Is the Fama French model robust in the Chinese stock market?

Hongxi Chen

Business, Economics & Law, The University of Queensland, Brisbane, Queensland, 4072, Australia

117020015@link.cuhk.edu.cn

Abstract. Many previous studies have analyzed the different stock return models in various segments of the Chinese stock market. This essay analyzes the Fama-French five factors and three factors model in the main board, the second board, and STAR market. It finds that the three-factor models are efficient in all three markets expect for the STAR market, while the five-factor model is inefficient in testing the second board market alone and overall efficient in testing the mixed market and the STAR market. Furthermore, it is also found that the MKT factors are the most significant contributors to the portfolio excess return in all six situations. The HML and SML factors also exhibit constant relations with the portfolio return. Besides, it is found that HML and SMB factors are not significantly correlated with the stock's performance on the STAR market. Overall, the relative efficiency of Fama French's three and five-factor model in explaining these markets is proved. Investors can use these two models to analyze the stock return and earn profit in the next period, and policymakers can also employ this model to understand the market and decide suitable fiscal and monetary policies.

Keywords: The main board market, second board, Fama-French, explanation power.

1. Introduction

The pricing problem of financial products has always been a core issue in the study of finance since it helps us to judge the degree of deviation between the current market price and the intrinsic value. One dominant way to solve this problem is asset pricing models, which incorporate mathematical and statistical methods into pricing. The mean-variance model proposed by Harry Markowitz in 1952 laid the theoretical foundation [1]. Sharp, Lintner, and Mossin later constructed the capital asset pricing model (CAPM), revealing the functional relationship between expected asset returns and systematic risks [2]. Ross in 1976 pointed out that stock returns are not only affected by systematic risks but also introduce other factors into the pricing model, constructing the arbitrage pricing theory (APT) [3]. In 1993, Fama and French proposed that market factors, size factors, and book-to-market ratio factors could explain the pricing problem of risky assets, which formed the Fama-French three factors models, and they later added investment style and profitability factors and updated to the five-factor model [4].

The Chinese stock market has unique characteristics in terms of development history, institutional background, investment groups, and financial system. Some previous articles have tried to analyze the effectiveness of the Fama-French five factors models in China concerning different markets, time horizons, and industries. Zhao, Yan, and Zhang found that the Fama French three-factor model is more suitable for China's stock market compared to the five-factor model [5]. The reason is that investors in the Chinese market are more concerned about the valuation level of listed companies rather than their development prospects and growth space. Liu and Yao found that the investment factor and scale factor were significant when verifying the applicability of the Fama French five-factor model in China, while the significance of the remaining three factors was weak [6]. Wu found that although the Fama French five-factor model cannot fully explain the changes in cross-sectional returns of stocks in China's A-share market, the five-factor model is more powerful in explaining the changes in cross-sectional returns of stocks in China's A-share market than the previous three-factor model [7]. Its explanatory power mainly comes from scale effects and profit effects, while value effects and investment effects are relatively weak overall. Ma proved that the Fama-French five-factor model can better explain the stock return rate of China's Internet industry with 53 Internet and related

service industry enterprises [8]. Yin compared the applicability of various asset pricing models in the Chinese stock market and found that the Fama French five-factor model is the most suitable so far [9]. Li and Yang Feng proved that the five-factor model in China has stronger explanatory power than CAPM, the three-factor model, and Carhart's four-factor model [10]. The reason is that China implemented the equity division reform in 2005. The reform restored the pricing function and resource allocation function of China's capital market, resulting in significant risk premiums for profitability, investment ability, and momentum factors.

Nevertheless, there are limited articles discussing the pricing mechanisms of the Chinese stock market. Moreover, due to the later development of the Chinese financial system, the stock pricing mechanism is not very similar to that of the foreign market. Thus, the articles focus on the specificity of the Chinese stock market and test whether Fama French's three and five factors model is effective and the explanatory power of the factors in the combination market of Shenzhen and Shangzheng's A shares (main board) and Second board. Further, the essay also places the model under the latest and newest segment of the Chinese stock market to further examine the models' efficiencies in the latest stock market.

