Review of Quantitative Finance and Accounting
https://doi.org/10.1007/s11156-024-01286-0

ORIGINAL RESEARCH

AN examination of linear factor models in U.K. stock returns in the presence of dynamic trading

Jonathan Fletcher[1]

Accepted: 18 April 2024

Abstract

This study uses the approach of Ferson and Siegel, Rev Financ Stud 22:2735–2758 (2009), and Ferson, Siegel and Wang, J Financ Quant Anal, forthcoming, (2024) to examine the unconditional mean–variance efficiency, in the presence of conditioning information (UMV), of ten linear factor models in U.K. stock returns. The study finds that the UMV efficiency of all the multifactor models is strongly rejected in U.K. stock returns in two different sets of test assets. This rejection is mainly driven by allowing dynamic trading in the test assets and factors. The optimal use of conditioning information also has a significant impact in relative model comparison tests. In relative model comparison tests based on UMV efficiency, the best performing model is the eight-factor model of Chib and Zeng, J Bus Econ Stat 38:771–783 (2020) model.

Keywords Multi-factor models · Asset pricing · Conditioning information · Dynamic trading

JEL Classification G11 · G12

1 Introduction

Mean–variance analysis developed by Markowitz (1952) has long played an important role in a number of areas in Finance. One of these areas is in the testing of asset pricing models. Roll (1977) shows that the central prediction of the capital asset pricing model (CAPM) is that the market portfolio lies on the ex ante mean–variance frontier. Chamberlain (1983) and Grinblatt and Titman (1987) show that for multifactor models, a combination of the K factor portfolios lie on the mean–variance frontier. Ferson (2019) points out that any candidate stochastic discount factor model, whether linear or nonlinear, implies that the portfolio with the maximum squared correlation portfolio to the stochastic discount factor lies on the mean–variance frontier (Hansen and Richard (1987)).

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

[1] Department of Accounting and Finance, University of Strathclyde, Stenhouse Wing, 199 Cathedral Street, Glasgow G4 0QU, UK

The classic test of mean–variance efficiency in the presence of a risk-free asset was developed by Gibbons et al. (1989) (GRS). The GRS test examines the mean–variance efficiency of a linear factor model relative to the efficient frontier where the optimal strategies are fixed-weight portfolios (passive mean–variance (PMV) frontier). The heart of the GRS test compares the maximum squared Sharpe (1966) performance of the factors to the maximum squared Sharpe performance of the factors and test assets to see if there is a significant shift. Barillas and Shanken (2017) extend this analysis and show that when it comes to relative model comparison tests, the choice of test assets is irrelevant and models can be compared in terms of the maximum squared Sharpe measures of the factors in each model. The better models are the ones with higher maximum squared Sharpe measures.

Ferson and Siegel (2009) extend the mean–variance efficiency tests of Gibbons et al. (1989) to allow dynamic trading strategies through the optimal use of conditioning information, building on the work by Hansen and Richard (1987), and Ferson and Siegel (2001). Hansen and Richard (1987) define the unconditional mean–variance frontier (UMV) in the presence of conditioning information where an investor can follow a dynamic trading strategy[1] to maximize the unconditional risk and return trade-off. Ferson and Siegel (2001, 2015) derive the closed-form solutions to UMV optimal portfolios.

Ferson and Siegel (2009) show that every asset pricing model makes a prediction about a portfolio (or combination of portfolios) that lie on the UMV frontier. Testing UMV efficiency represents a higher hurdle for asset pricing models to pass as models are required to correctly price not only fixed-weight portfolio strategies but also all portfolio strategies (satisfying the budget constraint) that can depend upon conditioning information. This approach compares the maximum squared Sharpe measure of the factors to the maximum squared Sharpe measure of the UMV frontier of the test assets and factors. Ferson et al. (2024) also extend the arguments of Barillas and Shanken (2017) and show that the maximum squared Sharpe measure of the UMV frontier of the factors of different models can be used in relative model comparison tests.

Ferson and Siegel (2009) use simulation analysis to test the UMV efficiency of factor models in U.S. stock returns and are able to reject unconditional and conditional versions of the CAPM and Fama and French (1993) models. Ferson et al. (2024) derive the asymptotic distribution of tests based on the maximum squared Sharpe measures of the UMV frontier, and the corresponding standard errors. These can be used to calculate *t*-statistics of the UMV efficiency tests of linear factor models, and to conduct relative model comparison tests.

This study examines the UMV efficiency of ten multifactor models in U.K. stock returns and to conduct relative model comparison tests. A focus on the U.K. is important for a number of reasons. A recent study by Pukthuanthong et al. (2023) find that the factors in the best model from a Bayesian model scan can be country specific. Dimson et al. (2015) find that the industrial compositions of the U.K. and U.S. markets can vary. The mining, oil, and gas sectors play a bigger role in the U.K. and the technology sector plays a smaller role relative to the U.S. market. The study of UMV efficiency of factor models is important in the evaluation of the performance of U.K. equity mutual funds. Ferson (2013) show that UMV efficient portfolios is an "Appropriate Benchmark" to use for clients with quadratic

[1] This approach allows the conditional expected returns and covariance matrix to change in response to lagged information variables.

utility functions.[2] The results of the study suggests whether any of the factor models are Appropriate Benchmarks in this context. This is the first study to examine the UMV efficiency of linear factor models in U.K. stock returns,[3] and complements the studies in U.S. stock returns such as Ferson and Siegel (2009), Penaranda (2016), and Ferson et al. (2024). Recent studies by Harvey (2019) and Hou et al. (2020) highlight the importance of replication studies in Finance.

The sample period is between July 1983 and December 2022. The models include the three-factor model of Fama and French (1993), the five-factor model of Fama and French (2015), the six-factor model of Fama and French (2018), the four-factor model of Hou et al. (2015), the three-factor model of Clarke (2022), the two-factor model of Frazzini and Pedersen (2014), the four-factor model of Stambaugh and Yuan (2017), and the best factor models drawn from Bayesian model scan studies of Barillas and Shanken (2018), Chib and Zeng (2020), and Chib et al. (2024). I use two sets of test assets in 16 size/book-to-market (BM) portfolios, and 15 volatility/momentum portfolios.

There are three main findings in my study. First, the UMV efficiency of all the factor models is rejected. When dynamic trading is allowed in the factors, the UMV efficiency is no longer rejected for some of the models using the volatility/momentum portfolios as the test assets. Second, the rejection of UMV efficiency is driven mainly by allowing dynamic trading in the test assets and factors. Third, the best performing model in the relative model comparison tests using the UMV frontiers is the Chib and Zeng (2020) model.

The paper is organized as follows. Section II presents the research method. Section III describes the data used in my study. Section IV reports the empirical results. The final 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) exists in financial markets, then there exists a stochastic discount factor[4] (m_{t+1}) such that:

$$P_t = E\left(m_{t+1}X_{t+1} \middle| Z_t\right) \tag{1}$$

where p_t are the costs of the N test assets at t, X_{t+1} are the payoffs of the N test assets at $t+1$, and Z_t is the information set of investors at time t. Where there are No Arbitrage (NA) opportunities available in financial markets, then $m_{t+1} > 0$ (Cochrane (2005)).[5] When the payoffs are gross returns (1 + returns), then Eq. (1) becomes:

[2] Ferson and Siegel (2009) find that hedge fund indexes are able to outperform the PMV frontier but not the UMV frontier.

[3] A partial list of prior U.K. studies includes Fletcher (1994, 2001, 2019), Davies, Fletcher and Marshall (2015), Gregory, Tharyan and Christidis (2013) among others.

[4] Cochrane (2005) and Ferson (2019) provide excellent textbook treatments of the stochastic discount factor approach to asset pricing.

[5] The analysis can be extended to incorporate market frictions such as no short selling constraints, and transaction costs (He and Modest (1995), Luttmer (1996), Hansen et al. (1995), De Roon, Nijman and Werker (2001), and Korsaye et al. (2021) among others).

$$1 = E\left(m_{t+1}R_{t+1}\middle|Z_t\right) \tag{2}$$

Ferson and Siegel (2009) show if we restrict portfolio strategies such that the weights sum to 1 at each point in time, and take unconditional expectations, then Eq. (2) becomes:

$$E\left(m_{t+1}x'\left(Z_t\right)R_{t+1}\right) = 1 \text{ for all } x'\left(Z_t\right)e = 1 \tag{3}$$

where $x'(Z_t)$ is a (N,1) vector of portfolio weights that can depend upon Z_t, and e is a (N,1) vector of ones. In Eq. (3) asset pricing models are required to price not only the test assets but also all dynamic trading strategies that trade on Z_t subject to the restriction that the weights sum to 1.

Ferson and Siegel (2009) show that if a candidate stochastic discount factor model satisfies Eq. (3), then it implies that a certain portfolio lies on the UMV frontier in the presence of conditioning information. The UMV frontier is defined as in Hansen and Richard (1987) as a portfolio (R_{pt+1}) such that:

$$\mathrm{VAR}\left(R_{pt+1}\right) \leq \mathrm{Var}\left(x'\left(Z_i\right)R_{i+1}\right) \text{ if } E\left(R_{pt+1}\right) = E\left(x'\left(Z_i\right)R_{i+1}\right) \text{ and } x'\left(Z_t\right)e = 1 \tag{4}$$

The candidate stochastic discount factor models examined in this study are linear factor models given by:

$$m_{t+1} = a + b_k f_{t+1} \tag{5}$$

where f_{t+1} is a (K,1) vector of the K excess factor returns at time t+1, and b_K is a (1,K) vector of slope coefficients on the K factors in the model. The individual slope coefficients in b_K tell us whether the factor is important for pricing the test assets given the other factors in the model. Proposition 2 in Ferson and Siegel (2009) show that if Eq. (3) is satisfied by a linear factor model in Eq. (5), then there will be a combination of the K factor portfolios that lie on the UMV frontier.

