THE EUROPEAN JOURNAL OF FINANCE
https://doi.org/10.1080/1351847X.2024.2312203

Model scan of factors in U.K. stock returns

Jonathan Fletcher, Andrew Marshall and Michael O'Connell

Department of Accounting and Finance, University of Strathclyde, Glasgow, Scotland

ABSTRACT

We use the Bayesian model scan approach of Chib, S., X. Zeng, and L. Zhao. 2020. 'On Comparing Asset Pricing Models.' *The Journal of Finance* 75 (1): 551–577. https://doi.org/10.1111/jofi.12854, and Chib, S., L. Zhao, and G. Zhou. 2023. 'Winners from Winners: A Tale of Risk Factors.' *Management Science* to examine which are the best performing models in a set of 12 candidate factors in U.K. stock returns. We find that a five-factor model has the highest posterior probability across the whole sample period but the posterior probability is low. The best factor model outperforms traditional factor models using a number of metrics. However the best model performs poorly in pricing a set of anomaly portfolios.

ARTICLE HISTORY
Received 20 December 2022
Accepted 22 January 2024

KEYWORDS
KEYWORDS: Bayesian analysis; model scan; factor models; U.K. stock returns

KEYWORDS
JEL CLASSIFICATION: G11; G12

1. Introduction

Ever since the development of the capital asset pricing model (CAPM) in the mid-1960s, linear factor models have played a significant role in many practical applications such as evaluating managed fund performance, estimating cost of equity capital, estimating expected returns and the covariance matrix, and factor investing. In recent years, a number of new factor models have been proposed including Fama and French (2015, 2018), Hou, Xue, and Zhang (2015), Stambaugh and Yuan (2017), and Daniel, Hirshleifer, and Sun (2020) among others. Given the increasing number of factor models, it becomes less clear which are best models to use in practical applications such as those mentioned above.

Barillas and Shanken (2017) show that when it comes to comparing linear factor models, the choice of the test assets is irrelevant for a number of different metrics. The important issue in relative model comparison tests is how well the model prices factors not included in the model. Barillas and Shanken (2018) develop a Bayesian approach for model comparison tests.[1] The Bayesian approach is useful as it can consider a large number of models, both nested and non-nested. The Bayesian approach evaluates factor models on the basis of their Marginal Likelihoods (ML), from which the posterior probabilities of the models can be calculated.

Chib, Zeng, and Zhao (2020) provide a critique of the Barillas and Shanken (2018) approach and show that their method of calculating ML is not appropriate.[2] Chib et al propose an alternative approach to calculating ML that can be used for relative model comparison tests. The attraction of their approach is that the ML can be solved analytically as relies on multivariate normality. Chib, Zhao, and Zhou (2023) use this approach to run a Bayesian model scan for comparing models among a set of factors from the models of Fama and French (2018), Hou, Xue, and Zhang (2015), Stambaugh and Yuan (2017), and Daniel, Hirshleifer, and Sun (2020) as these models have performed well in empirical studies. We will term these models as traditional factor models. Chib et al find that the best performing model is a seven-factor model, and that the four models mentioned above perform poorly in the Bayesian model scan tests.[3]

CONTACT Jonathan Fletcher j.fletcher@strath.ac.uk

Supplemental data for this article can be accessed online at https://doi.org/10.1080/1351847X.2024.2312203.

We use the Bayesian model scan approach of Chib, Zeng, and Zhao (2020) and Chib, Zhao, and Zhou (2023) to examine which are the best models that can be formed from a wide range of factors in U.K. stock returns. The factors that are included in the model scan come from the Fama and French (2018), Frazzini and Pedersen (2014), Hou, Xue, and Zhang (2015), Stambaugh and Yuan (2017), and Daniel, Hirshleifer, and Sun (2020) models. Our sample period is between July 1983 and December 2022. We use the first 60 observations of the sample as a training sample as in Chib, Zhao, and Zhou (2023) to estimate the hyperparameters for the prior distribution, and we use the remaining observations to conduct the model comparison tests.

There are three main findings in our study. First, we find that the best factor model across the whole sample period is a five-factor model, including the Market, value (HML), profitability (RMW), momentum (MOM), and financial (FIN) factors. However, the posterior probability of this model is low and is statistically indistinguishable from the next best six factor models. Second, the best model significantly outperforms the traditional factor models using a number of metrics. Third, the best model performs poorly in pricing a set of anomaly portfolios. Our study suggests that it is likely to be more useful in combining factor models in practical applications, e.g. through Bayesian model averaging, rather than using a single factor model.

There are two main contributions of our study. First, we complement the Bayesian model scan studies of Barillas and Shanken (2018), Chib, Zhao, and Zhou (2021, 2023) in US stock returns by conducting a Bayesian model scan of factors in a different market. Studies by Harvey (2017, 2020), and Hou, Xue, and Zhang (2020) highlight the importance of replication studies in Finance. Second, we also extend the prior evidence of asset pricing studies in U.K. stock returns including Fletcher (1994, 2001, 2019), Clare, Smith, and Thomas (1997), Antoniou, Garrett, and Priestley (1998), Gregory, Tharyan, and Christidis (2013), Michou and Zhou (2016), and Foye (2018). We extend this literature by running the Bayesian model scan among a large number of factors, comparing how well the best model performs relative to the traditional factor models, and conducting pricing tests of a set of anomaly portfolios.

The paper is organized as follows. Section 2 presents the research method. Section 3 describes the data used in my study. Section 4 reports the empirical results. The last section concludes.

2. Research method

Ross (1978), Harrison and Kreps (1979), and Hansen and Richard (1987) show that if the Law of One Price (LOP) holds in financial markets, then a stochastic discount factor[4] (m_{t+1}) exists such that:

$$E(m_{t+1}X_{it+1}|Z_t) = p_{it} \quad \text{for } i = 1, \dots . N \tag{1}$$

where X_{it+1} is the payoff of asset i at time $t+1$, p_{it} is the cost of asset i at time t, Z_t is the information set used by investors at time t, and N is the number of primitive assets. If $m_{t+1} > 0$, then financial markets also satisfy the No Arbitrage (NA) condition in financial markets (Cochrane 2005).[5] If the asset payoffs are excess returns, and there is no conditioning information, then equation (1) implies that:

$$E(m_{t+1}r_{it+1}) = 0 \quad \text{for } i = 1, \dots .., N \tag{2}$$

where r_{it+1} is the excess return of asset i at time $t+1$.

Most asset pricing models specify a candidate model for the stochastic discount factor (y_{t+1}). The most popular models are linear factor models, where the candidate stochastic discount factor is given by:

$$y_{t+1} = a + \Sigma_{k=1}{}^{K} b_k f_{kt+1} \tag{3}$$

where a and b_k are the constant and slope coefficients in the stochastic discount factor, f_{kt+1} are the values of the factors at time $t+1$, and K is the number of factors in the model. The slope coefficient (b_k) tells us whether factor k is important in pricing the primitive assets given the other factors in the model (Cochrane 2005). In this study we only consider models with traded factors.

