JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 43, No. 2, June 2008, pp. 355–380
COPYRIGHT 2008, MICHAEL G. FOSTER SCHOOL OF BUSINESS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

The Poor Predictive Performance of Asset Pricing Models

Timothy Simin*

Abstract

This paper examines time-series forecast errors of expected returns from conditional and unconditional asset pricing models for portfolio and individual firm equity returns. A new result that increases predictive precision concerning model specification and forecasting is introduced. Conditional versions of the models generally produce higher mean squared errors than unconditional versions for step ahead prediction. This holds for individual firm data when the instruments are firm specific. Mean square forecast error decompositions indicate that the asset pricing models produce relatively unbiased predictions, but the variance is severe enough to ruin the step ahead predictive ability beyond that of a constant benchmark.

I. Introduction

Evidence that returns are predictable has encouraged interest in conditional versions of asset pricing models that incorporate predictor variables as instruments for the information set used by investors. At least partially reconciling evidence of predictability and the efficient market hypothesis, conditional versions of asset pricing models capture a significant portion of the measured predictability in asset returns (e.g., see Ferson and Harvey (1991), (1993), and (1999)). Traditionally, asset pricing tests compare the model's fitted value with the expected returns measured by the arithmetic average or some conditional estimate of the mean.[1] While academics search for a model that matches mean returns well, practitioners require a model that provides accurate expectations of future returns. There is both direct and anecdotal evidence that unconditional versions of the capital

*Simin, tsimin@psu.edu, Department of Finance, Pennsylvania State University, 345 Business Building, University Park, PA 16802. I am grateful to Amit Goyal (the referee), Wayne Ferson, Ravi Jagannathan, Harold Mulherin, Gordon Hanka, Michael McCracken, and seminar participants at the Financial Management Association Doctoral Student Consortium, University of Washington, Pennsylvania State University, Syracuse University, Southern Methodist University, Arizona State University, Texas A&M, and University of North Carolina at Chapel Hill for helpful comments. I also gratefully acknowledge financial support from the Northern Finance Association through their Best Investments Paper award. All errors are my own.

[1] Although there are some studies focusing on model predictions over time (e.g., Ferson and Harvey (1991), (1993), Ghysels (1998), Ferson and Korajczyk (1995), and Evans (1994)), most of these studies are done within sample using the same data to estimate the models and to evaluate predictable variation.

asset pricing model (CAPM) and multifactor models such as the arbitrage pricing theory (APT) are most often used to calculate expectations of returns in practice.[2]

The purpose of this paper is to examine whether the asset pricing models in use provide accurate expectations of future returns. Motivated by the success of conditional asset pricing models within sample, I investigate whether expectations of returns generated by conditional versions of the models provide better step ahead forecasts then those from unconditional versions currently popular among practitioners. Focusing on asset pricing models used in practice and evaluating the models based on "forecast errors" rather than "pricing errors" differentiates this paper from extant predictability research that attempts to determine if returns are in fact predictable.

Evaluating model performance based on step ahead predictive ability has several advantages. I avoid the useless factor bias (Kan and Zhang (1999a), (1999b)) and the errors in variables problem (Kim (1995)) that appear in more conventional cross-sectional comparisons. Step ahead evaluation helps address data snooping biases (Lo and MacKinlay (1990a)), weak instrument issues in conditional versions of the models (Foster, Smith, and Whaley (1997)), and avoids the possibly misleading use of the sample R^2 as a measure of predictability (Kirby (1997)). Finally, focusing the evaluations on step ahead performance allows circumventing the efficiency differences between estimating risk premiums within the traditional beta framework rather than within the stochastic discount factor framework (e.g., Kan and Zhou (1999), (2001), Jagannathan and Wang (2002), Cochrane (2000), and Farnsworth, Ferson, Jackson, and Todd (2002)).

The test assets include size and book-to-market portfolios, as well as firm level data. Model-generated expectations of returns are compared to four simple benchmark predictors. Even though I implement a new result concerning model specification and forecasting that increases predictive precision under the quadratic loss function, the results do not bode well for any of the asset pricing models. While the expected returns provided by the models do not exhibit systematic over- or underpredictions, they are rarely as accurate as simply using a constant of 6% per annum. At the individual firm level, there is little evidence that any of the asset pricing models produce expectations that are significantly better than the historical average of the market return.

This research makes several contributions of interest to both academics and practitioners. I prove that estimation of parameters not used in creating forecasts is detrimental to the predictive performance of linear factor models. My test results indicate that linear factor models commonly used by practitioners produce poor expectations of one month ahead returns and moving to conditional versions of the models worsens forecast performance. Versions of the conditional models that allow for nonlinearity in expected returns and beta underperform their linearly specified counterparts. I evidence the deficiencies of asset pricing model expec-

[2]Evidence that either the CAPM or APT is used in industry can be inferred from MBA level corporate finance and practitioners' texts (e.g., Brealey and Myers (2000) and Grinold and Kahn (1995)). Surveys by Graham and Harvey (2001), Gitman and Mercurio (1982), and Bruner, Eades, Harris, and Higgins (1998) present evidence of the CAPM's dominance for estimating the cost of equity. Bruner et al. (1998) note that managers use betas primarily from commercial providers and Martin and Simin (2001) present evidence that commercial betas are invariably from unconditional models.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

tations of individual stock returns whether I adjust for firm-specific attributes in a cross-sectional or a time-series framework. Overall, I find relatively unbiased asset pricing model expectations have enough variance to ruin step ahead predictive ability beyond that of a constant benchmark.

Section II of the paper outlines the models I estimate, the metrics used to evaluate predictive ability, and a proposition for reducing squared error loss when the models are used for forecasting. The data and results of portfolio expected returns are described in Section III. Section IV includes the results for the firm level samples. Section V concludes.

II. Asset Pricing Models and Prediction Metrics

A. The Models

The set of models is based on those noted in footnotes 1 and 2. This restriction eliminates many other possible choices such as the family of nonlinear consumption models, nonlinear models incorporating higher moments (Harvey and Siddique (2000)), linear models using improved factors (Jagannathan and Wang (1996), Lettau and Ludvigson (2001)), a host of possible time-series specifications, and purely predictive regressions. The main question is whether the unconditional versions of the models used in practice provide better expectations of future realized returns than do their conditional counterparts.

I examine conditional and unconditional versions of the Sharpe (1964) and Lintner (1965) CAPM and the three-factor model of Fama and French (1993).[3] The general specification for the expected return of firm i is

$$
\begin{aligned}
E\{r_{i,t+1}|Z_t\} &= \alpha_{i,t+1} + \beta'_{i,t+1} E\{r_{p,t+1}|Z_t\}, \\
\alpha_{i,t+1} &= a_{i,0} + a'_{i,1} Z_t, \\
\beta_{i,t+1} &= b_{i,0} + b'_{i,1} Z_t.
\end{aligned} \tag{1}
$$

For example, in the conditional version of the Fama-French model b_0 is 3×1, b_1 is $L \times 3$, a_1 is $L \times 1$, and a_0 is a scalar, L being the number of instrumental variables and r_p is a 3×1 vector containing the excess market proxy return and the return on the portfolios high minus low (HML) and small minus big (SMB). The unconditional version of the Fama-French model simply sets Z_t to the null set.

I consider all possible variations of the general form in equation (1). The various specifications are:

[3] I also examined the predictive performance of the APT of Ross (1976) using the statistical factors of Connor and Korajczyk (1988), (1993) and a set of economic factors resembling those of Chen, Roll, and Ross (1986). The results of these models are uniformly worse than the CAPM and Fama and French models and have been omitted.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

$$
\begin{aligned}
(2)\quad & r_t = a_0 + b_0' r_{pt} + e_t,\\
(3)\quad & r_t = + b_0' r_{pt} + e_t,\\
(4)\quad & r_t = (a_0 + a_1' Z_{t-1}) + b_0' r_{pt} + e_t,\\
(5)\quad & r_t = a_0 + (b_0 + b_1' Z_{t-1})' r_{pt} + e_t,\\
(6)\quad & r_t = (a_0 + a_1' Z_{t-1}) + (b_0 + b_1' Z_{t-1})' r_{pt} + e_t,\\
(7)\quad & r_t = (a_0 + a_1' Z_{t-1}) + (b_0 + b_1' Z_{t-1})' (c_0 + c_1 Z_{t-1}) + e_t,\\
(8)\quad & r_t = + (b_0 + b_1' Z_{t-1})' r_{pt} + e_t,\\
(9)\quad & r_t = + (b_0 + b_1' Z_{t-1})' (c_0 + c_1' Z_{t-1}) + e_t.
\end{aligned}
$$

While the true specification of the CAPM does not contain an intercept, specifications (2) and (4)–(7) allow me to answer questions about whether time variation in the intercept or in beta helps or hurts predictive performance.

Following standard practice, the one month step ahead forecasts are created using 61-month rolling windows, where the model parameters are estimated using a training period of 60 months and the vector of forecast errors is formed in the obvious way.[4] Expectations of returns from specifications (2)–(6) and (8) can be formed using either the conditional or unconditional expectations of the factor risk premium. For the unconditional expectation of the factor risk premium, I use the time-series average of r_p from the training period when forecasting r_{61}.[5] For the conditional expected factor risk premium, I replace $r_{p,61}$ with $E_t(r_{p,t+1})$ formed by regressing $r_{p,t}$ on Z_{t-1}, $t = 1, \ldots, 60$, then multiply the coefficients from that regression by Z_t.[6] For specifications (7) and (9), I avoid singularity issues when estimating these two specifications by using a different set of instruments to generate the conditional factor risk premia.

I compare the asset pricing models to four benchmark predictors of returns. These include the historical average return of the asset, the historical average return of the market index, a constant, and the conditional mean of the market index. For the step ahead comparisons, the historical means are the training period sample averages while the conditional mean is formed using the slopes from regressing the market index on the lagged instruments in the training period and the prediction period values of the lagged instruments. The constant of 0.005 per month (6% per year) represents the average monthly return of the value-weighted index for the five-year period 1926–1930, a period prior to any of the forecasts.[7]

[4]Fama and French (1997) show that the length of the estimation window has little impact on the forward-looking cost of capital estimates.

[5]Note that using the historical mean of r_p in specification (2) when predicting is identical to using the asset's time-series average as a predictor, which is one of the benchmark predictors discussed below.

[6]Note that using conditional expectations of the factor premia in specification (4) yields identical predictions for both the CAPM and the Fama-French model. A proof is available from the author.

