Modified Fama-French Three-Factor Model and the Equity Premium: An Empirical Test in the Chinese Stock Market

Xiaoyuan Wang

College of Arts & Science, New York University, New York, NY 1000, USA.
Email: xw1158@nyu.edu

Abstract. Fama-French model is proposed as an alternative to the CAPM model to explain the equity premium found in stock market. This paper applies the original and modified Fama-French three-factor model in three industries (Finance & Insurance, Real Estate, and Pharmaceutical) of China's Shanghai Stock Exchange market. The study examines the relation between risk premium and the three Fama-French factors using six portfolios sorted by market capitalization and book-to-market ratio (inverse of price-to-book ratio). The result shows that Fama-French factors have consistent explanatory power of the risk premium. Using the original Three-Factor Model and CAPM Model as reference, the modified Three-Factor Model, with cross and quadratic terms added, significantly improves the performance of the model.

1. Introduction

Equity premium is the difference between the average return on equity and the average return on a virtually riskless asset. Based on historical U.S. stock market data, it is observed that the risk premium is too high (Mehra & Scott, 1985).

Various models have been used to explain the equity premium from many perspectives. Using the first-order risk aversion and recursive utility, the equity premium can only be partially solved; the equity premium under this model is reduced, but not effectively eliminated (Epstein & Zin, 1990). The continuous-time delayed-adjustment models potentially allows high equity premium, but fails long-horizon Euler equation test (Gabaix & Laibson). The incorporation of loss aversion and narrow framing into the traditional utility function showed promising initial results under some specific assumptions (Barberis & Ming, 2006).

Given the much shorter time span since the establishment of the Chinese stock market, there were only a few comprehensive verification results of the Fama-French three-factor model over the years. In Shanghai Stock Exchange, size and book-to-market ratio are highly correlated to the return on stocks (Yu, 2001); in Shenzhen Stock Exchange, the Fama-French three factor model holds true (Deng & Ma, 2005). However, as the stock market in China has become more mature and regulated, along with a quickly growing economy in terms of gross domestic product (GDP), the data from the last decade used in these researches cannot fully reflect the present situation of the Chinese stock market. Our goal is to examine whether the Fama-French three-factor model is still a good measure of the risk premium in the Chinese stock market in recent years.

Fama-French model is first proposed in 1992, as a potential alternative to the one factor Sharpe-Lintner model, for assessing the expected stock returns. Fama and French notice that the one factor Sharpe-Lintner model is unable to explain the relation between the average return and earning-to-price ratio (the inverse of price-earnings ratio), as they find the intercept in the one factor model increases monotonically while the market βs for all positive earning-to-price ratio portfolios are close to 1.0. In another word, earning-to-price ratio has a significant residual effect in average returns. They conclude that the failure is due to the nearly constant market βs for most portfolios, while other potential explanatory variables are not included. On the other hand, the Fama-French three-factor model performs well on the same set of data, with no residual effect in average return caused by earning-to-price ratio (Fama & French, 1992).

Later researches have proven that the coefficients in the three-factor model differ greatly across different countries. Fama-French factors are country-specific; domestic economic factors have higher influence than global ones (Griffin, 2002). In a U.S. regional test, it is shown that the HML factor has a significant impact in the risk premium, but not the SMB factor (Abhakorn, Smith & Wickens, 2013).

The second part of this paper is related works. The third part explains the models used in detail. The fourth part is dataset and the sources. The fifth part is result and conclusion, and the last part is future work.

2. Related Works

The Fama-French three-factor model can be express as follows:

$$\mathrm{r} = \mathrm{R_f} + \beta_3(\mathrm{K_m} - \mathrm{R_f}) + \beta_s \mathrm{SMB} + \beta_r \mathrm{HML} + \alpha \quad (1)$$

In (1), r is the expected return, and K_m is the return of the market portfolio. β_3 is analogous, but not equal, to the β value from the Sharpe-Lintner-Black model, which is shown to be unable to explain the average stock return (Fama & French, 1992). SMB and HML are the historic excess returns of "Small (market capitalization) Minus Big" and "High (book-to-market ratio) Minus Low", respectively. The coefficients β_s and β_r are determined using linear regression.

