Frontiers in Business, Economics and Management
ISSN: 2766-824X | Vol. 4, No. 2, 2022

A Comparative Analysis of the Fama-French Five-Factor Model

-- Based on the Data of China's SME Board

Ziyue Hua

Southwestern University of Finance and Economics, Chengdu 611130, China

Abstract: The five-factor model by Fama-French was obtained by adding the profitability factor (RWM) and the investment factor (CMA) to the three-factor model by Fama-French. It is still the first paradigm for asset pricing model research to imitate and compare. In the case of the Chinese stock market, there is no consensus on whether the five-factor model explains the anomalies better than other models, and there is a lack of comparative studies on asset pricing models for specific markets such as the SME. Based on the empirical analysis of trading data and financial data of the SME baord, this paper compares the explanatory abilities of the Fama-French five-factor model with four other mainstream models on factors and market anomalies, and provides new and feasible development ideas for the research of asset pricing models for specific markets. The findings of this paper are valuable for the construction of value investment strategies for Chinese small and mid-cap stocks, the empirical study of the effectiveness of market-specific asset pricing models, and the prediction of development trends in ChiNext and NEEQ.

Keywords: Fama-French five-factor model, Asset pricing, SME, Value strategy.

1. Introduction

The founding of Chinese SME Board in 2004 symbolizes the structural reformation of Chinese security market. As an important component to construct multi-level capital market, SME Board was founded to specifically expand financing channels for technological innovations of SMEs, and the development of emerging industries of strategic importance. Compared with companies on Main Board, companies on SME Board bear characteristics like small market size, high potential in growth, high technology, good risk management etc. In recent years, under the influence of the logic behind inclusive finance, to solve the financing for SMEs has been a key direction for policies. SME Board has, to a large extent, pushed the implementation and practice of related policies. While SME Board is taking more proportion in security market and the market quotations in small-cap stocks keeps improving, more attention should be put on the research over asset pricing in SME Board.

Asset pricing has always been the focus and difficulty in financial research, as well as an essential question in modern financial discussion. Based on the portfolio theory from Markowitz (1952), Sharpe (1964) proposed Capital Asset Pricing Model (CAPM), which captured the relation between expected rate of return and risk assets, the formation mechanism of equilibrium price, and became the basis and pillar for pricing theories in financial market. Later, Roll(1977), Stattman(1980) and Banz(1981) conducted researches from different perspectives to further the development of asset pricing model. Based on predecessors' researches, the three-factor model, with better explanatory power, was proposed by Fama and French(1993). In 2015, Fama and French extended the three-factor model into the five-factor model by adding two factors -- profitability and investment. Beside the above models, scholars across the world attempt to construct factors from different indexes, or discover new factors and construct new models from various theories.

This paper takes Chinese SME Board as the research object to compare, from multiple dimensions and evaluation indicators, the explanatory power of different asset pricing models on SME Board and analyze their differences. It also proposes new and practical thoughts for the researches on asset pricing models in specific markets such as the SME Board, and aims, through the above analysis, to provide certain assistance to the understanding of the development of empirical study, investment, and multi-level capital market.

2. Data Selection and Model Construction

2.1. The Selecting and Processing of Sample Data

This paper takes the monthly data of companies listed on SME Board of A-shares from June 2004 to December 2021 as the sample data. (Due to "2 united" and "4 unchanged", the operation of SME Board has not been essentially changed, while it has not been a while since the merger of SZSE's Main Board and SME Board, so this paper does not exam the data before and after the merger date respectively.) 1 year lump-sum deposit rate is chosen as risk-free interest rate, and the market capitalization of individual stocks is chosen as the market capitalization index. The monthly market return is comprehensively calculated by adopting the weighted average of the market capitalization on the basis of the reinvestment of cash dividends. According to data processing done in other Chinese researches, the data 6 months before and after IPO are excluded to avoid "First Day Effect" and "IPO Effect". ST, ST* and financial stocks are excluded to avoid interference. All data of this paper are from CSMAR.

2.2. Factor Calculation and Model Construction

This paper compares 4 main-stream factor models with Fama-French five-factor model (FF-5). The 4 models are Fama-French three-factor model (FF-3), Carhart four-factor model (CH-4), Novy-Marx four-factor model (NV-4), Hou-Xue-Zhang four-factor model/q-factor model (HXZ-4/q).