[7]The historical average benchmarks are motivated by responses in Graham and Harvey (2001) indicating the use of historical average returns in the cost of capital calculations. Ferson and Locke (1998) find that a large part of the estimation error in cost of equity calculations is attributable to market risk premiums based on historical averages and suggest modeling market returns as a linear function of the instruments. The conditional market benchmarks follow this suggestion. The constant is the naïve low cost forecast.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

B. Measuring Forecast Accuracy

To assess the predictive ability of the asset pricing models, I use two metrics: the square root of the mean of the squared forecast error (RMSFE), and the test of conditional predictive ability (CTP) promoted by Giacomini and White (2003). I focus on the RMSFE rather than alpha because step ahead alpha is identical to the difference between using the historical average return as a predictor and the model prediction. The RMSFE measures the average squared distance between the expected return produced by the model and the realized return through time. It is a direct measure of the quality of the model's expectation and can yield a rejection that would not occur based on evaluating alpha.[8]

An advantage of the mean squared error is that it admits decomposition into the variance and squared bias of the forecast errors. The tradeoff between variance and bias is important since two different models producing equivalent RMSFEs may have dramatically different amounts of bias and variance in their forecast error distributions. Let $\hat{r}_t$ be the forecast of the return r_t at time t, so that the forecast error at time t is $e_t = r_t - \hat{r}_t$. The mean of the squared forecast error (MSFE) is defined as

$$\text{MSFE}(\hat{r}_t) \quad = \quad E[(r_t - \hat{r}_t)^2] \quad = \quad \text{var}(e_t) + [\text{bias}(e_t)]^2. \tag{10}$$

A caveat is that the MSFE is a quadratic function, applying equal weight to positive and negative errors of the same magnitude. Users with asymmetric preferences for under- versus overpredictions will find the direction of bias, rather than the aggregate MSFE, is important.[9]

A wide variety of tests are available to compare step ahead forecasts from competing models. Unfortunately, most require stringent assumptions on the forecast error series, prohibit structural changes in the data or parameters, disallow variation in model estimation technique, and depend on whether the models are nested. The CTP proposed by Giacomini and White (2003) is unique in that it avoids the above requirements while at the same time accounting for the estimation method.[10] This conditional approach allows evaluation of the forecasting method, including the model, estimation procedure, and size of the estimation window rather then just testing the forecasting model.

The null hypothesis of the CTP is of equal predictive ability:

$$\text{H}_0 : E\left[\hat{u}^2_{1,t+1} - \hat{u}^2_{2,t+1} | F_t\right] \quad = \quad 0, \tag{11}$$

where $\hat{u}_i$ is the time t forecast error for model i. Define $\Delta U_{m,t+1}$ as the difference between the squared error loss functions where m observations are used in the estimation window for both models. Under the null, the step ahead sequence

[8]For example, in a given sample of data it is possible that the CAPM expectation is typically close to the arithmetic average (I cannot reject that alphas are jointly zero), while the expectations of returns produced by both the historical mean and the CAPM are horrible forecasts of the eventually realized returns.

[9]Leitch and Tanner (1991) footnote an account of a prominent forecaster who negatively biases forecasts believing his clients want to protect themselves from "worst case" scenarios indicating users may have preferences for bias versus variance in their expectations of returns.

[10]I thank Michael McCracken for helpful discussions on the proper test to employ.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

$\{\Delta U_{m,t+1}, F_t\}$ is a martingale difference sequence, meaning the null can be written as $E[h_t \Delta U_{m,t+1}] = 0$ for all F_t measurable functions h_t.

Giacomini and White (2003) show that for one step ahead forecasts, the test statistic can be computed as nR^2 where n is the number of forecasts and R^2 is the uncentered multiple correlation coefficient from the regression of $\Delta U_{m,t+1}$ on h_t'. The α level test rejects the null of equal conditional predictive ability when nR^2 is greater than the $(1-\alpha)$ quantile of the χ_2^2 distribution. As in the empirical example in Giacomini and White (2003), I use $h_t = (1, \Delta U_{m,t})'$. Rejection of the null does not specify which model produces better forecasts. The suggested decision rule is based on the fitted values from the test regression of $\Delta U_{m,t+1}$ on h_t'. In particular, the proportion of times that the fitted values are positive indicates which forecast method the decision rule would choose.

C. Asset Pricing Models versus Regression Equations

A researcher wishing to predict excess returns based on a theoretic model such as the unconditional CAPM will form predictions as $\hat{r}_{t+1} = bE(rm_{t+1})$.[11] Standard protocol dictates estimating b from the regression $r_t = a + brm_{1,t} + v_{1,t}$ and imposing the model restriction that the intercept is zero by not including it in the forecast. Wheatley (1998) derives distributions of forecast errors for forecasts including the intercept and for forecasts that estimate the intercept but restrict it to zero when forecasting, a method that Fama and French (1997) show marginally improves step ahead estimates of costs of capital. Wheatley's first result is interesting given that practitioners sometimes use historical mean returns in cost of capital estimation, which is identical to including a constant in the forecast.

Alternatively, b could be estimated in accordance with the APT using the restricted regression $r_t = brm_{1,t} + u_{1,t}$. While even introductory econometrics texts warn against suppressing the intercept in regression equations, there are efficiency gains from imposing the null implied by the CAPM. Omission of a relevant variable, such as the intercept, can produce a more precise albeit biased slope estimate. Proposition 1 addresses the issue of whether to include an intercept when estimating the forecasting parameters.

Proposition 1. Let r_p be the pth period return to be predicted. Define two sets of prediction errors $\hat{e} = r_p - \hat{r}_p$ and $\tilde{e} = r_p - \tilde{r}_p$, where $\hat{r}_p = X_{2p}'\hat{b}$ with $\hat{b}$ estimated from the regression $r_t = X_{1,t}\hat{a} + X_{2,t}\hat{b} + v_t$, while $\tilde{r}_p = X_{2p}'\tilde{b}$, where $\tilde{b}$ results from the restricted regression $r_t = X_{2,t}\tilde{b} + u_t$. Here r is $T \times 1$, where $p > T$, X is $T \times K$, $\hat{a}$ is a K_1-vector, $\hat{b}$ and $\tilde{b}$ are K_2-vectors such that $K_1 + K_2 = K$ and ε is $T \times 1$, $\varepsilon \sim (0, \sigma^2)\ \forall t$, $E(\varepsilon_t \varepsilon_s) = 0$ for $t \neq s$. Then $E\{\text{MSE}(\hat{e}) - \text{MSE}(\tilde{e})\} \geq 0$.

A proof of Proposition 1 and an evaluation of the magnitude of the difference by simulation is in the Appendix. The denouement of the derivation and the simulations is that, if the intercept is not going to be used to forecast, suppressing the intercept during estimation of the parameters provides a lower MSFE. The two

[11] I am not suggesting that using an asset pricing model is the minimum MSFE method for forecasting returns, only that the manager forms expectations based on the CAPM following the procedure outlined in most standard corporate finance texts such as Brealey and Myers (2000).

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

methods produce equivalent MSFEs when the true intercept is zero. Proposition 1 differs from Wheatley (1998) by suggesting that the intercept should be restricted to zero prior to estimation. Proposition 1 holds for the conditional setting, where $X_{1,t}$ contains a vector of ones, as well as the observations of the lagged instrumental variables and $X_{2,t}$ contains the time t observations on the relevant risk factors and the product of the instrumental variables and the risk factors and in the unconditional setting, where $X_{1,t}$ contains a vector of ones and $X_{2,t}$ contains the time t observations on the relevant risk factors.[12]

III. Model Expectations of Portfolio Returns

A. Test Assets and Instruments

I evaluate the predictive capabilities of the models and benchmarks using 25 size and book-to-market portfolios spanning the period 01/1931 to 12/2004.[13] The instrumental variable data consists of five series representing those typically used in asset pricing studies. These include a dummy for the month of January (Jan.), U.S. one-month T-bill from the CRSP Risk-Free Rates file (T-bill), the dividend yield of the Standard & Poor's 500 index (DivYld), the spread between the lagged Moody's Composite Average of Yields on Corporate Bonds, and the U.S. one-month T-bill from the CRSP Risk-Free Rates file (Term), and the difference between the Moody's BAA and AAA corporate bond yield (Junk). The details on the construction of all the instrumental variables can be found in the Appendix.

To estimate specifications (7) and (9) using OLS, I select a different set of instruments to determine the conditional market risk premium to avoid perfect multicollinearity. In this case, I use the 90-day bid for the U.S. T-bill from the CRSP Risk-Free Rates file, the dividend yield of the NYSE, AMEX, and NASDAQ composite, the spread between the five-year and one-month discount Treasury yield from the Fama-Bliss files as a proxy for the term instrument, and the bond default premium from the Ibbotson files as a measure of the instrument Junk.[14]

The market proxy reported throughout is the value-weighted NYSE, NASDAQ, and AMEX market index from CRSP. All returns are in excess of the 30-day Treasury bill return found in the CRSP CTI file (T30RET). For the multifactor models, I use the SMB and HML spread portfolio returns from Ken French's web page (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/).

[12]The issue is not confined to unconditional versions of the CAPM. Harvey (1989) estimates the constant reward to risk conditional CAPM with and without an intercept. It is clear from the distribution of the average pricing errors that the specification of the model, including an intercept, produces larger errors.

[13]In unreported results, I use a set of market capitalization decile portfolios; 12 industry-based portfolios, four portfolios following dynamic trading strategies including the Jegadeesh and Titman (1993), (2001) 6/6 spread momentum portfolio, a contrarian portfolio via Lo and MacKinlay (1990b), and two portfolios similar to strategies available from Dimensional Fund Advisors. The results are qualitatively similar across the different sets of portfolios. Details on formation and results for the unreported portfolios are available from the author.

[14]I experimented with several other instrumental variables when determining the conditional market risk premium for specifications (7) and (9) based on the set of instruments outlined in Ferson and Qian (2004). These included measures of short-term corporate illiquidity, inflation, industrial production, and convexity of term structure. The choice of the instrument set had little impact on the qualitative performance of the models.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

B. Model Comparisons

Table 1 contains the one step ahead RMSFEs and bias of the forecast errors scaled by 100 for the OLS specifications of the CAPM and Fama and French models estimated for the 25 size and book-to-market portfolios. The rolling window estimations produce 875 forecast errors for each asset. The large number of model specifications, benchmarks, and test assets produce 750 RMSFEs. I present the results for the small/high, small/low, large/high, and large/low portfolios as a representative sample. Ranking the portfolios by model RMSFE reveals that the small/low portfolio is consistently the most difficult to predict while the large/low portfolio is consistently the second easiest to predict. For all the models and benchmarks, there is a strong negative relation between RMSFE and size. There is no discernable relation between the magnitude of the RMSFE and book-to-market for any of the models or benchmarks with the exception that the highest book-to-market firms tend to have higher then average RMSFEs.[15]

Table 2 contains the relative performance of each model and benchmark based on the CPT described in equation (11). In all panels, each cell contains the number of times, across the 25 size and book-to-market portfolios, the test for CPT rejected equal predictive ability at the 10% level, and the model in the column was chosen over the model in the row based on the decision rule described above. Panel A shows the relative performance of the different versions of the CAPM and the benchmarks. Panel B has the same results for the Fama and French 3-factor model. Fa and Fb stand for fixed alpha and beta, Tb, and Ta stand for time-varying beta and alpha, and (C) indicates that a conditional expected risk premium is used to forecast while (U) indicates that the unconditional risk premium is used. For instance, TaTb(C) represents specification (6) using the conditional expectation of the factor risk premium.

