Asset Pricing Models Comparison Based on Bayesian Method

ZijunWang[1,a,*]

[1]*School of Economics, Renmin University of China, Zhongguan Cun Street, Beijing, China*
a. zijunwang@ruc.edu.cn
**corresponding author*

Keywords: asset pricing model, Bayesian method, candidate factors, categorical models, factor models

Abstract: To figure out a proper asset pricing factor model among a bunch of candidate factors and anomalies has been always a heated field. This paper applies the Bayesian method proposed by Barillas and Shanken (2018) to investigate the optimal factor models under a more comprehensive collection of candidate factors as well as candidate models. Given a collection of 13 candidate nonmarket factors and asset returns data of U.S. stock market from 1980 to 2018, this paper compares a total of 810 individual models and 64 category models under the same prior specification of a maximum Sharp ratio multiple. This paper also ranks the candidate individual factors and categorical factors based on their posterior probabilities. The results show that the best individual model is the six-factor model {Mkt, UMD, SMB, HMLm, ROE IA} and the optimal categorical models is {Mkt SIZE VALUE PROF INV}. In addition, factors IA and ROE showcased highest predictive power. The results are robust under different prior assumption.

1. Introduction

Asset pricing literature has explored various factors since the three-factor model of Fama and French (1993) that includes market, size, and value factors. Factors to explain stock excess return are always identified firstly by anomalies recognition. Harvey et al. (2016) presented a list of 316 anomalies as the potential factors for asset pricing. Fama and French (2015) categorized those anomalies as five categories as to Valuation, Size, Investment, Profitability and Momentum. To find more factors to explain risk premium, anomalies are broadly utilized to recognize potential factors for asset pricing. Based on anomalies analysis, additional factors of Investment, Profitability, Momentum and Reversal has been developed and proved to be significantly correlated with stock excess return in asset pricing models. Two challenges arose at the same time. One is that how to combine those factors to formulate a best model. The other is that which variables are the best proxies with highest predictive power for each category factors to make a best prediction. Given a variety of candidate factors which might be extended with more anomalies recognized in the future, a satisfactory statistical methodology of

identifying the best models and comparing pairwise models is of great significance (Barillas and Shanken, 2018).

Two approaches are available for investigation on these two questions. As for determination of proper combination for factors, a classic test to evaluate factor models is to test whether the alphas are significant given the implication that market portfolio is efficient. This way to evaluate factor models was standardized by Gibbons, Ross and Shanken (1989) by proposing a joint F-test, henceforth GRS, based on which the hypothesis is that intercept of models should be zero. Traditional statistical method as to select the most effective factors is to use the remain-to-test factor to regress on other factors, if the alpha is significantly far away from zero, then this factor should be added into the models (Fama and French, 2015, 2016, 2017, 2018).

However, with more factors to be considered into models are developed and larger data set to be tested, GRS suggests "rejection" for hypothesis more often (Fama and French, 2016). Therefore, the efficiency of GRS test should be considered again. Barillas and Shanken (2018) proposed that a very large p-value may imply more about the imprecision about alpha than the model's accuracy. De Moor et al. (2015) suggested that the p-value estimation might be influenced by estimation precision across models, which implies that utilization of GRS to compare precision of models would be faced with unexpected obstacles.

Another approach of Bayesian method to test alpha has been proposed firstly by Shanken (1987b). Shanken developed a Bayesian test in his paper to evaluate the mean-variance efficiency of a portfolio by directly linking the prior belief about relative efficiency to the odds in favour of efficiency. Harvey and Zhou (1990) complimented the procedure of Shanken by using both a diffuse prior and an informative prior to calculate posterior-odds ratio which therefore can provide a prior belief for all parameters, as opposed to Shanken's method which can only impose a prior belief on the function. Barillas and Shanken (2018) adopted suggestion of Harvey and Zhou to utilize the diffuse prior and derived a Bayesian asset pricing test that requires a prior judgment about the magnitude of plausible model deviations. They applied Bayesian test to the model comparison based on a given variety of candidate factors proposed by Hou, Xue, and Zhang (2015a, 2015b) and Fama and French (2015, 2016) and the best model with the highest posterior probability is the six-factor model {Mkt IA ROE SMB HMLm UMD}.