This paper does not include Stambaugh-Yuan four-factor model or Daniel-Hirshleifer-Sun three-factor model, because the former is constructed on the anomalies in the USA stock market, while the latter could not get enough data about share repurchase and so on in A-share to support its factor calculation. These 2 models provide little help to asset pricing in SME Board and are thus abandoned.

2.2.1. Fama-French Three-factor Model

Beside CAPM, FF-3 has considered size and value effects, and is formulated as the following:

$$R_{it} - R_{ft} = a_i + b_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + e_{it}$$

In this equation, R_{it} is the return on portfolio i for period t; R_{it} is the risk-free return; R_{mt} is the return on the value-weight (VW) market portfolio; $R_{mt} - R_{ft}$ is the market risk premium, which is the market factor represented by MKT in latter part. SMB_t is the size factor, and stands for the return on a diversified portfolio of small stocks minus the return on a diversified portfolio of big stocks for period t, while the size is the market cap for SME stocks at the end of June of year t; HML_t, the value factor, is the difference between the returns on diversified portfolios of high and low B/M stocks for period t, and B/M equals the book value in year $t-1$ divided by the market cap at the end of December of year $t-1$. e_{it} is a zero-mean residual.

Like Fama and French (1993), the size and value factors in this paper use independent sorts of stocks into two size groups and three B/M groups (independent 2 × 3 sorts). According to median market cap, all sample stocks are sorted into two size groups, small (S) or big (B). The B/M breakpoints are the 30th and 70th percentiles of B/M for SME stocks, and all stocks are sorted into 3 B/M groups, high (H), neutral (N), or low (L). Different sorts would intersect with each other. The delayed disclosure of financial statements would cause financial data to mismatch with market data, so the fiscal year starts from July of year t to June of year $t-1$. The construction of the factors are as the following:

Table 1. Construction of Size, B/M factors

Sort	Factors and their components
2×3	$SMB = \frac{SH+SL}{2} - \frac{BH+BL}{2}$ $HML = \frac{SH+BH}{2} - \frac{SL+BL}{2}$

2.2.2. Carhart Four-factor Model

In addition to the FF3 factors, CH-4 includes the momentum factor, and the model is as below:

$$R_{it} - R_{ft} = a_i + b_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + m_i MOM_t + e_{it}$$

In the above equation, MOM_t is the momentum factor. At the end of each month, the stock samples are sorted according to the individual stock return rate of 11 months from $t-12$ to $t-1$, and the momentum factor is built by longing the first 30% and shorting the last 30% of stocks. The momentum factor is constructed by univariate sorting, and the return of long and short portfolios is calculated by arithmetic average.

2.2.3. Novy-Marx Four-factor Model

NV-4 is slightly different from FF-3 in data processing and includes profitability anomalies. The following is the model:

$$R_{it} - R_{ft} = a_i + b_i(R_{mt} - R_{ft}) + h_i HML_t + u_i UMD_t + p_i PMU_t + e_{it}$$

UMD_t is the momentum factor, which, like in CH-4, uses the cumulative return of individual stocks in the past 11 months as the grouping index and updates in months. HML_t and PMU_t are value and profitability factors, which respectively uses log(B/M) and GP as proxy variables (GP is calculated as annual gross profit divided by total assets) and are replaced according to the annual report data of the previous year at the end of June of each year. HML, UMD and PMU are all constructed by double sorting.

It should be emphasized that industry neutrality is required when calculating the portfolio return, which is, for each stock in the portfolio, the return of the industry, to which the stock belongs, is subtracted according to its equal weight.

2.2.4. Fama-French Five-factor Model

On the basis of the three-factor model, Fama and French added profitability and investment factors, and expanded the three-factor model into a five-factor model. The model is shown as follows:

$$R_{it} - R_{ft} = a_i + b_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it}$$

RMW_t, the profitability factor, is the difference between the portfolios with robust profitability (R) and the portfolios with weak profitability (W). Following relevant researches in China on asset pricing models, the operating profitability (OP) is calculated by dividing total operating profit by shareholders' equity. CMA_t is the investment factor and is the difference between the returns on diversified portfolios of the stocks of low and high investment firms, which we call conservative (C) and aggressive (A). Investment (Inv) is the increase in total assets from the fiscal year ending in year $t-2$ to the fiscal year ending in year $t-1$, divided by the total assets at the end of year $t-2$.