Tables 1 and 2 include the following results: i) the constant benchmark provides the lowest RMSFE for all of the 25 size and book-to-market sorted portfolios and most often outperforms the other benchmarks and models; ii) all versions of the models that employ conditional expected factor risk premia when predicting are dominated by the versions of the models that use unconditional expected risk premia; iii) focusing on the versions using unconditional expected risk premia when predicting, allowing for time variation in beta tends to marginally improve performance, and allowing for time variation in the intercept typically worsens performance; iv) based on the total number of times the model was chosen over the other models and benchmarks, the best performing asset pricing model is the unconditional factor risk premium CAPM with fixed alpha and time-varying beta, closely followed by the same version of the Fama-French model; and v) the bias in the forecasts is small and typically positive. Across all model predictions, the bias is never greater than 15% of the RMSFE and is negative for only 12 of the 875 (1.4%). On average, bias accounts for 6% of the RMSFE.

The results based on the asset pricing models are emphatically feeble. None of the asset pricing models outperforms using a constant of 6%. Even though the versions of the asset pricing models using unconditional expected factor risk

[15]Firms with very high book-to-market ratios may be distressed and have returns that are difficult to predict.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

TABLE 1

Step Ahead Root Mean Squared Forecast Errors: One-Month Returns

Table 1 contains the square root of the MSFE∗100 and bias∗100 of 875 one-month forecasts of the extreme size/BM portfolios with the parameters of each model estimated using rolling 60-month windows. Forecast errors for Mean are the one month ahead asset returns minus the mean asset return from the training period. Forecast errors for Six are the one month ahead asset returns minus 0.06/12. Forecast errors for the unconditional market mean (Umkt) are the one month ahead asset returns minus the mean of the value-weighted index from the training period. Forecast errors for the conditional market mean (Cmkt) are the one month ahead asset returns minus the product of the betas from regressing the value-weighted index on the lagged instruments and the test period value of the instruments. Construction of the factors and instruments is described in the text. The market proxy is the CRSP value-weighted index. The lagged predetermined instruments include a January dummy, the one-month T-bill, the NYSE dividend yield, a measure of quality in corporate bonds, and a measure of the term structure. The sample covers the period 01/1931 to 12/2004. Fa and Fb stand for fixed alpha and beta, Tb, and Ta stand for time-varying beta and alpha. (C) indicates that a conditional expected risk premium is used to forecast while (U) indicates that the unconditional risk premium is used.

	RMSFE∗100				Bias∗100			
Model	Small/Low	Small/High	Big/Low	Big/High	Small/Low	Small/High	Big/Low	Big/High
Mean	12.52	9.74	5.38	9.87	0.22	0.20	0.06	1.17
Six	12.34	9.69	5.32	8.31	0.54	1.43	0.47	0.92
Umkt	12.41	9.75	5.38	8.39	0.40	1.29	0.33	0.78
Cmkt	12.60	10.10	5.80	8.78	0.55	1.43	0.48	0.93
	CAPM				CAPM			
Fb(C)	12.97	10.34	5.78	9.07	0.34	1.37	0.48	1.05
FaFb(C)	13.05	10.28	5.77	10.55	0.40	0.45	0.21	1.54
TaFb(C)	13.58	10.38	5.81	11.25	0.91	0.35	0.22	1.66
FaTb(C)	12.85	10.00	5.86	9.95	0.35	0.47	0.13	1.40
TaTb(C)	13.35	10.09	5.86	10.83	0.75	0.34	0.17	1.45
Tb(C)	12.80	10.12	5.85	8.69	0.27	1.15	0.36	0.90
Fb(U)	12.46	9.78	5.38	8.44	0.17	1.11	0.32	0.67
TaFb(U)	12.95	9.79	5.42	10.54	0.76	0.14	0.06	1.33
FaTb(U)	12.50	9.70	5.38	9.94	0.25	0.43	0.06	1.24
TaTb(U)	12.85	9.83	5.41	10.90	0.62	0.33	0.10	1.31
Tb(U)	12.41	9.74	5.38	8.39	0.24	1.24	0.30	0.69
TaTbTrp	14.05	11.76	6.40	13.83	0.56	0.56	0.17	1.65
TbTrp	13.13	10.70	5.94	12.57	0.84	1.48	0.49	1.39
	FF 3-Factor				FF 3-Factor			
Fb(C)	13.16	10.31	5.83	9.04	0.25	0.70	0.63	0.72
FaFb(C)	13.22	10.29	5.80	10.78	0.49	0.31	0.23	1.51
TaFb(C)	13.58	10.38	5.81	11.25	0.91	0.35	0.22	1.66
FaTb(C)	13.53	10.22	5.82	10.71	−0.06	0.28	0.17	1.11
TaTb(C)	13.81	10.34	5.81	11.62	0.12	0.33	0.19	1.46
Tb(C)	13.52	10.26	5.85	9.02	−0.36	0.47	0.39	0.43
Fb(U)	12.46	9.76	5.40	8.42	−0.02	0.59	0.45	0.38
TaFb(U)	12.83	9.80	5.40	10.51	0.62	0.25	0.05	1.35
FaTb(U)	12.43	9.74	5.38	10.73	0.50	0.34	0.09	1.50
TaTb(U)	12.82	9.91	5.40	12.30	0.64	0.39	0.12	1.56
Tb(U)	12.36	9.75	5.39	8.40	0.02	0.66	0.46	0.37
TaTbTrp	20.02	15.77	8.97	17.63	0.09	0.14	−0.04	0.06
TbTrp	17.20	13.49	7.99	15.24	0.45	0.46	0.24	0.64

premia outperform their counterparts using conditional expected factor premia, none of the models clearly outperforms the unconditional average of the asset. Together with the fact that few of the time-varying parameter models outperform Fb(U), this indicates that conditional versions of the models underperform unconditional versions since the unconditional average of the asset is equivalent to using a fixed alpha, fixed beta version of the model with the unconditional factor premium. Comparing the different versions of the models to the unconditional average of the asset also indicates that allowing for time variation in the intercept more often increases the RMSFE while the impact of allowing for time variation in beta is much more mixed.

The only good news for the model-generated expectations is that they are relatively unbiased. The different specifications of the Fama and French model

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

TABLE 2

Relative Performance Based on Squared Error Loss: Size/BM Portfolios

In all panels, each cell contains the number of times, across the 25 size and book-to-market portfolios, the test for conditional predictive ability rejected equal predictive ability at the 10% level and the model in the column was chosen over the model in the row based on the decision rule described in the text. Panel A contains the comparisons of the different specifications of the CAPM and the benchmarks. Panel B contains the comparisons of the different specifications of the Fama and French model. Fa and Fb stand for fixed alpha and beta, Tb, and Ta stand for time-varying beta and alpha. (C) indicates that a conditional expected risk premium is used to forecast while (U) indicates that the unconditional risk premium is used.

	Mean	Six	Umkt	Cmkt	Fb(C)	FaFb(C)	TaFb(C)	FaTb(C)	TaTb(C)	Tb(C)	Fb(U)	TaFb(U)	FaTb(U)	TaTb(U)	Tb(U)
Panel A. The CAPM and the Benchmarks															
Mean	—	25	6	1	1	0	0	0	0	1	4	8	9	3	6
Six	0	—	0	0	0	0	0	0	0	0	0	1	0	0	0
Umkt	17	25	—	0	0	0	0	0	0	0	1	10	14	8	4
Cmkt	9	23	13	—	2	2	1	4	4	2	11	14	10	15	11
Fb(C)	18	24	19	23	—	17	9	15	8	6	17	20	19	21	18
FaFb(C)	16	24	19	22	6	—	3	10	10	9	18	19	16	18	17
TaFb(C)	23	25	22	23	9	19	—	16	11	11	22	21	23	21	21
FaTb(C)	15	23	16	14	5	8	5	—	3	2	16	18	16	17	15
TaTb(C)	20	23	21	15	9	9	4	17	—	6	21	16	22	17	21
Tb(C)	19	24	19	19	10	11	11	20	16	—	18	21	18	17	20
Fb(U)	17	25	19	0	0	0	0	0	0	0	—	11	16	9	14
TaFb(U)	13	20	8	2	1	1	0	2	0	2	6	—	15	6	9
FaTb(U)	9	25	7	1	1	0	0	0	0	1	5	8	—	4	6
TaTb(U)	13	16	12	2	1	1	0	2	0	2	9	10	15	—	13
Tb(U)	15	25	13	0	0	0	0	0	0	0	8	10	16	8	—
Panel B. The FF 3-Factor Model and the Benchmarks															
Mean	—	25	6	1	1	0	0	0	0	1	8	4	8	4	10
Six	0	—	0	0	0	0	0	0	0	0	0	0	0	0	0
Umkt	17	25	—	0	0	0	0	0	0	0	6	6	10	6	8
Cmkt	9	23	13	—	1	2	1	4	2	2	10	12	11	14	11
Fb(C)	21	25	22	23	—	15	3	7	3	3	20	22	17	20	19
FaFb(C)	18	25	19	23	6	—	0	3	4	4	17	23	19	19	18
TaFb(C)	23	25	22	23	14	16	—	12	6	7	23	20	22	19	21
FaTb(C)	24	25	23	20	9	15	5	—	3	2	24	24	23	22	23
TaTb(C)	23	25	25	19	15	14	11	17	—	11	23	23	25	23	25
Tb(C)	23	25	24	23	17	17	10	18	13	—	24	23	22	22	23
Fb(U)	13	25	11	0	0	0	0	0	0	0	—	7	11	7	14
TaFb(U)	11	24	9	2	1	0	0	0	0	1	12	—	17	11	13
FaTb(U)	10	24	9	1	1	0	0	0	0	1	6	6	—	5	9
TaTb(U)	13	22	14	2	1	1	1	1	0	1	10	11	14	—	14
Tb(U)	11	25	9	0	0	0	0	0	0	0	7	8	12	7	—

tend to be less biased then the CAPM specifications and the specifications using an unconditional expectation of the risk premiums tend to have less bias then those utilizing the conditional expected risk premium. It is interesting that the worst performing benchmark is the conditional estimate of the market. Indeed, the specifications of the models that utilize a conditional factor risk premium during estimation (specifications (7) and (9)) are completely dominated by the other models and benchmarks and are not included in Table 2 to save space.[16] Along with the underperformance of the models that use a conditional expectation of the factor risk premium, these results indicate that either the instrumental variables are doing a poor job of capturing the relevant predictive information or that

[16]There is high volatility early in the sample for all portfolio returns used here. Repeating the forecasting exercise over the period 08/1969 to 07/1999 leaves the results unchanged except for a minor reordering of the models. Restricting the sample eliminates the largest outliers in the forecast series as seen by a dramatic decrease in the RMSFEs and biases, but there is little change in the relative performance of the asset pricing models.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

the linear specification of the conditional expectation of the factor risk premia is incorrect.