The UMV efficiency of a linear factor model can be tested by comparing squared Sharpe (1966) measures. Define r as the N test assets, f as the K factors in a model, and Sh2 as the squared Sharpe measure. The null hypothesis is given by:

$$\mathrm{Sh2Diff} = \mathrm{Sh2}_{umv}(r, f) - \mathrm{Sh2}_{umv}(f) = 0 \tag{6}$$

where $\mathrm{Sh2}_{umv}(r,f)$ is the maximum squared Sharpe measure from the UMV frontier of the test assets and factors, and $\mathrm{Sh2}_{pmv}(f)$ is the maximum squared Sharpe measure from the PMV frontier of the factors. To estimate the Sharpe measures, a zero-beta return is required[6] and it is assumed in this study to be equal to the average return of the one-month Treasury Bill as in Ferson and Siegel (2009) and Ferson et al. (2024).

Ferson and Siegel (2001, 2015) derive the closed-form solutions of the optimal weights of the UMV frontier. Define u_t as a (N,1) vector of the conditional expected returns of the assets based on information at time t, V_t is the (N,N) conditional covariance matrix, Z_t is a (L,1) vector of lagged information variables (including a constant), and L_t is the (N,N) inverse conditional second moment matrix and is equal to $(V_t+u_tu_t')^{-1}$. Define $D_t=L_t - (L_tee'L_t)/(e'L_te)$, $\alpha_1=E(1/e'L_te)$, $\alpha_2=E((e'L_tu_t)/(e'L_te))$, and $\alpha_3=E((u_t'D_tu_t)/(e'L_te)$. The optimal weights are given by:

[6] See Ferson (2019).

$$x(Z_t) = (L_t e / e\prime L_t e) + ((u_p - \alpha_2)/\alpha_3) D_t u_t \tag{7}$$

where u_p is the target expected return.[7]

The first term in Eq. (7) is the minimum conditional second moment portfolio where the weights sum to 1. The second term are the excess returns on the mean–variance component where the weights sum to zero. By varying the target u_p, any point on the UMV frontier can be selected.[8] Ferson and Siegel (2001) show that investors with a quadratic utility function will select UMV portfolios. Ferson and Siegel (2001) point out that the UMV optimal portfolio weights are conservative for extreme values of Z_t. For a client who does not observe the information of the fund manager they would want the fund manager to hold the UMV portfolio.[9]

Given a model of conditional moments, the optimal weights can be estimated, and the corresponding squared Sharpe measures can be calculated. Ferson et al. (2024) show that the squared Sharpe measure on the UMV frontier can be calculated as $Sh2 = a - 2br_z + cr_z^2$, where $a = [(\alpha_2^2 + \alpha_1\alpha_3)/(\alpha_1(1-\alpha_3) - \alpha_2^2)]$, $b = [\alpha_2/(\alpha_1(1-\alpha_3) - \alpha_2^2)]$, and $c = [(1-\alpha_3)/(\alpha_1(1-\alpha_3) - \alpha_2^2)]$. Ferson and Siegel (2009), and Ferson et al. (2024) use a predictive regression of the asset returns on Z_t to model the conditional moments. The conditional expected returns are the fitted values from the regression, and the conditional covariance matrix is assumed constant and given by the residual covariance matrix from the regression. Ferson and Siegel (2009) point out the tests are robust to using the wrong model of conditional moments. The UMV portfolio is still a valid portfolio strategy but no longer the optimal one. Ferson and Siegel (2009) note that this leads to a loss in power. Ferson and Siegel (2009) use simulations to test the UMV efficiency of a factor model. Ferson et al. (2024) derive the asymptotic distribution of the test of Eq. (6) through Theorem I and Corollary I.[10] These can be used to calculate the standard error of the Sh2 Diff measure, and the corresponding *t*-statistic to evaluate the null hypothesis.

To provide further insight into the UMV efficiency tests, Ferson et al. (2024) consider two decompositions. The first decomposition is given by:

$$\text{Sh2Diff} = \text{Sh2Diff} + \text{Sh2Diff2} \tag{8}$$

where Sh2 Diff1 = $Sh2_{pmv}$(r,f)–$Sh2_{pmv}$(f), Sh2 Diff2 = $Sh2_{umv}$(r,f)–$Sh2_{pmv}$(r,f), where $Sh2_{pmv}$(r,f) is the maximum squared Sharpe measure from the PMV frontier of the test assets and factors. The first term in the decomposition in Eq. (8) is a test of PMV efficiency of the factor model, and the second term captures the impact of allowing dynamic trading in the test assets and factors. The second decomposition is given by:

$$\text{Sh}^2\text{Diff} + \text{Sh2Diff3} + \text{Sh2Diff4} \tag{9}$$

where Sh2 Diff3 = Sh^2_{umv}(r,f)–Sh^2_{umv}(f), Sh2 Diff4 = Sh^2_{umv}(f)–Sh^2_{pmv}(f), where $Sh2_{umv}$(f) is the maximum squared Sharpe measure from the UMV frontier of the factors. The first term in the decomposition in Eq. (9) is the UMV efficiency test of the factor model, where dynamic trading is allowed in the factors. Ferson and Siegel(2009) point out that this is

[7] The target expected return depends upon the choice of the zero-beta return.

[8] Hansen and Richard (1987) show that portfolios on the UMV frontier also lie on the conditional mean–variance frontier but the converse is not generally true.

[9] Dybvig and Ross (1985) show that a manager who selects a portfolio on the conditional mean-wherevariance frontier can appear inefficient from the perspective of the client.

[10] The relevant Corollary for the PMV frontier is in the Appendix of Ferson et al. (2024).

a test of dynamic mean–variance intersection along the lines of Huberman and Kandel (1987). The second term captures the impact in dynamic trading in the factors, and estimates the increase in the maximum squared Sharpe performance of the factors through the optimal use of conditioning information. The Sh2 Diff measures can be calculated for the four terms, and the corresponding *t*-statistics using Theorem I and relevant Corollaries in Ferson et al. (2024).

Ferson et al. (2024) extend the analysis of Barillas and Shanken (2017) to relative model comparison tests using the UMV fronter. Better models have higher maximum squared Sharpe measures from the UMV frontier. Define two factor models f_A, and f_B. The null hypothesis in relative model comparison tests is:

$$\mathrm{Sh2Diff} = \mathrm{Sh2}_{\mathrm{umv}}\left(f_A\right) - \mathrm{Sh2}_{\mathrm{umv}}\left(f_B\right) = 0 \tag{10}$$

The *t*-statistic of the Sh2 Diff measure in the null hypotheses in Eq. (10) can be calculated using Theroem I and the relevant Corollaries in Ferson et al. (2024).

One issue that arises when using the maximum squared Sharpe measures to test and compare factor models is that there is a large upward bias in the sample maximum squared Sharpe measure (Jobson and Korkie (1980)). Ferson and Siegel (2003) in their study of Hansen and Jagannathan (1991) volatility bounds use a bias adjusted maximum squared Sharpe measure given by:

$$\mathrm{Sh2}_b = \mathrm{Sh2}^*((T - N - 2)/T) - N/T \tag{11}$$

The adjusted $\mathrm{Sh2}_b$ works well when evaluating models using the PMV frontier. Ferson and Siegel (2003) find that the bias adjustment works less well when using the UMV frontier. Proposition II in Ferson et al. (2024) derives a bias adjustment of the maximum squared Sharpe measures of UMV portfolios based on the method of statistical differentials (see Siegel and Woodgate (2007)). Simulation evidence in Ferson et al. (2024) suggests that their bias adjustment works well in testing factor models in U.S. stock returns, and performs better than alternative bias adjustment methods. In this study, I use the adjusted maximum squared Sharpe measures for UMV portfolios based on Proposition II in Ferson et al. (2024), and Sh^2_b for the PMV frontier.

3 Data

3.1 A) Test assets

The sample period covers between July 1983 and December 2022. I use two sets of test assets in the study. Details on the formation of the test assets are included in the Appendix. The first set is 16 size/BM portfolios, where the stocks are sorted by market value (Small to Big), and the BM ratio (Growth to Value). The portfolios are reformed annually and are value weighted portfolio returns. The data for forming the size/BM portfolios is collected from the London Share Price Database (LSPD) provided by the London Business School, and Refinitiv Worldscope.

The second set of test assets is motivated from Kirby and Ostdiek (2012a, b), and is 15 portfolios sorted by volatility (Low to High), and momentum (Losers to Winners). The volatility/momentum portfolios are formed monthly and are value weighted portfolio returns. The data for forming the volatility/momentum portfolios is collected from LSPD. I use the return on the one-month U.K. Treasury Bill as the risk-free asset, which I collect from LSPD and Datastream.

3.2 B) Factor models

I consider the performance of ten different linear factor models. Fama and French (2018) argue for using a small number of linear factor models in relative model comparison tests to mitigate the impact of data dredging issues. The factors are formed using data on LSPD and Worldscope. Details on how the factor models are formed are included in the Appendix. The following factor models are used.