FF-5 also uses 2×3 sorts, and except market risk premium factor (MKT), the equation for the other four factors is as follows:

Table 2. Construction of size, B/M, profitability, and investment factors

Sort	Factors and Their Components
2×3	$SMB_{EP} = \frac{SH+SN+SL}{3} - \frac{BH+BN+BL}{3}$ $SMB_{OP} = \frac{SR+SN+SW}{3} - \frac{BR+BN+BW}{3}$ $SMB_{INV} = \frac{SC+SN+SA}{3} - \frac{BC+BN+BA}{3}$ $SMB = \frac{SMB_{EP}+SMB_{OP}+SMB_{INV}}{3}, HML = \frac{SH+BH}{2} - \frac{SL+BL}{2}$ $RMW = \frac{SR+BR}{2} - \frac{SW+BW}{2}, CMA = \frac{SC+BC}{2} - \frac{SA+BA}{2}$

2.2.5. Hou-Xue-Zhang Four-factor Model

The q-factor model is based on the q-theory from the economist James Tobin, and is essentially different from the

previous models. However, the construction of factors and the model form of q-factor model have not changed significantly. The model is shown as below:

$$R_i - R_{ft} = a_i + b_{i,M}(R_m - R_f) + \beta_{i,ME}R_{ME} + \beta_{i,I/A}R_{I/A} + \beta_{i,ROE}R_{ROE} + e_i$$

I/A_t is the investment factor, which is constructed by the change rate of total assets as the proxy variable like the investment factor (CMA) in FF-5. ROE_t is profit index, and ROE is calculated from net profit / net assets. The market factor and investment factor are updated annually with reference to the fiscal year in FF-5 and other models, and the ROE is adjusted according to quarters, which means 1-quarter-lagged ROE is taken as the sorting standard. Following Hou (2015) and other researches, this paper uses the independent sorts, 2×3×3, to calculate the portfolio return.

3. Comparison of Effectiveness of Factor Models

3.1. Descriptive Statistics and Correlation Matrix

Before formally comparing the effectiveness of the model, it is necessary to analyze the similarities and differences between the factors through descriptive statistics and correlation matrix to preliminarily understand the overall situation of the data.

The following table shows the descriptive statistical results of 14 factors, including sample number (N), arithmetic mean (mean), standard deviation (sd), minimum (min), 50th percentile (p50) and maximum (max).

Table 3. Descriptive statistics of factors

Descriptive Statistics of Factors						
	N	Mean	sd	min	p50	max
MKT	186	1.024	8.280	-26.835	1.132	29.604
SMB(FF-3)	186	0.406	5.219	-18.246	0.492	23.114
HML(FF-3)	186	0.315	4.320	-13.704	0.106	18.621
MOM	186	0.474	1.949	-1.861	0.165	5.993
SMB(FF-5)	186	0.516	3.712	-11.274	0.582	12.794
HML(FF-5)	186	0.342	3.716	-14.787	0.127	16.731
RMW	186	-0.426	4.214	-15.831	-0.311	18.217
CMA	186	0.109	2.723	-7.388	0.004	11.294
HML(NV-4)	186	0.317	4.314	-13.655	0.104	18.626
UMD	186	1.290	6.256	-22.675	0.422	30.135
PMU	186	-0.380	4.696	-16.790	-0.128	17.071
ME	150	0.939	3.910	-6.554	0.554	30.225
I/A	150	-0.002	4.176	-37.496	-0.038	8.945
ROE	150	0.065	4.868	-13.644	0.080	41.856

Through descriptive statistics, it can be inferred that the ability of different factors to capture excess returns varies greatly.

According to the fact that the statistical characteristics of the factors constructed by the same proxy index do not converge, it shows that the factor return is sensitive to the data frequency and sorting method.

Therefore, factors constructed by different methods could have significantly different explanatory power to anomalies, and the degree of deviation needs to be further checked.

In order to observe the commonness between factors, we calculate the correlation coefficient between each two factors and integrate the result into a correlation matrix, as shown in the following table:

Table 4. Correlation matrix of factors

Correlation Matrix of Factors							
	RMW	CMA	MOM	PMU	UMD	I/A	ROE
RMW	1.000						
CMA	-0.480***	1.000					
MOM	0.057	-0.013	1.000				
PMU	0.702***	-0.369***	0.034	1.000			
UMD	0.148**	0.095	0.372***	-0.009	1.000		
I/A	-0.378***	0.551***	-0.010	-0.190**	-0.318***	1.000	
ROE	0.513***	-0.313***	0.098	0.270***	0.455***	-0.582***	1.000
MKT	-0.213***	0.045	-0.058	-0.167**	0.002	0.217***	-0.297***
SMB	-0.396***	0.282***	-0.209***	-0.260***	-0.174**	0.142*	-0.262***
HML	-0.387***	0.191***	0.011	-0.561***	0.080	0.044	-0.035
*** p<.01, ** p<.05, * p<.1							

The correlation matrix shows: 1. in general, there is a certain connection between the factors of SME board; 2. the connection between momentum factors, MOM and UMD, and other factors is relatively insignificant, and the

momentum effect is weakly affected by other effects.

