balance sheet item "TOT_ASSETS") for the latest fiscal quarter after its announcement. We sort stocks based on *ebpq or ndpq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.9 Hiring Rate (hn)
Following Belo, Lin and Bazdresch (2014), at the end of June of year t, the hiring rate (*hn*) is $(\mathrm{N}_{t-1} - N_{t-2})/(0.5N_{t-1} + 0.5N_{t-2})$, in which $\mathrm{N}_{\mathrm{t-j}}$ is the number of employees for the fiscal year ending in calendar year t-j. We exclude firms with zero *hn*. We sort stocks based on *hn* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.10 Firm Age (age)
Jiang, Lee and Zhang (2005) find that young firms earn lower returns than old firms. Following Jiang, Lee and Zhang (2005), firm age, *age*, is the number of months between the portfolio formation date and the firms' IPO date. We sort stocks based on *age* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.11 % Change in Sales minus % Change in Inventory (dsi)
According to Abarbanell and Bushee (1998), we define the %d(.) operator as the percentage change in the variable in the parentheses from its average over the prior two years, for example, %d(Sales)=[Sales(t)-E[Sales(t)]]/E[Sales(t)], in which E[Sales(t)]=[Sales(t-1)+Sales(t-2)]/2. *dsi* is %d(sales)-%d(inventory), in which sales is operating revenue (Wind income statement item "OPER_REV") for the fiscal year ending in calendar year t-1, and inventory is net inventory (WIND balance sheet item "INVENTORIES") for the fiscal year ending in calendar year t-1. We sort stocks based on *dsi* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.12 % Change in Sales minus % Change in Accounts Receivable (dsa)
Following Abarbanell and Bushee (1998), we define the %d(.) operator as the percentage change in the variable in the parentheses from its average over the prior two years, for example, %d(Sales)=[Sales(t)-E[Sales(t)]]/E[Sales(t)], in which E[Sales(t)]=[Sales(t-1)+Sales(t-2)]/2. *dsa* is %d(Sales)-%d(accounts receivable), in which sales is operating revenue (Wind income statement item "OPER_REV") for the fiscal year ending in calendar year t-1, and accounts receivables is total receivables (WIND balance sheet item "ACCT_RCV") for the fiscal year ending in calendar year t-1. We sort stocks based on *dsa* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.13 % Change in Gross Margin minus % Change in Sales (dgs)
Following Abarbanell and Bushee (1998), we define the %d(.) operator as the percentage change in the variable in the parentheses from its average over the prior two years, for example, %d(Sales)=[Sales(t)-E[Sales(t)]]/ E[Sales(t)], in which E[Sales(t)]=[Sales(t-1)+Sales(t-2)]/2. *dgs* is %d(gross margin)-%d(sales), in which sales is operating revenue (Wind income statement item "OPER_REV") for the fiscal year ending in calendar year t-1, and gross margin is sales minus cost of goods sold (operating profit, WIND

income statement item “OPER_PROFIT”) for the fiscal year ending in calendar year t-1. We sort stocks based on *dgs* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.14 % Change in Sales minus % Change in SG&A (dss)

Following Abarbanell and Bushee (1998), we define the %d(.) operator as the percentage change in the variable in the parentheses from its average over the prior two years, for example, %d(Sales)=[Sales(t)-E[Sales(t)]]/E[Sales(t)], in which E[Sales(t)]=[Sales(t-1)+Sales(t-2)]/2. *dss* is %d(Sales)-%d(SG&A), in which sales is operating revenue (Wind income statement item “OPER_REV”) for the fiscal year ending in calendar year t-1, and SG&A is selling, general, and administrative expenses (WIND income statement item “LESS_GERL_ADMIN_EXP” plus the item “LESS_SELLING_DIST_EXP”). We sort stocks based on *dss* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.15 Effective Tax Rate (etr)

Following Abarbanell and Bushee (1998), we measure effective tax rate, *etr*, as

$$\mathrm{etr(t)} = \left[\frac{\mathrm{TaxExpense(t)}}{\mathrm{EBT(t)}} - \frac{1}{3}\sum_{\tau=1}^{3}\frac{TaxExpense(t-\tau)}{EBT(t-\tau)}\right] * \mathrm{dEPS(t)} \tag{B.2.9}$$

in which TaxExpense(t) is total income taxes (WIND income statement “INC_TAX”) paid in year t, EBT(t) is EBIT (WIND income statement “EBIT”) minus interest expense (WIND income statement “LESS_INT_EXP”), and *dEPS* is the change in split-adjusted earnings per share between years t−1 and t, deflated by stock close price at the end of t−1. Earnings per share is net profit (Wind income statement item “NET_PROFIT_EXCL_MIN_INT_INC”) minus nonrecurrent gains/losses (“PLUS_NON_OPER_REV” minus “LESS_NON_OPER_EXP”) for the fiscal year ending in calendar year *t*-1 divided by the market equity at the end of December in year *t*-1. The market equity is unadjusted close price multiplying by total shares. At the end of June of each year t, we measure *etr* for the fiscal year ending in calendar year t-1. We sort stocks based on *etr* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.16 Labor Force Efficiency (lfe)

Following Abarbanell and Bushee (1998), we measure labor force efficiency, *lfe*, as

$$\mathrm{Lfe(t)} = \left[\frac{\mathrm{Sales(t)}}{Employee(t)} - \frac{Sales(t-1)}{Employee(t-1)}\right] / \frac{Sales(t-1)}{Employee(t-1)} \tag{B.2.10}$$

in which Sales(t) is net sales (Wind income statement item “OPER_REV”) in year t, and Employees(t) is the number of employees. At the end of June of each year t, we measure *lfe* for the fiscal year ending in calendar year t−1. We sort stocks based on *lfe* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.17 Tangibility (tan)

We measure tangibility, *tan*, as cash holdings (Wind balance sheet item “MONETARY_CAP”) + 0.715*accounts receivable (WIND balance sheet item “ACCT_RCV”) +0.547*inventory (WIND balance sheet item “INVENTORIES”) + 0.535*gross property, plant, and equipment, all scaled by total assets (Wind balance sheet item “TOT_ASSETS”). Since there is no item for gross property, plant, and equipment and they are in fixed assets (Wind balance sheet item "FIX_ASSETS"). We use fixed assets as a proxy for gross property, plant, and equipment. At the end of June of each year t, we measure *tan* for the fiscal year ending in calendar year t-1. We sort stocks based on *tan* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.18 Quarterly Tangibility (tanq)

We measure quarterly tangibility, *tanq*, as cash holdings (Wind balance sheet item "MONETARY_CAP", zero if missing)+0.715*accounts receivable (WIND balance sheet item "ACCT_RCV", zero if missing)+0.547*inventory (WIND balance sheet item "INVENTORIES") + 0.535*gross property, plant, and equipment, all scaled by total assets (Wind balance sheet item "TOT_ASSETS"). Since there is no item for gross property, plant, and equipment and they are in fixed assets. We use fixed assets (Wind balance sheet item "FIX_ASSETS") as a proxy for gross property, plant, and equipment. At the end of each month *t*, we measure *tanq* for the latest fiscal quarter after its announcement data, sort stocks based on *tanq* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.19 Cash Flow Volatility (vcf)

Cash flow volatility, *vcf*, is the standard deviation of the ratio of operating cash flows (Wind cash flow statement item "NET_CASH_FLOWS_OPER_ACT") to sales (Wind income statement item "OPER_REV") during the past 16 quarters (a minimum of eight nonmissing quarters). At the end of each month t, we measure *vcf* for the latest fiscal quarter after its announcement date, sort stocks based on *vcf*, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.20 Cash to Assets (cta)

Palazzo (2012) implies a positive relation between expected equity returns and cash holdings (cash-to-assets). Following Palazzo (2012), we measure the cash-to-assets ratio, *cta*, as cash and cash equivalents (WIND balance sheet item "MONETARY_CAP") divided by total assets (Wind balance sheet item "TOT_ASSETS") for the latest fiscal quarter after its announcement date. We sort stocks based on *cta* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

.

