Proceedings of the 2nd International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/14/20230808

Examining the Q-Factor Model in the Chinese Stock Market

Zirui Feng[1,a,*]

[1]Jinan University, Jinan University-University of Birmingham Joint Institute, No.855 East Xingye Avenue, Guangzhou, PRC
a. fengzr@stu2020.jnu.edu.cn
**corresponding author*

Abstract: The aim of this paper is to compare model the q-factor model with the Fama-French three-factor using multiple regression techniques and tests. This is done by constructing the market, size, value, investment, and return-on-equity factors with features tailored for the Chinese stock market using 60-month pre-COVID Chinese stock market data. One specific feature tailored for the Chinese stock market, which is eliminating the bottom 30% size of stocks to avoid shell value contamination of returns, is explained. Differences and similarities between the Chinese and the U.S. stock market that can be observed in the process and the results of the regressions and tests are discussed in the paper. It can be concluded that both models behave well in the chosen period of the Chinese stock market, but the q-factor model works slightly better. Other explanatory explorations are done in the paper as well. Limitations of the models due to practical reasons, e.g., the inconsistent changing dynamic of the Chinese stock market institution, and potential further explorations, e.g., potential shell stock pricing strategies, are also discussed in the paper.

Keywords: asset pricing, Chinese stock market, q-factor model

1. Introduction

Asset pricing models have been a hot topic in financial research since their inception. Sharpe, Lintner, and Black introduced the capital asset pricing model (CAPM) to explain the relationship between asset returns and risks [1-3]. Fama and French expanded the CAPM to the Fama-French three-factor (FF3) model with a size factor based on the size of the firm and a value factor based on the book-to-market (B/M) ratio in 1992 [4]. They argued that value stocks with high B/M ratios had outperformed growth stocks with low B/M ratios and small-capitalization stocks had outperformed large-capitalization stocks. Thus, the value and size factors were also called the high-minus-low and small-minus-big factors. After that, many factor models based on the FF3 model have been raised by researchers, e.g., the four-factor model with a momentum factor expanded by Carhart [5]. In 2011, Chen, Novy-Marx, and Zhang constructed another model with the original market factor, the ROE factor based on the return-on-equity (ROE) ratio, and the investment factor based on the investment-to-assets (I/A) ratio [6]. They argued that this model has a more solid economic foundation and called it the q-factor model. Later in 2015, Fama and French constructed a five-factor model with an additional profitability and investment factor added to their original three-factor model [7].

On the other hand, China has the second largest stock market with a 12 trillion market capitalization of listed domestic companies and has long been the focus of researchers. Many pieces

DOI: 10.54254/2754-1169/14/20230808

of literature have examined asset pricing models with empirical data from the stock market in China. For example, Liu, Stambaugh, and Yuan established a modified version of the FF3 model, where they constructed the value factor based on the earning-price ratio instead of the B/M ratio [8]. They also examined other important Chinese stock market characteristics such as the impact of the reverse merger phenomenon on the value of small stocks and the documented anomalies. They argued that due to the inefficiency of getting listed in China, reverse merger becomes a good choice for companies willing to get listed in China, and because of this, the market values of the potential firms for reverse merger are contaminated by the expected reverse merger return. They thus excluded firms the smallest 30% of firms to avoid the shell value contamination, since the volume of those stocks is small.

This paper examines the q-factor model using the Chinese stock market data by multiple regressions and tests, at the same time comparing the results with the traditional FF3 model. The q-factor model, constructed by Chen, Novy-Marx, and Zhang, consists of three factors, the market factor, an investment factor, and a ROE factor [6]. The factors are constructed based on market excess return, I/A ratio, and the ROE ratio. The market factor, denoted by MKT , is the market excess return. The I/A ratio is defined as the firm's annual change in property, plant, and equipment (PP&E) plus the firm's annual change in inventories divided by the lagged book value of assets. Thus, the investment factor, denoted by r_{INV}, is the difference between the return of the high and low I/A portfolios constructed. Moreover, the ROE factor, denoted by r_{ROE} , is the difference between the return of the high and low ROE portfolios. Chen, Novy- Marx, and Zhang formally stated the model as formula (1).

$$E\left[r^i\right] - r^f = \beta^i_{MKT} E[MKT] + \beta^i_{ROE} E[r_{ROE}] + \beta^i_{INV} E[r_{INV}] \quad (1)$$

where β^i_{MKT}, β^i_{ROE} and β^i_{INV} are the factor loadings, r^f is the risk-free interest rate, and $E[MKT]$, $E[r_{ROE}]$ and $E[r_{INV}]$ are the expected risk premiums. Section 2 explains the methodologies. Section 3 reports the results from implementing the methodologies introduced in the previous section. Section 4 discusses the limitations of the models. Finally, Section 5 concludes.

