Empirical analysis on construction of mimicking factors and its effect on model adequacy in the Fama-French three-factor model

Jiajing Yao*

Faculty of Business and Economics, Centre for Actuarial Studies, The University of Melbourne, Melbourne, VIC 3000, Australia

* Corresponding author: jiajingy@student.unimelb.edu.au

Abstract. Since the empirical failure of the CAPM, the Fama-French three-factor model that attempts to address some of its greatest issues has become widely popular. However, the choice of the two "mimicking" factors in the model are based mostly on empirical merit, and their construction is purely arbitrary. This paper investigates whether the model is robust with respect to changes in how the two mimicking factors are constructed, and its model adequacy, using US stocks data over a longer period (including more recent years) from July 1926 to June 2022. The conclusion is that the model is largely robust to an alternative formulation of the two factors, though the model is clearly inadequate particularly when tested using data over the longer period. Model fit is especially poor for portfolios formed by small stocks with low book-to-market equity ratio. Hence, caution needs to be taken when using the model for these stocks, and alternative models need to be sought.

Keywords: Fama-French; three-factor model; CAPM; mimicking factors; alternative construction; empirical analysis; model adequacy.

1. Introduction

The Capital Asset Pricing Model (CAPM) first derived by Sharpe (1964) and Linter (1965), and the later more realistic Black (1972) version of it (which restricts borrowing), have always been a topic of controversy and debate [1, 2, 3]. On one hand, the model provides a convenient and simple linear relationship between expected returns and risks that is rather intuitive. However, on the other hand, many such as Banz (1981) [4] and Rosenberg et al. (1985) [5] have pointed out its empirical failings – other factors such as the size, in terms of market equity (ME), and the book-to-market equity ratio (BE/ME), have explanatory power on expected return that the market β of the CAPM is unable to capture, while β is too "flat" (and hence lack significance in explaining returns) even when it is the only explanatory variable used to model returns of stocks (Black, Jensen and Scholes (1972) [6], Fama and French (1992) [7]).

Although some such as Roll and Ross (1994) point out the difficulties in testing the CAPM, particularly in the choice of a suitable market index proxy, and hence conclude that these findings are not themselves enough to completely reject the CAPM's theory, even they conclude that these findings are an "indictment" of the model [8]. It does not matter whether the model's theory is technically rejected or not if the model's empirical performance is so poor that it is of little practical value. As such, many alternative models attempting to address the empirical shortcomings of the CAPM emerged, and perhaps one of the most prevalent and widely used alternatives is the three-factor model proposed in Fama and French (1993) (referred to as FF model hereafter) [9].

In addition to the excess return of the market proxy, the FF model proposes two additional factors that can be used to explain the expected return of stocks in their model: "Small Minus Big" (SMB), and "High Minus Low" (HML), which capture the difference in return between small and large firms and firms with high and low BE/ME ratios respectively. The choice of these two factors is based mostly on the two aforementioned empirical anomalies for the CAPM, though Fama and French (1993) argue that these two factors could also be "mimicking" factors that potentially capture risk premiums due to unrecognised general risk factors.

In any case, the FF model is for the most part an empirically motivated model, supported by their later admission in Fama and French (2004) [10] that these two factors are indeed “brute force constructs”, though they argue that it is “not fatal”. Moreover, the construction of these two factors is ultimately based on arbitrarily chosen breakpoints/splits based on the ME and ME/BE of firms， as Fama and French even pointed out in their own 1993 paper – “the splits are arbitrary…and [they] have not searched over alternatives”, with hopes that the model’s effectiveness does not hinge on this choice alone. Not many have investigated this potential oversight in much detail. Of course, since the FF model is very much an empirically motivated and driven model, the fact that these factors are constructed arbitrarily is itself of no concern as long as the model is, empirically, robust with respect to this choice, and is an adequate approximation for the returns of assets.

Therefore, this paper will re-evaluate the FF model’s empirical adequacy using time-series regression on data over an even longer period than the 68-year period of Fama and French (2000) (hereafter FF00), including data in more recent years [11]. It will also investigate model robustness with respect to the construction of factors by examining the effects of alternative constructions (based on another set of arbitrarily chosen portfolio splits) of the SMB and HML factors on the model’s performance and adequacy, by comparing the fitted time-series regression results with that of the original FF model specification.

Section 2 begins with an introduction and summary statistics on the data used for analysis and provides the specification of the models and notations used throughout the piece. Section 3 presents and compares the fitted models under both constructions of the two factors SMB and HML. Section 4 then concludes with a summary, and the implications of these results.

