Stock Return Predictability and Asset Pricing Models

Doron Avramov
University of Maryland

This article develops an asset allocation framework that incorporates prior beliefs about the extent of stock return predictability explained by asset pricing models. We find that when prior beliefs allow even minor deviations from pricing model implications, the resulting asset allocations depart considerably from and substantially outperform allocations dictated by either the underlying models or the sample evidence on return predictability. Under a wide range of beliefs about model pricing abilities, asset allocations based on conditional models outperform their unconditional counterparts that exclude return predictability.

Financial economists have identified economic variables such as the dividend yield and the term spread that predict future stock returns through time. Subsequently, several studies have examined whether this return predictability is explained by rational pricing or whether it is due to asset pricing misspecification [see, e.g., Campbell (1987), Ferson and Korajczyk (1995), and Kirby (1998)]. Studies such as these approach finance theory by focusing on two polar viewpoints: rejecting or not rejecting a pricing model based on hypothesis tests. However, such an approach fails to capture many aspects of both the model and the data that could be potentially useful for a decision maker [see, e.g., Pastor (2000)]. This article studies return predictability from a different perspective. In particular, it incorporates pricing restrictions on predictive regression parameters as a reference point for a hypothetical investor's prior belief. The investor uses the sample evidence about the extent of predictability to update various degrees of belief in a pricing model and then allocates funds across cash and stocks. Pricing models are expected to exert stronger influence on asset allocation when the prior confidence in

I thank Rob Stambaugh for numerous discussions and insightful suggestions. Comments of two anonymous referees and Cam Harvey (the editor) have greatly improved the focus and content of this work. I also thank the comments of Rella Avramov, Gurdip Bakshi, John Chao, Mark Chen, Tarun Chordia, Wayne Ferson, Jeremy Gold, Levant Guntay, Gergana Jostova, Chris Kirby, Robert Korajczyk (WFA discussant), Kai Li, Dilip Madan, Maureen O'Hara (the executive editor), Lubos Pastor, Chris Polk (EFA discussant), Nagpurnanand Prabhala, Arnold Zellner, seminar participants at Bocconi, University of British Columbia, Emory University, University of Maryland, the 2001 meetings of the Western Finance Association and European Finance Association, and the 2001 Accounting and Finance Conference at Tel Aviv University. Financial support from the General Research Board at the University of Maryland is gratefully acknowledged. All errors are mine. Address correspondence to Doron Avramov, Finance Department, University of Maryland, Van Munching Hall, College Park, MD 20742, or e-mail: davramov@rhsmith.umd.edu.

The Review of Financial Studies Vol. 17, No. 3
DOI: 10.1093/rfs/hhg059 Advance Access publication October 15, 2003

their validity is stronger and when they explain much of the sample evidence on predictability. Thus the investor's asset allocation provides a useful metric for assessing the economic significance of pricing model misspecification.

Our design has enough flexibility to incorporate realistic constraints on portfolio holdings. This feature is especially desirable given that U.S. investors are not allowed to establish unlimited long and short positions per Regulation T of the Federal Reserve Board. In the presence of portfolio constraints, equilibrium expected returns need not be linear in factor sensitivities. Thus it might be difficult to implement traditional asset pricing tests such as that of Gibbons, Ross, and Shanken (1989). Indeed, Hansen, Heaton, and Luttmer (1995) develop asymptotic tests for evaluating asset pricing misspecification in the presence of portfolio constraints. Ours is a finite sample metric. In addition, our framework does not attempt to deliver a binary (reject or not) decision about the validity of an asset pricing model. Instead, it analyzes how the sample evidence about model misspecification leads to updating of various degrees of prior belief about model accuracy, as reflected in asset allocation decisions.

We apply the approach to examine the restrictions on return predictability implied by the three-factor Fama and French model (1993; henceforth FF) and by two extended versions of FF. The first extension accounts for momentum documented by Jegadeesh and Titman (1993). The second follows Chen, Roll, and Ross (1986) in incorporating risk exposures related to interest rates and credit conditions. Momentum is captured by the winners-minus-losers portfolio (WML); interest rate risk is proxied by the return differential between long- and short-term government bonds (TERM); and risk related to credit conditions is captured by the return differential between low- and high-rated corporate bonds (DEF). We consider two sets of investable assets. The first set contains cash, the value-weighted Center for Research in Security Prices (CRSP) index, SMB (size premium), HML (value premium), WML, TERM, DEF, and 25 portfolios sorted on size and book-to-market. The second replaces the size and book-to-market portfolios by 25 industry portfolios. We first analyze the economic significance of deviations from pricing model implications, then evaluate performance of asset allocations based on various prior beliefs about the extent of predictability explained by asset pricing models. We also compare performance of asset allocations based on conditional pricing models versus unconditional ones that disregard conditioning information. Such performance is studied under a wide range of prior beliefs about pricing abilities of conditional and unconditional models, ranging from perfect confidence to complete skepticism.

Our approach yields several insights about the influence of conditional pricing models and the data upon financial decision making. For one, we find that asset allocation is extremely sensitive to the imposition of model restrictions on predictive regressions. Indeed, an investor who believes those restrictions are perfectly valid but is forced to allocate funds disregarding model implications faces an enormous utility loss. Furthermore, asset allocations depart considerably from those dictated by the pricing models when the prior allows even minor deviations from the underlying models. In the presence of short-sale constraints or the 50% margin requirement per Regulation T of the Federal Reserve Board, the departure from model implications declines substantially, and sometimes disappears, yet it often remains economically significant. On the other hand, when the prior reflects a fairly skeptical view about model pricing abilities, the resulting asset allocations could be much closer to those implied by the model alone than to those implied by the sample evidence on predictability alone. This suggests that pricing model implications for investment decisions could be important even when the possibility of model mispricing is recognized. We also demonstrate that asset allocations based on conditional pricing models and the sample evidence combined are sensible in that long and short positions are much less extreme than those obtained when model implications are completely disregarded.

Ex post out-of-sample analysis gives further insights about asset allocation with conditioning information. We demonstrate that optimal portfolios based on dogmatic beliefs in conditional pricing models deliver the lowest Sharpe ratios. In addition, completely disregarding pricing model implications results in the second lowest Sharpe ratios. Remarkably, much higher Sharpe ratios are obtained when asset allocations are based on the so-called shrinkage approach, in which inputs for portfolio optimization combine the underlying pricing model and the sample evidence on predictability. Indeed, a shrinkage approach for asset allocation has been entertained in earlier articles such as Black and Litterman (1992) and Pastor (2000). None of these, however, accounts for expected return variation, as we do. In fact, we show that asset allocations that reflect such variation considerably outperform allocations based on models that assume time-invariant expected returns. These performance-based findings suggest that expected returns do indeed vary over time and that the shrinkage approach to asset allocation helps performance even under expected return variation.

In related work, McCulloch and Rossi (1990) develop a utility-based metric to test asset pricing models. They assume that returns are unpredictable, and they focus on two polar views about the usefulness of asset pricing theories: perfect confidence versus complete skepticism. Our study differs in that it accounts for predictability and entertains some intermediate views about model misspecification. Kandel and Stambaugh (1996) investigate the predictive regression evidence from the perspective of a Bayesian investor who allocates funds between a single risky security and cash. Unlike their data-based approach, our multiple-security paradigm accounts for pricing model implications for predictive regressions. Pastor and Stambaugh (1999, 2000) and Pastor (2000) suggest that finance theory can be used to form informative prior beliefs in financial decision making. In their framework, one uses the sample evidence to update different degrees of prior belief in model mispricing. The idea of having various degrees of beliefs in a pricing model has also been suggested by Shanken (1987) in the context of testing portfolio efficiency. Shanken formalizes model beliefs as prior beliefs about the ratio of an index's Sharpe ratio to the Sharpe ratio of the tangency portfolio. Relative to Pastor and Stambaugh and Shanken, we account for expected return variation, which, as noted earlier, considerably improves ex post portfolio performance. Shanken and Tamayo (2001) examine the interaction between data and prior beliefs about the source of market index predictability. In this work we entertain prior degrees of belief about the usefulness of pricing models in order to explain why expected returns are different across securities and through time.

The remainder of the article proceeds as follows. Section 1 sets forth an econometric framework to address asset allocation when a hypothetical investor incorporates prior beliefs about the extent of stock return predictability explained by asset pricing models. Section 2 describes the data used in the empirical analysis and Section 3 presents the findings. Conclusions and avenues for future research are offered in Section 4. Unless otherwise noted, all derivations are presented in the appendix.

1. Predictability and Finance Theory: An Investment Perspective

1.1 Pricing model implications for predictive regressions

Financial economists have shown that asset pricing models imply restrictions on the intercept, slope, and goodness-of-fit in a time-series predictive regression. Under pricing models, the component of returns explained by forecasting variables is attributed to changes in risk premia, risk exposures, or both. Otherwise, return predictability is due to model misspecification, the extent of which may vary in response to changing economic conditions. The approach developed here is to assess asset pricing misspecification using a utility-based metric. We derive and compare portfolio decisions of three types of hypothetical Bayesian investors. The first exhibits a completely skeptical view about the usefulness of asset pricing theories. The second has some intermediate views. He neither takes a pricing model as a dogma nor discards it as useless. The third has complete faith in the accuracy of the asset pricing restrictions. These investors' initial views are updated by the data, and portfolio decisions with estimation risk are made on the basis of the updated beliefs. Optimal portfolios are then mapped onto riskfree certainty equivalent returns. The difference between these returns constitutes a powerful metric for assessing the economic significance of model mispricing, as reflected by both prior beliefs and the sample evidence on predictability.

In the following, we develop the econometric framework. First, we model excess returns (returns in excess of the one-month Treasury bill rate) on N investable assets as

$$r_t = \alpha(z_{t-1}) + \beta f_t + u_{rt}, \tag{1}$$

$$\alpha(z_{t-1}) = \alpha_0 + \alpha_1 z_{t-1}, \tag{2}$$

where f_t is a set of K monthly excess returns on portfolio-based factors, z_{t-1} is an M-vector of predictive variables observed at the end of time $t-1$, α_0 stands for an N-vector of the fixed component of asset mispricing, α_1 is an $N \times M$ matrix of the time-varying component, β is an $N \times K$ matrix of beta, and u_{rt} is an N-vector of a zero-mean security-specific event whose variance-covariance matrix, denoted by Σ_{RR}, is assumed constant over time. Modeling α as a linear function of instruments goes back to Rosenberg and Marathe (1979). Shanken (1990) and Ferson and Harvey (1999) simultaneously model α and β as linear functions of information variables. In our framework, β is time invariant, as in Campbell (1987), Connor and Korajczyk (1989), and Kirby (1998). Earlier work [see, e.g., Ferson and Harvey (1991) and Evans (1994)] shows that predictability is primarily driven by changes in risk premia over time, whereas the impact of beta variation is second order. Ferson and Harvey (1999), however, provide evidence of changes in beta, and Tamayo (2002) allows beta variation. Of course, a pricing model implies that $\alpha(z_{t-1})$ in Equation (1) is equal to zero. Below we derive the implications of this no mispricing for predictive regressions.

We model the evolution of factors and predictive variables using the specifications

$$f_t = \lambda(z_{t-1}) + u_{ft}, \tag{3}$$

$$\lambda(z_{t-1}) = \lambda_0 + \lambda_1 z_{t-1}, \tag{4}$$

$$z_t = \alpha_z + a_z z_{t-1} + u_{zt}, \tag{5}$$

where $\lambda(z_{t-1})$ is the expected value of f_t conditioned on z_{t-1}, u_{ft} is a K-vector of zero-mean unexpected factor realizations with a variance-covariance matrix, denoted by Σ_{FF}, that is assumed constant over time, λ_0 is a K-vector, λ_1 is a $K \times M$ matrix, α_z is an M-vector, and a_z is an $M \times M$ matrix. As noted earlier, in this work factors are portfolio based. Thus the conditional expected value $\lambda(z_{t-1})$ stands for risk premia, or market price of beta risks. Risk premia can vary over time if $\lambda_1 \neq 0$. In addition, observe from Equation (5) that the predictive variables are modeled using vector autoregression of order one. That specification has been used in an investment context also by Kandel and Stambaugh (1996), Stambaugh (1999), Barberis (2000), and Avramov (2002).

Now, a conditional version of an asset pricing model implies the relation

$$\mathbb{E}(r_t \mid z_{t-1}) = \beta\lambda(z_{t-1}) \tag{6}$$

for all t, where $\mathbb{E}$ stands for the expected value operator. The pricing model in Equation (6) implies restrictions on parameters and goodness-of-fit in the time-series predictive regression

$$r_t = \mu_0 + \mu_1 z_{t-1} + v_t, \tag{7}$$

where μ_0 is an N-vector and μ_1 is an $N \times M$ matrix. In particular, note that by adding to the right hand-side of Equation (7) the quantity $\beta(f_t - \lambda_0 - \lambda_1 z_{t-1})$, subtracting the (same) quantity βu_{ft}, and decomposing the residual in Equation (7) into two orthogonal components $v_t = \beta u_{ft} + u_{rt}$, we reparameterize the return-generating process in Equation (7) as

$$r_t = (\mu_0 - \beta\lambda_0) + (\mu_1 - \beta\lambda_1) z_{t-1} + \beta f_t + u_{rt}. \tag{8}$$

Matching the right-hand side coefficients in Equation (8) with those in Equation (1) yields

$$\mu_0 = \alpha_0 + \beta\lambda_0, \tag{9}$$

$$\mu_1 = \alpha_1 + \beta\lambda_1. \tag{10}$$

Equation (10) indicates that when a model fails to explain why average returns differ across stocks and over time, that is, when $\alpha(z_{t-1}) \neq 0$ in Equation (1), return predictability, if it exists, is due to the security-specific model mispricing component ($\alpha_1 \neq 0$) or the common component in risk premia that varies ($\lambda_1 \neq 0$). When mispricing is precluded, the regression parameters that conform to asset pricing models are

$$\mu_0 = \beta\lambda_0, \tag{11}$$

$$\mu_1 = \beta\lambda_1. \tag{12}$$

The pricing restrictions of Equations (11) and (12) capture the fact that the intercepts and slopes in a multivariate time-series predictive regression are equal to beta times the intercepts and slopes in the regression of portfolio-based factors on lagged forecasting variables.

Earlier studies [see, e.g., Campbell (1987) and Kirby (1998)] test pricing restrictions on predictability using statistical procedures such as the general method of moments (GMM) of Hansen (1982). In essence, asset pricing tests are founded on the two polar viewpoints of rejecting or not rejecting a model. This work studies predictability from a different angle. In particular Equations (11) and (12) serve as a reference point for informative prior beliefs in an investment context. Our aim is to assess the economic significance of deviations from pricing model implications by comparing asset allocations based on various degrees of confidence in the underlying model.