2. Methodology

Many classical methods in asset pricing research are used in this paper, including the Fama-MacBeth two-step regression, Newey and West procedure of adjustment to the standard error for heteroskedasticity and autocorrelation, and the Gibbons, Ross, and Shanken (GRS) F test for portfolio efficiency [9-12]. Some specific features are tailored for the stock market in China. The construction of the factor portfolios is similar to the construction that Chen, Novy-Marx, and Zhang did in their paper [6].

2.1. Data Source and Processing

The accounting data for the stocks are from the WIND database. The one-year Chinese government bond yield data are from the China Stock Market & Accounting Research Database (CSMAR), accessed via Experimental Education Center of Economics and Management at Jinan University. The duration of the data is 60 months, from January 2015 through December 2019. Data after 2019 are excluded due to the COVID pricing phenomenon. Similar to what Liu, Stambaugh, and Yuan did, the smallest 30% of stocks are eliminated first to avoid shell value contamination of returns on asset pricing [8]. Robust analysis of this number is also conducted by performing the same procedures eliminating the smallest 25% and 35%. Furthermore, firms that are half-year (including the sixth month) after IPO are excluded, in addition to financial firms.

2.2. Metrics Construction

The main metrics used in this paper are the market capitalization of the stocks, B/M ratio of the firm, returns of the stocks, one-year government bond yield, ROE ratio, and I/A ratio. The construction of the first two metrics follows the definition in the Fama and French 1992 paper, and the construction of the last two metrics follows the definition in the Chen, Novy-Marx, and Zhang 2011 paper. These are illustrated by formulas (2), (3), and (4) [4, 6].

$$B/M = \frac{Shareholders'Equity}{Market\ Capitalization} \tag{2}$$

$$I/A = \frac{\Delta(PP\&E)+\Delta(Inventories)}{Book\ Value\ of\ Assets} \tag{3}$$

$$ROE = \frac{Income\ Before\ Extraordinary\ Items}{Book\ Equity} \tag{4}$$

2.3. Portfolio and Factors Construction

Three factors, which are the market factor, investment factor (low-minus-high I/A), and ROE (high-minus-low ROE) factor, are used to construct the q factors. The market factor is the market excess return compared to the risk-free interest rate, that is the one-year Chinese government bond yield. The other two factors are constructed by dividing portfolios. Twenty-seven (HSP, HSD, HSZ, HBP, HBD, HBZ, HNP, HND, HNZ, MSP, MSD, MSZ, MBP, MBD, MBZ, MNP, MND, MNZ, LSP, LSD, LSZ, LBP, LBD, LBZ, LNP, LND, LNZ) value-weighted portfolios for the q-factor model are constructed according to the ranked I/A, ROE, and market capitalization. There are three I/A groups (H, M, and L), divided on an annual basis and three ROE and size (market capitalization) groups (S, M, B, and P, D, Z) respectively divided monthly. The breakpoints for all groups are low at 30%, medium at 40%, and high at 30%. The factors are constructed with the simple average returns according to the following formulas (5) and (6) and are rebalanced monthly.

$$r_{INV} = \frac{1}{9}(LSP + LSD + LSZ + LBP + LBD + LBZ + LNP + LND + LNZ) - \frac{1}{9}(HSP + HSD + HSZ + HBP + HBD + HBZ + HNP + HND + HNZ) \tag{5}$$

$$r_{ROE} = \frac{1}{9}(HSP + HBP + HNP + LSP + LBP + LNP + MSP + MBP + MNP) - \frac{1}{9}(LSZ + LBZ + LNZ + HSZ + HBZ + HNZ + MSZ + MBZ + MNZ) \tag{6}$$