To examine more comprehensively the predictive power of individual factors among same categorical factors and to investigate the optimal combination of pricing factors, this paper adopts the relative comparison method proposed by Barillas and Shaken (2018) to further test a total of 13 candidate nonmarket factors and compare 810 candidate models based on U.S. stock market data over thirty years period. The rest of paper is organized as follows. The second section presents the methodology for calculating the posterior probabilities and comparing models. We conduct the Bayesian method on U.S. stock market to compare different individual models and categorical models in section three. Section four presents the empirical results. Section five gives the final conclusion.

2. Methodology

The empirical application of Bayesian approach lagged behind classical statistical method for many years was the result of two critical obstacles facing Bayesian estimation method (Harvey and Zhou, 1990). One difficulty is how to choose a proper prior distribution, and the other is how to calculate posterior distribution. In this section, we introduced in turn the prior specification and posterior probabilities for calculating model unconditional probabilities.

2.1. Prior Specification

Given a multivariate pricing model

$$r_t = \alpha + \beta f_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \Sigma) \tag{1}$$

ε_t is independent over time. The diffuse prior for β and sample F is

$$P(\beta, F) \propto |\Sigma|^{-\frac{N+1}{2}} \tag{2}$$

We assume that alpha is normally distributed conditional on β, F

$$P(\alpha|\beta, F) \sim MVN(0, k\Sigma) \tag{3}$$

where k reflects our prior belief about the extent the alpha correlated with residual variance. In addition,

$$\widehat{\alpha}'\Sigma^{-1}\widehat{\alpha} = {sh_{max}}^2 - sh^2 \tag{4}$$

and $\widehat{\alpha}'(k\Sigma)^{-1}\widehat{\alpha}$ is distributed as Chi~Squre with freedom degrees of N. Therefore, given a target maximum Sharp ratio multiple, the required k can be calculated with

$$k = E\left(\frac{{sh_{max}}^2 - sh^2}{N}\right) \tag{5}$$

2.2.Posterior Probabilities

Posterior probabilities for each model in this paper are utilized to compare models with different factors. Given a collection of candidate factors, the (L-1) test nonmarket factors are denoted by f. The models with (Mkt, f) are denoted by M. The factors excluded by M is denoted by w^*. For each model with prior model probability of $P(M_j)$, the posterior probability conditional on data F is given by

$$P(M_j|D) = \{ML_j \times P(M_j)\} \times \left\{\sum_i ML_i \times P(M_i)\right\}, \tag{6}$$

where ML_j is the marginal likelihoods and F is sample for all factors and assets returns. The ML is given by

$$ML = ML_U(f|Mkt) \times ML_R(f^*|Mkt, f) \times ML_R(r|Mkt, f, f^*) \tag{7}$$

Marginal likelihoods of model M conditional on w:

$$ML(M|w) = ML_U(f|Mkt) \times ML_R(f^*|Mkt, f) \times ML_R(r|Mkt, f, f^*) \tag{8}$$

M|w refers to uniform prior over models in w. Unconditional probability of data:

$$P(F) = E_w\{ML(w)\} \tag{9}$$

By averaging ML(w) over uniform prior P(w). The posterior probability for w is:

$$\mathrm{P(w|F)} = \mathrm{ML(w)} \times \frac{\mathrm{P(w)}}{\mathrm{P(F)}} \tag{10}$$