The metric developed here is attractive for several reasons. For one, it applies to finite samples. In contrast, asset pricing tests with conditioning information, like the GMM, are valid only asymptotically. In addition, the proposed metric flexibly accommodates portfolio constraints. Most importantly, our framework is sufficiently general for analyzing scenarios where an investor neither takes a pricing model as a dogma nor discards model implications as completely worthless. Under such prior beliefs, inputs for asset allocations are based on both the pricing model and the sample evidence on predictability. Indeed, a shrinkage approach to asset allocation has been suggested by Black and Litterman (1992) and Pastor (2000). However, these studies assume that expected return is constant over time. In this work, expected return could vary due to time-varying risk premia and (potential) model mispricing.

We note that this study focuses on a conditional version of beta pricing models, whereas recent works [see, e.g., Lettau and Ludvigson (2001)] have analyzed the empirical performance of asset pricing models using the pricing kernel representation

$$\xi_{t+1} = a_t + b_t' f_{t+1}, \tag{13}$$

where ξ is the stochastic discount factor. The parameters a and b in Equation (13) are often modeled as linear functions of forecasting variables. Indeed, Kan and Zhou (1999) have shown that the beta pricing setup is at least as efficient as the stochastic discount factor approach. In addition, we note that our constant beta framework is entirely consistent with time-varying pricing kernel parameters. To illustrate, using the equivalence [see Cochrane (1996)] between the representations in Equations (6) and (13) yields

$$a_t = 1/r_{ft}(1 + \lambda(z_t)' \Sigma_{FF}^{-1} \lambda(z_t)), \tag{14}$$

$$b_t = -1/r_{ft} \Sigma_{FF}^{-1} \lambda(z_t), \tag{15}$$

where r_{ft} is the riskfree rate. Observe from Equations (14) and (15) that variation in risk premiums or riskfree rates implies time-varying a_t and b_t, even when beta is constant over time.

In what follows, it will be convenient to work with the multivariate forms of the dynamics of excess returns, factors, and predictive variables. These are given by

$$R = W\Gamma + U_R, \tag{16}$$

$$F = XA_F + U_F, \tag{17}$$

$$Z = XA_Z + U_Z, \tag{18}$$

where $R = [r_1, r_2, \ldots, r_T]'$, $W = [X, F]$, $X = [\iota_T, Z_{-1}]$, ι_T is a $T \times 1$ vector of ones, $Z_{-1} = [z_0, z_1, \ldots, z_{T-1}]'$, z_0 is an $M \times 1$ vector of initial values of the predictive variables, $F = [f_1, f_2, \ldots, f_T]'$, $Z = [z_1, z_2, \ldots, z_T]'$, $\Gamma = [\alpha, \beta]'$, $\alpha = [\alpha_0, \alpha_1]$, $A_F = [\lambda_0, \lambda_1]'$, and $A_Z = [\alpha_Z, a_Z]'$. The dimensions of the matrices are as follows: R and U_R are $T \times N$, F and U_F are $T \times K$, Z and U_Z are $T \times M$, X is $T \times (M+1)$, W is $T \times (M+K+1)$, Γ is $(M+K+1) \times N$, A_F is $(M+1) \times K$, and A_Z is $(M+1) \times M$. We assume that the system of residuals $V = [U_R, U_F, U_Z]$ has the distribution $\mathrm{vec}(V) \sim N(0, \Sigma \otimes I_T)$, where $\mathrm{vec}(V)$ is the column-wise vectorization of V, the symbol $\otimes$ stands for the Kronecker product, and I_T is a $T \times T$ identity matrix.

The rest of this section proceeds as follows. Subsection 1.2 formulates various degrees of prior belief about the validity of the asset pricing restrictions in Equations (11) and (12). Prior beliefs are combined with the data to produce posterior beliefs. Based on posterior beliefs one can derive the so-called predictive distribution, introduced by Zellner and Chetty (1965), as an input in computing asset allocation with estimation risk. Subsection 1.3 describes the utility-based metric for analyzing model mispricing.

1.2 Prior

1.2.1 Polar views concerning the validity of pricing restrictions on predictive regressions. We consider two polar viewpoints. First, asset mispricing is part of the parameter space, denoted by Θ_U. Second, conditional pricing models hold exactly, meaning that α_0 and α_1 are excluded from the set of parameters, denoted by Θ_R. The first specification could be viewed as one where investors believe that pricing models do not explain return predictability. The second reflects perfect confidence about pricing model implications for predictability, as formulated in Equations (11) and (12). Investors' prior beliefs about Θ_U are represented by the standard noninformative diffuse prior

$$P(\Theta_U) \propto |\Sigma|^{-\frac{N+K+M+1}{2}}. \tag{19}$$

Initial beliefs about all parameters in Θ_R are noninformative as well. The appeal of the noninformative prior stems from the close correspondence that results between the posterior densities of the unknown parameters and classical confidence intervals for the parameters. The noninformative prior has been used in several other studies in financial economics including McCulloch and Rossi (1990) and Barberis (2000).

1.2.2 Intermediate views. What if, prior to observing the data, investors believe that asset pricing restrictions on stock return predictability could be useful, but that no pricing model is perfect? As proposed by Pastor and Stambaugh (1999, 2000) and Pastor (2000) one can formulate prior beliefs about finance theory in a Bayesian framework. We use two prior specifications of beliefs concerning the usefulness of finance theory.

The first draws on Pastor and Stambaugh. Prior beliefs about time-varying model mispricing are concentrated at zero, that is, $\alpha_1 = 0$ in Equation (2). Prior beliefs about α_0 are

$$P(\alpha_0 \mid \Sigma_{RR}) \propto |\Sigma_{RR}|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\alpha_0' \left(\frac{\sigma^2}{s^2}\Sigma_{RR}\right)^{-1} \alpha_0 \right\}, \tag{20}$$

where σ indicates the degree of confidence about the pricing abilities of a model, and s^2 is computed as the cross-sectional average of the sample variance of the residuals in Equation (1). Note that the prior density implies a positive relation between asset mispricing and the residual covariance matrix. This relation is motivated by MacKinlay (1995) as a means of reducing the probability of high Sharpe ratios. For priors that treat asset mispricing and the residual covariance matrix as independent, high Sharpe ratios could simply result from combining the test assets with the benchmark portfolios.

The second specification entertains informative prior beliefs about α_1 as well. As in Kandel and Stambaugh (1996), we assume that an optimizing investor observes a hypothetical prior sample of length T_0 weighted against return predictability. On the basis of the hypothetical sample, regressing r_t on one, z_{t-1}, and f_t yields a zero estimate for α_1, and regressing f_t on one and z_{t-1} yields a zero estimate for λ_1. In addition, it is assumed that the estimate of α_0, like that of α_1, is taken to be zero in the hypothetical return regression. That is, not only is the hypothetical prior sample weighted against return predictability, but it is weighted against model misspecification, both fixed and time varying, as well. As T_0 approaches infinity, risk premiums are constrained to be time invariant and the fixed and time-varying asset mispricing components disappear.

It is assumed that the mean and variance estimates of r_t, f_t, and z_t based on the hypothetical sample are equal to the actual sample counterparts, which are given by

$$\bar{r}=\frac{1}{T}\sum_{t=1}^{T}r_t,\quad \bar{f}=\frac{1}{T}\sum_{t=1}^{T}f_t,\quad \bar{z}=\frac{1}{T}\sum_{t=0}^{T-1}z_t, \tag{21}$$

$$\hat{V}_r=\frac{1}{T}\sum_{t=1}^{T}(r_t-\bar{r})(r_t-\bar{r})', \tag{22}$$

$$\hat{V}_f=\frac{1}{T}\sum_{t=1}^{T}(f_t-\bar{f})(f_t-\bar{f})', \tag{23}$$

$$\hat{V}_z=\frac{1}{T}\sum_{t=0}^{T-1}(z_t-\bar{z})(z_t-\bar{z})'. \tag{24}$$

Using sample statistics to specify some of the parameters of the prior distribution is commonly termed an "empirical Bayes" approach [see, e.g., Robbins (1955, 1964)]. Combining the noninformative prior in Equation (19) with the hypothetical sample yields the prior densities

$$\mathrm{vec}(\Gamma)|\Sigma_{RR}\sim N(\mathrm{vec}(\Gamma_0),\Sigma_{RR}\otimes(W_0'W_0)^{-1}), \tag{25}$$

$$\Sigma_{RR}\sim IW(T_0(\hat{V}_r-\beta_0\hat{V}_f\beta_0'),T_0-1), \tag{26}$$

$$\mathrm{vec}(A_F)|\Sigma_{FF}\ \sim\ N\left(\mathrm{vec}(A_{F0}),\frac{1}{T_0}\Sigma_{FF}\otimes\begin{bmatrix}1+\bar{z}'\hat{V}_z^{-1}\bar{z} & -\bar{z}'\hat{V}_z^{-1}\\ -\hat{V}_z^{-1}\bar{z} & \hat{V}_z^{-1}\end{bmatrix}\right), \tag{27}$$

$$\Sigma_{FF}\ \sim\ IW(T_0\hat{V}_f,T_0+N-1), \tag{28}$$

where W_0 is the hypothetical sample analog of $W=[X,F]$, IW is the inverted Wishart distribution [see, e.g., Zellner (1971)], $\Gamma_0=[0_{N,M+1},\beta_0]'$, $A_{F0}=[\bar{f},0_{K,M}]'$, $0x,y$ denotes an $x\times y$ matrix of zeros, β_0 is (following the assumptions about the means of r and f) the slope in the zero-intercept regression of r_t on f_t, and

$$(W_0'W_0)^{-1}=\frac{1}{T_0}\begin{bmatrix}1+\bar{z}'\hat{V}_z^{-1}\bar{z}+\bar{f}'\hat{V}_f^{-1}\bar{f} & -\bar{z}'\hat{V}_z^{-1} & -\bar{f}'\hat{V}_f^{-1}\\ -\hat{V}_z^{-1}\bar{z} & \hat{V}_z^{-1} & 0\\ -\hat{V}_f^{-1}\bar{f} & 0 & \hat{V}_f^{-1}\end{bmatrix}. \tag{29}$$

This second prior specification has an interesting interpretation in terms of shrinkage of the posterior means. To illustrate, consider the case where $\hat{\lambda}_1$, the sample estimate of λ_1, is zero. Then, one can show that the posterior means of α and β are weighted averages of prior means and sample estimates with weights $\frac{T_0}{T+T_0}$ and $\frac{T}{T+T_0}$, respectively. That is, the posterior mean of α is shrunk toward zero, a value implied by the pricing restrictions on predictability. When $\hat{\lambda}_1 \neq 0$, the weights $\frac{T_0}{T+T_0}$ and $\frac{T}{T+T_0}$ hold approximately. In addition, the posterior means of λ_0 and λ_1 are shrunk toward $\bar{f}$ and zero, respectively. These values are implied by time-invariant risk premiums.

How much asset mispricing is allowed in the second informative prior specification? Obviously less is allowed when T_0 is larger, but how much less in absolute terms? To get a sense of the magnitude, we rely on the case where returns are i.i.d. distributed. Following Pastor (2000), under the first prior specification the posterior mean of α_0 is equal to $\frac{\hat{\alpha}_0}{1+\frac{s^2}{T\sigma^2}(1+SR^2)}$, where $\hat{\alpha}_0$ is the sample estimate of α_0 and SR^2 is the squared Sharpe ratio of a tangency portfolio formed using benchmark assets only. Under exact pricing, this Sharpe ratio is equal to that formed using both test and benchmark assets. In the second prior specification, the posterior mean of α_0 is equal to $\frac{T\hat{\alpha}_0}{T+T_0}$. Matching the posterior mean of α_0 under the two prior specifications yields

$$T_0 = \frac{s^2}{\sigma^2}(1+SR^2), \tag{30}$$

which establishes the inverse relation between the mispricing uncertainty and the hypothetical sample size. To illustrate, let $s^2 = 0.0011$, as in Pastor (2000), and let $SR^2 = 0.09$. The values $\sigma = 1\%$, $\sigma = 0.5\%$, and $\sigma = 0.1\%$ correspond to $T_0 = 12$, $T_0 = 48$, and $T_0 = 1200$. The relation in Equation (30) holds also under more general conditions where risk or risk premiums or both are time varying but $\alpha_1 = 0$. In particular, when beta varies with the predictive variables, as in Shanken (1990), the largest admissible Sharpe ratio is the one obtained based on both the K fundamental factors f_t and the KM-scaled factors described by the Kronecker product $f_t \otimes z_{t-1}$.

1.3 A utility-based metric

The investors considered in this work find historical evidence useful in assessing investment opportunities. However, observing only a finite sample, they do not know the true parameter values. As a result, part of the risk they rationally perceive arises from parameter uncertainty. Investment opportunities in the presence of parameter uncertainty are expressed by the Bayesian predictive distribution. The predictive distribution has been extensively used in asset allocation in the context of both i.i.d. distributed stock returns [Bawa, Brown, and Klein (1979), Pastor (2000), Pastor and Stambaugh (2000), and Li (2003)] and time-varying expected stock returns [Kandel and Stambaugh (1996), Shanken and Tamayo (2001), and Avramov (2002)].

The Bayesian predictive distribution integrates over the posterior distribution that summarizes the uncertainty about the parameters given the sample observed and given the model that generates stock returns, factors, and predictors. The appendix derives the first two moments of the predictive distribution under several prior beliefs about pricing model implications for predictability. Specifically, Appendix A analyzes the case where investors disregard asset pricing restrictions. Appendix B studies the case where investors fully incorporate those restrictions. Appendix C (D) addresses asset allocation under the first (second) informative prior specification. Under all these scenarios, predictive moments with estimation risk obey analytic expressions.

Denoting the first and second predictive moments by $\mathbb{E}\{r_{T+1}|\mathcal{D}_T\}$ and $\mathbb{V}(r_{T+1}|\mathcal{D}_T)$, where $\mathcal{D}_T$ stands for the data observed up to time T, the optimal portfolio ω solves

$$\max_{\omega}\left(\omega'\mathbb{E}\{r_{T+1}|\mathcal{D}_T\} - \frac{\gamma}{2}\omega'\mathbb{V}(r_{T+1}|\mathcal{D}_T)\omega\right), \tag{31}$$

where γ is the coefficient of relative risk aversion defined with respect to the investor's utility of terminal wealth. In the empirical examination that follows, the optimization in Equation (31) is conducted under three scenarios. In the first, no portfolio constraints are imposed. The resulting portfolio is given by $\omega = \frac{1}{\gamma}\mathbb{V}(r_{T+1} \mid \mathcal{D}_T)^{-1}\mathbb{E}(r_{T+1} \mid \mathcal{D}_T)$. In the second, short positions and leveraging are precluded. In the third, a 50% margin requirement per Regulation T of the Federal Reserve Board is accounted for. Specifically, Regulation T requires a margin of 50% for purchasing securities using a loan given by the broker/dealers, and a margin of 150% for short selling (of which 100% can come from the proceeds of the short sale, with investors depositing the remaining 50%). Brokers are allowed to set the margin even higher. When borrowing and lending rates are equal, the constraint in the presence of ϕ% margin is $\sum_{i=1}^{N} \mid \omega_i \mid <= \frac{100}{\phi}$.

