A New Four-Factor Model for Pricing China's Small Stocks

by

Yuezhou Cai

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Business and Economics Honors Program

NYU Shanghai

May 2021

Professor Marti G. Subrahmanyam Professor Christina Wang Professor Wendy Jin	Professor Jennifer N. Carpenter
Faculty Advisers	Thesis Adviser

Abstract

China's financial market is developing rapidly, and its stock market has grown around fivefold over the last decade. The establishment of an asset pricing model tailored to China's stocks is fundamental for research on the Chinese financial market. Liu, Stambaugh and Yuan (2018) propose a CH3 model that explains the cross-sectional variation of the biggest 70% of Chinese stocks. However, there doesn't exist a proper factor pricing model for the smallest 30% of stocks in China. This study aims to fill the gap by extending the CH3 model to accommodate the pricing of small stocks in China.

Keywords: Asset pricing, China's stock market, Factors, Anomalies

Acknowledgment

First of all, I would like to thank my thesis advisor, Professor Jennifer Carpenter, for your help and guidance throughout the whole year. Thank you for agreeing to be my supervisor and teaching me essential research skills during the DURF program. Thank you for accepting me to be your advisee for the Business Honors Program and always sharing your expertise and resource with me. I am especially grateful for your willingness to arrange zoom meetings and skype calls whenever necessary. I learned a lot from you, and I very much appreciate all your guidance and encouragement.

I also want to thank Professor Robert Whitelaw for your help during the whole research process. Despite the time zone difference and your busy schedule, you have been generous in sharing your ideas and knowledge with me. Your clear explanation and detailed feedback helped me a lot when I find it particularly challenging to determine the next steps.

I greatly appreciate the dedication of Professor Marti Subrahmanyam, Professor Christina Wang and Professor Jens Leth Hougaard for coordinating this program. I also want to thank Xinyi Yang, the teaching assistant of the program. Thank you for organizing seminars and make sure the program run smoothly.

Finally, I would like to thank my friends and family for the warmth and support throughout the year.

1. Introduction

A number of scholars have tested the goodness of fit of different factor pricing models for China's stock market. For example, Liu, Stambaugh & Yuan (2018) question the feasibility of applying the US three-factor model by Fama and French (1993) on the Chinese stock market because China and the US have different financial systems. They propose a CH3 factor model that adjusts the construction of the size and value factors, and they claim that their model better explains the cross-section of stock returns in China than FF3. They eliminate the smallest 30% of firms in China when constructing the size and value factors because they think these stocks are more likely to be targeted as shell companies whose prices are contaminated by the prospect of reverse mergers. Although their CH3 model explains many of the well-documented Chinese anomalies, there remain a number of other potential anomalies yet to be explored. The most serious limitation of the CH3 model is that it is restricted to pricing the biggest 70% of stocks in China and fails to price the smallest 30% of stocks.

China's A-share market mainly consists of four boards. The Shenzhen stock exchange market (SZSE) and Shanghai stock exchange market (SSE), established respectively in December 1990 and July 1991, are two main boards designed for large firms. The Small and Medium Enterprises Board (SME), established in May 2004, is primarily for firms of small and medium-size, which usually have stable profitability with a high growth rate. The other board, ChiNext, is a NASDAQ-style board within SZSE established in October 2009. ChiNext has relatively looser listing standards in terms of profitability, revenue, cash flow and net assets. By the end of January 2021, 905 companies were listed on ChiNext, with a total market value of 11073.44 billion RMB.

In contrast to the US stock market, which is dominated by a limited number of large-cap companies, China's stock market features a large amount of small-cap stocks. Chinese small-caps have expanded exponentially over the past decade. These stocks take the largest proportion of the global small-cap asset class, but research coverage for them remains minimal. Over the past twenty years, small-cap firms have been driving China’s economic shift away from traditional government-supported manufacturing industries into more innovative sectors. A significantly larger proportion of small stocks in China are listed on ChiNext. According to SZSE’s official website[1], “ChiNext serves to promote the development of innovative enterprises and other growing start-ups. It has helped encourage entrepreneurship, inspire creativity and popularize innovative business models”. The small-cap stocks reflect China’s entrepreneurial spirit and have the potential to be the country’s underlying growth drivers. Furthermore, the small-stock market in China is a relatively isolated market that can help to diversify the portfolios of domestic and foreign investors.

Table 1: The number of small and big stocks listed on different boards in China.

	Smallest 30% stocks				Biggest 70% stocks			
	SSE	SZSE	SME	ChiNext	SSE	SZSE	SME	ChiNext
2004	216	147	5	0	546	303	5	0
2005	191	137	21	0	515	272	23	0
2006	194	138	24	0	529	269	28	0
2007	220	110	69	0	576	289	64	0
2008	200	127	113	0	598	290	135	0
2009	210	131	111	0	599	295	156	0
2010	209	136	143	42	613	284	288	47
2011	183	129	196	117	675	282	384	114
2012	152	107	253	183	744	317	415	141
2013	191	131	234	139	704	303	425	187
2014	189	107	222	158	679	302	425	169
2015	245	124	203	157	700	276	460	261
2016	267	106	218	192	753	304	506	262
2017	324	88	223	268	889	319	567	329
2018	364	100	252	312	1028	343	625	399

[1] http://www.szse.cn/English/

Table 2: Listing standards for main boards, SME and ChiNext.

<table>
<tr><th colspan="2"></th><th>Main Board & SME Board</th><th colspan="2">ChiNext Market</th></tr>
<tr><td colspan="2">Targets</td><td>Large and mature firms & SMEs</td><td colspan="2">Innovative enterprises and other growing start-ups</td></tr>
<tr><td colspan="2" rowspan="2">Business integrity</td><td rowspan="2">In the last 3 years, no major adverse changes have occurred in the main business, directors and senior managers</td><td colspan="2">In the last 2 years, no major adverse changes have occurred in the main business and directors, senior managers and core technical personnel</td></tr>
<tr><td>Criteria I</td><td>Criteria II</td></tr>
<tr><td rowspan="4">Finance and Accounting</td><td>Net Profit</td><td>Positive net profit in the last 3 consecutive years and the accumulated net profit no less than RMB 30 million</td><td>positive net profit in the past 2 consecutive years, and accumulated profit no less than RMB 10 million</td><td>Positive net profit in last one year</td></tr>
<tr><td>Revenue</td><td>Accumulated revenue in past 3 years no less than RMB 300 million</td><td></td><td>Positive revenue in last one year and no less than RMB 50 million</td></tr>
<tr><td>Cash Flow</td><td>Accumulated revenue in past 3 years no less than RMB 300 million, or with accumulative net operating cash flow no less than RMB 50 million</td><td colspan="2"></td></tr>
<tr><td>Net asset</td><td>Intangible assets do not exceed 20% of net assets at the end of the latest period; There is no undistributed deficit at the end of the latest period</td><td colspan="2">Net assets of no less than RMB 20 million at the end of the most recent reporting period with no uncovered losses</td></tr>
</table>

Source: From Public Offering and Listing Standards for the ChiNext, the Shenzhen Stock Exchange and Shanghai Stock Exchange official website

In this paper, in addition to testing the robustness of the CH3 model in the context of some new potential anomalies, I provide a new CHS4 model that explains the pricing of the smallest 30% of stocks. I explore three potential anomalies: (1) the role of state ownership in the pricing of stocks, (2) stock performance subsequent to debt issuance, and (3) stock performance subsequent to equity issuance. In this analysis, I test the robustness of both the CH3 model and the new CHS4 model.

2. Data sources and sample period

I use the Wind Information Inc. (WIND) database to obtain returns, trading information, financial statements and debt and equity issuance data for all A-share stocks on the mainboards, Shenzhen Small and Medium Enterprises Board (SME) and Shenzhen Growth Enterprises Market Board (ChiNext). The risk-free rate is the one-month deposit rate from the China Stock

Market and Accounting Research (CSMAR) database. The sample period is January 2000 to December 2018.

I construct calendar-time value-weighted portfolios of treated stocks and analyze their time-series alphas with respect to various factor models. I apply the filters adopted in the CH3 model to our data. In particular, I exclude (1) stocks that have become public within the past six months, (2) stocks that have less than 120 days of trading records during the past 12 months and (3) stocks with less than 15 days of trading records during the most recent month.

3. Baseline pricing ability of alternative factor models

I first examine the ability of the existing factor pricing models to price the bottom 30% of stocks in China. I rank the stocks by market capitalization and then divide the stock universe into the bottom 30% of stocks and top 70% of stocks. Market capitalization is calculated as the A-share price times the total number of shares outstanding at the end of the previous month, including A shares, B shares, H shares, and non-tradable shares. I alternately apply the CH3, FF3, and FF5 models to the two sets of stocks. The following is a list of the factors that I use. FFSMB, FFHML and FFMKT refer to the size, value, and market factors from the FF3 model. CHMKT, CHSMB, CHVMG refers to the same set of factors from the CH3 model. FFRMW (Robust Minus Weak) is the return difference between firms with robust operating profitability and weak operating profitability. FFCMA (Conservative Minus Aggressive) is the return difference between firms that invest conservatively and invest aggressively, based on 2x2 sorts. Below is a summary of these factors:

CHMKT: The excess return on the value-weighted portfolio of top 70% of stocks

CHSMB = 1/3 (S/H+S/M+S/L) – 1/3(B/H+B/M+B/L) for the top 70% of stocks

CHVMG = 1/2 (S/V+B/V) – 1/2 (S/G + B/G) for top the 70% of stocks

FFMKT: The excess return on the value-weighted portfolio of all the stocks

FFSMB = 1/3 (S/H+S/M+S/L) – 1/3(B/H+B/M+B/L) for all the stocks

FFHML = 1/2 (S/H+B/H) – 1/2 (S/L + B/L) for all the stocks

FFRMW = 1/2 (Small Robust + Big Robust) - 1/2 (Small Weak + Big Weak).

FFCMA =1/2 (Small Conservative + Big Conservative) - 1/2 (Small Aggressive + Big Aggressive)

The regression results show that the value-weighted portfolio of the bottom 30% of stocks has large positive CH3 and FF alphas and smaller positive alphas with FF5. It is perhaps interesting to see that the CH3 alpha is particularly large and significant compared with other models.

Table 3: Regression results of the value-weighted portfolio of the bottom 30% of stocks on CH3 and FF factor models.

For the bottom 30% stocks and the top 70% respectively, the regressions estimated are:

CH3: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{CHMKT} (R_{CHMKT,t} - R_{f,t}) + \beta_{CHSMB}\ CHSMB_t + \beta_{CHVMG}\ CHVMG_t + \in_{i,t}$

FF3: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{FFSMB}\ FFSMB_t + \beta_{FFHML}\ FFHML_t + \in_{i,t}$

FF5: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{FFSMB}\ FFSMB_t + \beta_{FFHML}\ FFHML_t + \beta_{FFRMW}\ FFRMW_t + \beta_{FFVMG}\ FFVMG_t + \in_{i,t}$

The sample is from 2000 to 2018. White (1980) heteroskedasticity-consistent t-statistics are in parentheses.

	Panel A : Bottom 30% Stocks			Panel B : Top 70% Stocks		
	CH3	FF3	FF5	CH3	FF3	FF5
MKT	0.962***	0.967***	0.963***	0.994***	0.973***	0.969***
	(61.11)	(34.69)	(33.17)	(120.72)	(39.83)	(37.94)
SMB	1.098***	1.159***	1.072***	-0.02	0.023	-0.044
	(30.37)	(22.91)	(18.08)	(-0.97)	(0.59)	(-1.33)
VMG	-0.396***			0.075**		
	(-9.38)			(3.18)		
HML		0.092	0.149		0.127**	0.105*
		(1.41)	(1.95)		(3.24)	(2.42)
RMW			-0.232			-0.037
			(-1.96)			(-0.50)
CMA			0.500***			0.159
			(4.16)			(1.92)
Alpha	0.751***	0.315*	0.283*	0.012	0.047	0.09
	(7.61)	(2.42)	(2.2)	(0.16)	(0.41)	(0.79)
N	228	228	228	228	228	228

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001

4. Size and value effects within the bottom 30% of stocks

Considering that the CH3, FF3 and FF5 models all fail to price the smallest 30% of stocks, I check to see if there exist cross-sectional patterns in the returns of the smallest 30% stocks. Size and value effects are two important cross-sectional patterns for stock returns. To capture the value effect, Fama and French (1992) consider three valuation ratios: earnings-to-price (E/P), book-to-market (B/M), and assets-to-market (A/M) and find that B/M can best capture the value effect in the US stock market. Liu, Stambaugh and Yuan (2018) follow their method and find that E/P dominates the other two in capturing the value effect in the top 70% of stocks in China's A-share market. Following them, I run a Fama–MacBeth regression including these three valuation ratios to determine the one that can best capture the value effects for the bottom 30% of stocks in China.

Table 4 reports average slopes from the standard Fama–MacBeth regressions. From column (1), we can see that, similar to the US, the loadings on β are not significant. The size variable, logMC, has a very significant negative loading. From column (2) and (3), we can see that the size slopes are -0.0185 and -0.0184, with t-statistics of -6.413 and -6.417 respectively, without and with β included in the regression. These results show a particularly significant size effect for the bottom 30% of stocks in China.

Columns (4) through (7) report the regression results when each valuation ratio is included in the regression. All three ratios demonstrate significant explanatory power for return variation for those stocks. In column (8), B/M dominates other value ratios when all the ratios are included simultaneously. Thus, I use B/M to construct the value factor for those small stocks.

Table 4: Average Slopes (t-statistics) from month-by-month Fama–MacBeth regressions of Stock Returns on β, Size, Book-to-Market Equity, Leverage, and E/P for smallest 30% stocks: January 2000 to December 2018. Following Fama and French (1992), pre-ranking CAPM βs of individual stocks are estimated using the past two to five years of monthly returns (as available). logBM is the the log of book-to-market, logAM is the log of assets-to-market. E(+)/P is the ratio of earnings-to-price if the earnings are positive, and it equals zero otherwise. D(E/P) is a dummy variable that equals one if the earning is negative, and it equals zero otherwise. The average slope is the time-series average of the monthly slopes. The t-statistic is the average slope divided by the standard error from the time series. To limit the influence of outliers on the regression results, the smallest and largest 0.5% of the observations on E(+)/P, BE/ME, A/ME are set to be equal to the next largest or smallest values of the ratios of 0.005 and 0.995 fractiles, following Fama and French (1992).