This essay has four sections. The first section is a review of the development of asset pricing theory and a literature review. In the section two, the different segments of the Chinese stock market is introduced. Then, the process for the preparation of data and factors' explanation in the model will be emphasized. In section three, the statistical summary, the analysis, and an explanation of regression results will be given. In section four, the conclusion, implication, and limitations of the results will be analyzed and discussed.

2. Methodology

2.1. Market selection

The following four segments in the Chinese stock market are introduced and selected. There are some differences between them.

The first market is the main board market. The main board market contains Shenzheng's and Shangzheng's A shares. Its companies are always mature and large [7]. The main board also has stricter requirements regarding the ability of management, profitability, capital, and financial situation [7]. Besides, the risk and return rates of companies are relatively low.

The second segment is the second board market, who contains small and medium-sized enterprises and emerging companies that cannot be listed on the main board temporarily [9]. Contrary to the main board market, the companies on the second board are always unmature, innovative and relatively small in size and capital amount [9]. The second board market also has fewer limitation. Thus, the risk and return rate on the second board market are usually higher.

Lastly is The Science and Technology Innovation Board (STAR Market) board, which exhibits several distinct characteristics, including a strong focus on high-tech, high-growth, and high-risk enterprises. [10]. It implements stringent issuance conditions and limits the eligibility of investors. STAR Market also boasts a flexible pricing mechanism and a listing process that is both simplified and more market-oriented [10]. It primarily targets emerging industries and employs hierarchical listing standards

Thus, the stock market of the main board, the second board, and STAR market are chosen to analyze. The first two markets are the main points since the mixed market can represent a large portion of the whole market of the Chinese stock market. The main board and second board market can reflect the current situation of the whole economy. On the other hand, the STAR are used to test the efficiencies of model in the relative new and updated markets since the companies on these two market are later listed and with very different characters compared with those on the main board and second board market.

2.2. Data Selection

Since the high representativeness of the mixed market in the Chinese stock market, it is used as the main sample to test the model efficiency. Then the test on the main board and second board market will be conducted separately. For the STAR market, it will also be tested to further examine the model's efficiencies. The Fama French three factors and five factors model will be used on each situation.

Thus, the first step is to find the value of factors of the model from CSMAR (Fama French five factors research). The frequency is daily and the period is 10 years from the first day of 2012 to the last day of 2021 for the main board and second board market. The 10-year horizon and high frequency make the data ample and make the regression results more precise and representative. For STAR market, their data time period is limited in recent 2.5 years, which is not a problem since these recent data are enough to test the model in latest market situation.

After that, the data Ri, which is the daily return rate of the different markets, is also found from CSMAR (Mixed and independent stock market daily return). The portfolio construction method selected in the data is 2 (the classification of 2*2). The risk-free rate is also gathered from CSMAR (stock market daily risk-free return rate). However, since the frequency of it is not incorporated into the frequency of other data, the average of daily risk-free rate is calculated and used to calculate the excess market return.

2.3. Modeling

The main model selected is the Fama French five factors model. The E[Ri]-Rf represents the excess return of the stock market. There are five different factors in the formula: MKT, SMB, HML, RMV, and CMA, which represent factors of market risk premium, size, Book-to-Market ratio, Profitability, and Investment Pattern, respectively.

$$E(R_i) - R_f = \beta_{i,mkt}\left(E[R_m] - R_f\right) + \beta_{i,smb}E[R_{smb}] + \beta_{i,hml}E[R_{hml}] + \beta_{i,rmv}E[R_{rmv}] + \beta_{i,cma}E[R_{cma}] \quad (1)$$

MKT is equal to the stock market returns rate including reinvestment minus the risk-free rate. SMB is the difference between the yields of small-cap portfolios and large-cap portfolios. The median market value is used to divide all firms into small and big groups. The HML is the difference in the yields between the high book-to-market ratio portfolio and the low book-to-market ratio portfolio. The 30th and 70th percentiles of the Book-to-Market ratio are used to divide all firms into three groups. RMV is the difference in the yields between the high-profit stock portfolio and the low-profit portfolio. The 30th and 70th percentiles serve as data breakpoints. CMA is the difference between the returns of low-investment-proportion stock portfolios and high-investment-proportion stock portfolios. The 30th and 70th percentiles serve as data breakpoints. All factors are calculated through total value-weighted.

