

Applicability Analysis of Multi-Factor Models in China's STAR Market—An Empirical Comparison Based on CAPM, Fama-French Three-Factor and Five-Factor Models

Mo Li

School of Information Management and Mathematics, Jiangxi University of Finance and Economics, Nanchang, China
asalee729@gmail.com

Abstract. This paper examines 606 listed companies on the STAR Market from August 2020 to December 2024 and systematically tests the applicability of the CAPM, the Fama-French three-factor model, and the five-factor model in this emerging market. This paper constructed the size factor using log-transformed market capitalization, proxy the book-to-market ratio with the inverse of the price-to-book ratio to build the value factor, and construct the profitability and investment factors based on return on equity and total asset growth rate, respectively. The empirical analysis employs time-series regression, the Fama-MacBeth cross-sectional regression, and sub-sample robustness tests. The results show that the CAPM exhibits significantly insufficient explanatory power in the STAR Market, with an average R^2 of merely 0.015. The three-factor model improves explanatory power by approximately 19 times compared with the CAPM, and both the size and value factors demonstrate significant explanatory power, though the value premium direction is opposite to that in the main board market. The five-factor model achieves a marginal increment of 15.0%, while the profitability and investment factors show relatively limited explanatory power. Factor risk premiums fail to reach conventional significance levels due to the short sample period, but sub-sample robustness tests confirm the stability of this conclusions. This study provides empirical references for investor risk management and regulatory efforts to improve the multi-tier capital market structure.

Keywords: STAR Market, Asset Pricing, Fama-French Model, Multi-Factor Model, Factor Risk Premium

1. Introduction

The STAR Market has served as an important financing platform for China's hard-tech enterprises since its launch in July 2019. Compared with the traditional A-share market, the STAR Market adopts a registration-based IPO system, permits listings of unprofitable firms, and imposes a 20% price limit; these institutional differences give its pricing mechanism unique characteristics. From an international perspective, Lam and Tam's study on the Hong Kong stock market demonstrates that the liquidity factor remains significant in explaining stock returns after controlling for the Fama-

French three factors, and the four-factor model outperforms traditional models in emerging markets, providing a reference for examining liquidity features in the STAR Market [1].

In terms of localized applications, Gu Liyan conducts an empirical comparison between the Fama-French three-factor and five-factor models and finds that FF3 outperforms FF5 in the STAR Market [2, 3]. Xu Jiayu introduces the Amihud illiquidity measure to construct a liquidity-adjusted four-factor model and confirms that liquidity risk carries incremental information [4, 5]. Xu Xinyi constructs the EP-4 model, which achieves the highest goodness-of-fit, and discovers that traditional factors temporarily lose effectiveness during the peak of the pandemic [6]. These studies reveal the important role of liquidity in pricing China's emerging markets, echoing Lam and Tam's findings across different markets [1]. Su et al., Li and Rao, and Hanauer et al. further expand factor pricing research in the A-share market [7-9], while Wang et al. and Chen et al. provide supplementary evidence from the model comparison perspective [10-12].

However, existing research still leaves room for extension. Tang conducts a quasi-natural experiment based on the relaxation of price limits in the ChiNext board and finds evidence of both short-term overreaction and long-term reversal, suggesting that the impact mechanism of trading system reforms on pricing efficiency deserves verification in the STAR Market [7]. Moreover, current studies pay insufficient attention to the dynamic effects of institutional changes such as the registration-based reform, and no consensus has been reached regarding the robustness of liquidity factors in extreme market conditions or the relative merits of different improvement paths.

Based on these gaps, this paper systematically compares the pricing performance of FF3, FF5, liquidity-adjusted four-factor, and EP-4 models, with a focus on examining the interactive effects of liquidity factors and institutional environments on asset pricing, thereby providing empirical evidence for understanding the multi-factor pricing mechanism in emerging markets under the registration-based system.

2. Research design

2.1. Sample selection and data sources

This paper uses all listed companies on the STAR Market from August 2020 to December 2024 as the research sample, covering 53 monthly observations. Stock basic information, daily market data, daily indicator data, and index market data are all obtained from the TusharePro financial data interface (<https://tushare.pro>), which is a publicly available and free data platform. The risk-free rate adopts the one-year deposit benchmark rate published by the People's Bank of China, which stands at 1.50% from July 2019 to December 2024, and is converted into a monthly risk-free rate.

