Assessing Asset Pricing Anomalies

Michael J. Brennan
University of California, Los Angeles

Yihong Xia
University of Pennsylvania

The optimal portfolio strategy is developed for an investor who has detected an asset pricing anomaly but is not certain that the anomaly is genuine rather than merely apparent. The analysis takes account of the fact that the parameters of both the underlying asset pricing model and the anomalous returns are estimated rather than known. The value that an investor would place on the ability to invest to exploit the apparent anomaly is also derived and illustrative calculations are presented for the Fama and French SMB and HML portfolios, whose returns are anomalous relative to the CAPM.

An asset pricing anomaly is a statistically significant difference between the realized average returns associated with certain characteristics of securities, or on portfolios of securities formed on the basis of those characteristics, and the returns that are predicted by a particular asset pricing model. What is anomalous with respect to one model may be consistent with the predictions of other asset pricing models. For example, an excess return associated with a security's dividend yield is anomalous with respect to the basic capital asset pricing model (CAPM) but is consistent with extensions that incorporate investor taxes. Some anomalies are inconsistent with any known rational asset pricing model; they appear to represent "money left on the table." Such examples include the NASDAQ anomaly [Brennan, Chordia, and Subrahmanyam (1998)], the apparent slow adjustment of stock prices to earnings announcements [Ball and Brown (1968), Bernard and Thomas (1990)], and the existence of momentum (as well as longer-term reversals) in individual security returns, which was documented by Jegadeesh and Titman (1993). Sometimes these anomalies may be explicable within a model that posits certain "nonrational" behavior on the part of investors.[1]

We would like to thank two anonymous referees and seminar participants at the Western Finance Association 2000 Annual Meeting for helpful comments. All remaining errors are ours. Earlier versions of the article were finished when the second author was at the University of California, Los Angeles. Address correspondence to Michael J. Brennan, Anderson School of Management, UCLA, 110 Westwood Plaza, Box 951481, Los Angeles, CA 90095-1481, or e-mail: michael.brennan@anderson.ucla.edu.

[1] Daniel, Hirshleifer, and Subrahmanyam (1998) attempt to account for short-run momentum and long-run reversals in stock returns in terms of two classical psychological biases in judgement: overconfidence and "biased self-attribution."

The Review of Financial Studies Winter 2001 Vol. 14, No. 4, pp. 905–942

The most prominent anomalies in the contemporary asset pricing literature are those that are related to firm size[2] and the book-to-market value ratio. Fama and French (1996) have constructed two zero net investment portfolios that are designed to capture these anomalies. The SMB portfolio is a zero net investment portfolio that is long in small firms and short in large firms. The HML portfolio is a zero net investment portfolio that is long in high book-to-market value ratio firms and short in low book-to-market value ratio firms. Fama and French report that the mean returns on these zero net investment portfolios over the period 1963–1993 are 0.28% and 0.46% per month. These returns are inconsistent with the CAPM, although, as Fama and French argue, they may be consistent with certain versions of the arbitrage pricing theory or intertemporal CAPM.

This article is concerned with both normative and positive issues surrounding anomalies. From a normative standpoint it is concerned with optimal dynamic portfolio strategies for exploiting apparent asset pricing anomalies. An investor who wishes to exploit an apparent anomaly must address at least three issues. The first is whether or not the apparent anomaly is the result of "data mining."[3] To the extent that a given anomaly is inconsistent with any asset pricing model, it lacks an explanation. Under such circumstances an investor is likely to withhold his full assent to the genuineness of the anomaly. A reasonable strategy is to assign some probability to the anomaly being genuine. As time passes and more returns are observed the investor will revise his probability of the genuineness of the anomaly; this introduces an element of learning into the investor's portfolio problem. The second issue is that, even if the anomaly is genuine, the investor must decide how much to invest in it, bearing in mind that the anomalous expected return has been *estimated*, and is not a known parameter. Since the investor will learn more about the anomalous return as time passes, this introduces a second element of learning into the portfolio decision. A third issue that is important for the investor is whether or not the anomaly, if genuine, can be expected to persist in the future, or if not, the rate at which it is likely to be eliminated by the trading of other investors.

We address the first two normative issues: how an investor should revise his probability assessment of the existence of an anomaly as further returns are observed; and how he should revise his conditional distribution of asset returns, taking into account his changing assessment of the probability that the anomaly is genuine. And how this learning and the prospect of future

[2] Brennan, Chordia, and Subrahmanyam (1998) argue that the size anomaly is actually associated with the volume of trading in a security rather than size—they then argue that this is consistent with liquidity effects in returns.

[3] For one view see Black (1993): "Most of the so-called anomalies that have plagued the literature on investments seems likely to be the result of data-mining." Lo and MacKinlay (1990) present a formal analysis of the effects of data mining on the significance levels of standard statistical tests. Haugen (1995) adopts a more robust view: "In the course of the last 10 years, economists have been struggling to explain...the *huge predictable premiums in the cross-section of equity returns."*

learning should affect his optimal portfolio strategy. We do not address the issue of uncertainty about the rate at which an anomaly will be arbitraged away by other investors. Nor are we concerned with other issues of equilibrium and our investor is not intended to be a representative agent.

From a positive standpoint, we show how to assess the economic value of an anomaly to an investor with given risk attitudes and horizon. The economic value of an anomaly may help to assess the probability of an anomaly being genuine. To adapt the metaphor that is frequently used in connection with efficient markets, if genuine dollar bills are to be found on the street, it is more likely that they will be found in denominations of one dollar than of one hundred dollars. We shall use the Fama and French MKT, SMB, and HML portfolios to illustrate our approach.

Our study is an application of the analysis of the effects of learning on optimal portfolio behavior in a setting in which asset prices follow diffusion processes. Several authors[4] have studied the effects of learning about unobservable state variables that affect asset returns on the optimal portfolio choice problem in this setting. For the most part the class of information structures that have been considered is based on a Gaussian prior distribution over the state variables which, combined with the assumption of a linear relation between the state variables and the drifts of the asset return process, leads to a Gaussian conditional distribution that is completely characterized by the vector of conditional means of the state variables, the covariance matrix evolving in a deterministic fashion. Brennan (1998) and Xia (2001) have considered specializations of the general model in which the unobservable state variable is a (fixed) parameter of the return generating model; in one case the mean return on the risky asset (Brennan), and in the other a regression parameter relating the stochastically evolving mean return to a vector of observable predictor variables. Detemple (1991), extending the work of Benes and Karatzas (1983), relaxes the assumption of a Gaussian prior distribution and shows that this leads to a conditional distribution of the state variables which is characterized by two sets of sufficient statistics, the conditional mean vector and a vector of parameters that describe the covariance matrix. A major impediment to the implementation of learning models is that, for realistic problems in which asset drifts are both unobservable and evolving stochastically, the number of state variables easily becomes too large for tractability. Detemple's analysis demonstrates that a general non-Gaussian prior exacerbates this problem of dimensionality.

