Replication: Size and Value in China

Wei Ling

School of Finance, Renmin University of China

June 28, 2023

1 Introduction

Liu et al. (2019) excludes the smallest 30% of firms and construct size and value factors in China, which explains most Chinese anomalies. This paper guides the data filter procedures and anomalies construction in later literature which concerns on Chinese stock market. The original data used in Liu et al. (2019) comes from WIND database and terminal, which is not available in the library or difficult to download, so I use data from CSMAR instead.

2 CH3 and CH4 Factors

I first replicate CH3 and CH4 factors, whose official version can be downloaded on the personal website of Robert F. Stambaugh.

Table 1: Correlation between official CH3 factors and replicated CH3 factors
The suffix stands for official data from the personal website site of Robert F. Stambaugh while variables without suffix are replicated variables. Correlations between replicated variables and official variables are 1, 0.99 and 0.96 respectively.

		mktrf	smb	vmg	mktrf_ref	smb_ref	vmg_ref
MEAN		0.66	0.57	1.23	0.57	0.68	1.10
STD		7.43	4.61	3.81	7.57	4.36	3.76
N		264	248	248	259	264	264
CORR	mktrf	1.00	0.11	-0.25	1.00	0.14	-0.29
CORR	smb	0.11	1.00	-0.50	0.12	0.99	-0.56
CORR	vmg	-0.25	-0.50	1.00	-0.25	-0.50	0.96
CORR	mktrf_ref	1.00	0.12	-0.25	1.00	0.15	-0.28
CORR	smb_ref	0.14	0.99	-0.50	0.15	1.00	-0.54
CORR	vmg_ref	-0.29	-0.56	0.96	-0.28	-0.54	1.00

Table 2: Correlation between official CH4 factors and replicated CH4 factors
The suffix stands for official data from the personal website site of Robert F. Stambaugh while variables without suffix are replicated variables. Since mkt and vmg factors are the same in CH3 and CH4 factors, so this table ignores them. Correlations between replicated variables and official variables are 0.98 and 0.96 respectively.

		smb	pmo	smb_ref	pmo_ref
MEAN		0.23	0.83	0.51	0.80
STD		4.64	3.50	4.50	3.55
N		265.00	265.00	264.00	264.00
CORR	smb	1.00	0.11	0.98	0.08
CORR	pmo	0.11	1.00	0.11	0.96
CORR	smb_ref	0.98	0.11	1.00	0.11
CORR	pmo_ref	0.08	0.96	0.11	1.00

3 Main Results

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

	China			US		
Quantity	Smallest	Middle	Largest	Smallest	Middle	Largest
Panel A: $k=0$						
b	0.03	<0.01	0.01	0.16	<0.01	<0.01
	(2.88)	(1.25)	(1.95)	(4.74)	(0.51)	(0.63)
R^2	0.001	<0.001	0.001	0.003	<0.001	<0.001
Panel B: $k=3$						
b	0.21	<0.01	-0.18	0.12	0.01	<0.01
	(1.54)	(-6.54)	(-3.14)	(1.41)	(0.98)	(1.19)
R^2	<0.001	<0.001	<0.001	0.001	<0.001	<0.001

Table 3 attempts to replicate Table 1 in Liu et al. (2019) but fail to obtain the same perfect pattern.

Table 4 replicates Table 2 in Liu et al. (2019). In the original table, the slope coefficients of $logME$ is significant through all specifications under the significant level 5%, while they become insignificant in Table 4. EP^+ dominates other valuation ratios in specification (8) in the original table, but Table 4 shows $D(EP<0)$ dominates other valuation ratios instead in specification (8). Remind results are similar between Table 2 in Liu et al. (2019) and 4.

Table 5 replicates Table 4 in Liu et al. (2019) with the same pattern but slightly bigger average R-square.

Here I use official SMB and VMG variables to produce Table 6, and obtain similar results with respect to Table 5 in Liu et al. (2019). If I use replicated SMB and VMG variables, the t-statistics of FFHML alpha reaches -2.03 and other results remain similar.

Table 7 show close pattern with Table 6 in Liu et al. (2019).

Table 8 show close pattern with Table 7 in Liu et al. (2019).

Table 9 show close pattern with Table 8 in Liu et al. (2019).