2. Data and Methodology

2.1 Model Specification and Original Construction of Factors

The data used in this paper is based on the US stock returns data provided in the data library of the Kenneth R. French website, which is itself based on data from the centre for Research in Security Prices (CRSP). The data includes all NYSE, AMEX and NASDAQ firms that have the necessary data, and spans from July 1926 to June 2022 (a period of 1152 months).

The FF model for time series regression based on monthly returns is specified by:

$$R_{it} - R_{ft} = \alpha_i + b_i MER_t + s_i SMB_t + h_i HML_t + \epsilon_{it} \tag{1}$$

For $i = 1, \dots, N$ and $t = 1, \dots, T$, where R_{it} and R_{ft} respectively denote the return on portfolio i and the risk-free rate of return (the one-month US treasury bill return), in month t; MER_t represents the excess return of the market index (in this case it is the value-weighted return of all CRSP monitored US firms that have valid data) over the risk-free rate in month t; and ϵ_{it} are error terms.

In the original specification of the three-factor model of Fama and French (1993), for each month t, the factor SMB_t is constructed as follows:

$$SMB_t = \frac{[(S/L)_t + (S/M)_t + (S/H)_t]}{3} - \frac{[(B/L)_t + (B/M)_t + (B/H)_t]}{3}, \tag{2}$$

With S and B respectively denoting “small” and “big” stocks in terms of ME respectively, L, M and H denoting stocks with “low”, “medium” and “high” BE/ME. These groupings are based on sorts of all stocks in the sample at the end of June every year. “Big” stocks are stocks with ME higher than the median ME of NYSE firms only, and “small” stocks are below the median. “Low” BE/ME stocks have BE/ME that are below the 30th percentile in terms of BE/ME of NYSE firms, “high” BE/ME stocks are greater than the 70th percentile, and “medium” BE/ME stocks are stocks in the middle 40 percent. The BE/ME for a stock at the end of June of every year is found as the book value at the end of the last fiscal year by the end of the previous calendar year, divided by the ME of the firm until the end of December the prior year.

Based on these groupings, six portfolios $S/L, S/M, S/H, B/L, B/M, B/H$ are formed as the intersection of the two size and three BE/ME groups (hereafter a "2×3 portfolio split"), with the weight of each stock in the portfolio being their value. Then the subscript t denotes the returns on these portfolios in month t. For example, $(B/H)_t$ denotes the value-weighted monthly return in month t of a portfolio of stocks that are both above the 50th and 70th percentiles respectively of the ME and BE/ME of NYSE stocks. Therefore, SMB_t is constructed to be the difference in the equal-weighted average of the returns from the three portfolios consisting of small stocks, and the three portfolios of big stocks.

Similarly, HML_t is constructed as:

$$HML_t = \frac{[(S/H)_t + (B/H)_t]}{2} - \frac{[(S/L)_t + (B/L)_t]}{2}, \tag{3}$$

Which is the difference between the average returns on the two portfolios of low BE/ME stocks, and the two portfolios consisting of high BE/ME stocks (again with equal weights), in month t.

The reason why SMB_t and HML_t are constructed in this manner, with equal weights, is so that these factors capture mostly only the effect of ME or BE/ME on the returns, without also including the effect of the other factor (BE/ME or ME) on returns.

Do note that the portfolios constructed are using all stocks from NYSE, AMEX and NASDAQ, whilst the breakpoints for the groups are based purely on the relevant percentiles of the NYSE firms, and that the portfolios are formed annually in end of June of years 1926 to 2022 based on these criteria, then held constant for a year.

Summary statistics for MER, SMB, HML and the returns of each of the six portfolios are provided in table 1 below:

Table 1. Summary Statistics for MER, SMB, HML and 2×3 Portfolio Returns

	MER	SMB	HML	S/L	S/M	S/H	B/L	B/M	B/H
Mean	0.669	0.191	0.357	0.974	1.233	1.430	0.926	0.954	1.183
SD	5.342	3.174	3.562	7.469	6.965	8.114	5.289	5.632	7.136
t-stat	4.251	2.039	3.398	4.424	6.007	5.982	5.945	5.750	5.629

The mean and SD represent the average and standard deviations respectively over months $t = 1, \dots, 1152$ (i.e. time series means and standard deviations) of the corresponding values over the entire observation period (July 1926 to June 2022, 1152 months). Values of MER_t, SMB_t, HML_t and the returns of the six portfolios are found using the method outlined above. The t-stat is the test statistic for a one-sample t-test with null hypothesis that the mean is zero.

The average value for the market excess return (MER) over the entire sample period is 0.669% per month, with a rather large t-statistic of 4.251, indicating that there is an extremely significant market premium. In fact, based on the t-statistic, it appears to be the most significant of the three factors. The SMB return premium is, however, rather small, at an average value of just 0.191% per month, and it is borderline significant at the 5% level of significance – it has a p-value of just 0.0416. This is similar to the weak size effect observed in FF00, and is a troubling sign that may suggest that the size factor of the FF model is insignificant. The HML value premium is also significant, like the MER, averaging an excess return of 0.357% per month, though its t-value is lower, at 3.398.