1. Fama and French (1993) (FF3).
   The FF3 model is a three-factor model, which includes the excess market returns, and two zero-cost portfolios that capture the size (SMB), and value (HML) effects in stock returns.
2. Fama and French (2015) (FF5).
   The FF5 model is a five-factor model, which includes the FF3 factors and adds two zero-cost portfolios that capture the profitability (RMW), and investment (CMA) effects in stock returns.
3. Fama and French (2018) (FF6).
   The FF6 model is a six-factor model, which includes the FF5 factors, and a zero-cost portfolio that captures the momentum (MOM) effect in stock returns.
4. Clarke (2022) (LSC).
   The LSC model is a three-factor model, which includes the excess returns of a Level, Slope, and Curve factors in stock returns.[11]
5. Hou et al. (2015) (HXZ).
   The HXZ model is a four-factor model, which includes the excess market return and three zero-cost portfolios that capture the size (ME), profitability (ROE), and investment (IA) effects in stock returns.
6. Frazzini and Pedersen (2014) (FP).
   The FP model is a two-factor model which includes the excess market returns, and the Betting against Beta (BAB) factor.
7. Stambaugh and Yuan (2017) (SY).
   The SY model is a four-factor model, which includes the excess market return, and zero-cost portfolios for the size, (SIZE), management (MGMT), and performance (PERF) factors.
   The final three models are selected from recent Bayesian model scan studies of Barillas and Shanken (2018), Chib and Zeng (2020), and Chib et al. (2024) in U.S. stock returns. The model scan searches for the best model which has the highest posterior probability (log Marginal Likelihood) among a set of factors.
8. Barillas and Shanken (2018) (BS).
   The BS model is a six-factor model, and includes the excess market return, and zero-cost portfolios for the size (SMB), value (HMLT),[12] profitability (ROE), investment (IA), and momentum (MOM) factors.

[11] The LSC model is an equity version of the corresponding factors in bond returns as in Litterman and Scheinkman (1991).

[12] The value factor here is the more timely value factor of Asness and Frazzini (2013).

9. Chib and Zeng. (2020) (CZ).
   The CZ model is an eight-factor model.[13] The model includes the excess market returns, and zero-cost portfolios including the SMB, HMLT, RMW, ROE, MOM, BAB, and the Quality minus Junk (QMJ)[14] factors.
10. Chib et al. (2024) (CZZ).
   The CZZ model is a seven-factor model. The model includes the excess market returns, and zero-cost portfolios including the SMB, ROE, MOM, MGMT, PERF, and Post Earnings Announcement Drift (PEAD)[15] factors.

Table 1 reports summary statistics of the test assets and the factors between July 1983 and December 2022. Panel A of Table 1 includes 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 for the different factors. Panel B of Table 1 reports the average excess returns (%) of the size/BM, and volatility/momentum portfolios.

Panel A of Table 1 shows that most of the factors have significant positive average excess returns. The main exception are the size factors (SMB, ME, Size), the ROE and HMLT factors. The MOM factor has the largest average excess return across factors at 0.760%, highlighting the strong momentum effect in U.K. stock returns, followed by the BAB factor at 0.590%. The MGMT and PERF factors in the SY model also have substantial average excess returns. There is a significant investment effect in the FF5 and HXZ models, using the CMA, and IA factors. It is only the CMA and MOM factors that have a *t*-statistic larger than 3, which is the recommended cut-off *t*-statistic by Harvey et al. (2016) to control for multiple testing.

Panel B of Table 1 shows that there is a wide spread in average excess returns for both sets of test assets. The average excess returns of the size/BM portfolios range between 0.141% (Small/Growth), and 0.722% (Small/Value). The value effect is stronger in smaller companies, which is consistent with Fama and French (2012). There is a small size effect in the Value portfolios, and a reverse size effect in the Growth portfolios.

The spread in average excess returns in panel B of Table 1 is a lot wider in the volatility/momentum portfolios compared to the size/BM portfolios. The average excess returns of the volatility/momentum portfolios range between -0.751% (High/Losers), and 0.914% (4/Winners). There is a large momentum effect in average excess returns across all volatility groups. There is likewise a volatility effect in average excess returns, for the Losers and 2 portfolios, where the Low volatility portfolio has a much higher average excess return than the High volatility portfolio. The volatility effect is a lot stronger when we look the standard deviations of the volatility/momentum portfolios.

[13] Chib and Zeng (2020) use a multivariate *t*-distribution in their model scan and Barillas and Shanken (2018) assume multivariate normality. Barillas and Shanken allow only one type of each factor in the model. Chib, Zeng and Zhao (2020) provide a critique of the Barillas and Shanken approach. See Barillas and Shanken (2022) for a response.

[14] The QMJ factor is developed in Asness, Frazzini and Pedersen (2019).

[15] The PEAD factor is proposed in the recent behavioral factor model of Daniel, Hirshleifer and Sun (2020). Bryzgalova et al. (2023a, b) highlight the importance of the PEAD factor in U.S. stock returns.

3.3 C) Lagged information variables

I use four lagged information variables that earlier studies have found to have some predictive ability of future stock returns.[16] The lagged information variables include the lag one-month annualized dividend yield (DY) of the U.K. market index (Fama and French (1988)), lag return on the one-month U.K. Treasury Bill return (Rf) (Fama and Schwert (1977), Ferson (1989)), the lag one-month term spread (Term) given by the difference in the annualized yields of the long-term government bonds (International Financial Statistics), and the three-month U.K. Treasury Bill (LSPD), and the lag one-month excess return on U.K. market index. The lag DY is formed using data from LSPD.

To examine the predictive ability of the lagged information variables, I run predictive regressions for both sets of test assets of the excess asset returns on a constant and the four lagged information variables in unreported tests.[17] It is only for the size/BM portfolios that the Wald test rejects the null hypothesis that the slope coefficients on the lagged information variables are jointly equal to zero. The magnitude of the predictability is small in statistical terms with the highest adjusted R^2 is 8.13% (Small/Value) in the size/BM portfolios, and 2.74% (High/2) in the volatility/momentum portfolios.

4 Empirical results

I begin the empirical analysis by testing the UMV efficiency of the linear factor models. Table 2 reports the difference in adjusted maximum squared Sharpe measures (Sh2 Diff) between the UMV frontier of the test assets and factors and the PMV frontier of the factors, and the corresponding *t*-statistics. An earlier version of Ferson et al. (2024) point out that the increase in squared Sharpe performance can be interpreted in economic terms using maximum quadratic utilities (Kan and Zhou (2007)). The Certainty Equivalent (CE) excess returns is given by (1/2γ)Sh2 Diff where γ is risk aversion level. The CE in Table 2 assumes a risk aversion level of 5 as in Ferson et al. (2024). Panel A includes the results using the size/BM portfolios as the test assets, and panel B includes the results using the volatility/momentum portfolios as the test assets.

Table 2 shows that the UMV efficiency of each factor model is strongly rejected in both sets of test assets. There is a large significant increase in the adjusted maximum squared Sharpe performance between the UMV frontier of the test assets and factors, and the PMV frontier of the factors. The Sh2 Diff measures range between 0.28 (BS) and 0.377 (CZ) for the size/BM portfolios, and 0.117 (SY) and 0.268 (CZ) for the volatility/momentum portfolios. The magnitude of the CE excess returns is greater than 2.79% for all models using the size/BM portfolios, and greater than 1.17% for all models using the volatility/momentum portfolios. The rejection of the UMV efficiency of the factor models is consistent with Ferson and Siegel (2009), and Ferson et al. (2024) in U.S. stock returns.

Table 2 shows that the UMV efficiency of all the factor models is rejected in both sets of test assets. I next explore what drives the rejection in UMV efficiency by estimating the decompositions of Ferson et al. (2024) in Eqs. (8) and (9). Tables 3 and 4

[16] Rapach and Zhou (2013, 2022), Ferson (2019) chap 32 provide excellent reviews of stock return predictability.

[17] Results are available on request.

Table 1 Summary statistics of test assets and factors

Panel A:				
Factors	Mean	Std Dev	*t*-statistic	
Market	0.437	4.245	2.24[1]	
SMB	0.053	2.970	0.39	
HML	0.261	2.813	2.02[1]	
RMW	0.212	2.131	2.16[1]	
CMA	0.285	1.945	3.19[1]	
MOM	0.760	3.231	5.12[1]	
ME	-0.138	3.341	-0.90	
ROE	-0.031	2.596	-0.26	
IA	0.400	3.636	2.39[1]	
BAB	0.590	5.698	2.25[1]	
Size	-0.259	3.980	-1.41	
MGMT	0.434	3.452	2.74[1]	
PERF	0.431	4.410	2.13[1]	
HMLT	0.163	3.301	1.08	
QMJ	0.249	2.427	2.24[1]	
PEAD	0.205	2.216	2.01[1]	
Panel B:				
Test Assets				
Size/BM	Growth	2	3	Value
Small	0.141	0.408	0.556	0.722
2	0.207	0.329	0.576	0.695
3	0.418	0.411	0.493	0.684
Big	0.352	0.450	0.518	0.560
Volatility/Momentum	Losers	2	Winners	
Low	0.066	0.377	0.558	
2	0.153	0.711	0.713	
3	0.025	0.153	0.871	
4	-0.672	0.246	0.914	
High	-0.751	-0.360	0.622	

[1]Significant at 5%

report the two decompositions of UMV efficiency tests using the size/BM portfolios as the test assets (Table 3), and volatility/momentum portfolios as the test assets (Table 4). The first decomposition is in panel A of each table, and the second decomposition is in panel B. The table reports the differences in adjusted maximum squared Sharpe measures (Sh2 Diff1, Sh2 Diff2, Sh3 Diff3, Sh4 Diff4), and the corresponding *t*-statistics.

Panel A of Table 3 shows that using the size/BM portfolios as the test assets, it is the dynamic trading in both the test assets and factors that drives the rejection of UMV efficiency of the factor models. The Sh2 Diff2 measures are a lot larger than the Sh2Diff1 measures and all are highly statistically significant. The Sh2 Diff1 measures reject the passive mean–variance efficiency of all models at the 10% level, except for the BS and CZ models. The finding that the dynamic trading drives the rejection of UMV efficiency of the factor models is similar to Ferson et al. (2024).