On the other hand, besides the q-factor model, a traditional FF3 model is also constructed for comparison and explanation. Besides the market factor, two factors (small-minus-big size and value-minus-growth) are constructed in this model by constructing portfolios. Six value-weighted portfolios (SH, SM, SL, BH, BM, BL) are constructed based on market capitalization and B/M ratio. Two groups (B and S) are divided according to their size and three groups (H, M, and L) by B/M ratio. The construction is illustrated by formulas (7) and (8).

$$SMB = \frac{1}{3}(SH + SM + SL) - \frac{1}{3}(BH + BM + BL) \tag{7}$$

$$HML = \frac{1}{2}(SH + BH) - \frac{1}{2}(SL + BL) \tag{8}$$

2.4. Regression Models and Tests

Regressions and tests are used to examine the models in this paper. One important regression is the Fama-MacBeth regression. Fama and MacBeth introduced the Fama-MacBeth two-step regression

DOI: 10.54254/2754-1169/14/20230808

for regressing panel data, which involves first the time-series regression and then the cross-sectional regression [9]. The regression model is specified in formula (9).

$$r_t^i - r_t^f = \alpha_t^i + \beta_{MKT} MKT_t + \beta_{INV} r_{INV,t} + \beta_{ROE} r_{ROE,t} + \epsilon_t^i \quad (9)$$

Newey and West adjusted standard errors for heteroskedasticity and autocorrelations are also used in this paper to adjust standard errors and the regression t-statistics for the loadings of the factors [11, 12]. Average adjusted R-square and t-statistics are used to evaluate the explanatory power and the mispricing of the model. Large values imply more excess return explained by the regression. T-statistics, on the other hand, measures how certain one can be to conclude that the variable is different from zero.

Moreover, Gibbons, Ross, and Shanken introduced the GRS test, as an F test, to test the mispricing of the factors [10]. The null hypothesis is that in the model, $\alpha = 0$, which means there is no mispricing in the model. This test is used in this paper to test and compare the q-factor model and the FF3 model. Besides, robust analysis of the percentage of stocks eliminated to avoid shell value contamination of returns is performed by eliminating the smallest 25% and 35% of stocks in the market and then performing the same regressions and tests.

3. Results

This section reports the main result by performing the methods introduced in section 2. The robustness test for the cutoff for the shell-value contamination of returns avoidance shows that the results are robust. It can be concluded in this section that the stock markets in China and the U.S. have many differences and similarities, and the q-factor model behaves better in the chosen period of the stock market in China

3.1. Summary Statistics

The Table 1 reports the summary statistics of the factors, and the Table 2 shows the correlation matrix for the factors resulting from performing the construction presented in subsection 2.3. The unit for the mean value and standard deviation in the table is percentage.

Table 1: Summary statistics of the traditional factors and q factors.

Factor	Mean	Std. Dev.	t-stat.
r_{INV}	-0.394	5.365	0.7461
r_{ROE}	0.675	6.393	0.6422
MKT	-1.773	9.410	0.1531
SMB	-0.157	10.918	0.9152
HML	2.586	13.889	0.1580

Table 2: Correlation coefficient matrix of the factors.

	r_{INV}	r_{ROE}	MKT	SMB	HML
r_{INV}	1.0000	-0.3951	-0.7132	0.3813	-0.0160
r_{ROE}	-0.3951	1.0000	0.4707	-0.2706	-0.2172
MKT	-0.7132	0.4707	1.0000	-0.0298	0.1353
SMB	0.3813	-0.2706	-0.0298	1.0000	-0.2785
HML	-0.0160	-0.2172	0.1353	-0.2785	1.0000

During the 60-month period, the mean value of the investment factor is -0.39% and 0.68% for the ROE factor. The mean value of the investment factor is different from the results in the paper of Chen, Novy-Marx, and Zhang, representing that the high investment stocks perform better than the low investment stocks in China [6]. The standard deviations are 5.37% and 6.39% respectively, with t-statistics of 0.75 and 0.64. Since the t-statistics of the market factor, size factor, and value factor are 0.15, 0.91, and 0.16 respectively, the combination of the investment and ROE premiums is more substantial in the chosen period of the Chinese market than the combination of the size and value factor, showing stronger q-factor effect. However, the premiums are not strong as the difference between the factors and zero is not statistically significant.