Quantity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.0157	0.1596	0.1566	0.1795	0.1656	0.1554	0.1732	0.1698
	(2.361)	(6.489)	(6.525)	(7.668)	(7.01)	(6.993)	(7.71)	(7.628)
β	0.0018		0.0021	0.0025	0.0024	0.0019	0.0022	0.0019
	(1.287)		(1.566)	(1.847)	(1.771)	(1.501)	(1.729)	(1.546)
$\log MC$		-0.0185	-0.0184	-0.0208	-0.0192	-0.0184	-0.02	-0.0194
		(-6.413)	(-6.471)	(-7.476)	(-6.865)	(-7.043)	(-7.577)	(-7.498)
$\log BM$				0.0037			0.004	0.0024
				(3.904)			(4.26)	(2.238)
$\log AM$					0.005			0.0032
					(3.563)			(1.615)
E(+)/P						0.1052	0.0531	0.0222
						(2.22)	(1.157)	(0.464)
$D\,(EP < 0)$						0.0022	0.0031	0.0022
						(1.156)	(1.638)	(1.271)
R Squared	0.0063	0.0139	0.0198	0.0297	0.0324	0.0391	0.0473	0.0579

5. A four-factor model to price the smallest 30% of stocks in China

In this section, I present the new model, CHS4, which combines the market factor from the FF3 model, the size factor from the CH3 model, and two new size and value factors constructed from the bottom 30% of stocks.

5.1. Size and value factors

To examine the size and value effects in small stocks, I sort the bottom 30% of stocks into nine portfolios with equal numbers of stocks, based on a 3x3 sort on their size and value. To better see the characteristics of the bottom 30% stocks compared to the top 70% stocks, I also sort the top 70% of stocks into 25 portfolios, based on a 5x5 sort on their size and value. Table 5 reports the summary statistics for these portfolios. We can see that average returns decrease almost uniformly with size and growth. The same patterns are maintained with size when crossing from the 70% largest stocks to the 30% smallest stocks. However, the size effect seems more dramatic in the bottom 30%. As we move from the top 70% to the bottom 30%, volatilities decrease with size, but the shift is not dramatic. There is also some evidence that volatility increases in growth, but it is much less dramatic and consistent than the size effect. Skewness is small and seems to decrease in size. The excess kurtosis seems to increase with size.

The following is a list of the factors that I use for pricing the bottom 30% stocks. FFSMB, FFHML and FFMKT refer to the size, value, and market factors from the FF3 model. CHMKT, CHSMB, CHVMG refers to the same set of factors from the CH3 model. I follow Fama-French's methodology to construct two zero-cost portfolios: SMB30 and HML30 from the 3x3 sorted portfolios to mimic the risk factors associated with the special size and value effect in the bottom 30% stocks. Furthermore, I construct a market factor for the bottom 30% of stocks. Below is a summary of these factors:

MKT30: The excess return on the value-weighted portfolio of bottom 30% of stocks

SMB30 = 1/2 (S/H+ S/L) – 1/2 (B/H+B/L) after 2x2 sort on size and value

HML30 = 1/2 (S/H+ B/H) – 1/2 (S/L + B/L) after 2x2 sort on size and value

Table 5: Descriptive statistics for 3x3 and 5x5 sorted portfolios for bottom 30% stocks and top 70% stocks.

This table reports standard deviations, range, quantiles, excess kurtosis for the nine portfolios constructed from the bottom 30% stocks and the 25 portfolios constructed from the top 70% of stocks. The numbers are expressed in percent per month except for excess kurtosis and skewness. The sample period from January 2000 to December 2018 (228 months).

	Mean	Std. Dev	Min	p5	p25	Median	p75	p95	Max	kurtosis	skewness
Bottom 30% of stocks											
p1	1.95	11.40	-32.43	-14.29	-5.50	1.13	8.40	23.24	36.96	3.61	0.23
p2	2.27	11.15	-30.47	-15.49	-5.01	1.68	8.25	23.08	37.75	3.38	0.27
p3	2.41	10.75	-30.50	-13.43	-4.97	1.97	8.63	22.04	36.29	3.49	0.24
p4	1.21	10.92	-30.49	-14.40	-5.69	1.57	7.32	20.36	32.38	3.59	0.15
p5	1.83	10.43	-30.29	-13.97	-5.16	1.50	7.86	21.53	31.49	3.54	0.26
p6	1.97	10.48	-28.13	-15.17	-4.53	1.77	7.66	20.25	37.67	3.77	0.30
p7	0.89	10.61	-33.06	-15.18	-5.82	0.90	6.25	21.28	32.65	3.92	0.27
p8	1.28	10.59	-30.49	-13.95	-5.45	0.83	7.29	20.85	32.02	3.85	0.25
p9	1.62	10.20	-28.57	-13.73	-4.44	1.98	6.87	19.39	39.89	4.07	0.17
Top 70% of stocks											
p1	0.65	10.55	-32.06	-15.37	-5.77	0.56	6.61	18.91	37.94	3.90	0.15
p2	0.97	10.38	-32.34	-15.08	-5.34	1.18	6.03	18.67	33.89	3.88	0.17
p3	1.32	10.24	-31.24	-13.77	-4.48	1.24	7.03	20.21	39.67	4.30	0.27
p4	1.49	10.12	-30.52	-13.50	-4.83	1.34	7.05	19.91	38.73	4.09	0.15
p5	1.60	10.08	-29.29	-14.61	-4.12	1.18	7.18	18.73	38.54	4.25	0.12
p6	0.72	10.63	-33.59	-14.66	-5.25	0.35	6.50	20.78	31.25	3.97	0.17
p7	0.93	10.09	-31.96	-14.63	-5.45	0.44	6.65	18.38	34.04	3.98	0.18
p8	1.12	9.90	-31.29	-14.64	-4.84	0.60	6.98	19.68	31.48	3.78	0.14
p9	1.26	10.42	-30.98	-14.41	-5.00	0.89	6.58	21.36	39.11	4.26	0.34
p10	1.26	9.88	-28.14	-13.13	-4.23	1.03	6.48	18.49	44.45	4.74	0.23
p11	0.48	9.87	-31.54	-15.31	-5.41	0.36	5.52	17.96	34.19	4.08	0.19
p12	0.80	9.64	-29.44	-13.44	-5.09	0.81	6.24	18.13	34.74	4.07	0.15
p13	0.77	9.71	-30.82	-14.25	-5.26	0.57	5.35	18.92	29.44	4.08	0.06
p14	0.90	9.79	-31.28	-15.45	-5.20	0.51	5.53	19.68	34.30	3.86	0.09
p15	1.21	9.78	-27.80	-13.31	-4.53	1.11	5.88	19.64	40.75	4.37	0.23
p16	0.34	9.36	-34.54	-13.85	-4.93	0.03	5.37	16.00	29.41	4.61	0.11
p17	0.69	8.99	-28.83	-13.45	-4.19	0.77	5.64	17.70	24.31	4.11	-0.17
p18	0.69	9.90	-31.16	-15.05	-5.01	0.40	5.32	18.83	40.98	4.60	0.16
p19	0.92	9.22	-29.74	-13.22	-4.74	0.96	5.63	16.65	36.70	4.27	0.11
p20	1.20	9.52	-27.86	-11.79	-4.72	0.83	5.54	17.68	46.46	5.38	0.50
p21	0.34	8.87	-28.37	-13.90	-3.70	0.07	4.43	15.12	32.16	4.66	0.13
p22	0.07	8.40	-25.20	-12.55	-4.67	0.18	4.66	14.58	23.65	4.10	-0.08
p23	0.45	8.49	-26.85	-12.01	-4.40	0.34	5.58	14.43	25.67	4.10	-0.14
p24	0.54	7.95	-23.46	-11.51	-4.26	0.41	4.60	12.74	31.31	4.86	0.12
p25	1.27	8.27	-25.40	-11.42	-3.72	1.13	4.70	16.11	32.14	4.59	0.31

note:

Bottom 30% of stocks			
	Growth	Neutral	Value
Small	p1	p2	p3
Medium	p4	p5	p6
Large	p7	p8	p9

Top 70% of stocks					
	Low	Book-to-Market			High
Small	p1	p2	p3	p4	p5
	p6	p7	p8	p9	p10
Size	p11	p12	p13	p14	p15
	p16	p17	p18	p19	p20
Big	p21	p22	p23	p24	p25

Table 6 reports summary statistics for the nine factors in the 228-month sample period. For means, MKT30 has a much higher mean return than CHMKT, which is around three times as large. In contrast, the size and value effects measured by the factor portfolios are relatively close in magnitude. In terms of volatility, MKT30 is more volatile than CHMKT. It may be surprising to see that SMB30 is less than half as volatile as CHSMB, given the relative magnitudes of their means. The value factors have similar volatility. For higher moments, skewness is, in general, small across the board. There is some excess kurtosis, which suggests that there exist fatter tails than a normal distribution, but the number is not extreme.

Table 7 reports the correlation matrix for the nine factors. MKT30 has a relatively low correlation of 0.83 to CHMKT. Large, well-diversified stock portfolios usually have very high correlations. Therefore, this is evidence that there is independent variation in these small stocks. MKT30 has a high correlation of 0.66 with CHSMB. There exists a low correlation of 0.32 between SMB30 and CHSMB, which suggests that the size effect within MKT30 is not the same as the size effect within the top 70% of stocks. Similarly, HML30 has a relatively low correlation of 0.34 with CHVMG. The value effect is not the same in the bottom 30% and top 70% of stocks, although CH3 does a better job on the value effect than it does on the size effect. We can also see MKT30, SMB30, and HML30 have low pairwise correlations of smaller than 0.25 in magnitude, which is good because factor models are much easier to interpret when the factors are close to orthogonal. In contrast, there is a problematic high correlation of -0.62 between CHSMB and CHVMG, which makes factor loadings in the CH3 model for the top 70% of stocks hard to interpret. Value and size are not orthogonal, making size and value effects hard to disentangle.

Table 6. Summary statistics for the nine factors.

This table reports the means, standard deviations, range, quantiles, excess kurtosis for the nine factors listed above. The numbers are expressed in percent per month except for excess kurtosis and skewness. The sample period from January 2000 to December 2018 (228 months).

	Mean	Std. Dev.	Min	p5	p25	Median	p75	p95	Max	kurtosis	skewness
MKT30	1.63	10.51	-30.46	-13.69	-5.10	1.35	7.44	21.24	33.95	3.67	0.22
SMB30	0.66	1.78	-6.96	-1.77	-0.45	0.62	1.75	3.47	8.25	5.81	0.18
HML30	0.52	1.82	-6.87	-2.16	-0.51	0.41	1.65	3.63	6.64	4.52	0.14
CHMKT	0.53	7.78	-25.47	-13.45	-4.59	0.72	4.59	14.63	24.58	3.99	-0.24
CHSMB	0.77	4.40	-17.35	-5.64	-1.41	0.64	2.72	8.68	17.73	5.06	0.06
CHVMG	1.20	3.73	-10.28	-4.46	-0.91	1.13	3.66	7.01	15.17	4.42	0.20
FFMKT	0.53	7.78	-25.09	-13.12	-4.59	0.81	4.66	14.16	25.55	4.09	-0.20
FFSMB	0.67	5.36	-23.54	-7.50	-2.52	0.40	3.67	9.06	24.17	5.77	-0.11
FFHML	0.25	3.29	-14.48	-4.77	-1.49	0.27	2.07	4.86	15.87	6.91	-0.17

Table 7. Correlation matrix for the nine factors.

This table reports the pairwise correlations for all nine factors. The sample period is January 2000 through December 2018 (228 months).

	MKT30	SMB30	HML30	CHMKT	CHSMB	CHVMG	FFMKT	FFSMB	FFHML
MKT30	1.00								
SMB30	0.25	1.00							
HML30	-0.16	-0.22	1.00						
CHMKT	0.83	-0.04	0.00	1.00					
CHSMB	0.66	0.32	-0.20	0.16	1.00				
CHVMG	-0.63	-0.28	0.34	-0.29	-0.62	1.00			
FFMKT	0.80	-0.02	-0.01	0.99	0.12	-0.27	1.00		
FFSMB	0.68	0.41	-0.30	0.17	0.95	-0.76	0.15	1.00	
FFHML	-0.37	-0.25	0.50	-0.10	-0.50	0.49	-0.11	-0.54	1.00

Table 8 reports alpha and factor loadings for the nine portfolios from the bottom 30% stocks regressed on different factor models. From Panel A, we can see that there exist large alphas not explained by the CAPM model. There exist both size and value effects, with the size effect larger in magnitude. There exists little variation across portfolios for betas. From Panel B, we can see that there exist large average alphas for the bottom 30% stocks. Furthermore, there exists both large size and value effects of about the same magnitude after controlling for the MKT30. The magnitude of the size effect is larger than the value effect. From Panel C, most alphas are close to zero. The size (SMB30) betas are ordered as they should be and change signs from positive to negative as we go from small to large firms. The value (HML30) betas are ordered as they should be and change signs from positive to negative as we go from value to growth firms. The portfolios sorted based on characteristics line up with the betas of these portfolios with respect to the size and value factors.