C. Modeling Time Variation in Beta

In conditional asset pricing models, time variation may be present in the betas and/or the factor risk premiums. In the prediction literature, several authors attribute poor predictive performance to the static nature of the models claiming that allowing for time variation in the parameters would produce better forecasts.[17] While time variation in the betas has generally been found to be statistically significant, several papers (e.g., Pástor and Stambaugh (1998), Ferson and Locke (1998), and Fama and French (1997)) argue that uncertainty in the risk premiums is more important than errors in estimating the model parameters. Ghysels (1998) makes the stronger claim that conditional versions of several asset pricing models with time variation in beta produce larger pricing errors than conditional versions with fixed betas.[18]

I examine how the predictive capabilities of the models are affected by the restriction that time variation in beta is a linear function of the model parameters in two ways. For the conditional CAPM (CCAPM), I compare a conditional version of the model with beta fixed to a version where the conditional covariance and conditional variance that make up beta are modeled separately allowing for a nonlinear function for beta. Separately estimating the conditional moments that compose time-varying beta comes at a cost of greatly increasing the number of parameters estimated. If estimation error in the parameters is large, the performance of the model will be adversely affected. Because of the increase in the number of parameters involved in separately estimating the conditional covariances and variances of the betas, for the multi-beta models I simply compare versions of the models where the betas are fixed to versions where they are linear functions of the instruments.

The fixed beta version of the CCAPM follows Gibbons and Ferson (1985) and Ghysels (1998). In this model, the conditional expected excess return on the asset is fixed in proportion to the conditional market risk premium,

$$E\{r_{t+1}|Z_t\} = \beta E\{r_{M,t+1}|Z_t\}. \tag{12}$$

This differs from the unconditional version of the model by letting the market risk premium vary through time with Z_t.[19] The time-varying beta version of the

[17] See Ghysels (1998), Bossaerts and Hillion (1999), and Goyal and Welch (1999).

[18] Note that specification of equation (1) constrains the conditional expectation of the asset returns to a linear function of the model parameters. The specification also implies the restriction that the parameters defining the conditional expectation of the asset return in the model must reduce to a linear function. Assuming that time-varying beta is a linear function of the instruments is equivalent to assuming that time variation in the ratio of the conditional covariance to conditional variance follows a linear function of the instruments. These restrictions would not generally be satisfied.

[19] Ghysels (1998) uses this version of the CCAPM as the base model in comparing several different time-varying beta models. His tables label this model as an unconditional CAPM instead of a constant beta conditional CAPM. It is important to distinguish between time-varying versus constant parameter models and conditional versus unconditional models. Here, a conditional model is simply one that uses Z_t, whereas an unconditional model does not incorporate lagged instruments. There is no need to associate conditional asset pricing models with time-varying parameter models.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

CCAPM considered follows from Harvey (1989) and Ellis (1996). This version allows for time variation in all the conditional moments. Here, both the conditionally expected excess market returns and second moments are modeled as linear functions of the instruments. The moment conditions and forecast error for this model are

$$\text{(13)} \qquad \begin{aligned} \mathrm{E}\left(\begin{array}{c} r_{m,t+1} - Z_t\delta_m \\ (r_{m,t+1} - Z_t\delta_m)\, r_{t+1} - Z_t\delta_{\text{cov}} \\ (r_{m,t+1} - Z_t\delta_m)^2 - Z_t\delta_{\text{var}} \\ (Z_t\delta_{\text{var}})\, r_{t+1} - (Z_t\delta_{\text{cov}})\,(Z_t\delta_m) \end{array} \right) \otimes Z_t &= 0, \\ e_{t+1} &= r_{t+1} - (Z_t\delta_{\text{cov}})(Z_t\delta_m)(Z_t\delta_{\text{var}}), \end{aligned}$$

where Z_t is a $1 \times K$ matrix of instrumental variables, δ_m, δ_{cov}, δ_{var} are $K \times 1$ vectors of parameters.

The first of the two versions of the conditional multi-beta models follows from Ferson and Korajczyk (1995). Here beta is assumed constant. The moment conditions and forecast errors for this model are defined by

$$\text{(14)} \qquad \begin{aligned} \mathrm{E}\left(\begin{array}{c} F_{t+1} - Z_t\gamma \\ (F_{t+1} - Z_t\gamma)\,(F_{t+1} - Z_t\gamma)'\,\beta - (F_{t+1} - Z_t\gamma)\, r_{t+1} \end{array} \right) \otimes Z_t &= 0, \\ e &= r_{t+1} - (Z_t\gamma)\,\beta, \end{aligned}$$

where Z_t is a $1 \times K$ matrix of instrumental variables, F_t is a $1 \times L$ vector of factor-mimicking portfolios, β and γ are $L \times 1$ and $K \times L$ vectors of parameters. Ferson and Harvey (1993) modify equation (14) to allow for time variation in beta by modeling beta as a linear function of the predetermined instruments. This changes the previous set of moment conditions and forecast errors by replacing β with (κZ_t) where κ is $K \times L$. In this section, all the models have been estimated using Hansen's (1982) GMM with the Newey-West (1987) heteroskedasticity and autocorrelation consistent weighing matrix.

An important feature of the model specifications I choose is that none require the contemporaneous market or factor returns in the prediction of the next period return for a particular asset. This ensures that the comparisons across model forecast errors are "fair" in that the time $t + 1$ forecasts are formed using only the estimated model parameters and information known at time t. The test assets for these experiments are restricted to the four extreme small/low, small/high, large/low, and large/high size and book-to-market portfolios. Table 3 contains the one step ahead forecast performance for these four specifications when estimated using both expanding samples and rolling window samples.

In both the rolling window and expanding window exercises, the nonlinear beta version of the CAPM produces higher RMSFEs for all four portfolios. Larger differences exist between the two CAPMs for the small size portfolios. The results for the multi-beta models are much closer in both exercises where the time-varying beta version typically has a slightly lower RMSFE. As in Tables 1 and 2, the amount of bias in the forecast errors is nominal, although the Fama and French model tends to underpredict. A consistent feature throughout Table 3 is that the small portfolios are harder to predict. Other than marginal impacts on

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

TABLE 3

Time-Varying Beta versus Fixed Beta Conditional Models
Step Ahead Forecast Errors: One-Month Returns

Table 3 contains the RMSFE∗100 and forecast error bias∗100 for the four extreme size and book-to-market sorted portfolios. FF stands for the Fama-French Model. Construction of the factors and instruments is described in the text. The Fixed models have constant parameters and Time-Varying models allow for time variation in the betas. The market proxy is the CRSP value-weighted index. The lagged predetermined instruments include a January dummy, the one-month T-bill, the NYSE dividend yield, a measure of quality in corporate bonds, and a measure of the term structure. The sample (1/1931 to 12/2004) allows for 827 predictions. All forecast error series are Winsorized at the 5% level. For the rolling window exercise, the training period for each model is restricted to 60 months of data and the parameters are used to predict the following months' return. In the expanding sample exercise, each model is estimated starting with data from the first 60 months to forecast the 61st month. The estimation period data is then expanded by one period and the process is repeated. A * next to the RMSFE indicates that the conditional test of equal predictive ability rejected the null and the decision rule described in the text chose that model.

Panel A. Rolling Window

	RMSFE			
	CAPM		FF	
Portfolio	Fixed	Time-Varying	Fixed	Time-Varying
Small/Low	9.05*	13.73	9.01	9.01
Small/High	6.94*	11.92	6.92	6.93
Big/Low	4.77*	6.12	4.70	4.71
Big/High	5.998	8.36	5.84	5.84
	Bias			
Small/Low	0.16	0.67	−0.29	−0.29
Small/High	0.78	1.84	0.56	0.55
Big/Low	0.29	0.22	0.14	0.14
Big/High	0.68	0.89	0.02	0.02

Panel B. Expanding Window

	RMSFE			
	CAPM		FF	
Portfolio	Fixed	Time-Varying	Fixed	Time-Varying
Small/Low	9.03	13.51	9.07	9.05*
Small/High	7.19	9.17	6.90	6.90*
Big/Low	4.78	4.99	4.77	4.76
Big/High	8.11	11.47	5.93	5.91
	Bias			
Small/Low	0.99	0.49	−1.15	−1.16
Small/High	1.09	1.43	−0.30	−0.31
Big/Low	0.61	0.93	0.08	0.08
Big/High	2.69	0.74	−0.14	−0.10

the bias, estimating the models using rolling versus expanding windows has little impact on predictive performance.

Table 3 also contains the results of the CTPs for both exercises. The test rejects equal predictive ability between the fixed and time-varying beta versions of the CAPM for all four assets except in the expanding window case for the large capitalization portfolios. For the CAPM, the decision rule consistently selects the fixed beta version of the model. The CTP rejects equal predictive ability between the fixed and time-varying beta versions of the Fama-French model only for small portfolios in the expanding window case. For the Fama-French model, the decision rule results are mixed. Overall, the results in these tables offer little evidence that the predictive capabilities of the models are improved by allowing beta to vary as either a linear or nonlinear function of the instruments.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

IV. Model Expectations of Individual Firm Returns

Firm-specific characteristics have a long history of being used (successfully) to refute the ability of asset pricing models to explain the cross-sectional variation of expected returns.[20] The more popular attempts to exploit the role of firm-specific characteristics have resulted in creating factor-mimicking portfolios that capture cross-sectional dispersion in expected returns. Beyond the polar views that portfolios formed by sorting on firm characteristics are proxies for risk factors (e.g., Fama and French (1993), (1996)) or that expected returns are determined solely by firm-specific characteristics and not risk (e.g., Daniel and Titman (1997)), others have tried to utilize the information content of firm-specific attributes within the context of time-varying parameter models.[21] I perform two experiments to determine if including firm-specific information increases the accuracy of the expectations produced by commonly used asset pricing models.

A. Model Comparisons Using Firm Attributes: A Cross-Sectional Approach

In the first exercise, I employ the Fama-MacBeth cross-sectional procedure following Brennan, Chordia, and Subrahmanyam (1998). Here I use a 60-month training period to estimate the model coefficients, β, for my candidate models for all available stocks. In the following month, I form fitted values of each model using the risk factor(s) and, if needed the instruments, and run the following cross-sectional regression to obtain the parameter γ_t, where X_t is the matrix of firm characteristics,

$$r_{i,t} - \hat{\beta}_t r_{p,t} = \gamma_t X_{t-1} + \varepsilon_{i,t}. \tag{15}$$

In the subsequent month, I form forecast errors as

$$e_{t+1} = r_{i,t+1} - \hat{\beta}_t E(r_{p,t}) - \hat{\gamma}_t X_t. \tag{16}$$

Rolling the entire procedure through the dataset one month at a time produces a time series of forecast errors for each firm. Given the poor results in the previous tables when using the conditional expectation of the factor risk premia, I only examine the predictions based on the unconditional premia and for similar reasons do not estimate specifications (7) and (9).