When we evaluate linear factor models in terms of equation (2), we are unable to identify all the stochastic discount factor coefficients in equation (3), and so must choose some normalization for the stochastic discount factor. We choose the normalization followed by Chib and Zeng (2020) so that the expected value of the stochastic discount factor is set equal to 1, and equation (3)[6] becomes:

$$y_{t+1} = 1 - \Sigma_{k=1}{}^{K} b_k f_{dkt+1} \tag{4}$$

where f_{dkt+1} is the demeaned value of factor k at time $t + 1$.

Dybvig and Ingersoll (1982) and Ferson and Jagannathan (1996) show that linear factor models imply an equivalent expected return and beta relation. Cochrane (2005) and Ferson (2019) show that stochastic discount factors, expected return/betas, and mean-variance frontiers are equivalent frameworks. Cochrane (2005) shows that if the linear factor model in equation (4) satisfies the pricing restrictions in equation (2), then:

$$E\ (r_{it+1}) = \Sigma_{k=1}{}^{K} \beta_{ik} \lambda_k \tag{5}$$

where β_{ik} is the factor beta of asset i relative to factor k, and λ_k is the factor risk premium of factor k. Define b is a (K,1) vector of stochastic discount factor coefficients (b_k), λ is a (K,1) vector of factor premiums (λ_k), and V_f is the (K,K) covariance matrix of the factors, Cochrane shows that:

$$b = V_f{}^{-1} \lambda \tag{6}$$

Barillas and Shanken (2017) show that for relative model comparison tests for traded linear factor models, the choice of test assets is irrelevant for a number of metrics. Any linear factor model should be able to correctly price the test assets, and any excluded factors from the model. When the union of all the factors in each model is included in the investment universe, then the role of test assets drops out. To illustrate, consider two models A and B with factors f_A, and f_B, and a set of test asset excess returns r. Using the maximum squared Sharpe (1966) (Sh^2) ratio to compare models, then for model A the metric is $Sh^2(f_A,f_B,r)$-$Sh^2(f_A)$, and for model B $Sh^2(f_A,f_B,r)$-$Sh^2(f_B)$. Given the $Sh^2(f_A,f_B,r)$ is fixed across models, then we can compare the relative performance of the models using the $Sh^2(f_A)$ and $Sh^2(f_B)$ measures alone. The better models are the ones with the highest squared Sharpe performance.[7]

Chib and Zeng (2020) extend this test asset irrelevance arguments to the stochastic discount factor approach in relative model comparison tests. Since equations (2) and (5) hold for linear factor models in the form of equation (4), this implies that:

$$\begin{aligned} &E(F_{t+1}) = \lambda \text{ for factors included in the model,} \\ &\text{and } E(F^*{}_{t+1}) = \beta\lambda \text{ for factors excluded from the model} \end{aligned} \tag{7}$$

where β is a (L,K) matrix of factor betas, F_{t+1} is the K factors included in the model, and $F*_{t+1}$, represents the L factors excluded from the model. The restrictions of equation (7) can be cast into a regression framework, where the excess returns of the factors included in the model are regressed on a constant, and the excess returns of the excluded factors are regressed on the excess returns of the included factors in the model with the intercept suppressed. The exclusion of the intercept[8] in the second regression imposes the zero pricing error restriction on the excluded factors.

Barillas and Shanken (2018) derive a Bayesian model comparison test on the basis of Marginal Likelihoods (ML) that can be used to compare the performance of a large number of traded factor models simultaneously.[9] Chib, Zeng, and Zhao (2020) propose a framework that can be applied within the context of the regression equations in (7).[10] Chib and Zeng (2020), and Chib, Zeng, and Zhao (2020) start with a model that includes all $K + L$ factors among the set of factor models being compared. They propose a prior distribution for the model with all factors, and then the prior distribution can be derived for any model with a subset of the factors. The log ML of a candidate model (m_j) is given by:

$$\text{Log ML}(m_j) = \log \text{ML}(F) + \log \text{ML}(F^*) \tag{8}$$

Under multivariate normality, Chib et al. show that the log ML of each model can be solved analytically in Proposition 5 of their paper.[11]

Chib, Zhao, and Zhou (2023) apply the results of Chib, Zeng, and Zhao (2020) to evaluate model comparison tests of all models formed from factors included in Fama and French (2018), Hou, Xue, and Zhang (2015), Stambaugh and Yuan (2017), and Daniel, Hirshleifer, and Sun (2020).[12] There are $J = 2^{K+L} - 1$ potential factor models that can be constructed by the model scan. The models can be compared using their posterior probabilities, assuming that each factor model has an equal prior probability of 1/J as:

$$\text{Posterior Probability}_j = ML_j / \Sigma_{j=1}{}^{J} ML_j \tag{9}$$

where ML_j is the marginal likelihood of model j. The approach of Chib, Zhao, and Zhou (2023) uses a training sample to estimate the hyperparameters for the prior distributions. Follow this we use the first 60 observations of our sample period as the training sample. They point out that we can test whether models are statistically indistinguishable from one another by using the difference in the log ML of two models. For example, if the model with the highest posterior probability has a difference with the log ML of another model higher than 1.15, then according to the Jeffreys's rule (Jeffreys 1961) the two models are significantly different.[13]

Chib, Zhao, and Zhou (2023) also derive the posterior distribution of the various parameters in the regression framework, from which we can then derive the posterior distribution of the stochastic discount factor coefficients in equation (6). Chib, Zhao, and Zhou (2023) also propose an approach to examine whether a given factor model is able to correctly price excluded factors from the model.[14] Define r_{it+1} as the excess return of the excluded factor i at time t + 1. We run two regressions as follows:

$$r_{it+1} = \sum_{k=1}^{K} \beta_{ik} f_{kt+1} + e_{it+1} \tag{10}$$

$$r_{it+1} = \alpha_i + \sum_{k=1}^{K} \beta_{ik} f_{kt+1} + e_{it+1} \tag{11}$$

where e_{it+1} are random error terms for asset i at time t + 1, β_{ik} are the factor beta of asset i relative to factor k, and α_i is the pricing error. When the factor model is able to price the excluded factor, then α_i will equal zero.

To test the hypothesis that the factor model prices the excluded factor, Chib, Zhao, and Zhou (2023) use the difference in the log ML of the unrestricted regression (equation (11)), and the restricted regression (equation (10)). If the difference in the log ML exceeds 1.15, then the factor model is not able to correctly price the excluded factor. Chib, Zhao, and Zhou (2023) use a Bayesian Markov Chain Monte Carlo (MCMC) to estimate the two regressions, and compute the log ML following Chib (1995). A normal-inverse gamma prior is assumed for the regression coefficients and the residual variance, which is estimated from the training sample (see Greenberg 2013). A Gibbs sampler is used with a burn-in sample of 1000 trials and 10,000 usable trials. Chib et al also propose an adjusted alpha where the posterior mean alpha is multiplied by the posterior probability of the alternative hypothesis. We use the regressions in (10) and (11) to examine the pricing of excluded factors and anomaly portfolios.