Table 2 UMV Efficiency tests of linear factor models

Panel A:			
Size/BM	Sh2 Diff	*t*-statistic	CE
FF3	0.3	5.5[1]	2.996
FF5	0.297	5.36[1]	2.966
FF6	0.299	5.23[1]	2.989
LSC	0.298	5.53[1]	2.984
HXZ	0.297	5.42[1]	2.971
FP	0.3	5.6[1]	3.003
SY	0.337	5.8[1]	3.366
BS	0.28	5.01[1]	2.799
CZ	0.377	6.14[1]	3.769
CZZ	0.340	5.61[1]	3.400
Panel B:			
Volatility/Momentum	Sh2 Diff	*t*-statistic	CE
FF3	0.167	3.3[1]	1.67
FF5	0.196	3.71[1]	1.962
FF6	0.155	3.16[1]	1.553
LSC	0.166	3.23[1]	1.663
HXZ	0.156	3.15[1]	1.558
FP	0.19	3.87[1]	1.897
SY	0.117	2.49[1]	1.174
BS	0.146	3[1]	1.457
CZ	0.268	4.77[1]	2.681
CZZ	0.138	2.89[1]	1.390

[1]Significant at 5%

Panel B of Table 3 shows that allowing dynamic trading in the factors, there is a significant increase in the maximum adjusted squared Sharpe measures of the factors, as reflected in the significant positive Sh2 Diff4 measures. This is especially the case for the CZ model. This result provides support for the optimal use of conditioning information in the factors, which is consistent with Ferson and Siegel (2009), Abhyankar et al. (2012), Penaranda (2016), and Ferson et al. (2024). Although allowing dynamic trading in the factors leads to a significant increase in squared Sharpe performance of the factors, the UMV efficiency of each model is still rejected. The Sh2 Diff3 measures are all large in economic terms and highly statistically significant. This result rejects the dynamin mean–variance intersection of all the models (Huberman and Kandel (1987)).

When using the volatility/momentum portfolios as the test assets, panel A of Table 4 shows again that it is the dynamic trading in the test assets and factors that drives the rejection of the UMV efficiency of the models in most cases. This is especially the case for the FF6, BS, CZ, and CZZ models. All of the Sh2 Diff2 measures are significantly positive at the 10% level. In contrast, the Sh2 Diff1 measures are only significantly positive for FF3, FF5, LSC, HXZ, FP, and SY models. It is interesting to note that the PMV efficiency is not rejected for the BS and CZ models in either set of test assets. Allowing dynamic trading in the factors in panel B of Table 4 shows that the UMV efficiency of the FF6, BS, CZ, and CZZ models is no longer rejected. For these models, the hypothesis of dynamic mean–variance intersection is not rejected.

Table 3 UMV Efficiency decomposition tests: size/bm portfolios

Panel A:				
First Decomposition	Sh2 Diff1	*t*-statistic	Sh2 Diff2	*t*-statistic
FF3	0.052	1.91[2]	0.248	5.29[1]
FF5	0.053	1.88[2]	0.244	5.17[1]
FF6	0.051	1.77[2]	0.247	5.07[1]
LSC	0.051	1.81[2]	0.247	5.35[1]
HXZ	0.048	1.81[2]	0.249	5.23[1]
FP	0.054	1.97[1]	0.246	5.45[1]
SY	0.06	2.05[1]	0.277	5.58[1]
BS	0.027	1.07	0.253	5.19[1]
CZ	0.04	1.41	0.337	6.38[1]
CZZ	0.061	2.01[1]	0.279	5.41[1]
Panel B:				
Second Decomposition	Sh2 Diff3	*t*-statistic	Sh2 Diff4	*t*-statistic
FF3	0.239	4.66[1]	0.061	2.77[1]
FF5	0.192	3.83[1]	0.104	3.93[1]
FF6	0.184	3.58[1]	0.115	4.03[1]
LSC	0.254	4.91[1]	0.045	2.29[1]
HXZ	0.238	4.61[1]	0.059	2.6[1]
FP	0.194	3.96[1]	0.106	4.24[1]
SY	0.3	5.42[1]	0.037	1.89[2]
BS	0.187	3.72[1]	0.093	3.44[1]
CZ	0.162	3.27[1]	0.215	5.74[1]
CZZ	0.236	4.34[1]	0.103	3.59[1]

[1]Significant at 5%

[2]Significant at 10%

Tables 3 and 4 provide some support for the FF6, BS, CZ, and CZZ models when allowing dynamic trading in the factors. I next examine the relative model comparison tests using the maximum adjusted squared Sharpe measures from the PMV and UMV frontiers of the factors. Tables 5 and 6 report the difference between the adjusted maximum squared Sharpe measures (Sh2 Diff) of two factor models (panel A), and corresponding *t*-statistics (panel B). Table 5 reports the relative model comparison tests using the PMV frontier, and Table 6 reports the relative model comparison tests using the UMV frontier. The Sh2 Diff measures in Tables 5 and 6 is the difference between the adjusted maximum squared Sharpe measures of the model in the column and the model in the row.

Table 5 shows that there are a large number of significant differences in the adjusted maximum squared Sharpe measures using the PMV frontier between the factor models. The FF3 model has a significant lower adjusted squared Sharpe measure then the FF5, FF6, BS, CZ, and CZZ models. The FF3, LSC, HXZ, FP, and SY models have similar adjusted squared Sharpe measures. These models all significantly underperform the FF6, BS, CZ, and CZZ models. Among the FF6, BS, CZ, and CZZ models, there are no significant differences in the adjusted squared Sharpe measures. These are the best performing models in the relative model comparison tests based on the PMV frontiers.

Table 6 shows that allowing dynamic trading in the factors has an impact on the relative model comparison tests. There is a sizeable increase in the magnitude of the Sh2

Table 4 UMV Efficiency decomposition tests: volatility/momentum portfolios

Panel A				
First Decomposition	Sh2 Diff1	*t*-statistic	Sh2 Diff2	*t*-statistic
FF3	0.09	2.59[1]	0.077	2.17[1]
FF5	0.094	2.69[1]	0.102	2.72[1]
FF6	0.035	1.24	0.121	3.09[1]
LSC	0.088	2.49[1]	0.078	2.16[1]
HXZ	0.078	2.41[1]	0.078	2.16[1]
FP	0.066	2.12[1]	0.123	3.36[1]
SY	0.059	1.95[2]	0.058	1.66[2]
BS	0.031	1.19	0.115	2.92[1]
CZ	0.039	1.32	0.229	5.01[1]
CZZ	0.020	0.79	0.119	2.92[1]
Panel B				
Second Decomposition	Sh2 Diff3	*t*-statistic	Sh2 Diff4	*t*-statistic
FF3	0.106	2.31[1]	0.061	2.77[1]
FF5	0.092	2.02[1]	0.104	3.93[1]
FF6	0.04	0.98	0.115	4.03[1]
LSC	0.122	2.56[1]	0.045	2.29[1]
HXZ	0.096	2.17[1]	0.059	2.6[1]
FP	0.083	1.93[2]	0.106	4.24[1]
SY	0.08	1.87[2]	0.037	1.89[2]
BS	0.053	1.29	0.093	3.44[1]
CZ	0.053	1.22	0.215	5.74[1]
CZZ	0.035	0.89	0.103	3.59[1]

[1]Significant at 5%

[2]Significant at 10%

Diff measures between models and a larger number of significant Sh2 Diff measures. The FF3 model continues to perform poorly in relative model comparison tests with a significant lower adjusted maximum squared Sharpe measures relative to the FF5, FF6, FP, BS, CZ, and CZZ models. The FP, LSC, HXZ, and SY models have similar performance to one another. The FF5 model significantly outperforms the LSC, HXZ, and SY models but significantly underperforms the FF6, and CZ models. Among the FF6, BS, CZ, and CZZ models, the CZ model has a significant higher adjusted squared Sharpe measure than the FF6 and CZZ models. The CZ model does have a sizeable higher adjusted squared Sharpe measure than the BS model but the difference is not statistically significant. The findings in Table 6 suggest that the CZ model is the best performing model, and complements the empirical results in Chib and Zeng (2020).

My study has used a standard set of portfolios as the test assets. However even in these test assets, the UMV efficiency of the models are rejected. It is likely that if a more challenging set of test assets such as the anomaly portfolios of Jensen et al. (2023) or the approach used by Bryzgalova et al. (2023b), the rejection of the UMV efficiency of the factor models would be even stronger. I have also used a standard set of lagged information variables. One of the attractions of the Ferson and Siegel (2009) approach is that the dimensions of the conditional covariance matrix remains fixed no matter how many lagged information variables are used. I repeat the tests by replacing the lagged excess market

Table 5 PMV Efficiency model comparison tests

Panel A:

Shp2 Diff	FF5	FF6	LSC	HXZ	FP	SY	BS	CZ	CZZ
FF3	0.038	0.11	0.014	0.006	0.001	0.03	0.099	0.137	0.081
FF5		0.072	-0.024	-0.032	-0.037	-0.008	0.061	0.098	0.043
FF6			-0.096	-0.104	-0.109	-0.08	-0.011	0.026	-0.029
LSC				-0.008	-0.013	0.016	0.085	0.123	0.067
HXZ					-0.006	0.024	0.093	0.13	0.075
FP						0.029	0.098	0.136	0.081
SY							0.069	0.106	0.051
BS								0.037	-0.018
CZ									-0.055

Panel B:

t-statistic	FF5	FF6	LSC	HXZ	FP	SY	BS	CZ	CZZ
FF3	2.12[1]	3.39[1]	1.03	0.54	0.09	1.47	3.3[1]	3.94[1]	2.46[1]
FF5		2.64[1]	-1.08	-1.8[2]	-1.92[2]	-0.38	1.81[2]	2.83[1]	1.29
FF6			-2.86[1]	-3.21[1]	-3.48[1]	-2.45[1]	-0.54	0.98	-1.02
LSC				-0.44	-1.04	0.65	2.67[1]	3.36[1]	2.06[1]
HXZ					-0.43	1.35	1.93[2]	3.3[1]	2.55[1]
FP						1.49	3.24[1]	3.54[1]	2.70[1]
SY							2.12[1]	3[1]	1.52
BS								0.34	-0.74
CZ									-1.16

[1] Significant at 5%

[2] Significant at 10%

returns with a lagged default spread. The benefits of the optimal use of conditioning information is a lot weaker in all the factor models, and the CZ model no longer significantly outperforms the FF6, BS, and CZZ models. Although the results can be sensitive to the choice of lagged information variables, the results in the paper are likely to be conservative given that a much broader set of lagged information variables can be used.