Although Regulation T permits limited short positions, the case of no short selling is especially relevant. As Sharpe (1991) notes, many institutional investors are prohibited from taking short positions either through explicit rules or via the implicit threat of lawsuit for violating fiduciary standards. Institutional investors account for a major fraction of equity holding in domestic exchanges. According to the Federal Reserve Board "Flow of Funds," at the end of September 1999, institutional investors, particularly pension and mutual funds, held 50.2% of outstanding equities in the NYSE.

Next, we present the metric for analyzing the economic significance of prior beliefs in the extent of return predictability explained by pricing models. Let ω_0 and ω_σ denote asset allocations corresponding to two distinct investors. The first has complete faith about the validity of pricing models, that is, $\sigma = 0$. The second has prior mispricing uncertainty $\sigma > 0$.

The case $\sigma \to \infty$ corresponds to complete skepticism about model implications. Let CER_0 and CER_σ denote certainty equivalent returns associated with $\sigma = 0$ and $\sigma > 0$, respectively. Certainty equivalent returns are

$$CER_0 = \omega_0' \mathbb{E}\{r_{T+1} \mid \mathcal{D}_T, \sigma = 0\} - \frac{\gamma}{2} \omega_0' \mathbb{V}\{r_{T+1} \mid \mathcal{D}_T, \sigma = 0\} \omega_0, \quad (32)$$

$$CER_\sigma = \omega_\sigma' \mathbb{E}\{r_{T+1} \mid \mathcal{D}_T, \sigma = 0\} - \frac{\gamma}{2} \omega_\sigma' \mathbb{V}\{r_{T+1} \mid \mathcal{D}_T, \sigma = 0\} \omega_\sigma. \quad (33)$$

The proposed metric is the difference $CER_0 - CER_\sigma$, which reflects the utility loss perceived by a hypothetical investor who has complete confidence that return predictability is entirely explained by asset pricing models, but is forced to hold the allocation of one who recognizes the possibility of model mispricing. Our metric builds on Kandel and Stambaugh (1996). They propose using a single predictive distribution to compute certainty equivalent returns of both portfolios instead of using the portfolio-specific predictive distribution approach of McCulloch and Rossi (1990).

2. Data

The sample contains monthly observations of predictive variables and excess returns on test and benchmark assets spanning 35.5 years from July 1963 to December 1998. The investment universe consists of a one-month Treasury bill, six benchmark assets, and 25 size and book-to-market portfolios identified by a combination of letters and numbers designating increasing values of size (S1, ..., S5) and book-to-market (B1, ..., B5). For instance, S1 (S5) refers to the smallest (largest) 20% in terms of market capitalization. Following Lo and MacKinlay (1990), we also examine an investment universe in which 25 industry portfolios replace the size and book-to-market portfolios. Lo and MacKinlay argue that if the test assets are formed on the basis of characteristics that have been found to be related to average returns, then the hypothesis in favor of the asset pricing model will be rejected too often due to data-mining biases.

The set of benchmark assets includes the market (MKT), size premium (SMB), and value premium (HML) studied by Fama and French (1993). We also examine the restrictions on predictability implied by two extended versions of FF. The first extension accounts for documented momentum. Following Carhart (1997), momentum is captured by the return differential between high- and low-momentum stocks (WML). The second accounts for momentum and for risk exposures related to interest rates and credit conditions. Interest rate risk is captured by the return differential between long- and short-term government bonds (TERM). Credit condition risk is captured by the return differential between low- and high-rated corporate bonds (DEF). Chen, Roll, and Ross (1986) show that a variable similar to DEF is a statistically powerful factor, and TERM has some power as well. In contrast, Shanken and Weinstein (1990) find that another estimate of the risk premium, the sample mean of the factor, is insignificantly negative over the same sample period studied by Chen, Roll, and Ross. TERM and DEF have also been included in the time-series regressions of FF, who, in turn, conclude that risk exposures related to interest rates and credit conditions are too small to explain much variation in the cross section of expected stock returns.

We select five information variables. These are the dividend yield (*Div*), the term spread (*Term*), the default spread (*Def*), the one-month Treasury bill rate (*Tbill*), and the excess return on the equal-weighted NYSE index (*Ret*). Dividend yield is constructed as the total payment of dividends on the value-weighted NYSE portfolio over the recent 12 months divided by the contemporaneous level of the index. *Term* is formed as the yield differential between long- and short-term government bonds. *Def* is formed as the yield differential between Moody's Baa and Aaa bond portfolios.

The sources of data are as follows. Monthly returns on the NYSE index and on the one-month Treasury bill are provided by the Center for Research in Security Prices (CRSP) at the University of Chicago. Yields on Moody's Aaa and Baa bond portfolios are from the *Federal Reserve Bulletin*. Term and default premia are from Ibbotson and Associates. Kenneth French has generously provided the returns on 25 size and book-to-market portfolios, the three FF factors, and returns on the WML portfolio. Chris Kirby has generously provided returns on 25 industry portfolios.

3. Results

3.1 The statistical evidence on asset pricing restrictions

Table 1 summarizes the time-series multivariate predictive regression of excess returns on 25 size and book-to-market portfolios on lagged values of the five information variables: dividend yield, excess return on the equal-weighted NYSE index, Treasury bill rate, default spread, and term spread. Panel A reports intercept and slope coefficients. Panel B displays the corresponding t-statistics. Regression parameters and t-statistics are presented both for the case where return predictability is unrestricted by asset pricing models and for the case where the three-factor FF model explains return predictability. The former suggests that the intercepts and slopes in the predictive regression in Equation (7) are the sample analogs of the values described in Equations (9) and (10). Under FF restrictions,

Table 1
Panel A—Predictive regression coefficients; Panel B—*t*-ratios

Portfolio:	Intercept	Div	Ret	Tbill	Def	Term	Intercept	Div	Ret	Tbill	Def	Term
	Unrestricted regressions						Restricted regressions					
Panel A: Predictive regression coefficients												
S1/B1	0.00	2.26	0.27	−18.11	−1.22	3.57	−0.01	1.96	0.27	−14.77	−0.88	3.19
S1/B2	0.00	1.72	0.24	−13.26	−0.74	2.96	0.00	1.71	0.24	−12.83	−0.74	2.86
S1/B3	0.00	1.42	0.22	−12.16	−0.70	3.15	0.00	1.53	0.21	−11.48	−0.65	2.61
S1/B4	0.01	1.22	0.22	−11.07	−0.62	2.76	0.00	1.42	0.20	−10.57	−0.58	2.43
S1/B5	0.01	1.27	0.26	−11.46	−0.66	2.80	0.00	1.46	0.21	−10.66	−0.58	2.47
S2/B1	0.00	1.78	0.16	−14.59	−0.88	3.33	0.00	1.71	0.22	−14.07	−0.82	3.24
S2/B2	−0.01	1.66	0.16	−12.33	−0.66	3.02	0.00	1.46	0.19	−11.78	−0.66	2.80
S2/B3	0.00	1.16	0.15	−10.27	−0.53	2.71	0.00	1.29	0.17	−10.38	−0.55	2.54
S2/B4	0.00	1.02	0.12	−8.43	−0.45	2.51	0.00	1.10	0.14	−9.09	−0.46	2.35
S2/B5	0.00	1.20	0.14	−10.19	−0.54	2.66	0.00	1.22	0.16	−9.73	−0.49	2.49
S3/B1	0.00	1.46	0.10	−12.65	−0.68	3.28	0.00	1.42	0.16	−12.46	−0.70	3.04
S3/B2	−0.01	1.33	0.10	−10.91	−0.55	3.00	0.00	1.15	0.14	−10.11	−0.54	2.57
S3/B3	0.00	1.00	0.09	−9.03	−0.40	2.54	0.00	0.99	0.12	−8.73	−0.44	2.32
S3/B4	0.01	0.91	0.07	−8.91	−0.44	2.49	0.00	0.84	0.09	−7.68	−0.36	2.16
S3/B5	0.00	1.16	0.11	−7.87	−0.39	1.67	0.00	0.99	0.12	−8.57	−0.40	2.37
S4/B1	0.00	1.17	0.06	−10.58	−0.60	2.85	0.00	1.04	0.10	−10.31	−0.55	2.73
S4/B2	0.00	1.07	0.06	−9.53	−0.43	2.70	0.00	0.86	0.08	−8.78	−0.43	2.49
S4/B3	0.00	0.84	0.03	−9.01	−0.47	2.83	0.00	0.75	0.07	−7.74	−0.35	2.31
S4/B4	0.01	0.83	0.00	−9.75	−0.52	2.85	0.00	0.65	0.06	−6.84	−0.28	2.13
S4/B5	0.00	0.90	0.04	−8.71	−0.42	2.75	0.00	0.80	0.08	−7.76	−0.33	2.35
S5/B1	0.02	0.11	0.02	−6.34	−0.34	2.42	0.01	0.45	0.00	−6.85	−0.31	2.20
S5/B2	0.00	0.49	0.01	−5.78	−0.20	2.04	0.01	0.42	0.00	−6.36	−0.26	2.17
S5/B3	0.00	0.67	−0.03	−5.60	−0.10	1.26	0.01	0.28	−0.02	−5.16	−0.18	1.92
S5/B4	0.01	0.38	−0.06	−6.30	−0.28	2.30	0.01	0.27	−0.01	−4.76	−0.14	1.86
S5/B5	0.01	0.07	−0.02	−5.38	−0.18	2.39	0.01	0.32	0.01	−4.73	−0.12	1.84
Panel B: *t*-ratios												
S1/B1	0.08	3.92	3.91	−5.37	−3.95	2.70	−0.30	3.52	4.00	−4.53	−2.95	2.50
S1/B2	−0.22	3.35	3.85	−4.43	−2.70	2.52	−0.27	3.43	3.94	−4.40	−2.75	2.50
S1/B3	−0.03	3.07	3.82	−4.48	−2.82	2.96	−0.25	3.35	3.86	−4.30	−2.64	2.49
S1/B4	0.41	2.78	4.21	−4.31	−2.61	2.74	−0.25	3.28	3.82	−4.17	−2.50	2.45
S1/B5	0.51	2.78	4.60	−4.28	−2.69	2.66	−0.26	3.21	3.81	−4.01	−2.36	2.37
S2/B1	−0.10	3.15	2.30	−4.43	−2.92	2.58	−0.11	3.14	3.24	−4.42	−2.80	2.60
S2/B2	−0.57	3.47	2.71	−4.42	−2.57	2.76	−0.11	3.13	3.33	−4.32	−2.62	2.62
S2/B3	0.20	2.64	2.72	−4.01	−2.26	2.69	−0.08	3.01	3.23	−4.16	−2.42	2.60
S2/B4	−0.10	2.51	2.37	−3.53	−2.04	2.68	−0.01	2.76	2.88	−3.91	−2.14	2.57
S2/B5	0.20	2.68	2.55	−3.89	−2.24	2.59	−0.06	2.79	3.03	−3.81	−2.07	2.49
S3/B1	−0.21	2.82	1.53	−4.19	−2.44	2.77	0.03	2.84	2.70	−4.27	−2.61	2.66
S3/B2	−0.39	3.10	2.00	−4.35	−2.38	3.05	0.06	2.77	2.67	−4.14	−2.39	2.69
S3/B3	−0.03	2.54	1.85	−3.90	−1.89	2.80	0.09	2.60	2.51	−3.91	−2.13	2.66
S3/B4	0.47	2.45	1.49	−4.10	−2.20	2.92	0.17	2.33	2.15	−3.64	−1.84	2.62
S3/B5	−0.07	2.70	2.03	−3.13	−1.71	1.69	0.10	2.41	2.40	−3.56	−1.80	2.52
S4/B1	0.07	2.55	1.00	−3.95	−2.46	2.71	0.23	2.34	1.83	−3.97	−2.31	2.69
S4/B2	−0.33	2.60	1.28	−3.95	−1.96	2.85	0.30	2.18	1.65	−3.79	−2.01	2.74
S4/B3	0.35	2.16	0.56	−3.95	−2.24	3.16	0.34	2.00	1.47	−3.55	−1.73	2.70
S4/B4	1.01	2.25	0.00	−4.50	−2.62	3.35	0.36	1.83	1.32	−3.30	−1.49	2.62
S4/B5	0.23	2.09	0.75	−3.46	−1.82	2.79	0.27	1.98	1.66	−3.29	−1.50	2.55
S5/B1	1.50	0.28	0.43	−2.83	−1.64	2.75	0.61	1.22	0.01	−3.15	−1.57	2.58
S5/B2	0.09	1.31	0.12	−2.66	−0.99	2.40	0.65	1.17	0.00	−3.06	−1.34	2.66
S5/B3	−0.05	1.95	−0.72	−2.77	−0.52	1.59	0.73	0.85	−0.41	−2.74	−1.01	2.60

Table 1
(continued)

Portfolio:	Intercept	Div	Ret	Tbill	Def	Term	Intercept	Div	Ret	Tbill	Def	Term
	Unrestricted regressions						Restricted regressions					
S5/B4	0.89	1.14	−1.36	−3.19	−1.55	2.97	0.70	0.83	−0.28	−2.54	−0.79	2.53
S5/B5	1.12	0.18	−0.49	−2.43	−0.88	2.76	0.59	0.94	0.11	−2.40	−0.68	2.39

The table summarizes the time-series multivariate regression

$$r_t = \mu_0 + \mu_1 z_{t-1} + v_t,$$

where r_t stands for excess returns on 25 size and book-to-market portfolios and z_{t-1} contains lagged values of dividend yield (Div), return on the equal-weighted NYSE index in excess of the return on the one-month T reasury bill (Ret), one-month T reasury bill rate (Tbill), default spread (Def), and term spread (Term). Panel A exhibits regression intercepts and slopes. Panel B displays t-ratios. The unrestricted estimates are the sample analogs of

$$\mu_0 = \alpha_0 + \beta\lambda_0,$$
$$\mu_1 = \alpha_1 + \beta\lambda_1.$$

The restricted estimates are the sample analogs of

$$\mu_0 = \beta\lambda_0,$$
$$\mu_1 = \beta\lambda_1.$$

The size and book-to-market portfolios are identified by a combination of letters and numbers designating increasing values of size (S1, ..., S5) and book-to-market (B1, ..., B5). For instance, S1 refers to the lowest 20% of the market capitalization, whereas S5 refers to the largest.

the intercepts and slopes in the predictive regression in Equation (7) are the sample analogs of Equations (11) and (12). In particular, β, a 25×3 matrix, denotes loadings on MKT, SMB, and HML, and λ_0 and λ_1, a 3×1 vector and a 3×5 matrix, respectively, denote the intercepts and slopes in the multivariate regression of MKT, SMB, and HML on lagged predictive variables. The standard error associated with the t-statistics under pricing model restrictions is derived in Appendix B. One caveat is in order. Stambaugh (1999) shows a small sample bias in an ordinary least squares (OLS) estimator of a predictive regression slope coefficient. Thus the statistical evidence is merely suggestive.