We also consider two out-of-sample performance tests following the approach of Chib, Zhao, and Zhou (2023). The first test is based on the Bayesian predictive likelihood. We use 120 months as the out-of-sample period. The training sample remains the same and we use the rest of the sample period as the in-sample period. We use 10 different contiguous divisions of the sample data to estimate the predictive likelihood as the difference in the log ML for the whole sample period minus the log ML using the in-sample period. We then calculate the predictive likelihood for the factor models as the log average of the 10 predictive likelihoods.

The second test is based on the out-of-sample Sharpe (1966) performance of the tangency portfolio of each factor model similar to Chib, Zhao, and Zhou (2023). The training sample is set to 60 months, and the initial estimation window is set to 240 months. We use the following steps. First, the posterior distribution of the parameters are estimated using the training sample and estimation window using 10,000 simulation draws. Second, we then calculate the predictive expected excess returns and covariance matrix of the factors in the model. Third, given the predictive excess returns and covariance matrix, the optimal tangency portfolio of the factors in the model can be calculated.[15] Fourth, given the optimal weights the portfolio excess return of the tangency portfolio is calculated in the next month. Steps 1–4 are repeated each month until the end of the sample period using a recursive estimation window. Given the time-series of the out-of-sample portfolio excess returns, we then calculate the Sharpe performance as the average excess return divided by the standard deviation.

3. Data

Our focus in this study is on U.K. factors. All of the data is collected from the London Share Price Database (LSPD) run by London Business School, and Refinitiv Worldscope. The sample period is between July 1983 and December 2022. Full details on the formation of the factors is included in an online Appendix. We use the monthly return on the one-month U.K. Treasury Bill as the risk-free asset, which we collect from LSPD and Refinitiv Datastream. When the factors are formed once each year, the annual revision takes place at the start of July each year as is common in the literature.[16]

The following factors are used;

1. Fama and French (1993, 2015, 2018) (FF3, FF5, FF6)

The factors are the excess returns on the market index, and zero-cost portfolios of the size (SMB), value (HML), profitability (RMW), investment (CMA), and momentum (MOM) effects in stock returns. From these factors, we get the single factor CAPM (Market), FF3 (Market, SMB, HML), FF5[17] (Market, SMB, HML, RMW, CMA), and FF6 (Market, SMB, HML, RMW, CMA, MOM).

2. Hou, Xue, and Zhang (2015) (HXZ)

The factors are the zero-cost portfolios of the profitability (ROE), and investment (IA) effects in stock returns. The HXZ model includes the Market, SMB[18], ROE, and IA factors.

3. Frazzini and Pedersen (2014)

Frazzini and Pedersen (2014) propose the Betting against Beta (BAB) factor.

4. Stambaugh and Yuan (2017) (SY)

The factors are the two mispricing factors termed Management (MGMT), and Performance (PERF). The SY model includes the Market, SMB, MGMT, and PERF factors.

5. Daniel, Hirshleifer, and Sun (2020) (DHS)

Daniel, Hirshleifer, and Sun (2020) propose the behavioral financial (FIN) factor.

Table 1 reports summary statistics of the excess factor returns between July 1983 and December 2022. The summary statistics include the average excess return (%), standard deviation (Std Dev), and the *t*-statistic of the null hypothesis that the average excess factor returns are equal to zero.

Table 1 shows that all of the factors have significant positive average excess returns, except for the SMB and ROE factors. The MOM, and BAB factors have the largest average excess returns at 0.760%, and 0.591% respectively. The profitability factors (RMW, ROE) have different average excess returns, with only the RMW factor having a significant positive mean excess return. The mispricing factors in the Daniel, Hirshleifer, and Sun (2020), and Stambaugh and Yuan (2017) models have significant positive average excess returns. The CMA, MOM, and FIN factors are the only factors with a *t*-statistic higher than 3, which is the cut-off *t*-statistic recommended by Harvey, Liu, and Zhu (2016) to control for multiple testing.

4. Empirical results

We begin our empirical analysis by running the model scan using all 12 factors. There are 4095 possible models, and we assign an equal prior probability to each model. Table 2 reports the empirical results. Panel A of the Table reports the results for the top 7 models in terms of the highest posterior probability. The panel includes the log ML, posterior probability (Posterior prob) of each model, and the ratio of the posterior probability to the prior

Table 1. Summary statistics of factors.

	Mean	Std Dev	*t*-statistic
Market	0.438	4.25	2.243[a]
SMB	0.054	2.974	0.394
HML	0.262	2.816	2.022[a]
RMW	0.212	2.134	2.166[a]
CMA	0.286	1.948	3.193[a]
MOM	0.76	3.235	5.116[a]
ROE	−0.032	2.599	−0.264
IA	0.401	3.641	2.396[a]
BAB	0.591	5.705	2.254[a]
FIN	0.452	1.755	5.607[a]
MGMT	0.435	3.456	2.739[a]
PERF	0.432	4.416	2.129[a]

Notes: The table reports summary statistics of factors between July 1983 and December 2022. The summary statistics include the average excess returns (%) and standard deviation (Std Dev) of the factors. The *t*-statistic column is the *t*-statistic of the null hypothesis that the average excess factor returns are equal to zero.
[a]Significant at 5%.

Table 2. Model scan of 12 factors.

Panel A:			
Best Models	Log ML	Posterior prob	Postpriorratio
1	11069.14	0.073	300.917
2	11068.52	0.04	162.485
3	11068.5	0.039	159.682
4	11068.5	0.039	159.188
5	11068.31	0.032	131.492
6	11068.17	0.028	114.456
7	11068.02	0.024	98.952
Panel B:	Log ML	Posterior prob	Postpriorratio
CAPM	11042.66	0	0
FF3	11038.42	0	0
FF5	11042.58	0	0
FF6	11058.39	0	0.006
HXZ	11039.6	0	0
SY	11048.42	0	0
Panel C:			
Best Models	Mkt,HML,RMW,MOM,FIN		
1	Mkt,RMW,MOM,FIN,MGMT		
2	Mkt,RMW,MOM,FIN		
3	Mkt,SMB,HML,RMW,MOM,FIN		
4	Mkt,MOM,FIN,MGMT		
5	Mkt,HML,RMW,MOM,FIN,MGMT		
6	Mkt,SMB,RMW,MOM,FIN		
7	Mkt,HML,RMW,MOM,FIN		

Notes: The table reports the results of the Bayesian model scan of 12 factors in U.K. stock returns. The sample period is July 1983 and December 2022. The first 60 observations of the sample period is used for the training sample, and the model scan is then conducted on the remaining observations of the sample period. Panel A reports the log marginal likelihood (Log ML), the posterior probability (Posterior prob), and the ratio of posterior probability to prior probability (postpriorratio) for the top 7 models. Panel B includes the log marginal likelihood, posterior probability, and the ratio of posterior probability to prior probability for the traditional factor models. Panel C presents the identity of the factors in the top 7 models from the Bayesian model scan.

probability (postpriorratio). Panel B of Table 2 reports the same information for the traditional factor models as CAPM, FF3, FF5, FF6, HXZ, and SY. Panel C reports the identity of the factors in the top 7 models from the model scan.