The returns data represent all monthly common stock returns on CRSP from January 1962 through 1999 excluding financial firms having at least 25 returns and firm characteristic observations in the training period. On average, there are 3,139 firms in the monthly cross sections. I use the same firm characteristics as those in Brennan, Chordia, and Subrahmanyam (1998). These include: i) the

[20] See Basu (1977), He, Kan, Ng, and Zhang (1996), Fama and French (1996), and Brennan, Chordia, and Subrahmanyam (1998), among others.

[21] Early work by Rosenberg and McKibben (1973) and Rosenberg and Marathe (1979) posit model parameters as linear functions of firm-specific attributes. Fama and French (1997) attempt to capture time variation in the loadings on their SMB and HML factors in a similar way. In an international setting, Ferson and Harvey (1997) define time variation in beta as a function of country-specific attributes.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

lagged natural logarithm of the market value of equity of the firm as of the end of the previous month; ii) the natural logarithm of the ratio of the book value of equity plus deferred taxes to the market value of equity, using the end of the previous year's market and book values; iii) the lagged natural logarithm of the reciprocal of the share price at the end of the previous month; iv) the lagged dividend yield as measured by the sum of all dividends paid over the previous 12 months divided by the share price at the end of the previous month; v) the lagged natural logarithm of the dollar volume of trading in the security in the last month; vi) the natural logarithm of the cumulative return over the two months ending at the beginning of the previous month; vii) the natural logarithm of the cumulative return over the three months ending three months previously; and viii) the natural logarithm of the cumulative return over the six months ending six months previously.

Unlike in the previous sections, the number of observations for each asset is not constant. Firm forecast error series vary from one to 395. There is an average of 98 forecast error observations with 90% of the firms providing 19 or more. The top half of Table 4 contains summary statistics of the RMSFEs and the biases for the 10,010 firms that allow for at least one forecast error. The bottom portion of the table contains the relative performance of each model and benchmark based on the CPT described in equation (11). As before, each cell contains the number of times across all assets, the CPT rejected equal predictive ability at the 10% level, and the model in the column was chosen over the model in the row based on the decision rule described above.

As might be expected, the RMSFEs are larger for individual firm returns than for portfolio returns. Similar to the portfolio exercise, the constant benchmark provides the lowest RMSFE and most often outperforms the other benchmarks and models. Based on median RMSFEs and the CPTs, all four benchmark expectations outperform the asset pricing models. Unlike in the portfolio exercise, allowing for time variation in the parameters does not improve performance for either the CAPM or the Fama and French 3-factor model. For both models, the version without an intercept and a fixed beta is the best performing version, while the worst performing version allows for time variation in both the intercept and the slope coefficient(s). The bias in the forecasts is again small accounting for 3.8% of RMSFE on average and typically positive.[22]

B. Model Comparisons Using Firm Attributes: A Time-Series Approach

An alternative means of incorporating firm-specific information into the asset pricing models is to include firm characteristics as lagged instrumental variables. In my second experiment, I choose eight firm-specific characteristics as instrumental variables for the conditional models in the evaluation of individual firm data. These include the firm's dividend yield, book value to price ratio, and earnings to price ratio as instruments related to price; the ratio of trading volume to market capitalization as a characteristic related to liquidity; the return on equity as a characteristic related to growth potential; and the previous one-, two-, and 12-

[22] In unreported results, I repeat this exercise setting $\gamma = 0$. The forecast error series for most of the models shows a slight increase in bias on average. Because the bias component is such a small portion of the forecast errors there is little difference in RMSFEs.

TABLE 4

Individual Returns Accounting for Firm-Specific Attributes in the Cross Section

Panel A contains summary statistics for the distributions of RMSFE∗100 and Panel B the bias∗100 of the one month ahead predictions for the 10,010 firms sample. For each firm, the model parameters are estimated from 60-month rolling windows. The forecast errors for the unconditional means are the difference between the asset return value in the test period and the average of the value-weighted market proxy (asset return) during the training period. The conditional mean uses the slopes from regressing the value-weighted index on the lagged instruments in the training period and the prediction period values of the lagged instruments. In all cases, the forecast error distributions have been Winsorized by setting the values beyond the 5th and 95th percentiles to the values at the 5th and 95th percentiles. Each cell of Panel C contains the number of times, across the 527 firms, the test for conditional predictive ability rejected equal predictive ability at the 10% level and the model in the column was chosen over the model in the row based on the decision rule described in the text. Rank is the ranking across all sums of the column and represents the number of times the model in the column outperformed the other models and benchmarks.

	Benchmarks				CAPM					FF 3-Factor				
	Mean	Six	Umkt	Cmkt	Fb(U)	TaFb(U)	FaTb(U)	TaTb(U)	Tb(U)	Fb(U)	TaFb(U)	FaTb(U)	TaTb(U)	Tb(U)
Panel A. RMSFE														
Mean	18.84	18.68	18.72	18.77	19.05	20.11	19.48	20.93	19.38	18.93	20.07	21.49	23.93	21.31
Median	16.32	16.18	16.20	16.27	16.51	17.41	16.93	17.98	16.83	16.37	17.30	18.29	20.18	18.13
Std. Dev.	12.17	12.07	12.04	12.01	12.18	13.17	12.53	13.89	12.44	12.27	13.22	14.77	16.62	14.63
Panel B. Bias														
Mean	0.75	1.01	1.31	1.48	0.58	0.57	0.71	0.57	0.59	0.61	0.58	0.44	0.53	0.41
Median	0.59	0.89	1.36	1.53	0.62	0.62	0.63	0.59	0.66	0.51	0.53	0.43	0.49	0.42
Std. Dev.	4.67	4.85	4.82	4.81	4.72	5.08	4.88	5.68	4.68	5.01	5.37	6.72	8.99	6.25
	Benchmarks				CAPM					FF 3-Factor				
	Mean	Six	Umkt	Cmkt	Fb	TaFb	FaTb	TaTb	Tb	Fb	TaFb	FaTb	TaTb	Tb
Panel C. Model Comparisons														
Benchmarks														
Mean	—	2,302	1,751	1,331	446	143	222	82	270	416	125	50	24	57
Six	610	—	792	966	321	127	191	77	222	289	100	40	26	46
Umkt	1,077	2,381	—	1,121	367	140	224	84	245	399	117	48	24	56
Cmkt	1,682	2,784	2,535	—	657	209	357	115	414	867	221	45	29	65
CAPM														
Fb	2,827	3,534	3,112	2,533	—	263	414	127	483	1,831	431	134	38	154
TaFb	4,847	5,247	4,987	4,353	4,197	—	2,368	463	2,605	4,126	1,445	481	84	567
FaTb	3,852	4,313	4,000	3,422	2,738	663	—	276	1,787	2,896	818	175	56	250
TaTb	5,521	5,884	5,705	5,192	5,009	3,114	4,157	—	4,308	4,985	2,800	972	140	1,129
Tb	3,536	4,179	3,848	3,264	2,538	565	998	239	—	2,619	731	171	52	217
FF 3-Factor														
Fb	2,626	3,434	2,900	2,349	962	252	382	125	470	—	235	68	34	81
TaFb	5,015	5,433	5,122	4,399	3,299	1,241	2,008	553	2,257	4,480	—	425	62	489
FaTb	6,367	6,622	6,478	6,063	5,256	2,950	4,431	2,003	4,544	6,055	3,373	—	234	2,289
TaTb	7,690	7,828	7,801	7,632	7,192	6,239	6,685	5,505	6,789	7,513	6,729	5,117	—	5,322
Tb	6,025	6,442	6,259	5,732	4,942	2,582	3,912	1,742	4,269	5,862	3,025	909	210	—
Rank	3	1	2	4	6	10	8	11	7	5	9	13	14	12

month returns in excess of the return on the S&P 500 as technical or momentum-related instruments. Except for the dividend yield, these particular instruments are what Haugen and Baker (1996) determine are the seven "most important factors" out of more than 50 for explaining monthly return differentials across a set of individual securities over the 1979 to 1993 period (see their Table I).

In this second experiment, I restrict the individual firm sample to firms listed on the S&P 500 index at any time during the same period as in the previous exercise. To reduce delisting or backfilling biases, the monthly time series for each firm is restricted to the period(s) it was listed in the index.[23] To be included in the

[23]Limiting each firm's time series to the period it was listed in the index exposes the sample to two events that may change prices for non-risk-based reasons. The exclusion event presents no problems since S&P announces exclusions after the close of trading. Price changes at inclusion due to increased trading by institutions tracking the S&P are well documented, although whether these increases are

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

sample, a firm must have at least 61 contiguous data points of returns and one of the firm characteristics. The restrictions and merge process with Compustat leave 527 usable firm observations with time series ranging from 61 to 443. The sample contains slightly more positive than negative returns. Of the 97,842 individual observations, 52,449 are positive with an overall mean monthly return of 0.084 while the remaining 45,393 are less than or equal to zero with an overall monthly mean of −0.066.

Table 5 contains the summary statistics for the distributions of the step ahead RMSFEs, the bias components, and the relative performance of each model and benchmark based on the CPT. The number of observations for each asset varies with the amount of time it was included in the S&P index. Firm forecast error series vary from one to 383. There is an average of 120 forecast error observations with 90% of the firms providing 20 or more data points. Outliers in the conditional models forecast series are especially prevalent. To deal with these, all the forecast error distributions have been Winsorized at the 5% level prior to calculating the RMSFEs.

As in the previous section, the unconditional versions of the models consistently have smaller RMSFEs than do their conditional counterparts. Again, there is little evidence that any of the models have predictive capability beyond using the benchmarks, given the similarity of the distributions. Based on the CTP statistic, the unconditional models clearly produce more rejections than their conditional counterparts. The unconditional Fama and French model has the overall smallest average RMSFE with the smallest standard deviation, but only 66 of its forecasts are statistically different and smaller than those produced by using 6% as a predictor. There is little indication that using the firm-specific instruments help the conditional versions of the models. Overall the results mirror those above. The evidence clearly indicates that when forecasting one step ahead, unconditional versions of the asset pricing models outperform the conditional versions in terms of the RMSFE. While the amount of bias in the forecast error series is again small, there is no evidence that any version of either asset pricing model produces forecasts significantly better then using the constant of 6% as a predictor of portfolio returns.[24]

permanent or temporary is debatable (Dhillon and Johnson (1991) and Erwin and Miller (1998)). While temporary price increases (less then a month) impact one data point in the training period of the models, the comparisons between forecasts are conditional on the extent that price changes are permanent.

