Comparing Global Factor Models with Sharpe Ratios

Shiqi Wang[1,a,*]

[1]*College of Electrical and Control Engineering, North China University of Technology, Jinyuanzhuang Road, Shijingshan District, Beijing, China*
a.892137457@qq.com
**corresponding author*

Keywords: sharpe ratio test, factor model, model comparison

Abstract: Global factors today matter in the age of globalization and the associated integrated global capital markets. In this paper, we use the GRS test and the Sharpe ratio approach of Barillas et al. (2019) to compare the global versions of ten prominent traded-factor models. We find that the best performing models are the three six-factor models of Fama and French (2018) and Aseness et al. (2015), which all include market excess return, size, value, investment, profitability, and momentum factors. This paper contributes to the literature by comparing more models based on the Sharpe ratio tests that provide more economic significance.

1. Introduction

Since the world has become an integrated capital market, compared with the domestic factors, the global factors are the relevant risk factors. There are developed global versions of domestic factor models such as the Capital Asset Pricing Model (CAPM) and empirical factor models such as Fama and French (1998) so on [1]. Recently, numerous different factor modes are advocated by recent studies. Thus, the purpose of our paper is comparing the global version of several of the most prominent factor models.

In terms of comparing models, differentiated models need to be treated in different ways[2]. According to Barillas and Shanken (2017), the ability to explain the factor in each other is the most important and test assets are irrelevant. As for nested models, we only need to concentrate on testing the excluded factor restriction, which could be formally evaluated by the basic alpha test. When it comes to the non-nested models, it is more complex, and appropriate methods for comparison are required. Fletcher (2018) used the Bayesian method of Barillas and Shanken (2018) to conduct model comparison for global factor models. We complemented his comparison by taking a Sharpe ratio approach. Barillas et al. (2019) derived the asymptotic distribution of the difference between two models' sample squared Sharpe ratios, based on which we can conduct hypothesis testing for model comparison. Sharpe ratio is widely used in finance and tests based on Sharpe ratio add more economic significance compared to the Bayesian method.

Fletcher (2018) motivated us when selecting models used in the comparison. The models selected in the models comparison contains the CAPM, notable three-factor model (FF3) proposed by Fama and French (1993), the four-factor model of Carhart (1997) extended momentum factor to FF3, the five-factor model of Fama and French (2015), and the six-factor models of Fama and French (2018)

(FF6) and Asness et al. (2015) (AFIM) that combine momentum and the Fama-French factors[3]. Besides, Fama and French (2018) modified their models to use only the small ($FF6_s$) or the big spread factors to replace several primary factors, such as the profitability, value, momentum and investment factors.

We could conclude that the best performing models were the FF6, $FF6_s$ and AFIM models. They outperform other competing models siginicantly while the differences between their Sharpe ratios were not significant.

The rest of the paper is organized as follows. Section 2 describes the research method used in the study. Section 3 discusses the factors and global models and empirical results. Section 4 concludes the paper with a conclusion.

2. Comparing Sharpe Ratios for Models with Traded Factors

Firstly, let us take a look at some definitions and notations of the GRS test. M can represent a factor model, which is a linear regression with N excess returns, R, and K traded factors, f, which includes several variates[3]. And this factor model is also with T observations on f_t and R_t:

$$R_t = \alpha_R + \beta f_t + \epsilon_\tau, \tau = 1, \dots, T \tag{1}$$

where R_t, ϵ_t and α_R are N-vectors, β is an $N \times K$ matrix, and f_t is a K-vector. GRS test could tell the improvement of the squared Sharpe ratio when the investment universe, including test assets R, is a quadratic form in the test-asset alphas:

$$\alpha'_R \Sigma^{-1} \alpha_R = Sh^2(f, R) - Sh^2(f) \tag{2}$$

where the zero-mean disturbance ϵ_t 's invertible population covariance matrix is Σ. Then, related F-statistic is proportionate to the statistics coming from replacing the sample quantities in (2) and divided by one plus the sample estimate of Sh^2(f). Therefore, the test of $\alpha_R = 0_N$, where 0_N is an N-vector of zeros, shows whether f yields the maximum squared Sharpe ratio[3].

And then, focus on the pricing restrictions of the nested models and how to carry out the GRS test when the factors are not included in the nested models and serve as left-hand-side returns at the same time.

2.1. Model Comparison

Let A represent the pricing model that consists of factors $[f'_{1t}, f'_{2t}]'$. Model A nests model B with factors f_{1t}. α_{21} denote the alphas of regressing f_{2t} on f_{1t}, where f_{1t} and f_{2t} are K1 and K2-vectors, respectively. We could conclude the proposition 1 in Barillas and Shanken (2017) that only testing the excluded-factor restriction, which could be formally evaluated by the basic alpha test, could help us compare nested models, $\alpha_{21} = 0_{K_2}$, where test assets are irrelevant[4].