Interestingly, the significance of the value premiums due to the market and higher BE/ME seems to have swapped around compared to FF00 – the market premium is now more significant, and the HML premium less so – when data over a longer period, including more recent data, is used (FF00 uses data from 1926 to 1997).

The correlation between the two factors SMB and HML is 0.1157, which is relatively low, and seems to indicate that the factors are indeed mostly constructed to be independent of each other.

2.2 The Independent Returns to be Explained

Then, after the values of SMB_t and HML_t are found as above, a time series regression of the form proposed in Black, Jensen and Scholes (1972) is fitted for the value-weighted returns of 25 portfolios (i.e. a time series regression of the form in equation (1) is fitted using Ordinary Least Squares (OLS)).

These 25 portfolios are formed using a similar way to the 6 portfolios in section 2.1 above, except with breakpoints at each quintile for size and BE/ME of the NYSE stocks instead. For the rest of this paper, the following notation is used: Sx denotes stocks with ME in the xth size quintile of NYSE stocks, and By denotes stocks with BE/ME in the yth quintile in terms of BE/ME of NYSE stocks. The 25 portfolios are constructed based on the intersection of $S1, \dots, S5$ and $B1, \dots, B5$. A "$/$" is used to denote the intersection much like previously. For example, stocks in the portfolio $S1/B5$ have ME in the first size quintile (i.e. the smallest 20%) of NYSE stocks, and BE/ME in the fifth quintile (i.e. the highest 20%) for BE/ME of NYSE stocks.

The average return of these 25 constructed portfolios are shown in table 2 below:

Table 2. Average Monthly Returns (Over 1152 Months) of the 25 Portfolios Used as the Response

Size quintile	Book-to-market equity (BE/ME) quintiles				
	B1	B2	B3	B4	B5
	Average (time-series) monthly return				
S1	0.836	0.955	1.274	1.420	1.627
S2	0.897	1.196	1.231	1.298	1.520
S3	0.979	1.166	1.184	1.277	1.368
S4	0.997	1.050	1.119	1.214	1.317
S5	0.927	0.897	0.981	0.917	1.202

From Table 2, there is an overall increasing trend in average monthly returns moving across the BE/ME quintiles, confirming the BE/ME return premium observed in previous studies. However, the trend is less clear for the largest of stocks. Moving across from $S5/B1$ to $S5/B4$, there is no clear increasing trend in returns, and it is not until S5/B5 that there is an obvious return premium for higher BE/ME firms. Nonetheless, the trend is quite clear, and there does seem to be a return premium for higher BE/ME firms, justifying the use of the BE/ME factor of the FF model.

The size effect, however, is much more debatable. For portfolios in the 3 largest BE/ME quintiles, there is indeed a clear return premium for smaller stocks – the average monthly return decreases when moving down to larger size quintiles. For portfolios in the 2 smallest BE/ME quintiles, however, there is no clear size effect. In fact, portfolio S1/B1 has the lowest average monthly return out of the 5 portfolios formed from stocks in the first BE/ME quintile B1, despite it being comprised of stocks with the smallest ME. For the 5 portfolios formed with stocks in the second BE/ME quintile B2, there is a similar story – S1/B2 has the second lowest monthly return out of the five, and the 3 portfolios with size in the middle 3 size quintiles actually have the highest returns.

Therefore, it seems that not only is the size premium for small stocks not significant for portfolios formed from stocks falling into the two smallest BE/ME quintiles, but there is also a slight reversal of the size effect for those portfolios – the portfolios with the smallest stocks are in fact yielding lower average returns, though not by much. This proves problematic for the model fit, and is discussed in further detail in the results section.

2.3 An Alternative Construction of the Two Factors

In addition to the original construction of the SMB and HML factors in FF93, this paper also considers an alternate construction of the two mimicking factors to investigate whether how the two factors are constructed has an effect on model robustness and the adequacy of its fit. The following alternative construction in proposed: The same data outlined in section 2.1 is used. However, instead of the 2×3 portfolio split used to construct the two factors in the original FF93 model, a 3×5

portfolio split (i.e. 3 groupings based on size, and 5 groupings based on BE/ME) will be used. The three size groupings are denoted small (S), medium (M) and big (B), and respectively represent stocks below the 40th percentile, in the middle 20 percent, and above the 60th percentile in terms of size of the NYSE stocks. The BE/ME groupings are based on the BE/ME quintiles of NYSE stocks, and denoted B1, B2, B3, B4, B5 much like in section 2.2 above. Then 15 portfolios are formed using stocks in the intersections of these groupings (again using a "/" to denote the intersection). For example, stocks in the portfolio B/B1 have size above the 60th percentile of NYSE stocks in terms of size, and are in the lowest quintile in terms of BE/ME. The portfolios are sorted in July of every year.