Combine the above functions, we get:

$$\begin{aligned} \mathrm{P(M|F)} &= \sum_w P(w|F) \times P(M|w,F) \\ &= \sum_w \{ML(w) \times P(w)/P(F)\} \times P(M|w,F) \\ &= \sum_w \frac{ML(w)}{\sum_w ML(w)} \times \frac{ML}{\sum_w ML} \end{aligned} \tag{11}$$

Where

$$\mathrm{ML} = \mathrm{ML_U(f|Mkt)} \times \mathrm{ML_R(f^*|Mkt, f)} \tag{12}$$

Therefore, $\sum_{\mathrm{M}} \mathrm{P(M|F)} = 1$, this means that the probabilities for a total of test asset pricing models under the same prior specification would sum up to one.

3. Comparing Models with Categorical Factors

Barillas and Shanken (2018) refers category factors to the factors that include more than one measurement opposed to standard factors that only have one measurement. The models with category factors are category models opposed to individual models that only consider individual factors. For example, for Value factor, there are two measurements of HML and HMLm (Asness and Frazzini, 2013) to calculate Value. A bunch of test models can be formulated by the combination between category factors and individual factors. We consider a total of 14 candidate factors including one market factor and 13 non-market factors which can be categorized into five categories. For *Size*, we consider SMB and ME (Hou, Xue, and Zhang, 2015a); for *Value*, we consider HMLm (Asness and Frazzini, 2013) and three types of HML constructed respectively by B/M, E/P and CF/P; for *Profitability*, we consider RMW and ROE; for *Investment*, we consider CMA and IA; for *Reversal*, we consider LR (long-term reversal factor) and SR (short-term reversal factor). We also consider the standard momentum factor of UMD. We construct factors in a way similar to Fama and French (2015). This paper creates each factor by interacting them with size based on 2X3 sorts of portfolio. For example, HML is constructed by interacting size with value to create a 2X3 portfolio and HML is the average of two high-value portfolios returns minus the average of two low-value portfolio returns; RMW is the average of two high-profitability portfolio returns minus two low-profitability portfolio returns; the same construction method goes to the rest of factors respectively.

In all, we consider two standard factor of Mkt and UMD, and five categorical factors of *Size*, *Value, Profitability, Investment* and *Reversal*. Since each model has up to seven factors and Mkt is always included, a total of 64 categorical models can be formulated. Consider each categorical model can adjust their categorical factors; therefore, we have a total of 810 individual models. Our benchmark of prior specification assumes that maximum Sharp ratio equals 1.5 times market sharp ratio.

4. Empirical Results

In this section, we present the result of model comparison for the sample period from 1980 to 2019. Model probabilities are shown at each point in time to provide a historical perspective on how posterior beliefs of model have evolved over time. Therefore, we use monthly data from January 1980 up to the given point of time to estimates unconditional probabilities for each model. The unconditional probabilities for each model are presented in the figure1. Since we test up to 810 models for each period of sample, the sum of unconditional probabilities for all test models equals one, each model therefore finally got a very low probability. However, the empirical result can still rank out the fit model with the highest unconditional probability.

