Financial Engineering and Risk Management (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/ferm.2023.060201
ISSN 2523-2576 Vol. 6 Num. 2

Examination of Fama-French Four-Factor Model: Creating Factors in Quantitative Finance

Yikun Li*

Department of Financial Management, Metropolitan College, Boston University, Taiyuan, Shanxi, 030001, China

**Corresponding author*

Keywords: Fama-French Four-Factor Model, SMB ("Small-minus-Big"), HML ("High-minus-Low"), Quantitative Finance, Anomalies, Equity Returns

Abstract: The paper consists of two parts. The first part creates the SMB ("Small-minus-Big") and HML ("High-minus-Low") factors according to the rules laid out in "Common risk factors in the returns on stocks and bonds" (Fama, E. F., & French, K. R, 1992) through the application of Python in quantitative finance. It also creates the momentum factor in Fama-French Four-Factor Model proposed by "On Persistence in Mutual Fund Performance," (Mark M. Carhart, 1997). The second part examines the effectiveness of factors in explaining anomalies in equity returns. The paper concludes that these three factors to some extent explain anomalies in equity returns.

1. Introduction

The Fama-French three-factor model expands on the capital asset pricing model (CAPM) by adding size risk and value risk factors to the overall market risk factor in CAPM model [1].

The reason for this model is that CAPM [2] has proved less reliable in practice. However, due to its simplicity, it is still widely used. It remains one of the easiest tools to predict return when investing. However, one problem with this model is that when we include beta in the formula, we assume that risk can be measured solely by the price volatility of the stock. But the risks of moving prices in two different directions are not the same. The CAPM also assumes that the risk-free rate remains constant during the discount period. In real life, when the portfolio is held for more than ten years, the ratio is unlikely to stay the same for the entire period. When the risk-free rate increases, stocks may end up overvalued because the cost of capital also increases. In summary, unfortunately, CAPM is not flexible enough and it uses only one variable to describe stock returns.

Professors Eugene Fama and Kenneth French [3] designed this model back in the 1990s to describe stock returns in portfolio management and asset pricing. The Fama-French three-factor model [4] focus on three major factors: the overall market risk, company size factor which represents outperformance of small versus big companies, and value factor which represents outperformance of high book-to-market equity versus low book-to-market equity companies. In general, small-cap high-value companies tend to do better than the overall market [5]. Valued companies outperform growth companies.

According to Carhart [6], much of what appeared to be the alpha of many mutual funds could in fact be explained as due to their loadings or sensitivities to market momentum. The original Fama-French model augmented with a momentum factor has become a common four-factor model used to evaluate abnormal performance of a stock portfolio.

2. Methods

2.1. Process of Creating Factors

2.1.1. Creation of SMB and HML Factor

We will use stock data from 1991 to 2020 to recreate the SMB ("Small-minus-Big") and HML ("High-minus-Low") factors. Firstly, we download stock data from June 1991 to December 2020 with total of 342 months in stock-security files, CRSP section at Wharton research data services website. We select price, common shares outstanding and holding period return without dividends. We need first two to calculate market equity of stocks and returns to calculate value-weighted returns of portfolios. This file is named "stock-ME". Then in CRSP/Compustat section, we download fundamentals annual data also from June 1991 to December 2020 which includes book value of stockholders' equity, balance sheet deferred taxes, investment tax credits and the book value of preferred stock. We will use these data to form a new variable, book equity. This file is named "stock-BE".

Since when we calculate the value-weighted return of six portfolios later, we will use annualized return, first we use preceding twelve months return in "stock-ME" file to calculate the annualized return. The formula shows below:

$$\text{Annualized Return}=((1+r1)*(1+r2)*(1+r3)*\ldots*(1+r12))^{1/12}-1 \quad (1)$$

Then we append the annualized return as a new variable in the "stock-ME" data frame. According to the paper, the market equity should be measured at the end of June. Therefore, we filter all data in June and form a new data frame in order to access the data we need later. Then we merge it with "stock-BE" file on ticker symbol and now we have all ME and BE data in one data frame. It is named "df". We use slicing to access calendar year data in "data date" column so that we could do the same operation for every year's data. We then group data frame "df" by the calendar year and then form thirty-one data frames which includes all stocks' data every year from 1991-2020 and ranked by the calendar year order.

In every separate data frame, we create three new variables: "ME" which is the market value, "BE" which is the book value and "BE/ME" which is book-to-market equity. The market equity equals to stock price times shares outstanding in calendar year t. The book equity of stock equals to book value of stockholders' equity plus balance sheet deferred taxes and investment tax credits, minus the book value of preferred stock in calendar year t. And book-to-market equity is BE for the fiscal year ending in calendar year t-1 divided by ME ending in calendar year t-1.