However, the correlation analysis can only provide a reference for the connection between factors. How the factors of the compared models replace and explain each other should be further tested and observed.

3.2. The Comparison of Model Explanatory Power

Since FF-5 is the "axis" in the asset models, FF-5 will be used as the benchmark to compare all models. The explanatory power of FF-5 and other models will be compared from the following three dimensions: 1. regression of the other models to the anomalies contained in this model; 2. joint regression of this model to all factors of the other models; 3. the regression of FF-5 and the other models to the anomalies excluded from all models. The following three problems are meant to be solved: whether the chosen model contains excess returns that cannot be explained by other models, whether the other models contain excess returns that cannot be explained by this model, and how many kinds of market abnormal returns can be explained by each model.

3.2.1. Regression of Other Models to Factors of This model

Table 5. Regression of the other models to RMW, CMA

Regression of the other models to RMW, CMA								
	RMW				CMA			
	FF-3	CH-4	NV-4	HXZ-4	FF-3	CH-4	NV-4	HXZ-4
MKT	-.068*	-.068*	-.049	-.084**	-.002	-.002	-.007	.011
	(.039)	(.039)	(.034)	(.033)	(.034)	(.034)	(.032)	(.03)
SMB*	-.32***	-.323***		-.252***	.147**	.15**		.296***
	(.099)	(.101)		(.091)	(.06)	(.06)		(.072)
HML*	-.357**	-.357**	.074		.117	.117	-.059	
	(.141)	(.141)	(.151)		(.088)	(.088)	(.091)	
MOM		-.045				.051		
		(.122)				(.118)		
UMD			.099**				.044	
			(.049)				(.042)	
PMU			.658***				-.248***	
			(.059)				(.067)	
I/A				-.143				.413***
				(.098)				(.063)
ROE				.269***				-.017
				(.101)				(.058)
α	-.114	-.092	-.278	-.145	.015	-.011	-.016	-.074
	(.259)	(.264)	(.189)	(.204)	(.18)	(.19)	(.173)	(.118)
Standard errors are in parentheses *** p<.01, ** p<.05, * p<.1								

The above regressions test how the profitability and investment factors in FF-5 are explained by other models, and SMB* and HML* are the size and value factors in each model. From the regression results adjusted by Newey-West, it can be seen that α in each regression is not significant, and RMW and CMA do not contain excess returns that cannot be explained by other models.

In order to compare with RMW and CMA of FF-5, further regression is done toward MOM, UMD, PMU, I/A and ROE. The following table demonstrates α obtained from each regression.

Table 6. Regression of the other models to MOM, UMD, PMU, I/A and ROE

Regression of the Other Models to MOM, UMD, PMU, I/A and ROE					
	FF-3	CH-4	NV-4	FF-5	HXZ-4
MOM	.509***		.343***	.531***	.517***
	(.09)		(.074)	(.094)	(.119)
UMD	1.289***	.691		1.392***	1***
	(.406)	(.435)		(.407)	(.348)
PMU	-.044	-.035		.062	-.179
	(.299)	(.308)		(.248)	(.333)
I/A	.117	.146	.164	-.056	
	(.277)	(.268)	(.241)	(.232)	
ROE	.134	.017	-.215	.503*	
	(.343)	(.352)	(.261)	(.272)	
Standard errors are in parentheses *** p<.01, ** p<.05, * p<.1					

Similar to RMW and CMA, I/A, which stands for investment style, PMU and ROE, which stand for profitability,

can be well explained by other models; however, the regressions that take MOM and UMD as variables to be explained mostly produce significant alphas, while only NV-4 and CH-4 can explain the momentum premium to a greater extent.