5 Conclusions

This paper examines the UMV efficiency of ten multifactor models in U.K. stock returns, and conducts relative model comparison tests. There are three main findings in the study.

First, the UMV efficiency of all the factor models is strongly rejected in both sets of test assets. There is a significant increase in the maximum adjusted squared Sharpe performance in moving from the PMV frontier of the factors to the UMV frontier of the test assets and factors. Allowing dynamic trading in the factors, the UMV efficiency is rejected for all factor models using the size/BM portfolios as the test assets. This result implies the dynamic mean–variance intersection hypothesis (Huberman and Kandel (1987)) is rejected for each factor model. In contrast, using the volatility/momentum portfolios the dynamic mean–variance intersection is only rejected for the FF3, FF5,

Table 6 UMV Efficiency model comparison tests

Panel A:									
Sh2 Diff	FF5	FF6	LSC	HXZ	FP	SY	BS	CZ	CZZ
FF3	0.082	0.165	-0.002	0.005	0.046	0.006	0.131	0.29	0.124
FF5		0.083	-0.084	-0.077	-0.035	-0.076	0.049	0.209	0.042
FF6			-0.167	-0.16	-0.118	-0.158	-0.033	0.126	-0.040
LSC				0.007	0.049	0.009	0.134	0.293	0.126
HXZ					0.041	0.001	0.126	0.286	0.119
FP						-0.04	0.085	0.244	0.077
SY							0.125	0.284	0.118
BS								0.159	-0.007
CZ									-0.166
Panel B:									
t-statistic	FF5	FF6	LSC	HXZ	FP	SY	BS	CZ	CZZ
FF3	3.32[1]	4.31[1]	-0.08	0.18	1.66[2]	0.21	3.78[1]	6.25[1]	1.92[2]
FF5		2.87[1]	-2.49[1]	-2.36[1]	-0.93	-2.29[1]	1.32	4.12[1]	1.02
FF6			-3.82[1]	-3.6[1]	-2.57[1]	-3.75[1]	-1.02	2.65[1]	-1.10
LSC				0.24	1.4	0.27	3.21[1]	5.62[1]	2.89[1]
HXZ					1.27	0.05	2.29[1]	5.24[1]	3.02[1]
FP						-1.17	1.95[2]	4.32[1]	1.73[2]
SY							3.05[1]	5.65[1]	2.64[1]
BS								1.3	-0.22
CZ									-2.66[1]

[1]Significant at 5%

[2]Significant at 10%

LSC, HXZ, FP, and SY models. The rejection of UMV efficiency of the multifactor models is consistent with Ferson and Siegel (2009), and Ferson et al. (2024).

Second, the rejection of the UMV efficiency of the factor models is driven mainly by the dynamic trading in the test assets and factors. All of the Sh2 Diff2 measures are significantly positive and in most cases a lot higher than the Sh2 Diff1 measures. For the BS and CZ models, the PMV efficiency cannot be rejected in either set of test assets. Allowing dynamic trading in the factors leads to a significant increase in the maximum adjusted squared Sharpe measures, with a significant positive Sh2 Diff4 measures for all models. This is especially the case with the CZ model. The importance of dynamic trading is in evaluating factor models is consistent with Ferson and Siegel (2009) and Ferson et al. (2024).

Third, allowing dynamic trading in the factors has a significant impact on the relative model comparison tests. In most cases, there is a sizeable increase in the Sh2 Diff measures using the UMV frontier relative to the PMV frontier, and more of the Sh2 Diff measures are statistically significant. The CZ model has the highest maximum adjusted squared Sharpe measure and significantly outperforms all models, except the BS model. The difference in adjusted squared Sharpe measures of the CZ and BS models is sizeable but not statistically significant. The superior performance of the CZ model stems from the large increase in maximum adjusted squared Sharpe measure in moving from the PMV to UMV frontier, and complements the evidence in Chib and Zeng (2020).

My study suggests that testing UMV efficiency of linear factor models represents a much greater challenge for asset pricing models to pass, and it also has a significant impact on relative model comparison tests. The rejection of UMV efficiency suggests that none of the factor models are an "Appropriate Benchmark" to evaluate U.K. equity fund managers for clients with a quadratic utility function. My study has assumed the zero-beta return is given by the average return of the one-month U.K. Treasury Bill. An interesting extension would be to conduct model comparison tests where the optimal zero-beta return is estimated along the lines suggested by Ferson et al. (2024). My study has focused on multifactor models. An interesting extension would be to look at alternative stochastic discount factor models based on nonlinear models like the consumption CAPM, or the use of conditional factor models following the approach in Ferson et al. (2024). Recent studies by Ehsani and Linnainmaa (2022), and Chib et al. (2023) suggest alternative ways of forming the factors. An examination of the UMV efficiency of these factor models is also of interest. I leave these issues to future research.

The table reports summary statistics of test assets and factors between July 1983 and December 2022. Panel A of the table includes 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. Panel B of the table includes the average excess returns (%) of the 16 size/BM portfolios, and 15 volatility/momentum portfolios.

The table reports the UMV efficiency tests of ten linear factor models in U.K. stock returns, and corresponding *t*-statistics. The Sh2 Diff measure is given by the difference between the adjusted maximum squared Sharpe measures of the UMV frontier of the test assets and factors and the PMV frontier of the factors. The *t*-statistic comes from Ferson et al. (2024). The CE is the Certainty Equivalent excess return and is given by $(1/2\gamma)$Sh2 Diff, and γ is set equal to 5. In panel A, the test assets are 16 size/BM portfolios, and in panel B the test assets are 15 volatility/momentum portfolios. The zero-beta return is assumed to be given by the average returns of the one-month U.K. Treasury Bill. The sample period is between July 1983 and December 2022.

The table reports the decompositions of the UMV efficiency tests of Ferson et al. (2024). The first decomposition in panel A reports Sh2 Diff1 = $\mathrm{Sh2}_{\mathrm{pmv}}(\mathrm{r,f}) - \mathrm{Sh2}_{\mathrm{pmv}}(\mathrm{f})$, Sh2 Diff2 = $\mathrm{Sh2}_{\mathrm{umv}}(\mathrm{r,f}) - \mathrm{Sh2}_{\mathrm{pmv}}(\mathrm{r,f})$, and the corresponding *t*-statistics. The second decomposition in panel B reports Sh2 Diff3 = $\mathrm{Sh2}_{\mathrm{umv}}(\mathrm{r,f}) - \mathrm{Sh2}_{\mathrm{umv}}(\mathrm{f})$, and Sh2 Diff4 = $\mathrm{Sh2}_{\mathrm{umv}}(\mathrm{f}) - \mathrm{Sh2}_{\mathrm{pmv}}(\mathrm{f})$, and the corresponding *t*-statistics. The test assets are 16 size/BM portfolios, and the zero-beta return is given by the average return of the one-month U.K. Treasury Bill. The *t*-statistics are estimated from Ferson et al. (2024). The sample period is July 1983 and December 2022.

The table reports the decompositions of the UMV efficiency tests of Ferson et al. (2024). The first decomposition in panel A reports Sh2 Diff1 = $\mathrm{Sh2}_{\mathrm{pmv}}(\mathrm{r,f}) - \mathrm{Sh2}_{\mathrm{pmv}}(\mathrm{f})$, Sh2 Diff2 = $\mathrm{Sh2}_{\mathrm{umv}}(\mathrm{r,f}) - \mathrm{Sh2}_{\mathrm{pmv}}(\mathrm{r,f})$, and the corresponding *t*-statistics. The second decomposition in panel B reports Sh2 Diff3 = $\mathrm{Sh2}_{\mathrm{umv}}(\mathrm{r,f}) - \mathrm{Sh2}_{\mathrm{umv}}(\mathrm{f})$, and Sh2 Diff4 = $\mathrm{Sh2}_{\mathrm{umv}}(\mathrm{f}) - \mathrm{Sh2}_{\mathrm{pmv}}(\mathrm{f})$, and the corresponding *t*-statistics. The test assets are 15 volatility/momentum portfolios, and the zero-beta return is given by the average return of the one-month U.K. Treasury Bill. The *t*-statistics are estimated from Ferson et al. (2024). The sample period is July 1983 and December 2022.

The table reports relative model comparison tests between the linear factor models using the PMV frontier. Panel A includes the difference between the adjusted maximum squared Sharpe measures of the model in the column and the model in the row (Sh2 Diff). Panel B includes the corresponding *t*-statistics which are estimated from Ferson et al. (2024). The zero-beta return is given by the average return of the one-month U.K. Treasury Bill. The sample period is July 1983 and December 2022.

The table reports relative model comparison tests between the linear factor models using the UMV frontier. Panel A includes the difference between the adjusted maximum squared Sharpe measures of the model in the column and the model in the row (Sh2 Diff). Panel B includes the corresponding *t*-statistics which are estimated from Ferson et al. (2024). The zero-beta return is given by the average return of the one-month U.K. Treasury Bill. The sample period is July 1983 and December 2022.

Appendix

All of the data is collected from LSPD and Refinitiv Worldscope. The accounting data comes from Worldscope, and all the remaining data come from LSPD. In forming the test assets and factors, the following corrections are made. Foreign companies, investment trusts,[18] and secondary shares are excluded. The delisting bias of Shumway (1997) is corrected following the approach of Dimson, Nagel and Quigley (2003). Where a company dies valueless according to LSPD, then we allocate a -100% return on the death event month. When calculating the portfolio returns used to form the factors, I allocate missing values, for things such as temporary suspension, to 0 as in Liu and Strong (2008). The prior month-end market values are used to calculate value weights in the portfolios used to form the factors.