Table 2 shows that the best factor model in the model scan is a five-factor model with a posterior probability of 0.073. The next six best models have a posterior probability that ranges between 0.024 and 0.040. The difference in the log ML of the best model and the next six best models are all below 1.15 and so the models are statistically indistinguishable from one another (Chib, Zhao, and Zhou 2023). The ratio of the posterior probability to prior probability shows a substantial increase for the seven best models. The results in panel A of Table 2 differ from Chib et al. They find that the top 3 models in US stock returns have a combined posterior probability of 0.3201. After the top 3 models, there is a sharp drop in the posterior probabilities, and there are significant differences between the top 3 models, and the remaining models. The difference in results might stem from the longer sample period that Chib et al use or the use of a different market. A recent study by Pukthuanthong, Qiao, and Wang (2023) find that the best factor model in international markets from the Bayesian model scan is country specific.

Panel B of Table 2 shows that the traditional factor models perform poorly in the model scan. The posterior probabilities of each model are essentially zero, and all of the models have a much lower posterior probability than the prior probability assigned to each model. Likewise the best models significantly outperform the traditional factor models as the difference in the log ML between the models all exceed 1.15. The inferior performance of the traditional factor models is consistent with Chib, Zhao, and Zhou (2023).

Panel C of Table 2 shows that the best factor model includes the Market, HML, RMW, MOM, and FIN factors. The Market, MOM, and FIN factors are present in all of the top models. The key role of the Market factor in factor models is consistent with Harvey and Liu (2021), Fletcher (2019), and Bryzgalova, Huang, and Julliard (2023). The RMW factor is included in six of the top models. The HML, SMB, and MGMT factors are all included in at least one of the top seven models.

Table 2 suggests that the best model in terms of posterior probability is a five-factor model. Chib, Zhao, and Zhou (2023) also derive the posterior distribution of the factor premiums for a given factor model in equations (17) to (20) of their paper. We use 10,000 simulation draws for generating the posterior distribution of the factor premiums, and the corresponding stochastic discount factor coefficients. Table 3 reports the summary statistics of the posterior distribution of the factor premiums (panel A), and stochastic discount factor coefficients (panel B) of the best model. The summary statistics include the mean, standard deviation (Std Dev), median, and 2.5% and 97.5% percentiles of the posterior distribution.

Panel A of Table 3 shows that the MOM factor has the largest posterior mean factor premium at 0.770%, followed by the FIN factor at 0.437%. It is only the RMW, MOM, and FIN factors which have significant positive

Table 3. Summary statistics of the posterior distribution of factor premiums and stochastic discount factor coefficients of the best model risk factors.

Panel A:					
Premiums	Mean	Std Dev	Median	2.5%	97.5%
Market	0.382	0.196	0.381	−0.005	0.769
HML	0.162	0.144	0.161	−0.124	0.442
RMW	0.225	0.107	0.224	0.017	0.435
MOM	0.770	0.169	0.771	0.441	1.101
FIN	0.437	0.088	0.438	0.267	0.612
Panel B:					
SDF Coefficients	Mean	Std Dev	Median	2.5%	97.5%
Market	6.932	1.494	6.921	3.029	10.871
HML	6.564	2.604	6.516	−0.003	13.370
RMW	10.530	3.193	10.446	2.633	19.055
MOM	9.608	1.814	9.616	5.045	14.345
FIN	14.282	3.423	14.222	5.538	23.288

Notes: The table reports the summary statistics of the posterior distribution of the factors in the best model from the Bayesian model scan of 12 factors in U.K. stock returns. The sample period is July 1983 and December 2022. The first 60 observations of the sample period is used for the training sample, and the model scan is then conducted on the remaining observations of the sample period. Panel A reports the summary statistics of the posterior distribution of the factor premiums (%), and panel B reports the summary statistics of the posterior distribution of the stochastic discount factor coefficients. The summary statistics include the mean, standard deviation (Std Dev), median, 2.5% and 97.5% percentiles using 10,000 simulation draws.

factor premiums where the Bayesian credibility intervals do not straddle zero. In panel B of Table 3 all of the stochastic discount factor coefficients are significantly positive for each factor, except for the HML factor, using the 95% Bayesian credibility intervals. This finding suggests that four of the five factors play an important role in the stochastic discount factor in pricing assets given the other factors in the model (Cochrane 2005). The Market factor plays an important role even where the corresponding factor premium is not significantly positive. Cochrane points out that when the factors are correlated with each other, a significant positive factor premium does not tell us whether the factor is useful in pricing other assets given the other factors in the model but looking at the stochastic discount factor coefficients does.[19]

To compare the best factor model in Table 2 to the traditional factor models in more detail, we estimate the fitted stochastic discount factor values for each factor model using the posterior mean values of the corresponding stochastic discount factor coefficients. Table 4 reports the summary statistics of the fitted stochastic discount factor values for each model. The summary statistics include the standard deviation (Std Dev), minimum, maximum, and the proportion of fitted stochastic discount factor values below zero (Prop < 0). We do not report the mean values as by construction they equal one. The final column reports the correlations between the traditional factor models, and the best factor model.

Table 4 shows that the best factor model has the largest stochastic discount factor volatility across the factor models, closely followed by the FF6 model. This finding suggests that the best factor model is more likely to pass the Hansen and Jagannathan (1991) volatility bounds. The CAPM and FF3 stochastic discount factors have the lowest volatility reflecting the inferior performance of these models. The best model, HXZ and FF6 models have a tiny proportion of negative fitted stochastic discount factor values. This result suggests that the better performance of the best factor model does not come at the expense of violating the NA restriction. This result differs from Li, Xu, and Zhang (2010), and Wang and Zhang (2012) who find that multifactor models have better pricing performance at the expense of a greater proportion of negative fitted stochastic discount factor values.[20] The best factor model has the highest correlation with the FF6 stochastic discount factor due to the inclusion of the HML, RMW, and MOM factors in the best model.

We next examine how well the best factor model, and the traditional factor models are able to correctly price the excluded factors from the model. Chib, Zhao, and Zhou (2023) find that the traditional factor models do a poor job in pricing the excluded factors in US stock returns, but the best factor model is able to price the excluded factors. Table 5 reports the adjusted alphas (%) for the excluded factors relative to each model. Highlighted in bold is where the difference in the log ML for the unrestricted regression (equation (11)), and restricted regression (equation (10)) is greater than 1.15.