Table 8. Alpha and factor loadings for the nine portfolios under different factor models. This table reports the alpha and factor loadings for the nine size-value sorted value-weighted portfolios regressed on different factor models. White (1980) heteroskedasticity-consistent t-statistics are reported in parentheses. Below is a list of these factor models:

Panel A: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{CHMKT} (R_{CHMKT,t} - R_{f,t}) + \epsilon_{i,t}$

Panel B: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{MKT30} (R_{MKT30,t} - R_{f,t}) + \epsilon_{i,t}$

Panel C: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{MKT30} (R_{MKT30,t} - R_{f,t}) + \beta_{SMB30}\, SMB30_t + \beta_{HML30}\, HML30_t + \epsilon_{i,t}$

Panel D: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{CHMKT} (R_{CHMKT,t} - R_{f,t}) + \beta_{CHSMB}\, CHSMB_t + \beta_{CHVMG}\, CHVMG30_t + \epsilon_{i,t}$

Panel A: Bottom 30% on CHMKT							
	Alpha			MKT			
	Low BM	Medium BM	Large BM		Low BM	Medium BM	Large BM
Small	1.359**	1.675***	1.825***	Small	1.120***	1.114***	1.097***
	(2.86)	(3.70)	(4.32)		(14.76)	(15.12)	(15.83)
Medium	0.616	1.241**	1.382***	Medium	1.123***	1.102***	1.114***
	(1.45)	(3.25)	(3.64)		(16.34)	(17.18)	(17.48)
Large	0.307	0.674	1.027**	Large	1.096***	1.140***	1.121***
	(0.76)	(1.81)	(3.00)		(15.70)	(17.91)	(18.47)

Panel B: Bottom 30% on MKT30							
	Alpha			MKT30			
	Low BM	Medium BM	Large BM		Low BM	Medium BM	Large BM
Small	0.285	0.582***	0.766***	Small	1.056***	1.047***	1.010***
	(1.63)	(4.67)	(7.02)		(58.60)	(58.35)	(77.32)
Medium	-0.444***	0.191*	0.342**	Medium	1.023***	0.987***	0.981***
	(-3.87)	(2.08)	(3.23)		(85.63)	(102.71)	(59.73)
Large	-0.728***	-0.322***	0.0214	Large	0.999***	1.001***	0.954***
	(-6.09)	(-3.41)	(0.17)		(63.20)	(74.68)	(55.60)

Panel C: Bottom 30% on CHS3				Panel D: Bottom 30% on CH3			
Alpha				Alpha			
	Low BM	Medium BM	Large BM		Low BM	Medium BM	Large BM
Small	-0.002	0.216*	0.137	Small	1.341***	1.430***	1.548***
	(-0.01)	(2.24)	(1.54)		(5.06)	(6.74)	(8.39)
Medium	-0.149	0.167	-0.047	Medium	0.530**	0.830***	0.838***
	(-1.47)	(1.60)	(-0.55)		(2.65)	(6.06)	(5.69)
Large	-0.073	-0.075	-0.048	Large	0.098	0.351*	0.642***
	(-0.84)	(-0.82)	(-0.47)		(0.65)	(2.45)	(4.82)
MKT30				CHMKT			
	Low BM	Medium BM	Large BM		Low BM	Medium BM	Large BM
Small	0.999***	1.011***	0.996***	Small	0.938***	0.939***	0.938***
	(64.30)	(74.42)	(93.35)		(23.97)	(32.45)	(32.31)
Medium	1.011***	0.986***	0.999***	Medium	0.950***	0.960***	0.982***
	(93.53)	(103.98)	(86.99)		(38.70)	(50.78)	(43.99)
Large	0.990***	1.021***	0.989***	Large	0.933***	0.996***	0.997***
	(117.51)	(87.14)	(83.48)		(46.45)	(50.64)	(51.82)
SMB30				CHSMB			
	Low BM	Medium BM	Large BM		Low BM	Medium BM	Large BM
Small	0.918***	0.720***	0.617***	Small	1.098***	1.205***	1.129***
	(8.07)	(10.68)	(9.75)		(12.07)	(21.98)	(18.18)
Medium	0.003	-0.002	0.087	Medium	1.085***	1.116***	1.145***
	(0.04)	(-0.03)	(1.62)		(17.04)	(27.98)	(22.90)
Large	-0.393***	-0.530***	-0.441***	Large	1.109***	1.070***	0.991***
	(-7.17)	(-6.44)	(-6.59)		(20.02)	(25.00)	(21.70)
HML30				CHVMG			
	Low BM	Medium BM	Large BM		Low BM	Medium BM	Large BM
Small	-0.549***	-0.150**	0.456***	Small	-0.608***	-0.491***	-0.423***
	(-7.67)	(-2.69)	(8.02)		(-6.28)	(-6.55)	(-6.18)
Medium	-0.562***	0.01	0.644***	Medium	-0.548***	-0.311***	-0.223***
	(-8.43)	(1.79)	(11.86)		(-7.82)	(-6.55)	(-3.55)
Large	-0.763***	0.077	0.674***	Large	-0.465***	-0.353***	-0.261***
	(-14.99)	(1.44)	(11.86)		(-7.85)	(-6.56)	(-4.33)

Panel D shows that the CH3 model does not explain large average alpha and fails to explain the size effect. CHVMG betas are almost all negative, which hurts the ability of the CH3 model to explain average alpha. All CHSMB betas are large and positive, but do not exhibit much variation across portfolios. Thus, as Stambaugh et al. conclude, it appears that the bottom 30% of stocks are different. A model that explains the top 70% of stocks does not explain the bottom 30%. The big puzzles are the high average alphas. These small stocks have high returns, and the alphas seem to depend on size in a way that is not captured by the size factor in the top 70% of stocks.

Table 9. Alpha and factor loadings for the 25 portfolios regressed on different factor models. This table reports the alpha and factor loadings for the 25 size-value sorted value-weighted portfolios regressed on different factor models. Below is a list of these factor models. White (1980) heteroskedasticity-consistent t-statistics are reported in parentheses.

Panel D: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{CHMKT} (R_{CHMKT,t} - R_{f,t}) + \in_{i,t}$

Panel E: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{MKT30} (R_{MKT30,t} - R_{f,t}) + \beta_{SMB30}\, SMB30_t + \beta_{HML30}\, HML30_t + \in_{i,t}$

Panel F: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{CHMKT} (R_{CHMKT,t} - R_{f,t}) + \beta_{CHSMB}\, CHSMB_t + \beta_{CHVMG} CHVMG30_t + \in_{i,t}$

Panel E: Top 70% on CHMKT											
	Alpha						CHMKT				
	Low		Book-to-market		High		Low		Book-to-market		High
Small	0.0657	0.377	0.727*	0.897*	1.015**	Small	1.093***	1.117***	1.109***	1.120***	1.108***
	(0.16)	(1.03)	(2.05)	(2.55)	(2.98)		(17.95)	(18.00)	(17.32)	(17.38)	(18.37)
	0.131	0.345	0.544	0.641	0.666*		1.107***	1.099***	1.083***	1.160***	1.124***
	(0.32)	(1.00)	(1.64)	(1.92)	(2.24)		(17.74)	(19.15)	(18.80)	(18.55)	(19.75)
Size	-0.0816	0.245	0.194	0.302	0.613*	Size	1.064***	1.052***	1.093***	1.128***	1.131***
	(-0.23)	(0.74)	(0.64)	(1.07)	(2.21)		(19.79)	(20.15)	(20.58)	(23.27)	(22.23)
	-0.180	0.155	0.0828	0.339	0.606*		0.978***	1.009***	1.150***	1.094***	1.116***
	(-0.51)	(0.55)	(0.30)	(1.47)	(2.44)		(17.53)	(22.47)	(23.91)	(26.10)	(22.47)
Big	-0.173	-0.467*	-0.0948	0.0449	0.776**	Big	0.959***	1.009***	1.021***	0.928***	0.933***
	(-0.55)	(-2.39)	(-0.49)	(0.21)	(3.03)		(22.82)	(32.26)	(33.39)	(23.94)	(22.56)

Panel F: Top 70% on CHS3						Panel G: Top 70% on CH3					
	Alpha						Alpha				
	Low		Book-to-market		High		Low		Book-to-market		High
Small	-0.415*	-0.104	-0.0988	-0.0310	-0.248	Small	-0.026	0.0960	0.250	0.489***	0.645***
	(-2.51)	(-0.68)	(-0.55)	(-0.27)	(-1.44)		(-0.13)	(0.63)	(1.86)	(3.53)	(3.41)
	0.003	0.0344	0.0350	-0.263	-0.480**		0.037	0.0216	0.224	0.115	0.193
	(0.01)	(0.19)	(0.23)	(-1.67)	(-2.76)		(0.14)	(0.13)	(1.59)	(0.83)	(1.04)
Size	-0.0842	-0.0676	-0.217	-0.139	-0.286	Size	0.002	0.0402	-0.309	-0.0312	0.087
	(-0.36)	(-0.29)	(-1.13)	(-0.74)	(-1.55)		(0.01)	(0.19)	(-1.76)	(-0.20)	(0.45)
	0.338	0.325	-0.225	-0.125	-0.257		-0.089	-0.0811	-0.0460	0.0489	0.259
	(0.98)	(1.20)	(-0.90)	(-0.56)	(-1.05)		(-0.30)	(-0.35)	(-0.22)	(0.32)	(1.04)
Big	0.580	-0.418	0.0289	-0.221	0.419	Big	0.279	-0.0539	-0.0361	-0.371	0.451
	(1.23)	(-1.09)	(0.07)	(-0.60)	(0.99)		(0.84)	(-0.27)	(-0.19)	(-1.90)	(1.59)
	MKT30						CHMKT				
	Low		Book-to-market		High		Low		Book-to-market		High
Small	0.966***	0.988***	0.974***	0.976***	0.967***	Small	0.938***	0.980***	0.985***	1.003***	0.995***
	(45.91)	(65.45)	(53.71)	(65.06)	(63.40)		(36.94)	(46.93)	(57.88)	(58.62)	(40.04)
	0.960***	0.946***	0.948***	1.002***	0.944***		0.965***	0.977***	0.961***	1.047***	1.039***
	(39.42)	(41.10)	(57.82)	(57.41)	(41.76)		(30.57)	(42.98)	(53.55)	(57.08)	(35.46)
Size	0.891***	0.892***	0.925***	0.939***	0.927***	Size	0.936***	0.942***	1.002***	1.036***	1.064***
	(38.73)	(37.74)	(39.91)	(53.03)	(41.67)		(35.86)	(39.88)	(45.03)	(53.26)	(35.76)
	0.797***	0.812***	0.924***	0.869***	0.872***		0.876***	0.934***	1.063***	1.028***	1.066***
	(24.98)	(29.40)	(36.54)	(37.81)	(27.50)		(21.37)	(28.33)	(41.39)	(44.74)	(30.16)
Big	0.640***	0.626***	0.603***	0.555***	0.557***	Big	0.932***	1.014***	1.066***	0.998***	1.001***
	(15.76)	(16.10)	(14.15)	(12.54)	(13.73)		(20.27)	(33.47)	(42.93)	(29.53)	(25.54)
	SMB30						CHSMB				
	Low		Book-to-market		High		Low		Book-to-market		High
Small	-0.438***	-0.810***	-0.497***	-0.572***	-0.423***	Small	0.981***	1.000***	1.052***	0.964***	0.913***
	(-4.48)	(-7.75)	(-3.72)	(-7.51)	(-4.25)		(15.80)	(20.60)	(25.46)	(22.95)	(15.42)
	-0.814***	-0.889***	-0.860***	-0.700***	-0.577***		0.909***	0.937***	0.939***	1.017***	0.816***
	(-4.94)	(-8.95)	(-11.08)	(-7.08)	(-4.98)		(14.80)	(17.84)	(22.73)	(21.60)	(14.45)
Size	-0.969***	-0.829***	-1.028***	-1.070***	-0.826***	Size	0.710***	0.792***	0.873***	0.761***	0.744***
	(-6.58)	(-7.30)	(-8.97)	(-9.77)	(-7.22)		(9.14)	(14.25)	(16.70)	(16.85)	(11.13)
	-1.359***	-1.331***	-1.123***	-1.001***	-0.866***		0.552***	0.602***	0.606***	0.584***	0.522***
	(-5.76)	(-8.05)	(-7.28)	(-6.90)	(-5.25)		(6.38)	(10.66)	(10.93)	(12.33)	(6.99)
Big	-1.542***	-1.057***	-1.243***	-0.890***	-0.967***	Big	-0.132	-0.300***	-0.309***	-0.147*	-0.198*
	(-5.17)	(-5.09)	(-5.02)	(-4.43)	(-3.96)		(-1.28)	(-4.94)	(-5.74)	(-2.33)	(-2.14)
	HML30						CHVMG				
	Low		Book-to-market		High		Low		Book-to-market		High
Small	-0.313***	-0.0819	0.0527	0.205***	0.431***	Small	-0.484***	-0.347***	-0.224***	-0.227***	-0.228**
	(-5.83)	(-1.87)	(1.07)	(5.40)	(9.01)		(-6.64)	(-5.72)	(-4.33)	(-4.27)	(-2.68)
	-0.376***	-0.129**	0.0302	0.245***	0.473***		-0.442***	-0.277***	-0.282***	-0.165**	-0.0915
	(-4.37)	(-2.69)	(0.56)	(4.86)	(8.85)		(-5.13)	(-4.22)	(-5.31)	(-3.25)	(-1.10)
Size	-0.294***	-0.102	0.0947	0.147**	0.438***	Size	-0.468***	-0.289***	-0.101	-0.170**	-0.010
	(-3.90)	(-1.64)	(1.49)	(2.81)	(6.66)		(-5.47)	(-4.18)	(-1.68)	(-3.03)	(-0.11)
	-0.411***	-0.107	0.0918	0.221***	0.516***		-0.384***	-0.156*	-0.243**	-0.104	-0.023
	(-4.49)	(-1.50)	(1.19)	(3.78)	(6.64)		(-3.46)	(-2.11)	(-3.28)	(-1.78)	(-0.24)
Big	-0.256	0.133	0.235	0.401**	0.538***	Big	-0.279*	-0.153*	0.129	0.409***	0.367**
	(-1.79)	(1.09)	(1.64)	(2.79)	(3.95)		(-2.26)	(-2.29)	(1.95)	(5.76)	(3.32)

Table 9 reports alphas and factor loadings for the 25 portfolios from the top 70% stocks under different factor models. From panel E, we can see that there are clear size and value effects. Average alphas across portfolios appear positive, and the alphas are very significant under the CAPM model. From panel G, we can see that after regressing on CH3, alphas become close to zero and insignificant. Both the size and value betas show the correct ordering in general, but the magnitudes are problematic. The size betas are predominantly large and positive, while the value betas are predominantly large and negative. Although the CH3 model appears to price the top 70% of stocks, it is not an ideal model because the factor loadings are difficult to interpret in an economic sense.