Table 1 shows ample evidence in favor of in-sample predictability of excess returns, as many of the t-statistics are greater (in absolute value) than two. Ferson and Harvey (1999) find similar evidence. Also, it appears that small stocks are more predictable than large stocks. To illustrate, dividend yield, excess return on the NYSE index, and default spread do not appear to be useful in forecasting returns on the largest size portfolio (S5), having t-statistics less than two. However, these predictors appear to be useful in forecasting excess returns on portfolios belonging to the smaller size groups. Avramov (2002) finds similar cross-sectional dispersion in return predictability in a framework that accounts for uncertainty about the return-forecasting model.

Comparing the restricted and unrestricted regression parameters and t-statistics, we note that the magnitudes are similar and the signs are virtually identical. In particular, whether the pricing restrictions on predictability are imposed or ignored, higher dividend yield, higher excess return on the NYSE index, and higher term spread are accompanied by higher predicted returns. In contrast, a higher Treasury bill rate and higher default spread are followed by lower predicted returns. In addition, explanatory variables that are significant (insignificant) in the unrestricted regression are, in most cases, significant (insignificant) in the regression that conforms to FF.

3.2 The economic content of asset pricing misspecification

The pricing restrictions on return predictability formulated in Equations (11) and (12) are used as a reference point for informative prior beliefs about the intercept and slope parameters in the predictive regression in Equation (7). Prior beliefs are updated by the data, which contain returns on test assets, returns on benchmark assets, and predictive variables. The updated beliefs are then used for asset allocation. Asset allocation thus reflects pricing model restrictions and the sample evidence on predictability. Pricing models are expected to exert stronger influence on asset allocation when the prior confidence in their validity is stronger and when the sample indicates that much of the evidence on return predictability is explained by pricing models. Thus the investor's asset allocation provides a useful metric for assessing the economic significance of model misspecification.

In conducting the optimization in Equation (31), we choose a risk aversion parameter of $\gamma = 10$, the upper bound on relative risk aversion estimates described by Mehra and Prescott (1985). We choose this value to mitigate extreme long and short positions in risky assets and cash. We note that when portfolio holdings are unconstrained, the risk aversion level has no impact on the weights of risky securities in the tangency portfolio. In particular, readers that are comfortable with other levels of γ, say, $\gamma = \tilde{\gamma}$, can multiply the unconstrained investments in risky securities and the corresponding certainty equivalent returns, presented below, by $\frac{10}{\tilde{\gamma}}$. In addition, the current values of the predictive variables (z_T) take on the actual values realized on December 31, 1998, the end-of-sample period. Those values would be the natural choice of real-time investors seeking sensible portfolio decisions when the signal z_T becomes available.

3.2.1 The evidence based on size/book-to-market portfolios. Tables 2 and 3 exhibit conditional mean-variance efficient portfolios as a fraction of the total invested wealth when prior beliefs about the restrictions on predictability implied by asset pricing models range from perfect confidence ($\sigma = 0$) to complete skepticism ($\sigma = \infty$). The intermediate case ($0 < \sigma < \infty$) is based on the first prior specification formulated in

Table 2
Panel A—Unconstrained portfolios; Panel B—Excluding short selling and leveraging; Panel C—Incorporating 50% margin requirements

	$\mathcal{M}_1$			$\mathcal{M}_2$			$\mathcal{M}_3$			NI
	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = \infty$
Panel A: Unconstrained portfolios										
MKT	80.99	−171.24	−346.71	79.81	−190.32	−345.01	122.05	−164.64	−304.69	−769.78
SMB	−158.39	−148.59	−141.18	−143.29	−157.95	−164.44	−166.72	−179.31	−183.93	−500.92
HML	108.27	73.80	57.33	115.95	68.90	46.58	139.81	89.46	68.97	−402.73
WML	0.00	56.73	97.14	35.46	40.60	47.56	37.92	42.46	48.27	65.50
TERM	0.00	4.26	7.29	0.00	4.61	7.28	−93.61	−101.18	−104.19	−190.80
PREM	0.00	−11.09	−18.99	0.00	−11.99	−18.94	−139.73	−143.14	−145.37	−161.12
S1/B1	0.00	−67.82	−116.12	0.00	−73.32	−115.86	0.00	−77.13	−115.33	−150.63
S1/B2	0.00	44.74	76.60	0.00	48.36	76.42	0.00	50.88	76.07	−92.66
S1/B3	0.00	−22.10	−37.85	0.00	−23.90	−37.76	0.00	−25.14	−37.59	−6.50
S1/B4	0.00	52.92	90.61	0.00	57.21	90.40	0.00	60.18	89.99	352.49
S1/B5	0.00	37.06	63.45	0.00	40.06	63.31	0.00	42.14	63.02	352.80
S2/B1	0.00	−12.93	−22.14	0.00	−13.98	−22.09	0.00	−14.71	−21.99	17.00
S2/B2	0.00	26.96	46.16	0.00	29.14	46.05	0.00	30.66	45.84	−210.77
S2/B3	0.00	22.31	38.21	0.00	24.12	38.12	0.00	25.38	37.95	178.49
S2/B4	0.00	15.48	26.50	0.00	16.73	26.44	0.00	17.60	26.32	82.97
S2/B5	0.00	−10.23	−17.51	0.00	−11.06	−17.47	0.00	−11.63	−17.39	4.31
S3/B1	0.00	−39.73	−68.03	0.00	−42.95	−67.88	0.00	−45.19	−67.56	−24.39
S3/B2	0.00	6.63	11.35	0.00	7.17	11.33	0.00	7.54	11.28	−105.10
S3/B3	0.00	−9.86	−16.89	0.00	−10.66	−16.85	0.00	−11.22	−16.77	141.30
S3/B4	0.00	16.94	29.01	0.00	18.32	28.95	0.00	19.27	28.81	113.92
S3/B5	0.00	−5.47	−9.37	0.00	−5.91	−9.35	0.00	−6.22	−9.30	−244.11
S4/B1	0.00	64.67	110.72	0.00	69.91	110.47	0.00	73.54	109.96	−41.93
S4/B2	0.00	−34.07	−58.34	0.00	−36.84	−58.21	0.00	−38.75	−57.94	−21.16
S4/B3	0.00	−25.32	−43.36	0.00	−27.38	−43.26	0.00	−28.80	−43.06	74.66
S4/B4	0.00	20.51	35.11	0.00	22.17	35.03	0.00	23.32	34.87	153.62
S4/B5	0.00	−3.02	−5.17	0.00	−3.26	−5.16	0.00	−3.43	−5.14	−27.72

S5/B1	0.00	93.76	160.55	0.00	101.37	160.18	0.00	106.63	159.45	326.15
S5/B2	0.00	39.12	66.98	0.00	42.29	66.83	0.00	44.49	66.52	−52.43
S5/B3	0.00	24.28	41.57	0.00	26.25	41.47	0.00	27.61	41.28	−150.51
S5/B4	0.00	34.31	58.75	0.00	37.10	58.62	0.00	39.02	58.35	118.55
S5/B5	0.00	−1.11	−1.90	0.00	−1.20	−1.90	0.00	−1.26	−1.89	198.67
CER	1.12	0.58	−0.47	1.19	0.85	0.32	1.67	1.30	0.82	−5.95
SR	0.47	0.39	0.316	0.49	0.43	0.38	0.58	0.54	0.48	0.27
Panel B: Excluding short selling and leveraging										
MKT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SMB	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HML	44.46	23.38	10.83	30.78	33.59	38.11	29.07	31.70	36.41	30.26
WML	9.13	46.31	66.00	32.72	32.03	28.38	33.02	32.59	29.34	16.55
TERM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PREM	4.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1/B1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1/B3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1/B4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1/B5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 2
(continued)

	$\mathcal{M}_1$			$\mathcal{M}_2$			$\mathcal{M}_3$			NI
	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = \infty$
S5/B1	0.00	0.00	0.00	0.00	1.32	14.22	0.00	0.00	10.06	35.64
S5/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S5/B3	10.52	0.00	0.00	0.82	0.00	0.00	0.00	0.00	0.00	0.00
S5/B4	26.11	27.71	22.66	25.68	33.05	19.30	21.37	29.63	24.18	0.00
S5/B5	5.50	2.59	0.52	10.01	0.00	0.00	16.55	6.08	0.00	17.55
CER	0.42	0.29	0.09	0.50	0.50	0.47	0.52	0.51	0.48	0.35
SR	0.31	0.24	0.18	0.36	0.36	0.36	0.36	0.36	0.36	0.27
Panel C: Incorporating 50% margin requirements										
MKT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SMB	−84.80	0.00	0.00	−62.38	0.00	0.00	−57.87	−2.47	0.00	0.00
HML	11.23	0.00	0.00	7.77	0.00	0.00	0.00	0.00	0.00	0.00
WML	0.00	37.95	60.96	21.42	18.22	13.63	16.43	15.13	11.38	0.00
TERM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PREM	0.00	0.00	0.00	0.00	0.00	0.00	−33.19	−29.93	−26.77	−1.14
S1/B1	−13.32	−48.38	−49.80	−18.41	−51.58	−58.68	−18.09	−50.51	−56.00	−64.58
S1/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1/B3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1/B4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S1/B5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B1	−9.86	−8.91	0.00	−8.96	−13.65	−9.83	−1.70	−0.40	0.00	−5.27
S2/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S2/B5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S3/B5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

S4/B1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S4/B5	6.87	9.07	14.73	4.96	8.89	13.07	0.00	0.00	0.00	0.00
S5/B1	0.00	16.46	16.08	0.00	26.14	38.54	0.00	16.35	29.40	73.06
S5/B2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S5/B3	8.99	3.41	0.00	8.77	4.80	0.00	0.73	1.78	0.00	0.00
S5/B4	38.03	53.42	45.06	38.71	54.37	50.95	36.80	54.78	52.02	0.00
S5/B5	26.91	22.40	13.38	28.62	22.36	15.30	35.21	28.65	24.42	55.96
CER	1.01	0.79	0.55	1.04	0.96	0.91	1.12	1.05	1.00	0.74
SR	0.46	0.40	0.34	0.47	0.45	0.44	0.53	0.50	0.48	0.39

The table reports conditional mean-variance portfolios based on scenarios where prior beliefs about the extent of return predictability explained by asset pricing models range from dogmatic beliefs ($\sigma = 0$) to no faith at all ($\sigma = \infty$), and where the risk aversion parameter is 10. Under consideration are conditional versions of the Fama and French (FF) three-factor model ($\mathcal{M}_1$), FF augmented with momentum ($\mathcal{M}_2$), and FF augmented with momentum, a term premium, and a default premium ($\mathcal{M}_3$). The case of $\sigma = \infty$ corresponds to completely disregarding asset pricing implications for return predictability. The universe of equity portfolios contains the market (MKT), size (SMB), and value (HML) premiums, winners-minus-losers (WML) one-year momentum, TERM, PREM, and 25 portfolios sorted on size and book-to-market. Negative asset allocations represent short selling. CER is a certainty equivalent riskfree rate, and SR is a conditional Sharpe ratio. Certainty equivalent rates and Sharpe ratios are computed from the $\sigma = 0$ perspective. Panel A describes unconstrained portfolio holdings. Panel B displays optimal portfolios constrained by no short selling and leveraging. Panel C accounts for 50% margin requirements per Regulation T of the Federal Reserve Board. The current values of the predictive variables (z_T) are those realized on December 31, 1998.

Table 3
The case of industry portfolios

	$\mathcal{M}_1$			$\mathcal{M}_2$			$\mathcal{M}_3$			NI
	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = \infty$
MKT	80.99	48.74	−25.87	79.81	47.90	−24.90	122.05	87.71	13.96	263.10
SMB	−158.39	−127.76	−48.83	−143.29	−123.64	−72.28	−166.72	−145.01	−92.19	−149.30
HML	108.27	139.66	234.60	115.95	144.58	223.46	139.81	169.02	245.05	391.46
WML	0.00	26.96	95.73	35.46	36.29	46.21	37.92	38.31	46.93	71.01
TERM	0.00	5.34	18.97	0.00	5.55	18.93	−93.61	−93.79	−92.58	−70.78
PREM	0.00	−3.67	−13.02	0.00	−3.81	−12.99	−139.73	−139.28	−139.45	−114.41
Aerospace	0.00	3.74	13.30	0.00	3.89	13.27	0.00	4.04	13.21	−122.29
Transportation	0.00	−5.94	−21.10	0.00	−6.18	−21.05	0.00	−6.41	−20.96	−42.91
Banking	0.00	−0.22	−0.78	0.00	−0.23	−0.78	0.00	−0.24	−0.78	−4.29
Building materials	0.00	−3.31	−11.75	0.00	−3.44	−11.72	0.00	−3.57	−11.67	−29.09
Chemicals	0.00	4.66	16.56	0.00	4.85	16.52	0.00	5.03	16.44	−28.58
Construction	0.00	−5.36	−19.02	0.00	−5.57	−18.97	0.00	−5.77	−18.89	−23.57
Entertainment	0.00	−1.26	−4.47	0.00	−1.31	−4.46	0.00	−1.36	−4.44	−44.34
Food & beverages	0.00	11.41	40.53	0.00	11.86	40.44	0.00	12.31	40.25	−89.72
Health care	0.00	18.17	64.53	0.00	18.89	64.38	0.00	19.59	64.09	117.28
Indus. Machinery	0.00	4.76	16.91	0.00	4.95	16.87	0.00	5.13	16.79	118.83
Insurance	0.00	1.01	3.59	0.00	1.05	3.58	0.00	1.09	3.56	−102.15
Investments	0.00	0.83	2.95	0.00	0.86	2.94	0.00	0.90	2.93	−13.71
Metal	0.00	−11.45	−40.68	0.00	−11.91	−40.58	0.00	−12.35	−40.40	4.51
Mining	0.00	0.65	2.31	0.00	0.68	2.30	0.00	0.70	2.29	−17.54
Motor vehicles	0.00	0.47	1.65	0.00	0.48	1.65	0.00	0.50	1.64	13.18
Paper	0.00	1.61	5.72	0.00	1.67	5.71	0.00	1.74	5.68	−39.92
Petroleum	0.00	12.53	44.49	0.00	13.02	44.39	0.00	13.51	44.19	−70.91
Publishing	0.00	2.68	9.52	0.00	2.79	9.50	0.00	2.89	9.46	−94.67
Prof. services	0.00	9.06	32.17	0.00	9.42	32.10	0.00	9.77	31.96	148.90
Retailing	0.00	12.45	44.21	0.00	12.94	44.11	0.00	13.42	43.91	195.30