Next, we will sort all stocks, select particular groups and create the six portfolios formed by two size groups and three value groups [7]. First, we sort stocks by the market value. According to Fama-French three-factor model, market value represents the company's size and implies that the company tends to behave more like a big or small company. We add top 30% stocks of rank to the big size company group and the last 30% stocks of rank to the small size company group. Then we sort stocks again by the book-to-market equity. The book-to-market equity represents the value of a company and conveys that the company tend to behave more like a valued or growth company. Similarly, we take out top 30% stocks of rank into the high-value company group, 30%-70% stocks into the neutral-value company group, and the last 30% stocks into the low-value company group.

We select the intersections of the small-high-value group, small-neutral-value group, small-low-value group, big-high-value group, big-neutral-value group, big-low-value group and form corresponding six portfolios [8]. Then we need to calculate the returns of stocks in each portfolio, which is the weighted average return of each portfolio that calculated according to the weights of stocks in the ME column in "stock-ME" file. Now we can create the SMB and HML factor according to the following formula.

SMB = 1/3 (Small-high-Value+ Small-Neutral -value+ Small-low-value) – 1/3 (Big -high-Value + Big-Neutral-value + Big-low-value) (2)

HML=1/2(Small-high-Value + Big-high-Value) -1/2(Small-low-value+ Big-low-value) (3)

2.1.2. Creation of Momentum Factor

In addition to the market factor, Fama and French have proposed the "Small-minus-Big" (SMB) and "High-minus-Low" (HML) factors to explain anomalies in equity returns. Carhart has argued that there is evidence for a "momentum" (MOM) factor as well, leading to an enhanced CAPM model with a total of four factors [9]. The reasoning is as follows. The CAPM predicts that portfolios only earn return based on their market risk exposure, therefore, in

$$E[rP-rf]=\alpha CAPM+\beta M,CAPME[rMkt-rf] \quad (4)$$

The intercept $\alpha CAPM$=0. However, Fama and French showed that there are firms for which this is not true, or that there are "anomalous returns". But once the model is extended to

$$[rP-rf]=\alpha FF3+\beta M,\ F3\lambda M+\beta SMB\lambda SMB+\beta HML\lambda HML \quad (5)$$

They argued $\alpha FF3$=0. In other words, including the risk premia λSMB and λHML in addition to λM=$[rP-rf]$ should solve the problem. However, when researchers and practitioners have analyzed mutual fund performance, $\alpha FF3\neq0$ in some cases. Carhart showed that the strategy of "buying past winners, selling past losers", mimicked by the momentum factor and using the "risk premium" λMOM is able to explain this away and give $\alpha FF3$, MOM=0.

We are going to attempt to build this momentum factor [10]. Again, the details of construction are more involved than the guidelines in this problem, in particular to remove correlation and biases that exist between the factors. The model we want to set up is:

$$rP-rf=\alpha+\beta M(rMkt-rf)+\beta SMBSMBt+\beta HMLHMLt+\beta MOMMOMt+eP,t \quad (6)$$

Yielding

$$E[rP]-rf=\alpha+\beta M\lambda M+\beta SMB\lambda SMB+\beta HML\lambda HML+\beta MOM\lambda MOM \quad (7)$$

To construct the momentum factor, we need to do the following. First, we find the return of each stock in the preceding 11 months. We need to compute the rolling sum over 11 months of data. (We are going to use log-returns, and therefore, the total return over a period is the sum of the monthly returns.) Then it shifts the data by one month. This way, we put ourselves in the position of the investor who looks at how stocks have performed over the previous 11 months to pick winners and losers. Then we need to rank the returns to find the winners and the losers. The entries with the largest rank have the highest positive returns, and vice versa. Save the result in a new data frame, named ranked df. Next we find the average returns of the best 30% and worst 30% of performers over the preceding 11 months. This step corresponds to forming an equal-weighted portfolio of past winners and past losers. Applying the mean function to compute the average, we ignore all N/A values and are left with the returns of the portfolio according to our selection. Finally, to get the value for each time of the factor, find the difference between the top performers and the bottom performers.

And also plot the equity curve of the momentum factor, as shown in Figure 1.