5.2 Comparing size and value factors

Table 10 reports the results of the market, size and value factors from the bottom 30% of stocks when regressing on the set of factors from the top 70% of stocks. In column (1), the model says that MKT30 is primarily comprised of small and growth stocks. The regression leaves large and significant alpha. From column (2), we can see that the betas are all close to zero. The model has no explanatory power for the size effect in the bottom 30%. This is consistent with the earlier results that the CH3 betas do not vary much across the 3x3 sort of bottom 30% stocks. The size effect within the bottom 30% of stocks seems to be independent. From column (3), we can see that the HML30 loads significantly on CHVMG, as it should if the value effects are common to both sets of stocks. However, the loading is not sufficient to eliminate the value effect, and the alpha is still significant, albeit only about half the size of the factor return. The CH3 model R-squared is low for SMB30 and HML30. Size and value in the top 70% do not completely explain variation in size and value in the bottom 30%. In other words, while the top 70% size may explain the average level of returns in the bottom 30%, there may still be cross-sectional variations in the bottom 30% that are left unexplained. In contrast, despite the alpha still being large and statistically significant, the R-squared for the regression of MKT30 on the CH3 factors is very high. The R-squared is close to 98%, but the alpha is still 0.75% per month. The model explains most of the variation, but there is still excess return. One way to see this is to compare the monthly Sharpe ratio of the MKT30 portfolio to its information ratio. The Sharpe ratio is:

$$SR = \frac{1.63\%}{10.51\%} = 0.155$$

The information ratio is

$$IR = \frac{\alpha}{\sigma_\varepsilon} = \frac{0.751\%}{\sqrt{(1-R^2)\sigma^2}} = \frac{0.751\%}{\sqrt{(1-0.9786)10.51\%^2}} = 0.488$$

We can see that the information ratio is about three times as large as the original SR.

Table 10: (1) OLS regression of MKT30 on CH3, (2) OLS regression of SMB30 on CH3 (3) OLS regression of HML30 on CH3 (4) OLS regression of MKT30 on CHSMB (5) OLS regression of MKT30 on CHMKT. White (1980) heteroskedasticity-consistent t-statistics are reported in parentheses.

	(1)	(2)	(3)	(4)	(5)
	MKT30	SMB30	HML30	MKT30	MKT30
CHMKT	0.962***	-0.030	0.026	1.000***	1.115***
	(61.11)	(-1.77)	(1.48)	(49.63)	(17.24)
CHSMB	1.098***	0.093*	0.007	1.295***	
	(30.37)	(2.47)	(0.16)	(36.22)	
CHVMG	-0.396***	-0.084*	0.184***		
	(-9.38)	(-1.99)	(3.57)		
Alpha	0.751***	0.708***	0.276	0.102	1.040**
	(7.61)	(5.92)	(1.87)	(0.81)	(2.72)
Adj R-squared	97.86%	11.58%	11.16%	96.72%	67.93%

t statistics in parentheses
* p<0.05 ** p<0.01 *** p<0.001

From column (4), we can see that if we drop CHVMG and run the 2-factor model, the R-squared declines very little, but the alpha is eliminated. In other words, CHVMG offers little explanatory power, but including it in the factor model blows up the alpha. The size factor in the top 70% does explain the average returns on the bottom 30%, as long as we don't include the CH3 value factor. Perhaps the reason is that the CHVMG is highly negatively correlated with the CH3 size factor. As a result, it destroys the ability of the model to price the bottom 30%. From column (5), when we drop CHSMB, the R-squared goes down, and the alpha goes back up, just as it should. Thus, the size factor CHSMB explains the return variation and the alpha.

Given that CHVMG seems to be a problem for pricing small stocks, I run regressions of the nine size-value sorted portfolios on a 2-factor model with CHMKT and CHSMB. Table 11 below reports the alpha and factor loadings from the regression. From table 11, we can see that there is a clear unexplained cross-sectional variation in average returns. The size premium is about 1% across all columns, and the value premium is about the same magnitude. There are more positive alphas than negative alpha portfolios, but large stock portfolios have lower negative alphas. CHSMB does little in explaining the size premium within the bottom 30% of stocks. However, including CHSMB in the regression reduces the average alpha very significantly. Thus, all stocks in the bottom 30% have a return premium associated with their size that is consistent with those in the top 70%, but the size premium within the bottom 30% is not explained by CHSMB.

Table 11: Alpha and factor loadings from nine portfolios on CHMKT and CHSMB (CH2). The table reports alpha and slope coefficients for the nine size-value sorted portfolios regressed on CH2 factors. The regression estimated is:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{CHMKT} (R_{CHMKT,t} - R_{f,t}) + \beta_{CHSMB}\, CHSMB_t + \epsilon_{i,t}$$

t-statistics based on the heteroskedasticity-consistent standard errors of White (1980) are reported in parentheses.

Panel E: Bottom 30% on CHMKT,CHSMB			
Alpha			
	Low BM	Medium BM	Large BM
Small	0.344	0.625**	0.855***
	(1.25)	(3.08)	(4.41)
Medium	-0.368	0.321*	0.473**
	(-1.94)	(2.19)	(3.21)
Large	-0.663***	-0.228	0.214
	(-4.15)	(-1.70)	(1.55)
CHMKT			
	Low BM	Medium BM	Large BM
Small	0.996***	0.986***	0.979***
	(23.32)	(32.46)	(31.28)
Medium	1.003***	0.989***	1.003***
	(35.60)	(47.37)	(41.56)
Large	0.977***	1.030***	1.022***
	(39.68)	(43.40)	(45.62)
CHSMB			
	Low BM	Medium BM	Large BM
Small	1.401***	1.450***	1.340***
	(18.49)	(28.21)	(26.60)
Medium	1.358***	1.271***	1.256***
	(24.55)	(34.71)	(30.96)
Large	1.340***	1.246***	1.122***
	(25.87)	(34.03)	(25.03)

The size and value effects within the bottom 30% of stocks present a challenge to evaluating the performance of anomaly portfolios and call for an extended model that eliminates these size and value effects. To this end, I incorporate the SMB30 and HML30 into the FF or CH2 models to price these small stocks. Fama and French (1993) mentioned that applying different sorting methods to construct size and value factors should not change the results significantly. However, considering the relatively small data sample, I adopt different sorting methods to construct the size and value factors for the bottom 30% of stocks and develop different pricing models to price those stocks. Below are additional size and value factors that I later incorporate into the models for testing:

FFSMB30 = 1/3 (S/H+S/M+S/L) – 1/3(B/H+B/M+B/L) after 3x3 sort on size and value

FFHML30 = 1/2 (S/H+ B/H) – 1/2 (S/L + B/L) after 3x3 sort on size and value

In unreported results, I explore numerous models that fail to eliminate the alphas of the nine size-value portfolios of the small stocks. I report results only for the following four models as they eliminate the alphas and explain the cross-sectional return variation. Here are the details of these four models:

Model A: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{CHSMB}\, CHSMB_t + \beta_{FFSMB30}\, FFSMB30_t$
$+ \beta_{FFSMB30}\, FFHML30_t + \in_{i,t}$

Model B: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{CHSMB}\, CHSMB_t + \beta_{SMB30}\, SMB30_t$
$+\beta_{SMB30}\, HML30_t + \in_{i,t}$

Model C: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{SMB30}\, SMB30_t + \beta_{SMB30}\, HML30_t + \in_{i,t}$

Model D: $R_{i,t} - R_{f,t} = \alpha_i + \beta_{CHMKT} (R_{CHMKT,t} - R_{f,t}) + \beta_{CHSMB}\, CHSMB_t + \beta_{SMB30}\, SMB30_t$
$+\beta_{SMB30}\, HML30_t + \in_{i,t}$

For each model, Table 12 below presents the alphas of the nine size-value sorted portfolios of small stocks from the time-series regressions described above. The table shows that while all four models produce insignificant alphas for the nine size-value portfolios, Model B and Model D outperform Model A and Model C, as they produce smaller alphas with smaller t-statistics. Model B and Model D rival each other in their ability to price the bottom 30% of stocks. They are constructed in similar ways except that they use different stock universes to construct the market factor. In the following analysis, I use Model B, a four-factor model with FFMKT, CHSMB, SMB30 and HML30 as the benchmark pricing model.

Table 12: Alphas of the nine size-value sorted portfolios on different pricing models

Model A				Model B			
Intercept				Intercept			
	Low BM	Medium BM	High BM		Low BM	Medium BM	High BM
Small	-0.150	0.009	-0.093	Small	0.051	0.237	0.176
	(-0.59)	(0.04)	(-0.37)		(0.20)	(1.14)	(0.75)
Medium	-0.343	0.068	0.033	Medium	-0.083	0.184	-0.009
	(-1.42)	(0.28)	(0.13)		(-0.40)	(0.91)	(-0.04)
Big	-0.068	-0.041	-0.125	Big	-0.059	0.040	0.009
	(-0.30)	(-0.14)	(-0.56)		(-0.29)	(0.15)	(0.04)
Model C				Model D			
Intercept				Intercept			
	Low BM	Medium BM	High BM		Low BM	Medium BM	High BM
Small	0.333	0.557	0.424	Small	0.018	0.207	0.145
	(0.70)	(1.11)	(0.88)		(0.08)	(1.16)	(0.71)
Medium	0.180	0.589	0.547	Medium	-0.114	0.155	-0.042
	(0.37)	(1.20)	(1.10)		(-0.62)	(0.84)	(-0.22)
Big	0.458	0.488	0.367	Big	-0.090	0.007	-0.023
	(0.96)	(0.95)	(0.78)		(-0.49)	(0.03)	(-0.12)

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001"

6. Factor model explanations of anomalies

6.1 State ownership

Carpenter, Lu and Whitelaw (2020) stress the influence of state ownership on stocks' price informativeness and argue that the returns of firms with greater state ownership are more challenging to predict, especially after the postcrisis stimulus in 2009. I use their definition of the fraction of shares owned by the state (SOE) in our analysis, and I run both full sample period (2000-2018) and sub-period regressions (2009-2018). Before the split-share reform in 2005, the variable SOE is calculated as the proportion of non-tradable state-owned shares out of the total number of shares. After 2005, SOE is defined as the number of shares held by the "top ten shareholders which are state entities" divided by the total number of shares. I calculate SOE for each stock for each year in our sample based on data from annual reports on WIND.

For the top 70% of stocks, I construct one portfolio with SOE equal to zero and then divide all remaining stocks into ten decile portfolios based on firms' SOE, each containing an equal number of stocks. Decile 1 stands for the portfolio of firms with the lowest SOE, and decile 10 stands for the portfolio of firms with the highest SOE. For the smallest 30% stocks, to ensure a sufficiently large number of stocks in each portfolio, I divide the stocks with non-zero SOE into five quintile portfolios instead of ten decile portfolios. The following two tables show the results of time-series regressions of the returns on these SOE-sorted portfolios using the CH3 model for the largest 70% of stocks and the CHS4 model for the smallest 30% of stocks.

As Table 13 shows, there is no particular pattern in the alphas across the SOE-portfolios in either period 2000-2018 or 2009-2018. The alphas are insignificant and close to zero for almost all portfolios. In other words, the CH3 model and the CHS4 model explain the variation of returns of stocks with different proportions of state ownership.

Table 13: Alphas and factor loadings for each decile portfolio under CH3 and CHS4 factor models.

The table reports alphas and factor loadings for each decile portfolio under CH3 and CHS4 factor models, respectively. After separating out one portfolio with stocks whose SOE equal to zero, the remaining stocks are sorted into ten equal decile portfolios based on the SOE variable. The regression estimated are:

CH3: $R_t = \alpha + \beta_{CHMKT}\ (R_{CHMKT,t} - R_{f,t}) + \beta_{CHSMB}\ CHSMB_t + \beta_{CHVMG}\ CHVMG_t + \epsilon_t$

CHS4: $R_t = \alpha + \beta_{CHMKT}\ (R_{CHMKT,t} - R_{f,t}) + \beta_{CHSMB}\ CHSMB_t + \beta_{SMB30}\ SMB30_t +$

$\beta_{HML30}\ HML30_t + \epsilon_t$

All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Top 70% stocks on CH3											
Panel A: 2000-2018											
	SOE=0	decile 1	2	3	4	5	6	7	8	9	decile 10
MKT	0.969***	1.012***	1.081***	1.086***	1.016***	1.100***	1.077***	1.032***	1.059***	1.029***	0.966***
	(33.56)	(32.73)	(32.02)	(28.90)	(28.28)	(36.87)	(38.01)	(38.76)	(31.44)	(25.64)	(27.98)
SMB	0.310***	0.341***	0.171*	0.167*	0.106	0.228***	0.208**	0.267***	0.114	-0.063	-0.215**
	(5.70)	(4.29)	(2.07)	(2.30)	(1.29)	(3.61)	(2.92)	(4.06)	(1.57)	(-0.84)	(-2.68)
VMG	-0.0768	-0.129	-0.054	0.056	-0.004	0.053	0.013	-0.026	0.094	0.216*	0.260**
	(-1.07)	(-1.85)	(-0.57)	(0.60)	(-0.05)	(0.69)	(0.16)	(-0.35)	(1.20)	(2.40)	(2.65)
Constant	0.082	0.111	-0.070	0.186	-0.154	-0.182	0.049	0.064	-0.277	0.099	0.183
	(0.46)	(0.54)	(-0.26)	(0.88)	(-0.73)	(-0.98)	(0.25)	(0.35)	(-1.34)	(0.53)	(1.11)
Panel B: 2009-2018											
MKT	0.948***	0.966***	1.114***	1.113***	1.133***	1.117***	1.089***	1.081***	1.154***	1.059***	0.871***
	(33.74)	(30.92)	(33.21)	(26.19)	(26.12)	(34.76)	(23.19)	(33.54)	(26.85)	(28.87)	(47.23)
SMB	0.423***	0.484***	0.212**	0.161*	-0.025	0.038	0.026	0.051	-0.036	-0.359***	-0.405***
	(8.03)	(7.19)	(3.12)	(2.04)	(-0.32)	(0.64)	(0.26)	(0.72)	(-0.37)	(-6.60)	(-11.51)
VMG	-0.264***	-0.274***	-0.08	0.118	0.066	0.162*	-0.013	-0.045	0.096	0.390***	0.324***
	(-3.83)	(-3.96)	(-0.95)	(1.29)	(0.83)	(2.38)	(-0.13)	(-0.59)	(0.91)	(5.99)	(7.16)
Constant	0.189	0.138	0.120	-0.064	0.152	-0.08	0.131	0.258	-0.278	0.202	0.006
	(1.07)	(0.66)	(0.57)	(-0.29)	(0.70)	(-0.44)	(0.51)	(1.17)	(-1.10)	(1.24)	(0.04)