Semiconductors	0.00	11.65	41.37	0.00	12.11	41.28	0.00	12.56	41.09	77.18
Telecomunications	0.00	5.34	18.95	0.00	5.55	18.90	0.00	5.75	18.82	8.96
Textiles	0.00	−12.42	−44.12	0.00	−12.91	−44.02	0.00	−13.40	−43.82	−77.03
Utilities	0.00	−9.93	−35.26	0.00	−10.32	−35.18	0.00	−10.71	−35.02	−13.34
Wholesaling	0.00	−16.84	−59.80	0.00	−17.50	−59.66	0.00	−18.16	−59.39	−17.85
CER	1.12	0.99	−0.46	1.19	1.12	0.33	1.67	1.60	0.83	−5.95
SR	0.47	0.39	−0.14	0.49	0.41	0.17	0.58	0.51	0.22	0.22

The table reports unconstrained conditional mean-variance portfolios based on scenarios where prior beliefs about the extent of return predictability explained by asset pricing models range from dogmatic beliefs ($\sigma = 0$) to no faith at all ($\sigma = \infty$), and where the risk aversion parameter is 10. Under consideration are conditional versions of the Fama and French (FF) three-factor model ($\mathcal{M}_1$), FF augmented with momentum ($\mathcal{M}_2$), and FF augmented with momentum, a term premium, and a default premium ($\mathcal{M}_3$). The case of $\sigma = \infty$ corresponds to completely disregarding asset pricing implications for return predictability. The universe of equity portfolios contains the market (MKT), size (SMB), and value (HML) premiums, winners-minus-losers (WML) one-year momentum, TERM, PREM, and 25 industry portfolios. Negative asset allocations represent short selling. CER is a certainty equivalent riskfree rate, and SR is a conditional Sharpe ratio. Certainty equivalent rates and Sharpe ratios are computed from the $\sigma = 0$ perspective. The current values of the predictive variables (z_T) are those realized on December 31, 1998.

Equation (20). Specifically, $\sigma = 0.1\%$ implies that prior beliefs about α_1 are concentrated at zero and uncertainty about α_0 is 0.1% per month. Asset allocations are based on the conditional versions of FF ($\mathcal{M}_1$), FF augmented with momentum ($\mathcal{M}_2$), and FF augmented with momentum, a term premium, and a default premium ($\mathcal{M}_3$). The case of $\sigma = \infty$ corresponds to noninformative diffuse priors (**NI**), in which asset pricing restrictions are completely disregarded by investors. Table 2 studies an investment universe consisting of MKT, SMB, HML, WML, TERM, DEF, and 25 portfolios sorted on size and book-to-market. In Table 3, industry portfolios replace the size and book-to-market portfolios.

In Tables 2 and 3, negative asset allocations represent short selling, CER is a certainty equivalent return computed using Equations (32) and (33), and SR is a conditional Sharpe ratio. When the prior uncertainty about model validity is 0%, 0.1%, or 5% per month, CER and SR are computed using the predictive moments $\mathbb{E}\{r_{T+1} \mid \mathcal{D}_T, \sigma = 0, \mathcal{M}_i\}$ and $\mathbb{V}\{r_{T+1} \mid \mathcal{D}_T, \sigma = 0, \mathcal{M}_i\}$, where $\mathcal{M}_\mathrm{i}$ is the corresponding asset pricing specification ($\mathcal{M}_1$, $\mathcal{M}_2$, or $\mathcal{M}_3$). When $\sigma = \infty$, CER and SR are computed using the moments $\mathbb{E}\{r_{T+1} \mid \mathcal{D}_T, \sigma = 0, \mathcal{M}_3\}$ and $\mathbb{V}\{r_{T+1} \mid \mathcal{D}_\mathrm{T}, \sigma = 0, \mathcal{M}_3\}$. That is, certainty equivalents and Sharpe ratios are computed from the $\sigma = 0$ perspective. There are three panels in Table 2. Panel A presents unconstrained portfolios. Panel B excludes short selling and leveraging. Panel C incorporates a 50% margin per Regulation T of the Federal Reserve Board. Table 3 displays unconstrained portfolio holdings only.

Several interesting results emerge from Table 2. Notably, unconstrained asset allocation is considerably sensitive to prior confidence about the accuracy of pricing model implications for predictive regressions. With complete confidence, the fraction of wealth invested in the market portfolio is 80.99% ($\mathcal{M}_1$), 79.81% ($\mathcal{M}_2$), and 122.05% ($\mathcal{M}_3$). When the prior uncertainty about the fixed component of asset mispricing is $\sigma = 0.1\%$ per month, the corresponding quantities are -171.24%, -190.32%, and -164.64%. In addition, we note that if asset pricing models explain return predictability, the unconstrained asset allocation must involve the benchmark assets only. Some combination of the benchmarks delivers the highest ex ante conditional Sharpe ratio. In particular, if the FF model explains return predictability, the investment in risky securities other than MKT, SMB, and HML must be zero. Table 2 shows that with even a small degree of mispricing ($\sigma = 0.1\%$), there are substantial deviations from the model implications. Long and short positions in nonbenchmark assets considerably depart from zero, ranging between -67.82% (S1/B1) and 93.76% (S5/B1).

We map asset allocations onto certainty equivalent returns and compute a utility loss. That loss is perceived by an investor, who believes asset pricing restrictions on return predictability are perfectly valid but is forced to hold the asset allocation of one whose prior mispricing uncertainty is

$\sigma = 0.1\%$. The monthly utility loss appears to be economically meaningful. It is 0.54% [1.12% − 0.58%] under $\mathcal{M}_1$, 0.34% [1.19% − 0.85%] under $\mathcal{M}_2$, and 0.37% [1.67% − 1.30%] under $\mathcal{M}_3$. The large loss underscores the economic significance of prior beliefs in the extent of predictability explained by asset pricing models. The case where an investor with $\sigma = 0$ is forced to hold asset allocation based on prior beliefs that reflect the recognition of uncertainty about time-varying asset mispricing as well, that is, $\alpha_1 \neq 0$, is analyzed in Section 3.2.4.

As described earlier, the weight on the market portfolio dramatically changes when just a modest amount of model mispricing ($\sigma = 0.1\%$) is allowed for in the prior. How can this change be explained? Under $\sigma = 0\%$, asset allocation is based on pricing model restrictions, whereas under $\sigma = 0.1\%$, asset allocation is impacted by both the model and the data. Observe that model-based asset allocation differs greatly from its data-based counterpart, that is, the case of $\sigma = \infty$. The certainty equivalent loss corresponding to these polar scenarios exceeds 7% per month for all models examined. Under FF restrictions, for example, the utility loss is 1.12% − (−5.95%) = 7.07%.

Interestingly, asset pricing restrictions exert substantial influence on asset allocation under $\sigma = 5\%$, let alone under $\sigma = 0.1\%$. Consider an investor with mispricing uncertainty of 5%. If such an investor is forced to hold asset allocation based on the model alone, she faces a monthly utility loss equal to 1.59% [1.12% − (−0.47)] under $\mathcal{M}_1$, 0.87% under $\mathcal{M}_2$, and 0.85% under $\mathcal{M}_3$. Instead, if the same investor is forced to hold asset allocation based on the data alone, the corresponding losses are 5.48% [−0.47 − (−5.95)], 6.27%, and 6.77%—substantially larger. Similar utility losses are perceived also when certainty equivalents are computed from the $\sigma = \infty$ perspective. The evidence thus suggests that asset allocation based on $\sigma = 5\%$ per month is much closer to that implied by the model alone than to that implied by the data alone.

It also emerges from Table 2 that data-based optimal portfolios exhibit extreme long and short positions unlikely to be adopted in reality by fund managers. For example, a hypothetical investor takes an unrealistic short position (769.78%) in the market portfolio. In addition, the investments in size/book-to-market portfolios are rather extreme, ranging between −244.11% (S3/B5) and 352.80% (S1/B5). On the other hand, optimal portfolios based on the model and the data combined seem relatively more sensible. Focusing on $\mathcal{M}_1$ and $\sigma = 0.1\%$, the largest long and short positions in size/book-to-market portfolios are 93.76% (S5/B1) and −67.82% (S1/B1).

Indeed, the idea of combining model and data to improve asset allocation goes far back. It has been known that data-based portfolios behave badly. For example, Jobson, Korkie, and Ratti (1979), Best and Grauer (1991), and Green and Hollifield (1992) discuss the extreme sensitivity of mean variance portfolios to variation in the mean and the covariance matrix. In response, Black and Litterman (1992) advocate using Black's (1972) version of the capital asset pricing model (CAPM) as a benchmark for expected return. Recently Pastor and Stambaugh (2000) and Pastor (2000) have proposed a Bayesian framework for combining a model and data in deriving asset allocation. However, relative to those important innovations, our framework allows risk premiums and model mispricing to vary with the business cycle. We show below that incorporating such time variation substantially improves ex post performance of optimal portfolios.

Last, we demonstrate that model uncertainty, or uncertainty about model mispricing and benchmark assets, is important. Consider an investor who must allocate funds across cash and stocks. To derive expected returns, she can use any of the specifications. If the pair $\mathcal{M}_1$ and $\sigma = 5\%$ is chosen, the monthly certainty equivalent is -0.47%. Instead, if the pair $\mathcal{M}_3$ and $\sigma = 0$ is preferred, the certainty equivalent is 1.67%. This suggests that asset allocation is considerably sensitive to the model that generates inputs for asset allocation. To incorporate model uncertainty one can use bayesian model averaging, first computing posterior probabilities for each model under consideration and then using the probabilities to average across models. Avramov (2002) accounts for model uncertainty in a predictive regression framework that disregards pricing model implications. Posterior probabilities have also been derived by Shanken (1987), Harvey and Zhou (1990), and McCulloch and Rossi (1991) for testing pricing models. In these studies, however, returns are assumed unpredictable. Recently Avramov and Chao (2003) derive posterior probabilities in a conditional asset pricing framework. Their analysis could be a key step toward incorporating uncertainty about common factors and lagged variables in a portfolio choice context.

3.2.2 Imposing portfolio constraints. We now analyze asset allocation in the presence of portfolio constraints. Observe from panels B and C of Table 2 that optimal portfolios are not well diversified, regardless of initial beliefs in model validity. The notion is that there are many corner solutions, in which most of the investments in risky securities are zero. Thus it is not surprising that imposing portfolio constraints considerably reduces, yet does not eliminate, the discrepancies between the various specifications. Focusing on the $\mathcal{M}_3$ specification and no short selling, the difference in certainty equivalent returns corresponding to $\sigma = 0$ and $\sigma = \infty$ is only 0.17% per month. Discrepancies between the views are considerably larger when short positions are allowed subject to a 50% margin per Regulation T of the Federal Reserve Board. In that case, the certainty equivalent loss is 0.38% per month. Below, we examine the impact of portfolio constraints on ex post performance of asset allocation under various beliefs regarding pricing model validity.

3.2.3 The case of industry portfolios. Table 3 exhibits portfolio decisions for an investment universe that consists of 25 industry portfolios and the six benchmark assets noted earlier. Qualitatively, the evidence based on the industry portfolios does not differ that much from that based on size and book-to-market portfolios. Thus we briefly describe unconstrained portfolio holdings only. As in Pastor (2000), entertaining even mild prior mispricing uncertainty yields asset allocations that depart considerably from the model implications. Nonetheless, the pricing restrictions still have substantial influence on portfolio selection. Consider an investor with mispricing uncertainty of $\sigma = 5\%$. If this investor is forced to hold the FF-based asset allocation, the resulting certainty equivalent loss is 1.58% [1.12% $-$ ($-$0.46%)] under $\mathcal{M}_1$. If, instead, the investor is forced to hold data-based asset allocation, the loss is much larger, given by 5.49% [5.95%–0.46%].

3.2.4 Informative prior beliefs: the second specification. We investigate the economic content of changes in risk premiums and model mispricing using the second prior specification noted earlier, which recognizes the different economic roles of the fixed versus time-varying asset mispricing components. That specification establishes an intuitive metric for understanding the roles played by changes in risk premiums and model mispricing. One could simply ask: How large does the hypothetical sample need to be to call into question the sample evidence about time-varying risk premiums and model mispricing? To pursue the analysis we derive asset allocation on the basis of the following hypothetical sample lengths: $T_0 = T$, $T_0 = 2 \times T$, $T_0 = 5 \times T$, $T_0 = 10 \times T$, and $T_0 = \infty$. Restricted (RE) and unrestricted (UN) optimal portfolios are reported in Table 4. The unrestricted portfolios are based on the prior formulated in Equations (25)–(28). The restricted analogs are based on a similar prior, with the exception that α_0 and α_1 are concentrated at zero. Appendix D derives predictive moments for both restricted and unrestricted asset allocations.