Table 10 show close pattern with Table 9 in Liu et al. (2019).

Table 11 show close pattern with Table 10 in Liu et al. (2019).

4 Code

The replication code package contains 7 sas code file, 1 stata code file and 1 matlab code file.

The "Step1_filter_monthly_return_file.sas" file prepares and filters Chinese trading data and financial statements data from CSMAR. The "Step2_ch3_factors.sas" and "Step3_ch4_factors.sas" files construct CH-3 and CH-4 factors respectively. The "Table_1.sas", "Table_2.sas", "Table_4.sas" and "Table_6-10.sas" files prepares panel data for regression.

The "table.do" file run regressions to produce Table 3-11 in sequence.

The "grstest.m" file computes the unadjusted column in Table 10.

Table 4: Fama-MacBeth regressions of stock returns on beta, size and valuation ratios

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

Quantity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.012	0.052	0.035	0.035	0.036	0.067	0.048	0.037	0.061
	(1.44)	(1.29)	(0.88)	(0.94)	(0.99)	(1.81)	(1.25)	(0.67)	(1.73)
β	-0.001		-0.001	-0.001	-0.001	-0.000	0.003	0.003	-0.001
	(-0.50)		(-0.40)	(-0.67)	(-0.71)	(-0.14)	(0.81)	(0.86)	(-0.36)
$logME$		-0.002	-0.001	-0.001	-0.001	-0.002	-0.002	-0.001	-0.002
		(-1.00)	(-0.55)	(-0.42)	(-0.55)	(-1.62)	(-0.96)	(-0.47)	(-1.34)
$logBM$				0.004				-0.001	0.003
				(2.51)				(-0.48)	(1.75)
$logAM$					0.004			0.004	
					(2.53)			(1.47)	
EP^+						0.351		0.027	0.287
						(4.36)		(0.12)	(3.59)
$D(EP < 0)$						-0.006		-0.010	-0.006
						(-3.97)		(-3.98)	(-4.26)
CP^+							-0.010	-0.018	
							(-0.24)	(-0.09)	
$D(CP < 0)$							-0.001	-0.000	
							(-0.80)	(-0.03)	
R^2	0.015	0.022	0.036	0.053	0.054	0.045	0.073	0.111	0.062

Table 5: Average R-squares for individual stocks in China and the US.
The table compares the average R-squares in regressions of monthly individual stocks' returns on factors in China's and the US stock markets. Regressions are estimated for four models: one with just the excess market return (MKT); one with MKT plus the size factor; one with MKT plus the value factor; and the three-factor model with MKT plus the size and value factors. In China's stock market, we use our CH-3 model's market (MKT), size (SMB) and value (VMG) factors, while in the US market, we use FF-3's three factors: market, SMB and BM-based HML. For each stock, we run rolling-window regressions of each stock's monthly returns on factors over the past three years (36 months). We average the R-square across time for each stock and then compute the mean of these averages across all stocks. Panel A reports average R-squares across all individual stocks on China's main boards and the Growth Enterprise Market (GEM), including the smallest 30% of stocks. Panel B reports average R-squares of all but the smallest 30% of stocks. Panel C reports average R-squares of all common stocks from the NYSE, Amex, and Nasdaq for the US. The sample periods for both China and the US are from January 2000 through December 2016 (204 months).

Factors	Avg. R-square
Panel A: All individual stocks in China	
MKT	0.417
MKT, SMB	0.542
MKT, VMG	0.509
MKT, SMB, VMG	0.574
Panel B: All but the smallest 30% of stocks in China	
MKT	0.427
MKT, SMB	0.533
MKT, VMG	0.510
MKT, SMB, VMG	0.568
Panel C: All individual stocks in the US	
MKT	0.186
MKT, SMB	0.255
MKT, HML	0.250
MKT, SMB, HML	0.299

Table 6: Abilities of models CH-3 and FF-3 to explain each other's size and value factors
Panel A reports a factor's estimated monthly alpha (in percent) with respect to the other model (with White, 1980, heteroskedasticity-consistent t-statistics in parentheses). Panel B computes the Gibbons-Ross-Shanken (1989) F-test of whether a given model produces zero alphas for the factors of the other model (p-value in parentheses). The sample period is January 2000 through December 2016.