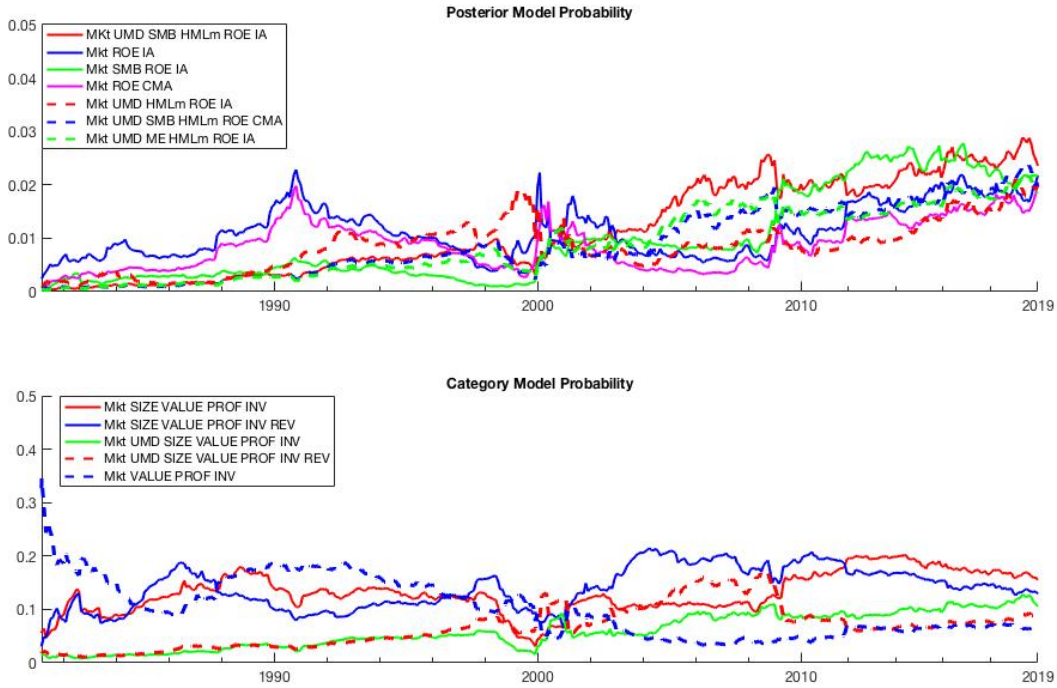

Figure 1: Individual Model Probabilities and Categorical Model Probabilities.

Under the prior belief that maximum sharp ratio is 1.5 times market sharp ratio, the best two models with highest unconditional probability is a six-factor model with{Mkt, UMD, SMB, HMLm, ROE IA} and a four-factor model with {Mkt, SMB, ROE IA} which excludes value categorical factor and momentum categorical factor compared to the six-factor model. The above two models stand out with a much higher probability than the other models. These two models have no significant difference since 2010, while the six-factor model beat the four-factor model with a much higher probability from 2003 to 2009. The remaining five models of top seven are one four-factor model with {Mkt, SMB, ROE, IA}, one three-factor model with {Mkt, ROE, CMA}, one five-factor model with {Mkt, UMD, HMLm, ROE, IA}, and two six factor models with {Mkt, UMD, SMB, HMLm, ROE, CMA} and {Mkt, UMD, ME, HMLm, ROE, IA} respectively. Among the best seven models, all include Profitability and Investment categorical factors; for *Profitability*, all models include ROE; for *Investment,* five of seven include IA; four out of seven models include *Value* factors and they all choose HMLm. All of models don't include any *Reversal* factors. The second panel of figure 1 gives another perspective on probabilities of categorical models by aggregating posterior probabilities for

different versions of each categorical model. The posterior probabilities for categorical models are much more stable than that of individual models. Standing at the latest point, the best category model is the four factor model with {Mkt, SIZE, VALUE, PROF, INV}, which is closely followed by two five factor categorical models each adding REV and UMD respectively.

Figure 2 present the exploration of factor probabilities. The first panel of figure 2 shows the cumulative categorical factor probabilities, that is, the sum of posterior probability of model that include this categorical factor. The cumulative probabilities for *Profitability* and *Investment* close to one, with *Size*, *Value* and *Reversal* around 0.8, respectively, while *Momentum* only about 0.5. The second panel of figure 2 gives cumulative probabilities for each factor. ROE and IA ranked first two both with much higher cumulative probabilities than other factors, which corresponds to the model ranking showing that all seven models include ROE and five models include IA. Combined with first panel of figure 2, we can conclude that categorical factors of *Profitability* measured as ROE and categorical factor *Investment* measured as IA is significant for explaining pricing model. For *Size*, the probabilities of SMB and ME are close to each other; for *Value*, HMLm beat over other three factors less than 0.3 with a much higher probability; for *Reversal*, the probability of LR is higher than SR, but both of them are lower than 0.3.