The Fama French three factor model is also used to conduct regression. The result of it will be compared with that of five factor's model and see whether the deduction of some factors will affect the efficiencies of model. The factors in the three factors model is totally the same with that in the five factors model but without RMV and CMA factor. The three factors model is listed below.

$$E(R_i) - R_f = \beta_{i,mkt}\left(E[R_m] - R_f\right) + \beta_{i,smb}E[R_{smb}] + \beta_{i,hml}E[R_{hml}] \quad (2)$$

After the daily data of Rf, Ri, and five factors are collected, the regression analysis will be implemented and find the value and statistical significance of the each factor's coefficient (β).

3. Empirical results

3.1. Descriptive Statistics- mixed market

Before doing the regression, the distribution character of the data of five factors in the combination of the main board and the second board market is analyzed and shown in the table1. The MKT factor has the highest mean value of 0.0550%, the highest standard deviation of 0.0292%, and the largest range. The SMB factor has the highest median value 0.1135%. The HML has the lowest mean value of -0.0067% and the lowest median value of -0.0262%. The RMW has the second lowest median value -0.0240% and standard deviation 0.0088%. It also has the highest kurtosis, suggesting that its data concentration is the highest. The CMA has the lowest standard deviation 0.0062% smallest kurtosis value 2.6965 and smallest range. The data of MKT, SMB, and CMA are negatively skewness, and the extent of SMB is the highest. On the contrary, the HML and RMW are positively skewness.

Table 1. Summary statistics (%)

Variable	Mean	Median	Min	Max	SD	Kurtosis	Skewness
MKT	0.0550	0.0928	-9.4484	7.3827	0.0292	6.5728	-0.8684
SMB	0.0383	0.1135	-7.1706	5.3307	0.0182	6.7828	-1.0320
HML	-0.0067	-0.0262	-2.9935	3.6975	0.0107	4.8654	0.6478
RMW	0.0001	-0.0240	-2.5113	3.9406	0.0088	7.1760	0.9128
CMA	-0.0018	0.0022	-1.8908	1.3929	0.0062	2.6965	-0.3537

Table 2 represents the correlation matrix of these five factors. The table suggests that when MKT increases by 1 unit, SMB, HML, RMW, and CMA increase by 0.2562, -0.3524,-0.2576, and 0.0546 units respectively, when SMB increases by 1 unit, the HML, RMW, and CMA increase by -0.6356, -0.7692, and 0.2788 units respectively. When HML increases by 1 unit, the RMW, and CMA increase by 0.3947 and 0.0532 units respectively. When RMW increases by 1 unit, CMA increases by -0.4987 units.

Table 2. Correlation matrix

Variable	MKT	SMB	HML	RMW	CMA
MKT	1				
SMB	0.2562	1			
HML	-0.3524	-0.6356	1		
RMW	-0.2576	-0.7692	0.3947	1	
CMA	0.0547	0.2788	0.0532	-0.4987	1

3.2. Result analysis of the main board and the second board

The regression results are shown in the table 3. The odd number columns represent the results of the five-factor model, and the even number columns represent the three-factor model. The second board market results are in (1) and (2) columns. The main board market results are in (3) and (4) columns. The mixed market (combination of previous two market) results are in (5) and (6) columns. The Y represents the portfolio excess return $R_i - R_f$.

Table 3. Regression results of two models in three different stock markets.

	(1)	(2)	(3)	(4)	(5)	(6)
	The second board		The main board		Mixed market	
	5-factor model	3-factor model	5-factor model	3-factor model	5-factor model	3-factor model
MKT	0.976***	0.977***	1.007***	1.007***	1.005***	1.006***
	(306.01)	(309.78)	(812.81)	(818.32)	(777.01)	(783.34)
SMB	-0.012	0.011	0.055***	0.055***	0.051***	0.056***
	(-0.87)	(0.95)	(15.08)	(21.85)	(14.00)	(22.16)
HML	-0.005	0.009	-0.106***	-0.101***	-0.102***	-0.095***
	(-0.25)	(0.48)	(-23.74)	(-23.45)	(-21.94)	(-21.59)
RMW	-0.050**		0.016*		0.003	
	(-2.74)		(2.41)		(0.41)	
CMA	0.020		0.034***		0.031***	
	(1.14)		(5.22)		(4.47)	
R-sq	0.976	0.976	0.997	0.997	0.997	0.997
Constant	-0.0003***	-0.0003***	-0.0056***	-0.0056***	-0.0056***	-0.0056***
	(-4.23)	(-4.39)	(-334.61)	(-333.39)	(-322.87)	(-322.24)
N	2431	2431	2431	2431	2431	2431