3.2.2. Regression of This Model to All Factor of the Other Models

The following compares the explanatory power of this model to factors in the other models, in which GRS is used to test. In the comparisons, factors contained in FF-3, CH-4, NV-4, FF-5 and HXZ-4 are successively selected as the explanatory variables, while all the factors of all other models are chosen as variables to be explained, the following tests are conducted to check whether a given model produces jointly zero alphas across anomalies from other models.

Table 7. The results of GRS test

The Result of GRS					
Explanatory variables	FF-3	CH-4	NV-4	FF-5	HXZ-4
GRS	3.0098351	2.3648265	5.296778	4.6150692	1.6055693
P	.00128051	.01305157	1.432e-06	.00002472	.11117979

According to the GRS regression, the factors of the other models, adjusted by FF-5, still have mispricing in which coefficient is significant; but for q-factor model, the GRS p-value of 0.11117979 fails to reject the joint hypothesis that all factors produce zero q-factor alphas, which indicates that q-factor model has better explanatory power on the factors of the other models.

3.2.3. Regression of Each Model to Market Anomalies

In order to find the most suitable asset pricing model for the SME board, it is better to compare the explanatory power of different models on anomalies.

There is a lack of researches on the anomalies of China's SME board, so the 6 categories commonly used in researches on China A-Shares are selected, including Size, Value, Quality, Risk, Past Returns and trading, as well as 24 anomalies to be tested. The anomalies, references and short descriptions are shown in the following table.

Table 8. Category, reference and description of anomalies

Anomalies		
Category	Reference	Short description
Size		
SIZE	Hsu et al. (2018)	SIZE is calculated as the natural logarithm of a firm's end-of-month total A-share market capitalization
Value		
Book-to-Market (BM)	Fama and French(1992)	Book-to-market is calculated as the ratio of book value to market value. Book value is defined as common shareholder's equity (total shareholder's equity minus the book value of preferred stocks) excluding minority interests. Market value refers to a proxy of end-of-month total market capitalization; the number of shares outstanding (including B- and H-shares) times the A-share's price. We exclude firms with negative book values.
Earnings-to-Price (EP)	Basu (1977)	EP. Following Liu et al. (2019), earnings-to-price is calculated as the ratio of total earnings to price. Total earnings is defined as net profit excluding minority interest income. Price refers to a proxy of end-of-month total market capitalization; the number of shares outstanding (including B- and H-shares) times the A-share's price. We exclude firms with negative earnings.
Sales-to-Price (SP)	Barbee et al.(1996)	SP. Following Hsu et al. (2018), sales-to-price is calculated as the ratio of total sales to price. Total sales is equal to operating revenue. Price refers to a proxy of end-of-month total market capitalization; the number of shares outstanding (including B- and H-shares) times the A-share's price
Dividend-to-Price (DP)	Hsu et al. (2018)	DP. Following Hsu et al. (2018), dividend-to-price is calculated as the ratio of total dividends to price. Total dividends is calculated as the total monetary value of dividends paid out to shareholder's over the previous reporting period. Price refers to a proxy of end-of-month total market capitalization; the number of shares outstanding (including B- and H-shares) times the A-share's price. We exclude firms that do not pay dividends.
Cash Flow-to-Price (CP)	LSV(1994)	CP. Following Liu et al. (2019), cash-flow-to-price is calculated as the ratio of total cash flow to price. Total cash flow is calculated as the net increase in cash or cash equivalents over