The Chinese market is different from the U.S. market in other aspects. The mean value of the size factor is negative, which is not the same as the one in the U.S. [4]. It can be deduced that the big stocks in China perform better than the small stocks in China. However, the gap is not big. One reason related to the process of portfolio construction is that the smallest 30% of stocks are eliminated, so the real small stocks are not considered. From the correlation matrix reported in the Table 2, it can also be deduced that the correlations of the traditional factors and the q factors are not strong, except that the market factor has a strong correlation with the investment and ROE factor, though the factors still have stronger correlations in China than in the U.S. The difference in signs of the correlation coefficient between the results in the original paper and the results in this paper exhibit a big difference in internal correlations between the U.S. and the Chinese stock market (only three in ten pairs of correlations are with the same sign) [6].

3.2. Fama-MacBeth Two-Step Regressions

The Table 3 reports the regression coefficients of the Fama-MacBeth two-step procedure with multiple regressors, with t-statistics in the parentheses adjusted for heteroskedasticity and autocorrelations based on the Newey and West standard errors with two lags [9, 11, 12]. The included regressors in the regression are indicated by the nonempty rows in the Table 3. Regressors in the Table 3 include the CAPM β, the log of B/M ratio (log BM), the log of end-of-month market capitalization (log MC), investment-to-assets ratio (I/A), and ROE ratio (ROE). The α row in the Table 3 reports the results of the intercept term of the regression. The last row (R^2 row) of the Table 3 reports the average adjusted R-square for the regression. The result of the robust test for the 25% and 35% cutoffs shows that the result is robust.

From the above regression results, it can be concluded that both models work well in the chosen data of the stock market in China. The FF3 model has a larger average adjusted R-square in regression than the q-factor model, but the intercept term of the q-factor model is more statistically significantly indistinguishable from zero (t-statistics of -0.45 compared to -3.62). This shows that the q factors have more explanatory power than the Fama-French factors. However, both models have small average adjusted R-squares, leaving large unexplained variations of the excess return. Though both do not significantly misprice, there exists a big portion unexplained by both.

On the other hand, in columns (1), (2), (4), (6), (7), and (8), the CAPM β does not significantly enter. One essential observation is that the market factor can be well explained by other factors, since, as reported in the last column of the Table 3, even if the CAPM β is excluded from the regression, though the average adjusted R-square is smaller than other results, the t-statistics of the intercept term, -0.21, is still small. Besides the market factor, other metrics selected for regression can also explain each other in some way because their t-statistics are big only if some metrics are excluded from the regression. Furthermore, as reported in the second last row, the ROE ratio outperforms other metrics, and this confirms a significant ROE effect in China.

What can be concluded in this section is that though both pricing models behave well, the q factors have more explanatory power, and the ROE effect is relatively more significant in China.

DOI: 10.54254/2754-1169/14/20230808

Table 3: Results of the Fama-MacBeth regressions.

Quantity	(1)	(2)	(3)	(4)
α	-0.0055	-0.0237	-0.0003	0.0046
	(-0.45)	(-3.62)	(-1.59)	(-0.36)
β	0.0005	0.0036	0.0127	-0.0002
	(0.09)	(0.31)	(1.37)	(-0.03)
$\log BM$		- 0.0139		-0.0046
		(-1.00)		(-0.83)
$\log MC$		0.0126		0.0132
		(0.93)		(0.73)
I/A	0.0002		-0.0003	0.0001
	(1.29)		(-1.59)	(0.62)
ROE	0.0876			
	(1.45)			
R^2	0.0334	0.0510	0.0460	0.0598
	(5)	(6)	(7)	(8)
α	-0.0215	-0.0105	-0.0051	-0.0036
	(-2.63)	(-0.93)	(-0.37)	(-0.21)
β	0.0155	0.0057	0.0003	
	(1.54)	(0.55)	(0.04)	
$\log BM$		-0.0131	-0.0131	-0.0117
		(-0.59)	(-0.85)	(-0.76)
$\log MC$		0.0098	0.0110	0.0117
		(0.49)	(0.79)	(0.71)
I/A			0.0001	-0.0002
			(0.57)	(-2.43)
ROE	0.0645	0.0130	-0.0538	-0.0239
	(1.43)	(0.39)	(-0.93)	(-0.67)
R^2	0.0651	0.0814	0.0485	0.0174