Therefore in this article we follow Brennan (1998) in assuming that the vector of proportional asset price drifts is constant but unobservable. However, rather than imposing a Gaussian prior, we consider a situation in which the investor's prior distribution over the drifts can be expressed as a mixture of (two) normal densities. This prior is appropriate in a situation in

[4] In particular, see Detemple (1986) and Gennotte (1986).

which the investor is not sure whether a particular asset pricing model holds. Under the null hypothesis of the asset pricing model, the prior distribution over the drift vector is assumed to be Gaussian, and under the alternative hypothesis that the model does not hold, the prior distribution is assumed also to be Gaussian but with different parameters. The two sets of parameters may be thought of as the constrained and unconstrained estimates of the mean return vector, where the constraint is that the particular asset pricing model holds. For example, an investor may be unsure whether the CAPM holds or whether the expected returns on the Fama and French SMB and HML portfolios violate the model. Assuming that the covariance matrix of asset returns is known, the CAPM constrains the mean vector so that there is only a single unknown parameter to be estimated—the market risk premium; this may be estimated from historical data. On the other hand, if the CAPM is not imposed, the expected returns of MKT, SMB, and HML are unconstrained and their prior distribution may be estimated from historical data. Thus, depending on whether or not the CAPM is imposed, there are two different distributions over the asset drift vector. Under standard assumptions these may be assumed to be Gaussian. But an investor who places a $y\%$ probability on the CAPM holding will have a prior distribution over the asset drift vector that is a mixture of normals. A prior that is a mixture of two normals may also arise if the investor does not take a strong position on the model holding, but simply uses the model as one element of an algorithm for constructing his prior. For example, we may not believe that the CAPM holds exactly because we recognize that we can never observe the relevant market portfolio; nevertheless we might wish to assign some weight to an approximate CAPM-type relation with a given market proxy in forming our prior distribution over the mean return vector. We demonstrate that when the prior is a mixture of normals, the posterior distribution is also a mixture of normals, and the vector of realized asset returns is a sufficient statistic for the posterior distribution. This means that, given the investor's information, his perceived opportunity set is described by a set of n state variables if there are n securities or portfolios, as opposed to the $2n$ variables that are required to describe the investor's posterior distribution for a general non-Gaussian prior and unobservable but stochastically evolving drifts. When the prior is a mixture of an exact asset pricing model with estimated coefficients and an unconstrained normal distribution, the investor assigns a prior probability to the validity of some asset pricing model,[5] and, over time, updates not only the estimated parameters of the model, but also the probability that the model is correct.

We solve (numerically) the investor's dynamic optimal control problem when there are three risky assets, corresponding to the three Fama and French

[5] Although we explicitly consider only a single asset pricing model, the approach is readily adapted to allow for several alternative asset pricing models.

portfolios. For the pure Gaussian prior case in which the investor assigns probability one or zero to the validity of the asset pricing model,[6] we characterize the indirect utility function and portfolio strategy of an investor with an isoelastic utility function and show that the indirect utility function and the optimal portfolio strategy are determined by a system of ordinary differential equations.

While we use a mixed normal prior over a given asset pricing model (CAPM) and an unconstrained (normal) alternative to assess anomalous returns with respect to that specific asset pricing model (CAPM), this mixed normal prior setup is also applicable in assessing the relative merits of two nonnested asset pricing models such as the CAPM and the consumption capital asset pricing model (CCAPM).[7] In this latter case, the mixed normal prior is based on two asset pricing models, and the investor assigns a prior probability to the validity of one asset pricing model (CAPM) versus another (CCAPM). Moreover, the odds ratio, which is the ratio of the posterior probability that one model holds over the posterior probability that the other model holds, provides a natural test of the relative merits of each model.

This article is related to recent articles by Pastor (1999) and Pastor and Stambaugh (1999). Pastor also analyzes the portfolio decision of an investor who can invest in the three Fama and French portfolios and is uncertain about the parameters of the joint distribution of asset returns. Our analysis differs from his in three major respects. First, Pastor considers the investor's decision in a single period or myopic context, whereas the focus of this article is on the effect of future learning on the current portfolio decision of a long-lived investor. It is worth noting that, in a myopic context in which asset prices follow diffusion processes, the effect of parameter uncertainty on the investor's decision becomes negligible as the decision horizon shrinks to zero; that is, the investor behaves as though the current assessment of the mean return vector is known for sure [see Feldman (1992)]. Secondly, while in Pastor's analysis the investor believes that (an approximate version of) the CAPM holds with probability one, and the strength of the investor's prior belief in the model is represented by the (inverse of) the covariance matrix of deviations from the model, in this article the investor assigns a probability to the validity of (a possibly approximate version of) the CAPM. Thirdly, while Pastor allows for uncertainty about the variance-covariance matrix of asset returns, this article, in keeping with the diffusion assumption for asset prices, takes the variance-covariance matrix as known [see Williams (1977)]. Pastor and Stambaugh (1999) apply a similar analysis to prior beliefs formed on the basis of (approximate versions) of two other asset pricing models, and allow for margin constraints. Finally, both articles rely implicitly on quadratic

[6] Pastor and Stambaugh (1999) label such a prior as "dogmatic."

[7] We are grateful to a referee for suggesting this.

utility, while in this article the investor is assumed to have a power utility function.

In Section 1 of the article we summarize the conflicting evidence on the size and book-to-market pricing anomalies in order to motivate our analysis of the decision problem of an investor who is uncertain about the existence of these anomalies. In Section 2 we analyze the investor's inference problem when the asset drifts are constant but unknown. We consider three cases in turn: first, the case of a general non-Gaussian prior; second, a mixture of normals prior; and third, the normal prior. The normal prior corresponds to the standard Kalman–Bucey filter and is included for completeness. We also show that the investor's inference problem is affected by whether his prior is based on an *exact* or an *approximate* asset pricing model. In Section 3 we analyze the investor's dynamic investment problem for mixture of normal priors and for the pure Gaussian prior. Section 4 applies the analysis to the Fama–French anomaly.