B.2.4.21 Earnings Persistence, Earnings Predictability (eper, eprd)

Following Francis, LaFond, Olsson and Schipper (2004), earnings persistence and earnings predictability are from a first-order autoregressive model for annual earnings per share in the 10-year rolling window. Earnings per share are net profit (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses (Wind income statement item "PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP") scaled by total shares. At the end of June of each year t, we estimate the autoregressive model in the 10-year rolling window up to the fiscal year ending in calendar year t−1. Only firms with a complete 10-year history are included. *eper* is measured as the slope coefficient, and *eprd* is measured as the residual volatility. We sort stocks based on *eper, eprd* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.22 Earnings Smoothness (esm)

Following Francis, LaFond, Olsson and Schipper (2004), we measure earnings smoothness, *esm*, as the ratio of standard deviation of earnings scaled by 1-year lagged total assets to the standard deviation of cash flow from operations (Wind cash flow statement item "NET_CASH_FLOWS_OPER_ACT") scaled by 1-year lagged total assets (Wind balance sheet item "TOT_ASSETS"). Earnings are net profit (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses ("PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP"). At the end of June of each year t, we estimate *esm* over the 10-year rolling window up to the fiscal year ending in calendar year t−1. Only firms with a complete 10-year history are included. We sort stocks based on *esm* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.23 Value Relevance of Earnings (evr)

Following Francis, LaFond, Olsson and Schipper (2004), we measure value relevance of earnings, *evr*, as the R^2 from the following rolling-window regression:

$$\mathrm{R_{it}} = \delta_{i0} + \delta_{i1} EARN_{it} + \delta_{2t} dERARN_{it} + \epsilon_{it} \qquad \text{(B.2.11)}$$

In which $\mathrm{R_{it}}$ is firms i's 15-month stock return ending three months after the end of fiscal year. $EARN_{it}$ is net profit (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses ("PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP") for the fiscal year ending in t, scaled by the fiscal year-end market equity. $dERARN_{it}$ is the 1-year change in earnings scaled by the market equity. At the end of June of each year t, we calculate *evr* over the 10-year rolling window up to the fiscal year ending in calendar year t−1. Only firms with a complete 10-year history are included. We sort stocks based on *evr* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.24 Earnings Timeliness, Earnings Conservatism (etl, ecs)

Following Francis, LaFond, Olsson and Schipper (2004), we measure earnings timeliness, *etl*, and earnings conservatism, *ecs*, from the following rolling-window regression:

$$\mathrm{EARN_{it}} = \delta_{i0} + \alpha_{i1} NEG_{it} + \beta_{i1} R_{it} + \beta_{i2} NEG_{it} R_{it} + \epsilon_{it} \qquad \text{(B.2.12)}$$

in which $EARN_{it}$ is net profit (Wind income statement item "NET_PROFIT_EXCL_MIN_INT_INC") minus nonrecurrent gains/losses ("PLUS_NON_OPER_REV" minus "LESS_NON_OPER_EXP") for the fiscal year ending in calendar year *t*, scaled by the fiscal year-end market equity. R_{it} is firm *i*'s 15-month stock return ending three months after the end of fiscal year ending in calendar year t. NEG_{it} equals one if R_{it}<0, and zero otherwise. We measure *etl* as the R^2 and *ecs* as $(\beta_{i1} + \beta_{i2})/\beta_{i1}$ from the regression in (A2.12). At the end of June of each year t, we measure *etl* and *ecs* over the 10-year rolling window up to the fiscal year ending in calendar year t−1. Only firms with a complete 10-year history are included. We sort stocks based on *etl, ecs* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.25 Asset Liquidity (ala, alm)

We measure asset liquidity as cash + 0.75 × noncash current assets + 0.50 × tangible fixed assets, cash as cash and short-term investments (WIND balance sheet item "MONETARY_CAP"), noncash current assets as current assets (WIND balance sheet item "TOT_CUR_ASSETS") minus cash (WIND balance sheet item "MONETARY_CAP"), and tangible fixed assets as total assets (Wind balance sheet item "TOT_ASSETS") minus current assets (WIND balance sheet item "TOT_CUR_ASSETS"), minus goodwill (WIND balance sheet item "GOODWILL", zero if missing), and minus intangibles (WIND balance sheet item "INTANG_ASSETS", zero if missing). *ala* is asset liquidity scaled by 1-year-lagged total assets. *alm* is asset liquidity scaled by 1-year-lagged market value of assets. The market value of assets is total assets plus market equity minus book equity. At the end of June of each year t, we measure ala and alm for the fiscal year ending in calendar year t−1. We sort stocks based on *ala, alm* at the end of June of each year t, wait for 0 month (m=0), and hold portfolios over the next t+m+n months (n=12).

B.2.4.26 Quarterly Asset Liquidity (alaq, almq)

We measure asset liquidity as cash + 0.75 × noncash current assets + 0.50 × tangible fixed assets, cash as cash and short-term investments (Compustat annual item CHE) (WIND balance sheet item "MONETARY_CAP"), noncash current assets as current assets (WIND balance sheet item "TOT_CUR_ASSETS") minus cash (WIND balance sheet item "MONETARY_CAP"), and tangible fixed assets as total assets (Wind balance sheet item "TOT_ASSETS") minus current assets (WIND balance sheet item "TOT_CUR_ASSETS"), minus goodwill (WIND balance sheet item "GOODWILL", zero if missing), and minus intangibles (WIND balance sheet item "INTANG_ASSETS", zero if missing). *alaq* is asset liquidity scaled by 1-quarter-lagged total assets. *almq* is asset liquidity scaled by 1-quarter-lagged market value of assets. The market value of assets is total assets plus market equity minus book equity. At the end of each month t, we measure *alaq* and *almq* for the latest fiscal quarter after its announcement date, sort stocks based on *alaq* or *almq*, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.27 Standard Unexpected Earnings (sue)

Following Foster, Olsen and Shevlin (1984), we calculate standardized unexpected earnings, *sue*, as the annual change in split-adjusted quarterly earnings per share in quarter *t* from its value in quarter *t*-4 divided by the standard deviation of its change in quarterly earnings over the prior 8 quarters (at least 6 quarters are required) from quarter *t*-7 to quarter *t*. At the end of each month t, we estimate sue for the latest fiscal quarter after its announcement date, sort stocks based on *sue*, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.28 Revenue Surprises (rs)

Following Jegadeesh and Livnat (2006), the revenue surprise, *rs*, in quarter *t* is calculated as the annual change in revenue (operating revenue, Wind income statement item "OPER_REV") in quarter *t* divided by total shares from its value in quarter *t*-4 divided by the standard deviation of its change in quarterly revenue over the prior eight quarters (at least 6 quarters are required) from quarter *t*-7 to quarter *t*. At the end of each month t, we estimate *rs* for the latest fiscal quarter after its announcement date, sort stocks based on *rs*, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.29 Tax Expense Surprises (tes)

Thomas and Zhang (2011) find that seasonally differenced quarterly tax expense, a proxy for tax expense surprise, is related positively to future returns. According to Thomas and Zhang (2011), we measure the tax expense surprise, *tes*, as the annual percent change in total taxes (WIND cash flow statement item "PAY_ALL_TYP_TAX") from quarter *t*-4 to quarter *t*. At the end of each month t, we estimate *tes* for the latest fiscal quarter after its announcement, sort stocks based on *tes*, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.30 Revisions in Analyst Earnings Forecasts (re)

Following Chan, Jegadeesh and Lakonishok (1996), we measure earnings surprise, *re*, as the revision in analysts' forecasts of earnings. Because analysts' forecasts are not necessarily revised each month, we construct a 6-month moving average of past changes in analysts' forecasts:

$$\mathrm{re_{it}} = \sum_{\tau=1}^{6} \frac{f_{it-\tau} - f_{it-\tau-1}}{p_{it-\tau-1}} \quad \text{(B.2.13)}$$

where $f_{it-\tau}$ is the consensus mean forecast issued in month $\mathrm{t} - \tau$ for firm *i*'s current fiscal year earnings, and $p_{it-\tau-1}$ is the prior month's share price. We also adjust for any stock splits and require a minimum of four monthly forecast changes when constructing *re*. We sort stocks based on *re* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.31 Changes in Analyst Earnings Forecasts (def)

Following Hawkins, Chamberlin and Daniel (1984), we measure changes in analyst earnings forecasts, *def*, as

$$\mathrm{def} = (\mathrm{f_{it-1}} - f_{it-2})/(0.5|f_{it-1}| + 0.5|f_{it-2}|) \quad \text{(B.2.14)}$$

where f_{it-1} is the consensus mean forecast in which issued in month t−1 for firm *i*'s current fiscal year earnings. We sort stocks based on *def* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.32 Earnings Forecast–to-Price (efp)

Following Elgers, Lo and Pfeiffer (2001), we define analysts' earnings forecast-to-price, *efp*, as the consensus median forecasts for the current fiscal year (the forecasts within one year before annual earnings announcement date) divided by share price. We sort stocks based on *efp* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.33 Analyst Coverage (ana)

Following Elgers, Lo and Pfeiffer (2001), we measure analysts coverage, *ana*, as the number of analysts' earnings forecasts for the current fiscal year (the forecasts within one year before annual earnings announcement date). We sort stocks based on *ana* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

B.2.4.34 Dispersion in Analyst Forecasts (dis)

We measure dispersion in analyst earnings forecasts, *dis*, as the ratio of the standard deviation of earnings forecasts to the absolute value of the consensus mean forecast. We use the earnings forecasts for the current fiscal year (the forecasts within one year before annual earnings announcement date). We sort stocks based on *dis* at the end of each month t, wait for 0 month (m=0), and hold portfolios over the next [t+m+1, t+m+n] months (n=1, 6, 12).