In our case, the return of the market portfolio, K_m, reflects the situation in the entire market of Shanghai Securities Exchange. The rate of return is determined using the change of closing price (P_t) at the end of each trading day from the previous day, calculated as the following equation:

$$\mathrm{K_m} = \frac{\mathrm{I_{t+1}}}{\mathrm{I_t}} - 1 \quad (2)$$

The factor SMB measures the difference in return due to market capitalization. It is calculated as the difference between the mean return on portfolios consisting small market capitalization stocks and the mean return on portfolios consisting big market capitalization stocks, as the following equation:

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

The factor HML measures the difference in return due to book-to-market ratio. It is calculated as the difference between the mean return on portfolios consisting high book-to-market ratio stocks and the mean return on portfolios consisting low book-to-market ratio stocks, as the following equation:

$$\mathrm{HML} = \frac{1}{2}(\mathrm{SH} + \mathrm{BH}) - \frac{1}{2}(\mathrm{SL} + \mathrm{BL}) \quad (4)$$

3. The Model

$E[r]$ refers to expected return of all stocks. R_f is the daily risk free return, calculated based on the demand deposit rate shown in Table 1 by dividing the annual rate by 365. K_m is for the market performance over the period, using equation (2). SMB, HML are the two Fama-French factors determined using equation (3) and (4) resectively. If the model fits well, these explanatory variables should explain most of the risk premium after linear regression on β_3, β_s, and β_r, and α should be close to 0.

When running regression on the risk premium for the Fama-French three-factor model, equation (1) is modified in the following way, to show risk premium on one side and the three factors on the other side:

$$E[r] - R_f = \beta_3(K_m - R_f) + \beta_s SMB + \beta_r HML + \alpha \quad (5)$$

In equation (5), r is the return of the stocks in the portfolio, as shown in table 2. R_f is the risk-free return rate on the day based on the term deposit return rate. K_m reflects the market trends using the percent change in the SHSE index, as shown in equation (2). The factors SMB and HML, as in equation (3) and (4), are also adjusted at the end of each trading day, using the mean percent change in stock prices for each of the six combinations:

$$SH = E[\frac{P_{t+1}}{P_t}] - 1 \text{ (Similar for the other five types)} \quad (6)$$

The coefficients of the three factors are determined using linear regression. If the explanatory variables, i.e. the three Fama-French factors, can explain the excess return of the portfolio, the model should fit well into the dataset. The coefficients should have a relatively high R2 and T-statistic.
Since the original Fama-French Three-Factor Model holds an implicit assumption that the size (SMB) and book-to-market ratio (HML) factors have constant effect at all levels, which may not be realistic for the chosen dataset, equation (5) is further modified to include two quadratic terms (SMB^2, HML^2) and a cross term ($SMB \cdot HML$) in order to account for the possibly varying effects of the two factors at different levels:

$$E[r] - R_f = \beta_3(K_m - R_f) + \beta_s SMB + \beta_r HML + \beta_c SMB \cdot HML + \beta_p HML^2 + \beta_q SMB^2 + \alpha \quad (7)$$

4. Dataset

All data for the stocks and their respective companies, selected from the Shanghai Securities Exchange (SHSE), is daily (trading day) spanning from January 2011 to December 2016 taken from Wind Financial Terminal. R_f, the return rate on risk free asset, is calculated using the annual yield of demand deposit saving rate, which is subject to infrequent adjustments. Although termed deposit and national bond provides a consistently higher return, they apparently have significantly less liquidity and are subject to greater interest rate risk. The return on market portfolio, K_m, is calculated using the daily percentage change of Shanghai Securities Composite Index. SMB and HML are determined using the difference between expected returns due to market capitalization and book-to-market ratio from the stocks in the sample.