1. Stock Market Anomalies and Conflicting Evidence

It seems that no sooner had Michael Jensen (1978) proclaimed the "end of history" in the debate about the efficiency of stock market prices than disturbing new anomalous results began to appear, at first in a trickle, and more recently in a torrent.[8] However, these results, while apparently offering profitable investment opportunities for investors, have not gone unchallenged.[9]

The small-firm effect, the apparent abnormal returns on small-firm stocks, was first discovered by Banz (1981) and Reinganum (1981) in the United States, and subsequently confirmed in the United Kingdom and a large number of other countries. Dimson and Marsh (1999) report that a small-cap portfolio in the United States outperformed a large-firm portfolio by 4.10% per year over the period 1955–1983, while the corresponding result for the United Kingdom was 5.90% for the period 1955–1988. The terminal dates of these sample periods correspond to the end of the "launch periods" in the two countries for small-cap funds that were designed to exploit the apparent anomaly. However, over the periods 1983–1997 and 1988–1997 the small-cap premium was actually *minus* 2.4% and *minus* 5.6% for the United States and the United Kingdom, respectively. Dimson and Marsh label this commercially unfortunate phenomenon "Murphy's Law";[10] however, the negative coefficient in the "postlaunch" era is only significant at the 10% level in the

[8] For a recent summary of the evidence see Hawawini and Keim (1999).

[9] Roll (1994), principal of an investment management firm, confesses that "Over the past decade, I have attempted to exploit many of the seemingly most promising 'inefficiencies' by actually trading significant amounts of money... Many of these effects are surprisingly strong in the reported empirical work, but I have never yet found one that worked in practice."

[10] One of the authors notes that ethnic allusions are still apparently acceptable, provided that they are aimed at the right target groups!

United Kingdom, so that it is by no means clear that we should reject the earlier evidence that the premium is positive.[11]

The book-to-market effect appears to have first been identified by Rosenberg, Reid, and Lanstein (1984). Fama and French (1992, 1993) show that the book-to-market and size effects subsume many of the other anomalies.[12] In Fama and French (1998), they confirm that the effects are present in markets around the world, at least during the period 1975–1995. On the other hand, Breen and Korajcyk (1995) find no evidence of a book-to-market or size effect when they examine "real-time" Compustat tapes that avoid the biases created by "backfilling" the data on firms that meet the criterion for inclusion on the tapes [see also Kothari, Shanken, and Sloan (1995)]. However, Barber and Lyon (1997) find that the size and book-to-market effects are present in financial as well as nonfinancial firms which suggests that they cannot be accounted for by data-snooping biases; they also discount the "backfilling" bias.[13] More recently, La Porta et al. (1997) report that a significant portion of the relative performance of "value stocks" over the period 1971–1993 is attributable to systematic expectational errors which are revealed in the abnormal returns around quarterly earnings announcements. On the other hand, Loughran (1997) has found that, for large firms, the book-to-market ratio explains none of the variation in the cross section of returns outside the month of January and argues that the phenomenon (in the United States) is driven by the poor performance of small growth firms. Fama and French (1992, 1993, 1995, 1996) argue that the book-to-market and size effects are consistent with rational risk pricing, while Lakonishok, Shleifer, and Vishny (1994) suggest that the book-to-market effect is the result of investor overreaction, and Daniel, Hirshleifer, and Subrahmanyam (1998) provide an explanation that is based on systematic biases in investor decision making. Brennan, Chordia, and Subrahmanyam (1998) argue that the size effect is really a liquidity effect that is associated with the volume of trading, while Berk (1995) argues that it is a manifestation of an empirically inadequate asset pricing model.

Hawawini and Keim (1999) summarize their comprehensive discussion of the empirical evidence by saying that they believe that proposals to replace the CAPM by the Fama and French three-factor model "may be premature," and remark that the fact "that many of these (anomalous) effects have persisted for nearly 100 years in no way guarantees their persistence in the future . . . Research over the next 100 years will, we hope, settle many of these issues."

---

[11] Knez and Ready (1997) show that for the period 1963–1990 the size effect for NYSE stocks is attributable to a small number of small firms and disappears when the 1% most extreme observations are removed.

[12] But not the Jegadeesh and Titman (1993) momentum effect.

[13] Chan, Jegadeesh, and Lakonishok (1995) also argue that the survivorship bias in Compustat data is likely to be small.

Thus, after assessing the available scientific evidence, an investor is likely to agree with these authors that the issue remains in doubt and, if he has a long-term horizon, to agree that over the course of his investment horizon further light will be shed on these CAPM anomalies. The issue we address in this article is how the investor should take account of this uncertainty in designing his investment strategy today. We turn first to the investor's inference problem, and then to how he should incorporate this into his investment strategy.

2. The Investor's Inference Problem

Consider a setting in which an investor can invest in n risky assets as well as a riskless asset. The values of the risky assets follow possibly correlated geometric Brownian motions with constant coefficients. However, the investor is assumed not to be able to observe the vector of drifts or expected asset returns; he must then update his prior distribution over the drift vector from observations of the asset returns.

Then let $\mathbf{S} \equiv \ln \mathbf{P}$ where the $(n \times 1)$ vector of asset prices, $\mathbf{P}$, includes reinvested dividends, and denote the stochastic process for $\mathbf{S}$ by

$$d\mathbf{S} = \mathbf{x}\,dt + \sigma\,\mathbf{dz}, \tag{1}$$

where $\mathbf{x}$ is a constant but unobservable $(n \times 1)$ vector, σ is a known $(n \times n)$ matrix, dz is an $(n \times 1)$ vector of independent Brownian increments, and $\mathbf{\Omega} \equiv \sigma\sigma'$ is the variance-covariance matrix of asset returns.

Let $\Im_0$ denote the investor's prior information about the unobserved vector of asset drifts and let $F_0(\mathbf{x})$ denote the investor's prior distribution over the unknown vector $\mathbf{x}$. We shall consider first the general case in which the prior distribution has finite first and second moments and can be characterized by a distribution function with density $f_0(\mathbf{x})$, then the special case in which the prior distribution is a mixture of Gaussian distributions, and finally the pure Gaussian prior case.