Table 5 shows that the best factor model does a good job pricing the excluded factors. None of the excluded factors have significant alphas relative to the best model. All of the factors, except SMB, have tiny adjusted alphas. In contrast, the traditional factor models have large significant mispricing in at least one of the excluded factors. The CAPM and FF3 models have large significant positive adjusted alphas for most of the excluded factors. The

Table 4. Summary statistics of candidate stochastic discount factor models.

	Std Dev	Minimum	Maximum	Prop ($y < 0$)	Correlation
Best	0.464	−0.646	3.058	0.017	
CAPM	0.098	0.734	1.623	0	0.308
FF3	0.104	0.604	1.557	0	0.331
FF5	0.243	0.057	1.923	0	0.576
FF6	0.412	−0.578	2.978	0.008	0.873
HXZ	0.231	−0.536	2.111	0.006	0.202
SY	0.293	0.187	2.259	0	0.38

Notes: The table reports the summary statistics of the fitted stochastic discount factor values of the best factor model from the Bayesian model scan, and for the traditional factor models. The fitted values are estimated using the posterior mean of the stochastic discount factor coefficients. The sample period is July 1983 and December 2022. The first 60 observations of the sample period is used for the training sample, and the model scan is then conducted on the remaining observations of the sample period. The summary statistics include the standard deviation (Std Dev), the minimum, maximum, and the proportion (Prop $y < 0$) of fitted values that are below zero. The final column includes the correlations between fitted values of the traditional factor models, and the best model. The posterior distribution of the stochastic discount factor coefficients is estimated from 10,000 draws.

Table 5. Adjusted alphas of excluded factors.

	Best	CAPM	FF3	FF5	FF6	HXZ	SY
SMB	0.183	0.002					
HML		0.006				0.003	0.027
RMW		**0.242**	**0.327**			**0.22**	0.027
CMA	0.012	**0.241**	**0.248**			0.01	0.023
MOM		**0.826**	**0.896**	**0.911**		**0.842**	**0.743**
ROE	−0.002			−0.003	−0.002		−0.01
IA	0.01	0.246	0.193	0.007	0.016		0.029
BAB	0.013	0.017	0.048	0.067	−0.005	0.042	0.005
FIN		**0.49**	**0.46**	**0.385**	**0.352**	**0.459**	**0.433**
MGMT	0.046	**0.493**	**0.448**	0.079	0.12	0.282	
PERF	0.031	**0.637**	**0.717**	**0.479**	0.03	**0.608**	

Notes: The table reports the adjusted alphas (%) for the excluded factors relative to the best factor model in the model scan, and relative to the traditional factor models. The sample period is July 1983 and December 2022. The first 60 observations of the sample period is used for the training sample and the remaining observations of the sample is used to estimate the posterior distribution of the alphas. A Gibbs sampler is used with a burn-in sample of 1000 trials and 10,000 simulation trials are used to estimate the posterior distribution of the regression coefficients. Highlighted in bold are where the difference in log marginal likelihoods of the unrestricted and restricted regressions is larger than 1.15.

FF5, HXZ, and SY models have massive adjusted alphas on the MOM factor. All three models are also unable to price the FIN factor. The FIN factor is the only factor that the FF6 model is not able to correctly price. The results in Table 5 are consistent with Chib, Zhao, and Zhou (2023).

We next examine how well the best factor model and the traditional factor models are able to price a set of test assets. We use 34 anomaly portfolios as our set of test assets[21] using a subset of the characteristics used by Chib et al. (2023)[22] (see also Kozak, Nagel, and Santosh 2020).[23] Table 6 reports the empirical results. Panel A reports the mean absolute adjusted alpha (%), and the number of significant alphas for each factor model. Panel B reports the adjusted alphas (%) of the anomaly portfolios relative to the best model, and FF6 and SY models. Highlighted in bold are where the difference in log ML of the unrestricted and restricted regressions exceeds 1.15.

Panel A of Table 6 shows that the best factor model and the traditional factor models are unable to correctly price a number of anomaly portfolios. The number of significant alphas range between 12 (SY), and 16 (Best). The mean absolute adjusted alphas are similar across models ranging between 0.217% (FF5) and 0.256% (FF3). In this set of test assets, the best factor model is not able to outperform the traditional factor models.

Panel B of Table 6 shows that the majority of the significant alphas on the anomaly portfolios are negative. All three models are not able to price Asset Growth, Cash Productivity, Cash/Price, Growth in CAPEX, CAPEX and Inventory, 1-month Reversal, Size, Sales Growth, and Sales/Market Value. The 1-month Reversal has the largest mispricing with adjusted alphas of −1.653% (Best), −1.586% (FF6), and −1.453% (SY). The SY model is able to correctly price the Beta, BM, Cash/Total Assets, and Net Operating Assets anomaly portfolios that the Best and FF6 models cannot. The best and FF6 models can price the Momentum and Long-Term Reversals anomaly portfolio that the Best and SY models cannot.

Table 6 shows the best factor model does not outperform the traditional factor models in pricing the anomaly portfolios. A recent study by Chib et al. (2023) find that their best model from a model scan in a much larger set of factors has much better pricing performance than the FF6 model in characteristic portfolios of stocks, exchange-traded funds, and individual US stocks. They find that their best model is unable to correctly price 24 anomaly portfolios out of 125 but does not examine the pricing performance of the traditional models. The better performance of the best models in these studies might stem from the higher posterior probabilities that these models have in US stock returns. Our final set of tests examines the out-of-sample performance of the best factor model, and traditional factor models using the predictive likelihood and Sharpe performance. Table 7 reports the log mean predictive likelihood and annualized Sharpe performance for the best model and the traditional factor models.

Table 7 shows that the best factor model outperforms the traditional factor models in both out-of-sample tests. The best factor model has the highest predictive likelihood across models. The differences in the log predictive likelihoods of the best model to the other models all exceed 1.15. Among the traditional factor models,

Table 6. Adjusted alphas of anomaly portfolios.