However, in terms of comparing non-nested models, it is less straightforward. Thus, we would use direct asymptotic tests to compare non-nested models. When we focus on A and B, which represent two non-nested models, consisting of f_{At} and f_{Bt}, respectively, we could suppose that every time series is jointly invariable with finite fourth moments. This part consists of returns from traded-factor and subsequently, returns from nontraded factors and other basis-assets[5]. The two equations, $\theta_A^2 = \mu'_A V_A^{-1} \mu_A$ and $\theta_B^2 = \mu'_B V_B^{-1} \mu_B$, represent the highest squared Sharpe ratios that can achieve from the two sets of factors , where V_A, V_B, μ_A and μ_B respectively represent the two sets of factors' invertible covariance matrices and nonzero means. Similarly, $\ddot{\theta}_A^2 = \mu'_B \hat{V}_B^{-1} \mu_A$ and $\ddot{\theta}_B^2 = \mu'_B \hat{V}_B^{-1} \mu_B$

denote corresponding sample quantities. And thus, we can draw the conclusion that PROPOSITION 1 in the Model Comparison with Sharpe Ratios that the asymptotic distribution of the difference in sample squared Sharpe ratios is given by

$$\sqrt{T}\,([\ddot{\theta}_A^2 - \ddot{\theta}_B^2] - [\theta_A^2 - \theta_B^2]) \overset{A}{\sim} N(0, E[d_t^2]) \qquad (3)$$

provided that $E[d_t^2] > 0$, where

$$d_t = 2(\mu_{At} - \mu_{Bt}) - (\mu_{At}^2 - \mu_{Bt}^2) + (\theta_A^2 - \theta_B^2) \qquad (4)$$

with $\mu_{At} = \mu_A' V_A^{-1}(f_{At} - \mu_A)$ and $\mu_{Bt} = \mu_B' V_B^{-1}(f_{Bt} - \mu_B)$. And we could use this distribution to do hypothesis testing, and thus, examine the different value and the p-value of the global version.

3. Factor Models

There are ten global factor models used in our main empirical test. The test also includes 15 (K =15) distinct factors for all the models, thus, each model can be thought as a subset of the 15-factor model. The models consist of.

i. CAPM

The CAPM model consists of a single factor, Market. The Market factor is the value of the value-weighted market return minus the one-month U.S. Treasury bill rate.

ii. Fama and French (1993) (FF3)

The FF3 model consists of three factors. In addition to the factor in the CAPM model, the model adds the SMB factor, which is the value of the small minus big size factor, and the HML factor, which is the value of high minus low book-to-market value factor of Fama and French (1993)[6].

iii. Carhart (1997)

The Carhart model consists of four factors. The factors extend from the WML factor, which is the value of the high minus low momentum factor of Carhart (1997), to factors in the FF3 model.[7]

iv. Fama and French (2015) (FF5)

The FF5 model consists of five factors. The factors include the factors in the FF3 model and RMW, which is the value of the robust minus weak cash profitability factor of Fama and French (2017) and CMA, which is the investment factor of conservative minus aggressive investment factor of Fama and French (2015). We have to mention that the SMB factor in the FF5 model is used as the size factor across all models.[8]

v. Fama and French (2018) ($FF5_s$)

The $FF5_s$ model consists of five factors. The factors include factors used in FF5, but use the small ends of the CMA, RMW and HML[9]. Thus, the factors are presented as CMA_S, RMW_S and HML_S.

vi. Fama and French ($FF5_b$)

The FF5b model consists of five factors. Similar to the construction of the $FF5_s$ model, it contains the same five factors used in FF5 but uses the big ends of the CMA, RMW and HML[8]. Thus, the factors are listed as CMA_b, RMW_b and HML_b.

vii. Fama and French (2018) (FF6)

The FF6 model consists of six factors. The factors include the factors in the FF5 model and add the WML factor into it.

viii. Fama and French (2018) ($FF6_s$)

The $\mathrm{FF6}_s$ model consists of six factors. The factors include the factors in the $\mathrm{FF5}_s$ model added with the WML_s factor.

ix. Fama and French ($\mathrm{FF6}_b$)

The $\mathrm{FF6}_b$ model consists of six factors. It uses the same factors used in FF6 but includes the big ends of the HML, RMW, CMA, and WML, thus, the factors are denoted as HML_B, $\mathrm{RMW_B}$, $\mathrm{CMA_B}$, and WML_B.

x. Asness et al. (2015) (AFIM)

The AFIM model consists of six factors. Our final model includes the more-timely value factor, HML_T, from Asness and Frazzini (2013), instead of the standard HML. This factor is different from the usual HML factor which would use annually updated lagged prices value[10]. It is on the foundation of book-to-market rankings which use the most recent monthly stock price in the denominator. The HML_T factor is collected from the AQR data library.