Then, based on the 3×5 portfolio split, an alternative formulation of the factor capturing the size premium on small stocks is denoted and given as follows:

$$\widetilde{SMB}_t = \frac{[(S/B1)_t + (S/B2)_t + \cdots + (S/B5)_t]}{5} - \frac{[(B/B1)_t + \cdots + (B/B5)_t]}{5} \tag{4}$$

Where again the subscript t denotes the return of the portfolio in month t. Thus, under this alternative construction, the size factor $\widetilde{SMB}_t$ still represents the equally-weighted average difference in the returns of portfolios consisting of small stocks, and portfolios of big stocks in month t. However, the breakpoints used to construct the portfolios are different – there are now 3 groupings in terms of size, and 5 in terms of BE/ME – and so this is the average difference in returns of the 5 portfolios consisting of small stocks (in the smallest 40th percentile) and the returns of the 5 portfolios consisting of big stocks (in the largest 40th percentile).

Similarly, an alternative specification of the factor capturing the premium due to high BE/ME is denoted and given as follows:

$$\widetilde{HML}_t = \frac{[(S/B5)_t + (M/B5)_t + (B/B5)_t]}{3} - \frac{[(S/B1)_t + (M/B1)_t + (B/B1)_t]}{3}, \tag{5}$$

Therefore $\widetilde{HML}_t$ is the average difference in returns of the 3 portfolios consisting of high BE/ME stocks (in the highest 20 percent of NYSE stocks) and the 3 portfolios comprised of stocks in the lowest 20 percent.

Then, the model is still of the form in equation (1), and still fitted on the 25 portfolios outlined in section 2.2 above, except with SMB_t replaced with $\widetilde{SMB}_t$ and HML_t replaced with $\widetilde{HML}_t$. Table 3 below shows the summary statistics for the explanatory variables used in this modified version of the model:

Mean and SD are the time series averages and standard deviations of the factors respectively. t-stat is the test statistic for the t-test for whether the true mean is zero.

Table 3. Summary statistics for MER, $\widetilde{SMB}$ and $\widetilde{HML}$

	MER	$\widetilde{SMB}$	$\widetilde{HML}$
Mean	0.669	0.217	0.442
SD	5.342	3.742	4.214
t-stat	4.251	1.965	3.563

With the alternative formulation, the average value of the factor capturing the size premium, $\widetilde{SMB}$, is higher compared to the average value of SMB in section 2.1, though its t-statistic is marginally lower. The mean of the factor capturing the BE/ME premium, $\widetilde{HML}$, is noticeably (about 24%) higher than under the original construction, with also a larger value of the t-statistic.

The correlation between the two factors $\widetilde{SMB}$ and $\widetilde{HML}$ is now only 0.0051, which indicates almost no correlation between the two constructed mimicking factors, and is desired – the two are intended to capture variation in returns that is unexplained by the other, so it is desirable for their collinearity to be low. This is much lower than that observed in section 2.1, and can potentially improve the model fit.

3. Results

3.1 Fitted Model Under Original Construction

The results of the model fit under the original FF93 construction of the factors is summarised in table 4. The model fitted is of the form in equation (1). i.e.:

$$R_{it} - R_{ft} = \alpha_i + b_i MER_t + s_i SMB_t + h_i HML_t + \epsilon_{it}$$

The time series regression estimates for b_i for all 25 portfolios all have extremely large t-statistics, and are all positive, suggesting that there is indeed a significant market premium, and that the excess return of the market is a significant predictor of the excess returns of the portfolios, which is as predicted by both the FF model and CAPM.

The rows and columns represent the size and BE/ME quintiles that the stocks in the portfolio fall into respectively. The table gives the fitted time-series regression coefficient estimates for each of the portfolios and their t-statistics. The fitted model for each of the portfolios is of the form:

$$R_{it} - R_{ft} = \alpha_i + b_i MER_t + s_i SMB_t + h_i HML_t + \epsilon_{it}$$

R_a^2 represents the coefficient of determination, adjusted for degrees of freedom (adjusted R^2). The RSE is the residual standard error of the fitted model.