Panel A	Best	CAPM	FF3	FF5	FF6	HXZ	SY
|Alpha|	0.245	0.253	0.256	0.217	0.228	0.229	0.228
N Sig	16	15	15	14	15	14	12

Panel B	Best	FF6	SY
Accruals	−0.066	−0.018	−0.045
Asset growth	**−0.455**	**−0.447**	**−0.549**
Beta	**0.343**	**0.162**	0.006
BM	**0.4**	**0.471**	0.314
Cash/Total assets	**0.352**	**0.268**	0.064
Cash/Debt	0.002	−0.004	−0.001
Cash productivity	**−0.514**	**−0.558**	**−0.484**
Cash/Price	**0.644**	**0.643**	**0.541**
Changes in outstanding shares	0.016	−0.005	−0.019
Dividend yield	0.025	0.034	0
Earnings/Price	0.004	0.001	0.006
Gross profitability	0.054	0.043	0.075
Growth in CAPEX	**−0.225**	**−0.25**	**−0.268**
Idiosyncratic volatility	**0.443**	0.165	0.043
CAPEX and inventory	**−0.532**	**−0.557**	**−0.595**
Leverage	0.197	**0.265**	0.167
% Change in total liabilities	**−0.192**	**−0.269**	**−0.272**
Momentum	0.005	0.023	**0.66**
1-month reversal	**−1.653**	**−1.586**	**−1.453**
Size	**−0.61**	**−0.544**	**−0.527**
Operating profitability	0.001	0.017	0.015
% change in gross margin-% change in sales	−0.007	−0.001	−0.004
ROE	−0.002	−0.002	−0.01
Sales/Cash	**−0.219**	−0.091	−0.01
Sales/Inventory	0.039	0.08	0.06
Sales growth	**−0.401**	**−0.448**	**−0.519**
Sales/market value	**0.272**	**0.331**	**0.519**
Asset turnover	0.001	0.004	0.02
Gross margin	0.1	0.057	0.029
Net operating assets	**−0.387**	**−0.322**	−0.056
Investment/Capital	0.008	−0.006	−0.049
Momentum-reversal	−0.03	−0.005	−0.063
Long-term reversals	−0.123	−0.071	**−0.326**
Seasonality	−0.003	−0.002	0.001

Notes: The table reports the pricing tests using 34 anomaly portfolios. The sample period is July 1983 and December 2022. The first 60 observations of the sample period is used for the training sample and the remaining observations of the sample is used to estimate the posterior distribution of the alphas. A Gibbs sampler is used with a burn-in sample of 1000 trials and 10,000 simulation trials are used to estimate the posterior distribution of the regression coefficients. Panel A reports the average absolute adjusted (%) (|alpha|), and the number of significant alphas (N sig) for each model. Panel B reports the adjusted alphas (%) of 34 anomaly portfolios relative to the best factor model in the model scan, and relative to the FF6 and SY models. Highlighted in bold are where the difference in log marginal likelihoods of the unrestricted and restricted regressions is larger than 1.15.

the FF6 model has the highest predictive likelihood. The best model also has the largest annualized Sharpe performance at 1.355, which is considerably higher than the traditional models. The SY and FF6 have the next highest annualized Sharpe performance at 1.078 and 0.922, respectively.

The results in Tables 2–7 show that although the best factor model outperforms the traditional factor models in a number of tests, it does perform poorly in pricing the anomaly portfolios. Likewise, the top seven factor models in the model scan are statistically indistinguishable from one another. Related empirical evidence in Bayesian model scan studies such as Barillas and Shanken (2018, 2022), Chib and Zeng (2020), Chib, Zhao, and Zhou (2023), and Chib et al. (2023) highlight that the best factor model varies between the different studies. Pukthuanthong, Qiao, and Wang (2023) highlight the best factor models across international markets varies considerably. These studies suggest that the Bayesian model scan studies are sensitive to time period chosen, factors used, and financial market studied. Given this, it suggests the importance of combining factor models in practical applications along the lines of Bayesian model averaging, of which examples are the recent studies by Avramov et al. (2023) and Bryzgalova, Huang, and Julliard (2023).

Table 7. Out-of-sample tests.

	Predictive Likelihood	Sharpe
Best	3406.89	1.355
CAPM	3391.19	0.431
FF3	3390.55	−0.051
FF5	3391.76	0.777
FF6	3402.29	0.922
HXZ	3389.57	0.636
SY	3392.67	1.078

Notes: The table reports out-of-sample tests of the best factor model from the Bayesian model scan and traditional factor models. The sample period is July 1983 and December 2022. The first 60 observations of the sample period is used for the training sample. The first test is the predictive likelihood. The out-of-sample period is set to 120 months. Using 10 contiguous divisions of the sample period, the predictive likelihood is calculated and the log of the average predictive likelihood is reported. The second test is the out-of-sample Sharpe performance of the optimal tangency portfolio of each model. The optimal weights are estimated each month using a recursive estimation window, where the initial window is set to 240 months. The predictive expected excess returns and covariance matrix are estimated using 10,000 simulation draws from the posterior distribution of the factor premiums and factor covariance matrix.

5. Conclusions

We use the Bayesian model scan approach of Chib, Zeng, and Zhao (2020) and Chib, Zhao, and Zhou (2023) to examine model comparison tests among a set of U.K. factors. We examine how well the best model from the model scan performs relative to traditional factor models based on the CAPM, Fama and French (1993, 2015, 2018), Hou, Xue, and Zhang (2015), and Stambaugh and Yuan (2017). There are three main findings in our study.

First, we find that the best factor model during the whole sample period is a five-factor model, which includes the Market, HML, RMW, MOM, and FIN factors. However, the posterior probability is small at 0.073, and the performance of the best model is statistically indistinguishable from the next best six factor models in terms of posterior probability. All of the factors in the best model, except the HML factor, have significant positive stochastic discount factor coefficients and so play an important role in the stochastic discount factor given the other factors in the model. The finding that the top models are statistically indistinguishable and have relative low probabilities suggesting that it might be better to combine models in practical applications rather than simply using a single model.

Second, the best factor model significantly outperforms the traditional factor models along a number of dimensions. The posterior probabilities of the traditional factor models are essentially zero, and the best model has a much higher significant log ML than the traditional models. The best model does a better job pricing the excluded factors. The best model also outperforms in out-of-sample tests providing both higher predictive likelihood and higher Sharpe performance. This finding is consistent with Chib, Zhao, and Zhou (2023).

Third, the best model struggles to correctly price a number of anomaly portfolios. Out of 34 anomaly portfolios, there are 16 significant alphas for the best model, and the adjusted posterior mean alphas are substantial. The best model also does not outperform the traditional factor models in pricing the anomaly portfolios. This finding suggests that even using the best factor model from a model scan can still perform poorly in pricing test assets.

Our empirical results suggest in terms of practical applications that it might be better to combine factor models rather than use a single model. One application could be in the area of optimal portfolio choice. Our analysis has used the Bayesian model scan under the multivariate normal distribution of Chib, Zeng, and Zhao

(2020), and Chib, Zhao, and Zhou (2023). The analysis could be extended to use the multivariate t distribution along the lines of Chib and Zeng (2020). Another interesting extension to our study would be to look at impact how the factors are formed such as in Ehansi and Linnainmaa (2021), or a wider range of factors as in Chib et al. (2023). A final topic is looking at whether the best factor model changes with using more recent data following Chib, Zhao, and Zhou (2021). We leave these issues to future research.