C. Multiple Tests

C1. The procedure of Harvey and Liu (2020) bootstrap method

Suppose we have N strategies/anomalies and D time periods. We arrange the long-short time-series data into a D*N data matrix $\mathrm{X_0}$. Suppose one believes that a fraction, $\mathrm{p_0}$ of the N strategies/anomalies are true. This is to control type II error. The choice of $\mathrm{p_0}$ is inherently subjective. It is likely driven by both her previous experience and the data. For a given $\mathrm{p_0}$, we start by choosing $p_0 * N$ strategies that are believed to be true (true anomalies). We first rank the strategies by their t-statistics and then choose the top $p_0 * N$ with the highest t-statistics.

Step I: Each time, we bootstrap the time periods and create an alternative panel of Long-Short returns, $\mathrm{X_i}$. For $\mathrm{X_i}$, we calculate the corresponding *1*N* vector of t-statistics, t_i.

Step II: We rank its strategies based on their t-statistics, t_i. For the top $\mathrm{p_0} * N$ strategies with the highest t-statistics, we find the corresponding strategies in $\mathrm{X_0}$. We adjust these strategies in $\mathrm{X_0}$ so that their means are the same means for the top $\mathrm{p_0} * N$ in $\mathrm{X_i}$. We denote the data matrix of these adjusted strategies by $X_{0,1}^{(i)}$. For the remaining strategies in X_0, we adjust them so they have a zero in-sample mean. We denote the data matrix for these adjusted strategies by $X_{0,0}^{(i)}$. Finally, we arrange $X_{0,1}^{(i)}$ and $X_{0,0}^{(i)}$ into a new data matrix Y_i by concatenating the two data matrices. We use the new data matrix Y_i to calculate error rates.

Table C1. Classifying testing outcomes

Decision	H_0 true	H_1 true
Reject	False discovery (Type I error) $FP^{i,j}$	True discovery $TP^{i,j}$
Accept	$TN^{i,j}$	Miss discovery (Type II error) $FN^{i,j}$

The **false discovery rate (FDR)** is

$$FDR^{i,j} = \frac{FP^{i,j}}{FP^{i,j} + TP^{i,j}}$$

It is the proportion of rejection in $X_{0,0}^{(i)}$ divided by all rejections.
The **realized rate of misses (MISS)** is

$$MISS^{i,j} = \frac{FN^{i,j}}{FN^{i,j} + TN^{i,j}}$$

It is the proportion of not rejection in $X_{0,1}^{(i)}$ divided by all not rejections.
The ratio of false discoveries to misses, **odds ratio (RATIO)** is

$$RATIO^{i,j} = \frac{FP^{i,j}}{FN^{i,j}}$$

Step III: We bootstrap Y_i the time periods *J* times. For each bootstrapped sample, we calculate the error rates for Y_i for a statistical procedure, such as a fixed t-statistic threshold (e.g., a conventional t-statistic threshold of 2.0) or the range of multiple-testing approaches detailed in Harvey, Liu, and Zhu (2016).

Step IV: Repeat Step I-III *I* time. Calculate the final bootstrapped error rate as $\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} f_{i,j}$

C2. Other multiple test methods

We also consider other multiple test methods that control for false discovery rates, which include Benjamini and Hochberg (1995) (BH), Benjamini, Hochberg, and Yekutieli (2001) (BHY), Barras, Scaillet, and Wermers (2010, JF) (BSW) (based on the method of Storey (2002)). The process for these multiple test (BH, BHY, and BSW) is as follows:

Suppose M denotes the number of tests, N_0 is the number of true null hypothesis. The hypothesis (two-sided): H_0: $HL_i=0$, against the alternative hypothesis H_1: $HL_i \neq 0$ for i=1, 2,…, M.
Type I error is defined as: If the null is the truth, we reject the null. This is also called false discoveries. The significance level α is the probability of Type I error.
First, we calculate the p-values based on two-sided t-statistics in the data.
Second, order the original p-values from low to high such that $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(b)} \leq \cdots \leq p_{(M)}$.
Third, calculate the adjusted critical p-value (adj-p) based on different methods.
Finally, find the maximum index k, such that $p_{(b)} < adj\_p$, then $p_{(k)}$ the new p-value for multiple tests.

The BH, BHY, and BSW methods have the same procedure above. The differences are how to calculate the adjusted p-value in the third step. They are as follows:
For the BHY method, the adjusted p-value is

$$adj\_p = \frac{b * \alpha}{M * c(M)}$$

where $c(M) = \sum_{j=1}^{M} \frac{1}{j}$
For the BH method, the adjusted p-value is

$$adj\_p = \frac{b * \alpha}{M}$$

For the BSW method, the adjusted p-value is

$$adj\_p = \frac{b * \alpha}{N_0}$$

where $\widehat{N}_0 = \frac{1}{1-\lambda}\sum_{\mathrm{i}=1}^{\mathrm{N}}\{\mathrm{p}_{(\mathrm{i})} > \lambda^*\}$, where $\lambda \in (0,1)$ is a tuning parameter. Under the zero-alpha nulls, the p-values are uniformly distributed on (0,1), therefore, one would expect N_0(1-λ) of the p-values to lie within the interval (λ,1) for any sufficiently large λ.

The proportion of zero Long-Short anomalies is

$$\hat{\pi}_0(\lambda^*) = \frac{\hat{N}_0}{M} = \frac{\sum_{i=1}^{N}\{p_{(i)} > \lambda^*\}}{M * (1-\lambda)}$$

Since N_0 is not observable, how to estimate it?
We use the bootstrap procedure proposed by Storey (2002) and Storey, Taylor, and Siegmund (2004). This resampling approach chooses λ from the data such that an estimate of the mean-squared error (MSE) of $\hat{\pi}_0(\lambda)$ is minimized.
First, we compute $\hat{\pi}_0(\lambda)$ across a range of λ values ($\lambda = 0.3, 0.35, \dots 0.7$).
Second, for each possible value of λ, we form 1000 bootstrap replications of $\hat{\pi}_0(\lambda)$ by drawing with replacement from M*1 vector of p-values. These are denoted by $\hat{\pi}_0^b(\lambda)$ from b=1, 2, …1000.
Third, we compute the estimated MSE for each possible value λ:

$$\widehat{MSE}(\lambda) = \frac{1}{1000}\sum_{b=1}^{1000}\left[\hat{\pi}_0^b(\lambda) - \min_{\lambda}\hat{\pi}_0(\lambda)\right]^2$$

We choose λ^* such that $\lambda^* = \arg\min_{\lambda}\widehat{MSE}(\lambda)$.
For Chinese 426 anomalies, BSW λ^* is 0.65 and N_0=223 (<426). In the US fund data, λ^* is 0.6. Bajgrowicz and Scaillet (2012) suggest λ= 0.6.

Table C2. The robust tests with different values of λ in 426 L-S in China

	λ=0.3	λ= 0.6	λ= 0.65	λ= 0.8
t-cutoff	2.68	2.58	2.55	2.60

Table C3 reports the multiple t-cutoff from different methods. For the BHY method, it is too conservative to find the multiple t-cutoff. The BH and BSW methods can produce a t-cutoff, but they do not allow dependence and cannot solve short sample. The multiple t-cutoff of the BH method is 3.16. It is 2.55 for the BSW method.

Table C3: Multiple tests with different methods with 426 VW anomalies (significance level $\alpha = 5\%$)

	BHY	BH	BSW(λ= 0.65)
Multiple t-cutoff	all p>adjusted p	3.16	2.55

D. More Empirical Results

Table D1. The Number of Significant Anomalies in Long-Leg Portfolios

The table reports the number of significant anomalies for long-leg raw returns and returns relative to the CAPM and CH3 factor models in each category and in total at the 5% level with the *t*-cutoffs of 1.96. The anomalies are categorized into 2 groups, which are trading-related anomalies and accounting-related anomalies. We adopt single portfolio analysis to test whether these anomalies are significant with all A-share but 30% of microcap firms from 2000 to 2018. Panel A shows the result for value-weighted portfolios, and Panel B provides the result for equally-weighted portfolios.