Table 4. Summary Statistics for Model Fit Using FF93 Specification

Size quintile	Book-to-market equity (BE/ME) quintiles									
	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5
			a_i					$t(a_i)$		
S1	-0.698	-0.400	-0.147	0.083	0.133	-3.338	-3.280	-1.505	1.265	1.807
S2	-0.237	0.013	0.023	0.027	0.050	-3.338	0.224	0.418	0.562	0.886
S3	-0.121	0.102	0.060	0.053	-0.061	-2.038	2.019	1.137	0.993	-0.879
S4	0.068	0.013	0.018	0.012	-0.143	1.424	0.253	0.329	0.194	-1.886
S5	0.095	0.006	0.001	-0.240	-0.182	2.844	0.147	0.012	-4.419	-1.766
			b_i					$t(b_i)$		
S1	1.282	1.077	1.079	0.943	0.981	30.722	44.223	55.317	72.005	66.578
S2	1.091	1.020	0.988	0.965	1.070	76.918	88.414	90.529	100.705	94.497
S3	1.134	1.017	0.985	0.998	1.116	95.895	100.676	94.250	93.150	80.777
S4	1.078	1.030	1.007	1.038	1.194	112.412	102.366	91.843	87.249	78.655
S5	1.028	0.980	0.961	1.033	1.180	153.507	119.162	93.992	95.471	57.353
			s_i					$t(s_i)$		
S1	1.461	1.538	1.245	1.224	1.310	21.247	38.325	38.736	56.755	53.920
S2	1.136	0.992	0.823	0.815	0.919	48.606	52.167	45.769	51.592	49.246
S3	0.809	0.544	0.444	0.471	0.578	41.525	32.698	25.787	26.707	25.392
S4	0.333	0.232	0.207	0.203	0.312	21.067	13.981	11.479	10.383	12.476
S5	-0.151	-0.190	-0.235	-0.187	-0.174	-13.702	-14.036	-13.964	-10.481	-5.135
			h_i					$t(h_i)$		
S1	0.366	0.208	0.547	0.578	0.900	6.117	5.964	19.575	30.829	42.641
S2	-0.224	0.125	0.348	0.570	0.876	-11.002	7.571	22.270	41.521	54.017
S3	-0.225	0.036	0.320	0.561	0.856	-13.280	2.511	21.392	36.542	43.288
S4	-0.345	0.105	0.340	0.568	0.942	-25.090	7.270	21.642	33.327	43.316
S5	-0.263	0.016	0.324	0.658	1.013	-27.446	1.317	22.129	42.417	34.395
			R_a^2					RSE		
S1	0.659	0.823	0.873	0.929	0.929	7.020	4.096	3.282	2.202	2.479
S2	0.910	0.932	0.935	0.952	0.952	2.387	1.940	1.835	1.612	1.904
S3	0.928	0.931	0.926	0.933	0.923	1.989	1.700	1.757	1.801	2.323
S4	0.932	0.923	0.914	0.915	0.912	1.614	1.692	1.843	2.000	2.553
S5	0.955	0.931	0.905	0.923	0.836	1.126	1.383	1.720	1.821	3.460

A similar statement can be made for the estimates of s_i, and so despite SMB being near insignificant (p-value of 0.0416) as a stand-alone variable (see section 2.1), when combined with the other two factors in the model, it does still have significant explanatory power on excess returns.

Moreover, the coefficient estimates for SMB do seem to be related to size – it always decreases from smaller to bigger-size quintiles (given a BE/ME quintile), which is as predicted by the model (the size premium should be lower as stocks get larger). This is consistent with the findings in FF93.

The same can mostly be said about the estimates for h_i, except for the 5 portfolios with stocks with BE/ME in the second BE/ME quintile – they seem to have much smaller (in terms of magnitude) t-statistics, especially for portfolio S5/B2, which has a t-statistic for h_i of 1.317, which is not significant at the 5% level of significance. The coefficient estimates for HML do also seem to be related to BE/ME (for the most part). It increases when moving from lower to higher BE/ME quintiles (for a given size quintile), again predicted by the model (BE/ME premium should be higher as BE/ME gets higher), with the exception of the portfolio S1/B1 – it is the only positive estimate for h_i out of all 5 estimates for the B1 quintile. This is again similar with the findings in FF93.

The adjusted R^2 values for the models fitted for each portfolio are for the most part high – 21 of them are greater than 0.9, suggesting that the 3 Fama-French factors do indeed explain a large proportion of the variation in the excess returns of portfolios. However, the portfolios S1/B1 (R^2 of 0.659) and S1/B2 (R^2 of 0.823) do indeed prove problematic for the model, giving the two lowest adjusted R^2 values out of all of the fitted models for the 25 portfolios (and S1/B1 by an extremely large margin). This diverges from the findings in FF93, which observed great fits for these two portfolios. This may be due to the observations made in section 2.2 – the size premium is non-existent and perhaps even reversed for these two portfolios in the data used. Despite them being formed by the smallest stocks in their respective BE/ME quintiles, they offer lower returns than portfolios formed from larger stocks in the same BE/ME quintiles. This differs from the predictions of the model.