We first assess the sample evidence about variation in risk premiums. Asset allocation derived when FF holds exactly and when risk premiums can (cannot) vary over time is described in the first column of Table 2 (last column of Table 4). When risk premiums can vary, investments in benchmark assets are 80.99% (MKT), $-$158.39% (SMB), and 108.27% (HML), and the certainty equivalent rate, computed based on predictive moments conforming to FF, is 1.12%. When risk premiums are fixed, the corresponding investments are 47.65% (MKT), 4.94% (SMB), and 92.26% (HML), and the certainty equivalent return is only 0.32%. Returns on benchmark assets appear to be predictable. Nonetheless, entertaining a hypothetical sample of length $T_0 = T$ weighted against predictability calls into question the sample evidence in favor of predictability in benchmark asset returns. Specifically, under FF and $T = T_0$, asset allocation is similar

Table 4
Asset allocation under time-varying model mispricing

	$T_0 = T$		$T_0 = 2 \times T$		$T_0 = 5 \times T$		$T_0 = 10 \times T$		
	RE	UN	RE	UN	RE	UN	RE	UN	$T_0 = \infty$
MKT	63.33	517.15	57.92	327.82	52.70	179.20	50.38	117.49	47.65
SMB	−66.24	−63.87	−41.21	−48.63	−17.68	−23.74	−7.32	−11.01	4.94
HML	98.69	−72.13	96.36	−25.86	94.24	27.79	93.33	53.97	92.26
WML	0.00	6.71	0.00	4.17	0.00	2.02	0.00	1.10	0.00
TERM	0.00	−112.18	0.00	−73.97	0.00	−37.10	0.00	−20.35	0.00
PREM	0.00	−66.74	0.00	−44.63	0.00	−22.54	0.00	−12.38	0.00
S1/B1	0.00	−90.15	0.00	−57.65	0.00	−28.42	0.00	−15.52	0.00
S1/B2	0.00	−49.83	0.00	−34.31	0.00	−17.58	0.00	−9.69	0.00
S1/B3	0.00	−5.05	0.00	−2.40	0.00	−0.95	0.00	−0.50	0.00
S1/B4	0.00	166.40	0.00	109.57	0.00	54.82	0.00	30.03	0.00
S1/B5	0.00	105.49	0.00	71.95	0.00	36.61	0.00	20.12	0.00
S2/B1	0.00	5.61	0.00	4.19	0.00	2.24	0.00	1.25	0.00
S2/B2	0.00	−119.03	0.00	−79.07	0.00	−39.77	0.00	−21.82	0.00
S2/B3	0.00	75.07	0.00	50.07	0.00	25.23	0.00	13.84	0.00
S2/B4	0.00	−6.21	0.00	−2.18	0.00	−0.57	0.00	−0.25	0.00
S2/B5	0.00	−24.91	0.00	−14.78	0.00	−6.99	0.00	−3.78	0.00
S3/B1	0.00	−25.59	0.00	−15.68	0.00	−7.57	0.00	−4.12	0.00
S3/B2	0.00	−63.91	0.00	−42.28	0.00	−21.25	0.00	−11.67	0.00
S3/B3	0.00	37.93	0.00	26.45	0.00	13.58	0.00	7.48	0.00
S3/B4	0.00	14.05	0.00	10.66	0.00	5.67	0.00	3.14	0.00
S3/B5	0.00	−134.21	0.00	−88.70	0.00	−44.46	0.00	−24.36	0.00
S4/B1	0.00	−68.18	0.00	−44.71	0.00	−22.32	0.00	−12.22	0.00
S4/B2	0.00	−25.45	0.00	−16.20	0.00	−8.00	0.00	−4.38	0.00
S4/B3	0.00	49.21	0.00	33.26	0.00	16.89	0.00	9.29	0.00
S4/B4	0.00	67.04	0.00	45.45	0.00	23.09	0.00	12.70	0.00
S4/B5	0.00	−17.19	0.00	−10.95	0.00	−5.38	0.00	−2.94	0.00
S5/B1	0.00	−37.09	0.00	−19.79	0.00	−8.76	0.00	−4.66	0.00
S5/B2	0.00	−124.56	0.00	−80.59	0.00	−40.00	0.00	−21.88	0.00
S5/B3	0.00	−169.31	0.00	−109.79	0.00	−54.56	0.00	−29.84	0.00
S5/B4	0.00	−27.94	0.00	−15.44	0.00	−6.98	0.00	−3.73	0.00
S5/B5	0.00	72.18	0.00	49.22	0.00	25.07	0.00	13.78	0.00
CER	0.49	−1.41	0.39	−0.41	0.34	0.14	0.32	0.26	0.32

The table reports unconstrained mean-variance efficient portfolios under several scenarios in which a hypothetical investor observes a prior sample of length T_0 weighted against predictability and asset mispricing, and where the risk aversion parameter is 10. The universe of equity portfolios contains the market (MKT), size (SMB), and value (HML) premiums, winners-minus-losers (WML) one-year momentum, TERM, PREM, and 25 portfolios sorted on size and book-to-market. Negative asset allocations represent short selling. CER is a certainty equivalent riskfree rate. All certainty equivalent rates are computed from the $\sigma = 0$ perspective. The current values of the predictive variables (z_T) are those realized on December 31, 1998.

to that obtained when risk premiums are constrained to be fixed ($T_0 = \infty$), in that the difference in certainty equivalents is only 0.07% [0.39% − 0.32%].

We next assess the utility loss perceived by an investor who believes the restrictions on predictive regressions implied by FF hold exactly, but is forced to allocate funds accounting for model mispricing. Table 4 shows that such a utility loss is 1.90% [0.49 − (−1.41)] under $T_0 = T$, 0.80% under $T_0 = 2 \times T$, 0.20% under $T_0 = 5 \times T$, and 0.06% under $T_0 = 10 \times T$. In essence, when the prior sample size is equal to T or $2 \times T$, asset allocation considerably departs from model implications. This departure declines substantially, yet remains economically large when $T_0 = 5 \times T$, and virtually disappears when $T_0 = 10 \times T$. Recall that under the second informative prior specification, the weights on the model and the data can be approximated by $\frac{T_0}{T_0+T}$ and $\frac{T}{T_0+T}$. Such weights are about 80% and 20% under $T_0 = 5 \times T$, suggesting that even when the underlying model accounts for 80% of the overall weight, the resulting asset allocation greatly differs from that dictated by the model.

3.3 Ex post out-of-sample outlook

This section takes on an ex post out-of-sample outlook to approach two questions:

1. Does the shrinkage approach to asset allocation improve performance when the possibility of return predictability is recognized?
2. Do conditional asset pricing models outperform their unconditional counterparts that exclude return predictability?

The first question is motivated by earlier evidence that shrinking the sample mean return toward a grand mean [see Jobson, Korkie, and Ratti (1979)] or values implied by asset pricing models [see Pastor (2000)] improves asset allocation. Whether shrinkage helps in the presence of time-varying expected return is an open issue. In asking the second question, we aim to assess the gain from using forecasting variables in asset pricing models. This is a natural question, given the fact that we have documented in-sample predictability and have then analyzed pricing model implications for predictability. Nonetheless, the fact remains that some academics and practitioners believe that the documented in-sample predictability is not evident out of sample or in real time.

To approach these two questions, we implement a rolling scheme. (A recursive scheme yields similar results.) Specifically, we fix the estimation window size and drop distant observations as recent ones are added. Asset allocation is first computed using data from 1 to P, and is then rebalanced using data from 2 to $P + 1, \ldots,$ and it is finally rebalanced using data from $T - P$ to $T - 1$. This produces a time series of $T - P$ "ex post" excess returns for various dynamic trading strategies that differ with respect to the prior confidence about model validity. To illustrate, let $\omega_{\sigma,t}$ be the tangency portfolio at time t under mispricing uncertainty σ, and let r_{t+1} denote the excess return on investable securities realized at time $t + 1$. The realized excess return on the selected portfolio is $r_{\sigma,t+1} = \omega'_{\sigma,t} r_{t+1}$. We then compute a Sharpe ratio by dividing the average value of the $T - P$ realized returns by the standard deviation.

Table 5 summarizes ex post out-of-sample performance. The first four columns display Sharpe ratios when prior beliefs about the conditional version of FF are expressed by $\sigma = 0$, $\sigma = 0.1\%$, $\sigma = 5\%$, and $\sigma = \infty$.

Table 5
Ex post Sharpe ratios

	With predictability				No predictability				Second prior	
	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = \infty$	$\sigma = 0$	$\sigma = 0.1\%$	$\sigma = 5\%$	$\sigma = \infty$	$T_0 = T$	$T_0 = 2 \times T$
Size B/M portfolios + FF + momentum + DEF + TERM										
Unconstrained portfolios										
$P = 108$	0.28	0.33	**0.38**	0.29	0.16	0.27	0.31	0.30	0.26	0.26
$P = 216$	0.33	0.47	0.54	0.62	0.28	0.48	0.47	0.47	**0.63**	0.63
$P = 324$	0.33	0.46	0.50	0.63	0.32	0.52	0.50	0.50	**0.65**	0.65
50% margin										
$P = 108$	0.36	0.38	**0.42**	0.40	0.22	0.23	0.34	0.33	0.41	0.41
$P = 216$	0.34	0.40	**0.48**	0.47	0.31	0.40	0.45	0.45	0.47	0.47
$P = 324$	0.28	0.41	0.46	**0.49**	0.29	0.41	0.45	0.45	0.42	0.40
No short selling										
$P = 108$	0.32	0.33	0.34	0.32	0.18	0.25	0.36	0.36	0.38	**0.39**
$P = 216$	0.34	0.40	0.43	0.44	0.28	0.46	0.51	0.51	0.59	**0.60**
$P = 324$	0.30	0.40	0.42	0.44	0.32	0.53	0.56	0.56	0.62	**0.65**
Industry portfolios + FF + momentum + DEF + TERM										
Unconstrained portfolios										
$P = 108$	0.28	0.30	**0.35**	0.21	0.16	0.20	0.29	0.28	0.19	0.19
$P = 216$	0.33	0.37	**0.46**	0.35	0.28	0.36	0.41	0.41	0.37	0.39
$P = 324$	0.33	0.38	0.47	0.47	0.32	0.44	0.46	0.46	0.49	**0.51**
50% margin										
$P = 108$	0.35	0.36	**0.38**	0.34	0.18	0.18	0.29	0.29	0.36	0.37
$P = 216$	0.32	0.33	0.43	0.45	0.27	0.28	0.40	0.40	**0.46**	0.44
$P = 324$	0.26	0.32	0.43	**0.50**	0.24	0.28	0.34	0.34	0.43	0.41
No short selling										
$P = 108$	0.34	0.34	**0.35**	0.28	0.18	0.20	0.30	0.30	0.30	0.29
$P = 216$	0.31	0.33	0.38	0.35	0.27	0.34	0.45	0.45	0.44	**0.46**
$P = 324$	0.28	0.32	0.39	0.39	0.32	0.41	0.45	0.45	**0.50**	0.48

The table reports ex post out-of-sample Sharpe ratios generated by dynamic trading strategies corresponding to various degrees of prior belief about the validity of conditional (with predictability) and unconditional (no predictability) versions of FF. Boldface figures denote the highest Sharpe ratios within a row. Trading strategies are based on a rolling scheme. In particular, asset allocation is first computed using data from 1 to P, and is then rebalanced using data from 2 to $P + 1, \ldots,$ and it is finally rebalanced using data from $T - P$ to $T - 1$. This produces $T - P$ "ex post" excess returns for various dynamic strategies that differ with respect to the prior confidence about model validity. Sharpe ratios are obtained by dividing the average value of the $T - P$ realized returns by the standard deviation.

The next four columns correspond to the same beliefs in model validity, but the beliefs pertain to the version of FF that does not use conditioning information, as in Pastor (2000) and Pastor and Stambaugh (2000). The last two columns describe Sharpe ratios perceived by an investor who observes a hypothetical sample of lengths $T_0 = T$ and $T_0 = T \times 2$ as formulated in Equations (25)–(28). Unconditional Sharpe ratios are derived for unconstrained and constrained portfolio holdings (50% margin requirement and no short positions in equities and cash), for the two sets of investable securities, and for three cutoff points $P = 108$, $P = 216$, and $P = 324$. Boldface figures display the highest Sharpe ratio per scenario.

Focusing on the conditional case, we demonstrate that complete confidence in asset pricing restrictions on predictive regressions yields the lowest Sharpe ratios for 14 of 18 cases. In four other cases, the lowest Sharpe ratios are attributed to data-based optimal portfolios. To illustrate, consider the first three rows of Table 5. When $\sigma = 0$, Sharpe ratios are 0.28, 0.33, and 0.33. The next lowest Sharpe ratios are 0.29, 0.47, and 0.46. Strikingly, intermediate views about model mispricing, entertained either via specifying $0 < \sigma < \infty$ or by considering a prior sample of length T_0 weighted against asset mispricing and return predictability, deliver substantially larger Sharpe ratios in all 18 cases. For example, focusing on the first three rows of Table 5, the largest Sharpe ratios are 0.38, 0.63, and 0.65. The evidence thus suggests that shrinking the predictive moments toward values implied by asset pricing models considerably improves ex post performance even when expected return could vary over time.

In about half the cases examined, the second prior specification delivers the largest Sharpe ratios. For example, when the investment universe consists of 25 size/book-to-market portfolios and the benchmark assets MKT, SMB, HML, WML, TERM, and DEF, and when portfolio holdings are unconstrained, the largest Sharpe ratios based on the first four columns (last two columns) of Table 5 are 0.38 (0.26), 0.62 (0.63), and 0.63 (0.65). These findings imply that ex post performance can be improved when a hypothetical investor recognizes the possibility of time-varying model misspecification.

We now assess the economic gain from using conditioning information in forming optimal portfolios by comparing Sharpe ratios based on conditional versus unconditional pricing models. We first compare Sharpe ratios under scenarios where conditional and unconditional versions of FF hold exactly. It is evident from Table 5 that in 15 of 18 cases Sharpe ratios are larger in the presence of conditioning information. For example, under 50% margin requirements, Sharpe ratios are 0.36, 0.34, and 0.28 in the presence of predictability. The corresponding no-predictability quantities are 0.22, 0.31, and 0.29. Strikingly, under *all* cases studied, substantially larger Sharpe ratios are generated by optimal portfolios based on conditional models. Thus there is a remarkable economic gain to modeling asset mispricing, risk premiums, or both as linear functions of conditioning information as described in Equations (2) and (4).

In sum, the shrinkage approach to asset allocation greatly improves ex post performance under expected return variation. In addition, optimal portfolios based on conditional pricing models substantially outperform their unconditional counterparts that exclude predictability. These performance-based findings are robust to the investment universe chosen and inclusion of portfolio constraints. The gain of exploiting predictability is obtained under a wide range of beliefs about model pricing abilities.

4. Conclusion

This article studies pricing model implications for stock return predictability from an investment perspective. The analysis yields several insights about the influence of conditional pricing models and the sample evidence about predictability on financial decision making. For one, asset allocation is extremely sensitive to the imposition of pricing model restrictions on predictive regressions. Indeed, an investor who believes conditional asset pricing models are perfectly valid but is forced to allocate funds disregarding model implications faces an enormous utility loss. In fact, asset allocation departs considerably from model implications, even when the prior allows only a small degree of model mispricing. Nonetheless, when the prior reflects a fair amount of model mispricing, asset allocation could still be much closer to that implied by the model alone than to that implied by the data alone. This suggests that pricing restrictions on predictive regressions could have important implications for investment decisions even when the possibility of model mispricing is recognized. Ex post out-of-sample analysis shows that (i) optimal portfolios based on a pricing model and the sample evidence on predictability combined outperform allocations dictated by either complete dogma or complete skepticism about model implications for predictability; and (ii) under a wide range of prior beliefs about model pricing abilities, asset allocations based on conditional models substantially outperform their unconditional counterparts that exclude predictability. These performance-based findings suggest that expected returns do vary over time and that the shrinkage approach to asset allocation helps ex post performance even under expected return variation.