[24]In unreported results, I replaced the firm-specific instrument with the macroeconomic instruments used in Table 1. The use of firm-specific information in this time-series exercise shows a small improvement over using macroeconomic-based instruments.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

TABLE 5

Individual Returns Accounting for Firm-Specific Attributes through Time

Panel A contains summary statistics for the distributions of RMSFE∗100 and Panel B the bias∗100 of the one month ahead predictions for the 527 firms listed in the S&P 500 index. For each firm, the model parameters are estimated from 60-month rolling windows. The forecast errors for the unconditional means are the difference between the asset return value in the test period and the average of the value-weighted market proxy (asset return) during the training period. The conditional mean uses the slopes from regressing the value-weighted index on the lagged instruments in the training period and the prediction period values of the lagged instruments. In all cases, the forecast error distributions have been Winsorized by setting the values beyond the 5th, and 95th percentiles to the values at the 5th and 95th percentiles. Each cell of Panel C contains the number of times, across the 527 firms, the test for conditional predictive ability rejected equal predictive ability at the 10% level and the model in the column was chosen over the model in the row based on the decision rule described in the text. Rank is the ranking across all the sums of the column and represents the number of times the model in the column outperformed the other models and benchmarks.

	Benchmarks				CAPM					FF 3-Factor				
	Mean	Six	Umkt	Cmkt	Fb(U)	TaFb(U)	FaTb(U)	TaTb(U)	Tb(U)	Fb(U)	TaFb(U)	FaTb(U)	TaTb(U)	Tb(U)
Panel A. RMSFE														
Mean	8.87	8.87	8.87	9.14	8.87	9.38	8.92	9.47	8.89	8.86	9.36	9.03	9.61	9.00
Median	8.10	8.08	8.09	8.37	8.08	8.59	8.15	8.63	8.14	8.08	8.58	8.27	8.75	8.24
Std. Dev.	3.42	3.42	3.42	3.40	3.41	3.67	3.46	3.74	3.42	3.42	3.61	3.49	3.64	3.48
Panel B. Bias														
Mean	0.25	0.96	0.86	0.72	0.78	0.37	0.39	0.44	0.85	0.64	0.37	0.34	0.49	0.66
Median	0.02	0.86	0.79	0.55	0.73	0.25	0.19	0.30	0.82	0.53	0.27	0.19	0.37	0.54
Std. Dev.	1.75	1.87	1.85	2.01	1.84	1.80	1.82	1.95	1.81	1.82	1.76	1.93	2.09	1.95
	Benchmarks				CAPM					FF 3-Factor				
	Mean	Six	Umkt	Cmkt	Fb	TaFb	FaTb	TaTb	Tb	Fb	TaFb	FaTb	TaTb	Tb
Panel C. Model Comparisons														
Benchmarks														
Mean	—	59	52	33	52	3	33	1	38	61	2	16	3	23
Six	47	—	73	31	58	4	43	4	46	66	3	25	3	34
Umkt	44	69	—	44	64	2	39	3	60	68	2	22	4	31
Cmkt	134	182	173	—	165	25	125	14	146	173	24	70	13	93
CAPM														
Fb	40	76	79	41	—	2	50	3	60	90	2	23	4	33
TaFb	290	266	264	165	268	—	275	57	268	278	91	216	24	222
FaTb	121	87	78	37	78	4	—	2	72	92	2	29	6	42
TaTb	310	288	290	186	288	150	297	—	281	302	133	242	29	242
Tb	74	122	115	46	118	2	50	4	—	130	2	25	5	33
FF 3-Factor														
Fb	38	63	60	36	59	3	36	2	48	—	2	16	2	25
TaFb	295	265	266	161	267	76	276	53	264	266	—	216	24	211
FaTb	177	158	155	67	160	12	146	9	152	163	9	—	6	86
TaTb	335	322	317	228	323	167	323	150	313	331	186	302	—	298
Tb	144	160	167	62	164	12	122	8	155	180	9	45	6	—
Rank	5	2	3	10	4	12	7	13	6	1	11	9	14	8

V. Conclusion

This paper examines how closely expectations of equity returns generated by unconditional and conditional versions of two popular asset pricing models compare to the eventual realized returns. As Fama and MacBeth ((1973), p. 618) note, "As a normative theory the model only has content if there is some relation between future returns and estimates of risk that can be made on the basis of current information." Overall, I find little normative content in any of the specifications of the CAPM and the Fama-French 3-factor model. While the asset pricing models produce relatively unbiased time series of expectations, no version of either model consistently outperforms simple benchmarks such as the average market return or a constant of 6%. When estimating the expected returns of the size and

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

book-to-market portfolios, it is the unconditional versions that tend to achieve lower RMSFEs. When applied to individual equities, using firm-specific information to condition via either cross-sectional or time-series methods does little to narrow the gap in performance between unconditional and conditional versions of the models.

I do find evidence that incorporating the information content of commonly used instrumental variables worsens the quality of the model expectations. The results that the conditional expectation of the market is the worst performing benchmark and the poor performance of the specifications incorporating a conditional expectation of the factor premium when forecasting are consistent with the findings of Goyal and Welch (forthcoming) who show that none of the most common predictive variables have forecasting ability for average aggregate stock returns. Avramov and Chordia (2006) find that allowing for time variation in the intercept improves asset allocations when using linear factor models to generate the inputs to optimization algorithm. Even so, in my statistically-based examination, I find that allowing for time variation in betas is less detrimental to the expectations than allowing for time variation in the intercepts of the models.

Models employing nonlinear specifications of time-varying beta tend to have higher RMSFEs then their fixed beta counterparts. Whether time-varying betas help or hurt the performance of conditional models depends on how that time variation is modeled. The increased MSFE from estimating the conditional covariance and conditional variance of the time-varying beta as linear functions of the instruments coupled with the lack of a significant increase in bias implies that too much variance is introduced when estimating beta in this manner. The nonlinear specifications may simply be overstating the time variation in beta (Ghysels (1998), Braun, Nelson, and Sunier (1995)).

While it is tempting to view my results as pointing toward model uncertainty as the most prevalent source of uncertainty in estimating accurate expected returns, it is important to note that this exercise is not meant to be a formal test of the asset pricing models.[25] The failure of the conditional versions of the models may indicate a poor set of informational variables and hence incorrectly specified conditional versions of the models. The same conclusions can be reached comparing the unconditional and conditional versions of the benchmarks. Since the instrumental variables have traditionally been chosen based on their own ability to predict returns, the evidence here is consistent with the possibility of data mining discussed in Ferson, Sarkissian, and Simin (2003).

One possible interpretation of the results is that specifying the time variation in beta and the expected factor premia as linear functions of the instruments may not be correct. A possible solution would be to model the conditional beta using a set of firm-specific instruments while using macroeconomic instruments when estimating the factor risk premia (e.g., Ferson and Harvey (1995)). The linear specifications themselves may also be problematic. The rolling window analysis used with an unconditional model presents a crude method of allowing beta to vary through time, possibly nonlinearly. Comparing the unconditional and conditional models can be interpreted as comparing different ways of modeling time

[25] See Inoue and Kilian (2004) for general arguments against using out of sample experiments to test asset pricing models.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

variation in beta. Maheu and McCurdy (2006) discuss the non-optimality of using rolling window estimations in the presence of structural breaks in the data, leaving estimation error in the parameters due to structural change as a possible explanation for the failure of both the conditional and unconditional versions of the models.

Harvey and Siddique (2000) show a dramatic reduction of in sample RMSFE when skewness is added to their model of expected returns. In addition, they find evidence relating the momentum effect to skewness. Even though I find little indication of skewness in the forecast error series for the assets I examine, accounting for higher moments in asset returns may provide a solution to the weak performance of the asset pricing models. It is also possible that the relative rankings of the different versions of the asset pricing models change for longer run predictions. Fama and French (1997) explore the impact of mean reversion and nonstationarity in parameter estimates of the CAPM and their three-factor model in conjunction with longer run predictions of the cost of equity. Their mixed results leave open the potential for future work.

The lack of bias in the model error series and the fact the best performing benchmark of 6% has no variance in its predictions clearly indicates that reducing the variance of the model expectations would produce better forecasts. Alternatively, the benchmarks of Six and Umkt pull the forecasts toward different versions of the unconditional historical market average. The success of these ad hoc shrinkage estimators mirror that of the Bayesian methods used in Fama and French (1997). These results suggest that looking toward biased estimation techniques such as those suggested by James and Stein (1961) and Jorion (1986) may offer a means of producing better expectations from the popular linear asset pricing models.

The results here are related to work by Wheatley (1998) who develops a test for differences between the RMSFEs produced from using the asset historical mean and using forecasts from an asset pricing model. My results broadly confirm his test results. There is little evidence that using an asset pricing model provides better expectations of returns than simply using the mean. The results here also contribute to a discussion in the literature on which aspect of the process of estimating the cost of capital is the most important, i.e., a manager faces uncertainty about which model to use, the parameter estimates and the risk premium estimates. Ferson and Locke (1998) and Fama and French (1997) argue that uncertainty in the risk premiums is more important than errors in estimating the model parameters. Pástor and Stambaugh (1998) find that at the firm level, uncertainty about parameter estimates contributes substantially to the uncertainty in estimating the cost of equity, but to a lesser degree than the uncertainty of the risk premium. Finally, as Cooper and Gulen (2006) point out, my results are also indicative of the real time uncertainty faced by market participants when predicting returns.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

Appendix

A. Proof of Theorem 1

Suppose that the excess return r for an asset is generated by a linear function of K non-stochastic explanatory variables X and a random error term $\varepsilon \sim (0, \sigma^2)$. The regression model for returns is then

$$r = X\beta + \varepsilon, \tag{A-1}$$

where r is $T \times 1$, X is $T \times K$, β is $K \times 1$ and ε is $T \times 1$ and $\varepsilon \sim (0, \sigma^2) \forall t$, $E(\varepsilon_t \varepsilon_s) = 0$ for $t \neq s$. Partition $\beta' = (A', B')$, where A is a K_1-vector and B is K_2-vector such that $K_1 + K_2 = K$ and partition X accordingly as $X = (X_1, X_2)$ to get

$$r = X_1 A + X_2 B + \varepsilon. \tag{A-2}$$

Add to the model the pth period relation,

$$r_p = X'_{1p} A + X'_{2p} B + \varepsilon_p, \tag{A-3}$$

where $p > T$, r_p and ε_p are scalars, $\varepsilon_p \sim (0, \sigma^2)$, with ε_p and ε independent. Here X_p are the pth period observations on the regressors assumed to be random variables distributed independently of ε_p and ε, where $X'_{1p} = X'_{2p}\hat{C} + v$ with $\hat{C}$ being the $K_2 \times K_1$ coefficient matrix from regressing X_1 on X_2 and assume $E(X_{2p} v) = 0$. If theory states that predictors of r_p are of the form $\hat{r}_p = X'_{2p}\hat{B}$, the question that arises is whether it is better, in terms of mean squared error, to estimate $\hat{B}$ for this purpose from a restricted regression that imposes the constraint that $A = 0$ or to estimate the unrestricted regression and apply the constraint when forming the predicted values.