Note: t statistics in parentheses, *p<0.05,** p<0.01,*** p<0.001

For the second board market employed a five factors model, only the MKT and CMA factors are positively correlated with the dependent variable. The t-values suggest that the MKT and RMW factors have strong explanatory power while SMB HML and CMA factors are inefficient. Thus, whether the five factors model can explain well is dubious. If employing the factors model in this market, all coefficients except intercept are positive, suggesting the positive relations between the three factors and the dependent variable. The t-values indicate that the three factors model can explain the second board market well under a 95% confidence level. Both models employed in this market have an R-squared of 0.976, suggesting that the two models can explain about 97.6% of the return data. Further, the MKT factor always has the strongest relations with the dependent variable in both models.

For the main board market employed five factors model, the coefficients except MHL's and constant are all positive, suggesting the relative factors are positively correlated with the dependent variable. Besides, the t-values suggest that the model is significant in explanation under the 95% confidence level, and the factor having the lowest explanation power is RMW since it can be significant only at a confidence level no larger than 95%. Employing the three factors model in the same market, the coefficients of MKT and SMB are positive, exhibiting the positive relations between these two factors and the dependent variable. Further, the t-values suggest that it explains this market well under a 99% confidence level. Both models in this market have an R-squared of 0.997, indicating that these two models can explain about 99.7% of the return data. Compared with the second board market, the higher R-squared suggests that these two models explain the return better in the main board market. Besides, the MKT factor always has the strongest relations with the dependent variable in both models.

For the mixed market employing the five-factors model, only the constant and HML factors are negatively correlated with the dependent variable. Furthermore, the t-values suggest that all factors except RMW are statistically significant under a 99% confidence level. This suggests that the overall model may be efficient in explaining this market if getting rid of the RMW alone. Changing to the three-factors model, the coefficient of MKT and SMB are positive, exhibiting the positive relations between these two factors Further, the t-values suggest that this model can explain the market well under a 99% confidence level. The R-squared of both models in this market is the same as that in the main board market (0.997), suggesting that the explanatory power of the two models in this market

is also strong. Again, the MKT factor always has the strongest relations with the dependent variable in both models.

3.3. Result analysis of STAR market

For the STAR market employing the five-factors model, only the HML factors are negatively correlated with the dependent variable (please see Table 4). Furthermore, the t-values suggest that all factors except HML are statistically significant under a 95% confidence level. This suggests that the overall model may be efficient in explaining this market if getting rid of the HML alone. The factors having lowest explanatory power among them is SMB, which is significant only at the confidence level not higher than 95%. Changing to the three-factors model, the coefficient of MKT and SMB are positive, exhibiting the positive relations between these two factors. Further, the t-values suggest that this model can't explain the market well since the SMB and HML are not significant under 95% confidence level. The R-squared of both models in this market is very close to each other (around 0.93). Again, the MKT factor always has the strongest relations with the dependent variable in both models.

Table 4. The STAR market

	(1)	(2)
	5-factor model	3-factor model
MKT	0.937***	0.941***
	(-91.1)	(89.79)
SMB	0.034*	0.0293
	(-2.1)	(1.77)
HML	-0.054	-0.0284
	(-1.89)	(-0.98)
RMW	0.161***	
	-5.24	
CMA	0.072***	
	(-3.48)	
R-sq	0.937	0.934
Constant	0.005***	-0.00497***
	(-30.63)	(-30.03)
N	663	663

Note: t statistics in parentheses, *p<0.05,** p<0.01,*** p<0.001

3.4. Potential explanation

From the model's perspective, the Fama-French five factors model's explanation power is strong in the main board market and is overall meaningful in the mixed market and STAR market. One reason for the insignificant of employing the Fama-French five factors model in the second board market might be because of the lower efficiency of the pricing mechanism in that market since the stocks in the second board market face less monitoring, and fewer restrictions on information disclosure. The investors in the second board market may be more irrational, which furtherstimulates this trend.