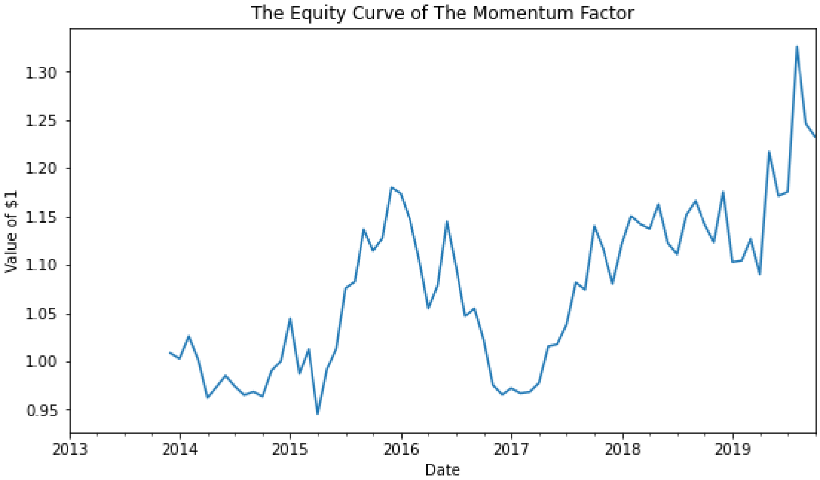

Figure 1: Equity curve of the momentum factor

3. Results and Discussions

3.1. Results of Creating Factors SMB and HML

Next, we are going to test how much of stock price volatility is explained by the created factors SMB and HML. Because there are more than ten thousand companies included in our data file, it will take much time to test them all. Select one stock AIR as a sample and run regressions of annualized return and the two factors separately. The results are shown in Figure 2 and Figure 3.

The R-squared number is 0.021 and 0.015 respectively. The result implies that created factors only explain a small portion in stock return variation. Our model or calculation method needs further improvement.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:      annualized_return   R-squared:                       0.021
Model:                            OLS   Adj. R-squared:                 -0.016
Method:                 Least Squares   F-statistic:                    0.5711
Date:                Fri, 07 May 2021   Prob (F-statistic):              0.456
Time:                        18:43:39   Log-Likelihood:                 58.313
No. Observations:                  29   AIC:                            -112.6
Df Residuals:                      27   BIC:                            -109.9
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      0.0060      0.008      0.792      0.435      -0.010       0.022
SMB           -0.2181      0.289     -0.756      0.456      -0.810       0.374
==============================================================================
Omnibus:                        0.909   Durbin-Watson:                   2.001
Prob(Omnibus):                  0.635   Jarque-Bera (JB):                0.824
Skew:                          -0.375   Prob(JB):                        0.662
Kurtosis:                       2.657   Cond. No.                         46.3
==============================================================================
```

Figure 2: Regression results of SMB