2.1 The general nonnormal prior

Let $\Im_t \equiv \{\mathbf{S}_\tau, \tau \preceq t\}$ denote the investor's information set at time t after observing the returns on the risky assets up to time t. Then, it follows from Theorem 5.1 of Benes and Karatzas (1983)[14] that, if the investor's prior distribution over $\mathbf{x}$ has finite first and second moments and can be characterized by a distribution function with density $f_0(\mathbf{x})$, then the investor's posterior

[14] See also Theorem 3.1 of Detemple (1991).

density function, $f_t(\mathbf{x})$, is given by the following lemma:

Lemma 1. *The conditional distribution of* $\mathbf{x}$, $F_t(\mathbf{x}) \equiv F(\mathbf{x}|\Im_t)$, *has the density*

$$f(\mathbf{x};\mathbf{q},t) = \frac{\exp\left[-\frac{1}{2}t\mathbf{x}'\boldsymbol{\Omega}^{-1}(\mathbf{x}-2\mathbf{q})\right]f_0(\mathbf{x})}{\int \exp\left[-\frac{1}{2}t\mathbf{x}'\Omega^{-1}(\mathbf{x}-2\mathbf{q})\right]f_0(\mathbf{x})\mathbf{dx}}, \tag{2}$$

where $\mathbf{q}$ *is* $(n \times 1)$ *vector of realized average continuously compounded rates of return up to time* t*:*

$$\mathbf{q}_t = \frac{1}{t}\left[\ln\mathbf{P}_t - \ln\mathbf{P}_0\right]. \tag{3}$$

The mean of the conditional density, $\mathbf{m}(\mathbf{q},t) \equiv E[\mathbf{x}|\Im_t] = \int \mathbf{x}f(\mathbf{x};\mathbf{q},t)\,\mathbf{dx}$ *follows the stochastic process,*

$$\mathbf{dm} = \mathbf{G}(\mathbf{q},t)\boldsymbol{\Omega}^{-1}\mathbf{dw}, \tag{4}$$

where $\mathbf{m}_0 = \int \mathbf{x}f_0(\mathbf{x})\,\mathbf{dx}$ *and* $\mathbf{G}(\mathbf{q},t)$ *is the covariance matrix of the conditional distribution of* μ*:*

$$\mathbf{G}(\mathbf{q},t) = \frac{\int \mathbf{x}'\mathbf{x}\exp\left[-\frac{1}{2}t\mathbf{x}'\boldsymbol{\Omega}^{-1}(\mathbf{x}-2\mathbf{q})\right]f_0(\mathbf{x})\,\mathbf{dx}}{\int \exp\left[-\frac{1}{2}t\mathbf{x}'\boldsymbol{\Omega}^{-1}(\mathbf{x}-2\mathbf{q})\right]f_0(\mathbf{x})\,\mathbf{dx}} - \mathbf{m}(t,\mathbf{q})\mathbf{m}(t,\mathbf{q})' \tag{5}$$

and

$$\mathbf{dw} = \mathbf{dS} - \mathbf{mdt}. \tag{6}$$

■

Note that in contrast to the case of a nonconstant drift vector that is analyzed by Benes and Karatzas (1983) and Detemple (1991), the conditional density of the asset price drifts, $f(\mathbf{x};\mathbf{q},t)$, depends only on the stochastic vector of realized average continuously compounded asset returns, $\mathbf{q}$.

2.2 The mixture of normals prior

An investor who is uncertain as to whether a particular anomaly is genuine or only apparent may be thought of as a first approximation having a prior distribution over the mean vector that is characterized by a mixture of normal distributions. One of the normal distributions will correspond to his estimates of the asset price drift parameters under the hypothesis that the apparent anomaly is of a purely statistical origin and the asset pricing model holds, while the other distribution will correspond to his estimate accepting the validity of the anomaly. The mixing parameter will depend on the strength of his prior belief that the anomaly is valid. The following theorem, whose proof is given in the appendix, shows that, when a mixture of normal priors is updated using the realized asset returns, the conditional distribution retains the mixture of normal characteristics of the prior, the additional information provided by observation of the realized returns modifying the mean and variance of both distributions, as well as the mixing parameter. Following the theorem, we shall consider how it is modified if the investor "takes seriously" the theory underlying his mixed normal prior.

Theorem 1. The mixture of normals case. *If the investor's prior information about the unknown vector* $\mathbf{x}$ *can be characterised by a distribution function with density* $f_0(\mathbf{x})$, *which can be written as a mixture of two normal densities,*

$$f_0(\mathbf{x}) = \frac{\pi_0}{(2\pi)^{\frac{n}{2}}|\boldsymbol{\Sigma}_1|^{\frac{1}{2}}} \exp^{\{-\frac{1}{2}(\mathbf{x}-\mu_1)'\,\boldsymbol{\Sigma}_1^{-1}(\mathbf{x}-\mu_1)\}} + \frac{1-\pi_0}{(2\pi)^{\frac{n}{2}}|\boldsymbol{\Sigma}_2|^{\frac{1}{2}}} \exp^{\{-\frac{1}{2}(\mathbf{x}-\mu_2)'\,\boldsymbol{\Sigma}_2^{-1}(\mathbf{x}-\mu_2)\}}, \tag{7}$$

then the posterior density can also be written as a mixture of two normal densities,

$$f(\mathbf{x};\mathbf{q},t) = \frac{\pi(\mathbf{q},t)}{(2\pi)^{\frac{n}{2}}|\widehat{\boldsymbol{\Sigma}}_1(t)|^{\frac{1}{2}}} \exp^{\{-\frac{1}{2}(\mathbf{x}-\hat{\mu}_1(\mathbf{q},t))'\widehat{\boldsymbol{\Sigma}}_1^{-1}(t)(\mathbf{x}-\hat{\mu}_1(\mathbf{q},)t)\}} + \frac{1-\pi(\mathbf{q},t)}{(2\pi)^{\frac{n}{2}}|\widehat{\boldsymbol{\Sigma}}_2(t)|^{\frac{1}{2}}} \exp^{\{-\frac{1}{2}(\mathbf{x}-\hat{\mu}_2(\mathbf{q},t))'\widehat{\boldsymbol{\Sigma}}_2^{-1}(t)(\mathbf{x}-\hat{\mu}_2(\mathbf{q},t))\}}, \tag{8}$$