	Alphas with respect to:	
Factors	CH-3	FF-3
Panel A: Alpha (t-statistic)		
FFSMB	0.003	
	(0.05)	
FFHML	-0.269	
	(-0.88)	
SMB		0.439
		(5.55)
VMG		1.217
		(7.93)
Panel B: GRS F-statistics (p-value)		
FFSMB, FFHML	0.524	
	(0.59)	
SMB, VMG		31.432
		(5.083×10^{-13})

Table 7: CAPM alphas and betas for anomalies

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

Category	Anomaly	$\bar{R}$	α	β	$t(\bar{R})$	$t(\alpha)$	$t(\beta)$
Panel A: Unconditional sorts							
Size	Market cap	1.25	1.15	0.20	2.10	2.03	2.08
Value	EP	1.28	1.41	-0.24	2.54	2.96	-3.28
Value	BM	1.13	1.12	0.01	2.12	2.14	0.15
Value	CP	0.56	0.56	0.01	1.23	1.26	0.08
Profitability	ROE	1.18	1.24	-0.12	2.64	2.86	-1.71
Volatility	1-Month vol.	0.95	1.12	-0.31	1.87	2.38	-4.89
Volatility	MAX	0.52	0.69	-0.34	1.15	1.73	-5.68
Reversal	1-Month return	1.16	1.17	-0.02	2.35	2.42	-0.19
Turnover	12-Month turn.	0.12	0.27	-0.29	0.21	0.51	-3.45
Turnover	1-Mo. abn. turn.	0.95	1.03	-0.17	2.02	2.33	-2.29
Panel B: Size-neutralsorts							
Value	EP	1.78	1.88	-0.18	3.93	4.34	-2.97
Value	BM	0.90	0.88	0.04	1.79	1.77	0.46
Value	CP	1.02	1.03	-0.01	2.66	2.71	-0.24
Profitability	ROE	1.38	1.41	-0.08	3.79	3.91	-1.33
Volatility	1-Month vol.	0.62	0.81	-0.37	1.25	1.85	-5.82
Volatility	MAX	0.29	0.42	-0.25	0.70	1.08	-3.92
Reversal	1-Month return	1.28	1.26	0.04	2.83	2.80	0.49
Turnover	12-Month turn.	0.40	0.58	-0.36	0.81	1.34	-5.39
Turnover	1-Mo. abn. turn.	0.87	0.92	-0.09	2.23	2.44	-1.49

Table 8: CH-3 alphas and factor loadings for anomalies.
For each of ten anomalies, the table reports the monthly long-short return spread's CH-3 alpha and factor loadings. For each anomaly, the regression estimated is $R_t = \alpha + \beta_{MKT} MKT_t + \beta_{SMB} SMB_t + \beta_{VMG} VMG_t + \epsilon_t$, where R_t is the anomaly's long-short return spread in month t, MKT_t is the excess market return, SMB_t is CH-3's size factor, and VMG_t is the EP-based value factor. In Panel A, for the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. In Panel B, long / short legs are neutralized with respect to size. That is, we first form size deciles by sorting in the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size group, again with value weighting. Panel B omits the size anomaly, whose alpha equals zero by construction with size-neutral sorts. Our sample period is January 2000 through December 2016. All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Category	Anomaly	α	β_{MKT}	β_{SMB}	β_{VMG}	$t(\alpha)$	$t(\beta_{MKT})$	$t(\beta_{SMB})$	$t(\beta_{VMG})$
Panel A: Unconditional sorts									
Size	Market cap	0.57	0.01	1.42	-0.59	4.37	0.43	32.87	-10.76
Value	EP	0.20	-0.04	-0.40	1.27	0.74	-1.1	-5.08	12.55
Value	BM	0.67	0.10	-0.22	0.52	1.21	1.21	-1.11	2.3
Value	CP	0.17	0.11	-0.29	0.56	0.36	1.44	-1.92	3.5
Profitability	ROE	-0.10	0.05	-0.18	1.21	-0.27	1.24	-1.57	8.98
Volatility	1-Month vol.	0.28	-0.21	-0.08	0.74	0.49	-3.17	-0.43	3.45
Volatility	MAX	0.13	-0.28	0.03	0.43	0.31	-4.26	0.19	2.84
Reversal	1-Month return	1.06	-0.08	0.36	-0.20	2.07	-0.88	2.35	-1.08
Turnover	12-Month turn.	-0.23	-0.14	-0.53	0.80	-0.5	-2.17	-3.4	4.19
Turnover	1-Mo. abn. turn.	1.19	-0.19	0.04	-0.15	2.57	-2.14	0.2	-0.67
Panel B: Size-neutral sorts									
Value	EP	-0.02	-0.02	0.21	1.36	-0.05	-0.46	1.93	12.30
Value	BM	0.56	0.11	-0.26	0.45	0.98	1.59	-1.45	2.03
Value	CP	0.57	0.06	-0.09	0.46	1.39	0.98	-0.78	3.70
Profitability	ROE	-0.35	0.03	0.45	1.07	-1.04	0.62	4.81	8.91
Volatility	1-Month vol.	-0.12	-0.28	0.07	0.69	-0.26	-4.47	0.40	3.70
Volatility	MAX	0.01	-0.21	0.03	0.30	0.01	-3.08	0.22	1.77
Reversal	1-Month return	1.27	-0.02	0.25	-0.21	2.59	-0.19	2.00	-1.32
Turnover	12-Month turn.	-0.12	-0.26	-0.12	0.65	-0.26	-4.45	-0.81	3.99
Turnover	1-Mo. abn. turn.	0.89	-0.11	0.16	-0.10	2.10	-1.65	1.02	-0.52