Of course, the assessment of the model's performance also lies in whether it is adequate in explaining the returns of the portfolios. Based on the works of Merton (1973) and Ross (1976), this is equivalent to testing whether the intercepts of the fitted model are indistinguishable from zero [12, 13]. Based on the 5% level of significance, the critical value is around 1.96, and so the null hypothesis that the intercept is zero should be rejected when the t-statistic exceeds 1.96 in absolute value. 7 of the 25 intercepts are greater than the critical value in absolute value, and so are statistically significant at the 5% level of significance. This again diverges from the findings in FF93, which only finds 3 of the intercepts as being statistically significant. Perhaps the more problematic part is how far the intercepts differ from zero in terms of absolute magnitude, particularly with portfolios S1/B1 and S1/B2 (again), which have their fitted intercepts being -0.698% and -0.400% in monthly returns. This is concerning, as the average monthly risk-free rate of return over the whole sample period is only 0.267%, and these intercepts have well surpassed that.

Examining the t-statistics only demonstrates which of the 25 fitted models have the most problematic estimates for the intercepts, but to test jointly whether the FF model adequately explains the variation in returns for all 25 models, a Gibbons, Ross, and Shanken (1989) (GRS) F-test is required [14]. The test has the null hypothesis that the intercepts of all $N = 25$ intercepts are equal to 0, and its test statistic is given by:

$$GRS = \frac{T(T-N-L)}{N(T-L-1)} \frac{\alpha'\Sigma^{-1}\alpha}{1+\mu'\Sigma_f^{-1}\mu} \sim F(N, T-N-L), \tag{6}$$

Where $T = 1152$ is the number of time periods in the data, $L = 3$ is the number of predictors in the model, α is a column vector of the estimated intercepts, $'$ denotes the transpose, Σ is the covariance matrix of the model residuals, μ is the mean column vector of the 3 factors, and Σ_f is the covariance matrix of the 3 factors. $F(N, T-N-L)$ denotes a central F-distribution with N and $T-N-L$ degrees of freedom, and the sampling distribution of the test statistic is $F(N, T-N-L)$ under the null hypothesis. The test also assumes that the error terms ϵ_{it} are normally distributed, are uncorrelated through time, and have constant variance (homoscedastic).

The value of the test statistic is 3.343, and this has a p-value of 6.78×10^{-8}, which firmly rejects the null hypothesis at any reasonable level of significance. Therefore, the hypothesis that the FF model suffices to explain variation in excess returns is rejected, and much more firmly so compared to in FF93, as the data used spans a much longer time period. Although the FF model may be able to explain a lot of the variation in excess returns of the portfolios (demonstrated by the high adjusted R^2 values), it is not adequate in explaining all of the variation – there is variation left uncaptured by the model, indicated by the significance of the intercepts.

3.2 Fitted Model Using Alternative Formulation

The results of the fitted model using the alternative construction of factors (set out in section 2.3 above) are summarised in table 5. The fitted model is of the form:

$$R_{it} - R_{ft} = \alpha_i + b_i MER_t + s_i \widetilde{SMB}_t + h_i \widetilde{HML}_t + \epsilon_{it}, \tag{7}$$

Where $\widetilde{SMB}_t$ and $\widetilde{HML}_t$ are as outlined in section 2.3.

Overall, the results of the fitted model as similar to that presented in table 4 – the conclusion about the significance and the coefficient estimates of the MER factor and the size factor remain largely the same. For the BE/ME related factor, the only thing of note is that portfolio S5/B2 now has a coefficient estimate h_i that is significant at the 95% level of significance (t-statistic 2.394 instead of 1.317 as in table 4).

With respect to the intercepts, the alternative model fit significantly improves for portfolio S3/B1 (absolute value of t-statistic for the intercept drops from 2.038 to 1.557, so that it is no longer significant at 95% level), portfolio S5/B5 (from -4.419 to -3.743, though still significant), and portfolio S3/B2 (though only marginally, again so it is no longer significant). However, it deteriorates greatly for portfolios in higher BE/ME quintiles – particularly for portfolio S5/B5 (the intercept is now 2.540 standard errors from 0 compared to 1.766 previously). Overall, there are still 7 intercept terms that are significant at the 95% level.

For the GRS F-test, the statistic is now 3.241, with p-value 1.62×10^{-7}. Though this still firmly rejects the null hypothesis, it is about 2.4 times larger than the p-value for the model fitted using the original specification of the two factors. Thus, arguably, based on the GRS test statistic, the model fit does improve (i.e., more variation in excess returns is captured) when using the alternative formulation, though not significantly (at least not enough to make the 3-factor model adequate in explaining excess returns).