where

$$\hat{\mu}_1(\mathbf{q},t) = (\boldsymbol{\Sigma}_1^{-1} + t\boldsymbol{\Omega}^{-1})^{-1}(\boldsymbol{\Sigma}_1^{-1}\mu_1 + t\boldsymbol{\Omega}^{-1}\mathbf{q}) \tag{9}$$

$$\hat{\mu}_2(\mathbf{q},t) = (\boldsymbol{\Sigma}_2^{-1} + t\boldsymbol{\Omega}^{-1})^{-1}(\boldsymbol{\Sigma}_2^{-1}\mu_2 + t\boldsymbol{\Omega}^{-1}\mathbf{q}) \tag{10}$$

$$\mathbf{q}_t \equiv \frac{1}{t}\left[\ln \mathbf{P}_t - \ln \mathbf{P}_0\right] \tag{11}$$

$$\widehat{\boldsymbol{\Sigma}}_1(t) = (\boldsymbol{\Sigma}_1^{-1} + t\boldsymbol{\Omega}^{-1})^{-1},\ \widehat{\boldsymbol{\Sigma}}_2(t) = (\boldsymbol{\Sigma}_2^{-1} + t\boldsymbol{\Omega}^{-1})^{-1} \tag{12}$$

$$\pi(\mathbf{q},t) = \frac{\pi_0 A_1(\mathbf{q},t)}{\pi_0 A_1(\mathbf{q},t) + (1-\pi_0)A_2(\mathbf{q},t)}, \tag{13}$$

and $A_i(\mathbf{q},t)$, $(i = 1, 2)$ *is given by*

$$A_i(\mathbf{q},t) = \frac{|\widehat{\boldsymbol{\Sigma}}_i(t)|^{\frac{1}{2}}}{|\boldsymbol{\Sigma}_i|^{\frac{1}{2}}} \exp^{\frac{-1}{2}\{\mu_i'\boldsymbol{\Sigma}_i^{-1}\mu_i - \hat{\mu}_i(\mathbf{q},t)'\widehat{\boldsymbol{\Sigma}}_i(\mathbf{t})^{-1}\hat{\mu}_i(\mathbf{q},t)\}}. \tag{14}$$

The mean of the investor's distribution over the expected continuously compounded rate of return vector $\mathbf{x}$, $\mathbf{m}_t \equiv E[\mathbf{x}|\Im_t] \equiv E[\mathbf{x}|\mathbf{q},t]$ *is given by*

$$\mathbf{m}_t = \pi(\mathbf{q},t)\hat{\mu}_1(\mathbf{q},t) + (1-\pi(\mathbf{q},t))\hat{\mu}_2(\mathbf{q},t), \tag{15}$$

and its stochastic evolution is given by

$$\mathbf{dm} = \mathbf{G}(\mathbf{q}, t)\boldsymbol{\Omega}^{-1}[\mathbf{dS} - \mathbf{m}\, dt], \tag{16}$$

where the $(n \times n)$ *matrix* $\mathbf{G}(\mathbf{q}, t)$ *has typical element*

$$\begin{aligned}\mathbf{G}(\mathbf{q}, t)_{ij} = {} & \pi(\mathbf{q}, t)(\hat{\sigma}_{ij,1} + \hat{\mu}_{i,1}\hat{\mu}_{j,1}) + (1 - \pi(\mathbf{q}, t)) \\ & \times (\hat{\sigma}_{ij,2} + \hat{\mu}_{i,2}\hat{\mu}_{j,2}) - \mathbf{m}(\mathbf{q}, t)_i \mathbf{m}(\mathbf{q}, t)_j\end{aligned} \tag{17}$$

and $\hat{\sigma}_{ij,l}$ *and* $\hat{\mu}_{i,l}$ $(l = 1, 2)$ *are elements of the matrices* $\widehat{\boldsymbol{\Sigma}}_l(\mathbf{q}, t)$ *and* $\hat{\mu}_l(\mathbf{q}, t)$*, respectively.*

The variance-covariance matrix of the innovations in the mean vector, $\mathbf{m}$*, is given by*

$$(\mathbf{dm})^2 = \mathbf{G}(\mathbf{q}, t)\boldsymbol{\Omega}^{-1}\mathbf{G}(\mathbf{q}, t)'dt \equiv M(\mathbf{q}, t)dt \tag{18}$$

$$\mathbf{dmdS} = \mathbf{G}(\mathbf{q}, t)dt. \tag{19}$$

■

While Theorem 1 shows how to update a general mixture of normal priors, the following theorem applies when one of the normal distributions corresponds to a particular factor pricing model and the mixing parameter, π_0, to the probability that that asset pricing model holds. We shall consider two model-based priors: the first is based on an *approximate* factor pricing model, while the second is based on an *exact* asset pricing model.[15] While the parameter updating described in Theorem 1 is appropriate for the *approximate* model-based prior, the updating for the *exact* model-based prior is constrained by the model: first, the asset returns are used to update the distribution of the parameters of the asset pricing model under the hypothesis that it holds; then they are used to revise the probability that the model holds; finally, the model-based parameter estimates are combined with those estimated under the alternative hypothesis to update the distribution of the asset price drifts. The following theorem, which uses the results of Theorem 1, describes this formally.

Theorem 2. *Factor pricing model-based prior*

(1) *Approximate model-based prior*

Consider a prior over the mean vector that is a mixture of two normal distributions. The first distribution is derived from the hypothesis of an approximate K*-factor asset pricing model, and the second distribution over* $\mathbf{x}$ *is a*

[15] The model may be approximate because the empirical factor portfolios may not correspond exactly to their theoretical counterparts; or the theoretical model itself may be only approximate, as in the Ross (1973) arbitrage pricing theory.

multivariate normal distribution $N(\mu_2, \Sigma_2)$. *The mixing parameter,* π_0, *is the probability that the approximate factor pricing model holds.*

The first K *assets are the factor portfolios and the asset price drifts,* x_i, *of the remaining* $N-K$ *assets satisfy*

$$x_i = c_i + \sum_{k=1}^{K} \beta_{ik} x_k + \eta_i, \tag{20}$$

where $c_i \equiv r - \frac{1}{2}\Omega_{ii} + \sum_{k=1}^{K} \beta_{ik}(\frac{1}{2}\Omega_{kk} - r)$, r *is the riskless interest rate,* $\frac{1}{2}\Omega_{ii}$ *is the* (i,i)*th element in the variance-covariance matrix of the sample returns* Ω, *and the factor loadings,* β_{ik}, *are defined by*

$$dS_i = \zeta_i dt + \sum_{k=1}^{K} \beta_{ik} dS_k + d\xi_i, \tag{21}$$

where ξ_i *is a Brownian motion.*

Under the approximate factor pricing model hypothesis, η_i *is the deviation from the exact factor pricing relation and its prior is distributed* $N(0, \sigma_\eta^2)$. *Under the unconstrained normal hypothesis,* $\eta_i \equiv x_i - c_i - \sum_{k=1}^{K} \beta_{ik} x_k$ *and its prior distribution, which is given in Appendix B, is derived from the prior on* $\mathbf{x}$.