Our work suggests some avenues for future research. First, following Merton (1973) and Fama (1996), one can study asset allocation patterns of intertemporal CAPM (ICAPM) optimizing investors as opposed to mean variance investors. In addition, this article studies pricing restrictions on return predictability when factors are portfolio based. Several asset pricing innovations describe nontraded factors as having a pervasive effect on portfolio consumption decisions. Major examples include the consumption-based CAPM of Breeden (1979) and the discrete time ICAPM of Campbell (1996). Future work could investigate the economic significance of the restrictions on stock return predictability that these asset pricing models imply.

Furthermore, our analysis deals with expected utility maximizing investors. Kahneman and Tversky (1979) argue that individuals systematically violate the axioms of expected utility theory. Thus they advocate defining utility over gains and losses relative to a reference point rather than over the level of wealth, which is a focal point in expected utility theories. This so-called prospect theory has motivated an alternative formulation of security price dynamics relative to traditional expected utility theories. Recently Barberis, Huang, and Santos (2001) find that prospect theory can resolve many of the so-called financial market anomalies documented in earlier studies. It would be of interest to consider prior beliefs about the dynamics of returns implied by prospect theory. One can then examine how the sample updates such beliefs.

An important contribution of this article is to provide evidence that conditional pricing models outperform their unconditional counterparts. Indeed, it seems worthwhile to study the economic value of return predictability in a framework that accommodates single investable stocks rather than portfolios, allowing stocks to enter and leave the sample periodically. In that framework, one could examine whether modeling beta as a linear function of predictive variables, as suggested by Shanken (1990), could improve ex post performance relative to models where beta is fixed. Beta could also be modeled as a linear function of security-specific characteristics, such as size and book-to-market, which are potentially correlated with expected return.

Finally, this article develops a conceptual framework for analyzing time-varying expected returns when prior beliefs about parameters in a predictive regression are informative. One can apply this framework to investigate topics, such as event studies and fund performance, in which benchmark assets are used to risk-adjust returns.

Appendix A: The Case of Unrestricted Predictive Regressions

First, we derive the likelihood function. It is useful to factor that function as

$$p(R, F, Z \mid \Theta_U, z_0) = p(R \mid F, Z, \Theta_U, z_0) P(F \mid Z, \Theta_U, z_0) p(Z \mid \Theta_U, z_0). \tag{34}$$

The first and second densities on the right-hand size are easily obtained using the processes of Equations (16) and (17). To work out the third density, it is useful to partition the variance-covariance matrix Σ as

$$\Sigma = \begin{bmatrix} \Sigma_{RR} & \Sigma_{RF} & \Sigma_{RZ} \\ \Sigma_{FR} & \Sigma_{FF} & \Sigma_{FZ} \\ \Sigma_{ZR} & \Sigma_{ZF} & \Sigma_{ZZ} \end{bmatrix}. \tag{35}$$

The disturbances in Equations (16) and (17) are orthogonal. Thus $\Sigma_{RF} = \Sigma_{FR} = 0$. Σ_{RZ} (Σ_{FZ}) is the conditional covariance between excess stock returns (factors) and contemporaneous values of the predictors, and Σ_{ZZ} is the conditional variance-covariance matrix of the predictors. In addition, it is useful to reparameterize the vector autoregression specification for the predictive variables. Specifically, let $W_Z = [X, U_F, U_R]$, let $\Phi = [A'_Z, \Sigma_{ZF}\Sigma_{FF}^{-1}, \Sigma_{ZR}\Sigma_{RR}^{-1}]'$, and let $V_Z = U_Z - U_F\Sigma_{FF}^{-1}\Sigma_{FZ} - U_R\Sigma_{RR}^{-1}\Sigma_{RZ}$. The vector autoregression in Equation (5) can be rewritten as $Z = W_Z\Phi + V_Z$. Note that the T rows in V_Z are orthogonal to the corresponding ones in U_R and U_F, and that $\text{vec}(V_Z) \sim N(0, \Sigma_{ZZ.R.F} \otimes I_T)$, where $\Sigma_{ZZ.R.F} = \Sigma_{ZZ} - \Sigma_{ZF}\Sigma_{FF}^{-1}\Sigma_{FZ} - \Sigma_{ZR}\Sigma_{RR}^{-1}\Sigma_{RZ}$. Now, let $\mathcal{M}_{\mathrm{U}}$ represent the unrestricted specification, let $\tilde{W}_Z = [X, F, R]$, let $Q_W = I_T - W(W'W)^{-1}W'$, let $Q_X = I_T - X(X'X)^{-1}X'$, and let $Q_Z = I_T - \tilde{W}_Z(\tilde{W}'_Z\tilde{W}_Z)^{-1}\tilde{W}'_Z$. The probability distribution of $\mathcal{D}_{\mathrm{T}}$

conditioned on the set of parameters and $\mathcal{M}_{\mathrm{U}}$ is

$$\begin{aligned}\mathcal{L} \propto |\Sigma_{RR}|^{-\frac{T}{2}}&\exp\{-\tfrac{1}{2}\mathrm{tr}[\Sigma_{RR}^{-1}(R'Q_W R+(\Gamma-\hat{\Gamma})'W'W(\Gamma-\hat{\Gamma}))]\}\\ \times|\Sigma_{FF}|^{-\frac{T}{2}}&\exp\{-\tfrac{1}{2}\mathrm{tr}[\Sigma_{FF}^{-1}(F'Q_X F+(A_F-\hat{A}_F)'X'X(A_F-\hat{A}_F)]\}\\ \times|\Sigma_{ZZ.R.F}|^{-\frac{T}{2}}&\exp\{-\tfrac{1}{2}\mathrm{tr}[\Sigma_{ZZ.R.F}^{-1}(Z'Q_Z Z+(\Phi-\hat{\Phi})'W_Z'W_Z(\Phi-\hat{\Phi}))]\},\end{aligned} \tag{36}$$

where $\hat{\Gamma}=(W'W)^{-1}W'R, \hat{A}_F=(X'X)^{-1}X'F$, and $\hat{\Phi}=(W_Z'W_Z)^{-1}W_Z'Z$. Posterior densities are obtained by combining the likelihood of Equation (36) with the prior of Equation (19) and using the relation $|\Sigma| = |\Sigma_{RR}||\Sigma_{FF}||\Sigma_{ZZ.R.F}|$. The posterior densities are

$$\mathrm{vec}(\Gamma)|\Sigma_{RR},\mathcal{D}_T \sim N[\mathrm{vec}(\hat{\Gamma}),\Sigma_{RR}\otimes(W'W)^{-1}], \tag{37}$$

$$\mathrm{vec}(A_F)|\Sigma_{FF},\mathcal{D}_T \sim N[\mathrm{vec}(\hat{A}_F),\Sigma_{FF}\otimes(X'X)^{-1}], \tag{38}$$

$$\Sigma_{RR}|\mathcal{D}_T \sim IW(R'Q_W R, T-1), \tag{39}$$

$$\Sigma_{FF}|\mathcal{D}_T \sim IW(F'Q_X F, T+N-1). \tag{40}$$

We show below that the mean and variance of predicted future returns are

$$\mathbb{E}\{r_{T+1}|\mathcal{D}_T,\mathcal{M}_U\}=\hat{\alpha}_0+\hat{\alpha}_1 z_T+\hat{\beta}[\bar{f}+\hat{V}_{fz}\hat{V}_z^{-1}(z_T-\bar{z})], \tag{41}$$

$$\mathbb{V}\{r_{T+1}|\mathcal{D}_T,\mathcal{M}_U\}=\delta_1\hat{\Sigma}_{RR}+(1+\delta_2)\hat{\beta}\hat{\Sigma}_{FF}\hat{\beta}', \tag{42}$$

where $\hat{\alpha}_0$, $\hat{\alpha}_1$, and $\hat{\beta}$ are the corresponding partitions of $\hat{\Gamma}'$,

$$\hat{V}_{fz}=\frac{1}{T}\sum_{t=0}^{T-1}(f_{t+1}-\bar{f})(z_t-\bar{z})', \tag{43}$$

$$\hat{\Sigma}_{RR}=\frac{1}{T-N-2}R'Q_W R, \tag{44}$$

$$\hat{\Sigma}_{FF}=\frac{1}{T+N-K-2}F'Q_X F, \tag{45}$$

$$\delta_1=1+(1+\delta_2)\mathrm{tr}(\hat{\Sigma}_{FF}W^{22})+\mathrm{tr}[(\hat{A}_F^{*\prime}x_T x_T'\hat{A}_F^*)(W'W)^{-1}], \tag{46}$$

$$\delta_2=\frac{1}{T}[1+(\bar{z}-z_T)'\hat{V}_z^{-1}(\bar{z}-z_T)], \tag{47}$$

W^{22} is the lower $K\times K$ block of the matrix $(W'W)^{-1}$, and $\hat{A}_F^*=[I_{M+1},\hat{A}_F]$.

The predictive moments follow by representing excess returns as

$$r_{T+1}=\Gamma'A_F^{*\prime}x_T+\beta u_{f,T+1}+v_{T+1}, \tag{48}$$

where $A_F^*=[I_{M+1},A_F]$. In particular, the predictive mean is obtained by taking expectations from both sides of Equation (48) conditioned on $\mathcal{D}_{\mathrm{T}}$. The predictive variance is

$$\begin{aligned}\mathbb{V}\{r_{T+1}|\mathcal{D}_T\}=&\underbrace{\mathbb{E}\{\Gamma'A_F^{*\prime}x_T x_T'A_F^*\Gamma\,|\,\mathcal{D}_T\}-\hat{\Gamma}'\hat{A}_F^{*\prime}x_T x_T'\hat{A}_F^*\hat{\Gamma}}_{\text{term 1}}\\ &+\underbrace{\mathbb{E}\{\beta u_{f,T+1}u_{f,T+1}'\beta'|\mathcal{D}_T\}}_{\text{term 2}}+\underbrace{\mathbb{E}\{v_{T+1}v_{T+1}'|\mathcal{D}_T\}}_{\text{term 3}}.\end{aligned} \tag{49}$$

We solve each of the terms in Equation (49). By the law of iterated expectations, the (i,j) element of the first factor in the first term is

$$\mathrm{tr}[\mathbb{E}\{A_F^{*\prime}x_T x_T' A_F^*|\mathcal{D}_T\}(\hat{\Gamma}_j\hat{\Gamma}_i' + \hat{\Sigma}_{RR}(i,j)(W'W)^{-1})], \tag{50}$$

where

$$\mathbb{E}\{A_F^{*\prime}x_T x_T' A_F^*|\mathcal{D}_T\} = \hat{A}_F^{*\prime}x_T x_T'\hat{A}_F^* + \begin{bmatrix} 0_{M+1,M+1} & 0_{M+1,K} \\ 0_{K,M+1} & \delta_2\hat{\Sigma}_{FF} \end{bmatrix}. \tag{51}$$

Thus

$$\text{term } 1 = \delta_2\hat{\beta}\hat{\Sigma}_{FF}\hat{\beta}' + \hat{\Sigma}_{RR}\{\mathrm{tr}[(\hat{A}_F^{*\prime}x_T x_T'\hat{A}_F^*)(W'W)^{-1} + \delta_2\hat{\Sigma}_{FF}W^{22}]\}. \tag{52}$$

Applying the law of iterated expectations to the second term yields

$$\text{term } 2 = \hat{\beta}\hat{\Sigma}_{FF}\hat{\beta}' + \hat{\Sigma}_{RR}\{\mathrm{tr}[\hat{\Sigma}_{FF}W^{22}]\}. \tag{53}$$

The variance is obtained by substituting $\hat{\Sigma}_{RR}$ (term 3), Equation (52), and Equation (53) into Equation (49).

Appendix B: Imposing Asset Pricing Restrictions on Predictive Regressions

When a multivariate predictive regression conforms to asset pricing restrictions, the predictive mean and variance are ($\mathcal{M}_R$ stands for the restricted specification)

$$\mathbb{E}\{r_{T+1}|\mathcal{D}_T,\mathcal{M}_R\} = \bar{\beta}[\bar{f} + \hat{V}_{fz}\hat{V}_z^{-1}(z_T - \bar{z})], \tag{54}$$

$$\mathbb{V}\{r_{T+1}|\mathcal{D}_T,\mathcal{M}_R\} = \bar{\delta}_1\bar{\Sigma}_{RR} + (1+\delta_2)\bar{\beta}\hat{\Sigma}_{FF}\bar{\beta}', \tag{55}$$

where

$$\bar{\beta} = R'F(F'F)^{-1}, \tag{56}$$

$$\bar{\Sigma}_{RR} = \frac{R'Q_F R}{T-N+M-1}, \tag{57}$$

$$Q_F = I_T - F(F'F)^{-1}F', \tag{58}$$

$$\bar{\delta}_1 = 1 + (1+\delta_2)\mathrm{tr}(\hat{\Sigma}_{FF}(F'F)^{-1}) + \mathrm{tr}[(\hat{A}_F'x_T x_T'\hat{A}_F)(F'F)^{-1}]. \tag{59}$$

The proof is similar to that shown in the unrestricted case.