The return of a stock after each trading day is calculated using the percentage change in price at the end of each trading day. The companies selected are from Finance & Insurance, Real Estate, and Pharmaceutical sectors, as categorized by China Securies Regulatory Commission; each is considered representative in their respective industry. Companies with negative equity value are excluded from the sample. When determining the factors, stocks not traded on a trading day are excluded from the mean for that day.

A few different ways were used to group the data of stocks according to their size and book-to-market ratio. Fama and French split the stocks into two groups by size, and three groups by book-to-market ratio; the cutoff percentage is 30%, 40%, and 30% (Fama & French, 1992). On the other hand, Deng & Ma split the stocks into three groups, with cutoff percentage of 35%, 40%, and 25% (Deng & Ma, 2005). Here we will divide the stocks having equal numbers in each group. The sample is first divided into two groups by size, big (B) and small (S), and each of the two subgroups is divided again into three groups by book-to-market ratio, high (H), medium (M), and low (L). The mean market capitalization and mean book-to-market ratio over time are used to categorize the stocks. Therefore, six different combinations are constructed: BH, BM, BL, SH, SM, and SL.

Table 1. Demand deposit annual risk-free return rate (R_f), 2011-2016.

Date of adjustment	2011.1.1	2011.2.9	2011.4.6	2012.6.8	2012.7.6
Risk free rate (annual)	0.36%	0.4%	0.5%	0.4%	0.35%

Table 2. Stocks Chosen for the Portfolio.(Each Column represents the categorization of stocks based on market capitalization and book-to-market ratio.)

BH	BM	BL	SH	SM	SL
Shanghai Pudong Development Bank (600000)	CITIC Securities Co., Ltd. (600030)	Jiangsu Hengrui Medicine Co., Ltd. (600276)	Nanjing Xingang High-Tech Co., Ltd. (600064)	China World Trade Center (600007)	Beijing Tongrentang Co., Ltd. (600085)
China Minsheng Banking Corp., Ltd (600016)	Poly Real Estate Group Co., Ltd. (600048)	Guangzhou Baiyunshan Pharmaceutical Holding Co., Ltd. (600332)	Beijing Vantone Real Estate Co., Ltd. (600246)	Wolong Real Estate Co., Ltd. (600173)	Chongqing Taiji Industry (Group) Co., Ltd. (600129)
Gemdale Corp. (600383)	China Pacific Insurance Group Co., Ltd. (601601)	Kangmei Pharmaceutical Co., Ltd. (600518)	Guangzhou Pearl River Industrial Development Co., Ltd. (600684)	Henan Taloph Pharmaceutical Co., Ltd. (600222)	Zhejiang Huahai Pharmaceutical Co., Ltd. (600521)
Industrial and Commercial Bank of China (601398)	China Life Insurance Co., Ltd. (601628)	Ping An Insurance Group Company of China, Ltd. (601318)	Huayuan Property Co., Ltd. (600743)	Zhejiang Hisun Pharmaceutical Co., Ltd. (600267)	Furen Medicines Group Co., Ltd. (600781)
Bank of China Ltd. (601988)	Huatai Securities Co., Ltd. (601688)	Industrial Securities Co., Ltd. (601377)	BEIH Property Co., Ltd. (600791)	Bright Real Estate Group Co., Ltd. (600708)	North China Pharmaceutical Co., Ltd. (600812)

5. Results and Conclusion

From the results in Table 3, it is apparent that the for the chosen dataset the CAPM component, in both the original and the modified Fama-French Three-Factor Model, still has the most dominant effect, with $\beta_3 \approx 1.1$ and a very high t-value. Using the CAPM model as a reference, it has $R^2 \approx 0.9$, but the intercept α is significantly nonzero, with a t-value of 3.60. In the original Fama-French Three-Factor model, even though SMB and HML factors are small in magnitude, both have a high t-statistic. The original Fama-French Three-Factor Model slightly improves (reducing α by about 1.64%) but is still far from fully resolving the issue of significantly nonzero α. However, with the introduction of additional terms, both the magnitude and t-value of α are reduced, and R^2 also increases slightly. The coefficients of newly introduced terms are all small in magnitude compared to β_3, but all have a relatively good t-value, with a minimum $t = 3.30$. In addition, α now can be easily rejected. Therefore, the modified Three-Factor Model has much more explanatory power over the original Fama-French Three-Factor Model and the CAPM model.