Then (i) *the vectors of the expected asset price drifts at time* t, $\mathbf{m}_{1,t} \equiv E(\mathbf{x}_1|\mathbf{q}, t)$ *for the first* K *assets, and* $\mathbf{m}_{2,t} \equiv E(\mathbf{x}_2|\mathbf{q}, t)$ *for the remaining* $N-K$ *assets, are given by*

$$\mathbf{m}_{1,t} = \pi(\mathbf{q}, t)\hat{\mu}_{\mathbf{x}_1,1} + (1 - \pi(\mathbf{q}, t))\hat{\mu}_{\mathbf{x}_1,2}; \tag{22}$$

$$\begin{aligned} \mathbf{m}_{2,t} &= \pi(\mathbf{q}, t)\Big[\hat{\mu}_{\eta,1} + \mathbf{c} + \beta\hat{\mu}_{\mathbf{x}_1,1}\Big] \\ &\quad + (1 - \pi(\mathbf{q}, t))\Big[\hat{\mu}_{\eta,2} + \mathbf{c} + \beta\hat{\mu}_{\mathbf{x}_1,2}\Big], \end{aligned} \tag{23}$$

where $\mathbf{c}$ *is a vector of* c_i, $\hat{\mu}_{\mathbf{x}_1,i}, \hat{\mu}_{\eta,i}, i = 1, 2$ *are the posterior means of the factor drifts,* $\mathbf{x}_1$, *and of the deviations from the factor pricing model,* η, *under the two hypotheses; expressions for them are given in Appendix B.*

(ii) *The variance-covariance matrix* $\mathbf{G} \equiv E((\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})'|\mathbf{q}, t)$ *has typical elements:*

$$\begin{aligned} \mathbf{G}(\mathbf{q}, t)_{ij} &= \pi(\mathbf{q}, t)(\hat{\sigma}_{ij,1} + \hat{\mu}_{i,1}\hat{\mu}_{j,1}) + (1 - \pi(\mathbf{q}, t)) \\ &\quad \times (\hat{\sigma}_{ij,2} + \hat{\mu}_{i,2}\hat{\mu}_{j,2}) - \mathbf{m}(\mathbf{q}, t)_i \mathbf{m}(\mathbf{q}, t)_j, \end{aligned} \tag{24}$$

where $\hat{\mu}_{j,k}$ *and* $\hat{\sigma}_{ij,k}$ $(k = 1, 2)$ *are elements of* $\hat{\mu}_k$ *and* $\widehat{\mathbf{\Sigma}}_k$, *the posterior mean vectors and covariance matrices of the asset drifts respectively.*

(iii) *The posterior probability of the asset pricing hypothesis,* $\pi(\mathbf{q}, t)$, *is given by Equations* (13) *and* (14).

(2) *"Exact" model-based prior*

Consider a prior over the mean vector that is a mixture of two normal distributions. The first distribution is derived from the hypothesis of an exact K-factor asset pricing model, and the second distribution over $\mathbf{x}$ *is a multivariate normal distribution $N(\mu_2, \Sigma_2)$. The mixing parameter, $\pi(\mathbf{q}, t)$, is the probability that the factor pricing model holds. Under the factor pricing model the first K assets are the factor portfolios and the asset price drifts, x_i, of the remaining $N - K$ assets satisfy the exact pricing model:*

$$x_i = c_i + \sum_{k=1}^{K} \beta_{ik} x_k, \tag{25}$$

where $c_i \equiv r - \frac{1}{2}\Omega_{ii} + \sum_{k=1}^{K} \beta_{ik}(\frac{1}{2}\Omega_{kk} - r)$.

Then (i) *the vector of posterior mean factor and nonfactor asset drifts is given by Equations* (22) *and* (23) *with $\hat{\mu}_{\eta,1} = 0$, which implies that $\hat{\mu}_{i,1}$, the posterior mean drift for asset i ($i = K+1, \ldots, N$), satisfies the asset pricing relation:*

$$\hat{\mu}_{1,i} + \frac{1}{2}\Omega_{ii} = r + \sum_{k=1}^{K} \beta_{ik}\left(\hat{\mu}_{1,k} + \frac{1}{2}\Omega_{kk} - r\right), \quad i = K+1, \ldots, n. \tag{26}$$

(ii) *The variance-covariance matrix $\mathbf{G} \equiv E((\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})'|\mathbf{q}, t)$ has typical elements given by Equation* (24).

(iii) *$\pi(\mathbf{q}, t)$, the posterior probability of the factor pricing model, is given by Equation* (13) *where $A_1(\mathbf{q}, t)$ is*

$$A_1(\mathbf{q}, t) = \frac{|\widehat{\mathbf{\Sigma}}_{11,1}(t)|^{\frac{1}{2}}}{|\mathbf{\Sigma}_{11,1}|^{\frac{1}{2}}} \exp^{\frac{-1}{2}\mu_{1,1}' \mathbf{\Sigma}_{11,1}^{-1} \mu_{1,1} - \hat{\mu}_{1,1}(\mathbf{q},t)' \widehat{\mathbf{\Sigma}}_{11,1}(t)^{-1} \hat{\mu}_{1,1}(\mathbf{q},t)\}} \tag{27}$$

and $A_2(\mathbf{q}, t)$ is given by Equation (14).

Proof. See Appendix B.