Test assets

A) Size/BM portfolios

I form 16 size/BM portfolios similar to Fama and French (2012). At the start of July each year between 1983 and 2022, all stocks are ranked on the basis of their size, and BM ratio. Stocks are allocated to one of four size groups using breakpoints of 2%, 5%, and 10% of aggregate market capitalisation (Small to Big), and stocks are allocated to four BM groups (Growth to Value) using quartile breakpoints from the BM ratios of Big stocks (largest 90% of market value). Sixteen portfolios are then formed using the intersection of companies of the 4×4 sorts. For each portfolio, the value weighted monthly returns are calculated for the next year. The BM ratio is calculated using the book value of equity (WC03501) during the previous calendar year and the year-end market value. Size is measured by the market value at the end of June. Companies with negative book values are excluded. The average number of stocks (rounded) across the portfolios ranges between 33 and 519.

b) Volatility/Momentum portfolios

The motivation for this set of portfolios stems from Kirby and Ostdiek (2012a, b). At the start of each month between July 1983 and December 2022, all stocks are grouped into quintile portfolios based on their average absolute returns during the past t-12 to t-2 months

[18] Investment trusts are equivalent to closed-end funds.

as in Kirby and Ostdiek (2012b). Within each quintile volatility portfolio, companies are further sorted into three portfolios on the basis of their cumulative buy and hold return during the past t-12 to t-2 months. All portfolios have an equal number of companies as an approximation. The value weighted portfolio return is then calculated during the next month. Companies with missing returns during the past 12 months are excluded. The average number of stocks within each portfolio is 114.

Factors

A) FF6 Factors

I form the market index following a similar approach to Dimson and Marsh (2001). At the start of July each year between 1983 and 2022, all stocks on LSPD that are alive are allocated to the market index. I then calculate the value weighted monthly returns during the next 12 months. I calculate the excess returns on the market index using the monthly returns of the one-month U.K. Treasury Bill.

I form the HML factor following a similar approach to Fama and French (2012). At the start of July each year between 1983 and 2022, all stocks are ranked on the basis of their market value and BM ratio. I allocate stocks into two size groups (Small and Big), where Big stocks are the top 90% by market value and small stocks are bottom 10% by market value. I also allocate stocks to three BM groups (Growth, Neutral, Value) using the BM breakpoints of 30% and 70% percentiles of the BM ratios of the Big stocks. I create 6 portfolios at the intersection of the independent 2×3 sorts (SG, SN, SV, BG, BN, and BV), and then calculate value-weighted monthly returns during the next 12 months.

The HML factor is calculated as the average of the HML_S and HML_B portfolios, where $HML_S = SV\text{-}SG$ and $HML_B = BV\text{-}BG$. I exclude companies with negative book values. From the 6 portfolios, I also calculate a SMB_{HML} factor as the average return of 3 Small stock portfolios minus the average return of 3 Big stock portfolios.

The RMW and CMA factors are formed as follows. At the start of July each year between 1983 and 2022, all stocks are ranked on the basis of their market value and either their gross profitability (GP) or investment growth (Inv). All stocks are then grouped into two size groups (Small and Big), three GP groups (Weak, Neutral, Robust), and three Inv groups (Conservative, Neutral, Aggressive). The GP and Inv groups are sorted by 30% and 70% percentiles of the GP and Inv measures of Big companies. Six size/GP (size/Inv) portfolios are formed at the intersection of the 2*3 sorts as SW, SN, SR, BW, BN, and BR (SC, SN, SA, BC, BN, and BA), and value weighted monthly returns on each portfolio is calculated over the next 12 months.

The RMW factor is then calculated as the average returns of RMW_S and RMW_B portfolios, where $RMW_S = SR - SW$, and $RMW_B = BR - BW$. The CMA factor is calculated as the average return of the CMA_S and CMA_B portfolios, where $CMA_S = SC - SA$, and $CMA_B = BC - BA$. I exclude companies with zero total assets. The GP of a company is calculated as the difference between sales revenue (WC01001) and the cost of goods sold (WC01051) divided by the total assets (WC02999) from the prior fiscal year t-1. Inv is defined as the annual change in total assets in years t-1 and t-2 divided by lagged total assets at t-2. From the 6 portfolios in each sort, I calculate SMB_{RMW}, and SMB_{CMA} factors, and the final SMB factor is the average return of the SMB_{HML}, SMB_{RMW}, and SMB_{CMA} factors.

The MOM factor is formed as follows. At the start of each month between July 1983 and December 2022, all stocks are ranked on the basis of their market value and cumulative buy and hold returns from months t-12 to t-2. All stocks are grouped into two size groups (Small and Big), and three momentum groups (Losers, Neutral, Winners) using breakpoints of 30% and 60% of the past returns of Big stocks. Six portfolios are formed at the intersection of the 2*3 sorts (SL, SN, SW, BL, BN, BW), and value weighted return is calculated over the next month for each portfolio. The MOM factor is calculated as the average returns of the MOM_S and MOM_B portfolios, where $MOM_S = SW - SL$, and $MOM_B = BW - BL$. Companies with incomplete returns during for the prior year are excluded.

b) LSC Factors

The factors in the LSC model of Clarke (2022) come from 16 portfolios formed on the basis of their expected excess return from a Fama and MacBeth (1973) cross-sectional regression of monthly excess returns on a set of stock characteristics. I use a similar set of characteristics to Clarke (2022), which is similar to the model 2 set of characteristics of Lewellen (2015). The stock characteristics at time t include:

1. Size – the log of prior month-end market value.
2. BM – the log of the monthly BM ratio. For each month between July of year t to June of year t + 1, the BM ratio is the book value of the company from the fiscal year-end of t-1 divided by the prior month-end market value.[19] Companies with negative book values are excluded.
3. Momentum – cumulative returns from month t-12 to t-2. Companies are only included if they have 12 past return observations.
4. Net stock issues – log of the one-year growth of split adjusted shares outstanding from year t to year t-1.
5. Accruals – is given as change in operating working capital per split-adjusted share from t-2 to t-1 divided by book value per split-adjusted share (WC05476) at time t-1. Operating working capital is current assets (WC02201) minus cash and short-term investments (WC02001) minus current liabilities (WC03101) plus debt in current liabilities (WC03051).
6. Gross profitability – the same as for the RMW factor.
7. Asset growth – the same as for the CMA factor.

The stock characteristics 1 to 4 are available monthly. The stock characteristics 5 to 7 are available annually and so the same characteristic is used between the July of year t to June of year t + 1.

For each month between July 1983 and Dec 2022, Fama and MacBeth (1973) cross-sectional regressions are estimated of individual stock excess returns on a constant and the seven characteristics. The characteristic premiums are calculated as the average slope coefficients over time. Given the characteristic premiums, expected excess returns of the stocks is calculated each month using the prior month stock characteristic values multiplied by the characteristic premiums. Fama and French (2015) point out that using monthly slope

[19] This is the same characteristic used in the more timely factor of Asness and Frazzini (2013).

coefficients instead of the average premiums would largely capture the unexpected variation in expected excess returns.[20] On the basis of their expected excess return, all stocks are allocated to 16 portfolios, with an equal number of stocks in each portfolio as an approximation and value weighted return of each portfolio is calculated for the next month.

From the monthly returns of the 16 portfolios, an eigenvalue decomposition is performed, and three factors are formed from the largest eigenvalues. The first factor is a Level factor, where the eigenvector of the largest eigenvalue is rescaled to sum to 1, and puts a similar weight across the 16 portfolios. The second factor is the Slope factor, and the third factor is the Curve factor. In addition to the exclusions already mentioned, financials are excluded and stocks are only included where company has the data on the seven characteristics from the prior month.

c) HXZ Factors

I form the ME, ROE, and IA factors as follows. At the start of July each year between 1983 and 2022, all stocks are ranked independently by size (ME), investment growth (IA), and the return on equity (ROE). Stocks are grouped into two size groups Small and Big using a breakpoint of 10% of aggregate market capitalization. Stocks are also grouped into three ROE and IA groups using breakpoints of 30% and 70% of the ROE and IA measures of Big companies. There are then 18 portfolios formed at the intersection of the 2*3*3 groups and value weighted portfolio returns are calculated for the next 12 months.

Size is the market value at the end of June in year t. IA is calculated as the change in total assets between the fiscal years t-1 and t-2 all divided by the total assets in year t-2. The ROE is calculated as net income (WC01551) divided by book value from the prior fiscal year t-1. I also exclude financials, companies with negative book values, or zero total assets.

The ME, ROE, and IA factors are estimated from the 18 portfolio returns. The ME factor is the difference between the average returns of 9 Small portfolios and the average returns of 9 Big portfolios. The IA factor is the difference between the average returns of the six low IA portfolios and the six high IA portfolios. The ROE factor is the difference between the average returns of the six high ROE portfolios and the six low ROE portfolios. My approach differs from Hou et al. (2015) who use a monthly portfolio revision, with a quarterly ROE. However a recent study by Hanauer (2020) also use an annual revision for forming the ROE factor in international stock markets.

d) BAB Factor

I form the BAB factor in Frazzini and Pedersen (2014) as follows. At the start of each month between July 1983 and December 2022, all stocks are ranked by the beta at the end of the previous month. The betas are provided by LSPD. All stocks are assigned ranks (zi) and z is the average rank. The weights in the low beta portfolios are given by $k(zi - z)^{-}$, and the weights in the high beta portfolio are given by $k(zi - z)^{+}$, where k is a normalizing constant.[21] The $^{-}$ and $^{+}$ take the min(weights,0) and max(weights,0) respectively. Given the

[20] There is a look-ahead bias in this approach but given the poor performance of the LSC model, it unlikely has a major impact.