A potential explanation for this noticeable improvement in model fit is the much lower correlation between the two mimicking factors under this alternative construction (correlation of 0.0051 compared to 0.1157 with the original specification). The reduced collinearity between factors means that more variation in the response can be captured by the model.

However, the portfolios S1/B1 and S1/B2 still remain problematic, though the estimates for the intercepts are marginally reduced compared to in section 3.1 (-0.698% to -0.685% and -0.400% to -0.378% in monthly returns respectively). The fitted models for these two portfolios would have contributed greatly to the GRS F-statistic, and it is a large reason why the null hypothesis of model adequacy is rejected in both cases.

The rows give the size quintiles and columns give the BE/ME quintiles that the stocks in the portfolio fall into. The table gives the fitted time-series regression coefficient estimates for each of the portfolios and their t-statistics. The fitted model for the portfolios is of the form:

$$R_{it} - R_{ft} = \alpha_i + b_i MER_t + s_i \widetilde{SMB}_t + h_i \widetilde{HML}_t + \epsilon_{it}$$

Where $\widetilde{SMB}_t$ and $\widetilde{HML}_t$ are the alternative formulations described in section 1.3. R_a^2 represents the adjusted R^2 of the resultant model. The RSE is the residual standard error of the model.

Table 5. Summary Statistics for Model Fitted Using Alternative Formulation

Size quintile	Book-to-market equity (BE/ME) quintiles									
	B1	B2	B3	B4	B5	B1	B2	B3	B4	B5
			a_i					$t(a_i)$		
S1	-0.685	-0.378	-0.151	0.091	0.124	-3.273	-3.334	-1.549	1.361	1.645
S2	-0.212	0.021	0.025	0.028	0.023	-3.179	0.380	0.472	0.548	0.439
S3	-0.094	0.102	0.054	0.054	-0.105	-1.557	1.903	0.999	0.915	-1.763
S4	0.079	0.010	0.013	0.004	-0.163	1.542	0.198	0.232	0.061	-2.043
S5	0.098	0.001	0.001	-0.240	-0.236	2.775	0.024	0.014	-3.743	-2.540
			b_i					$t(b_i)$		
S1	1.289	1.077	1.078	0.946	0.977	30.836	47.513	55.200	70.569	64.964
S2	1.101	1.022	0.986	0.967	1.057	82.756	92.951	94.041	93.105	102.192
S3	1.152	1.024	0.986	0.999	1.099	95.755	95.335	92.056	85.414	92.411
S4	1.088	1.031	1.005	1.038	1.186	106.095	101.518	90.058	82.534	74.271
S5	1.025	0.977	0.962	1.034	1.167	145.075	118.897	88.984	80.698	62.984
			s_i					$t(s_i)$		
S1	1.290	1.397	1.126	1.125	1.234	22.126	44.203	41.315	60.169	58.791
S2	0.953	0.872	0.764	0.767	0.895	51.335	56.874	52.255	52.961	61.983
S3	0.633	0.433	0.400	0.471	0.582	37.753	28.908	26.750	28.841	35.090
S4	0.210	0.197	0.221	0.231	0.386	14.685	13.873	14.188	13.197	17.324
S5	-0.141	-0.159	-0.162	-0.069	-0.086	-14.299	-13.875	-10.708	-3.840	-3.328
			h_i					$t(h_i)$		
S1	0.253	0.097	0.438	0.419	0.714	5.021	3.553	18.594	25.963	39.401
S2	-0.229	0.081	0.260	0.429	0.745	-14.293	6.109	20.598	34.322	59.788
S3	-0.230	0.043	0.266	0.422	0.780	-15.901	3.290	20.607	29.904	54.463
S4	-0.276	0.092	0.270	0.450	0.762	-22.352	7.532	20.071	29.688	39.584
S5	-0.210	0.024	0.236	0.483	0.924	-24.655	2.394	18.134	31.319	41.378
			R_a^2					RSE		
S1	0.659	0.847	0.873	0.926	0.926	7.019	3.805	3.280	2.251	2.526
S2	0.922	0.939	0.940	0.944	0.960	2.233	1.845	1.760	1.743	1.737
S3	0.926	0.922	0.923	0.920	0.943	2.019	1.803	1.798	1.964	1.996
S4	0.923	0.922	0.911	0.905	0.903	1.722	1.705	1.873	2.111	2.681
S5	0.950	0.931	0.894	0.892	0.867	1.186	1.380	1.816	2.151	3.112

4. Conclusion

Based on the results in section 3, the alternative construction outlined in this piece for the two factors capturing premiums related to size and BE/ME of firms respectively in the Fama-French three-factor model does manage to explain slightly more of the variation in excess returns of portfolios (based on the GRS F-test). However, this is based on the single sample considered in this piece, and so is by no means conclusive.