While Theorem 1 applies to any mixture of normal priors, Theorem 2 specializes it to a prior that is based on a factor pricing model such as the arbitrage pricing theory or the CAPM ($K = 1$).[16] When the prior is based on a mixture of an *approximate* factor pricing model and an unconstrained normal distribution, the application of Theorem 1 is direct: the posterior

[16] While we do not treat the CCAPM, the pure CCAPM-based prior and the mixed prior based on CCAPM and an unconstrained alternative would be similar to the CAPM-based prior, with uncertainty about the risk aversion of the representative agent replacing the uncertainty about the market risk premium. The treatment of a mixed normal prior based on two model-based priors such as CAPM and CCAPM is also similar: under CAPM, the prior market risk premium is assumed to be normally distributed; under CCAPM, it is derived from the prior distribution of the risk aversion of the representative agent. In this case the odds ratio, given by $\pi(\mathbf{q}, t)/(1 - \pi(\mathbf{q}, t))$, provides a natural test of the relative merit of one model against the other.

factor means, $\mathbf{m}_{1,t}$, are weighted averages of the posterior means under the two distributions, $\hat{\mu}_{\mathbf{x}_1,i}$, and the posterior means of the $N-K$ "non-factor" assets, $\mathbf{m}_{2,t}$, are constructed as the sum of the prediction from the factor pricing model using the posterior factor means, $\mathbf{c}+\beta\hat{\mu}_{\mathbf{x}_1,i}$, and the posterior means of η under the two distributions, $\hat{\mu}_{\eta,i}$. When the prior is based on a mixture of an *exact* factor pricing model and an unconstrained normal distribution, the posterior factor means and the posterior means of the $N-K$ "nonfactor" assets are constructed in the same way except that $\hat{\mu}_{\eta,1}=0$, and $\pi(\mathbf{q},t)$ is the posterior probability that the factor pricing model holds.[17]

Theorem 1 establishes that $\mathbf{q}$, the vector of realized average returns, and t are sufficient statistics for the investor's posterior distribution. We note that for the pure Gaussian case ($\pi_0 = 0, 1$) $\mathbf{m}$ is a linear function of $\mathbf{q}$ [see Equations (13] and (15)], so that $\mathbf{m}$, the vector of current assessments of the drift, and t are also sufficient statistics.

2.3 The Normal Prior

For completeness, and to relate the above to the classic Bucey–Kalman filtering problem, we consider the special case in which $\pi_0 = 1$, so that the prior distribution over $\mathbf{x}$ is Gaussian, and the filtering problem and the conditional distribution over $\mathbf{x}$ are also Gaussian, as shown in the following corollary. For simplicity we drop the subscript 1 from the parameters of $\mathbf{\Sigma}_1$ and μ_1 of the prior distribution.

Corollary 1 *(Kalman Filter). When $\pi_0 = 1$, the investor's posterior density over* $\mathbf{x}$ *simplifies to*

$$f(\mathbf{x};\mathbf{m},t) = \frac{1}{(2\pi)^{\frac{n}{2}}|\widehat{\mathbf{\Sigma}}(t)|^{\frac{1}{2}}}\exp^{\{-\frac{1}{2}(\mathbf{x}-\mathbf{m})'\widehat{\mathbf{\Sigma}}^{-1}(t)(\mathbf{x}-\mathbf{m})\}}, \tag{28}$$

where the variance covariance matrix of the conditional distribution over the drifts is

$$\widehat{\mathbf{\Sigma}}(\mathbf{t}) = (\mathbf{\Sigma}^{-1}+t\mathbf{\Omega}^{-1})^{-1} \tag{29}$$

and the conditional mean vector, $\mathbf{m}$*, is given by*

$$\mathbf{m} = (\mathbf{\Sigma}^{-1}+t\mathbf{\Omega}^{-1})^{-1}(\mathbf{\Sigma}^{-1}\mu+t\mathbf{\Omega}^{-1}q). \tag{30}$$

[17] Note that when an *approximate* factor pricing model is used to construct the prior, $\pi(\mathbf{q},t)$ cannot be interpreted as the posterior probability that the model holds, because the posterior means of the model deviations are not equal to zero.

The stochastic evolution of $\mathbf{m}$ *is given by*

$$\mathbf{dm} = \mathbf{G}(t)\boldsymbol{\Omega}^{-1}[\mathbf{dS} - \mathbf{m}\,dt] \tag{31}$$

$$(\mathbf{dm})^2 = \mathbf{G}(t)\boldsymbol{\Omega}^{-1}\mathbf{G}(t)'\,dt \equiv \mathbf{M}(t)\,dt, \tag{32}$$

where now the $(n \times n)$ *covariance matrix* $\mathbf{G}(t)$ *is*

$$\mathbf{G}(t) = \widehat{\boldsymbol{\Sigma}}(t) = (\boldsymbol{\Sigma}^{-1} + t\boldsymbol{\Omega}^{-1})^{-1}. \tag{33}$$

Hence the only difference between the Gaussian prior case and the mixture of normal priors case is in the definition of the matrix, $\mathbf{G}$, which is the variance-covariance matrix of the investor's conditional distribution over $\mathbf{x}$. In the Gaussian case the matrix is deterministic, being the solution of the familiar Ricatti equation. In the mixture of normals case, the matrix depends on the stochastic vector $\mathbf{q}$ as shown in Equation (17).

In the following section we consider the implications of the different priors for optimal dynamic portfolio strategies.

3. Optimal Portfolio Strategies for Anomalies

For simplicity, and to emphasize the role of the investment horizon, the agent is assumed to be concerned with maximizing the expected value of a monotone increasing concave von Neumann–Morgenstern utility function, defined over wealth at time T, which we denote by $U(W_T)$. As Gennotte (1986) has shown in a similar setting of incomplete information, the investor's decision problem may be decomposed into two separate problems: an inference problem such as we have described in Section 1, in which the investor updates his distribution over the current value of the unobserved vector, $\mathbf{x}$, and an optimization problem with a completely observed state in which the state is described by the sufficient statistics for the investor's distribution over the unobservable vector. Since the inference problem is different for the normal and mixed normal cases, we shall treat them separately.