Table 1: Summary statistics of factors.

Factor	Mean	Standard deviation	t-statistics
Market	0.51	4.16	2.28
SMB	0.09	1.90	0.88
HML	0.27	2.32	2.16
RMW	0.35	1.45	4.40
CMA	0.19	1.86	1.88
WML	0.60	3.83	2.88
$\boldsymbol{HML_T}$	0.28	2.87	1.82
$\boldsymbol{HML_s}$	0.48	2.61	3.42
$\boldsymbol{HML_b}$	0.06	2.54	0.42
$\boldsymbol{RMW_s}$	0.36	1.46	4.52
$\mathbf{RMW}_{\boldsymbol{b}}$	0.34	1.99	3.12
$\boldsymbol{CMA_s}$	-0.30	1.73	-3.17
$\boldsymbol{CMA_b}$	-0.08	2.27	-0.67
$\boldsymbol{MOM_s}$	0.79	3.69	3.97
$\boldsymbol{MOM_b}$	0.40	4.29	1.73

Table 1 lists the summary statistics for our monthly factor returns – means, standard deviations, and t-statistics in global stock returns. The sample period for our data is from November 1990 to July 2017. It shows that some factors have significant positive and sizable mean value and average returns.

Among all the values, the MOM_s factor has the highest average excess return. The MOM_s factor's higher mean excess return compared with the WML factor confirms that the momentum effect is stronger among smaller companies in global stock returns. Relatively, the MOM_b factor has an insignificant average excess return. And the SMB size factor also has an insignificant average excess return. The HML and HML_s factors both have significant positive average excess return, confirming the strong value effect in global stock returns. Relatively, the HML_T and HML_b factors are not significant. The HML_s factor's higher mean excess return compared with the HML factor confirms its stronger value effect among smaller companies in global stock returns. The significant mean excess returns of the RMW, RMW_s, RMW_b, and CMA_s factors confirm the significant profitability and investment effects in global stock returns. The profitability effect is less noticeable to reflect whether it would be stronger among smaller companies in the mean excess return of the RMW,

RMW_s and RMW_b factors. The CMA_s factor has a significant negative average excess return. Relatively, the CMA and CMA_b factor has an insignificant average excess return.

3.1. Empirical Results

Table 2: Test of mean-variance efficiency.

Panel A: The difference value of the squared Sharpe ratio									
	FF3	CARHART	FF5	$FF5_s$	$FF5_b$	FF6	$FF6_s$	$FF6_b$	AFIM
CAPM	-0.02	-0.07	-0.15	-0.12	-0.10	-0.18	-0.22	-0.10	-0.22
FF3		-0.05	-0.14	-0.10	-0.08	-0.16	-0.20	-0.08	-0.20
CARHART			-0.08	-0.05	-0.03	-0.11	-0.15	-0.03	-0.15
FF5				0.03	0.05	-0.02	-0.07	0.05	-0.07
$FF5_s$					0.02	-0.06	-0.10	0.02	-0.10
$FF5_b$						-0.08	-0.12	0.00	-0.12
FF6							-0.04	0.08	-0.04
$FF6_s$								0.12	0.00
$FF6_b$									-0.12
Panel B: p-value									
	FF3	CARHART	FF5	$FF5_s$	$FF5_b$	FF6	$FF6_s$	$FF6_b$	AFIM
CAPM	0.033	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FF3		0.000	0.000	0.001	0.025	0.000	0.000	0.030	0.000
CARHART			0.063	0.222	0.457	0.000	0.000	0.430	0.000
FF5				0.270	0.012	0.008	0.110	0.017	0.032
$FF5_s$					0.627	0.148	0.000	0.622	0.016
$FF5_b$						0.004	0.014	0.389	0.001
FF6							0.130	0.001	0.072
$FF6_s$								0.009	0.995
$FF6_b$									0.000

Table 2 reports pairwise tests of equality of the squared Sharpe ratios of the ten asset-pricing models. We report in Panel A the difference value of the squared Sharpe ratios of the models in row i and column j, $\ddot{\theta}_i^2 - \ddot{\theta}_j^2$, and in Panel B the associated p-value (in parentheses) for the test of H_0: $\theta_i^2 = \theta_j^2$.