[21] Frazzini and Pedersen (2014) set $k = 2/\text{sum}(\text{abs}(zi - z))$.

weights, I then calculate the portfolio return for the next month and portfolio beta for each group. The BAB factor is given by long the low beta portfolio and short the high beta portfolio given by $(1/\beta_{low})(r_{low} - Rf) - (1/\beta_{high})(r_{high} - Rf)$, where β_{low} and β_{high} are the portfolio betas using the weights above, r_{low} and r_{high} are the corresponding portfolio returns, and Rf is the one-month Treasury Bill return.

e) SY Factors

The Size, MGMT, and PERF mispricing factors are formed from market anomalies. Stambaugh and Yuan (2017) use 11 market anomalies. I follow a similar approach but exclude the O-score and distress measures. Lu, Stambaugh and Yuan (2017) also exclude these variables in their study of anomalies in international stock returns. To estimate the MGMT factor, at the start of July each year between 1983 and 2022, all stocks are ranked independently on the basis of six characteristics. The characteristics include net stock issues, composite equity issues, accruals, net operating assets, asset growth, and investment to assets. For each characteristic, all stocks are ranked between 0 and 1, and the average (P1) ranking is estimated for each stock. To be included, I require all stocks to have the relevant characteristic data.

All stocks are then ranked independently by size and their average P1 ranking. Two size groups are formed as before and three P1 groups are formed (Low, Medium, and High) based on the 30% and 70% percentiles average P1 ranking across all stocks. Six portfolios are formed at the intersection of the size and P1 groups (SL, SM, SH, BL, BM, and BH) and value weighted buy and hold portfolio returns are calculated over the next 12 months using prior month-end market values. The MGMT factor is given by the average return of the SL and BL portfolios minus the average return of the SH and BH portfolios. A SMB_{MGMT} factor is calculated as the average return of the three Small stock portfolios minus the average return of the three Big stock portfolios.

The stock characteristics used for the MGMT factor are:

1. Net Stock Issues—the log of the adjusted number of shares at end of June in year t divided by the adjusted number of shares at end of June in year t-1.
2. Composite Stock Issues—calculated as $\log(ME_t/ME_{t-1}) - r(t-1,t)$, where ME_t is market value at the end of June in year t, ME_{t-1} is the market value at the end of June in year t-1, and r(t-1,t) is the cumulative log returns over the prior 12 months. Stocks require continuous returns over the prior 12 months.
3. Accruals—calculated as in the LSC model.
4. Net operating assets—(Operating assets minus operating liabilities)/Total assets. Operating assets equals total assets minus cash and short-term investment (WC02001), operating liabilities equals total assets minus debt in current liabilities (WC03051) minus long-term debt (WC03251) minus common equity minus minority (non-controlling) interests (WC03426) minus preferred stocks (WC03451). Companies with zero total assets are excluded.
5. Asset growth—calculated as for the CMA factor.
6. Investment to assets—the change in gross property, plant and equipment (WC02501) and change in inventory (WC02101) between years t-1 and t-2 divided by total assets.

A similar approach is followed for the PERF factor, except only three stock characteristics are used based on momentum, gross profitability, and return on assets. All stocks are given a rank between 0 and 1 for each characteristic, and then the average P2 ranking is calculated for each stock. Two size groups and three P2 groups are formed as above and size portfolios are formed at the intersection of the size and P2 groups (SL, SM, SH, BL, BM, and BH) and value weighted buy and hold portfolio returns are calculated over the next 12 months using prior month-end market values. The PERF factor is given by the average return of the SH and BH portfolios minus the average return of the SL and BL portfolios. To be included in the six portfolios, I require all stocks to have the relevant characteristic data. A SMB_{PERF} factor is calculated as the average return of the three Small stock portfolios minus the average return of the three Big stock portfolios. The Size factor is given as the average return of the SMB_{MGMT} and SMB_{PERF} factors.

The stock characteristics used in the PERF factor are:

1. Momentum—calculated as for the MOM factor.
2. Gross profitability—calculated as for the RMW factor.
3. The return on assets (ROA)—net income before extraordinary items in year t-1 divided by total assets in year t-1.

f) HMLT Factor

Asness and Frazzini (2013) propose the more timely value (HMLT) factor. At the start of each month between July 1983 and Dec 2022, all stocks are ranked by their size and BM ratio. Big stocks are the top 90% by market value and small stocks are bottom 10% by market value (Small and Big). The BM breakpoints are 30% and 70% percentiles of the BM ratios of the Big stocks (Growth, Neutral, Value). Six portfolios are then formed at the intersection of the independent 2×3 sorts (SG, SN, SV, BG, BN, and BV). Value-weighted portfolio returns are then calculated for the next month.

Size is the market value at the end of the previous month. The monthly BM ratio from July of year t to June of year t+1 is given by the book value from the calendar year t-1 divided by the market value at the end of the previous month. The HML_T is equal to the average returns of HML_S minus the average return of HML_B, where HML_S = SV-SG and HML_B = BV-BG. Companies with negative book values are excluded.

g) QMJ Factor

Asness et al. (2019) propose the Quality Minus Junk (QMJ) factor. They build a quality score for each company using four composite proxies including Profitability, Growth, Safety, and Payout. The stock characteristics used in each of the four composite proxies are:

1. Profitability
    a. Gross profitability – same as for the RMW factor.
    b. ROE – same as for the ROE factor.
    c. ROA – same as for the PERF factor.
    d. Cash flow over assets—net income plus depreciation(WC01148) capital expenditures(WC04601) at year t-1 minus changes in working capital between years t-1 and t-2 all divided by total assets at year t-1. Working capital is current assets

minus current liabilities minus cash and short-term investments plus short-term debt (WC03051) and income taxes payable (WC01451).

   e. Gross margin—gross income (WC01100) at year t-1 divided by net sales at year t-1.
   f. Accruals—is minus (change in working capital-depreciation between years t-1 and t-2) divided by total assets at year t-1.

2. Growth is the growth rate on the profitability ratios above between years t-1 to t-6.
3. Safety

   a. Market beta – the market beta from LSPD at the end of June of year t.
   b. Residual volatility – the residual volatility from LSPD at the end of June of year t.
   c. Leverage—is total debt divided by total assets at year t-1.[22] Total debt is long-term debt plus debt in current liabilities plus minority interest plus preference shares at year t-1.

4. Payout

   a. Net equity issuance – same as used in the LSC model.
   b. Net debt issuance—is minus the log of total debt at time t divided by total debt at time t-1. c. Total net payout—is the sum of (net income minus changes in book equity) over past 5 years divided by total profits over last 5 years.

At the start of July each year between 1983 and 2022, all stocks are ranked independently by size and their Quality score. For each characteristic, all stocks are ranked and then standardized. The average rank is calculated for each of the four groups, and then the average rank is calculated across the four groups. Companies are allowed to have missing characteristic data. Stocks are allocated to two size groups (Small and Big), and three Quality groups (Junk, Medium, and Quality) using breakpoints of 30% and 70% of the Quality scores of Big companies. Six portfolios are then formed at the intersection of the size and Quality groups (SJ, SM, SQ, BJ, BM, and BQ) and value weighted portfolio returns are calculated over the next 12 months. The QMJ factor is calculated as the average return of the SQ and BQ portfolios minus the average return of the SJ and BJ factors.

h) PEAD Factor

To form the PEAD factor, all stocks at the start of each month between July 1983 and Dec 2022 are ranked independently by size and standardized unexpected earnings (SUE). Big stocks are top 90% by market value and Small stocks are the bottom 10% by market value. All stocks are allocated to three PEAD groups (Low, Medium, and High) using 30% and 70% breakpoints of the SUE measures of Big companies. Six portfolios are formed at the intersection of the size and PEAD groups (SL, SM, SH, BL, BM, and BH) and value weighted buy and hold portfolio returns are calculated over the next month.

The PEAD factor is given by (SH + BH)/2-(SL + BL)/2. The SUE is calculated in a similar way to Foster, Olsen and Shevlin (1984). For each month t, SUE is equal to the difference in annual adjusted EPS at month t-1 and month t-13 divided by standard deviation of

[22] I do not use O-score, Z-score, and earnings volatility.

the monthly annual difference in adjusted EPS over the past 24 months. The EPS comes from Datastream and financials are excluded.

References

Asness, C.S. and A. Frazzini, 2013, The devil in HML's details, *Journal of Portfolio Management*, 39, 49–68.

Asness, C.S., Frazzini, A. and L.H. Pedersen, 2019, Quality minus junk, *Review of Accounting Studies*, 24, 34–112.

Clarke, C., 2022, The level, slope and curve factor model for stocks, *Journal of Financial Economics*, 143, 159–187.

Dimson, E. and P.R. Marsh, 2001, U.K. financial market returns 1955–2000, *Journal of Business* 74, 1–31.

Dimson, P., Nagel, S. and G. Quigley, 2003, Capturing the value premium in the U.K. 1955–2001, *Financial Analysts Journal*, 5559, 35–45.

Fama, E.F. and K.R. French, 2012, Size, value, and momentum in international stock returns, *Journal of Financial Economics*, 105, 457–472.

Fama, E.F. and K.R. French, 2015, Incremental variables and the investment opportunity set, *Journal of Financial Economics*, 117, 470–488.

Fama, E.F. and J.D. MacBeth, 1973, Risk,return and equilibrium: Empirical tests, *Journal of Finance*, 81, 607–636.

Foster, G., Olsen, C. and T. Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, *Accounting Review*, 59, 574–603.

Frazzini, A. and L.H. Pedersen, 2014, Betting against beta, *Journal of Financial Economics*, 111, 1–25.

Hanauer, M.X., 2020, A comparison of global factor models, *Working Paper*, Techinical University of Munich.