		the previous reporting period. Price refers to a proxy of end-of-month total market capitalization; the number of shares outstanding (including B- and H-shares) times the A-share's price. We exclude firms with negative cash flows.
Quality		
Return-on-Equity (ROE)	Haugen and Baker (1996)	ROE. Following Haugen and Baker (1996) and Liu et al. (2019), return on equity is calculated as the ratio of total earnings to book equity value. Total earnings is defined as net profit excluding minority interest income. Book value is defined as common shareholder's equity (total shareholder's equity minus the book value of preferred stocks) excluding minority interests.
Gross Profitability (GP)	Novy-Marx(2013)	GP. Following Novy-Marx (2013), gross profitability is calculated as the ratio of gross profit to total assets. Gross profit is defined as operating revenue minus operating costs.
Operating Profit (OP)	Fama and French(2015)	OP. Following Fama and French (2015), operating profit is calculated as the ratio of operating profit excluding interest expense to book value. Interest expense is excluded by subtracting total interest expense from operating profit. Book value is defined as common shareholder's equity (total shareholder's equity minus the book value of preferred stocks) excluding minority interests.
Asset Growth (INV A)	Cooper et al. (2008); Hou et al. (2020)	INV ASSET. Following Hou et al. (2020), asset growth is calculated as the quarter-to-quarter asset growth rate, i.e., the difference between total assets in the previous two quarters divided by total assets two quarters ago.
Book Value Growth (INV B)	Hou et al. (2020)	INB BOOK. Following Hou et al. (2020), book value growth is calculated as the quarter-to-quarter book value growth rate, i.e., the difference between total book value in the previous two quarters divided by total book value two quarters ago. Book value is defined as common share-holder's equity (total shareholder's equity minus the book value of preferred stocks) excluding minority interests.
Accruals (ACC)	Sloan(1996);Liu et al.(2019)	ACC. Following Liu et al. (2019), Sloan (1996) firm-level accruals are calculated as $ACC = 2 \times Accrual_t / (TA_{t-1} + TA_t)$ $Accrual_t = (\Delta CA_t - \Delta Cash_t) - (\Delta CL_t - \Delta STD_t - \Delta TP_t) - DP_t$ Where TA_t is total assets, $Cash_t$ is the balance of cash and cash equivalents, CL_t is current liabilities, STD_t is the sum of notes payable and long-term debt due within one year, TP_t is taxes payable, DP_t is the sum of depreciation of fixed assets, oil and gas assets, and bearer biological assets, and intangible asset amortization. Δ denotes the year-on-year difference and t denotes the year.
Net Operating Assets (NOA)	Hirshleifer et al.(2004);Liu et al.(2019)	NOA. Following Liu et al. (2019), net operating assets is calculated as net operating assets in year t scaled by total assets in year t-1. The numerator is calculated as the difference between operating assets and operating liabilities. Operating assets is defined as total assets minus balance sheet cash, minus short-term investment. Operating liabilities represents total assets minus short-term loans, minus long-term loans, minus minority interest, minus common shareholder's equity excluding minority interest.
Profit Margin (PM)	Soliman(2008)	PM. Profit Margin = EBIT / Revenues. Updated annually.
Risk		
1M-Volatility (VOL 1M)	Ang et al.(2006)	VOL 1M. Following Ang et al. (2006), one-month volatility is calculated as the standard deviation of daily stock returns over the past twenty trading days.
3Y-Volatility (VOL 3Y)	Blitz et al.(2006)	VOL 3Y. Following Blitz and Van Vliet (2007), three-year volatility is calculated as the standard deviation of monthly stock

		returns over the past 36 months.
MAX	Bali et al.(2010)	MAX. Following Bali et al. (2011), the maximum daily return over the past twenty trading days equals the highest adjusted daily return over this period. Daily returns are adjusted in two ways. First, to account for a possible downward bias caused by trade suspensions, daily returns on suspended trading days are set equal to -99. Second, as suggested by Cheema et al. (2020), we aggregate the daily returns after a stock hits the upper price limit.
Past Returns		
MOM	Jegadeesh and Titman (1993)	MOM. Following Jegadeesh and Titman (1993), momentum at month t is calculated as the cumulative monthly stock return over the previous twelve months excluding the most recent month.
Lagged MOM (lag MOM)	Novy-Marx(2012)	Lagged Momentum: Lagged Momentum is calculated as the cumulative monthly stock return from $t-13$ through $t-8$. Updated monthly.
Short-term Reversals (REV ST)	Jegadeesh(1990)	REV ST. The short-term reversal is calculated as the cumulative monthly stock return over the past 20 trading days. Updated monthly.
Long-term Reversals (REV LT)	De Bondt and Thaler (1985)	REV LT. Following De Bondt and Thaler (1985), long-term reversal is calculated as the cumulative monthly stock return over the past five years excluding the previous year, i.e., month t = 60 to month $t-13$
52-Week High (52-HIGH)	George and Hwang(2004); Wang Mingtao et al.(2015)	52-Week High: Price scaled by the highest price during the last 12 months. Updated monthly.
Trading		
ILLIQ	Amihud(2002)	ILLIQ. Following Amihud (2002), illiquidity is calculated as the average stock illiquidity over the past twenty trading days. Stock illiquidity at day t is defined as $Illiqt = \|Ret_t\| / Volume_t$ Where $\|Ret_t\|$ is the absolute value of the stock's daily return and $Volume_t$ is the dollar trading volume on day t, i.e., total number of shares traded times close price.
Volume Variance (VV)	Chordia et al.(2001)	Volume Variance: standard deviation of monthly trading volume over the last 36 months. Updated monthly.