Panel A. Value-Weighted Portfolios

	Trading-Related Anomalies			Accounting-Related Anomalies				
	liquidity	risk	past returns	profitability	value	investment	Others	Total
	94	46	52	67	44	51	72	426
Raw returns	5	0	0	0	2	0	2	9
CAPM alphas	5	0	0	0	2	0	2	9
CH3 alphas	2	0	0	0	0	0	0	2

Panel B. Equally-Weighted Portfolios

	Trading-Related Anomalies			Accounting-Related Anomalies				
	liquidity	risk	past returns	profitability	value	investment	Others	Total
	94	46	52	67	44	51	72	426
Raw returns	11	2	0	7	3	0	5	28
CAPM alphas	11	2	0	7	3	0	5	28
CH3 alphas	5	0	0	0	0	0	3	8

Table D2. Insignificant Anomalies among All A-Share but 30% Microcap Firms

This table provides monthly value-weighted long-short raw returns, alphas in percentages relative to the CAPM and the CH-3 factors for insignificant anomalies with the t-cutoff of 1.96. Single portfolio analysis is among all A-share sample but 30% microcap firms from 2000 to 2018. It also provides the Newey-West *t*-statistics adjusted for heteroscedasticity and autocorrelation. The symbols and the definitions of these anomalies are described in Online Appendix B.

Panel A. Liquidity

	Raw return	t	CAPM alpha	t	CH3 alpha	t
size: firm size (market equity), n=1, 6, 12, Banz (1981)						
size	0.62	1.25	0.55	1.13	0.37	3.25
size1	0.73	1.40	0.63	1.25	0.22	1.42
size6	0.53	1.09	0.44	0.94	0.13	1.03
size12	0.55	1.15	0.46	1.01	0.18	1.43
turn: share turnover, k=1, 6, 12, m=0, n=1, 6, 12, Datar, Naik, and Radcliffe (1998)						
turn1-6	0.65	1.56	0.80	1.91	0.34	0.97
turn1-12	0.26	0.68	0.41	1.10	0.01	0.02
turn6-1	0.90	1.93	1.06	2.21	0.56	1.36
turn6-6	0.38	0.90	0.52	1.27	0.16	0.52
turn6-12	0.22	0.58	0.36	1.01	-0.03	-0.10
turn12-1	0.49	1.13	0.63	1.45	0.16	0.46
turn12-6	0.21	0.53	0.34	0.90	-0.07	-0.27
turn12-12	0.16	0.44	0.30	0.89	-0.12	-0.47
vturn: variation of share turnover, k=1, 6, 12, m=0, n=1, 6, 12, Chordia, Subrahmanyam, and Anshuman (2001)						
vturn1-6	0.62	1.52	0.76	1.87	0.40	1.23
vturn1-12	0.28	0.74	0.41	1.11	0.09	0.36
vturn6-6	0.57	1.37	0.69	1.66	0.31	1.03
vturn6-12	0.27	0.72	0.39	1.06	0.08	0.34
vturn12-1	0.75	1.67	0.88	1.98	0.47	1.29
vturn12-6	0.38	0.92	0.50	1.26	0.15	0.56
vturn12-12	0.18	0.50	0.31	0.89	-0.02	-0.09
cvturn: coefficient of variation of share turnover, k=1, 6, 12, m=0, n=1, 6, 12, Chordia, et al. (2001)						
cvturn1-1	0.62	1.61	0.63	1.69	0.58	1.79
cvturn1-6	0.17	0.82	0.16	0.83	0.20	0.99
cvturn1-12	0.03	0.17	0.03	0.17	0.11	0.60
cvturn6-1	0.66	1.91	0.63	1.85	0.86	2.54
cvturn6-12	0.21	0.87	0.18	0.78	0.30	1.24
cvturn12-6	0.29	0.99	0.26	0.92	0.39	1.27
cvturn12-12	0.07	0.25	0.04	0.15	0.21	0.75
abturn: abnormal return, k=12, m=0, n=6, 12, Liu, Stambaugh, and Yuan (2019)						
abturn6	0.32	1.23	0.38	1.46	0.39	1.21
abturn12	0.07	0.34	0.11	0.56	0.16	0.73
dtv: dollar trading volume, k=1, 5, 12, m=0, n=1, 6, 12, Brennan, Chordia, and Subrahmanyam (1998)						
dtv1-6	0.60	1.42	0.61	1.49	0.04	0.21
dtv1-12	0.55	1.32	0.55	1.37	0.03	0.17
dtv6-1	0.75	1.74	0.76	1.81	0.12	0.65

dtv6-6	0.62	1.47	0.61	1.50	0.04	0.22
dtv6-12	0.61	1.46	0.60	1.50	0.07	0.39
dtv12-1	0.70	1.65	0.70	1.73	0.14	0.84
dtv12-6	0.69	1.66	0.68	1.72	0.16	0.99
dtv12-12	0.62	1.50	0.61	1.55	0.10	0.61
vdtv: variation of share dollar trading volume, k=1, 6, 12, m=0, n=6, 12, Chordia, Subrahmanyam, and Anshuman (2001)						
vdtv1-6	0.71	1.75	0.72	1.82	0.14	0.82
vdtv1-12	0.60	1.49	0.59	1.54	0.09	0.57
vdtv6-6	0.67	1.61	0.68	1.69	0.13	0.77
vdtv6-12	0.64	1.56	0.65	1.64	0.15	0.88
vdtv12-12	0.71	1.76	0.72	1.87	0.24	1.33
cvd: coefficient of variation of dollar trading volume, k=1, 6, 12, m=0, n=1, 6, 12, Chordia, Subrahmanyam, and Anshuman (2001)						
cvd1-1	0.55	1.27	0.56	1.34	0.45	1.26
cvd1-6	0.16	0.73	0.16	0.77	0.16	0.80
cvd1-12	0.00	0.02	0.01	0.06	0.05	0.32
cvd6-12	0.29	1.09	0.29	1.13	0.44	1.64
cvd12-6	0.54	1.74	0.56	1.86	0.53	1.48
Ami: Amihud illiquidity, k=1, 6, 12, m=0, n=1, 6, 12, Amihud (2002)						
Ami1-1	0.64	1.39	0.61	1.38	0.28	1.48
Ami1-6	0.41	0.94	0.37	0.89	0.04	0.22
Ami1-12	0.38	0.88	0.34	0.82	0.03	0.16
Ami6-1	0.36	0.82	0.30	0.72	-0.04	-0.26
Ami6-6	0.37	0.88	0.31	0.77	-0.02	-0.10
Ami6-12	0.35	0.84	0.30	0.75	0.00	-0.01
Ami12-1	0.24	0.55	0.16	0.40	0.07	0.37
Ami12-6	0.30	0.71	0.24	0.59	0.07	0.40
Ami12-12	0.23	0.56	0.19	0.48	0.02	0.11
Lm: turnover-adjusted number of zero daily trading volume, k=1, 6, 12, m=0, n=1, 6, 12, Liu (2006)						
Lm1-6	0.61	1.71	0.74	2.06	0.25	0.77
Lm1-12	0.30	0.94	0.42	1.34	-0.03	-0.13
Lm6-1	0.06	0.20	0.12	0.43	0.07	0.20
Lm6-6	0.08	0.38	0.04	0.17	0.27	1.09
Lm6-12	0.00	0.01	0.04	0.22	-0.24	-1.17
Lm12-1	0.06	0.28	0.07	0.32	-0.24	-0.97
Lm12-6	0.07	0.35	0.08	0.40	-0.43	-2.13
Lm12-12	0.02	0.12	-0.01	-0.03	0.50	2.75
Liquidity betas, n=1, 6, 12, Acharya and Pedersen (2005)						
Bret1	0.65	1.76	0.87	2.31	0.01	0.02
Bret6	0.57	1.63	0.77	2.29	0.07	0.16
Bret12	0.49	1.43	0.69	2.15	-0.02	-0.06
Blcc1	0.09	0.27	0.14	0.41	0.12	0.47
Blcc6	0.12	0.37	0.15	0.50	0.10	0.39
Blcc12	0.21	0.69	0.24	0.80	0.28	1.07
Blrc1	0.56	1.48	0.62	1.60	0.43	0.91

Blrc6	0.41	1.38	0.44	1.52	0.23	0.60
Blrc12	0.35	1.29	0.39	1.48	0.07	0.19
Blcr1	0.18	0.53	0.20	0.61	0.43	1.74
Blcr6	0.24	0.72	0.25	0.78	0.44	1.81
Blcr12	0.25	0.76	0.26	0.80	0.44	1.79
Bnet1	0.22	0.58	0.20	0.55	0.14	0.47
Bnet6	0.04	0.13	0.04	0.11	0.01	0.02
Bnet12	0.07	0.21	0.06	0.19	0.20	0.67