3.3. The Explanation for Each Other

The Table 4 reports the regression coefficients resulting from regressing the investment and ROE factors with the traditional Fama-French factors, with t-statistics in the parentheses adjusted for heteroskedasticity and autocorrelations based on the Newey and West standard errors with two lags [11,12]. The included regressors in the regression are indicated by the nonempty rows in the table. The last column reports the average adjusted R-square for the regression. The intercept term in the regression is represented by α. The result of the robust test for the 25% and 35% cutoffs shows that the result is robust.

Table 4: Results of the regressions of the q-factors and the traditional factors.

Dep. Var.	α	$β_{MKT}$	$β_{SMB}$	$β_{HML}$	R^2
r_{INV}	-0.0029 (-0.24)	-0.3411 (-4.60)			0.509
	-0.0018 (-0.25)	-0.3476 (-8.62)	0.1844 (4.62)	0.0781 (1.99)	0.700
r_{ROE}	-0.0002 (-0.01)	0.3441 (2.66)			0.222
	0.0023 (0.17)	0.3634 (2.91)	-0.2721 (-1.60)	-0.2238 (-2.57)	0.460

The regression results illustrate strong relationships between the market factor and the investment and the ROE factor, which confirms the conclusion from the correlation analysis and the Fama-MacBeth regressions. Since the average adjusted R-square of the regression which includes the size and value factor is larger than the average adjusted R-square when regression only with the market factor (0.51 compared to 0.70 and 0.22 compared to 0.46 for the investment and ROE factor respectively), the traditional factor can explain the q factors better. Moreover, the result shows that the investment factor is negatively correlated with the market factor but positively correlated with the size and value factors when the Fama-French three factors are combined to explain the investment factor, but the reverse is true for the ROE factor. These results are similar to what Chen, Novy-Marx, and Zhang got in their paper in 2011 for the U.S. stock market [6].

The results of the GRSF-test statistics and the corresponding p-value in the parentheses for both models are reported in the Table 5 [10]. As shown in the Table, 5 the p-value of the test statistics of the q-factor model is significantly smaller than the value of the FF3 model (0.37 compared to 4.15×10^{-6}), thus the results of the GRS test show that the q-factor model misprices significantly less than the FF3 model.

Table 5: The values of the test statistics of the GRS F test.

Factors	F-statistics (p-values)
SMB, HML	1.11 (0.37)
r_{ROE}, r_{INV}	5.95 (4.15×10^{-6})

4. Discussion

Although the q-factor model behaves better in the stock market in China in the selected time period, the two model does not behave as well as in the U.S. Serval valid reasons, especially reasons about the efficiency of the market, can explain the situation.

First, it is generally assumed by the asset pricing models that the market is efficient. This is based on the efficient market hypothesis (EMH), but this is often not the case. Many papers have analyzed the efficiency of the stock market in China and concluded that the stock market in China is in the weak form or the semi-strong form of efficient market (e.g., Kiymaz in 2005) [13]. The U.S. stock market, in contrast, has been argued to be more efficient. This creates a great difference when doing empirical work to examine the asset pricing models. The information asymmetry and insider trading possibilities make the market different. Hence, the results of the tests of the models are different. Besides the common factors that are making the market inefficient, one particular essential factor that influences the pricing ability of the q-factor model is the potential incorrectness of the income statement.

One important assumption in the model is that the accounting data of the stocks in the market are accurate, which does not appear to be realistic, especially when it comes to the ROE data. Many companies have the incentive to utilize some accounting techniques to make the reported income in the income statement bigger to raise the stock price, however, in the long run, this would be corrected by the overall performance of the company. The pseudo-high ROE leads to the mispricing of the stock in the market, which is a common and important reason why the market is inefficient. This kind of market mispricing, incorrect metric, and cross-metric correction of price combine to make the empirical asset pricing inaccurate.