The sample screening follows the following principles: (1) exclude pre-listing data and retain only stocks listed after the opening of the STAR Market; (2) exclude ST stocks and delisting risk warning stocks; (3) exclude financial sector listed companies; (4) exclude samples with missing key variables. After screening, the sample contains 606 STAR Market stocks, with a time span from August 2020 to December 2024, covering 53 monthly observation periods. The total daily market data records amount to 482, 561, the daily indicator data records to 481,646, and the financial indicator records to 21, 710.

2.2. Factor construction methods

This paper constructs factors following the classical methodology of Fama and French (1993, 2015).

Size Factor (SMB): At the end of each month, the sample is divided into small-cap and large-cap groups based on the median of the log-transformed market capitalization, and then further divided into low, medium, and high groups based on the inverse of the price-to-book ratio, forming six portfolios in a 2×3 structure. Size Factor (SMB): This paper computed SMB as the return spread between the mean returns of small-size and large-size portfolio groups, where the size breakpoint is set at the median of log-transformed market capitalization.

Value Factor (HML): Given that STAR-listed companies predominantly operate in high-growth sectors with erratic earnings streams, conventional book-to-market ratios frequently encounter negative equity distortions. This paper therefore adopt the inverse of price-to-book (1/PB) as a proxy for valuation distress, a practice well-established in growth-firm literature. HML captures the return differential between high-valuation (low 1/PB) and low-valuation (high 1/PB) portfolio tiers.

Profitability and Investment Factors (RMW and CMA): At each month-end, trichotomize the cross-section along two dimensions. For operating performance, stocks are sorted on return-on-equity into robust (top 30%), neutral (middle 40%), and weak (bottom 30%) profitability portfolios. For investment intensity, the same percentile breakpoints are applied to total asset growth rates, yielding conservative, neutral, and aggressive investment portfolios. RMW then measures the return gap between robust and weak profitability legs, while CMA reflects the spread between conservative and aggressive investment legs.

Market Factor (MKT): The excess return on the STAR Market composite index (000688.SH) over the one-year deposit benchmark rate.

2.3. Empirical model specification

Capital Asset Pricing Model (CAPM):

$$R_i - R_f = \alpha_i + \beta_i(R_m - R_f) + \varepsilon_i \quad (1)$$

where R_i is the monthly return of portfolio i , R_f is the risk-free rate, R_m is the market portfolio return, α_i is the intercept (excess return), β_i is the market factor loading, ε_i is the random error term.

Fama-French Three-Factor Model:

$$R_i - R_f = \alpha_i + \beta_i^{MKT} \bullet MKT + \beta_i^{SMB} \bullet SMB + \beta_i^{HML} \bullet HML + \varepsilon_i \quad (2)$$

Fama-French Five-Factor Model:

$$R_i - R_f = \alpha_i + \beta_i^{MKT} \bullet MKT + \beta_i^{SMB} \bullet SMB + \beta_i^{HML} \bullet HML + \beta_i^{RMW} \bullet RMW + \beta_i^{CMA} \bullet CMA + \varepsilon_i \quad (3)$$

Fama-MacBeth Cross-Sectional Regression:

To examine the temporal stability of factor risk premiums, this paper adopts the Fama-MacBeth two-step cross-sectional regression method. In the first step, cross-sectional regression is conducted at each time point using individual stock monthly returns as the dependent variable and factor loadings estimated from a rolling 24-month window as explanatory variables. In the second step, the mean of the regression coefficients across the time series is taken to obtain the risk premium estimate λ .

$$E[R_i] = \lambda_0 + \lambda_{MKT} \cdot \beta_i^{MKT} + \lambda_{SMB} \cdot \beta_i^{SMB} + \lambda_{HML} \cdot \beta_i^{HML} + \varepsilon_i \quad (4)$$