The predicted value of returns when A is set to 0 prior to estimating the parameters can be written as

$$\begin{aligned} \hat{r}_p &= X'_{2p} \left(X'_2 X_2\right)^{-1} X'_2 r \\ &= X'_{2p} \left(X'_2 X_2\right)^{-1} X'_2 \left(X_1 A + X_2 B + \varepsilon\right), \end{aligned} \tag{A-4}$$

and the forecast error defined as the deviation of the forecast $\hat{r}_p$ from the true value of r_p is then

$$\begin{aligned} r_p - \hat{r}_p &= \left(X'_{1p} A + X'_{2p} B + \varepsilon_p\right) - X'_{2p} \left(X'_2 X_2\right)^{-1} X'_2 \left(X_1 A + X_2 B + \varepsilon\right) \\ &= X'_{1p} A + \varepsilon_p - X'_{2p} \left(X'_2 X_2\right)^{-1} X'_2 X_1 A - X'_{2p} \left(X'_2 X_2\right)^{-1} X'_2 \varepsilon. \end{aligned} \tag{A-5}$$

The MSFE conditional on X_p is given as

$$\begin{aligned} E\left\{(r_p - \hat{r}_p)^2 \,|X_p\right\} &= A' X_{1p} X'_{1p} A + A' X'_1 X_2 \left(X'_2 X_2\right)^{-1} X_{2p} X'_{2p} \left(X'_2 X_2\right)^{-1} X'_2 X_1 A \\ &\quad - 2A' X_{1p} X'_{2p} \left(X'_2 X_2\right)^{-1} X'_2 X_1 A + \sigma^2 \left[1 + X'_{2p} \left(X'_2 X_2\right)^{-1} X_{2p}\right]. \end{aligned} \tag{A-6}$$

The predicted value of returns when A is set to 0 *after* estimating the parameters without restriction can be written as

$$\begin{aligned} \hat{r}_p &= X'_{2p} \left(X'_2 M X_2\right)^{-1} X'_2 M r \\ &= X'_{2p} \left(X'_2 M X_2\right)^{-1} X'_2 M \left(X_1 A + X_2 B + \varepsilon\right), \end{aligned} \tag{A-7}$$

where $M = I - X_1 (X'_1 X_1)^{-1} X'_1$, I is a $T \times T$ identity matrix. Since $MX_1 = 0$, the forecast error defined as the deviation of the forecast $\tilde{r}_p$ from the true value of r_p is then

$$\begin{aligned} r_p - \tilde{r}_p &= \left(X'_{1p} A + X'_{2p} B + \varepsilon_p\right) - X'_{2p} \left(X'_2 M X_2\right)^{-1} X'_2 M \left(X_1 A + X_2 B + \varepsilon\right) \\ &= X'_{1p} A + \varepsilon_p - X'_{2p} \left(X'_2 M X_2\right)^{-1} X'_2 M \varepsilon. \end{aligned} \tag{A-8}$$

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

The MSFE conditional on X_p is given as

$$\text{(A-9)} \qquad E\left\{(r_p - \tilde{r}_p)^2 \,|X_p\right\} = A'X_{1p}X_{1p}'A + \sigma^2\left[1 + X_{2p}'\left(X_2'MX_2\right)^{-1}X_{2p}\right].$$

The difference between these conditional mean squared errors is then

$$\begin{aligned}\text{(A-10)}\; & E\left\{(r_p - \tilde{r}_p)^2 \,|X_p\right\} - E\left\{(r_p - \hat{r}_p)^2 \,|X_p\right\} \\ &= 2A'X_{1p}X_{2p}'\left(X_2'X_2\right)^{-1}X_2'X_1A - A'X_1'X_2\left(X_2'X_2\right)^{-1}X_{2p}X_{2p}'\left(X_2'X_2\right)^{-1}X_2'X_1A \\ &\quad + \sigma^2 X_{2p}'\left[\left(X_2'MX_2\right)^{-1} - \left(X_2'X_2\right)^{-1}\right]X_{2p}.\end{aligned}$$

The last line of (A-10) is a positive constant, call it Σ, since

$$\left(X_2'MX_2\right)^{-1} - \left(X_2'X_2\right)^{-1} = \left(X_2'X_2 - X_2'X_1\left(X_1'X_1\right)^{-1}X_1'X_2\right)^{-1} - \left(X_2'X_2\right)^{-1},$$

and

$$X_2'X_1\left(X_1'X_1\right)^{-1}X_1'X_2 \geq 0 \Leftrightarrow X_2'X_2 \geq X_2'X_2 - X_2'X_1\left(X_1'X_1\right)^{-1}X_1'X_2,$$

indicating that $[(X_2'MX_2)^{-1} - (X_2'X_2)^{-1}]$ is a positive semi-definite matrix. So,

$$\begin{aligned}\text{(A-11)} \qquad & E\left\{(r_p - \tilde{r}_p)^2 \,|X_p\right\} - E\left\{(r_p - \hat{r}_p)^2 \,|X_p\right\} \\ &= 2A'X_{1p}X_{2p}'\hat{C}A - A'\hat{C}'X_{2p}X_{2p}'\hat{C}A + \Sigma \\ &= 2A'\left(\hat{C}'X_{2p} + v\right)X_{2p}'\hat{C}A - A'\hat{C}'X_{2p}X_{2p}'\hat{C}A + \Sigma \\ &= A'\hat{C}'X_{2p}X_{2p}'\hat{C}A + A'vX_{2p}'\hat{C}A + \Sigma.\end{aligned}$$

By the *law of iterated expectations*, taking the unconditional expectation of the difference in conditional expectations produces the *unconditional* expected difference in the squared forecast errors as

$$\text{(A-12)} \qquad E\left\{(r_p - \tilde{r}_p)^2 - (r_p - \hat{r}_p)^2\right\} = E(A'\hat{X}_{1p}\hat{X}_{1p}'A + \Sigma) \geq 0,$$

where $\hat{X}_{1p}' = X_{2p}'\hat{C}$.

For the unconditional setting discussed in the text, X_1 will contain a vector of ones and X_2 will contain the time t observations on the relevant risk factors. In a conditional setting as described by equation (8) in the text, X_1 will contain a vector of ones as well as the observations of the lagged instrumental variables while X_2 will contain the time t observations on the relevant risk factors as well as the products of the instrumental variables and the risk factors. In either case the magnitude of the difference depends on the strength of the relation between the instruments and the factors.

Table A.1 contains comparisons of RMSFEs for conditional and unconditional versions of a single factor model estimated with and without an intercept. Each cell of the table contains the percentage increase or decrease in the RMSFE from including an intercept for 5,000 independent one-period out of sample forecasts when the models' parameters are estimated using 60 observations. Several different versions of the data are generated allowing for varying degrees of explanatory power of the factor and instrument indicated by the R^2s of 0.01, 0.1, 0.5, and for the different levels of autocorrelation in the instrumental variable indicated by ρ ranging from 0.0 to 0.999 in Table A.1. Returns are created exactly as in Ferson, Sarkissian, and Simin (2003).

The results indicate a general tendency for a rightward shift in the distribution of the mean of the squared forecast error distributions when an intercept is included in the estimation of the model parameters. This is especially true for the conditional models where the intercept includes a constant and the lagged instrumental variables. Although the percentage increases are rarely greater than 1%, the simulation indicates that when making a forecast without an intercept, estimating a model without an intercept will on average produce better forecasts within the context of a quadratic loss function.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

TABLE A.1

Percentage Increase or Decrease in RMSFE from Including an Intercept

	$R^2 = 0.01$		$R^2 = 0.1$		$R^2 = 0.5$	
ρ	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional
0	0.186	0.518	0.240	0.923	0.140	0.945
0.1	0.184	0.534	0.237	0.921	0.138	0.948
0.5	0.178	0.658	0.229	0.900	0.102	1.003
0.9	0.169	0.846	0.287	0.907	−0.089	1.005
0.95	0.165	0.862	0.319	0.849	0.141	1.011
0.98	0.166	0.821	0.344	0.736	0.462	0.964
0.99	0.170	0.821	0.345	0.687	0.455	0.860
0.995	0.174	0.834	0.345	0.674	0.361	0.744
0.999	0.181	0.858	0.352	0.667	0.249	0.566

B. Description of the Instrumental Variables

Macroeconomic Instruments

1. The risk-free rate used as an instrument is the average of the bid and ask for the U.S. one-month T-bill from the risk-free CRSP file. This yield is a continuously compounded 365-day rate de-annualized using rf = log(1 + (AVE_1)/100)/12. The Treasury bill selected in the one-month series has a minimum of 30 days to maturity, and is the closest T-bill to 30 days to maturity.
2. The Div. Yield for the period 1/27–12/99, is calculated by summing up 12 lags of $(I_{\text{NYSE},t} * (1 + R_{\text{NYSE},t+1})) - I_{\text{NYSE},t+1}$ and then dividing by $(I_{\text{NYSE},t+12})$, where I_{NYSE} is the level of the value-weighted NYSE index excluding dividends (VWINDX) and R_{NYSE} is the return on the value-weighted NYSE index including dividends (WRETD) found in the MSIA CRSP indices files.
3. Junk is the difference between the CITIBASE series FYBAAC: Moody's BAA corporate bond yield and FYAAAC: Moody's AAA corporate Bond Yield (%/annum) for the period 1/47–12/99 divided by 100. The series is updated from 6/96 to 12/99 and backdated from 1/26 to 12/46 using Moody's Industrial Manual.
4. Term is the spread between the lagged Moody's Composite Average of Yields on Corporate Bonds (industrial Manual)/1200 and U.S. three-month T-bill from the risk-free CRSP file described above.

Firm-Specific Instruments

The firm-specific instruments include the dividend yield and the seven most important "factors" from Haugen and Baker (1996). The Haugen and Baker factors are formed following Hanna and Ready (2005) and their descriptions follow.

1. One-month, two-month, and twelve-month excess return: Monthly returns on the S&P index are calculated by adding the dividend return on the value-weighted index to the S&P return in CRSP excluding dividends. The excess return is the difference between the cumulative S&P return and the cumulative firm return over the x-month interval ending with the month prior to the firm's observation. If the firm return is missing at either the beginning or the end of the interval, the excess return is missing.
2. Trading volume/Market cap.: Monthly trading volume divided by shares outstanding $(\times 100)$ is calculated for each of the previous 60 months. A good observation requires non-negative and nonmissing volume and positive shares outstanding. The variable is the average of the most recent 12 good values.
3. Earnings to Price: Most recently available annual earnings number deflated by the most recently available monthly closing price depending on the time between the most recently available financial statements and the current month.
4. Return on Equity: For the most recently available financial statements, calculate the ratio of annualized net income over total equity.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

5. Book to Price: Most recently available common equity number divided by the number of shares used to calculate earnings per share to get common equity per share. Deflate this ratio by the most recently available monthly closing price depending on the time between the most recently available financial statements and the current month.
6. Dividend Yield: The average of the most recent 12 months of the firm return with dividends less the firm return without dividends.

