Empirical Asset Pricing : HW1*

Maxime Borel

February 5, 2023

Introduction

Data

In order to construct the Fama-French 3-factor model (FF3) as proposed in Fama and French (1993), I obtained the fundamental data from Compustat and the price data from CRSP. The data range from January 1959 to December 2021. However, I focus on the analysis on the period from January 1963 to December 2021 (708 monthly observations). The data query is carried out using the WRDS python api and is inspired by Drechsler (2022).

I computed the book value of equity as shareholders' equity plus deferred taxes and investment tax credit minus preferred stock capital. The latter is measured at either current value, liquidation value, or redemption value. Only firms with strictly positive book value of equity are included in the sample. I calculated the market capitalization as the price of a share multiplied by the number of shares outstanding.

In the CRSP database, firms are identified by PERMNO and PERMCO. Some companies share the same PERMNO, so I aggregated them to get the holding company level. I merged the two data using the gvkey and PERMNO based linking table.

Factor construction

To construct the portfolios for the size and value factors, I applied the Fama and French (1993) method with a limited number of exceptions. In their seminal paper, Fama and French (1993) update the value-weighted weights and breakpoints of the portfolios on an annual basis, in June. They use a 2x3 double sort and calculate the breakpoints using only NYSE stocks.

I deviated from their methodology in the following ways. First, the portfolio weights and breakpoints are updated on a quarterly basis in March, June, September and December. Second, I use a double 3x3 sort based on the 25th and 75th percentiles of market capitalization and book-to-market. Finally, breakpoints are computed over the entire sample at hand. The portfolio returns are calculated as follows

$$R_t^{SMB^*} = \frac{1}{3}\left(R_t^{SH} + R_t^{SM} + R_t^{SL}\right) - \frac{1}{3}\left(R_t^{BH} + R_t^{BM} + R_t^{BL}\right) \tag{1}$$

$$R_t^{HML^*} = \frac{1}{3}\left(R_t^{SH} + R_t^{MH} + R_t^{BH}\right) - \frac{1}{3}\left(R_t^{SL} + R_t^{ML} + R_t^{BL}\right) \tag{2}$$

Thus, this leads to the new factors SMB* and HML*.

*The code is available on github here

Beta Pricing models

In this study, I analyse the 2 asset pricing models. The first is the traditional three-factor FF3 model by Fama and French (1993). The cross-section specification is

$$E[R_i] = \lambda_0 + \beta_{mkt-rf}\lambda_{mkt-rf} + \beta_{SMB}\lambda_{SMB} + \beta_{HML}\lambda_{HML} \tag{3}$$

I obtained the monthly data from Kenneth French's website[1].

The second asset pricing model specification is almost identical to the previous one. I replace the SMB and HML risk factors by the one proposed in the previous section. it gives

$$E[R_i] = \lambda_0 + \beta_{mkt-rf}\lambda_{mkt-rf} + \beta_{SMB^*}\lambda_{SMB^*} + \beta_{HML^*}\lambda_{HML^*} \tag{4}$$

We can notice that in the specifications (3) and (4), the first factor, namely the excess market return, is common to both models. It comes from the Fama-French library. I proceed as follows to further assess the impact of changing the construction of the last two components. In addition, I use the risk-free rate provided by Fama-French to calculate excess returns.

Test assets

Although there are various approaches to compare asset pricing models, it is very prevalent to use test assets to measure their pricing power. For traded risk factors, it is common to use test assets that are within the span of the models. For example, Fama and French (1993) used 25 portfolios based on size and book-to-market ratio. More recently, Kan et al. (2013) compared 8 asset pricing models using the 25 Fama-French portfolios based on size and book-to-market plus five industry portfolios. Moreover, Barillas and Shanken (2017) used portfolio based on the size, book-to-market, momentum, investment, operating profitability and 17 industries.

Since our final objective is to confront two models based on size and book-to-market and measure the effect of the difference in construction, I chose the 25 Fama-French size and book-to-market value-weighted portfolios as test assets. In addition, to follow the existing literature and add power to the test, I added 10 industry portfolios. I chose 10 industries because this number is closer to the first level of the widely accepted GICS industry classification.

Model comparison

in order to assess the pricing power and to compare the two models, I performed several tests. First, I followed Fama and French (1993) and performed time-series regression of both model specifications and look at the intercept of the regression. Furthermore, I performed the joint test on alpha proposed in Gibbons et al. (1989). This is not a statistical test that allows use to compare asset pricing model. To this aim, other tests are needed.