Bottom 30% stocks on CHS4												
	Panel A: 2000-2018						Panel B: 2009-2018					
	SOE=0	quantile 1	2	3	4	quantile 5	SOE=0	quantile 1	2	3	4	quantile 5
FFMKT	0.978***	0.998***	0.985***	0.982***	0.988***	1.007***	1.032***	1.055***	1.003***	1.025***	1.027***	1.034***
	(33.73)	(25.23)	(27.35)	(25.74)	(30.42)	(32.86)	(33.89)	(32.99)	(25.14)	(30.57)	(35.21)	(27.76)
CHSMB	1.304***	1.261***	1.307***	1.221***	1.231***	1.190***	1.206***	1.126***	1.118***	0.964***	1.040***	0.994***
	(28.05)	(22.04)	(19.76)	(16.34)	(18.49)	(16.39)	(23.61)	(21.07)	(14.95)	(12.78)	(14.52)	(9.33)
SMB30	0.497***	0.478***	0.531***	0.508***	0.306**	0.538***	0.690***	0.740***	0.822***	0.583***	0.326*	0.508**
	(3.65)	(3.35)	(4.40)	(4.81)	(2.72)	(4.17)	(6.77)	(6.95)	(6.98)	(3.47)	(2.42)	(2.98)
HML30	-0.362***	-0.131	-0.103	0.206	0.110	0.0967	-0.462***	-0.173	-0.197	0.207	0.119	0.108
	(-3.54)	(-1.25)	(-0.88)	(1.61)	(1.20)	(0.89)	(-4.47)	(-1.82)	(-1.83)	(1.35)	(1.11)	(0.77)
Alpha	0.183	0.259	0.201	-0.129	0.116	0.187	0.275	0.280	0.0337	-0.242	0.0109	0.310
	(0.94)	(1.06)	(0.97)	(-0.54)	(0.57)	(0.83)	(1.39)	(1.41)	(0.15)	(-0.99)	(0.05)	(1.18)

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001

6.2 Debt issuance and equity issuance

The other two potential anomalies I consider relate to stock return performance after debt issuance and equity issuance. The performance of equity prices in the US after seasoned equity offerings (SEO) are studied by Loughran and Ritter (1995), who find that issuing firms significantly underperform non-issuing firms in the five-year period after SEO. Spiess and Affleck-Graves (1995) argue that debt offerings in the US are similar to equity offerings in that they are also signals that the firm is overvalued. They find a significant long-run underperformance for firms that issue straight or convertible debt. They also document that the market seems to underreact around the debt issuance announcement, and the full impact is realized in the long run. However, only a few studies focus on equity performance after SEOs and debt issuance in China. The issuance of seasoned equity and debt in China has changed over time with the transformation of China's market and changes in regulation. Since the split-share reform in 2005, the size of total debt and seasoned equity issuance has soared fast as these have become the primary instruments for financing. The total size of the debt and seasoned equity offerings is heavily influenced by national policy and regulations. From 2005 to 2018, the overall trend of the debt and seasoned equity offerings increases, and the variations correspond to the change in monetary policy and interest rates. The sharp increase from 2014 to 2016 is largely due to the loosening of restrictions on debt issuance.

However, only a few studies focus on equity performance after SEOs and debt issuance in China. The issuance of seasoned equity and debt in China has changed over time with the transformation of China’s market and changes in regulation. Since the split-share reform in 2005, the size of total debt and seasoned equity issuance have soared fast as these have become the primary instruments for financing. The total size of debt and seasoned equity offerings is heavily influenced by national policy and regulations. From 2005 to 2018, the overall trend is increasing, and the variations correspond to the change in monetary policy and interest rates. The sharp increase from 2014 to 2016 is largely due to the loosening of restrictions on debt issuance.

There are three ways for firms to issue seasoned equity in China: rights issues, public offerings, and private placements. I omit the analysis of rights issues in our analysis, as the sample size is small, and rights issues have become less popular among firms in the recent decade.

Government intervention is quite apparent in the equity market. For example, after "The Administration of the Issuance of Securities by Listed Companies" that clarified the regulation of private placements was issued by the China Securities Regulatory Commission (CSRC) in 2006, private placement became the most popular capital-raising tool for listed firms. Following the release of several documents[2] that stimulated the equity market, there was a pronounced increase in equity issuance from 2013 to 2016. The subsequent decrease from 2016 to 2018 followed a tightened regulation in equity market.

[2] "Several opinions of the State Council on Further Promoting the Healthy Development of the Capital Market"（《国务院关于进一步促进资本市场健康发展的若干意见》)；

Table 14: Distribution of debt issuance and seasoned equity offerings by year (2000 –2018).

Year	Total debt issuance size (bn RMB)	Straight Debt		Convertible debt	
		Number of firms	Issue size (bn RMB)	Number of firms	Issue size (bn RMB)
2000	2.88	1	0.03	2	2.85
2001	0.40	1	0.40	0	0.00
2002	3.83	0	0.00	4	3.83
2003	16.55	0	0.00	15	16.55
2004	42.41	4	21.11	13	21.30
2005	60.20	27	60.20	0	0.00
2006	115.33	73	113.45	5	1.88
2007	297.33	90	274.58	12	22.75
2008	264.78	74	257.06	5	7.72
2009	595.78	119	591.12	6	4.66
2010	528.51	143	456.78	8	71.73
2011	836.85	265	795.53	9	41.32
2012	1070.78	354	1054.42	5	16.36
2013	955.02	303	900.54	8	54.48
2014	1933.81	327	1901.71	13	32.10
2015	5127.81	429	5118.01	3	9.80
2016	9467.49	468	9441.68	13	25.80
2017	12163.55	340	12068.93	40	94.62
2018	14385.49	376	14304.89	68	80.60
Total	47868.77	3394	47360.42	229	508.35

Year	Total SEO size (bn RMB)	Public offerings		Private placements	
		Number of firms	Capital raised (bn RMB)	Number of firms	Capital raised (bn RMB)
2000	22.68	24	22.68	0	0.00
2001	23.66	25	23.66	0	0.00
2002	17.38	28	17.38	0	0.00
2003	7.87	13	7.87	0	0.00
2004	6.84	11	6.84	0	0.00
2005	26.44	3	26.44	0	0.00
2006	98.73	6	10.03	48	88.70
2007	292.09	32	66.60	129	225.49
2008	192.50	25	44.37	100	148.13
2009	269.28	13	22.53	114	246.75
2010	331.14	10	36.85	149	294.29
2011	374.59	10	28.40	175	346.19
2012	326.28	6	11.00	149	315.28
2013	43.18	5	7.13	240	36.05
2014	521.01	1	0.33	370	520.68
2015	892.05	0	0.00	594	892.05
2016	1277.66	0	0.00	641	1277.66
2017	906.93	0	0.00	421	906.93
2018	689.68	0	0.00	236	689.68
Total	6319.97	212	332.09	3366	5987.88

Liu (2016) examines the short-term market reaction to different forms of SEOs in China and argues that investors react negatively to the announcement of both private placements and public offerings. I hypothesize that there also exists a negative long-term abnormal return after an SEO, as existing shareholders may sell overvalued shares to new shareholders at an inflated price.

Following Spiess and Affleck-Graves (1995), I hypothesize that there also exists a long-run post-issuance underperformance by firms that issue straight and convertible debt in China. Below are the regression results for the issuers and non-issuers of seasoned equity, categorized by whether the firm issued equity during the prior 1, 3 or 5 years. The column labeled Difference reports results for the portfolio that goes long issuers and short non-issuers.

Table 15: Alphas and factor loadings for issuers and non-issuers of seasoned equity under CH3 and CHS4 models, categorized by whether the firm issued equity during the prior 1, 3 or 5 years (2000 –2018).

2000--2018 total SEO for top 70% of stocks

	prior 1 year			prior 3 years			prior 5 years		
	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference
MKT	1.067***	0.987***	0.081	1.092***	0.974***	0.118**	1.090***	0.969***	0.121***
	(27.16)	(98.97)	(1.84)	(36.88)	(77.64)	(3.10)	(41.42)	(73.25)	(3.45)
SMB	0.029	-0.037	0.066	0.0223	-0.065*	0.0870	0.010	-0.069*	0.079
	(0.28)	(-1.56)	(0.60)	(0.30)	(-2.17)	(0.96)	(0.15)	(-2.25)	(0.97)
VMG	-1.018***	0.091***	-1.109***	-0.068	0.108**	-0.176	-0.090	0.121***	-0.211
	(-6.61)	(3.35)	(-6.91)	(-0.60)	(3.27)	(-1.41)	(-0.81)	(3.44)	(-1.71)
Constant	0.491	0.022	0.468	0.226	0.044	0.182	0.259	0.042	0.217
	(1.23)	(0.27)	(1.16)	(0.75)	(0.48)	(0.57)	(0.89)	(0.44)	(0.71)

2000--2018 total SEO for bottom 30% of stocks

	prior 1 year			prior 3 years			prior 5 years		
	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference
FFMKT	0.756***	0.989***	-0.232***	0.895***	0.988***	-0.0936*	0.978***	0.990***	-0.0118
	(13.52)	(32.30)	(-4.52)	(17.51)	(32.52)	(-2.24)	(21.05)	(32.75)	(-0.36)
CHSMB	1.094***	1.266***	-0.172	1.177***	1.267***	-0.0900	1.270***	1.273***	-0.00259
	(9.63)	(23.89)	(-1.52)	(10.63)	(24.25)	(-0.86)	(14.62)	(24.72)	(-0.04)
SMB30	-0.000	0.486***	-0.486	0.515	0.493***	0.0223	0.216	0.496***	-0.280
	(-0.00)	(4.44)	(-1.82)	(1.77)	(4.56)	(0.08)	(0.77)	(4.62)	(-1.15)
HML30	0.0772	-0.132*	0.209*	-0.006	-0.144**	0.144	-0.009	-0.157**	0.148
	(0.77)	(-2.35)	(2.05)	(-0.01)	(-2.61)	(1.47)	(-0.10)	(-2.86)	(1.74)
Constant	0.154	0.0397	0.114	-0.468	0.0474	-0.515	-0.0332	0.0509	-0.0841
	(0.36)	(0.20)	(0.26)	(-1.21)	(0.24)	(-1.35)	(-0.09)	(0.26)	(-0.24)

To separate the price effects of private placements and convertible debt, I run another two regressions for the issuers and non-issuers of private placements and convertible debt, and I report the results in the tables below.

Table 16: Alphas and factor loadings for issuers and non-issuers of private placements and convertible debt under the CH3 model during the prior 1, 3 or 5 years.

2007--2018 Private Placements for top 70% of stocks

	prior 1 year			prior 3 years			prior 5 years		
	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference
MKT	1.035***	1.00***	0.035	1.092***	0.989***	0.103**	1.107***	0.982***	0.125***
	(31.62)	(96.02)	(0.92)	(44.26)	(80.18)	(3.23)	(45.85)	(74.41)	(3.93)
SMB	0.197*	-0.0597*	0.256**	0.174***	-0.093**	0.267***	0.133**	-0.095**	0.228***
	(2.55)	(-2.28)	(2.96)	(3.47)	(-3.05)	(4.16)	(2.72)	(-3.02)	(3.64)
VMG	-0.113	0.123***	-0.236*	-0.073	0.157***	-0.230*	-0.056	0.171***	-0.227*
	(-1.15)	(3.86)	(-2.05)	(-0.91)	(4.00)	(-2.18)	(-0.71)	(4.09)	(-2.16)
Constant	-0.300	-0.014	-0.286	-0.124	0.028	-0.153	-0.08	0.034	-0.113
	(-1.28)	(-0.23)	(-1.05)	(-0.80)	(0.37)	(-0.73)	(-0.52)	(0.40)	(-0.54)

2000--2018 Convertible debt for top 70% of stocks

	prior 1 year			prior 3 years			prior 5 years		
	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference
MKT	1.060***	0.990***	1.033***	0.992***	1.006***	0.997***	0.0699	0.0409	0.009
	(15.17)	(115.53)	(18.67)	(102.36)	(26.27)	(89.78)	(0.98)	(0.69)	(0.21)
SMB	-0.048	-0.011	-0.433**	0.0108	-0.326***	0.0237	-0.0374	-0.444**	-0.350***
	(-0.30)	(-0.51)	(-3.28)	(0.43)	(-3.78)	(0.98)	(-0.23)	(-3.10)	(-3.70)
VMG	0.192	0.062*	0.227	0.065*	0.291**	0.0384	0.130	0.162	0.253*
	(1.10)	(2.55)	(1.46)	(2.30)	(2.85)	(1.28)	(0.72)	(0.96)	(2.28)
Constant	-0.143	0.005	0.436	-0.0344	0.189	0.0174	-0.148	0.470	0.172
	(-0.30)	(0.07)	(1.02)	(-0.41)	(0.61)	(0.19)	(-0.31)	(1.02)	(0.50)

Similarly, I run regressions for the issuers and non-issuers of debt, categorized by whether the firm issued debt during the prior 1, 3 or 5 years. I divide all the stocks into bottom 30% and top 70%. The results are reported in the tables below.

