

The Review of Fama-French Asset Pricing Model

Rong Qi

Department of Mathematical Sciences, University of Liverpool, Liverpool, United Kingdom, L69 3BX

rqi@gecedu.uu.me

Keywords: Asset pricing, Fama-french models

Abstract: Fama-French (FF) models are of great importance in capturing patterns in average stock returns and asset pricing. The aim of this paper is to provide a comprehensive overview of the FF models, including their factors, construction, estimation methods and their efficiency in markets of different regions. We elaborate the factors added to this model from three-factor model to six-factor model and the underlying theories, followed by an outline of the construction of each individual factor. Besides, we also group state-of-the-art researches to make comparisons between different model estimation methods. Furthermore, the performance of FF multifactor models is also distinguished in different local markets. Finally, we outline some issues remaining in the models, which are expected to be solved in the future.

1. Introduction

Asset pricing models, which play a significant role in the financial market, are those designed by economists and investors aimed at identifying the patterns in average stock returns. The most famous initial one is the capital asset pricing model (CAPM) of Sharpe [1] and Lintner [2], which, however, was unable to explain many anomalies in returns found later. Therefore, an increasing number of improved asset pricing models were documented by researchers in the following decades. Fama-French (FF) [3-5] are the most well-known among them, which capture many patterns in average stock returns. A comprehensive introduction of FF models will be provided in this paper for those who want to have a preliminary understanding of Fama-French asset pricing models. It presents a structured overview of FF multifactor models, including their factors, construction, estimation methods and efficiency in different markets. In the following sections, Section 2 introduces the prevailing Fama-French models and their difference, and Section 3 summarizes the construction of individual factors in the models. Section 4 and 5 respectively make comparisons on different model estimation methods and on the efficiency of FF models in different regions. The conclusion will be given in Section 6.

2. Fama-French Models

On account of two significant patterns in average returns, the size and value factor, which were left unexplained by the CAPM model, Fama and French [3] proposed a three-factor model on the basis of an empirical investigation of determinants of stock returns in America. The center of tests on this model is a time-series regression:

$$R_{it} - R_{ft} = a_i + b_i(R_{Mt} - R_{ft}) + s_iSMB_t + h_iHML_t + e_{it} \quad (1)$$

where $R_{it} - R_{ft}$ is the period t return on security or portfolio i in excess of the risk-free return of period t , $R_{Mt} - R_{ft}$ is the excess return of the value-weight (VW) market portfolio to the risk-free return, SMB_t is the difference value between returns on diversified portfolios of small stocks and big stocks, HML_t equals to the return on a diversified portfolio of high book-to-market equity (B/M) stocks minus that of low B/M stocks, and e_{it} is a zero-mean residual.

The dividend discount model provides an explanation why the value factor B/M is related to average returns. It says that the discounted value of expected per-share dividends represents the market value of a share of stock:

$$m_t = \sum_{\tau=1}^{\infty} E(d_{t+\tau}) / (1+r)^{\tau} \quad (2)$$

where m_t is the share price at time t , r is the average expected stock return in the long term, and $E(d_{t+\tau})$ is the expected dividend per share for period $t+\tau$. With a bit of manipulation, this equation can be transformed into:

$$\frac{M_t}{B_t} = \frac{\sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau}) / (1+r)^{\tau}}{B_t} \quad (3)$$

In Equation (3), $Y_{t+\tau}$ is the total net income for period $t+\tau$ and $dB_{t+\tau} = B_{t+\tau} - B_{t+\tau-1}$ is the change in book value of equity, where B_t is book equity at time t . It can be concluded that for fixed values of B_t and M_t , lower expected earnings imply a lower expected stock return. Furthermore, by fixing everything except the $dB_{t+\tau}$ and the expected return in Equation (3), lower expected growth in book equity – investment – will lead to a higher expected return. All evidence mentioned above illustrates how B/M is related to the return rate. In addition, since smaller companies tend to be less stable, and therefore riskier, which requires higher returns, the size factor SMB becomes another factor added to the three-factor model.

Fama and French found that the three-factor model in Equation (1) sorted in size and value factor captures most unexplained patterns left by CAMP in returns. Another research by Achola and Muriu [6] showed that the three-factor model can be applied to Nairobi Securities Exchange (NSE) as well, an emerging market, especially when they adjusted it for thin trading. Besides, for portfolios sorted in other variables including sales growth, cash flow/price, and earnings/price documented by Lakonishok, Shleifer, and Vishny [7] and other researchers, the three-factor model is also a good description for their returns. Whereas, there exist some drawbacks in the three-factor model as well. In the same research of Achola and Muriu [6], the premium is not statistically significant. It is also noticeable that the good performance of different three-factor models formed on other variables that replace the SMB factor indicates that SMB is not the unique effective factor to explain the excess returns [7]. Additionally, Bartholdy and Peare [8] found that the three-factor model does not perform well in the estimates for individual stock returns.