To complete the comparison of the models, I run two additional tests specifically built to compare asset pricing model. First, I compared the models based on the squared Sharpe ratio proposed in Barillas et al. (2020). This test is convenient because it does not rely on test assets. Second, I performed the test of equality of cross-sectional R^2 as in Kan et al. (2013).

[1]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Results

Factor returns

To begin with, I first check that the new construction of SMB* and HML* coincides with that of Fama and French (1993). Figure 1 shows the comparison of the performance of the two constructions for each factor. As expected, there are some discrepancies, but overall, the performances are relatively similar. Table 1 provides additional evidence. The correlation between the two constructs is relatively high for each factor. The size factor is less correlated with its Fama-French version, reaching 0.7141 while the HML* reaches 0.84. However, the results are in line with the change from 2 to 3 size categories. Overall, the new factors still capture the same underlying risks, respectively.

	SMB	HML
Correlation	0.7141	0.8371

Table 1: This table presents the correlation between the traditional size and value factor with the newly constructed. The correlation is measured over the full sample.

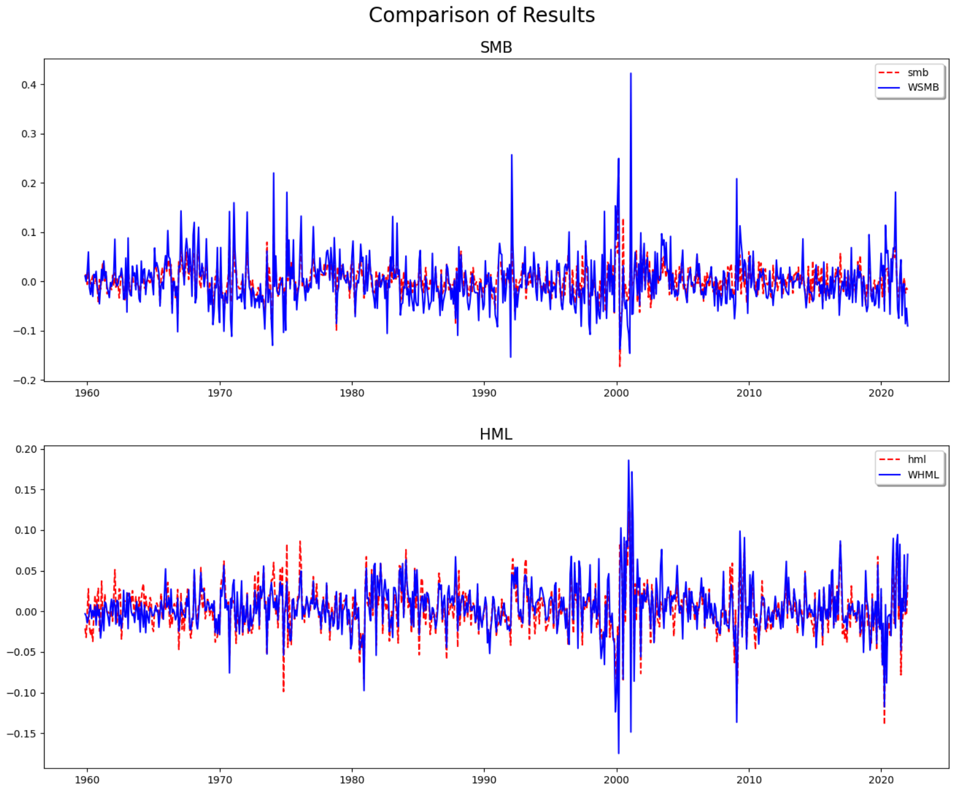

Figure 1: Comparison of the returns of SMB* and HML* with the construction of Fama and French (1993). WSMB and WHML are SMB* and HML*.

Time-series regression intercepts

a first step towards the comparison of the 2 asset pricing model is to compare their test assets intercepts from a time-series regression. Although this is not a proper way to statistically compare the two models, it provides some intuitions. I start the analysis by comparing panel A and B of Table 2. Then, I compare the α, in Table 3, from industry portfolio, which gives a better idea of the pricing power of the asset pricing model on assets that are not in the spirit of the factors.