Table 17: Alphas and factor loadings for issuers and non-issuers of total debt under CH3 and CHS4 models, categorized by whether the firm issued equity during the prior 1, 3 or 5 years (2000 –2018).

2000--2018 total Debt for top 70% of stocks

	prior 1 year			prior 3 years			prior 5 years		
	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference
MKT	1.031***	0.997***	0.0334	0.990***	1.010***	-0.0201	0.986***	1.023***	-0.0374
	(36.01)	(61.98)	(0.84)	(62.60)	(49.76)	(-0.64)	(71.80)	(54.66)	(-1.38)
SMB	-0.215***	0.123***	-0.338***	-0.230***	0.217***	-0.448***	-0.216***	0.273***	-0.489***
	(-3.45)	(3.35)	(-3.98)	(-5.02)	(4.68)	(-5.69)	(-5.30)	(6.31)	(-6.94)
VMG	0.332***	-0.0171	0.349***	0.306***	-0.103*	0.409***	0.301***	-0.114*	0.415***
	(4.70)	(-0.44)	(3.66)	(5.77)	(-2.01)	(4.63)	(6.29)	(-2.22)	(4.89)
Constant	0.0948	0.0339	0.0609	0.122	0.0918	0.0302	0.0838	0.0589	0.0249
	(0.41)	(0.31)	(0.21)	(0.90)	(0.75)	(0.15)	(0.72)	(0.51)	(0.15)

2000--2018 total Debt for bottom 30% of stocks

	prior 1 year			prior 3 years			prior 5 years		
	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference	Issuer	Non-issuer	Difference
FFMKT	0.731***	0.988***	-0.257**	0.882***	1.037***	-0.156	0.918***	0.989***	-0.0708
	(8.02)	(32.39)	(-2.72)	(10.09)	(32.53)	(-1.83)	(10.74)	(32.54)	(-0.85)
CHSMB	1.024***	1.270***	-0.246	1.040***	1.334***	-0.294	1.051***	1.279***	-0.227
	(6.19)	(24.06)	(-1.38)	(6.38)	(25.69)	(-1.74)	(6.43)	(24.68)	(-1.32)
SMB30	-0.676	0.486***	-1.163*	-0.557	0.502***	-1.059*	-0.561	0.489***	-1.051*
	(-1.68)	(4.45)	(-2.48)	(-1.36)	(4.43)	(-2.27)	(-1.44)	(4.54)	(-2.34)
HML30	-0.117	-0.134*	0.0169	-0.0724	-0.171**	0.0987	-0.0303	-0.152**	0.121
	(-0.95)	(-2.39)	(0.12)	(-0.61)	(-2.97)	(0.77)	(-0.26)	(-2.74)	(0.95)
Constant	0.415	0.0520	0.363	0.630	0.0459	0.584	0.290	0.0661	0.224
	(0.76)	(0.26)	(0.61)	(1.12)	(0.23)	(0.97)	(0.58)	(0.34)	(0.41)

The regression results show no abnormal return after either SEO or debt issuance in the long term, suggesting that equity prices reflect the available information and the market efficiently captures all the news around the announcement date.

6.3 Comparing the abilities of different models to explain anomalies in China

I use the CHS4 model to explore nine well-documented anomalies in China. The anomalies are listed below by category:

1. Size. Following the literature, size is measured as the natural logarithm of a stock's market capitalization.
1. Value. Earnings-price ratio (*E/P*). Following Liu, Stambaugh and Yuan (2018), value is the ratio of earnings to the product of last-month end's close price and total shares. Earnings equal the most recently reported net profit, excluding nonrecurrent gains/losses.
2. Volatility. MAX is used. Following Bali, Cakici and Whitelaw (2011), MAX is measured as the maximum daily stock return over the previous month.
3. Investment. Following Fama and French (2015), it is measured as a firm's total asset growth rate, updated annually.
4. Profitability. A common measurement is firm-level return-on-equity (ROE). Fama and French (2015) measure it as the ratio of a firm's net income to book equity. Liu, Stambaugh and Yuan (2018) measure it as net profit excluding non-recurring gains and losses divided by total shareholders' equity excluding minority interests. The CSRC has required public firms to disclose non-recurring gains and losses since 1999. Considering the accounting standards in China, I choose to follow the method of Liu, Stambaugh and Yuan (2018) to ensure the relevance of income information. ROE is updated quarterly.
5. Illiquidity. Following Amihud (2002), a stock's illiquidity measure for day t is calculated as $illiq_t = |return_t| / volume_t$

$|return_t|$ is the stock's absolute return, and $volume_t$ is the stock's trading volume.

It is measured as a stock's average daily illiquidity over the past month.

6. Reversal. Following Jegadeesh (1990) and Lehmann (1990), I measure short-term reversal as stock's cumulative return over the past month.
7. Turnover. Following Liu, Stambaugh and Yuan (2018), it is measured as 12-month turnover as the average daily share turnover over the past 12 months. A firm's daily turnover is calculated as share trading volume divided by total shares outstanding.
8. Momentum. Following Jegadeesh and Titman (1993), it is defined as the cumulative stock return over months t − 12 to t – 1.

For each of nine anomalies, the table below reports the monthly long-short return spread's CH3, FF3, FF5 and CHS4 alphas and corresponding t-statistics under unconditional sorts. The regressions estimated for different models are listed below:

CH3: $R_t = \alpha + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{CHSMB}\, CHSMB_t + \beta_{CHVMG}\, CHVMG_t + \in_t$

FF3: $R_t = \alpha + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{FFSMB}\, FFSMB_t + \beta_{FFHML}\, FFHML_t + \in_t$

FF5: $R_t = \alpha + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{FFSMB}\, FFSMB_t + \beta_{FFHML}\, FFHML_t +$
$\beta_{RMW}\, RMW_t + \beta_{CMA}\, CMA_t + \in_t$

CHS4: $R_t = \alpha + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{FFSMB}\, FFSMB_t + \beta_{SMB30}\, SMB30_t +$
$\beta_{HML30}\, HML30_t + \in_t$

R_t is the anomaly's long-short return spread in month t, $FFSMB_t$ is FF's size factor, $CHSMB_t$ is CH3's size factor, $FFHML_t$ *is* the BM-based value factor, $CHVMG_t$ is the EP-based value factor, and $SMB30_t$ *and* $HML30_t$ are the size and BM-based value factors constructed from the smallest 30% stocks, as defined before. The profitability factor (RMW_t) is the return spread between firms with robust operating profitability and weak operating profitability, the investment factor (CMA_t) is the return spread between firms that invest conservatively and aggressively, constructed using 2x3 sorts following Fama and French (2015). When constructing these anomaly portfolios, the sort uses the information in a firm's quarterly, semi-annual or annual financial report that gives the most recent month-end data released in the WIND database.

For almost all the anomalies, while FF3, FF5, and CH3 models usually leave both economically and statistically significant alphas across five quantiles and fail to reveal the pattern of alphas, the CHS4 model can successfully demonstrate the patterns of alphas across different quantiles for most of the anomaly categories.

It is not surprising to see that CHS4 explains the size anomaly. Furthermore, the CHS4 model that uses the B/M ratio to construct the value factor can successfully explain the alternative EP-based value factor. The long-short alpha on MAX is economically significantly negative. This is particularly noteworthy with a 10% cap on daily price movements by the Chinese government. Similar to the results by Bali, Cakici, and Whitelaw (2011) for the US stocks, the pattern is hump-shaped instead of strictly monotonic. This suggests that investors in China's small stocks market like lottery-type stocks and pay up for the chance to win a large amount of return. The table shows a monotonic illiquidity effect across the five quintiles that is statistically significant. As in the US, Chinese investors in small stocks require a large compensation for holding illiquid stocks. This could be due to transaction costs for trading those small stocks. The reversal effect is substantial and statistically significant. Similar to the biggest 70% stocks in China, small stocks that are past losers significantly outperform past winners. The presence of a strong turnover effect shows that small stocks' returns can be influenced by investor sentiment, just like big stocks. The return spread between firms with high and low profitability is economically large and significant. Perhaps more profitable firms are riskier because the aggregate demand shocks can influence their return to a larger degree. But we should treat this result with care because there exist large gaps between the ROE level when we use different measurements, for example, whether to exclude non-recurring profits and losses from the net income and whether to exclude minority interests from the shareholder's equity. The investment effect has a different sign compared to the US market. It is also interesting to see that the CHS4 model suggests the potential existence of a momentum effect that the other models fail to find.

Table 18: Anomaly alphas and factor loadings under CH3, FF3, FF5, and CHS4 models.

For each of nine anomalies, the table reports the monthly long-short return spread's CH3, FF3, FF5, and CHS4 alphas and factor loadings. For each anomaly, the regressions estimated are:

CH3: $R_t = \alpha_i + \beta_{CHMKT} (R_{CHMKT,t} - R_{f,t}) + \beta_{CHSMB} \, CHSMB_t + \beta_{CHVMG} \, CHVMG_t + \in_{i,t}$

FF3: $R_t = \alpha_i + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{FFSMB} \, FFSMB_t + \beta_{FFHML} \, FFHML_t + \in_{i,t}$

FF5: $R_t = \alpha_i + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{FFSMB} \, FFSMB_t + \beta_{FFHML} \, FFHML_t +$
$\beta_{FFRMW} \, FFRMW_t + \beta_{FFVMG} \, FFVMG_t + \in_{i,t}$

CHS4: $R_t = \alpha_i + \beta_{FFMKT} (R_{FFMKT,t} - R_{f,t}) + \beta_{CHSMB} \, CHSMB_t + \beta_{SMB30} \, SMB30_t +$
$\beta_{HML30} \, HML30_t + \in_{i,t}$

where R_t is the anomaly's long-short return spread in month t. The sample period is January 2000 through December 2018. All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Category	Anomaly	Model	(1) Quantile 1	(2) 2	(3) 3	(4) 4	(5) Quantile 5	(5) - (1) Difference
Volatility	MAX	CH3	0.804***	1.112***	1.156***	0.886***	0.295	-0.509*
			(4.23)	(7.24)	(7.94)	(5.28)	(1.94)	(-2.28)
		FF3	0.726***	0.814***	0.719***	0.360	-0.337	-1.063***
			(3.59)	(4.71)	(4.17)	(1.79)	(-1.78)	(-5.32)
		FF5	0.793***	0.875***	0.794***	0.404*	-0.284	-1.077***
			(3.98)	(5.19)	(4.81)	(2.04)	(-1.53)	(-5.26)
		CHS4	0.142	0.254	0.306	0.164	-0.620**	-0.763***
			(0.60)	(1.24)	(1.52)	(0.62)	(-2.85)	(-3.72)
Reversal	1-Month return	CH3	1.455***	1.318***	0.955***	0.700***	-0.118	-1.573***
			(8.09)	(8.11)	(6.46)	(4.83)	(-0.67)	(-6.52)
		FF3	1.025***	1.069***	0.646***	0.227	-0.668**	-1.693***
			(5.02)	(5.68)	(3.51)	(1.25)	(-3.04)	(-6.32)
		FF5	1.113***	1.128***	0.700***	0.283	-0.620**	-1.733***
			(5.48)	(6.03)	(3.86)	(1.60)	(-3.01)	(-6.54)
		CHS4	0.404	0.483	0.182	-0.121	-0.707**	-1.111***
			(1.45)	(1.96)	(0.80)	(-0.58)	(-3.04)	(-3.44)
Illiquidity	Amihud illiq.	CH3	0.342*	0.807***	1.135***	1.050***	0.977***	0.635*
			(2.01)	(5.50)	(7.32)	(6.88)	(4.60)	(2.26)
		FF3	-0.120	0.296	0.711***	0.815***	0.667***	0.787**
			(-0.54)	(1.57)	(3.94)	(4.64)	(3.60)	(3.17)
		FF5	0.0312	0.410*	0.764***	0.812***	0.624***	0.593*
			(0.15)	(2.43)	(4.52)	(4.92)	(3.52)	(2.51)
		CHS4	-0.634*	-0.154	0.313	0.391	0.447*	1.081***
			(-2.37)	(-0.64)	(1.43)	(1.92)	(2.11)	(3.99)
Momentum	MOM	CH3	0.596*	0.662***	0.620***	0.380*	0.0346	-0.562*
			(2.51)	(3.45)	(3.57)	(2.09)	(0.18)	(-2.08)
		FF3	0.966***	0.952***	1.008***	0.799***	0.580**	-0.386
			(4.15)	(6.01)	(7.81)	(5.63)	(2.95)	(-1.22)
		FF5	0.679**	0.760***	0.699***	0.421*	0.0371	-0.642*
			(2.97)	(4.03)	(4.20)	(2.41)	(0.20)	(-2.47)
		CHS4	-0.277	0.118	0.028	0.167	0.178	0.455
			(-1.03)	(0.51)	(0.14)	(0.72)	(0.75)	(1.66)
Size	Market Cap	CH3	1.982***	1.305***	1.049***	0.709***	0.679***	1.303***
			(8.91)	(7.73)	(8.12)	(5.38)	(5.76)	(6.23)
		FF3	1.420***	0.849***	0.733***	0.327*	0.329	1.091***
			(6.33)	(4.41)	(4.36)	(2.04)	(1.95)	(5.92)
		FF5	1.508***	0.923***	0.768***	0.394*	0.380*	1.128***
			(6.95)	(5.00)	(4.61)	(2.50)	(2.34)	(6.00)
		CHS4	0.442	0.179	0.259	0.075	0.312	0.130
			(1.86)	(0.85)	(1.32)	(0.38)	(1.39)	(1.17)
Investment	Asset Growth	CH3	0.862***	0.908***	0.905***	0.937***	0.656***	-0.206
			(4.13)	(6.07)	(7.37)	(7.27)	(3.75)	(-0.79)
		FF3	0.211	0.352*	0.507**	0.594***	0.546**	0.334
			(0.93)	(1.99)	(3.27)	(3.98)	(3.07)	(1.42)
		FF5	0.390	0.478**	0.558***	0.628***	0.477**	0.087
			(1.77)	(2.73)	(3.41)	(4.10)	(2.70)	(0.43)
		CHS4	-0.450	-0.174	0.116	0.322	0.304	0.754**
			(-1.76)	(-0.77)	(0.56)	(1.48)	(1.34)	(2.81)
Profitability	ROE	CH3	-0.257	0.142	0.184	0.653***	1.539***	1.796***
			(-1.00)	(0.84)	(1.12)	(4.05)	(8.39)	(6.10)
		FF3	0.407	0.659***	0.724***	0.886***	1.562***	1.155***
			(1.71)	(4.69)	(5.10)	(6.40)	(9.32)	(3.79)
		FF5	-0.0601	0.252	0.229	0.610***	1.523***	1.583***
			(-0.25)	(1.42)	(1.35)	(3.70)	(8.12)	(5.89)
		CHS4	-1.044***	-0.249	-0.128	0.382	1.364***	2.409***
			(-3.70)	(-1.15)	(-0.54)	(1.73)	(6.32)	(7.37)
Value	EP	CH3	0.930***	0.652***	0.861***	1.065***	0.798***	-0.132
			(6.69)	(4.45)	(5.57)	(5.09)	(4.98)	(-0.86)
		FF3	0.388*	0.207	0.735***	0.702**	0.230	-0.158
			(2.16)	(1.40)	(4.61)	(3.10)	(1.26)	(-1.15)
		FF5	0.472*	0.197	0.685***	0.835***	0.397*	-0.075
			(2.57)	(1.23)	(4.14)	(3.84)	(2.16)	(-0.60)
		CHS4	0.052	0.033	0.403	0.189	-0.07	-0.122
			(0.21)	(0.13)	(1.81)	(0.62)	(-0.24)	(-0.65)
Turnover	12-Month turn.	CH3	0.933***	0.945***	0.915***	0.770***	0.669**	-0.264
			(5.87)	(6.51)	(6.66)	(5.35)	(2.97)	(-0.96)
		FF3	0.838***	0.640***	0.430*	0.228	0.105	-0.733**
			(4.45)	(3.81)	(2.52)	(1.33)	(0.54)	(-3.20)
		FF5	0.927***	0.721***	0.519**	0.304	0.0721	-0.855***
			(4.93)	(4.31)	(3.13)	(1.89)	(0.37)	(-3.77)
		CHS4	0.391	0.270	-0.0202	-0.287	-0.145	-0.537*
			(1.73)	(1.18)	(-0.09)	(-1.28)	(-0.60)	(-2.13)