By looking at the $\beta \approx 1.12$ and $\beta_3 \approx 1.1$ in all three models, we can conclude that selected stocks are very sensitive to market fluctuations, namely changes in the SHSE Index from day to day. The SMB factor has positive coefficients, as expected; on the other hand, the HML factor has negative coefficients. This indicates that the effect of book-to-market ratio is reversed: stocks with lower book-to-market ratio actually have higher expected return than those with higher book-to-market ratio for this dataset. This may be due to the nature the industries, as real estate and pharmaceutical companies require much more capital than finance and insurance firms. Moreover, the result of the modified Three-Factor Model, with quadratic and cross terms included, significantly improves the original Fama-French Three-Factor Model. This also demonstrates that effects of SMB and HML are larger at higher levels. The distribution of α is more random in the modified model, while it is significantly nonzero in the original model and the CAPM model.

Table 3. Regression results using STATA. The CAPM model is also shown as a reference for the explanatory power of variables.

3-factor model: $E[r] - R_f = \beta_3(K_m - R_f) + \beta_s SMB + \beta_r HML + \alpha$						
Equation	Observations	Parameters	RMSE	R^2	F	P
$E[r] - R_f$	1458	4	0.5195277	0.9119	5016.993	0.000
$E[r] - R_f$	Coefficient	Std. Err.	t	P>|t|	95% Confidence Interval	
$K_m - R_f$	1.09869	0.0092929	118.23	0.000	1.080461	1.116919
SMB	0.040691	0.0031206	13.04	0.000	0.0345728	0.0468154
HML	-0.0243042	0.0053648	-4.53	0.000	-0.0348277	-0.0137806
α	0.0519413	0.0136081	3.82	0.000	0.0252476	0.0786349
3-factor model with quadratic terms added: $E[r] - R_f = \beta_3(K_m - R_f) + \beta_s SMB + \beta_r HML + \beta_c SMB \cdot HML + \beta_p HML^2 + \beta_q SMB^2 + \alpha$						
Equation	Observations	Parameters	RMSE	R^2	F	P
$E[r] - R_f$	1458	7	0.5140109	0.9139	2568.362	0.000
$E[r] - R_f$	Coefficient	Std. Err.	t	P>|t|	95% Confidence Interval	
$K_m - R_f$	1.099007	0.0091998	119.46	0.000	1.080961	1.117053
SMB	0.0415415	0.0031976	12.99	0.000	0.0352692	0.0478138
HML	-0.0257346	0.0053214	-4.84	0.000	-0.036173	-0.0152962
$SMB \times HML$	0.0038344	0.0009125	4.20	0.000	0.0020443	0.0056244
HML^2	0.0055115	0.0011528	4.78	0.000	0.0032502	0.0077728
SMB^2	0.000881	0.0002673	3.30	0.001	0.0003566	0.0014053
α	0.0076839	0.0156505	0.49	0.624	-0.0230161	0.038384
CAPM model: $E[r] - R_f = \beta(K_m - R_f)$						
Source	SS	df	MS		Observations = 1458	
Model	3997.18823	1	3997.18823		F(1, 1456) = 12716.79	
Residual	457.655305	1456	0.314323699		Probability > F = 0.0000	
Total	4454.84354	1457	3.05754532		R^2 = 0.8973	
					Adjusted R^2 = 0.8972	
					Root MSE = 0.56065	
$E[r] - R_f$	Coefficient	Std. Err.	t	P>|t|	95% Confidence Interval	
$K_m - R_f$	1.117395	0.0099087	112.77	0.000	1.097958	1.136832
α	0.0528094	0.0146838	3.60	0.000	0.0240058	0.081613

6. Future Work

The original research conducted by Fama and French used more data over a rather long period of time, and later they added two additional factors, profitability and investment, to their original Three-Factor Model. In a study conducted by Fama and French on the U.S. stock market, the Five-Factor Model explains 71% to 94% of the of the cross-section variance of expected returns for the portfolios examined (Fama & French, 2015). In Lin's work conducted on the Chinese A-share stock market over a long period, it is shown that the Five-Factor Model always outperforms the Three-Factor Model (Lin, 2017). In Guo's work on the Chinese stock market, four out of five factors are found highly correlated to expected returns, but the model need to be more realistic (Guo et al., 2017). Thus, the Five-Factor Model, possibly with some modifications to account for market or industry specific effects, would likely show promising results, making it a worthwhile direction for future work.