Panel A : FF3					
	LowBM	BM2	BM3	BM4	HighBM
Small	-0.0048^{***}	0.0000	-0.0004	0.0013^{***}	0.0018^{**}
	(-5.1337)	(0.0673)	(-0.8064)	(2.6969)	(2.4428)
ME2	-0.0020^{***}	0.0000	0.0005	0.0004	0.0001
	(-3.1148)	(0.0687)	(0.8841)	(0.8314)	(0.1624)
ME3	-0.0011^{*}	0.0007	-0.0001	0.0006	0.0003
	(-1.8802)	(1.1389)	(-0.2232)	(0.9655)	(0.3914)
ME4	0.0009	-0.0005	-0.0002	0.0004	-0.0007
	(1.5422)	(-0.7994)	(-0.3171)	(0.6222)	(-0.8769)
Big	0.0017^{***}	0.0001	0.0001	-0.0023^{***}	-0.0019^{**}
	(4.1089)	(0.1762)	(0.1281)	(-3.8712)	(-1.9773)
Panel B : FF3*					
	LowBM	BM2	BM3	BM4	HighBM
Small	-0.0027^{**}	0.0013	0.0000	0.0014	0.0015
	(-2.4810)	(1.2247)	(0.0329)	(1.6022)	(1.6393)
ME2	-0.0004	0.0005	0.0004	0.0001	-0.0005
	(-0.4098)	(0.5823)	(0.4113)	(0.0932)	(-0.5150)
ME3	0.0002	0.0008	-0.0006	-0.0002	-0.0006
	(0.2261)	(0.9825)	(-0.7685)	(-0.3262)	(-0.6732)
ME4	0.0021^{***}	-0.0008	-0.0010	-0.0005	-0.0019^{**}
	(2.7990)	(-1.1331)	(-1.4406)	(-0.6749)	(-2.0911)
Big	0.0019^{***}	-0.0004	-0.0007	-0.0035^{***}	-0.0035^{***}
	(3.8711)	(-0.6880)	(-1.0023)	(-4.8420)	(-3.2018)

Table 2: This table presents the intercept, α, of the time-series regression $R_{i,t} - R_{f,t} = \alpha + \beta_i^{mkt-R_f}(R_t^{mkt} - R_{f,t}) + \beta_i^{K1}K1_t + \beta_i^{K2}K2_t + \epsilon_{i,t}$ over the 25 Fama-French size and book-to-market portfolios. The value in parentheses is the t-statistics. Panel A shows the results using the traditional 3-factor model from Fama-French. Panel B shows the results for the new SMB* and HML*. $*, **, ***$ indicates 10%, 5%, 1% statistical significance respectively.

In panel A of Table 2, there are 8 statistically significant intercepts. They are located in the corners of the table. In panel B, there are only 6 statistically significant intercepts located in the same place as in the FF3 asset pricing model. The order of magnitude of intercepts is relatively close across the two models. The FF3* seems to do a better job for the small and low book-to-market portfolios as the intercept are lower. However, the α are slightly superior for large and high book-to-market in absolute terms. Nevertheless, FF3* performs a good job because the test assets are formed in the identical fashion as FF3.

	Panel A : FF3									
	NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils	Other
α	0.0018**	-0.0021	-0.0005	-0.0004	0.0019*	0.0000	0.0013	0.0036***	0.0007	−0.0019***
	(2.0013)	(-1.3426)	(-0.7245)	(-0.2222)	(1.7833)	(0.0114)	(1.3109)	(3.2288)	(0.6189)	(-3.0646)
	Panel B : FF3*									
	NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils	Other
α	0.0012	−0.0032**	-0.0010	-0.0014	0.0030***	-0.0001	0.0012	0.0035***	-0.0002	−0.0028***
	(1.3162)	(-2.1268)	(-1.6184)	(-0.8238)	(2.7971)	(-0.1156)	(1.2354)	(3.0820)	(-0.1398)	(-4.4142)

Table 3: This table presents the intercept, α, of the time-series regression $R_{i,t} - R_{f,t} = \alpha + \beta_i^{mkt-R_f}(R_t^{mkt} - R_{f,t}) + \beta_i^{K1}K1_t + \beta_i^{K2}K2_t + \epsilon_{i,t}$ over the 10 industry portfolios from Keneth French Library. The value in parentheses is the t-statistics. Panel A shows the results using the traditional 3-factor model from Fama-French. Panel B shows the results for the new SMB* and HML*. $*, ***, ***$ indicates 10%, 5%, 1% statistical significance respectively.