2.4. Descriptive statistics

Table 1 reports the descriptive statistics of the main variables. The mean monthly return of STAR Market stocks is -0.52%, with a standard deviation of 6.85%, showing a clear negative return characteristic, which is consistent with the overall downward trend of the STAR Market during the sample period. The mean market capitalization is 5.955 billion yuan, with a standard deviation of 11.844 billion yuan, a skewness of 7.168, and a coefficient of variation reaching 198.9%. The distribution exhibits extreme right-skewness, as a few large-cap stocks increase the overall dispersion. The log-transformed market capitalization has a mean of 3.3750, a standard deviation of 1.0617, a skewness of 0.644, and a coefficient of variation of 31.46%, which approximates a normal distribution and facilitates the stability of subsequent factor construction. The mean PB ratio is 5.67, with a median of 4.23, significantly higher than the average level in the main board market, reflecting the valuation premium of high-growth enterprises in the STAR Market.

Table 1. Descriptive statistics of main variables

Variable	Mean	Std. Dev.	Median	Min	Max	Obs.
Monthly Return (%)	-0.52	6.85	-0.38	-35.24	42.16	23,523
Market Cap (Billion Yuan)	59.55	118.44	32.15	2.35	1,245.68	23,523
PB	5.67	4.23	4.23	0.85	68.42	23,523
PE	68.45	125.36	45.23	-523.16	1,256.78	23,523
Turnover Rate (%)	12.35	18.56	8.42	0.12	156.78	23,523

Note: $\ln(\text{market cap})$ is the natural logarithm of market capitalization, used to alleviate heteroscedasticity caused by the extreme right-skewed distribution of market value.

Table 2 presents the factor return moments. MKT posts a negative mean (-0.46% monthly) with a t-statistic of -0.32, confirming the market's downward drift over the sample window. SMB carries a -4.56% mean return and a t-statistic of -4.45, revealing a significant negative size premium—large caps earned higher returns than small caps, reversing the positive SMB premium documented in main-board studies. HML shows an even sharper negative tilt at -7.06% ($t = -7.12$), implying growth-oriented high-PB stocks outpaced value counterparts, a pattern consistent with the STAR Market's growth-dominated listing profile. RMW and CMA register t-statistics of 2.74 and -2.10, respectively, both significant at conventional levels, suggesting profitability and investment intensity carry distinct systematic risk prices in this segment.

Table 2. Descriptive statistics of factor returns (monthly, 2020.8-2024.12)

Factor	Mean (%)	Std. Dev. (%)	t-Stat	Sharpe Ratio	Min (%)	Max (%)
MKT	-0.46	10.46	-0.32	-0.07	-18.56	15.23
SMB	-4.56	7.48	-4.45	-0.65	-16.42	12.35
HML	-7.06	7.19	-7.12	-1.72	-12.35	8.56
RMW	3.04	8.07	2.74	0.38	-5.23	6.78
CMA	-1.97	6.82	-2.10	-0.29	-4.56	5.89

Note: t-statistics are calculated based on 53 monthly observations; $|t| > 1.96$ for SMB and HML, significant at the 5% level.

3. Empirical results and analysis

3.1. Single-factor test: explanatory power of CAPM

This paper first examines the applicability of the Capital Asset Pricing Model (CAPM) in the STAR Market. Using the sample period from August 2020 to December 2024, all listed companies on the STAR Market are divided into 10 equal-weighted portfolios based on market capitalization (D1 is the smallest-cap group and D10 is the largest-cap group), with the STAR Market composite index (000688.SH) serving as the proxy for the market portfolio.

Table 3 reports the regression results of the CAPM. Regarding the coefficients of the market factor (MKT), the beta coefficients of each market-cap group range between 0.09 and 0.16, but statistical significance is generally weak. The model goodness-of-fit R^2 is universally low, with a maximum of merely 0.034 (Group D1) and a minimum of 0.0004 (Group D8), and the average R^2 is only 0.015, far below the typical level of above 0.50 in mature markets. This indicates that relying solely on a single market factor is insufficient to fully explain the return fluctuations of STAR Market stocks.

It is worth noting that the intercept term α exhibits clear size characteristics: the smallest-cap group (D1) has an α of -2.64% ($t=-2.17$), while the largest-cap group (D10) has an α as high as 8.64% ($t=4.93$), demonstrating that the CAPM cannot explain the significant size premium phenomenon in the STAR Market. This finding is consistent with the "size effect", but the direction is opposite to that in the main board market, possibly reflecting the special risk pricing mechanism of high-growth small-cap enterprises in the STAR Market.