Considering the anomalies of three-factor model, researchers manage to find some more factors to capture the patterns left unexplained by it. Aharoni, Grundy, and Zeng [9] discovered that investment has a statistically reliable relation with average return, and a proxy for expected profitability is documented to be strongly related to average return by Novy-Marx [10]. In addition, it is also implied that profit risk premium and investment risk premium will bring the excess return, and that many “anomaly” variables in average returns left unexplained by the previous three-factor model are related to investment and profitability. Therefore, Fama and French [4] developed a new multifactor asset pricing model augmented with two new factors, profitability and investment, called five-factor model:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_iSMB_t + h_iHML_t + r_iRMW_t + c_iCMA_t + e_{it} \quad (4)$$

In this equation, RMW_t and CMA_t are respectively the profitability and investment factors defined similarly as SMB_t and HML_t , that is, the difference in returns of diversified portfolios of stocks with robust and weak profitability, and that of low and high investment firms. The intercept a_i in (4) is zero for all portfolios and securities i , supposing that the five exposures, b_i , s_i , h_i , r_i

and c_i , capture all variations in expected returns.

Despite of four concerns casting doubt on the five-factor model by Hou, Xue and Zhang [11]: the negative correlation between internal rate of return and one-period-ahead expected return, the positive correlation between expected investment and one-period-ahead expected return, the redundancy of value factor and the poor proxy of past investment for the expected investment, in general, the five-factor model outperforms the FF three-factor model when it comes to different sets of left-hand-side (LHS) portfolios. In addition, several investigations indicate that the five-factor model also outperforms the three-factor one on many other markets all around the world [12][13].

In 2018, the five-factor model was augmented with a momentum factor [5]:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_iSMB_t + h_iHML_t + r_iRMW_t + c_iCMA_t + m_iUMD_t + e_{it} \quad (6)$$

During the study, models are simulated using $Sh^2(f)$, the max squared Sharpe ratio for model factors, as the index to be minimized. The results show that the six-factor model is reliably better on the index than other models. Additionally, it also performs best on the equal weight metrics and shows consistently strong performance in the split period tests.

3. Summary on the Construction of Individual Factors

To examine the significance of specified factor construction in tests of FF models, Fama and French [4] captured the patterns in average returns with different sets of factors. All of these factors are the difference values in average returns of portfolios constructed by various dimensions. Two main sorting methods are 2×3 and 5×5 .

In the 2×3 sorts, the stocks are allocated by Size and value factors independently into two Size groups and three B/M groups, where the 30th and 70th percentiles of B/M for NYSE stocks are the B/M breakpoints, and the Size breakpoint is the NYSE median market cap. Then, six value-weight (VW) portfolios are produced by intersections. The Size factor, $SMB_{B/M}$, is the difference value between the average return of the three small stock portfolios and that of the three big stock portfolios, that is, $SMB_{B/M} = (SH + SN + SL) / 3 - (BH + BN + BL) / 3$. The value factor HML is the average returns of the two high B/M portfolios of small stocks and big stocks minus that of the two low B/M portfolios, that is, $HML = (SH + BH) / 2 - (SL + BL) / 2$. The profitability and investment factors of the 2×3 sorts, RMW and CMA, are constructed similarly as HML. Two additional Size factors, SMB_{OP} and SMB_{Inv} , are produced by 2×3 sorts used to construct RMW and CMA, and thus the Size factor SMB from the three 2×3 sorts is equal to the average of SMB_{OP} , SMB_{Inv} , and $SMB_{B/M}$.

Another method is the 5×5 sorting, which is often used to construct the LHS portfolios. Taking 5×5 value-weight (VW) Size-B/M portfolios as an example, at the end of each period, stocks are sorted independently to five Size groups (Small to Big) according to NYSE market cap breakpoints. Five B/M groups (Low B/M to High B/M) are formed in the same way. 25 Size-B/M portfolios were produced by the intersections of the two sorts. The right-hand-side variables are $R_M - R_F$, SMB, HML, RMW and CMA, constructed using 2×3 sorts on Size and each of B/M, OP, and Inv. In each set of 25 regressions, the LHS variables are the monthly excess returns on the 25 Size-B/M portfolios. The five-factor regression equation is (2) mentioned above:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_iSMB_t + h_iHML_t + r_iRMW_t + c_iCMA_t + e_{it} \quad (7)$$

When developing the three-factor model, alternative definitions of SMB and HML are not considered. The choice of a 2×3 sort on Size and B/M is, however, arbitrary. Therefore, Fama and French [4] constructed individual factors in the five-factor model with several other methods including 2×2 and $2 \times 2 \times 2 \times 2$, which are similar to 2×3 . In general, despite the slight difference in accuracy, the selection of different construction of individual factors will not affect the result of

the estimation test.