Notes

1. See Fama and French (2018), and Barillas et al. (2020) for alternative approaches for relative model comparison tests.
2. See the response by Barillas and Shanken (2022).
3. Chib, Zhao, and Zhou (2023) also consider the benefits of adding principal components (formed from anomalies that the best model cannot price), and find that three principal components are included in the best model.
4. See Cochrane (2005) and Ferson (2019) for reviews of the stochastic discount factor approach to asset pricing.
5. The stochastic discount factor will only be unique if markets are complete.
6. See Cochrane (2005) for more details. Burnside (2016) examines the impact of different normalization methods in linear factor models in this situation.
7. Ferson, Siegel, and Wang (2022) extend these arguments in the presence of conditioning information.
8. If the intercept was included it would capture the Jensen (1968) alpha of the excluded factors. The Jensen alpha is identical to the stochastic discount factor alpha in this setup as $E(y_{t+1}) = 1$. See Ferson (2019).
9. Alternative Bayesian model comparison tests have been developed using the expected return/beta framework of Avramov et al. (2023).
10. See Bryzgalova, Huang, and Julliard (2023) for an alternative approach.
11. Chib and Zeng (2020) develop model comparison tests under the multivariate t distribution, which requires the use of Bayesian Markov Chain Monte Carlo simulation methods.
12. Chib, Zhao, and Zhou (2023) use the same size factor across all models.
13. Chib, Zhao, and Zhou (2023) point out that using 1.15 is equivalent to an odds ratio of 3.15 to 1. Alternative values of the difference between the log ML of two models can be set by selecting different odds ratio.
14. The same approach can be used to examine whether a given factor model correctly prices a set of test assets, which we use later in the paper.
15. We rescale the weights to sum to 1 as in Chib, Zhao, and Zhou (2023).
16. An alternative approach is used by Gregory, Tharyan, and Christidis (2013) where the factors are formed at the beginning of October each year. Given that we use accounting information from the prior calendar year, we stick with the July revision date.
17. Foye (2018) finds that both the Fama and French (1993; 2015) models perform poorly in U.K. stock returns. Foye also finds that using gross profitability to form the RMW factor performs better than alternative measures of profitability such as operating profitability used by Fama and French (2015).
18. Chib, Zhao, and Zhou (2023) use the same SMB factor across models in their study.
19. Kan, Robotti, and Shanken (2013) make a similar point in cross-sectional regression asset pricing tests.
20. See also Gospodinov, Kan, and Robotti (2016).
21. Details on how the anomaly portfolios are formed are detailed in the online Appendix.
22. The characteristics selected by Chib, Zhao, and Zhou (2023) are a subset of the characteristics used in Green, Hand, and Zhang (2017).
23. We thank one of the reviewers for suggesting an examination of this issue.

Acknowledgements

We are grateful for the comments of two anonymous reviewers, and the use of R code from S. Chib.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Jonathan Fletcher, Professors of Finance at the University of Strathclyde.

Andrew Marshall, Professors of Finance at the University of Strathclyde.

Michael O'Connell is finishing his PhD at the University of Strathclyde.

ORCID

Andrew Marshall http://orcid.org/0000-0001-7081-1296

References

Antoniou, A., I. Garrett, and R. Priestley. 1998. "Macroeconomic Variables as Common Pervasive Risk Factors and the Empirical Content of the Arbitrage Pricing Theory." *Journal of Empirical Finance* 5 (3): 221–240. https://doi.org/10.1016/S0927-5398(97)00019-4.

Avramov, D., S. Cheng, L. Metzker, and S. Voigt. 2023. "Integrating Factor Models." *The Journal of Finance* 78 (3): 1593–1646. https://doi.org/10.1111/jofi.13226.

Barillas, F., R. Kan, C. Robotti, and J. Shanken. 2020. "Model Comparison with Sharpe Ratios." *Journal of Financial and Quantitative Analysis* 55 (6): 1840–1874. https://doi.org/10.1017/S0022109019000589.

Barillas, F., and J. Shanken. 2017. "Which Alpha?" *Review of Financial Studies* 30 (4): 1316–1338. https://doi.org/10.1093/rfs/hhw101.

Barillas, F., and J. Shanken. 2018. "Comparing Asset Pricing Models." *The Journal of Finance* 73 (2): 715–754. https://doi.org/10.1111/jofi.12607.

Barillas, F., and J. Shanken. 2022. "Comparing Priors for Comparing Asset Pricing Models." *Journal of Finance, Replications and Comments*.

Bryzgalova, S., J. Huang, and C. Julliard. 2023. "Bayesian Solutions for the Factor Zoo: We Just Ran Two Quadrillion Models." *The Journal of Finance* 78 (1): 487–557. https://doi.org/10.1111/jofi.13197.

Burnside, C. 2016. "Identification and Inference in Linear Stochastic Discount Factor Models with Excess Returns." *Journal of Financial Econometrics* 14 (2): 295–330. https://doi.org/10.1093/jjfinec/nbv018.

Chib, S. 1995. "Marginal Likelihood from the Gibbs Output." *Journal of the American Statistical Association* 90 (432): 1313–1321. https://doi.org/10.1080/01621459.1995.10476635.

Chib, S., Y. C. Lin, K. Pukthuanthong, and X. Zeng. 2023. "Slope Factors Outperform: Evidence from a Large Comparative Study." Working Paper. Washington University in St Louis.

Chib, S., and X. Zeng. 2020. "Which Factors are Risk Factors in Asset Pricing? A Model Scan Framework." *Journal of Business & Economic Statistics* 38 (4): 771–783. https://doi.org/10.1080/07350015.2019.1573684.

Chib, S., X. Zeng, and L. Zhao. 2020. "On Comparing Asset Pricing Models." *The Journal of Finance* 75 (1): 551–577. https://doi.org/10.1111/jofi.12854.

Chib, S., L. Zhao, and G. Zhou. 2021. "Finding Risk-Factors . . . . Without Using Old Data." Working Paper. Washington University in St Louis.

Chib, S., L. Zhao, and G. Zhou. 2023. "Winners from Winners: A Tale of Risk Factors." *Management Science*.

Clare, A. D., P. N. Smith, and S. H. Thomas. 1997. "UK Stock Returns and Robust Tests of Mean Variance Efficiency." *Journal of Banking & Finance* 21 (5): 641–660. https://doi.org/10.1016/S0378-4266(96)00058-1.

Cochrane, J. H. 2005. *Asset Pricing*. Revised ed. Princeton, NJ: Princeton University Press.

Daniel, K. D., D. Hirshleifer, and L. Sun. 2020. "Short- and Long-Horizon Behavioral Factors." *The Review of Financial Studies* 33 (4): 1673–1736. https://doi.org/10.1093/rfs/hhz069.

Dybvig, P., and J. E. Ingersoll. 1982. "Mean-variance Theory in Complete Markets." *Journal of Business* 55 (2): 233–251. https://doi.org/10.1086/296162

Ehansi, S., and J. T. Linnainmaa. 2021. "Time-series Efficient Factors." Working Paper. Tuck School of Business.