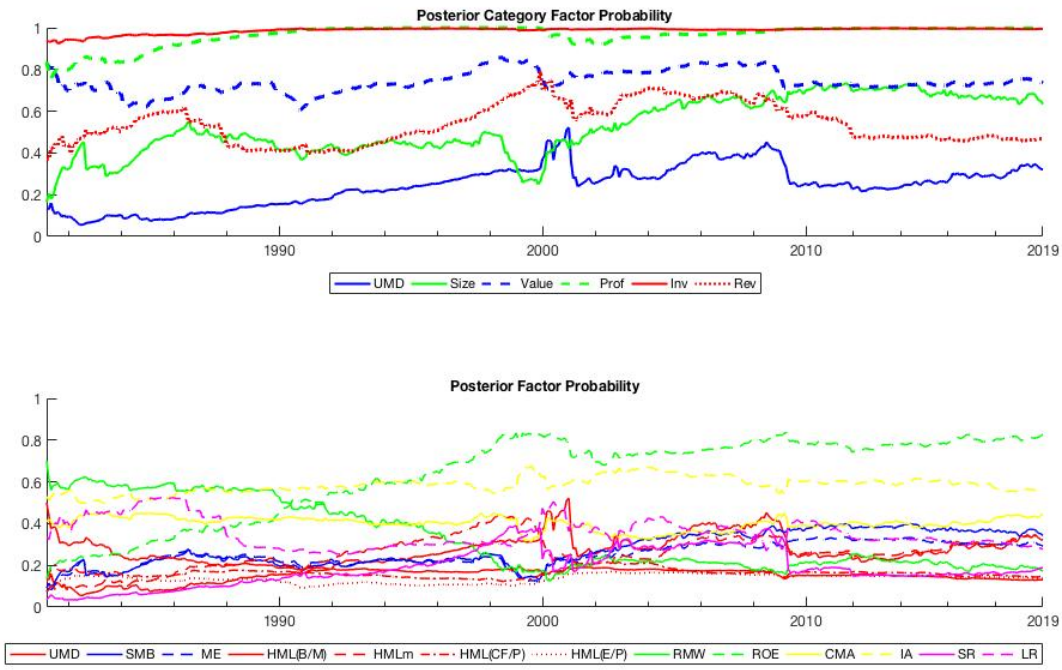

Figure 2: Cumulative Factor and Categorical Factor Probabilities.

The previous analysis is based on the prior assumption that maximum Sharp ratio is 1.5 times sample market Sharp ratio. We next examine the sensitivity to the prior Sharp multiples of posterior probabilities (in per thousand) for the top seven models under the assumption of Sharp multiples as 1.5. Table 1 present the result of posterior probabilities under Sharp multiples of 1.25, 1.75 and 2.0 for the top seven models when prior Sharp multiples is set as 1.5. The best model {Mkt, UMD, SMB, HMLm, ROE IA} still ranks highest under different multiples specification. However, the rankings of the other six top models changes when Sharp multiples rise to 1.75 and 2. The model {Mkt UMD

SMB HML ROE CMA} comes to the second place in opposed to {Mkt SMB ROE IA}. At the same time, when multiples increase, the probabilities for models are less spread out.

Table 1: Prior Sensitivity For Models Probabilities.