7. Principal Component Analysis on the bottom 30% of stocks

One natural question to ask is whether the size factors SMB30 and CHSMB, also reveal a pronounced seasonality pattern like the size factor in the US market. Figure 5 shows that the stock returns are higher in February and March, but Figure 6 demonstrates that, different from the US, the seasonality of the size effect in China isn't robust and prevails throughout the year. In other words, the seasonality pattern for the size effect in China is not as pronounced as it is in the US.

It is clear that the small and big stocks in China are different in their pricing, and there exists a strong small firm effect in China, but the economic reason for this phenomenon remains a puzzle. There is a debate on whether shell value is the primary reason for the difference in pricing between big and small stocks in China. On the one hand, Liu, Stambaugh and Yuan (2018) and Lee, Qu and Shen (2017) argue that the price of small stocks in China is significantly contaminated by the shell value associated with reverse merger prospects. On the other hand, Carpenter, Lu and Whitelaw (2020) emphasize the small number of reverse merger events in China and question the economic importance of shell value. The shell value story proposed by Liu, Stambaugh and Yuan (2018) is hard to verify because it is difficult to measure a firm's shell value. The existing literature doesn't suggest a proper approach to measuring it. But if this shell value story holds, it is expected that the bottom of 30% and the top of 70% stocks will respond to different kinds of shocks in the market and tend to move differently. However, Figure 3 illustrates the average monthly pairwise correlation between the bottom 30% stocks and between the bottom 30% and the top 70% stocks, which suggests small stocks in China are not fundamentally different from the big ones. Figure 4 shows the time-series plot of the average Herfindahl-Hirschman Index (HHI) for small stocks and big stocks. HHI is a measure of market concentration. In general, the HHI for big and small stocks change concurrently and follows an overall decreasing trend that reflects the privatization reform on the firm level. Following the split-share reform in 2005, there was a sharp decrease in overall market concentration for big and small stocks in China. Big and small stocks seem to respond to similar shocks consistently. Thus, the reason for the difference in pricing for China's big and small stocks seems to be more complicated than being a result of the impact from the prospects of reverse mergers.

To approach this question, I apply principal component analysis (PCA) to nine return series from the three by three sorts on size and value of the bottom 30% of stocks. From the table below, we can see a dominant first factor, which explains around 97% variance of the data. The second through fourth factors only explains 1%, 0.6% and 0.4%, respectively. The result is a bit surprising given the large variation in alphas across portfolios after we run a regression of these portfolios on MKT30. If MKT30 explains so much of the variation in returns, why doesn't it explain average returns?

Table 19: Eigenvalues and explained variance from a PCA on the 9-return series from the 3x3 sort.

Component	Eigenvalue	Difference	Proportion	Cumulative
Componet1	8.73614	8.64545	0.9707	0.9707
Componet2	0.0906866	0.0360716	0.0101	0.9808
Componet3	0.0546151	0.0185425	0.0061	0.9868
Componet4	0.0360726	0.0127338	0.004	0.9908
Componet5	0.0233387	0.00444596	0.0026	0.9934
Componet6	0.0188928	0.00300272	0.0021	0.9955
Componet7	0.0158901	0.00283588	0.0018	0.9973
Componet8	0.0130542	0.00174473	0.0015	0.9987
Componet9	0.0113095	.	0.0013	1

The table below reports the eigenvectors of the return covariance matrix. It seems that the first factor is close to a market factor. If we look at the 3x3 factor loadings of the second and third components (factors):

Panel A: Factor 1			
	Growth	Neutral	Value
Small	0.329	0.334	0.335
Median	0.334	0.335	0.334
Large	0.333	0.334	0.332
Panel B: Factor 2			
	Growth	Neutral	Value
Small	0.603	0.28	-0.003
Median	0.191	-0.117	-0.318
Large	0.19	-0.278	-0.542
Panel C: Factor 3			
	Growth	Neutral	Value
Small	0.411	0.077	0.39
Median	-0.292	-0.138	0.268
Large	-0.622	-0.275	0.188

It seems that the combination picks up the size and value factors. Factor 2 is a combination of small minus big and growth minus value. Factor 3 looks like small minus large plus value minus growth, except for the smallest stocks.

I also apply PCA to the regression residuals from the nine portfolios under CHMKT and CH2 models. The idea is to figure out the dimensionality of the remaining returns associated with the CAPM and CH2 models. Table 20 reports the eigenvalues and explained variance from a PCA of the regression residual under CHMKT. We can see that after pulling out CHMKT, the first factor is still dominant and explains 92% of the variance. MKT30 and CHMKT are positively correlated, but since the R-squared of regression of MKT30 on CHMKT is less than

70%, the results make sense, and the first factor is close to a market factor from the bottom 30% of stocks. The second through fourth factors explains 2.5%, 1.6%, and 1.1% of the variance. There is no second dominant factor.

Table 20: Eigenvalues and explained variance from a PCA for the regression residual under CHMKT.

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	8.28899	8.06516	0.9210	0.9210
Comp2	.223832	.0750746	0.0249	0.9459
Comp3	.148757	.0527269	0.0165	0.9624
Comp4	.09603	.0285896	0.0107	0.9731
Comp5	.0674405	.0118468	0.0075	0.9806
Comp6	.0555937	.00909628	0.0062	0.9867
Comp7	.0464974	.00482181	0.0052	0.9919
Comp8	.0416756	.0104949	0.0046	0.9965
Comp9	.0311807	.	0.0035	1.0000

The table below reports the eigenvectors of the return covariance matrix. The first factor is again something close to an equally weighted portfolio of the nine portfolios. The second factor looks like some combination of size and value factors. Factor 3 has a strong size component, but the value component is weaker.

Panel D: Factor 1			
	Growth	Neutral	Value
Small	0.325	0.339	0.337
Median	0.335	0.338	0.335
Large	0.332	0.335	0.327
Panel E: Factor 2			
	Growth	Neutral	Value
Small	0.472	0.221	-0.104
Median	0.265	-0.054	-0.377
Large	0.348	-0.175	-0.592
Panel F: Factor 3			
	Growth	Neutral	Value
Small	0.58	0.117	0.329
Median	-0.173	-0.189	0.184
Large	-0.514	-0.412	0.09

Table 21 reports the eigenvalues and explained variance from a PCA of the regression residual under the CH2 model. We can see that after removing both the CH3 market and size factors, the residual factors move in terms of explained variation. The first through fourth factors explain 56.9%, 14.5%, 7.1% and 6.1% of the variance, respectively.

Table 21: Eigenvalues and explained variance from a PCA for the regression residual under CH2

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	5.12258	3.8175	0.5692	0.5692
Comp2	1.30508	.666737	0.1450	0.7142
Comp3	.638341	.0868184	0.0709	0.7851
Comp4	.551523	.149877	0.0613	0.8464
Comp5	.401646	.096142	0.0446	0.8910
Comp6	.305504	.0276602	0.0339	0.9250
Comp7	.277844	.0494008	0.0309	0.9558
Comp8	.228443	.059397	0.0254	0.9812
Comp9	.169046	.	0.0188	1.0000

The table below reports the eigenvectors of the return covariance matrix. The first factor is again something close to an equal-weighted portfolio of the nine portfolios. It seems that the second factor picks up the combination of size and value factors. Factor 3 has a strong size component, but the value component is weaker.

Panel G: Factor 1			
	Growth	Neutral	Value
Small	0.364	0.374	0.384
Median	0.352	0.35	0.318
Large	0.29	0.305	0.237
Panel H: Factor 2			
	Growth	Neutral	Value
Small	-0.134	-0.156	0.187
Median	-0.254	-0.032	0.445
Large	-0.534	0.014	0.611
Panel I: Factor 3			
	Growth	Neutral	Value
Small	-0.369	-0.231	-0.090
Median	-0.116	0.079	-0.164
Large	0.147	0.854	0.073

8. Summary and Conclusions

While the CH3 model is superior to the FF3 and FF5 models in explaining equity returns in the Chinese stock market, all three of these models fail to explain the returns of the smallest 30% of stocks. Unlike small stocks in the US, small stocks in China require special size and value factors constructed from the same small stocks to explain the cross section of their returns. The new CHS4 model fully captures the return variation among the smallest 30% of stocks in China. At the same time, both the CHS4 and CH3 models explain the SOE, debt issuance, and

SEO anomalies. CHS4 can better reveal the pattern of the well-documented anomalies for small stocks in China compared to CH3, FF3, and FF5 models.

However, the CHS4 model cannot fully explain most of these anomalies. At the same time, it seems difficult to develop a parsimonious model for pricing the bottom 30% of China's stocks. The MKT30 seems to explain most of the variation but fails to explain the average return. The same situation happens when we apply the CH3 model to those small stocks. We seem to get high R-squared and also high alphas. As a result, we have very high information ratios, suggesting that the price of residual risk seems too high. In other words, investors seem to earn too high a return for taking on this residual risk. Since this risk is idiosyncratic relative to the models and potentially partially diversifiable, it is perhaps a bit surprising to see such high information ratios. The results seem to be more consistent with mispricing than a priced risk story. If such mispricing among the bottom 30% of stocks holds, I expect that very high returns of small stocks will vanish in the long term as the market becomes mature. Perhaps the bottom 30% of stocks may be too noisy a market right now, making statistical methods alone difficult to justify the choice of the best set of factors for pricing these stocks.

.

References

Amihud, Yakov, 2002, "Illiquidity and stock returns: cross-section and time-series effects," *Journal of Financial Markets* 5, Issue 1, 31-56.

Baker, M., Wurgler, J., 2006. "Investor sentiment and the cross-section of stock returns." *Journal of Finance* 61, 1645–1680.

Ball, R., 1992. "The earnings-price anomaly." *Journal of Accounting and Economics* 15, 319–345.

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.

Carpenter, Jennifer, Fangzhou Lu, and Robert Whitelaw, 2020, "The Real Value of China's Stock Market." *Journal of Financial Economics* 139, 679-696.

Carpenter, Jennifer, and Robert Whitelaw, 2017, "The Development of China's Stock Market and Stakes for the Global Economy", *Annual Review of Financial Economics* 9, Issue 1,1-22.

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, 2015. “A five-factor asset pricing model”, *Journal of Financial Economics* 116, 1–22.

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

Jegadeesh, Narasimhan, 1990, “Evidence of predictable behavior of security returns”, *Journal of Finance*, vol. 45, 881–898.

Lee, CharlesQu, Yuanyu and Shen, Tao, 2017, “Reverse Mergers, Shell Value, and Regulation Risk in Chinese Equity Markets”,*Stanford University Graduate School of Business Research Paper* 17-63, Available at SSRN: https://ssrn.com/abstract=3038446 or http://dx.doi.org/10.2139/ssrn.3038446.

Lehmann, Bruce N, 1990, “Fads, martingales, and market efficiency”, *Quarterly Journal of Economics* 105, 1–28.

Liu, Jianan, Robert F. Stambaugh and Yu Yuan, 2019, “Size and Value in China”, *Journal of Financial Economics* 134, Issue 1, 48-69.

Liu J, S. Akbar, S.Z.A. Shah, D. Zhang and D. Pang, 2016, “Market Reaction to Seasoned Offerings in China”. *Journal of Business Finance and Accounting* 43, 597-653.

Loughran, Tim, and Jay Ritter, 1995, “The new issues puzzle”, *Journal of Finance* 50, 23-52.

Newey, Whitney and Kenneth D. West, 1987, “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”, *Econometrica* 55, 703–708.

Spiess, D. Katherine and Affleck-Graves, 1999, "The long-run performance of stock returns following debt offerings", *Journal of Financial Economics* 54, 45-73.

White H, 1980, “A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity”, *Econometrica* 48, 817-838.