Table 9: FF-3 alphas and factor loadings for anomalies.
For each of ten anomalies, the table reports the monthly long-short return spread's FF-3 alpha and factor loadings. For each anomaly, the regression estimated is $R_t = \alpha + \beta_{MKT} MKT_t + \beta_{SMB} FFSMB_t + \beta_{HML} FFHML_t + \epsilon_t$, where R_t is the anomaly's long-short return spread in month t, MKT_t is the excess market return, $FFSMB_t$ is FF-3's size factor, and $FFHML_t$ is the BM-based value factor. In Panel A, for the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. In Panel B, long / short legs are neutralized with respect to size. That is, we first form size deciles by sorting in the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size group, again with value weighting. Panel B omits the size anomaly, whose alpha equals zero by construction with size-neutral sorts. Our sample period is January 2000 through December 2016. All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Category	Anomaly	α	β_{MKT}	β_{SMB}	β_{VMG}	$t(\alpha)$	$t(\beta_{MKT})$	$t(\beta_{SMB})$	$t(\beta_{VMG})$
Panel A: Unconditional sorts									
Size	Market cap	0.51	0.05	1.56	-0.15	3.64	2.14	32.72	-2.47
Value	EP	1.45	-0.17	-0.89	0.55	5.15	-3.82	-11.74	7.63
Value	BM	0.07	-0.01	0.00	1.45	0.29	-0.31	-0.03	29.72
Value	CP	0.31	0.04	-0.33	0.71	0.79	0.70	-3.16	6.70
Profitability	ROE	1.90	-0.03	-0.88	-0.15	5.37	-0.54	-9.88	-1.65
Volatility	1-Month vol.	0.90	-0.29	-0.30	0.60	2.06	-4.48	-2.19	4.44
Volatility	MAX	0.65	-0.33	-0.15	0.28	1.44	-5.00	-1.08	2.13
Reversal	1-Month return	1.27	-0.17	0.02	-0.22	2.54	-2.09	0.11	-1.05
Turnover	12-Month turn.	0.21	-0.24	-0.72	0.60	0.47	-4.12	-4.52	4.04
Turnover	1-Mo. abn. turn.	0.80	-0.10	0.48	0.23	1.49	-1.15	3.30	1.36
Panel B: Size-neutral sorts									
Value	EP	1.59	-0.17	-0.34	0.68	4.47	-3.28	-3.58	8.35
Value	BM	-0.14	0.02	-0.04	1.34	-0.61	0.41	-0.68	25.28
Value	CP	0.82	0.00	-0.18	0.49	2.16	0.04	-1.89	4.52
Profitability	ROE	1.94	-0.04	-0.28	-0.31	5.06	-0.65	-2.88	-3.26
Volatility	1-Month vol.	0.59	-0.36	-0.16	0.53	1.41	-5.59	-1.15	4.53
Volatility	MAX	0.33	-0.24	-0.07	0.25	0.77	-3.49	-0.53	2.44
Reversal	1-Month return	0.99	-0.03	0.39	0.18	1.89	-0.36	3.05	1.22
Turnover	12-Month turn.	0.40	-0.35	-0.27	0.53	0.97	-6.17	-1.95	4.41
Turnover	1-Mo. abn. turn.	1.03	-0.10	0.14	-0.17	2.48	-1.52	0.98	-1.10