Table 3 provides stronger intuition in comparing the two asset pricing model because the test asset portfolios are not directly related to the factors as the portfolios are based on criteria other than size and book-to-market sorts.

In both models, there are 4 intercepts that are statistically different from zero. They are the same except for the first two columns. The magnitude of the intercepts are slightly higher for the FF3* in absolute terms. However, they are very close. These results indicate that the FF3* model does a slightly less well than the traditional Fama-French 3-factor model. Overall, based on the individual asset test intercept comparison, there is not a model that performed clearly better than the other.

To obtain an overview of the pricing power of the model, the next step is to jointly test the hypothesis that all intercepts are equal to 0 for each model respectively. Table 4 shows the F-Statistics of Gibbons et al. (1989) for both models using only the 25 size and book-to-market portfolios in columns 1 and 3 and all test assets in columns 2 and 4.

Not surprisingly, the F-tests are rejected for both asset pricing models at the 99% level. Regardless of the set of test assets used to perform the statistical test, the F-statistic is weaker for the FF3* model. With only the 25 size and book-to-market portfolio, FF3 obtain a test statistic of 2.59 while FF3* reach 2.42. Augmenting the set of test assets with the industry portfolios increases the test statistic for both models and reduces the spread between them. FF3 reaches 3.58 and FF3* 3.49. These results suggest that FF3* generates the best-behaved intercepts. However, this is not a formal proof that one model is better than the other. The next section presents two tests for comparing asset pricing models.

	FF3 25x25	FF3 25x25+industry	FF3* 25x25	FF3* 25x25+industry
GRS	2.5905***	3.5828***	2.4215***	3.4887***
p-Values	0.0001	0.0000	0.0001	0.0000

Table 4: This table presents the GRS joint test on the intercepts. Under H_0, $\alpha_1 = ... = \alpha_N = 0$. The first two columns are the test statistics for the Fama and French (1993) factors. The last two columns are the test statistics for the SMB* and HML*. I distinguish between 25 Fama-French size and book-to-market portfolios and the full set of test assets. The second line shows the p-values of the test statistic. The distribution of the test is a $F(N, T-N-k)$. $*, ***, ***$ indicates 10%, 5%, 1% statistical significance respectively.

Model comparison tests

In this section, I show the results for two tests that allows model comparison. First, the test proposed by Barillas et al. (2020) based on squared Sharpe ratio. This test is convenient as it is agnostic to the choice of test assets. Second, I performed the CSR R^2 test proposed by Kan et al. (2013).

	Squared Sharpe Ratio
$\hat{\theta}_{FF3^*} - \hat{\theta}_{FF3}$	0.0279^{***}
p-Value	0.0037

Table 5: This table presents the pairwise test of equality of the squared Sharpe ratios of the two asset pricing model. The first row shows the test statistic proposed in Barillas et al. (2020) for nonnested models. The test is run using their sequential test. The first part is the GRS test proposed in Table 4. The second part compares the squared Sharpe ratios. $\theta_i^2 = \mu_i' \Sigma_i^{-1} \mu_i$. With$\mu_i$ and Σ_i^{-1} is the vector of means and covariance matrix of the factors, respectively $.*, ***, ***$ indicates 10%, 5%, 1% statistical significance respectively.

Table 5 presents the results for the first test. The test statistic is the difference in squared Sharpe ratio and is a sequential test for nonnested models. In our case, the difference is positive, it indicates that the FF3* model has higher squared Sharpe ratio than the FF3 model. The difference is statistically significant a 1%. Hence, based on this test, the newly constructed FF3* dominates the traditional FF3 model.

	Panel A: OLS
Model	FF3*
FF3	0.036
	(0.449)
	Panel B: GLS
Model	FF3*
FF3	0.031
	(0.475)

Table 6: This table shows the pairwise test of equality of the R^2 of the asset pricing model. Panel A shows the results using OLS while Panel B shows the results using GLS. The weighted matrix is the same as in Table 7, namely the covariance matrix of the factors. The test statistic is computed using the entire set of test assets. The test statistic is $\hat{\rho}_i^2 - \hat{\rho}_j^2$ and the associated p-value in parentheses. The null hypothesis is $H_0 : \hat{\rho}_i^2 = \hat{\rho}_j^2$. The p-value is computed under the assumption that the model is potentially misspecified. To perform the normal test as proposed in Kan et al. (2013). This approach rules out the scenario that the additional factors in both models are irrelevant and that both models are perfectly specified. $*, ***, ***$ indicates 10%, 5%, 1% statistical significance respectively.