The construction of the anomalies is similar to the construction of model factors, which is to sort the stock samples according to various indicators at the end of the holding-period, and long the top 10% and short the bottom 10%.

The construction of the anomalies adopts the univariate sorting, and the long-short portfolio returns are calculated by using the weighted average of the market capitalization.

Based on the above anomalies, the models will be comprehensively compared from the following 5 dimensions: the number of significant α, the mean of absolute value of α, the mean of absolute value of t statistic, F statistic of GRS test and its p-value.

Table 9. Comparison of the ability of factor models to explain anomalies

The Comparison of Model Explanatory Power					
	FF-3	CH-4	NV-4	FF-5	HXZ-4
Number Sig. at 5%	6	7	5	5	4
Number Sig. at 10%	7	7	5	7	4
Average|α| (%)	.16972251	.09420233	.13754319	.25525271	.16237252
Average|t|	0.508083	0.497875	0.455917	0.5255	0.474583
GRS	20.60629	18.929498	22.505342	26.871249	21.317151
P	0	0	0	0	0

FF-5 is not the best among 5 models when evaluated with 5 indexes together. At the 5% and 10% significance levels, regression of q-model to anomaly portfolios obtains the fewest α with significant coefficient; but |α|, obtained by the regression of CH-4 as the explanatory variable, and F statistics, tested by GRS, are relatively low. It can be

preliminarily concluded that the q-model and CH-4 are more suitable pricing models for stocks in SME board than FF-5, FF-3 and NV-4. These two have different advantages that q-model is more convincing in explaining Trading anomalies and Quality anomalies, while CH-4 is more trustworthy in explaining Past Returns anomalies. This is also in accordance with the results of mutual comparisons between the above models

Beside the comparison of explanatory power, Parsimony Index by Daniel et al. (2018) is introduced, which mainly uses factors and the number of indexes that construct the factors to calculate the complexity of the model and examine the efficiency of models in explaining. The formulations is: "0 - (the number of all factors + the number of all used indexes (reused indexes only count for once)". The Parsimony Indexes of FF-3, CH-4, NV-4, FF-5 and HXZ-4 are respectively: -5, -7, -9, -9, -7. The smaller the value of Parsimony Index is, the more complex the model. This shows that CH-4 and HXZ-4 are more efficient than other models.

The above analysis shows: for SME board, q-model and CH-4 outweigh FF-5 in explaining factors and anomalies.

4. Conclusions

This paper chooses annual and monthly data from 2004 to 2021 of stocks in China's SME board and refers to Fama and French (1993), Carhart (1997), Novy Marx (2013), Fama and French (2015) and Hou et al. (2015) to construct FF-3, CH-4, NV-4, FF-5, and HXZ-4 models. This paper also uses the above models to explain each model, tests the anomalies of SME board, and uses FF-5 as the comparison benchmark to comprehensively compare the effectiveness of the mainstream factor models.

The research shows:

1. There is a significant correlation between the factors of the mainstream models, but the factors are sensitive to the update frequency and construction mode of their indexes, and different factors have different explanatory power to anomalies;

2. Among all anomalies in Chinese SME board, Past Returns, Risk and Trading are significant anomalies, and Trading and Risk are the most significant, so it can be considered to build a portfolio based on the above indexes to obtain excess return in investing in SME stocks.

3. From the perspectives of explaining factors and market anomalies, FF-5 is not the most suitable pricing model for Chinese SME board, which is consistent with some researches on the full samples of China A-shares.

4. CH-4 and q-model are relatively better pricing models for SME board. These two have different advantages that CH-4 is more trustworthy in explaining momentums related to Past Returns, while q-model can explain more excess returns related to Trading and Quality. If two models could be combined, it is possible to have a more objective and comprehensive understanding of asset pricing in SME board, and have the opportunity to obtain more premiums.

References

[1] Weiwang Du, Shuguang Xiao: A Research on the Improvement of Fama-French Five-factor Model in Chinese Stock Market. Journal of Huaqiao University(Philosophy & Social Sciences). (2018) No. 3, p. 39-53.

[2] Zhibing Li, Guangyi Yang, Yongchang Feng, and Liang Jin: An Empirical Test of Fama-French Five-factor Model in Chinese Stock Market. Journal of Financial Research. (2017) No. 06, p. 191-206.