Fama, E. F., and K. R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1): 3–56. https://doi.org/10.1016/0304-405X(93)90023-5.

Fama, E. F., and K. R. French. 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 116 (1): 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010.

Fama, E. F., and K. R. French. 2018. "Choosing Factors." *Journal of Financial Economics* 128 (2): 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012.

Ferson, W. E. 2019. *Empirical Asset Pricing: Models and Methods*. Cambridge, MA: MIT Press.

Ferson, W. E., and R. Jagannathan. 1996. "Econometric Evaluation of Asset Pricing Models." In *Handbook of Statistics: Statistical Methods in Finance*, edited by G. S. Maddala and C. R. Rao, 1–33. Amsterdam: Elsevier Science.

Ferson, W. E., A. F. Siegel, and J. L. Wang. 2022. "Factor Model Comparisons with Conditioning Information." Working Paper. University of Southern California.

Fletcher, J. 1994. "The Mean-Variance Efficiency of Benchmark Portfolios: UK Evidence." *Journal of Banking & Finance* 18 (4): 673–685. https://doi.org/10.1016/0378-4266(94)00014-X.

Fletcher, J. 2001. "An Examination of Alternative Factor Models in UK Stock Returns." *Review of Quantitative Finance and Accounting* 16 (2): 117–130. https://doi.org/10.1023/A:1011270907471.

Fletcher, J. 2019. "Model Comparison Tests of Linear Factor Models in U.K. Stock Returns." *Finance Research Letters* 28: 281–291. https://doi.org/10.1016/j.frl.2018.05.005.

Foye, J. 2018. "Testing Alternative Versions of the Fama–French Five-Factor Model in the UK." *Risk Management* 20 (2): 167–183. https://doi.org/10.1057/s41283-018-0034-3.

Frazzini, A., and L. H. Pedersen. 2014. “Betting against Beta.” *Journal of Financial Economics* 111 (1): 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005.

Gospodinov, N., R. Kan, and C. Robotti. 2016. “On the Properties of the Constrained Hansen-Jagannathan Distance.” *Journal of Empirical Finance* 36: 121–150. https://doi.org/10.1016/j.jempfin.2015.10.001.

Green, J., J. R. Hand, and X. F. Zhang. 2017. “The Characteristics That Provide Independent Information About Average U.S. Monthly Stock Returns.” *The Review of Financial Studies* 30 (12): 4389–4436. https://doi.org/10.1093/rfs/hhx019.

Greenberg, E. 2013. *Introduction to Bayesian Econometrics* (2nd ed.). Cambridge: Cambridge University Press.

Gregory, A., R. Tharyan, and A. Christidis. 2013. “Constructing and Testing Alternative Versions of the Fama-French and Carhart Models in the UK.” *Journal of Business Finance & Accounting* 40 (1-2): 172–214. https://doi.org/10.1111/jbfa.12006.

Hansen, L. P., and R. Jagannathan. 1991. “Implications of Security Market Data for Models of Dynamic Economies.” *Journal of Political Economy* 99 (2): 225–262. https://doi.org/10.1086/261749.

Hansen, L. P., and S. F. R. Richard. 1987. “The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models.” *Econometrica* 55 (3): 587–613. https://doi.org/10.2307/1913601.

Harrison, M., and D. Kreps. 1979. “Martingales and Arbitrage in Multi-Period Securities Markets.” *Journal of Economic Theory* 20 (3): 381–408. https://doi.org/10.1016/0022-0531(79)90043-7.

Harvey, C. R. 2017. “Presidential Address: The Scientific Outlook in Financial Economics.” *The Journal of Finance* 72 (4): 1399–1440. https://doi.org/10.1111/jofi.12530.

Harvey, C. R. 2020. “Replication in Financial Economics.” *Critical Finance Review* 8: 1–9. https://doi.org/10.1561/104.00000080

Harvey, C. R., and Y. Liu. 2021. “Lucky Factors.” *Journal of Financial Economics* 141 (2): 413–435. https://doi.org/10.1016/j.jfineco.2021.04.014.

Harvey, C. R., Y. Liu, and H. Zhu. 2016. “ … and the Cross-Section of Expected Returns.” *Review of Financial Studies* 29 (1): 5–68. https://doi.org/10.1093/rfs/hhv059.

Hou, K., C. Xue, and L. Zhang. 2015. “Digesting Anomalies: An Investment Approach.” *Review of Financial Studies* 28 (3): 650–705. https://doi.org/10.1093/rfs/hhu068.

Hou, X., C. Xue, and L. Zhang. 2020. “Replicating Anomalies.” *The Review of Financial Studies* 33 (5): 2019–2133. https://doi.org/10.1093/rfs/hhy131.

Jeffreys, H. 1961. *Theory of Probability*. 3rd ed. Oxford: Clarendon Press.

Jensen, M. C. 1968. “The Performance of Mutual Funds in the Period 1945-1964.” *Journal of Finance* 23: 389–416.

Kan, R., C. Robotti, and J. Shanken. 2013. “Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology.” *The Journal of Finance* 68 (6): 2617–2649. https://doi.org/10.1111/jofi.12035.

Kozak, S., S. Nagel, and S. Santosh. 2020. “Shrinking the Cross-Section.” *Journal of Financial Economics* 135 (2): 271–292. https://doi.org/10.1016/j.jfineco.2019.06.008.

Li, H., Y. Xu, and X. Zhang. 2010. “Evaluating Asset Pricing Models Using the Second Hansen-Jagannathan Distance.” *Journal of Financial Economics* 97 (2): 279–301. https://doi.org/10.1016/j.jfineco.2010.03.002.

Michou, M., and H. Zhou. 2016. “On the Information Content of new Asset Pricing Factors in the UK.” Working Paper. University of Edinburgh.

Pukthuanthong, K., Z. Qiao, and Y. Wang. 2023. “No one-Size-Fits-all Tale: The Diversity and Complexity in Asset Pricing across Global Markets.” Working Paper. University of Missouri.

Ross, S. A. 1978. “A Simple Approach to the Valuation of Risky Streams.” *Journal of Business* 51: 153–475.

Sharpe, W. F. 1966. “Mutual Fund Performance.” *The Journal of Business* 39 (S1): 119–138. https://doi.org/10.1086/294846.

Stambaugh, R. F., and Y. Yuan. 2017. “Mispricing Factors.” *Review of Financial Studies* 30 (4): 1270–1315. https://doi.org/10.1093/rfs/hhw107.

Wang, Z., and X. Zhang. 2012. “Empirical Evaluation of Asset Pricing Models: Arbitrage and Pricing Errors in Contingent Claims.” *Journal of Empirical Finance* 19 (1): 65–78. https://doi.org/10.1016/j.jempfin.2011.11.001.