On the contrary, the Fama-French three factors model's explanation power is strong in all markets under a 95% confidence level expect for the STAR market. Compared with the five-factor model, the reason why it can explain the second board market and the mixed market might be because it gets rid of the factors that are weak in the explanation in the five-factor model and thus enhances the overall model significance. The reason about why the three-factor model cannot explain the STAR market while the five factor model can may because the companies on the STAR market are unique and their

stock performance more related to the CMA and RMW factors rather than SMB and HML factors, which can traditionally be used to explain the stock performance on the main board and second board market.

From the factors perspective, it can be found that the coefficient of MKT is always positive and the largest, and that of SMB is always positive (except the five-factor model in the second board market), while HML is always negative(except three factors model in the second board market) in all situations.

The character of MKT suggests that the market risk premium is always the most significant contributor to the portfolio excess return, compared with the other four factors, indicating that portfolio performance is mostly positively influenced by the overall market excess return. The SMB's character also suggests that the differences between the yields of small-cap portfolios and large-cap portfolios are always positively related to the portfolio's excess return. The HML character suggests that the differences in the yields between the high book-to-market ratio portfolio and the low book-to-market ratio portfolio are always negatively related to the portfolio's excess return. The reason might be the preference of the investors' group toward the low-quality stocks in the Chinese stock market.

4. Conclusion

Many previous researches have explored the efficiency of different stock return models in different markets in China. Some of them compare different models in a single market or industry to examine the differences in explanatory power and find the best model among them. Some of them test single models in different segments of the Chinese stock market and find which markets they can explain best.

The main purpose of this essay is to explore the efficiency of the Fama-French three and five factors models in the combination market of the main board and second board, and further test the efficiencies in the STAR market. Generally speaking, the three-factor model is statistically efficient in all different groups of markets except for STAR market, while the five-factor model is inefficient in the second board market. Moreover, it has been found that the MKT contributes to the portfolio excess return most all the time, and factors of SMB and HML always exhibit positive and negative relations with portfolio excess return respectively.

There are two limitations of this essay. The first point is that the momentum factor is not included when building the regression model. If the momentum factor is included, the stock markets' return can be better decomposed into more parts, which may also enhance the overall efficiency of the models. The second point is that the market focused in this article and the time horizon is limited due to the accessibility of data. Further research can include the momentous factor and expand to other Chinese stock markets, like B shares, Beizheng's A shares, and middle and small capital stocks with longer time horizons.

References

[1] Markowits H M. Portfolio selection[J]. Journal of finance, 1952, 7(1): 71-91.

[2] Lintner, J. Security prices, risk, and maximal gains from cross. The arbitrage theory of capital asset pricing [J]. Journal of Economic Theory, 1976, 13(3): 341-360.

[3] Ross. The arbitrage theory of capital asset pricing [J]. Journal of Economic Theory, 1976, 13(3), 341-360.

[4] Fama E F, French K R. Common risk factors in the returns on stocks and bonds[J]. Journal of financial economics, 1993, 33(1): 3-56.

[5] Zhao, S., Yan, H., & Zhang, K. Is the Fama-French five-factor model superior to the three-factor model? Empirical evidence from the Chinese A-share market [J]. Nankai Economic Research, 2016, (2): 41-59.

[6] Liu, L., & Yao, S. Asset growth effect and market efficiency: An empirical analysis based on risk factors [J]. Journal of Chongqing University (Social Science Edition), 2016, (5): 34-42.

[7] Wu, M. Empirical study of the Fama-French five-factor model in the Chinese A-share market. Doctoral dissertation, Jilin University, 2016, 1-48.

[8] Ma, Y. Empirical study of the Fama-French five-factor model in the valuation of internet listed companies. Master's thesis, Capital University of Economiscs and Business, 2018, 1-45.

[9] Yin, Y. (2016). Empirical study on asset pricing models in the Chinese A-share market. Doctoral dissertation, Dongbei University of Finance and Economics, 2016, 1-180.

[10] Li, Z., Yang, G., Feng, Y., & Jing, L. Empirical test of the Fama-French five-factor model in the Chinese stock market [J]. Financial Research, 2017, (6): 191-206.