Panel B. Risk

	Raw return	t	CAPM alpha	t	CH3 alpha	t
idv: idiosyncratic volatility per the CAPM, annual sort, k=12, n=12, Ali, Hwang, and Trombley (2003)						
idv	0.61	1.33	0.75	1.79	0.52	1.78
idvc: idiosyncratic volatility per the CAPM, k=1, m=0, n=6, 12						
idvc6	0.65	1.82	0.79	2.24	0.26	0.87
idvc12	0.30	0.89	0.44	1.38	0.10	0.42
idvff: idiosyncratic volatility per the CH3 model, k=1, m=0, n=12, Ang, Hodrick, Xing, and Zhang (2006)						
idvff12	0.37	1.10	0.50	1.62	0.17	0.74
idvq: idiosyncratic volatility per the q-factor model, k=1, m=0, n=12						
idvq12	0.36	1.04	0.49	1.57	0.18	0.83
tv: total volatility, k=1, m=0, n=12, Ang, Hodrick, Xing, and Zhang (2006)						
tv12	0.49	1.38	0.66	2.07	0.15	0.71
idsc: idiosyncratic skewness per the CAPM, k=1, m=0, n=6, 12						
idsc6	0.19	1.37	0.20	1.48	0.25	1.85
idsc12	0.01	0.13	0.00	0.04	0.11	0.92
idsff: idiosyncratic skewness per the CH 3-factor model, k=1, m=0, n=1, 6, 12, Boyer, Mitton, and Vorkink (2009)						
idsff1	0.46	1.76	0.49	1.91	0.85	3.34
idsff6	0.10	0.59	0.14	0.94	0.23	1.79
idsff12	0.05	0.39	0.02	0.15	0.07	0.66
idsq: idiosyncratic skewness per the q-factor model, k=1, m=0, n=1, 6, 12						
idsq1	0.30	1.35	0.33	1.54	0.64	3.12
idsq6	0.06	0.36	0.09	0.66	0.07	0.55
idsq12	0.10	0.79	0.07	0.60	0.14	1.25
ts: total skewness, k=1, m=0, n=6, 12, Amaya, Christoffersen, Jacobs, and Vasquez (2015)						
ts6	0.23	1.33	0.23	1.39	0.07	0.52
ts12	0.10	0.66	0.09	0.61	-0.10	-1.01
cs: co-skewness, k=1, m=0, n=1, 6, 12, Harvey and Siddique (2000)						
cs1	0.28	0.77	0.27	0.79	0.20	0.71
cs6	0.19	0.80	0.17	0.79	0.09	0.63
cs12	0.12	0.60	0.11	0.55	0.09	0.74
betam: market beta using monthly returns, k=60, m=0, n=1, 6, 12, Fama and MacBeth (1973)						
betam1	0.64	1.78	0.87	2.42	-0.01	-0.03
betam6	0.62	1.70	0.84	2.57	0.21	0.46
betam12	0.51	1.40	0.72	2.21	0.10	0.22
beta: market beta using daily returns, k=12, m=0, n=6, 12						

beta6	0.68	1.70	0.90	2.51	0.41	0.93
beta12	0.51	1.34	0.72	2.16	0.21	0.47
dbeta: downside beta, k=12, m=0, n=1, 6, 12, Ang, Chen, and Xing (2006)						
dbeta1	0.52	1.22	0.68	1.65	1.20	2.85
dbeta6	0.16	0.40	0.32	0.86	0.88	2.57
dbeta12	0.05	0.12	0.20	0.57	0.80	2.76
betaFP: the Frazzini-Pedersen beta, k=12, m=0, n=6, 12, Frazzini and Pedersen (2013)						
betaFP6	0.61	1.51	0.84	2.41	0.47	1.11
betaFP12	0.48	1.22	0.69	2.10	0.31	0.76
betaDM: the Dimson beta, k=1, m=0, n= 6, 12, Dimson (1979)						
betaDM6	0.12	0.52	0.27	1.25	0.06	0.22
betaDM12	0.20	0.93	0.34	1.76	0.11	0.45
tail: tail risk, k=120, m=0, n=1, 6, 12, Kelly and Jiang (2014)						
tail1	0.38	1.19	0.39	1.22	-0.06	-0.18
tail6	0.26	0.89	0.29	1.01	-0.02	-0.07
tail12	0.21	0.79	0.24	0.91	0.00	0.00

Panel C. Past Returns

	Raw return	t	CAPM alpha	t	CH3 alpha	t
m: price momentum, prior k-month returns, k=11, 9, 6, 3, m=1, n=1, 6, 12, Jegadeesh and Titman (1993)						
m11-1	0.38	0.86	0.43	1.01	0.54	0.92
m11-6	0.22	0.56	0.26	0.66	0.04	0.07
m11-12	0.12	0.35	0.14	0.41	-0.08	-0.15
m9-1	0.25	0.55	0.30	0.68	0.87	1.71
m9-6	0.42	1.12	0.46	1.25	0.34	0.67
m9-12	0.08	0.24	0.11	0.34	-0.08	-0.17
m6-1	0.00	0.00	-0.08	-0.18	-1.08	-2.49
m6-6	0.51	1.48	0.55	1.61	0.65	1.55
m6-12	0.16	0.59	0.19	0.68	0.13	0.37
m3-1	0.03	0.08	0.02	0.06	-0.92	-2.05
m3-6	0.27	0.96	0.31	1.11	0.71	2.35
m3-12	0.16	0.78	0.19	0.96	0.34	1.48
m24-6	0.52	1.56	0.53	1.65	0.20	0.55
m24-12	0.49	1.50	0.50	1.57	0.17	0.47
nm: industry momentum, k=6, m=0, n=1, 6, 12						
inm1	0.36	1.01	0.34	0.96	0.92	2.22
inm6	0.35	1.14	0.36	1.16	0.69	1.93
inm12	0.17	0.69	0.18	0.76	0.33	1.16
mchg: momentum change, k=12, m=1, n=1, 12, Gettleman and Marks (2006)						
mchg1	0.29	0.80	0.25	0.73	-0.64	-1.65
mchg12	0.24	1.54	0.25	1.68	0.33	1.91
im: residual momentum, prior k-month returns, k=11, 6, m=0, n=1, 6, 12, Blitz, Huij, and Martens (2011)						
im11-6	0.15	0.63	0.15	0.64	0.13	0.43
im11-12	0.04	0.21	0.05	0.25	0.02	0.09
im6-1	0.33	1.38	0.33	1.43	0.79	2.85

im6-12	0.09	0.49	0.10	0.55	0.19	0.94
srev: short-term reversal, k=1, m=0, n=1, 6, 12, Jegadeesh (1990)						
srev1	0.66	1.56	0.69	1.73	0.28	0.67
srev6	0.15	0.83	0.15	0.84	-0.30	-1.59
srev12	0.15	1.03	0.15	1.05	0.29	1.69
52wh: 52-week high, k=12, m=0, n=1, 6, George and Hwang (2004)						
52wh1	0.45	0.93	0.57	1.18	0.58	1.08
52wh6	0.65	1.78	0.78	2.26	0.56	1.09
mdr: maximum daily return, k=1, m=0, n=6, 12, Bali, Cakici, and Whitelaw (2011)						
mdr6	0.61	1.82	0.77	2.35	0.11	0.39
mdr12	0.28	0.88	0.43	1.49	-0.03	-0.13
pr: share price, k=1, m=0, n=1, 6, 12, Miller and Scholes (1982)						
pr1	0.32	1.03	0.38	1.26	-0.06	-0.19
pr6	0.22	0.69	0.27	0.89	-0.08	-0.29
pr12	0.24	0.79	0.29	0.98	-0.06	-0.22
ilr: industry lead-lag effect in prior returns, k=1, m=0, n=6, 12						
ilr6	0.13	0.80	0.11	0.74	0.48	2.35
ilr12	0.13	1.05	0.12	1.04	0.28	1.80
abr: cumulative abnormal returns around earnings announcement dates, n=1, 6, 12						
abr1	0.37	1.24	0.41	1.44	0.28	0.81
abr6	0.06	0.29	0.12	0.61	0.16	0.68
abr12	0.05	0.31	0.10	0.65	0.13	0.84
Ra1: Year 1-lagged return, annual, Heston and Sadka (2008)						
Ra1	0.35	1.15	0.36	1.20	0.06	0.18
Rn1: Year 1-lagged return, nonannual, Heston and Sadka (2008)						
Rn1	0.19	0.40	0.18	0.38	-0.30	-0.48