Table 6 shows the results based on the difference of R^2 using OLS in panel A and GLS in panel B. I report additional details of this test in Table 7, namely, the R^2 and the corresponding p-values.

The test statistic represents the difference in R^2 (row-column). It is positive as the R^2 in the Fama-McBeth cross-sectional regression using the full set of test assets is larger for

FF3 model than for FF3*. The difference slightly decrease using GLS as shown in panel B. The p-values are very large and clearly indicate that the null hypothesis of the equality of the R^2 is not rejected. Meaning that the two models are equivalent. This is in line with the results in Table 4 as the F-statistics were very close.

Thus, based on this test, we cannot say that one model is better than the other.

Conclusion

I evaluated the pricing power of the two asset pricing models based on three tests. The first is based on the time series regression intercept. As stated above, this is not an appropriate test for comparing models, but it does provide some insight on which model is the best. The results of this first part based on the individual test asset intercepts is ambiguous. Relying on the joint intercept test suggested that the FF3* model is likely to do a better job.

In the second part, on the one hand, the test based on the squared Sharpe ratio clearly indicates that the FF3* model dominates the FF3 model. On the other hand, the test based on the cross section R^2 indicates that the models are not statistically different from each other. However, the latter is dependent on the test assets choice. In my opinion, the former test is more robust because it does not rely on the test assets assumption.

Finally, one possible explanation for the better results of FF3* compared to the standard version could be due to the fact that the risk factors that are mimicked by the SMB and HML portfolios can vary over time. Thus, a portfolio that is updated more frequently is better able to capture this temporal variation. In addition, because the breakpoints are further apart in the corners for the FF3* model, this captures the same underlying source of risk but in a more accentuated manner.

However, one should keep in perspective that the test results comparing the two models are ambiguous. Although the Sharpe squared ratio indicates that the $textFF3^*$ is better than the FF3. However, the test based on the R^2 indicates that they are not different from each other. Thus, the two constructs may also be equivalent, suggesting that both capture risk factors in the same way.

References

F. Barillas and J. Shanken. Which alpha? *The Review of Financial Studies*, 30(4): 1316–1338, 2017.

F. Barillas, R. Kan, C. Robotti, and J. Shanken. Model comparison with sharpe ratios. *Journal of Financial and Quantitative Analysis*, 55(6):1840–1874, 2020.

Q. F. S. Drechsler. Python programs for empirical finance, 2022. URL https://www.fredasongdrechsler.com/data-crunching/fama-french.

E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.

M. R. Gibbons, S. A. Ross, and J. Shanken. A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152, 1989.

R. Kan, C. Robotti, and J. Shanken. Pricing model performance and the two-pass cross-sectional regression methodology. *The Journal of Finance*, 68(6):2617–2649, 2013.

Appendix

A

Panel A: OLS		
Model	FF3	FF3*
$\hat{\rho}^2$	0.496	0.461
$p(\hat{\rho}^2 = 1)$	0.001	0.001
$p(\hat{\rho}^2 = 0)$	0.08	0.096
$se(\hat{\rho}^2)$	0.203	0.196
No. of para.	3	3
Panel B: GLS		
Model	FF3	FF3*
$\hat{\rho}^2$	0.132	0.101
$p(\hat{\rho}^2 = 1)$	0.000	0.000
$p(\hat{\rho}^2 = 0)$	0.004	0.019
$se(\hat{\rho}^2)$	0.065	0.058
No. of para.	3	3

Table 7: This table presents the cross-sectional R^2 of both asset pricing model using the test assets. $p(\hat{\rho}^2 = 1)$ is the p-value for the test $H_0 : \hat{\rho}^2 = 1$; $p(\hat{\rho}^2 = 0)$ is the p-value for the test $H_0 : \hat{\rho}^2 = 0$; $se(\hat{\rho}^2)$ is the standard errors of $\hat{\rho}^2$ under the assumption $0 < \hat{\rho}^2 < 1$. Panel A shows the results using OLS. Panel B shows the results using GLS. The Weighted matrix is the covariance matrix of the factors.