4. Estimation Methods

There are different estimation methods for FF models. Ordinary Least Squares (OLS) regression method is the commonly used one in traditional tests of asset pricing theory, which focuses on the means of the distributions of covariates. Here, three other estimation methods will be introduced.

Quantile regression is a development of OLS model introduced by Koenker and Bassett [14]. Its central part is the median regression estimator which is designed to minimize a sum of absolute errors, while other conditional quantile functions are constructed to minimize an asymmetrically weighted sum of absolute errors. The combination of these estimated conditional quantile functions clearly shows how covariates affect the scale, location and shape of the distribution of the response variable. Allen, Powell and Singh [15] tested the FF three-factor model via this approach. This study shows that the factor models might follow a non-linear relationship. It also reveals the lack of effectiveness of the traditional method of OLS when applied to analyzing the extremes within a distribution, which always attracts investors' and risk managers' most attention.

Another method is the generalized method of moments minimizing a distance d (GMM_d), a method generating instruments with greater robustness in a panel data framework with fixed effects [16]. Bayesian averaging process is applied to the computation of these instruments [17]. This approach is parsimonious for its lack of requirements in computational power. Using the dataset same as FF, it leads to the conclusion that the market factor is the unique consistently significant factor. Moreover, no matter which method of OLS, GMM_d or $Haus_d$ (a modified Hausman artificial regression) [18] is used, the excess market return factor is significant at the 1% level for all twelve FF sectors for both the five-factor model and a six-factor model augmented with a liquidity factor.

One of the most popular method in recent years is machine learning presented by Diallo, Bagudu and Zhang [19]. It combines Bayesian optimization and support vector regression (SVR) to predict portfolio returns with the three and five factors proposed by Fama and French for the United States. Support vector regression (SVR) is a statistical machine learning toolbox that helps predict models according to the theory of structural risk minimization. For the search parameters of the SVR, they use a Bayesian algorithm. The combination of the two models make it possible to map both linear and non-linear relationships between the input factors and output dependent variables easily. As an algorithm of machine learning, researchers divide their dataset into training and testing datasets. Through cross-validation, the relationships between outputs and inputs can be established, and machine learning, which specializes in predictions, may help to get better predictions for the measurement of risk and return. Finally, the machine learning approach responds to collinearity issues among the factors as it has been found in the Finance literature. For five industries' portfolio returns in the United States over the period July 1926 to January 2019, the Bayesian support vector regression (BSVR) models perform very well compared with the existing literature. Specifically in the result, the new model, called the Fama-French BSVR three-factor model, outperformed the Fama-French BSVR five-factor model. More precisely, the Fama-French BSVR three-factor estimations attained correlation coefficients of 94% for portfolio returns in the testing dataset of the consumption and manufacturing industries. For the high-tech industry, a correlation of 92% between the predicted and experimental values of portfolio returns was found; 91% was found for the mining, construction, transportation, hotels, entertainment and finance industries. However, for the Fama-French BSVR five-factor model, the correlation coefficients seemed lower, which lay between 48% (health industry) and 89% (high-tech industry).

5. The Efficiency of Ff Models in Different Industries or Different Regions

Apart from the test of FF models on U.S. stock returns, several researches are aimed at the testing the efficiency of FF models in other regions. One of them focuses on the stock returns in

North America, Asia Pacific and Europe [12]. Factor spanning tests show that the investment factor, CMA, is redundant for stock returns in Europe and Japan from 1990 to 2015. One of the results common to previous work is that both the three-factor and five-factor model are unable to capture the “anomaly” on small stocks in U.S. the returns of which behave like those of firms that invest aggressively in spite of low profitability. In this test, they found similar evidence outside the U.S. market, which is even extreme in Europe and Asia Pacific.