In this study, the portfolios are not weighted; each stock has the same share. However, it is actually reasonable to assume that investors would initially weight them according to each stock's past performance before determining investment portfolio composition and may also make infrequent changes. Therefore, a weighted portfolio better reflects actual investment procedure, and may affect the risk premium associated with the stock portfolios.

In addition, an industry-specific study of the Fama-French Three-Factor Model can help determine whether the negative HML coefficient found in this study only applies to certain industries at certain time period, or it is a more common phenomenon in the Shanghai Securities Exchange and the more general Chinese stock market. In most studies, the data samples include stocks of many or all industries, and only a few are industry-focused studies. As companies in a certain industry may exhibit similarities, reflected through one or two factors in models, it would be easier to examine the other varying factors, in effect holding one or two factors relatively constant.

Moreover, stock markets are not isolated from the larger economic environment. It is shown that both domestic and global macroeconomic variables can effectively influence the stock market (Chen & Chiang, 2016). In particular, the U.S. economic variables have statistically significant influence over Chinese stock market volatility (Chen et al., 2016). Thus, a relatively accurate prediction of the Chinese stock market may involve many different, some seemingly unrelated factors, and the model may likely be industry-specific, from the factors to the coefficients.

7. References

[1] Abhakorn, Pongrapeeporn, Smith, Peter N., and Wickens, Michael R. What do the Fama-French Factors Add to the C-CAPM? Journal of Empirical Finance 22, 113-127.

[2] Baberis, Nicolas and Huang, Ming. The Loss Aversion / Narrow Framing Approach to the Equity Premium Puzzle. National Bureau of Economic Research No.12378.

[3] Chen, Jian., Jiang, Fuwei., Li, Hongyi and Xu, Weidong. Chinese Stock Market Volatility and the Role of U.S. Economic Variables.

[4] Chen, Xiaoyu and Chiang, Thomas C. Stock Returns and Economic Forces–an Empirical Investigation of Chinese Stock Markets. Global Finance Journal 30, 45-65.

[5] Deng, Changrong and Ma, Dongkai. Empirical Research of the Three-Factor Model in Chinese Stock Market. Chinese Journal of Management Vol.2 No.5.

[6] Epstein, Larry G. and Zin, Stanley E. “First Order” Risk Aversion and the Equity Premium Puzzle. Journal of Monetary Economics 26, 387-407.

[7] Fama, Eugene F. and French, Kenneth R. Common Risk Factors in the Returns on Stocks and bonds. Journal of Financial Economics 33, 3-56.

[8] Fama, Eugene F. and French, Kenneth R. A Five-Factor Asset Pricing Model. Journal of Financial Economics 116, 1-22.

[9] Guo, Bin., Zhang, Wei., Zhang, Yongjie., and Zhang, Han. The Five-Factor Asset Pricing Model Tests for the Chinese Stock Market. Pacific-Basin Finance Journal 43, 84-106.

[10] Lin, Qi. Noisy Prices and the Fama-French Five-Factor Asset Pricing Model in China. Emerging Market Review 31, 141-163.

[11] Mehra, Rajnish and Prescott, Edward C. The Equity Premium: a Puzzle. Journal of Financial Economics 15, 145-161.

[12] Xavier Gabaix, and David Laibson, "The 6D Bias and the Equity-Premium Puzzle," National Bureau of Economic Research Macroeconomics Annual 16 (2001)

[13] Yu, Shidian. An Empirical Test of Fama-French Three Factors Model in China Stock Market. BoHong Fund Management.