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

Measure	Unadjusted	CAPM	FF-3	CH-3
Panel A: Unconditional sorts				
Average $\|\alpha\|$	0.91	0.98	0.81	0.39
Average $\|t\|$	1.82	2.06	2.32	1.14
GRS_{10}	8.22	8.31	7.78	2.83
p_{10}	<0.0001	<0.0001	<0.0001	0.0027
GRS_7	3.30	4.05	7.30	1.55
p_7	<0.0001	0.0004	<0.0001	0.1515
Panel B: Size-neutral sorts				
Average $\|\alpha\|$	0.94	1.02	0.84	0.30
Average $\|t\|$	2.22	2.47	2.07	0.61
GRS_{10}	5.47	5.82	6.27	1.62
p_{10}	<0.0001	<0.0001	<0.0001	0.11
GRS_7	5.86	6.39	7.43	2.04
p_7	<0.0001	<0.0001	<0.0001	0.05

Table 11: Anomaly alphas under a four-factor model.
For each of ten anomalies, the table reports the monthly long-short return spread's CH-4 alpha and factor loadings. For each anomaly, the regression estimated is $R_t = \alpha + \beta_{MKT} MKT_t + \beta_{SMB} SMB_t + \beta_{VMG} VMG_t + \beta_{PMO} + \epsilon_t$, where R_t is the anomaly's long-short return spread in month t, MKT_t is the excess market return, SMB_t is CH-3's size factor, and VMG_t is the EP-based value factor, and PMO (pessimistic minus optimistic) is the sentiment factor based on abnormal turnover. In Panel A, for the unconditional sorts, the long leg of an anomaly is the value-weighted portfolio of stocks in the lowest decile of the anomaly measure, and the short leg contains the stocks in the highest decile, with a high value of the measure being associated with lower return. In Panel B, long / short legs are neutralized with respect to size. That is, we first form size deciles by sorting in the previous month's market value. Within each size decile, we then create ten deciles formed by sorting on the anomaly variable. Finally, we form the anomaly's decile portfolios, with each portfolio pooling the stocks in a given anomaly decile across the size group, again with value weighting. Panel B omits the size anomaly, whose alpha equals zero by construction with size-neutral sorts. Our sample period is January 2000 through December 2016. All t-statistics are based on the heteroskedasticity-consistent standard errors of White (1980).