In addition, an empirical test of the five-factor model in the Chinese equity market was given by Lin [13]. Using an extensive dataset of Chinese market over the period from 1997 to 2015, the test showed that compared with three-factor model, the five-factor model always has better performance on all metrics. More than 50% of the dispersion of average excess returns were left unexplained by the traditional three-factor model, while only approximately one third were left unexplained by the five-factor model. It is particularly striking that there is strong evidence that the investment factor, of great significance in Fama-French five factor model [4], seems to have little role in average returns description in China, while neither value factor nor profitability factor are redundant, consistent with what the factor spanning test shows in the results.

6. Conclusion

This paper discusses the great progress Fama-French models have made from three-factor, to five-factor, and to the current six-factor one. Overall, they have captured more patterns in average stock returns. In addition, by the original paper of five-factor model, it shows that the way to construct individual factors has little effect against the model estimation. Furthermore, different estimation methods of FF models can be applied to differently processed datasets, resulting in different efficiency. Several researches also demonstrate that FF models, especially the five-factor one, are valid in other international markets apart from U.S., in spite of different redundant factors. Last but not least, it is noticeable that the main problem of FF models is they fail to capture the patterns in low average returns of small stocks whose returns behave like those of firms with low profitability that invest aggressively. Therefore, a possible future work might be required to improve FF models, making it more efficient and reliable.

References

- [1] Sharpe, W.F. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk". *The Journal of Finance*, vol.19, no.3, pp.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*, vol.47, no.1, pp.13-37, 1965.
- [3] Fama, E.F., French, K.R. "Common risk factors in the returns on stocks and bonds". *Journal of Financial Economics*, vol.33, no.1, pp.3-56, 1993.
- [4] Fama, E.F., French, K.R. "A Five-Factor Asset Pricing Model". *Journal of Financial Economics*, vol.116, no.1, pp.1-22, 2015.
- [5] Fama, E.F., French, K.R. "Choosing Factors". *Journal of Financial Economics*, vol.128, no.2, pp.234-252, 2018.
- [6] Achola, N.K., Muri, P.W. "Testing The Three Factor Model Of Fama And French: Evidence From An Emerging Market." *European Scientific Journal*, vol.12, no.16, pp.211-236, 2016.
- [7] Lakonishok, J., Shleifer, A., Vishny, R.W. "Contrarian Investment, Extrapolation, and Risk". *The Journal of Finance*, vol.49, no.5, pp.1541-1578, 1994.
- [8] Bartholdy, J., Peare, P. "Estimation of expected return: CAPM vs. Fama and French". *International Review of Financial Analysis*, vol.14, no.4, pp.407-427, 2005.

- [9] Aharoni, G., Grundy, B., Zeng, Q. "Stock returns and the Miller Modigliani valuation formula: Revisiting the Fama French analysis". *Journal of Financial Economics*, vol.110, no.2, pp.347-357, 2013.
- [10] Novy-Marx, R. "The other side of value: The gross profitability premium". *Journal of Financial Economics*, vol.108, no.1, pp.1-28, 2013.
- [11] Hou, K., Xue, C., Zhang, L. "A Comparison of New Factor Models", [online] Available: <http://www.fmaconferences.org/Seoul/Papers/KeweiHou.pdf>
- [12] Fama, E.F., French, K.R. "International tests of a five-factor asset pricing model". *Journal of Financial Economics*, vol.123, no.3, pp.441-463, 2017.
- [13] Lin, Q. "Noisy prices and the Fama-French five-factor asset pricing model in China". *Emerging Markets Review*, vol.31, issue C, pp.141-163, 2017.
- [14] Koenker, R., Bassett Jr, G. "Regression Quantiles". *Econometrica*, vol.46, no.1, pp.33-50, 1978.
- [15] Allen, D.E., Singh, A.K., Powell, R. "Asset Pricing, the Fama-French Factor Model and the Implications of Quantile-Regression Analysis", [online] Available: <https://ssrn.com/abstract=1460470>
- [16] Racicot, F.E., Rentz, W.F. "A panel data robust instrumental variable approach: A test of the new Fama-French five-factor model". *Applied Economics Letters*, vol.24, no.6, pp.410-416, 2017.
- [17] Theil, H., Goldberger, A.S. "On pure and mixed statistical estimation in economics". *International Economic Review*, vol.2, no.1, pp.65-78, 1961.
- [18] Hausman, J.A. "Specification tests in econometrics". *Econometrica*, vol.46, no.6, pp.1251-1271, 1978.
- [19] Diallo, B., Bagudu, A., Zhang, Q. "A Machine Learning Approach to the Fama-French Three- and Five-Factor Models", [online] Available: <https://ssrn.com/abstract=3440840>