With the discussion presented in the 2.1, p-value has a significant difference in nested models and non-nested models. When the models are nested, we focus on the factor which is in the larger model and not included in the smaller model, having zero alphas when regressed on the smaller one. For example, the Fama and French (1993) model (FF3) which consists of the Market, SMB and HML factors is nested in the Carhart (1997) model which consists of the Market, SMB and WML factors, thus, the p-value showed in the panel B is in the regression of WML on the Fama and French (1993) model (FF3).[10]

When it comes to non-nested models, the sequential test would be used. Firstly, check the difference value of the squared Sharpe ratio between the model consisting of the common factors and the model consisting of all the factors in both models that are different from zero. The outcome can show whether it is the zero alphas of the non-common factors on the common factors. If the outcome is zero, there is evidence that the model which consists of common factors is as good as the model which is added with the non-overlapping factors. And then we have evidence that the two models are equivalent as well. However, if the outcome is significantly less than zero, it means that the row model has a weaker performance than the column model. Thus, we use the direct test in Proposition 1 from Model Comparison with Sharpe Ratios listing the p-value for the tests of equality of the squared Sharpe ratios to examine whether the squared Sharpe ratio of the non-nested model is different.

When the difference value presented in the panel A between the row model and the column value is negative, and the p-value presented in the panel B is less than 0.1, it means that the column model outperforms the row model. Especially when the p-value is less than 0.01, it means that the column model significantly performs better than row model. Thus, the CAPM model is outperformed by all other models. The FF3 model outperform all the other models except the CAPM model, with all the number of the row in panel B is less than 0.01, while the number with the $\mathrm{FF5}_s$ and $\mathrm{FF6}_b$ factors is less than 0.05. The FF5, FF6, $\mathrm{FF6}_s$, and AFIM model perform better than the Carhart model, since the results in panel B are all less than 0.01, except the one with FF5 is less than 0.1. And the difference between the Carhart model and the $\mathrm{FF5}_s$, $\mathrm{FF5}_b$, and $\mathrm{FF6}_b$ model is not statistically significant. The FF6 and AFIM model has better performance than the FF5 model since the number shown in panel B with the FF6 model is less than 0.01 and the one with the AFIM model is less than 0.05. The difference between the FF5 model and the $\mathrm{FF5}_s$ and $\mathrm{FF6}_b$ model are not statistically significant. The $\mathrm{FF6}_s$ and AFIM model could do better than the $\mathrm{FF5}_s$ model, for the number in the position of the column of the FF6 model and the same row is less than zero. And the one in the column of the AFIM model is less than 0.05. The differences between the $\mathrm{FF5}_s$, $\mathrm{FF5}_b$, FF6, and $\mathrm{FF6}_b$ models are not statistically significant. The FF5, FF6, $\mathrm{FF6}_s$, and AFIM models perform better than the $\mathrm{FF5}_b$ model because their results in panel B are all less than 0.01, except the results with the FF5 and $\mathrm{FF6}_s$ model are less than 0.05. And the difference between the $\mathrm{FF5}_b$ model and the $\mathrm{FF5}_s$ and $\mathrm{FF6}_b$ model is not statistically significant. The difference between the FF6 model and the $\mathrm{FF6}_s$ model and the one between the $\mathrm{FF6}_s$ model and the AFIM model are not statistically significant. The FF5, FF6, $\mathrm{FF6}_s$ and AFIM models perform better than the $\mathrm{FF6}_b$ model because all their results in panel B are less than 0.01.

4. Conclusion

With all the discussion above, we can conclude that the best performing models are the FF6, $\mathrm{FF6}_s$ and AFIM models. This paper made two contributions to the literature. Firstly, we extended the comparison in the Fletcher (2018) by adding the big spread factors for the value, profitability, investment and momentum factors. Second, we found the FF6, $\mathrm{FF6}_s$, and AFIM were the best performing models among others by using the Sharpe ratio method of Barillas et al. (2019), which added economic significance to the comparison and complements of related research.

References

[1] Asness, C., Frazzini, A., Israel, R., & Moskowitz, T. (2015). Fact, fiction, and value investing. The Journal of Portfolio Management, 42(1), 34-52.

[2] Barillas, F., Kan, R., Robotti, C., & Shanken, J. A. (2017). Model comparison with Sharpe ratios. Journal of financial and quantitative analysis, forthcoming.

[3] Barillas, F., & Shanken, J. (2018). Comparing asset pricing models. The Journal of Finance, 73(2), 715-754.

[4] Barillas, F., & Shanken, J. (2017). Which alpha? The Review of Financial Studies, 30(4), 1316-1338.

[5] Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of finance, 52(1), 57-82.

[6] Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of financial economics, 116(1), 1-22.

[7] *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.*
[8] *Fama, E. F., & French, K. R. (2018). Choosing factors. Journal of Financial Economics, 128(2), 234-252.*
[9] *Fama, E. F., & French, K. R. (1998). Value versus growth: The international evidence. The journal of finance, 53(6), 1975-1999.*
[10] *Fletcher, J. (2018). Bayesian tests of global factor models. Journal of Empirical Finance, 48, 279-289.*