In addition, we derive the mean and standard error of the intercepts and slopes in the restricted multivariate predictive regression. The mean and a ratio obtained by dividing the mean by the standard error are displayed in Table 1. Observe from Equations (11) and (12) that under pricing restrictions the intercept and slope in predictive regressions could be expressed jointly as $A_F\beta'$. The posterior density of that product does not correspond to any well-known density. Nevertheless, the posterior mean and standard error obey analytic expressions. In particular, the posterior mean is

$$\mathbb{E}\{A_F\beta'|\mathcal{D}_T\} = (X'X)^{-1}X'F(F'F)^{-1}F'R. \tag{60}$$

With some algebraic manipulation, we obtain the following expression for the posterior variance-covariance matrix of the ith column of $A_F\beta'$

$$\mathrm{tr}[\hat{\Sigma}_{FF}\bar{\beta}_i\bar{\beta}_i'](X'X)^{-1}+\bar{\Sigma}_{RR}(i,i)(\hat{A}_F(F'F)^{-1}\hat{A}_F'+\mathrm{tr}[\hat{\Sigma}_{FF}(F'F)^{-1}](X'X)^{-1}). \tag{61}$$

Appendix C: Intermediate Views About Mispricing, the First Specification

Let Δ be a $(K+M+1)$-vector whose first element is one and the rests are zero, let $\Upsilon=[\Delta\Delta'\frac{s^2}{\sigma^2}]$, and let $H=[\iota_T,F]$, where ι_T denotes a T-vector of ones. The prior of Equation (20) can be rewritten as

$$P(\alpha_0|\Sigma_{RR})\propto|\Sigma_{RR}|^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}\mathrm{tr}[\Sigma_{RR}^{-1}(\Gamma'\Upsilon\Gamma)]\}. \tag{62}$$

Combining Equation (62) with the prior of Equation (19) and the likelihood of Equation (36) yields the following posterior densities,

$$\mathrm{vec}([\alpha_0,\beta]')|\Sigma_{RR},\mathcal{D}_T\sim N[\mathrm{vec}(\tilde{\Gamma}),\Sigma_{RR}\otimes(H'H+\Upsilon)^{-1}], \tag{63}$$

$$\Sigma_{RR}|\mathcal{D}_T\sim IW[\Xi,T+M], \tag{64}$$

where

$$\tilde{\Gamma}=(H'H+\Upsilon)^{-1}H'R, \tag{65}$$

$$\Xi=R'R-R'H(H'H+\Upsilon)^{-1}H'R. \tag{66}$$

The mean and variance of predicted future returns are

$$\mathbb{E}\{r_{T+1}|\mathcal{D}_T\}=\tilde{\alpha}_0+\tilde{\beta}[\bar{f}+\hat{V}_{fz}\hat{V}_z^{-1}(z_T-\bar{z})], \tag{67}$$

$$\mathbb{V}\{r_{T+1}|\mathcal{D}_T\}=\frac{\tilde{\delta}_1}{T-N+M-1}\Xi+(1+\delta_2)\tilde{\beta}\hat{\Sigma}_{FF}\tilde{\beta}', \tag{68}$$

where $\tilde{\delta}_1=1+(1+\delta_2)+\mathrm{tr}(\hat{\Sigma}_{FF}\tilde{W}^{22})\mathrm{tr}[(\bar{A}^*\iota_F x_T x_T'\bar{A}_F^*)(H'H+\Upsilon)^{-1}]$, $\tilde{\alpha}_0$ and $\tilde{\beta}$ are the corresponding elements of $\tilde{\Gamma}$, $\tilde{W}^{22}$ is the lower $K\times K$ block of the matrix $(H'H+\Upsilon)^{-1}$, $\bar{A}_F^*=[i_{M+1},\hat{A}_F]$, and i_{M+1} is the first column of an identity matrix of order $M+1$.

Appendix D: Intermediate Views About Mispricing, the Second Specification

This part of the appendix analyzes two scenarios. The first, the unrestricted, is one where prior beliefs are formulated in Equations (25)–(28). The second, the restricted, relies on a similar prior, except that asset mispricing parameters are concentrated at zero. Focusing on the unrestricted case, the posterior densities are

$$\mathrm{vec}(\tilde{\Gamma})|\Sigma_{RR},\mathcal{D}_T\sim N[\mathrm{vec}(\tilde{\tilde{\Gamma}}),\Sigma_{RR}\otimes(W'W+W_0'W_0)^{-1}], \tag{69}$$

$$\Sigma_{RR}|\mathcal{D}_T\sim IW(S_R,T^*-1), \tag{70}$$

$$\mathrm{vec}(A_F)|\Sigma_{FF},\mathcal{D}_T\sim N\left[\mathrm{vec}(\tilde{A}_F),\frac{T}{T^*}\Sigma_{FF}\otimes(X'X)^{-1}\right], \tag{71}$$

$$\Sigma_{FF}|\mathcal{D}_T\sim IW(S_F,T^*+N-1), \tag{72}$$

where

$$\tilde{\tilde{\Gamma}}A = (W'W + W_0'W_0)^{-1}(W'R + W_0'W_0\Gamma_0), \tag{73}$$

$$S_R = \frac{T^*}{T}R'R - \tilde{\tilde{\Gamma}}'\ (W'W + W_0'W_0)\tilde{\tilde{\Gamma}}\ , \tag{74}$$

$$\tilde{A}_F = \frac{T}{T^*}\hat{A}_F + \frac{T_0}{T^*}A_{F0}, \tag{75}$$

$$S_F = \frac{T^*}{T}[F'F - \tilde{A}_F'(X'X)\tilde{A}_F], \tag{76}$$

$$W_0'W_0 = T_0\begin{bmatrix} 1 & \bar{z}' & \bar{f}' \\ \bar{z} & \hat{V}_z + \bar{z}\bar{z}' & \bar{z}\bar{f}' \\ \bar{f} & \bar{f}\bar{z}' & \hat{V}_f + \bar{f}\bar{f}' \end{bmatrix}, \tag{77}$$

and $T^* = T + T_0$.

Based on these posterior densities, the mean and variance of predicted future returns are

$$\mathbb{E}\{r_{T+1}|\mathcal{D}_T\} = \tilde{\tilde{\boldsymbol{\alpha}}}_0 + \tilde{\tilde{\boldsymbol{\alpha}}}_1 z_T + \tilde{\tilde{\boldsymbol{\beta}}}\left[\bar{f} + \frac{T}{T^*}\hat{V}_{fz}\hat{V}_z^{-1}(z_T - \bar{z})\right], \tag{78}$$

$$\mathbb{V}\{r_{T+1}|\mathcal{D}_T\} = \frac{\tilde{\tilde{\boldsymbol{\delta}}}_1}{T^* - N - 2}S_R + (1+\delta_2)\,\tilde{\tilde{\boldsymbol{\beta}}}\,\tilde{\Sigma}_{FF}\,\tilde{\tilde{\boldsymbol{\beta}}}\,', \tag{79}$$

where

$$\tilde{\tilde{\boldsymbol{\delta}}}_1 = 1 + (1+\delta_2)\mathrm{tr}(\tilde{\Sigma}_{FF}\ \tilde{\tilde{\boldsymbol{W}}}^{22}) + \mathrm{tr}[(\tilde{A}_F^{*'}x_T x_T'\tilde{A}_F^*)(W'W + W_0'W_0)^{-1}], \tag{80}$$

$$\tilde{\Sigma}_{FF} = \frac{S_F}{T^* + N - K - 2}, \tag{81}$$

$\tilde{A}_F^* = [I_{M+1}, \tilde{A}_F]$, and $\tilde{\tilde{\boldsymbol{W}}}^{22}$ is the lower $K \times K$ block of the matrix $(W'W + W_0'W_0)^{-1}$.

When model mispricing is precluded, the predictive mean and variance are

$$\mathbb{E}\{r_{T+1}|\mathcal{D}_T\} = \bar{\boldsymbol{\beta}}\left[\bar{f} + \frac{T}{T^*}\hat{V}_{fz}\hat{V}_z^{-1}(z_T - \bar{z})\right], \tag{82}$$

$$\mathbb{V}\{r_{T+1}|\mathcal{D}_T\} = \frac{\tilde{\tilde{\boldsymbol{\delta}}}_1}{T^* - N + M - 1}\frac{T^*}{T}R'Q_F R + (1+\delta_2)\bar{\boldsymbol{\beta}}\tilde{\Sigma}_{FF}\bar{\boldsymbol{\beta}}', \tag{83}$$

where

$$\tilde{\tilde{\boldsymbol{\delta}}}_1 = 1 + \frac{T}{T^*}\left\{(1+\delta_2)\mathrm{tr}(\tilde{\Sigma}_{FF}(F'F)^{-1}) + \mathrm{tr}[(\tilde{A}_F' x_T x_T'\tilde{A}_F)(F'F)^{-1}]\right\}. \tag{84}$$

References

Avramov, D., 2002, "Stock Return Predictability and Model Uncertainty," *Journal of Financial Economics*, 64, 423–458.

Avramov, D., and J. Chao, 2003, "An Exact Bayes Test of Asset Pricing Models with Application to International Markets," forthcoming in *Journal of Business*.

Barberis, N., 2000, "Investing for the Long Run When Returns are Predictable," *Journal of Finance*, 55, 225–264.

Barberis, N., M. Huang, and J. Santos, 2001, "Prospect Theory and Asset Prices," *Quarterly Journal of Economics*, 116, 1–53.

Bawa V., S. Brown, and R. Klein, 1979, *Estimation Risk and Optimal Portfolio Choice*, North Holland, Amsterdam.

Best, M. J., and R. R. Grauer, 1991, "On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results," *Review of Financial Studies*, 4, 315–342.

Black, F., 1972, "Capital Market Equilibrium with Restricted Borrowing," *Journal of Business*, 45, 444–454.

Black, F., and R. Litterman, 1992, "Global Portfolio Optimization," *Financial Analyst Journal*, September–October 28–43.

Breeden, D. T., 1979, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–296.

Campbell, J., 1987, "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, 373–399.

Campbell, J., 1996, "Understanding Risk and Return," *Journal of Political Economy*, 104, 572–621.

Carhart, M. M., 1997, "On the Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

Chen, N. F., R. Roll, and S. Ross, 1986, "Economic Forces and the Stock Market," *Journal of Business*, 59, 383–404.

Cochrane, J. H., 1996, "A Cross-Sectional Test of an Investment-Based Asset Pricing Model," *Journal of Political Economy*, 104, 572–621.

Connor, G., and R. Korajczyk, 1989, "An Intertemporal Equilibrium Beta Pricing Model," *Review of Financial Studies*, 2, 373–392.

Evans, M. D., 1994, "Expected Returns, Time-Varying Risk, and Risk Premia," *Journal of Finance*, 69, 655–679.

Fama, E., 1996, "Multifactor Efficiency and Multifactor Asset Pricing," *Journal of Financial and Quantitative Analysis*, 31, 441–465.

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

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

Ferson, W., and C. Harvey, 1999, "Conditioning Variables and the Cross Section of Stock Returns," *Journal of Finance*, 54, 1325–1360.

Ferson, W., and R. Korajczyk, 1995, "Do Arbitrage Pricing Models Explain the Predictability of Stock Returns?" *Journal of Business*, 68, 309–349.

Gibbons, M. R., S. A. Ross, and J. Shanken, 1989, "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 1121–1152.

Green, R., and B. Hollifield, 1992, "When Will Mean-Variance Efficient Portfolios be Well Diversified?," *Journal of Finance*, 47, 1785–1809.

Hansen, L. P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50, 1029–1054.

Hansen, L. P., J. Heaton, and E. G. J. Luttmer, 1995, "Econometric Evaluation of Asset Pricing Models," *Review of Financial Studies*, 8, 237–274.

Harvey, C., and G. Zhou, 1990, "Bayesian Inference in Asset Pricing Tests," *Journal of Financial Economics*, 26, 221–254.

Jegadeesh, N., and S. Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications for Stock-Market Efficiency," *Journal of Finance*, 48, 65–91.

Jobson, J. D., B. Korkie, and V. Ratti, 1979, "Improved Estimation for Markovitz Portfolios using James-Stein type Estimators," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 279–284.

Kahneman, D., and A. Tversky, 1979, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291.

Kan, R., and G. Zhou, 1999, "A Critique on the Stochastic Discount Factor Methodology," *Journal of Finance*, 54, 1221–1248.

Kandel, S., and R. Stambaugh, 1996, "On the Predictability of Stock Returns: An Asset allocation Perspective," *Journal of Finance*, 51, 385–424.

Kirby, C., 1998, "The Restriction on Predictability Implied by Rational Asset Pricing Models," *Review of Financial Studies*, 11, 343–382.

Lettau, M., and S. Ludvigson, 2001, "Resurrecting the (C)CAPM: A Cross Sectional Test When Risk Premia are Time-Varying," *Journal of Political Economy*, 109, 1238–1287.

Li, K., 2003, "Confidence in the Familiar: An International Perspective," forthcoming in *Journal of Financial and Quantitative Analysis*.

Lo, A. W., and A. C. MacKinlay, 1990, "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, 3, 431–468.

MacKinlay, A. C., 1995, "Multifactor Models do not Explain the CAPM," *Journal of Financial Economics*, 38, 3–28.

McCulloch, R., and P. E. Rossi, 1990, "Posterior, Predictive, and Utility-Based Approaches to Testing the Arbitrage Pricing Theory," *Journal of Financial Economics*, 28, 7–38.

McCulloch, R., and P. E. Rossi, 1991, "A Bayesian Approach to Testing the Arbitrage Pricing Theory," *Journal of Econometrics*, 49, 141–168.

Mehra, R., and E. Prescott, 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Merton, R. C., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867–887.

Pastor, L., 2000, "Portfolio Selection and Asset Pricing Models," *Journal of Finance*, 55, 179–223.

Pastor, L., and R. Stambaugh, 1999, "Costs of Equity Capital and Model Mispricing," *Journal of Finance*, 54, 67–121.

Pastor, L., and R. Stambaugh, 2000, "Comparing Asset Pricing Models: An Investment Perspective," *Journal of Financial Economics*, 56, 335–381.

Robbins, H., 1955, "An Empirical Bayes Approach to Statistics," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 157–164.

Robbins, H., 1964, "The Empirical Bayes Approach to Statistical Problems," *Annals of Mathematical Statistics*, 35, 1–20.

Rosenberg, B., and V. Marathe, 1979, "Tests of Capital Asset Pricing Hypotheses," *Research in Finance*, 1, 115–224.

Shanken, J., 1987, "A Bayesian Approach to Testing Portfolio Efficiency," *Journal of Financial Economics*, 19, 195–215.

Shanken, J., 1990, "Intertemporal Asset Pricing: An Empirical Investigation," *Journal of Econometrics*, 45, 99–120.

Shanken, J., and A. Tamayo, 2001, "Risk, Mispricing, and Asset allocation: Conditioning on Dividend Yield," working paper, University of Rochester.

Shanken, J., and M. Weinstein, 1990, "Macroeconomic Variables and Asset Pricing: Further Results," working paper, University of Rochester.

Sharpe, W., 1991, "Capital Asset Prices with and without Negative Holdings," *Journal of Finance*, 46, 489–509.

Stambaugh, R., 1999, "Predictive Regressions," *Journal of Financial Economics*, 54, 375–421.

Tamayo, A., 2002, "Stock Return Predictability, Conditional Asset Pricing Models, and Portfolio Selection," working paper, London Business School.

Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.

Zellner, A., and V. K. Chetty, 1965, "Prediction and Decision Problem in Regression Models from the Bayesian Point of View," *Journal of the American Statistical Association*, 60, 608–615.