Panel D. Profitability

	Raw return	t	CAPM alpha	t	CH3 alpha	t
roe: return on equity, n=12, Hou, Xue, and Zhang (2015)						
roe12	0.57	1.73	0.63	1.97	0.01	0.07
roa: return on assets, n=12, Balakrishnan, Bartov, and Faurel (2010)						
roa12	0.56	1.72	0.63	2.02	0.20	0.82
roa: return on assets, n=6, 12, Balakrishnan, Bartov, and Faurel (2010)						
rna	0.49	1.51	0.58	1.93	-0.13	-0.34
rnaq6	0.48	1.63	0.58	2.03	-0.20	-0.52
rnaq12	0.38	1.43	0.47	1.86	-0.19	-0.59
pm: profit margin, n=1, 6, 12, Soliman (2008)						
pm	0.43	0.96	0.52	1.25	0.12	0.53
pmq6	0.46	1.29	0.58	1.83	-0.12	-0.53
pmq12	0.43	1.25	0.55	1.71	-0.04	-0.17
atoq6	0.38	1.80	0.42	1.96	0.00	-0.01
atoq12	0.34	1.71	0.38	1.88	-0.03	-0.09
ct: capital turnover, n=1, 6, 12, Haugen and Baker (1996)						

ct	0.08	0.32	0.05	0.20	0.53	1.56
ctq1	0.14	0.49	0.09	0.32	0.24	0.82
ctq6	0.18	0.65	0.13	0.48	0.27	0.92
ctq12	0.17	0.63	0.12	0.44	0.31	1.02
gpa: gross profits-to-assets, Novy-Marx (2013)						
gpa	0.56	1.32	0.61	1.51	0.23	0.74
gpla: gross profits-to-lagged assets, n=12						
gpla	0.46	1.14	0.52	1.35	0.24	0.72
gplaq12	0.56	1.62	0.63	1.95	0.24	0.88
ope: operating profits-to-book equity, Fama and French (2015)						
ope	0.52	1.25	0.59	1.48	0.19	0.73
ople: operating profits-to-lagged book equity						
ople	0.57	1.38	0.65	1.67	0.33	1.03
opa: operating profits-to-assets, Ball, Gerakos, Linnaimma, and Nikolaev (2015)						
opa	0.54	1.33	0.60	1.55	0.30	1.14
opla: operating profits-to-lagged assets, n=12						
opla	0.43	1.07	0.50	1.29	0.24	0.83
oplaq12	0.53	1.60	0.61	1.97	0.16	0.66
tbi: taxable income-to-book income, n=1, 6, 12, Lev and Nissim (2004)						
tbi	0.17	0.52	0.10	0.32	0.13	0.38
tbiq1	0.21	0.76	0.28	1.09	0.29	0.89
tbiq6	0.18	0.71	0.25	1.01	0.17	0.67
tbiq12	0.17	0.74	0.23	1.06	0.19	0.87
bl: book leverage, n=1, 6, 12, Fama and French (1992)						
bl	0.29	0.76	0.27	0.72	0.62	1.64
blq1	0.05	0.12	0.08	0.22	-0.09	-0.28
blq6	0.05	0.13	0.08	0.23	-0.15	-0.45
blq12	0.00	0.01	-0.03	-0.09	0.16	0.48
sg: annual sales growth, n=12, Lakonishok, Shleifer, and Vishny (1994)						
sg	0.11	0.40	0.11	0.44	0.16	0.52
sgq12	0.34	1.88	0.33	1.89	0.11	0.50
Fscore: Fundamental score, Piotroski (2000)						
Fscore	0.15	0.60	0.20	0.84	-0.25	-1.09
Oscore: Ohlson’s (1980) O-score, n=1, 6, 12, Dichev (1998)						
Oscore	0.40	1.22	0.45	1.51	0.42	1.53
Oq1	0.31	1.07	0.41	1.56	0.29	1.10
Oq6	0.14	0.51	0.20	0.79	0.15	0.60
Oq12	0.19	0.71	0.24	0.97	0.16	0.61
Zscore: Altman’s (1968) Z-score, n=1, 6, 12, Dichev (1998)						
Zscore	0.19	0.48	0.14	0.37	-0.08	-0.17
Zq1	0.29	0.73	0.22	0.60	0.18	0.34
Zq6	0.30	0.78	0.24	0.68	0.09	0.19
Zq12	0.29	0.76	0.23	0.67	0.05	0.12

Panel E. Value

	Raw return	t	CAPM alpha	t	CH3 alpha	t

bm: book-to-market equity, Basu (1983)						
bm	0.79	1.92	0.76	1.90	0.62	1.14
bmq: quarterly book-to-market equity, n=6						
bmq6	0.79	1.87	0.77	1.89	0.42	0.73
dm: debt-to-market equity, n=1, 6, 12, Bhandari (1988)						
dm	0.34	0.84	0.34	0.88	0.03	0.07
dmq1	0.44	1.12	0.47	1.22	-0.01	-0.02
dmq6	0.35	0.89	0.37	0.98	-0.02	-0.05
dmq12	0.35	0.92	0.37	1.00	0.00	0.00
am: assets-to-market equity, n=1, 6, 12, Fama and French (1992)						
am	0.52	1.26	0.51	1.28	0.21	0.53
amq1	0.63	1.49	0.65	1.57	-0.07	-0.16
amq6	0.49	1.17	0.51	1.25	0.06	0.14
amq12	0.53	1.31	0.55	1.40	0.14	0.31
ep: earnings-to-price ratio, Basu (1983)						
ep	0.63	1.54	0.69	1.75	0.26	0.86
cfp: cash flow to price, n=1, 6, 12, Lakonishok, Shleifer, and Vishny (1994)						
cfp	0.17	0.72	0.20	0.83	0.09	0.27
cfpq1	0.01	0.04	0.03	0.15	0.34	1.31
cfpq6	0.20	1.21	0.20	1.24	0.53	2.13
cfpq12	0.07	0.47	0.07	0.52	0.31	1.42
sr: 5-year sales growth rank, Lakonishok, Shleifer, and Vishny (1994)						
sr	0.24	0.75	0.27	0.86	0.15	0.52
em: enterprise multiple, n=1, 6, 12, Loughran and Wellman (2011)						
em	0.13	0.42	0.14	0.47	-0.01	-0.03
emq1	0.00	0.01	-0.01	-0.05	0.00	0.01
emq6	0.06	0.42	0.06	0.45	0.24	1.42
emq12	0.01	0.07	0.01	0.11	-0.02	-0.10
ocfp: operating cash flow-to-price ratio, n=1, 6, 12, Desai, Rajgopal, and Venkatachalam (2004)						
ocfp	0.35	1.18	0.45	1.63	0.11	0.49
ocfpq1	0.32	1.56	0.38	2.03	0.16	0.67
ocfpq6	0.30	1.62	0.35	1.99	0.29	1.65
ocfpq12	0.18	1.03	0.23	1.45	0.19	1.18
de: debt-to-book equity, Bhandari (1988)						
de	0.21	0.54	0.19	0.50	0.56	1.50
ebp: enterprise book-to-price, n=6, 12, Penman, Richardson, and Tuna (2007)						
ebp	0.53	1.66	0.50	1.63	0.42	1.03
ebpq6	0.55	1.53	0.51	1.50	0.35	0.77
ebpq12	0.45	1.35	0.43	1.33	0.25	0.58
ndp: net debt-to-price, n=1, 6, 12, Penman, Richardson, and Tuna (2007)						
ndp	0.05	0.18	0.06	0.24	0.05	0.12
ndpq1	0.42	1.23	0.39	1.18	0.19	0.44
ndpq6	0.29	0.95	0.26	0.90	0.19	0.55
ndpq12	0.35	1.22	0.32	1.19	0.20	0.65

Panel F. Investment

	Raw return	t	CAPM alpha	t	CH3 alpha	t
aci: abnormal corporate investment, Titman, Wei, and Xie (2004)						
aci	0.06	0.30	0.07	0.41	0.20	0.66
ag: investment-to-assets, n=1, 6, 12, Cooper, Gulen, and Schill (2008)						
ag	0.19	0.66	0.15	0.55	0.09	0.26
agq1	0.49	1.80	0.50	1.86	0.28	0.94
agq6	0.27	1.04	0.29	1.11	0.25	0.80
agq12	0.07	0.27	0.08	0.34	0.08	0.31
dpia: change in property, plant, equipment and inventory scaled by book assets, Lyandres, Sun, and Zhang (2008)						
dpia	0.23	0.88	0.25	1.01	-0.01	-0.04
noa: net operating assets, Hirshleifer, Hou, Teoh, and Zhang (2004)						
noa	0.44	1.49	0.48	1.67	0.25	1.04
dnoa: change in net operating assets						
dnoa	0.32	1.43	0.31	1.42	0.37	1.27
ig: investment growth, Xie (2001)						
ig	0.08	0.48	0.04	0.28	-0.10	-0.51
ig2/3: 2/3 year investment growth, Anderson and Garcia-Feijoo (2006)						
ig2	0.06	0.29	0.05	0.23	0.23	1.01
ig3	0.24	1.15	0.25	1.21	0.00	0.02
nsi: net stock issues, Pontiff and Woodgate (2008)						
nsi	0.15	0.61	0.20	0.88	-0.04	-0.12
cdi: composite debt issuance, Lyandres, Sun and Zhang (2008)						
cdi	0.36	1.33	0.38	1.44	-0.06	-0.19
ivg: inventory growth, Belo and Lin (2011)						
ivg	0.13	0.77	0.15	0.87	0.01	0.03
ivc: inventory changes, Thomas and Zhang (2002)						
ivc	0.07	0.28	0.08	0.33	0.15	0.45
oacc: operating accruals, Sloan (1996)						
oacc	0.10	0.36	0.16	0.57	0.16	0.57
tacc: total accruals, Richardson, Sloan, Soliman, and Tuna (2005)						
tacc	0.01	0.04	0.04	0.21	-0.22	-0.78
dwc: change in net noncash working capital, Richardson, Sloan, Soliman, and Tuna (2005)						
dwc	0.05	0.29	0.03	0.22	-0.16	-0.73
dcoa: change in current operating assets, Richardson, Sloan, Soliman, and Tuna (2005)						
dcoa	0.25	0.97	0.24	0.98	0.20	0.61
dcol: change in current operating liabilities, Richardson, Sloan, Soliman, and Tuna (2005)						
dcol	0.07	0.33	0.08	0.36	-0.12	-0.44
dnco: change in net noncurrent operating assets, Richardson, Sloan, Soliman, and Tuna (2005)						
dnco	0.14	0.72	0.13	0.70	0.04	0.18
dnca: change in noncurrent operating assets, Richardson, Sloan, Soliman, and Tuna (2005)						
dnca	0.34	1.90	0.34	1.95	0.14	0.60
dncl: change in noncurrent operating liabilities, Richardson, Sloan, Soliman, and Tuna (2005)						
dncl	0.04	0.20	0.04	0.19	-0.05	-0.20
dfin: change in net financial assets, Richardson, Sloan, Soliman, and Tuna (2005)						