[3] Jifu Liu: On the Development of Asset Pricing Theories: from One-factor Model to Five-factor Model. Times Finance. (2020) No. 30, p. 80-82.

[4] Tong Lu: The Features and Improvement Direction of Corporate Governance in SME Board and GEM Board. Tsinghua Financial Review. (2017) No. 02, p. 81-84.

[5] Runping Ma, Jie Shen: A Study on the Effectiveness of Chinese GEM Board——an Analysis Based on Fama-French Five-factor Model. On Economic Problems. (2021) No. 09, p. 46-52.

[6] Lihui Tian, Guanying Wang: Five-factor Model for Stock Pricing in China: How does Trading Volume Affect Stock Return?. Nankai Economic Studies. (2014) No. 02, p. 54-75.

[7] Xiaoyue Wang: The Characteristics, Motivations and Suggestions of Dividend Policies of Companies in SME Board. Management Engineer. (2018) No. 01, p. 16-21.

[8] Shaohua Zhang, Huiling Chen: Is TFP an Effective Capital Asset Pricing Factor — a Test of Fama-French Model based on Chinese Stock Market. China Economic Studies. (2021) No. 02, p. 125-140.

[9] Zhaoqin Zhang, Qiuyu Wang, Shaohua Zhang: A Research on the Profit Effect and Pricing Power of A-share Market — Based on Fama-French Five-factor Model. Exploration of Financial Theory. (2022) No. 1, p. 1-18.

[10] Shengmin Zhao, Honglei Yan, Kai Zhang: Is the Fama-French Five-factor Model Better than the Three-factor Model -- Empirical Evidence from China's A-share market. Nankai Economic Studies. (2016) No. 02, p. 41-59.

[11] Barillas F. and J. Shanken: Comparing Asset Pricing Models. Journal of Finance. Vol. 73 (2018) No. 2, p. 715 – 754.

[12] Barillas, F. and J. Shanken: Which alpha? Review of Financial Studies. Vol. 30(2017) No. 4, p. 1316 – 1338.

[13] C. Han and Y. Shi.: Chinese stock anomalies and investor sentiment. Pacific-Basin Finance Journal. Vol. 73 (2022).

[14] Carhart, M. M.: On persistence in mutual fund performance. Journal of Finance. Vol. 52 (1997) No. 1, p. 57-82.

[15] Daniel, K. D., D. A. Hirshleifer, and L. Sun: Short- and long-horizon behavioral factors. Review of Financial Studies. Vol. 33(2020) No. 4, p. 1673-1736.

[16] Fama, E. F. and K. R. French: Comparing cross-section and time-series factor models. Review of Financial Studies. Vol. 33(2020) No. 5, p. 1891-1926.

[17] Fama, E. F. and K. R. French: Choosing factors. Journal of Financial Economics. Vol. 128 (2018) No. 2, p. 234-252.

[18] Fama, E. F. and K. R. French: A five-factor asset pricing model. Journal of Financial Economics. Vol. 116(2015) No.1, p. 1-22.

[19] Fama, E. F. and K. R. French: Common risk factors in the returns on stocks and bonds. Journal of Financial Economics. Vol. 33(1993) No. 1, p. 3-56.

[20] Fama, E.F. and K. R. French: The cross-section of expected stock returns. Journal of Finance. Vol. 47(1992) No. 2, p. 427-465.

[21] Hou, K., C. Xue, and L. Zhang: Digesting anomalies: an investment approach. The Review of Financial Studies. Vol. 28(2015) No. 3, p. 650-705.

[22] Jegadeesh, N. and S. Titman: Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance. Vol. 48(1993) No. 1, p. 65-91.

[23] Liao, Z. and Y. Liu: Optimal cross-sectional regression. Working paper, 2021, Information on: https://ssrn.com/abstract=3719299

[24] Liu, J., R. F. Stambaugh, and Y. Yuan. Size and Value in China. Journal of Financial Economics. Vol. 134(2019) No. 1, p. 48-69.

[25] Maarten, Jansen, Laurens, Swinkels, Weili, Zhou. Anomalies in the China A-share market. Pac. Basin Financ. Vol. 68(2021) Article 101607.

[26] Novy-Marx, R.. The Other Side of Value: the Gross Profitability Premium. Journal of Financial Economics. Vol. 108(2013) No. 1, p. 1-28.

[27] Stambaugh, R. F. and Y. Yuan. Mispricing Factors. Review of Financial Studies. Vol. 30(2016) No. 4, p. 1270-1315.