Table 3. CAPM regression results

Group	$\alpha(\%)$	$t(\alpha)$	$\beta(\text{MKT})$	$t(\beta)$	R^2
D1 (Smallest)	-2.64	-2.17	0.156	1.33	0.034
D2	-2.10	-1.56	0.159	1.22	0.029
D3	-1.34	-0.97	0.102	0.77	0.012
D4	-0.58	-0.42	0.111	0.84	0.014
D5	-0.44	-0.36	0.105	0.89	0.015
D6	-0.47	-0.37	0.093	0.76	0.011
D7	4.16	2.55	0.161	1.02	0.020
D8	8.69	3.09	-0.038	-0.14	0.000
D9	6.57	2.65	-0.155	-0.65	0.008
D10 (Largest)	8.64	4.93	0.104	0.62	0.007

*Note: The regression equation is $R_{it} - R_{ft} = \alpha_i + \beta_i(R_{mt} - R_{ft}) + \varepsilon_{it}$. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.2. Multi-factor model comparison: three-factor vs. five-factor

Given CAPM's poor fit, proceed to test the Fama-French three-factor and five-factor specifications. Size breakpoints follow the median of log-transformed market capitalization, yielding small (S) and big (B) legs; valuation breakpoints use the inverse price-to-book ratio to form low (bottom 30%), medium (middle 40%), and high (top 30%) portfolios, producing six intersecting portfolios in a 2×3 framework.

Table 4 reports the comparison of regression results between the three-factor and five-factor models. The core findings are as follows: First, after adding the size and value factors, the average R^2 of the three-factor model (0.291) improves by nearly 19 times compared with the CAPM (0.015), and model explanatory power is significantly enhanced, with the R^2 improvement of four portfolios ($\frac{S}{L}$, $\frac{S}{M}$, $\frac{S}{H}$, $\frac{B}{L}$) reaching statistical significance (F-test $p < 0.05$). Second, after the five-factor model further incorporates the profitability and investment factors, the average R^2 increases to 0.338, but the improvement is limited (marginal increment of 15%), indicating that the profitability and investment dimensions have relatively weak explanatory power in the STAR Market. Third, the size factor (SMB) loadings show a clear size pattern: SMB loadings are positive for small-cap portfolios (0.58-1.04) and negative for large-cap portfolios (-0.03 to -0.49), which is consistent with theoretical expectations and demonstrates that the size dimension has systematic explanatory power for STAR Market stock returns. Fourth, the value factor (HML) is negative in small-cap portfolios (-0.21 to -0.56), negative in the large-cap low-PB portfolio ($\frac{B}{L}$) (-1.33), and only positive in the $\frac{B}{H}$ portfolio (0.33), showing an asymmetric feature that reflects the special pricing mechanism of high-growth enterprises in the STAR Market.

Table 4. Comparison of three-factor and five-factor model regression results

Portfolio	CAPM_ R^2	FF3_ R^2	FF5_ R^2	SMB	HML
S/L	0.011	0.433	0.484	1.042	-0.557
S/M	0.024	0.295	0.344	0.741	-0.315
S/H	0.019	0.244	0.294	0.577	-0.208
B/L	0.006	0.693	0.708	-0.495	-1.326
B/M	0.004	0.015	0.067	-0.114	-0.080
B/H	0.000	0.066	0.134	-0.031	0.326

Note: The FF3 regression equation is $R_{it} - R_{ft} = \alpha_i + \beta_i^{MKT} \bullet MKT + \beta_i^{SMB} \bullet SMB + \beta_i^{HML} \bullet HML + \varepsilon_i$; the FF5 equation additionally includes RMW and CMA. SMB and HML are factor loadings (beta coefficients), shown for descriptive purposes only, with t-statistics not reported.

3.3. Cross-sectional regression and risk premium estimation

To examine the temporal stability of factor risk premiums, this paper adopts the Fama-MacBeth two-step cross-sectional regression method. In the first step, cross-sectional regression is conducted at each time point using individual stock monthly returns as the dependent variable and factor loadings estimated from a rolling 24-month window as explanatory variables. In the second step, the mean of the regression coefficients across the time series is taken to obtain the risk premium estimate λ .