```
                            OLS Regression Results
==============================================================================
Dep. Variable:      annualized_return   R-squared:                       0.015
Model:                            OLS   Adj. R-squared:                 -0.021
Method:                 Least Squares   F-statistic:                    0.4168
Date:                Fri, 07 May 2021   Prob (F-statistic):              0.524
Time:                        18:43:39   Log-Likelihood:                 58.232
No. Observations:                  29   AIC:                            -112.5
Df Residuals:                      27   BIC:                            -109.7
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      0.0044      0.007      0.654      0.519      -0.009       0.018
HML            0.1244      0.193      0.646      0.524      -0.271       0.520
==============================================================================
Omnibus:                        0.917   Durbin-Watson:                   1.992
Prob(Omnibus):                  0.632   Jarque-Bera (JB):                0.838
Skew:                          -0.377   Prob(JB):                        0.658
Kurtosis:                       2.647   Cond. No.                         30.8
==============================================================================
```

Figure 3: Regression results of HML

3.2 Results of Creating Momentum Factor

In order to examine how the newly created momentum factor helps explain anomalies in equity returns, use the momentum factor and the factors we created above to run regressions for the following stocks: AMZN and GE. Use the monthly returns from November 2014 to October 2019 for estimation.

The first part is estimating the $\beta\ M\ O\ M$ for both companies, shown in Figure 4.

	AMZN	GE
FF4_MKT	1.615776	0.609198
FF4_SMB	-0.578007	0.406866
FF4_HML	-0.897658	-0.231196
FF4_MOM	0.119774	-0.818458

Figure 4: βMOM for AMZN and GE

As a result, βMOM for AMZN equals to 0.120, and βMOM for GE equals to -0.818.

Next, compare the adjusted $R2$ number of the regression including the momentum factor with a regression with the original Fama-French 3-Factor model. If the adjusted $R2$ number is larger once we include the momentum factor, we may suspect it helps to explain the returns. The regression results are shown in Figure 5 and Figure 6.

OLS Regression Results

Dep. Variable:	GE	R-squared:	0.299
Model:	OLS	Adj. R-squared:	0.248
Method:	Least Squares	F-statistic:	5.866
Date:	Tue, 21 Feb 2023	Prob (F-statistic):	0.000526
Time:	18:44:15	Log-Likelihood:	72.313
No. Observations:	60	AIC:	-134.6
Df Residuals:	55	BIC:	-124.2
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-0.0148	0.010	-1.433	0.158	-0.035	0.006
MKT	0.6092	0.337	1.810	0.076	-0.065	1.284
SMB	0.4069	0.428	0.951	0.346	-0.451	1.265
HML	-0.2312	0.497	-0.465	0.644	-1.227	0.765
MOM	-0.8185	0.366	-2.235	0.030	-1.552	-0.084

Omnibus:	18.516	Durbin-Watson:	1.591
Prob(Omnibus):	0.000	Jarque-Bera (JB):	45.746
Skew:	-0.817	Prob(JB):	1.17e-10
Kurtosis:	6.953	Cond. No.	61.0

Figure 5: OLS regression results for Fama-French 4-Factor model

OLS Regression Results

Dep. Variable:	GE	R-squared:	0.235
Model:	OLS	Adj. R-squared:	0.194
Method:	Least Squares	F-statistic:	5.747
Date:	Tue, 21 Feb 2023	Prob (F-statistic):	0.00168
Time:	18:42:28	Log-Likelihood:	69.706
No. Observations:	60	AIC:	-131.4
Df Residuals:	56	BIC:	-123.0
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-0.0186	0.011	-1.759	0.084	-0.040	0.003
MKT	1.0033	0.297	3.381	0.001	0.409	1.598
SMB	0.4467	0.443	1.009	0.317	-0.440	1.333
HML	0.4880	0.392	1.246	0.218	-0.297	1.273

Omnibus:	15.717	Durbin-Watson:	1.634
Prob(Omnibus):	0.000	Jarque-Bera (JB):	34.599
Skew:	-0.708	Prob(JB):	3.07e-08
Kurtosis:	6.440	Cond. No.	46.1

Figure 6: OLS regression results for Fama-French 3-Factor model

4. Conclusion

This paper proves the effectiveness of SMB and HML factor presented in Fama-French four-factor model and the momentum factor proposed by Carhart. Through the regression analysis of historical stock returns and the values of three factors we created, we conclude that these three factors to some extent explain anomalies in equity returns.

Acknowledgement

Finding: The Adj.R^2 for FF3F is 0.194, and it increase to 0.248 if the momentum factor is included. So, it is suspected that the momentum factor helps to explain the returns.

References

[1] Amit Goyal, "Empirical Cross Sectional Asset Pricing: A Survey," Financial Markets and Portfolio Management 26 (2012), pp. 3–38

[2] Elbannan, Mona A. "The capital asset pricing model: an overview of the theory." International Journal of Economics and Finance 7.1 (2015): 216-228.

[3] Fama, E. F., & French, K. R. (1992). Common risk factors in the returns on stocks and bonds (pp. 392-449). University of Chicago Press

[4] Gil Aharoni, Bruce Grundy and Qi Zeng, "Stock returns and the Miller Modigliani valuation formula: Revisiting the Fama French analysis", Journal of Financial Economics • November 2013

[5] Grauer, Robert R., and Johannus A. Janmaat. "Cross-sectional tests of the CAPM and Fama–French three-factor model." Journal of banking & Finance 34.2 (2010): 457-470.

[6] Mark M. Carhart, "On Persistence in Mutual Fund Performance," Journal of Finance 52 (March 1997), pp. 57–82

[7] J. Liew and M. Vassalou, "Can Book-to-Market, Size and Momentum Be Risk Factors That Predict Economic Growth?" Journal of Financial Economics 57 (2000), pp. 221–45.

[8] Rehman, Alam, and Qadar Bakhsh Baloch. "Evaluating Pakistan's Mutual Fund Performance: Validating through CAPM and Fama French 3-Factor Model." Journal of Managerial Sciences 10.1 (2016).

[9] Bello, Zakri Y. "A statistical comparison of the CAPM to the Fama-French Three Factor Model and the Carhart's Model." Global Journal of Finance and Banking Issues 2.2 (2008).

[10] Rath, Subhrendu, and Robert B. Durand. "Decomposing the size, value and momentum premia of the Fama–French–Carhart four-factor model." Economics Letters 132 (2015): 139-141.