Moreover, the dynamic institutional changes of the Chinese stock market also make it hard for models to be persistent over time. Many reforms in the stock market in China have been proposed by the Chinese government in this century. In 2005, the Chinese government proposed a big reform that allowed the state-owned transaction-prohibited stock to circulate in the market. This first shocked the market, since at that time, the majority of the Chinese economy consists of the state-owned companies and so does the stock market, but then promoted the whole market, the Chinese stock market then accelerated to become the fourth largest stock market in the world. In 2013, the China Securities Regulatory Commission (CSRC), which is the controller of the Chinese Initial Public Offering (IPO) process, stopped reviewing the IPO applications to cool down the secondary market for almost a year. Since 2015, the Chinese government had proposed a new major institutional reform of the Chinese stock market, making short selling possible and deregulating the derivative market. Though many of these were later stopped or restricted, the process illustrates the institutionally changing feature of the Chinese market.

5. Conclusion

This paper tests the q-factor model with 60-month Chinese stock market data and compares the results with the traditional Fama-French three-factor model. Multiple regression techniques and tests are used to test and compare the models. The result shows that both the q-factor and the Fama-French three-factor model behave well in the Chinese stock market, but the q-factor model behaves slightly better than the traditional model in some statistical metrics. In the process, the differences and similarities between the Chinese and the U.S. stock market that can be observed are discussed. Major differences include the performance differences of small stocks compared to big stocks, high investment-to-assets ratio stocks compared to the lower ones, and the correlations differences between the factors. Similarities include the explaining abilities of the factors to explain each other, especially the market factor, and the ROE effect. Some practical limitations that are influencing the preciseness of the model are also discussed in the paper.

Further research on both models with the Chinese stock market is expected. Although this paper eliminates the bottom 30% of size of stocks to avoid shell-value contamination of returns, these stocks should have their unique pricing strategy, which can be explored further, for example, by adding a shell value factor in the model, since unless the inefficiency of getting listed in China vanishes, the potential reverse merger shell stocks should always be priced differently. Also, further explorations on the efficiency impact on models and the influence of institutional dynamics should be explored.

References

[1] *Sharpe, W. F.: Capital asset prices: A theory of market equilibrium under conditions of risk. The journal of finance, 19(3), 425-442 (1964).*

[2] *Lintner, J.: The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. The Review of Economics and Statistics, 47(1), 13-37 (1965).*

[3] *Black, F.: Capital market equilibrium with restricted borrowing. The Journal of business, 45(3), 444-455. (1972).*
[4] *Fama, E. F., French, K. R.: The cross-section of expected stock returns. The Journal of Finance, 47(2), 427-465 (1992).*
[5] *Carhart, M. M.: On persistence in mutual fund performance. The Journal of Finance, 52(1), 57-82 (1997).*
[6] *Long and Chen, Long and Novy-Marx, Robert and Zhang, Lu, An Alternative Three-Factor Model [Preprint]. Available at SSRN: https://ssrn.com/abstract=1418117 or http://dx.doi.org/10.2139/ssrn.1418117 (2011).*
[7] *Fama, E. F., French, K. R.: A five-factor asset pricing model. Journal of financial economics, 116(1), 1-22 (2015).*
[8] *Liu, J., Stambaugh, R. F., Yuan, Y.: Size and value in China. Journal of financial economics, 134(1), 48-69 (2019).*
[9] *Fama, E. F., MacBeth, J. D.: Risk, return, and equilibrium: Empirical tests. Journal of political economy, 81(3), 607-636 (1973).*
[10] *Gibbons, M. R., Ross, S. A., Shanken, J.: A test of the efficiency of a given portfolio. Econometrica: Journal of the Econometric Society, 1121-1152 (1989).*
[11] *Newey, W. K., & West, K. D.: A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica, 55(3), 703-708. (1987).*
[12] *Newey, W. K., & West, K. D. Automatic lag selection in covariance matrix estimation. The Review of Economic Studies, 61(4), 631-653 (1994).*
[13] *Kiymaz, H.: A performance evaluation of Chinese mutual funds. International Journal of Emerging Markets, 10(4), 820. (2015).*