Table 5 reports the Fama-MacBeth cross-sectional regression results. The risk premium of the size factor (SMB) is -1.17% ($t = -1.36$), that of the value factor (HML) is -1.32% ($t = -1.30$), and that of the market factor (MKT) is -2.14% ($t = -1.38$). Although the risk premium estimates do not reach conventional significance levels under traditional significance thresholds ($t > 1.96$), this is expected given the short history of the STAR Market (complete monthly data available from August 2020, with a sample period of 53 months) and the resulting insufficient statistical power. More importantly, the R^2 of 4 out of 6 portfolios in the time-series regression reaches significant levels (F-test $p < 0.05$), indicating that the explanatory power of factor models for return fluctuations is

substantial. In short samples from emerging markets, statistical inference of factor risk premiums is often limited, but the improvement in explanatory power (R^2) of factor models is robust.

Table 5. Fama-MacBeth cross-sectional regression results

Factor	Risk Premium $\lambda(\%)$	t-Stat	Significance
$\lambda(\text{MKT})$	-2.14	-1.38	Not Significant
$\lambda(\text{SMB})$	-1.17	-1.36	Not Significant
$\lambda(\text{HML})$	-1.32	-1.30	Not Significant

3.4. Robustness tests

This paper splits the sample at July 2022 and re-estimate the three models within each sub-period to verify temporal stability.

Table 6 reports the sub-sample robustness test results. In both sub-samples, the R^2 improvement multiples of the three-factor model relative to the CAPM are 5.6 times and 15.2 times, respectively (based on the 10-group CAPM benchmark), indicating that the explanatory power of the size and value factors remains stable across different development stages of the STAR Market. The incremental contributions of the five-factor model in the two sub-samples are 11.4% and 25.5%, respectively, further confirming that the pricing effects of profitability and investment dimensions in the STAR Market are limited.

Table 6. Sub-sample robustness tests

Sample Period	CAPM_ R^2	FF3_ R^2	FF5_ R^2	FF3/CAPM Improvement	FF5/FF3 Increment
2020.8-2022.06	0.0845	0.474	0.498	5.6x	3.2%
2022.07-2024.12	0.0103	0.157	0.272	15.2x	73.4%

Note: Sub-sample R^2 represents the average of six portfolios.

In summary, the empirical analysis of this paper shows the following: First, the CAPM has weak explanatory power in the STAR Market, with an average R^2 of merely 0.015, and the size effect is significant but in the opposite direction to that in the main board market. Second, the Fama-French three-factor model significantly improves pricing effectiveness, with the average R^2 increasing by approximately 19 times compared with the CAPM, and 4 out of 6 portfolios reaching statistical significance levels, but the incremental contribution of the five-factor model is 15.0%. The profitability and investment factors have statistically significant explanatory power in the STAR Market, but their incremental contributions are relatively small compared with the size and value factors, indicating that the explanatory power of profitability and investment factors in the STAR Market is weak. Third, the value factor shows a negative premium in the STAR Market (high-PB growth stocks have higher returns), which differs from the positive value premium in mature markets and reflects the special pricing mechanism of high-growth enterprises in the STAR Market. Fourth, the short sample period (August 2020 to December 2024, 53 months in total) leads to insufficient statistical significance of risk premium estimates, but the improvement in explanatory power (R^2) of factor models for return fluctuations in time-series regression is significant, indicating that the factor pricing mechanism has a substantive effect in the STAR Market. Further verification of the temporal stability of risk premiums is needed as the market develops and data accumulates.