Hou, K., Xue, C. and L. Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies*, 28, 650–705.

Kirby, C. and B. Ostdiek, 2012a, It's all in the timing: Simple active portfolio strategies that outperform naïve diversification, *Journal of Financial and Quantitative Analysis*, 47, 437–467.

Kirby, C. and B. Ostdiek, 2012b, Optimizing the performance of sample mean–variance portfolios, *Working Paper*, University of North Carolina at Charlotte.

Lewellen, J., 2015, The cross section of expected stock returns, *Critical Finance Review*, 4, 1–44.

Liu, W. and N. Strong, 2008, Biases in decomposing holding period portfolio returns, *Review of Financial Studies*, 21, 2243–2274.

Lu, X., Stambaugh, R.F. and Y. Yuan, 2020, Anomalies abroad: Beyond data mining, *Working Paper*, University of Pennsylvania.

Shumway, Tyler, 1997, The delisting bias in CRSP data, *Journal of Finance* 52, 327–340.

Stambaugh, R.F. and Y. Yuan, 2017, Mispricing factors, *Review of Financial Studies*, 30, 1270–1315.

References

Abhyankar A, Basu D, Stremme A (2012) The optimal use of return predictability: An empirical study. J Financial Quantit Anal 47:973–1001
Asness CS, Frazzini A (2013) The devil in HML's details. J Portf Manag 39:49–68
Asness CS, Frazzini A, Pedersen LH (2019) Quality minus junk. Rev Acc Stud 24:34–112
Barillas F, Shanken J (2022) Comparing priors for comparing asset pricing models. Journal of Finance, Replications and Comments
Barillas F, Shanken J (2017) Which alpha? Rev Financial Stud 30:1316–1388
Barillas F, Shanken J (2018) Comparing asset pricing models. J Finance 73:715–754
Bryzgalova S, Huang J, Julliard C (2023a) Bayesian solutions for the factor zoo: We just ran two quadrillion models. Journal of Financ 78:487–557
Bryzgalova S, Pelger M, Zhu J (2023b) Forest through the trees: building cross-sections of stock returns. J Financ, forthcoming
Chamberlain G (1983) Funds, factors, and diversification in arbitrage pricing models. Econometrica 51:1305–1323
Chib S, Zeng X (2020) Which factors are risk factors in asset pricing? A model scan approach. J Bus Econ Stat 38:771–783
Chib S, Zeng X, Zhao L (2020) On comparing asset pricing models. J Finance 75:551–577
Chib S, Zhao L, Zhou G (2024) Winners from winners: A tale of risk factors. Manage Sci 70:396–414
Chib S, Lin YC, Pukthuanthong K, Zeng X (2023) Slope factors outperform: evidence from a large comparative study, Working Paper, Washington University in St Louis
Clarke C (2022) The level, slope and curve factor model for stocks. J Financ Econ 143:159–187
Cochrane JH (2005) Asset pricing, Revised. Princeton University Press, Princeton NJ
Daniel KD, Hirshleifer D, Sun L (2020) Short- and long-horizon behavioral factors. Rev Financial Stud 33:1673–1736
Davies JR, Fletcher J, Marshall A (2015) Testing index-based models in U.K. stock returns. Rev Quant Financ Acc 45:337–362
De Roon FA, Nijman TE, Werker BJM (2001) Testing for mean-variance spanning with short sales constraints and transaction costs: The case of emerging markets. J Finance 56:721–742
Dimson E, Marsh P, Staunton M (2015) Industries: their rise and fall. In: Dimson E, Marsh P, Staunton M (eds) Credit Suisse Global Investment Returns 2015 Yearbook, Credit Suisse Research Institute, pp 5–16
Dybvig PH, Ross SA (1985) Differential information and performance measurement using a security market line. Journal of Finance 40:383–399
Ehsani S, Linnainmaa JT (2022) Time-series efficient factors, Working Paper, Tuck School of Business
Fama EF, French KR (1988) Dividend yields and expected stock returns. J Financ Econ 22:3–25
Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. J Financ Econ 33:3–56
Fama EF, French KR (2012) Size, value, and momentum in international stock returns. J Financ Econ 105:457–472
Fama EF, French KR (2015) A five-factor asset pricing model. J Financ Econ 116:1–22
Fama EF, French KR (2018) Choosing factors. J Financ Econ 128:234–252
Fama EF, Schwert GW (1977) Asset returns and inflation. J Financ Econ 5:115–146
Ferson WE (1989) Changes in expected security returns, risk, and the level of interest rates. Journal of Finance 44:1191–1217
Ferson WE (2019) Empirical asset pricing: models and methods. MIT Press
Ferson WE, Siegel AF (2001) The efficient use of conditioning information in portfolios. J Finance 56:967–982
Ferson WE, Siegel AF (2003) Stochastic discount factor bounds with conditioning information. Rev Financial Stud 16:567–595
Ferson WE, Siegel AF (2009) Testing portfolio efficiency with conditioning information. Rev Financ Stud 22:2735–2758
Ferson WE, Siegel AF (2015) Optimal orthogonal portfolios with conditioning information. In: Lee CF (ed) Handbook of Financial Econometrics. Springer Science and Business Media, New York, pp 978–1001

Ferson WE, Siegel AF, Wang JL (2024) Factor model comparisons with conditioning information. J Financ Quant Anal, forthcoming

Ferson WE (2013) Investment performance: a review and synthesis. In: Constantinides GM, Harris M, Stulz RM (eds) Handbook of the Economics and Finance, vol 2. part B. Elsevier, Amsterdam, pp 969–1000

Fletcher J (1994) The mean-variance efficiency of benchmark portfolios: UK evidence. J Bank Finance 18:673–685

Fletcher J (2001) An examination of alternative factor models in UK stock returns. Rev Quant Financ Acc 16:117–130

Fletcher J (2019) Model comparison tests of linear factor models in U.K. stock returns. Financ Res Lett 28:281–291

Frazzini A, Pedersen LH (2014) Betting against beta. J Financ Econ 111:1–25

Gibbons MR, Ross SA, Shanken J (1989) A test of the efficiency of a given portfolio. Econometrica 57:1121–1152

Gregory A, Tharyan R, Christidis A (2013) Constructing and testing alternative versions of the Fama-French and Carhart models in the UK. J Bus Financ Acc 40:172–214

Grinblatt M, Titman S (1987) The relation between mean-variance efficiency and arbitrage pricing. J Bus 60:97–112

Hansen LP, Jagannathan R (1991) Implications of security market data for models of dynamic economies. J Polit Econ 99:225–262

Hansen LP, Richard SFR (1987) The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. Econometrica 55:587–613

Hansen LP, Heaton J, Luttmer EGJ (1995) Econometric evaluation of asset pricing models. Rev Financ Stud 8:237–274

Harrison M, Kreps D (1979) Martingales and arbitrage in multi-period securities markets. J Econ Theor 20:381–408

Harvey CR (2019) Replication in Financial Economics 8:1–9

Harvey CR, Liu Y, Zhu H (2016) … and the cross-section of expected returns. Rev Financ Stud 29:5–68

He H, Modest DM (1995) Market frictions and consumption-based asset pricing. J Polit Econ 103:94–117

Hou K, Xue C, Zhang L (2015) Digesting anomalies: An investment approach. Rev Financial Stud 28:650–705

Hou K, Xue C, Zhang L (2020) Replicating anomalies. Rev Financial Stud 33:2019–2133

Huberman G, Kandel S (1987) Mean-variance spanning. J Finance 42:873–888

Jensen TI, Kelly B, Pedersen LH (2023) Is there a replication crisis in Finance? J Finance 78:2465–2518

Jobson JD, Korkie B (1980) Estimation of Markowitz Efficient Portfolios. J Am Stat Assoc 75:544–554

Kan R, Zhou G (2007) Optimal portfolio choice with parameter uncertainty. J Financial Quant Anal 42:621–656

Kirby C, Ostdiek B (2012a) It's all in the timing: Simple active portfolio strategies that outperform naïve diversification. J Financial Quant Anal 47:437–467

Kirby C, Ostdiek B (2012b) Optimizing the performance of sample mean-variance portfolios, Working Paper, University of North Carolina at Charlotte

Korsaye SA, Quaini A, Trojani F (2021) Smart stochastic discount factors, Working Paper, Swiss Finance Institute

Litterman RB, Scheinkman J (1991) Common factors affecting bond returns. J Fixed Income 1:54–61

Luttmer EGJ (1996) Asset pricing in economies with frictions. Econometrica 64:1439–1467

Markowitz H (1952) Portfolio selection. J Finance 7:77–91

Penaranda F (2016) Understanding portfolio efficiency with conditioning information. J Financial Quant Anal 51:985–1011

Pukthuanthong K, Qiao Z, Wang Y (2023) No one-size-fits-all tale: the diversity and complexity in asset pricing across global markets, Working Paper, University of Missouri

Rapach DE, Zhou G (2013) Forecasting stock returns. In: Elliott G, Timmermann A (eds) Handbook of Economic Forecasting, vol 2A. Elsevier, Amsterdam, pp 323–383

Rapach DE, Zhou G (2022) Asset pricing: Time-series predictability. In: Hamilton JH (ed) Oxford Research Encyclopaedia of Economics and Finance. Oxford University Press, Oxford, pp 1–34

Roll R (1977) A critique of the asset pricing theory's test Part i: on past and potential testability of the theory. J Financ Econ 4:129–176

Ross SA (1978) A simple approach to the valuation of risky streams. J Bus 51:153–475

Sharpe WF (1966) Mutual fund performance. J Bus 39:119–138

Siegel AF, Woodgate A (2007) Performance of portfolios optimized with estimation error. Manage Sci 53:1005–1015

Stambaugh RF, Yuan Y (2017) Mispricing factors. Rev Financial Stud 30:1270–1315