Category	Anomaly	α	β_{MKT}	β_{SMB}	β_{VMG}	β_{PMO}	$t(\alpha)$	$t(\beta_{MKT})$	$t(\beta_{SMB})$	$t(\beta_{VMG})$	$t(\beta_{PMO})$
Panel A: Unconditional sorts											
Size	Market cap	0.63	0.02	1.42	-0.45	0.03	3.63	1.01	30.92	-6.72	0.46
Value	EP	0.15	-0.04	-0.42	1.22	0.04	0.56	-0.91	-5.16	11.24	0.40
Value	BM	0.80	0.08	-0.19	0.51	-0.18	1.34	0.98	-1.04	2.16	-0.70
Value	CP	0.34	0.07	-0.24	0.55	-0.25	0.64	1.04	-1.66	3.25	-1.31
Profitability	ROE	0.10	0.03	-0.18	1.18	-0.20	0.27	0.61	-1.54	8.47	-1.72
Volatility	1-Month vol.	-0.35	-0.12	-0.20	0.67	0.77	-0.68	-2.13	-1.17	3.59	4.44
Volatility	MAX	-0.67	-0.17	-0.10	0.38	0.96	-1.94	-3.01	-0.96	3.25	8.00
Reversal	1-Month return	-0.12	-0.04	0.44	-0.10	0.76	-0.27	-0.61	3.39	-0.57	4.44
Turnover	12-Month turn.	-0.53	-0.11	-0.57	0.73	0.30	-1.11	-1.68	-4.06	3.50	1.30
Turnover	1-Mo. abn. turn.	-0.04	-0.02	-0.14	-0.21	1.45	-0.11	-0.44	-1.36	-1.63	12.50
Panel B: Size-neutral sorts											
Value	EP	0.11	-0.03	0.20	1.36	-0.09	0.37	-0.61	1.87	11.97	-0.67
Value	BM	0.68	0.10	-0.26	0.42	-0.13	1.13	1.38	-1.46	1.79	-0.61
Value	CP	0.62	0.05	-0.09	0.45	-0.07	1.35	0.80	-0.68	3.39	-0.50
Profitability	ROE	-0.31	0.03	0.45	1.12	-0.01	-0.84	0.63	4.59	8.56	-0.05
Volatility	1-Month vol.	-0.64	-0.20	-0.02	0.66	0.63	-1.43	-3.48	-0.10	3.81	3.65
Volatility	MAX	-0.60	-0.12	-0.07	0.27	0.73	-1.37	-1.81	-0.53	1.82	5.23
Reversal	1-Month return	0.96	0.03	0.20	-0.21	0.38	2.02	0.47	1.59	-1.33	2.48
Turnover	12-Month turn.	-0.49	-0.22	-0.16	0.62	0.42	-1.16	-3.49	-1.25	3.62	2.12
Turnover	1-Mo. abn. turn.	-0.12	0.03	0.02	-0.13	1.19	-0.41	0.76	0.21	-1.29	14.25

5 Referee Reports

5.1 Recommendation

Revise and Resubmit.

5.2 Summary and Analysis

This paper constructs Chinese version FF3 factors by exploiting the strong institutional regularity on IPO process. Inspired by the "reverse mergers" and "shell value" ideas, this paper excluding the smallest 30% of firms in the stock universe. Besides, this paper adopts the horse race model to find the most strong value factors in the way Fama and French do in 30 years ago. To diminish unexplained anomalies with respect to CH3, the PMO factor, constructed with an anomaly with largest risk-adjusted alpha, is proposed to construct the CH4 factors model. The major contribution of this paper is to identify the existence of shell values and to excluding the smallest 30% of firms to avoid contaminations of shell values.

There are three problems need to improve: the identification of shell values' effect on small stocks, the external validity of EP-based value factor and the choice of the fourth factor.

In section 3, this paper reviews the constraint IPO process and shows an anecdote of SF express to show the existence of reverse merges. Then they compute the shell values of stocks in an sketchy manner. Table 1 of the paper clearly shows distinct properties of stocks' reaction to fundamentals through different size group in China and US, but I fail to replicate the key result. The authors should provide further details for Table 1 in the paper or consider to provide original codes and data.

In section 4, this paper runs horse race contest Fama-MacBeth regressions to select the strongest value factors in China. Compared to Table 2 in the paper, I find $D(EP < 0)$ instead of EP^+ dominates other value factors, which implies $D(EP < 0)$ instead of EP^+ should be used to construct value factors, but $D(EP < 0)$ is a dummy variable and can not be sorted in deciles. Since the time period is January 2000 through December 2016, since the nontradable shares reform started in 2005, and ended in the end of 2006, Table 4 may be generated by the contaminated data because of nontradable constraints before 2005, therefore, the external validity is questionable. In addition, since 2019, registration-based IPO has been adopted in Chinese main board in sequence, which should relax the IPO constraints and thus reduce shell values of small stocks.Should we exclude the smallest 30% firms in empirical asset pricing research if we use data after 2019 ?

In Table 10 in the paper and my replication, the CH4-adjusted alphas of MAX and 1-Month return anomalies remain modest and marginal significant, e.g., the CH4-adjusted alpha of MAX is -0.59 with t-statistic -1.64 in Panel A, and -0.77 with t-statistic -2.05 in Panel B. The author provides solid reason for the choice of abnormal turnover as the fourth factor but the explanation power need to improve by considering a more concise factor to capture the sentiment.

References

Liu, J., Stambaugh, R. F., & Yuan, Y. (2019). Size and value in china. *Journal of financial economics*, *134*(1), 48–69.