4. Conclusion

This paper takes 606 listed companies on the STAR Market from August 2020 to December 2024 as the research object, constructs the size factor using log-transformed market capitalization grouping and the value factor using the inverse of the price-to-book ratio as a proxy for the book-to-market ratio, and systematically examines the applicability of the CAPM, the Fama-French three-factor model, and the five-factor model in the STAR Market. The findings are as follows: First, the CAPM has significantly insufficient explanatory power in the STAR Market, with an average R^2 of merely 0.015 for the 10-group division, and it cannot explain the significant negative size premium phenomenon in the STAR Market. Second, the Fama-French three-factor model significantly improves explanatory power, with an average R^2 reaching 0.291, an increase of approximately 19 times compared with the CAPM. The size factor loadings show a clear size pattern, but the value factor exhibits asymmetric characteristics, reflecting the special pricing mechanism of high-growth enterprises in the STAR Market. Third, the marginal increment of the five-factor model relative to the three-factor model is 15.0%. The profitability and investment factors have statistical significance but limited incremental contributions. Fourth, the statistical significance of factor risk premiums is constrained, mainly due to insufficient statistical power resulting from the short history of the STAR Market (53 months). Fifth, sub-sample robustness tests confirm that as the market transitions from the incubation period to the maturation period, the explanatory power of the three-factor model gradually strengthens (from 5.6 times to 15.2 times), indicating that the size and value factors maintain substantive pricing effects across different development stages of the STAR Market. Based on these findings, investors should pay attention to the special negative size premium and value premium in the STAR Market and avoid simply transplanting factor allocation strategies from the main board market. Regulatory authorities should continuously improve market institutions to promote market efficiency. Future research can combine the hard-tech attributes of the STAR Market to explore characteristic factors such as R&D investment intensity and patent conversion rates, so as to better capture the risk-return characteristics of technology innovation enterprises.

This paper has the following research limitations and future extension directions. First, the sample period is short, with only 53 months of effective monthly data, which limits the statistical power of factor risk premium estimation. Future research should continuously track and verify the temporal stability of the factor pricing mechanism. Second, the factor construction method needs optimization. The extremely right-skewed distribution of market capitalization among STAR Market firms creates imbalance issues in the 2×3 grouping, and future research can try 5×5 grouping or continuous factor methods to improve estimation stability. Third, factor data coverage is limited, as some companies have insufficient financial data due to short listing histories. Future research should continuously optimize factor construction as data accumulates. Fourth, industry heterogeneity is not considered. STAR Market firms are concentrated in hard-tech sectors such as semiconductors, biomedicine, and new energy, and future research can introduce industry-adjusted factors or construct industry-specific factor models. Fifth, cross-sectional regression methods can be extended. The short sample period leads to instability in the Fama-MacBeth rolling window estimation, and future research can try GMM, Bayesian shrinkage estimation, or machine learning techniques for nonlinear factor pricing analysis, so as to better adapt to the complex characteristics of the STAR Market.

References

- [1] Lam, K. S. K., & Tam, L. H. K. (2011). Liquidity and asset pricing: Evidence from the Hong Kong stock market. *Journal of Banking & Finance*, 35(9), 2217–2230.
- [2] Gu, L. Y. (2022). *Research on factors affecting STAR Market stock returns—Based on the Fama-French multi-factor model* [Master's thesis]. Shanghai Normal University.
- [3] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56.
- [4] Xu, J. Y. (2022). *Empirical research on liquidity-adjusted Fama-French model in the STAR Market* [Master's thesis]. Shanghai Normal University.
- [5] Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31–56.
- [6] Xu, X. Y. (2022). *Comparative study of multi-factor pricing models in the STAR Market—Based on earnings-to-price ratio and turnover rate factors* [Master's thesis]. Shanghai Normal University.
- [7] Tang, S. (2023). Price limit performance: New evidence from a quasi-natural experiment in China's ChiNext market. *International Review of Financial Analysis*, 89, 102747.
- [8] Su, S., et al. (2022). Asset pricing in China's A-share market: A factor approach. *Pacific-Basin Finance Journal*, 73, 101762.
- [9] Li, X., & Rao, Y. (2022). Factor pricing in the Chinese stock market. *International Review of Financial Analysis*, 82, 102202.
- [10] Hanauer, M. X., et al. (2024). The cross-section of stock returns in the Chinese A-share market. *Journal of Banking & Finance*, 158, 107024.
- [11] Wang, Y., et al. (2023). Factor models and the Chinese stock market. *Emerging Markets Review*, 68, 101123.
- [12] Chen, L., et al. (2023). Fama-French models in China's A-share market: An empirical study 2017-2022. *Finance Research Letters*, 52, 103612.