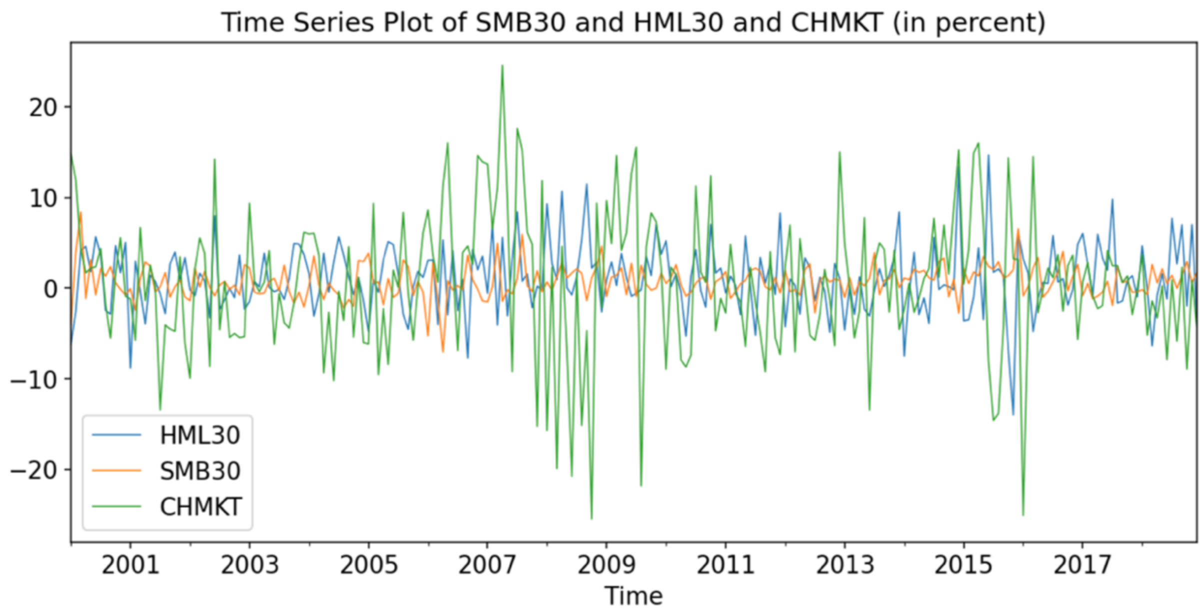
Appendix

Figure 1:

First: **Number of Listed Firms in China, 2000-2018**

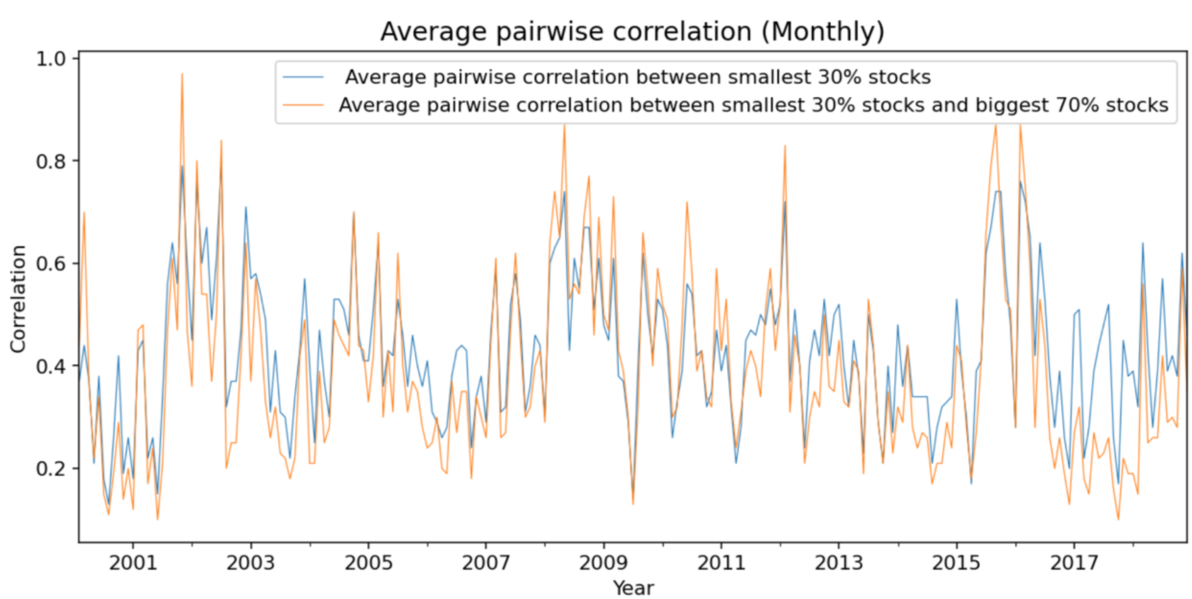
Second: **Market Capitalization of the Smallest 30% Stocks Over the Total Stocks, Monthly, 2000-2018**

Figure 2:



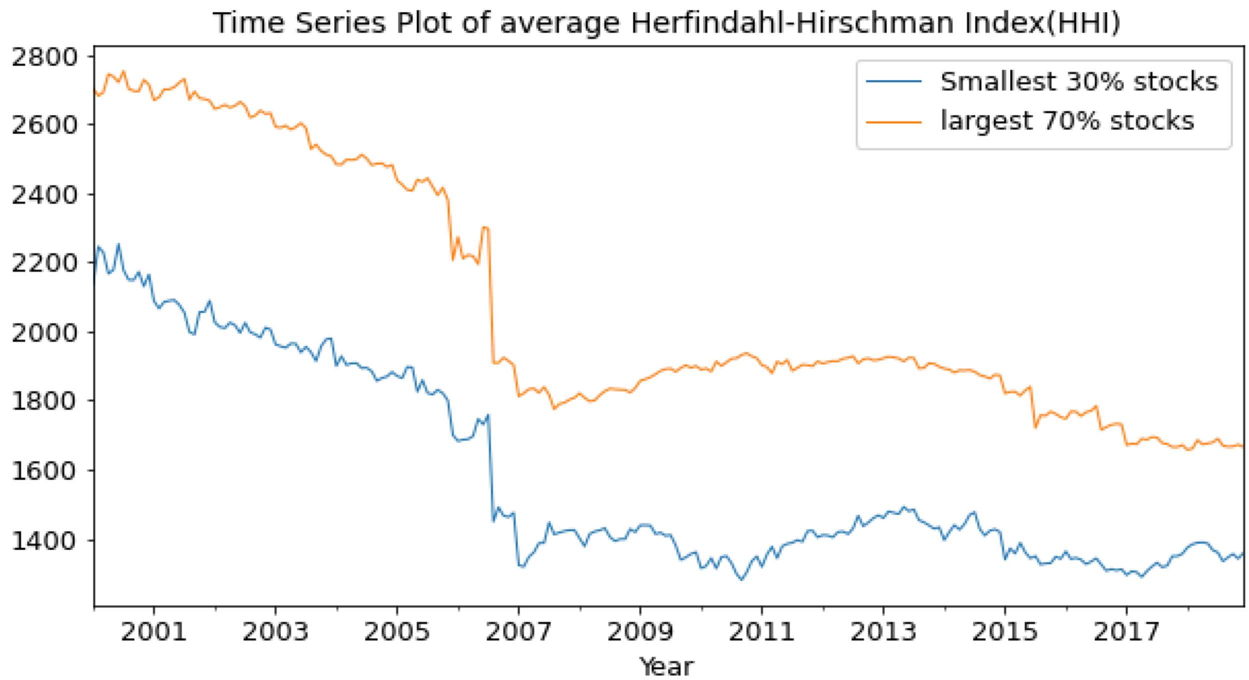
Time Series Plot of SMB30, HML30 and CHMKT, Monthly, 2000-2018

Figure 3:



Average Pairwise Correlation Between Smallest 30% Stocks, and Between Smallest 30% and Biggest 70% Stocks, 2000-2018

Figure 4:



Time series Plot of Average Herfindahl-Hirschman Index (HHI) for Smallest 30% stocks and Largest 70% Stocks, Monthly, 2000-2018

Note that the Herfindahl-Hirschman Index (HHI) is a measurement of ownership concentration. It calculated as the sum of the top 10 firm's market share of the stock:

HHI = $S_1^2 + S_2^2 + S_3^2 + \cdots + S_{10}^2$

where S_i is the percentage holdings by investor i.

Figure 5:

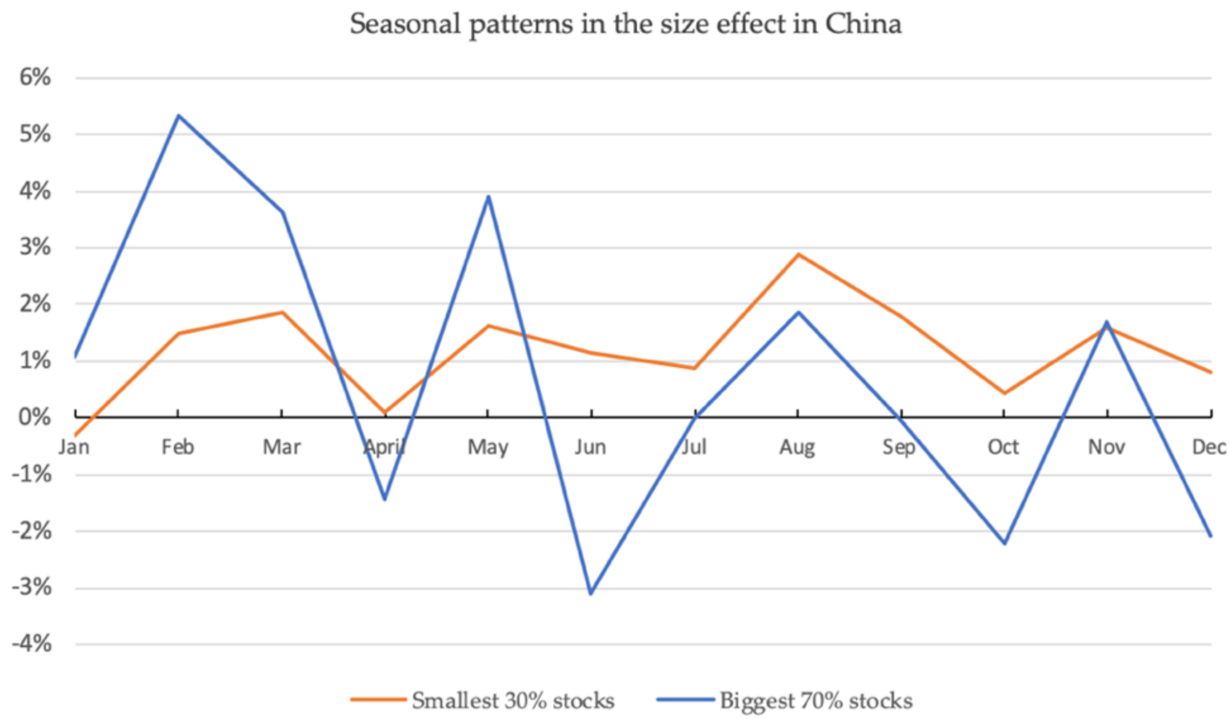
Biggest 70% stocks					
	Low	Book-to-market			High
Panel A: All Month					
small	0.64	0.97	1.31	1.49	1.60
	0.72	0.93	1.12	1.26	1.26
size	0.48	0.80	0.77	0.90	1.21
	0.34	0.69	0.69	0.92	1.20
big	0.34	0.07	0.45	0.54	1.27
Panel B: February only					
small	5.70	5.43	6.07	6.59	7.66
	5.11	5.63	5.95	7.07	6.18
size	5.08	4.71	5.64	6.03	6.00
	3.43	4.85	4.70	5.39	5.42
big	3.18	2.14	1.98	0.99	2.34
Panel C: Non-February only					
small	0.19	0.56	0.88	1.03	1.05
	0.32	0.50	0.68	0.73	0.81
size	0.06	0.45	0.33	0.43	0.78
	0.06	0.31	0.33	0.51	0.81
big	0.08	-0.12	0.31	0.50	1.17
Panel D: March only					
small	3.93	3.65	3.94	4.37	4.21
	3.68	3.09	3.57	3.12	3.77
size	2.42	2.73	3.32	2.93	3.13
	2.10	1.48	1.82	2.53	2.66
big	0.90	0.74	0.53	1.17	0.94
Panel E: Non-March only					
small	0.35	0.72	1.08	1.23	1.36
	0.45	0.73	0.90	1.09	1.03
size	0.31	0.63	0.54	0.71	1.04
	0.18	0.62	0.59	0.77	1.07
big	0.28	0.01	0.44	0.48	1.30

Smallest 30% stocks			
	Growth	Neutral	Value
Panel A: All Month			
Small	2.14	2.42	2.54
Medium	1.36	1.93	2.07
Big	1.03	1.44	1.70
Panel B: February only			
Small	7.57	7.76	7.78
Medium	7.07	6.89	7.75
Big	5.88	6.54	7.08
Panel C: Non-February only			
Small	1.65	1.94	2.07
Medium	0.84	1.48	1.55
Big	0.59	0.97	1.21
Panel D: March only			
Small	7.25	6.47	5.88
Medium	5.26	5.58	5.37
Big	4.97	4.38	4.86
Panel E: Non-March only			
Small	1.68	2.06	2.24
Medium	1.00	1.60	1.77
Big	0.67	1.17	1.41

Left: **Average Stock Returns for the Biggest 70% Stocks After a 5x5 Sort on Size and Value, Monthly, 2000-2018**

Right: **Average Stock Returns for the Smallest 30% Stocks After a 3x3 Sort on Size and Value, Monthly, 2000-2018**

Figure 6:



Seasonal Patterns in the Size Effect in China's Stock Returns, 2000-2018

Note that the smallest 30% stocks are cut into a total of 5 quantiles, the largest 70% stocks are cut into a total of 10 deciles based on market capitalization. This figure depicts the average value-weighted return differential between the smallest and largest quintiles of all A-shares in each month.