dfin	0.06	0.28	0.08	0.35	-0.31	-1.14
dsti: change in short-term investments, Richardson, Sloan, Soliman, and Tuna (2005)						
dsti	0.06	0.25	0.05	0.19	0.10	0.35
dlti: change in long-term investments, Richardson, Sloan, Soliman, and Tuna (2005)						
dlti	0.10	0.44	0.06	0.29	0.10	0.44
dfnl: change in financial liabilities, Richardson, Sloan, Soliman, and Tuna (2005)						
dfnl	0.34	1.75	0.36	1.88	-0.07	-0.32
dbe: change in common equity, Richardson, Sloan, Soliman, and Tuna (2005)						
dbe	0.00	0.00	0.03	0.09	-0.03	-0.08
dacc: discretionary accruals						
dacc	0.43	1.51	0.42	1.54	0.60	1.77
poacc: percent operating accruals, Hafzalla, Lundholm, and Van Winkle (2011)						
poacc	0.27	1.15	0.30	1.28	0.11	0.51
ptacc: percent total accruals, Hafzalla, Lundholm, and Van Winkle (2011)						
ptacc	0.19	0.76	0.21	0.90	0.19	0.53
pdacc: percent discretionary accruals						
pdacc	0.37	1.40	0.40	1.56	0.03	0.08
cagq: current asset growth, n=1, 6, 12						
cagq1	0.36	1.48	0.40	1.62	0.13	0.46
cagq6	0.22	0.94	0.25	1.05	0.15	0.62
cagq12	0.07	0.33	0.09	0.41	0.06	0.29
ncagq: noncurrent asset growth, n=1, 6, 12						
ncagq1	0.11	0.51	0.07	0.33	-0.11	-0.48
ncagq6	0.19	0.93	0.15	0.78	0.03	0.15
ncagq12	0.22	1.16	0.19	1.05	-0.02	-0.08
cashg: cash growth, n=1, 6, 12						
cashgq1	0.35	1.49	0.37	1.65	0.08	0.36
cashgq6	0.18	0.87	0.21	0.99	0.08	0.33
cashgq12	0.02	0.09	-0.01	-0.05	0.05	0.23
fagq: fixed asset growth, n=1, 6, 12						
fagq1	0.07	0.34	0.07	0.34	0.50	1.93
fagq6	0.06	0.31	0.06	0.30	-0.33	-1.56
fagq12	0.17	0.93	0.16	0.92	-0.20	-1.09
nccagq: non-cash current asset growth, n=1, 6, 12						
nccagq1	0.27	1.01	0.29	1.08	0.06	0.16
nccagq6	0.28	1.14	0.29	1.19	0.19	0.67
nccagq12	0.12	0.50	0.12	0.50	0.16	0.66
oagq: other asset growth, n=1						
oagq1	0.23	1.37	0.21	1.33	0.13	0.65

Panel G. Others

	Raw return	t	CAPM alpha	t	CH3 alpha	t
gad: growth in advertising expense						
gad	0.17	0.90	0.18	0.96	-0.05	-0.27
rds: R&D expense-to-sales, n=1, 6, 12, Chan, Lakonishok, and Sougiannis (2001)						

rds	0.04	0.16	-0.02	-0.06	-0.06	-0.21
rdsq1	0.05	0.17	0.12	0.45	0.35	1.10
rdsq6	0.10	0.38	0.17	0.66	0.37	1.18
rdsq12	0.02	0.08	0.09	0.38	0.24	0.85
ol: operating leverage, n=1, 6, 12, Novy-Marx (2011)						
ol	0.15	0.65	0.12	0.53	0.20	0.72
olq1	0.30	1.37	0.27	1.29	0.19	0.86
olq6	0.21	0.95	0.17	0.82	0.16	0.66
olq12	0.20	0.91	0.16	0.75	0.18	0.71
hn: hiring rate, Belo, Lin, and Bazdresch (2014)						
hn	0.15	0.42	0.12	0.34	-0.24	-0.58
age: firm age, n=1, 6, 12, Jiang, Lee, and Zhang (2005)						
age1	0.19	0.65	0.16	0.58	0.18	0.50
age6	0.12	0.48	0.09	0.37	0.20	0.64
age12	0.01	0.05	-0.02	-0.11	0.09	0.31
dsi: % change in sales minus % change in inventory, Abarbanell and Bushee (1998)						
dsi	0.21	1.19	0.19	1.13	0.08	0.26
dsa: % change in sales minus % change in accounts receivable, Abarbanell and Bushee (1998)						
dsa	0.80	1.33	0.79	1.29	0.49	0.89
dgs: % change in gross margin minus % change in sales, Abarbanell and Bushee (1998)						
dgs	0.03	0.13	0.01	0.05	0.11	0.54
dss: % change in sales minus % change in SG&A, Abarbanell and Bushee (1998)						
dss	0.05	0.20	0.05	0.21	-0.17	-0.49
etr: effective tax rate, Abarbanell and Bushee (1998)						
etr	0.26	1.14	0.28	1.23	0.00	-0.01
lfe: labor force efficiency, Abarbanell and Bushee (1998)						
lfe	0.08	0.34	0.04	0.19	0.02	0.05
vcf: cash flow volatility, n=1, 6, 12						
vcf1	0.05	0.21	0.06	0.23	0.31	1.04
vcf6	0.06	0.21	0.05	0.18	-0.13	-0.39
vcf12	0.02	0.06	0.01	0.03	-0.09	-0.28
cta: cash-to-assets, n=6, 12, Palazzo (2012)						
cta6	0.50	1.77	0.46	1.72	0.41	1.42
cta12	0.39	1.36	0.36	1.31	0.27	0.95
eper: earnings persistence, Francis, LaFond, Olsson, and Schipper (2004)						
eper	0.53	1.51	0.59	1.63	0.53	1.37
eprd: earnings predictability, Francis, LaFond, Olsson, and Schipper (2004)						
eprd	0.11	0.31	0.06	0.17	0.65	1.69
esm: earnings smoothness, Francis, LaFond, Olsson, and Schipper (2004)						
esm	0.20	0.44	0.20	0.47	0.40	1.00
evr: value relevance of earnings, Francis, LaFond, Olsson, and Schipper (2004)						
evr	0.45	1.26	0.43	1.24	0.43	1.07
etl: earnings timeliness, Francis, LaFond, Olsson, and Schipper (2004)						
etl	0.02	0.08	0.01	0.02	-0.05	-0.12
ecs: earnings conservatism						