References

Avramov, D., and T. Chordia. "Asset Pricing Models and Financial Market Anomalies." *Review of Financial Studies,* 19 (2006), 1001–1040.

Basu, S. "The Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: Test of the Efficient Markets Hypothesis." *Journal of Finance,* 32 (1977), 663–682.

Bossaerts, P., and P. Hillion. "Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?" *Review of Financial Studies,* 12 (1999), 405–428.

Braun, P. A.; D. B. Nelson; and A. M. Sunier. "Good News, Bad News, Volatility and Betas." *Journal of Finance,* 50 (1995), 1575–1603.

Brealey, R. A., and S. C. Myers. *Principles of Corporate Finance,* 6th ed. New York: McGraw Hill (2000).

Brennan, M. J.; T. Chordia; and A. Subrahmanyam. "Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns." *Journal of Financial Economics,* 49 (1998), 345–373.

Bruner, R. F.; K. M. Eades; R. S. Harris; and R. C. Higgins. "Best Practices in Estimating the Cost of Capital: Survey and Synthesis." *Financial Management,* 27 (1998), 13–28.

Chen, N. F.; R. Roll; and S. A. Ross. "Economic Forces and the Stock Market." *Journal of Business,* 59 (1986), 383–403.

Cochrane, J. H. "A Rehabilitation of the Stochastic Discount Factor Methodology." Working Paper #8533, NBER (2001).

Connor, G., and R. A. Korajczyk. "Estimating Pervasive Economic Factors with Missing Observations." Working Paper 34, Northwestern University (1988).

Connor, G., and R. A. Korajczyk. "A Test for the Number of Factors in an Approximate Factor Model." *Journal of Finance,* 48 (1993), 1263–1291.

Cooper, M., and H. Gulen. "Is Time-Series Based Predictability Evident in Real Time?" *Journal of Business,* 79 (2006), 1263–1292.

Daniel, K., and S. Titman. "Evidence on the Characteristics of Cross-Sectional Variation in Stock Returns." *Journal of Finance,* 52 (1997), 1–33.

Dhillon, U., and H. Johnson. "Changes in the Standard and Poor's 500 List." *Journal of Business,* 64 (1991), 75–85.

Ellis, D. M. "A Test of the Conditional CAPM with Simultaneous Estimation of the First and Second Conditional Moments." *Financial Review,* 31 (1996), 475–498.

Erwin, G. R., and J. M. Miller. "The Liquidity Effects Associated with Addition of a Stock to the S&P 500 Index: Evidence from Bid/Ask Spreads." *Financial Review,* 33 (1998), 131–146.

Evans, M. D. "Expected Returns, Time-Varying Risk, and Risk Premia." *Journal of Finance,* 49 (1994), 655–679.

Fama, E. F., and K. R. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics,* 33 (1993), 3–56.

Fama, E. F., and K. R. French. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance,* 51 (1996), 55–84.

Fama, E. F., and K. R. French. "Industry Costs of Equity." *Journal of Financial Economics,* 43 (1997), 153–193.

Fama, E., and J. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy,* 81 (1973), 607–636.

Farnsworth, H.; W. E. Ferson; D. Jackson; and S. Todd. "Performance Evaluation with Stochastic Discount Factors." *Journal of Business,* 75 (2002), 473–503.

Ferson, W. E., and C. R. Harvey. "The Variation of Economic Risk Premiums." *Journal of Political Economy,* 99 (1991), 385–415.

Ferson, W. E., and C. R. Harvey. "The Risk and Predictability of International Equity Returns." *Review of Financial Studies,* 6 (1993), 527–566.

Ferson, W. E., and C. R. Harvey. "Predictability and Time-Varying Risk in World Equity Markets." *Research in Finance*, 13 (1995), 25–85.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

Ferson, W. E., and C. R. Harvey. "Fundamental Determinants of National Equity Market Returns: A Perspective on Conditional Asset Pricing." *Journal of Banking and Finance,* 21 (1997), 1625–1665.
Ferson, W. E., and C. R. Harvey. "Conditional Variables and the Cross-Section of Stock Returns." *Journal of Finance,* 54 (1999), 1325–1360.
Ferson, W. E., and R. A. Korajczyk. "Do Arbitrage Pricing Models Explain the Predictability of Stock Returns?" *Journal of Business,* 68 (1995), 309–349.
Ferson, W. E., and D. H. Locke. "Estimating the Cost of Capital through Time: An Analysis of the Sources of Error." *Management Science,* 44 (1998), 485–500.
Ferson, W., and M. Qian. "Conditional Performance Evaluation Revisited." Monograph, The Research Foundation of the CFA Institute (2004).
Ferson, W. E.; S. Sarkissian; and T. T. Simin. "Spurious Regressions in Financial Economics?" *Journal of Finance,* 58 (2003), 1393–1413.
Foster, F. D.; T. Smith; and R. E. Whaley. "Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R^2." *Journal of Finance,* 52 (1997), 591–607.
Ghysels, E. "On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt?" *Journal of Finance,* 53 (1998), 549–573.
Giacomini, R., and H. White. "Tests of Conditional Predictive Ability." Working Paper #572, Boston College (2005).
Gibbons, M., and W. E. Ferson. "Testing Asset Pricing Models with Changing Expectations and an Unobservable Market Portfolio." *Journal of Financial Economics,* 14 (1985), 217–236.
Gitman, L. J., and V. A. Mercurio. "Cost of Capital Techniques Used by Major U.W. Firms: Survey and Analysis of Fortune's 1000." *Financial Management,* Winter (1982), 21–29.
Goyal, A., and I. Welch. "Predicting the Equity Premium." Working Paper, UCLA (1999).
Goyal, A., and I. Welch. "The Myth of Predictability: Does the Dividend Yield Forecast the Equity Premium?" *Management Science,* 49 (2003), 639–654.
Goyal, A., and I. Welch. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* (forthcoming).
Graham, J. R., and C. R. Harvey. "The Theory and Practice of Corporate Finance: Evidence from the Field." *Journal of Financial Economics,* 60 (2001), 187–243.
Grinold R. C., and R. N. Kahn. *Active Portfolio Management: Quantitative Theory and Applications.* New York, NY: McGraw-Hill (1995).
Hanna, D., and M. Ready. "Profitable Predictability in the Cross-Section of Stock Returns." *Journal of Financial Economics,* 78 (2005), 463–505.
Hansen, L. P. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica,* 50 (1982), 1029–1054.
Harvey, C. R. "Time-Varying Conditional Covariances in Tests of Asset Pricing Models." *Journal of Financial Economics,* 24 (1989), 289–317.
Harvey, C. R., and A. Siddique. "Conditional Skewness in Asset Pricing Tests." *Journal of Finance,* 55 (2000), 1263–1295.
Haugen, R. A., and N. L. Baker. "Commonality in the Determinants of Expected Stock Returns." *Journal of Financial Economics,* 41(1996), 401–439.
He, J.; R. Kan; L. Ng; and C. Zhang. "Tests of the Relations among Market-Wide Factors, Firm-Specific Variables, and Stock Returns Using a Conditional Asset Pricing Model." *Journal of Finance,* 51 (1996), 1891–1908.
Inoue, A., and L. Kilian. "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?" *Econometric Reviews,* 23 (2004), 371–402.
Jagannathan, R., and Z. Wang. "The Conditional CAPM and the Cross-Section of Expected Returns." *Journal of Finance,* 51 (1996), 3–53.
Jagannathan, R., and Z. Wang. "Empirical Evaluation of Asset Pricing Models: A Comparison of the SDF and Beta Methods." *Journal of Finance,* 57 (2002), 2337–2367.
James, W., and C. Stein. "Estimation with Quadratic Loss." In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, J. Neyman, ed. University of California Press (1961).
Jegadeesh, N., and S. Titman. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance,* 48 (1993), 65–91.
Jegadeesh, N., and S. Titman. "'Profitability of Momentum Strategies: An Evaluation of Alternative Explanations." *Journal of Finance*, 56 (2001), 699–720.
Jorion, P. "Bayes-Stein Estimation for Portfolio Analysis." *Journal of Financial and Quantitative Analysis*, 21 (1986), 279–292.
Kan, R., and C. Zhang. "Two-Pass Tests of Asset Pricing Models with Useless Factors." *Journal of Finance,* 54 (1999a), 203–235.

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press

Kan, R., and C. Zhang. "GMM Tests of Stochastic Discount Factor Models with Useless Factors." *Journal of Financial Economics*, 54 (1999b), 103–127.
Kan, R., and G. Zhou. "A Critique of the Stochastic Discount Factor Methodology." *Journal of Finance,* 54 (1999), 1221–1248.
Kan, R., and G. Zhou. "Empirical Asset Pricing: The Beta Method versus the Stochastic Discount Factor Method." Working Paper, University of Toronto (2001).
Kim, D. "The Errors in the Variables Problem in the Cross-Section of Expected Stock Returns." *Journal of Finance,* 52 (1995), 1605–1634.
Kirby, C. "Measuring the Predictable Variation in Stock and Bond Returns." *Review of Financial Studies,* 10 (1997), 579–630.
Leitch, G., and J. E. Tanner. "Economic Forecast Evaluation: Profits versus the Conventional Error Measures." *American Economic Review,* 81 (1991), 580–590.
Lettau, M., and S. Ludvigson. "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying." *Journal of Political Economy*, 109 (2001), 1238–1287.
Lintner, J. "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics,* 47 (1965), 13–37.
Lo, A., and A. C. MacKinlay. "Data-Snooping Biases in Tests of Financial Asset Pricing Models." *Review of Financial Studies,* 3 (1990a), 431–468.
Lo, A., and A. C. MacKinlay. "When Are Contrarian Profits Due to Stock Market Overreaction?" *Review of Financial Studies,* 3 (1990b), 175–205.
Maheu, J. M., and T. H. McCurdy. "How Useful Are Historical Data for Forecasting the Long-Run Equity Premium?" Working Paper, University of Toronto (2006).
Martin, R. D., and T. T. Simin. "Outlier Resistant Estimates of Beta." *Financial Analysts Journal,* 59 (2003), 56–69.
Newey, W. K., and K. D. West. "A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* 55 (1987), 703–708.
Pástor, L., and R. F. Stambaugh. "Cost of Equity Capital and Model Mispricing." *The Journal of Finance,* 54 (1999), 67–121.
Rosenberg, B., and V. Marathe. "Tests of Capital Asset Pricing Hypotheses." *Research in Finance,* 1 (1979), 115–223.
Rosenberg, B., and W. McKibben. "The Prediction of Systematic and Specific Risk in Common Stocks." *Journal of Financial and Quantitative Analysis,* 8 (1973), 317–333.
Ross, S. A. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory*, 13 (1976), 341–360.
Sharpe, W. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance,* 19 (1964), 425–442.
Wheatley, S. "Evaluating Asset Pricing Models." Working Paper, Australian Graduate School of Management (1998).

https://doi.org/10.1017/S0022109000003550 Published online by Cambridge University Press