This is also beside the point. The point is not to conduct an exhaustive optimality check for the "best" formulation (in terms of resultant GRS statistic) for the two factors – that would very much depend on the data, and is a classic example of "data snooping" – but rather to check whether the model is robust with respect to how these two factors are constructed (which could be an oversight, acknowledged in FF93), and whether the model is any more or less adequate compared to the conclusion in FF93 when data over a longer period (including more recent periods) is used to test the model empirically. This is why the alternative construction of the two factors is based also on arbitrarily chosen portfolio splits (the choice to use 3 groupings based on size and 5 based on BE/ME to construct the factors is purely arbitrary, much like the original FF93 construction is arbitrary).

The conclusion is that, based on the single alternative formulation, and single set of data considered in this piece, the fitted model does not differ greatly depending on how the two factors are constructed, with only slight improvements that may be down to sampling, and so the model is robust with respect to how the two factors are constructed. Future users of the model can proceed confidently knowing that it is indeed not an oversight of the model, as FF93 suspected.

However, when testing the model's adequacy under both specifications of the factors with US stocks data over July 1926 to June 2022, the hypothesis that the three Fama and French factors suffice to explain the variation in excess returns is firmly rejected, at a much more significant level than concluded in FF93 – the intercept terms are far too significant, and there must be other factors driving excess returns that are not captured by the model. Another point of note is that the model's fit suffers the most significantly for portfolios that are comprised of stocks that are both in the smallest size (S1) and two of the lowest BE/ME (B1, B2) quintiles.

Therefore, users of the model need to be wary that despite the model's popularity, it is not a perfect model – no models are. In fact, it is far from perfect, and it empirically does not suffice to explain the excess returns of portfolios, leaving some of the variation unexplained, despite it being a model born from empirical motivations. This is particularly the case for small firms with low BE/ME ratios, which the model struggles the most with, so caution needs to be taken when applying the model to these types of stocks.

Of course, the three-factor model proposed by Fama and French does still identify factors that are significant in explaining returns, and captures a lot (though not enough) of the variation. And as George Box once said, "all models are wrong", but some are useful. Even though the three-factor model is flawed, it should thus only be discarded in favour of a better one, and the fact that it empirically outperforms the CAPM (supported by findings of Banz (1981), Rosenberg et al. (1985) and Fama and French (1993)) is enough to warrant its use over the CAPM, and its recent popularity. However, the Fama-French three-factor model may be "wrong" by so much that its predictions need to be taken with a healthy pinch of salt, and better alternatives do need to be sought, and used, more often in practice.

References

[1] Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. The Journal of Finance, 19(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x.

[2] Lintner, J. (1965). 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. https://doi.org/10.2307/1924119.

[3] Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. The Journal of Business, 45(3), 444–455. Retrieved from https://www.jstor.org/stable/2351499.

[4] Banz, R. W. (1981). The relationship between return and market value of common stocks. Journal of Financial Economics, 9(1), 3–18. https://doi.org/10.1016/0304-405x(81)90018-0.

[5] Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. The Journal of Portfolio Management, 11(3), 9–16. https://doi.org/10.3905/jpm.1985.409007.

[6] Jensen, M. C., Black, F., & Scholes, M. S. (1972). The Capital Asset Pricing Model: Some Empirical Tests. Retrieved from SSRN website: https://ssrn.com/abstract=908569.

[7] Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. The Journal of Finance, 47(2), 427–465.

[8] ROLL, R., & ROSS, S. A. (1994). On the Cross-sectional Relation between Expected Returns and Betas. The Journal of Finance, 49(1), 101–121. https://doi.org/10.1111/j.1540-6261.1994.tb04422.x.

[9] 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.

[10] Fama, E. F., & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. Journal of Economic Perspectives, 18(3), 25–46. https://doi.org/10.1257/0895330042162430.

[11] Davis, J. L., Fama, E. F., & French, K. R. (2000). Characteristics, Covariances, and Average Returns: 1929 to 1997. The Journal of Finance, 55(1), 389–406. https://doi.org/10.1111/0022-1082.00209.

[12] Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. Econometrica, 41(5), 867. https://doi.org/10.2307/1913811.

[13] Ross, S. A. (1976). The arbitrage theory of capital asset pricing. Journal of Economic Theory, 13(3), 341–360. https://doi.org/10.1016/0022-0531(76)90046-6.

[14] Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A Test of the Efficiency of a Given Portfolio. Econometrica, 57(5), 1121. https://doi.org/10.2307/1913625.