Individual Models	1.25	1.5	1.75	2
Mkt, UMD, SMB, HMLm, ROE IA	22.82	23.57	30.79	35.21
Mkt, ROE and IA	22.82	21.79	22.66	24.82
Mkt, SMB, ROE, IA	16.25	21.62	25.77	29.33
Mkt, ROE, CMA	20.52	20.52	21.81	24.07
Mkt, UMD, HMLm, ROE, IA	13.98	19.99	23.03	25.39
Mkt, UMD, SMB, HMLm, ROE, CMA	10.10	19.79	25.62	29.12
Mkt, UMD, ME, HMLm, ROE, IA	9.96	18.41	23.21	25.95

The top model rankings for categorical models are also robust across different prior multiple specification. The best categorical model {Mkt SIZE VALUE PROF INV} always ranks the first under different prior multiples. The probabilities for the best categorical model are also stable with different prior Sharp multiples. Among the five top categorical models, {SIZE VALUE PROF INV} are included in each model except the fifth which excludes SIZE. The probabilities for categorical models with REV but without UMD decline as Sharp multiple rises. While the probabilities for the categorical models with UMD but without REV increase as multiple rises.

Table 2: Prior Sensitivity for Categorical Models Probabilities.

Categorical Models	1.25	1.5	1.75	2
Mkt SIZE VALUE PROF INV	15.11	15.50	15.70	15.58
Mkt SIZE VALUE PROF INV REV	13.92	12.89	11.61	10.46
Mkt UMD SIZE VALUE PROF INV	6.77	10.55	12.52	13.56
Mkt UMD SIZE VALUE PROF INV REV	5.80	8.24	8.87	8.84
Mkt VALUE PROF INV	10.28	7.74	6.88	6.53

5. Conclusions

This paper applies the method proposed by Barillas (2018) to utilize the Bayesian method to compare factor pricing models. Based on the data of American stock market from 1980 to 2019 and given a total of 13 nonmarket candidate traded factors, this paper compares 64 categorical models and 810 individual models based on their unconditional probabilities calculated with Bayesian method. Since Bayesian estimation method requires a prior specification, this paper uses Barillas' method to determine a prior specification by setting the expected maximum Sharp ratio multiples. The empirical result demonstrates that under the prior specification of 1.5 times Sharp multiple, the best performance model is {Mkt, UMD, SMB, HMLm, ROE IA}, the best categorical model is {Mkt SIZE VALUE PROF INV}, the unconditional probability rankings of the six categorical factors is: PROF, INV, VALUE, SIZE, REV, UMD, and the two individual factors with the highest predictive power are IA and ROE. The empirical result is fundamentally stable across different prior specification of maximum Sharp multiples.

References

[1] Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. The journal of Finance, 32, 663-682.

[2] Shanken, J. (1987). Multivariate proxies and asset pricing relations: Living with the Roll critique. Journal of Financial Economics, 18, 91-110.

[3] Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. Econometrica: Journal of the Econometric Society, 1121-1152.

[4] Harvey, C. R., & Zhou, G. (1990). Bayesian inference in asset pricing tests. Journal of Financial Economics, 26, 221-254.

[5] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33, 3-56.

[6] Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. The journal of finance, 49, 1541-1578.

[7] Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. Journal of Finance, 51, 55-84.

[8] Sheu, T. Y., & Rosenberg, M. (1998). Microstructure of microcapsules consisting of whey proteins and carbohydrates. Journal of Food Science, 63, 491-494.

[9] Asness, C., & Frazzini, A. (2013). The devil in HML's details. The Journal of Portfolio Management, 39, 49-68.

[10] De Moor, L., Dhaene, G., & Sercu, P. (2015). On comparing zero-alpha tests across multifactor asset pricing models. Journal of Banking & Finance, 61, S235-S240.

[11] Fama, E. F., & French, K. R. (2015). Incremental variables and the investment opportunity set. Journal of Financial Economics, 117, 470-488.

[12] Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. The Review of Financial Studies, 28, 650-705.

[13] Fama, K. R. F. E. F. (2016). Dissecting anomalies with a five-factor model. Review of Financial Studies, 29, 69-103.

[14] Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the cross-section of expected returns. The Review of Financial Studies, 29, 5-68.

[15] Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. Journal of financial Economics, 123, 441-463.

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

[17] Fama, E. F., & French, K. R. (2018). Choosing factors. Journal of Financial Economics, 128, 234-252.