ecs	0.05	0.17	0.07	0.28	-0.06	-0.21
ala: asset liquidity						
ala	0.09	0.35	0.03	0.14	0.07	0.23
alm	0.70	1.92	0.61	2.02	0.44	1.47
rs: revenue surprise, n=1, 6, Jegadeesh and Livnat (2006)						
rs1	0.41	1.45	0.44	1.59	0.19	0.76
rs6	0.41	1.80	0.43	1.91	0.18	0.73
tes: tax expense surprise, n=1, 6, 12, Thomas and Zhang (2011)						
tes1	0.34	1.31	0.34	1.40	0.17	0.77
tes6	0.25	1.07	0.22	1.06	0.37	1.40
tes12	0.10	0.52	0.07	0.41	0.20	0.93
def: change in analysts' earnings forecasts, n=1, 6, Hawkins, Chamberlin, and Daniel (1984)						
def1	0.68	1.71	0.57	1.81	0.43	1.25
def6	0.28	1.57	0.24	1.54	-0.01	-0.05
re: revisions in analysts' earnings forecasts, n=1, 6, 12, Chan, Jegadeesh, and Lakonishok (1996)						
re1	0.38	1.28	0.37	1.25	-0.34	-0.76
re6	0.26	0.84	0.29	0.94	-0.53	-1.34
re12	0.36	1.38	0.34	1.35	-0.30	-1.10
efp: analysts' earnings forecasts-to-price, n=6, 12, Elgers, Lo, and Pfeiffer (2001)						
efp6	0.81	1.43	0.90	1.68	-0.14	-0.40
efp12	0.51	0.96	0.60	1.19	0.11	0.34
ana: analyst coverage, n=1, 6, 12, Elgers, Lo, and Pfeiffer (2001)						
ana1	0.04	0.07	-0.08	-0.16	-0.15	-0.53
ana6	0.10	0.21	-0.03	-0.07	-0.15	-0.66
ana12	0.04	0.09	-0.09	-0.23	-0.14	-0.71
dis: dispersion in analyst forecasts, n=1, 6, 12						
dis1	0.27	0.71	0.34	0.89	-0.58	-1.45
dis6	0.10	0.43	0.20	0.93	-0.51	-1.90
dis12	0.15	0.64	0.26	1.20	-0.29	-1.11

References

Abarbanell, Jeffery S., and Brian J. Bushee, 1998, Abnormal returns to a fundamental analysis strategy, *Accounting Review* 73, 19-45.

Acharya, Viral V., and Lasse Heje Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375-410.

Ali, Ashiq, Lee-Seok Hwang, and Mark A. Trombley, 2003, Arbitrage risk and the book-to-market anomaly, *Journal of Financial Economics* 69, 355-373.

Altman, Edward I., 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance* 23, 589-609.

Amaya, Diego, Peter Christoffersen, Kris Jacobs, and Aurelio Vasquez, 2015, Does realized skewness predict the cross-section of equity returns?, *Journal of Financial Economics* 118, 135-167.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

Ang, Andrew, Joseph Chen, and Yuhang Xing, 2006, Downside risk, *Review of Financial Studies* 19, 1191-1239.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Asness, Clifford, and Andrea Frazzini, 2013, The devil in hml's details, *Journal of Portfolio Management* 39, 49-68.

Balakrishnan, Karthik, Eli Bartov, and Lucile Faurel, 2010, Post loss/profit announcement drift, *Journal of Accounting and Economics* 50, 20-41.

Bali, Turan G., Stephen J. Brown, Scott Murray, and Yi Tang, 2017, A lottery-demand-based explanation of the beta anomaly, *Journal of Financial and Quantitative Analysis* 52, 2369-2397.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427-446.

Ball, Ray, Joseph Gerakos, Juhani Linnaimma, and Valeri Nikolaev, 2015, Deflating profitability, *Journal of Financial Economics* 117, 225-248.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Barbee, William C., Sandip Mukherji Jr., and Gary A. Raines, 1996, Do sales–price and debt–equity explain stock returns better than book–market and firm size?, *Financial Analysts Journal* 52, 56-60.

Basu, Sanjoy, 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis, *Journal of Finance* 32, 663-682.

Basu, Sanjoy, 1983, The relationship between earnings' yield, market value and return for nyse common stocks: Further evidence, *Journal of Financial Economics* 12, 129-156.

Belo, Frederico, and Xiaoji Lin, 2011, The inventory growth spread, *Review of Financial Studies* 25, 278-313.

Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch, 2014, Labor hiring, investment, and stock return predictability in the cross section, *Journal of Political Economy* 122, 129-177.

Bhandari, Laxmi Chand, 1988, Debt/equity ratio and expected common stock returns: Empirical evidence, *Journal of Finance* 43, 507-528.

Blitz, David, Joop Huij, and Martin Martens, 2011, Residual momentum, *Journal of Empirical Finance* 18, 506-521.

Boyer, Brian, Todd Mitton, and Keith Vorkink, 2009, Expected idiosyncratic skewness, *Review of Financial Studies*

23, 169-202.

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 345-373.

Chan, Louis K. C., Josef Lakonishok, and Theodore Sougiannis, 2011, The stock market valuation of research and development expenditures, *Journal of Finance* 56, 2431-2456.

Chan, Louis K.C., Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *The Journal of Finance* 51, 1681-1713.

Chen, Xuanjuan, Kenneth A. Kim, Tong Yao, and Tong Yu, 2010, On the predictability of chinese stock returns, *Pacific-Basin Finance Journal* 18, 403-425.

Chordia, Tarun, Avanidhar Subrahmanyam, and V. Ravi Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3-32.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609-1651.

Daniel, Kent D., and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605-1643.

Datar, Vinay T., Narayan Y. Naik, and Robert Radcliffe, 1998, Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203-219.

De Bondt, Werner F. M., and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793-805.

Desai, Hemang, Shivaram Rajgopal, and Mohan Venkatachalam, 2004, Value-glamour and accruals mispricing: One anomaly or two?, *Accounting Review* 79, 355-385.

Dimson, Elroy, 1979, Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7, 197-226.

Elgers, Pieter T., May H. Lo, and Ray J. Pfeiffer, 2001, Delayed security price adjustments to financial analysts' forecasts of

annual earnings, *Accounting Review* 76, 613-32.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Foster, George, Chris Olsen, and Terry Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, *Accounting Review* 59, 574-603.

Francis, Jennifer, Ryan LaFond, Per M. Olsson, and Katherine Schipper, 2004, Cost of equity and earnings attributes, *Accounting Review* 79, 967-1010.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1-25.

George, Thomas J., and Chuan-Yang Hwang, 2004, The 52-week high and momentum investing, *Journal of Finance* 59, 2145-2176.

Gettleman, Eric, and Joseph M. Marks, 2006, Acceleration strategies, *working paper*.

Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263-1295.

Haugen, Robert A., and Nardin L. Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401-439.

Hawkins, Eugene H., Stanley C. Chamberlin, and Wayne E. Daniel, 1984, Earnings expectations and security prices, *Financial Analysts Journal* 40, 24-38.

Heston, Steven L., and Ronnie Sadka, 2008, Seasonality in the cross-section of stock returns, *Journal of Financial Economics* 87, 418-445.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650-705.

Hou, Kewei, Chen Xue, and Lu Zhang, 2019, Replicating anomalies, *Review of Financial Studies* forthcoming.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.

Jegadeesh, Narasimhan, and Joshua Livnat, 2006, Revenue surprises and stock returns, *Journal of Accounting and Economics* 41, 147-171.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jiang, Guohua, Charles M. C. Lee, and Yi Zhang, 2005, Information uncertainty and expected returns, *Review of Accounting Studies* 10, 185-221.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

Lev, Baruch, and Doron Nissim, 2004, Taxable income, future earnings, and equity values, *Accounting Review* 79, 1039-1074.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan, 2019, Size and value in china, *Journal of Financial Economics* forthcoming.

Loughran, Tim, and Jay W. Wellman, 2011, New evidence on the relation between the enterprise multiple and average stock returns, *Journal of Financial and Quantitative Analysis* 46, 16-29.

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2007, The new issues puzzle: Testing the investment-based explanation, *Review of Financial Studies* 21, 2825-2855.

Miller, Merton H., and Myron S. Scholes, 1982, Dividends and taxes: Some empirical evidence, *Journal of Political Economy* 90, 1118-1141.

Novy-Marx, Robert, 2011, Operating leverage, *Review of Finance* 15, 103-134.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1-28.

Ohlson, James A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109-131.

Palazzo, Berardino, 2012, Cash holdings, risk, and expected returns, *Journal of Financial Economics* 104, 162-185.

Penman, Stephen H., Scott A. Richardson, and Irem Tuna, 2007, The book-to-price effect in stock returns: Accounting for leverage, *Journal of Accounting Research* 45, 427-467.

Piotroski, Joseph D., 2000, Value investing: The use of historical financial statement information to separate winners from losers *Journal of Accounting Research* 38, 1-41.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance* 63, 921-945.

Soliman, Mark T, 2008, The use of dupont analysis by market participants, *Accounting Review* 83, 823-853.

Thomas, Jacob K., and Huai Zhang, 2002, Inventory changes and future returns, *Review of Accounting Studies* 7, 163-187.

Thomas, Jacob, and Frank X. Zhang, 2011, Tax expense momentum, *Journal of Accounting Research* 49, 791-821.

Valta, Philip, 2016, Strategic default, debt structure, and stock returns, *Journal of Financial and Quantitative analysis* 51, 1-33.