3.1 Normal prior distribution

Since $\mathbf{q}$ and t, or equivalently $\mathbf{m}$ and t under a normal prior [see Equation (30)], are sufficient statistics for the investor's conditional distribution over $\mathbf{x}$, define $J(W, \mathbf{m}, t)$ as the expected value at time $t \prec T$ of the utility of wealth at time T, under the optimal policy, when the investor's current wealth is W and his current assessment of the drift vector is $\mathbf{m}$. Then

$$J(W, \mathbf{m}, t) = \max_{\alpha} E[U(W_T)|\Im_t], \tag{34}$$

where α is the $(n \times 1)$ vector of portfolio allocations, and the maximization is subject to the dynamic budget constraint

$$dW = W[r + \alpha'(\mathbf{m}^* - r\mathbf{i})]\,dt + W\alpha'\sigma\mathbf{dz}, \tag{35}$$

where $\mathbf{m}^* \equiv \mathbf{m} + \frac{1}{2}\mathbf{diag}(\boldsymbol{\Omega})$ is the vector of expected instantaneous rates of return, r is the riskless interest rate, and $\mathbf{i}$ is a vector of units.

Then the Bellman equation for the optimal control problem can be written as

$$\begin{aligned}\max_{\alpha}\Big(&\frac{1}{2}J_{WW}W^2\alpha'\boldsymbol{\Omega}\alpha + \frac{1}{2}\mathrm{tr}\{\mathbf{MJ}_{mm'}\} + W\alpha'\mathbf{GJ}_{Wm},\\ &+ J_W W[r + \alpha'(\mathbf{m}^* - r\mathbf{i})] + J_t\Big) = 0,\end{aligned} \tag{36}$$

where $\mathbf{J}_{mm'}$ is an $(n \times n)$ matrix and J_{Wm} is an $(n \times 1)$ vector.

The vector of optimal proportional portfolio allocations, $\boldsymbol{\alpha}^*$, follows from the first-order conditions:

$$\boldsymbol{\alpha}^* = \frac{-J_W}{WJ_{WW}}\boldsymbol{\Omega}^{-1}(\mathbf{m}^* - r\mathbf{i}) - \frac{1}{WJ_{WW}}\boldsymbol{\Omega}^{-1}\mathbf{GJ}_{Wm}. \tag{37}$$

In the normal prior case, which corresponds to the dogmatic prior that the anomaly either does or does not exist, the matrices $\mathbf{M}$ and $\mathbf{G}$ are functions only of t, defined by Equations (32) and (33). Then there exist simple expressions for the investor's indirect utility function and optimal investment strategy when the investor's utility function is of the isoelastic family, as shown in the following theorem which may be verified by substitution in Equations (36) and (37).

Theorem 3. Optimal Strategies for Normal Priors. *If the investor's utility of terminal wealth may be written as*

$$U(W) = \frac{W^{1-\gamma}}{1-\gamma} \tag{38}$$

and the investor's prior distribution over the unobserved vector, $\mathbf{x}$, *is normal of the form of Equation* (28)*, then*

(i) *The investor's indirect utility function is given by*

$$J(W, \mathbf{m}, t) = \frac{W^{1-\gamma}}{1-\gamma}e^{\{\mathbf{A}(t) + \mathbf{B}(t)m + \frac{1}{2}\mathbf{m}'\mathbf{C}(\mathrm{t})\mathbf{m}\}}, \tag{39}$$

where $\mathbf{A}(t)$ *is a scalar,* $\mathbf{B}(t)$ *is a* $(1 \times n)$ *vector,* $\mathbf{C}(t)$ *is an* $(n \times n)$ *matrix, and* $\mathbf{A}(t), \mathbf{B}(t)$*, and* $\mathbf{C}(t)$ *satisfy the system of ordinary differential equations:*

$$\mathbf{C}'\mathbf{MC} + \frac{1-\gamma}{\gamma}\Big(\mathbf{\Omega}^{-1} + \mathbf{\Omega}^{-1}\mathbf{GC} + \mathbf{CG\Omega}^{-1} + \mathbf{C}'\mathbf{MC}\Big) + \mathbf{C}_t = \mathbf{0} \quad (40)$$

$$\mathbf{BMC} + \frac{1-\gamma}{\gamma}(\mathbf{BG\Omega}^{-1} - r\mathbf{i}'\mathbf{\Omega}^{-1}\mathbf{GC} - r\mathbf{i}'\mathbf{\Omega}^{-1} + \mathbf{BMC}) + B_t = \mathbf{0} \quad (41)$$

$$\frac{1}{2}\mathrm{tr}\{\mathbf{MC}'\} + \frac{1}{2}\mathbf{BMB}' + (1-\gamma)r + \frac{1-\gamma}{2\gamma} \times (r^2\mathbf{i}'\mathbf{\Omega}^{-1}\mathbf{i} - 2r\mathbf{i}'\mathbf{\Omega}^{-1}\mathbf{GB}' + \mathbf{BMB}') + A_t = 0 \quad (42)$$

with boundary conditions: $A(T) = 0$*;* $\mathbf{B}(T) = \mathbf{0}$*; and* $\mathbf{C}(T) = \mathbf{0}$.

(ii) *The vector of the investor's optimal proportional allocations to the risky assets is given by*

$$\alpha^* = \frac{1}{\gamma}\Big[\mathbf{\Omega}^{-1}(\mathbf{m}^* - r\mathbf{i}) + \mathbf{\Omega}^{-1}\mathbf{G}\,(\mathbf{B}(t) + \mathbf{C}(t)\mathbf{m})\Big]. \quad (43)$$

3.2 Mixture of normal priors

When the prior is a mixture of normals which, as we have argued, is likely to be the case when some of the portfolio returns are *anomalous*, the matrices $\mathbf{M}$ and $\mathbf{G}$ are functions of $\mathbf{q}$ as well as t; in this case there is no analytic solution to the control problem which must therefore be solved numerically. It follows from the definition of $\mathbf{q}$ that

$$\mathbf{dq} = \frac{\mathbf{m}(\mathbf{q}, t) - \mathbf{q}}{t}dt + \frac{\sigma}{t}dw. \quad (44)$$

Since $\mathbf{q}$ and t are sufficient statistics for the investor's conditional distribution over μ, define $J(W, \mathbf{q}, t)$ as the expected value at time $t \prec T$ of the utility of wealth at time T under the optimal policy, when the investor's current wealth is W and the vector of realized average returns on the portfolios is $\mathbf{q}$. The Bellman equation for the investor's optimal control problem is then

$$\max_{\alpha}\Bigg(\frac{1}{2}J_{WW}W^2\alpha'\mathbf{\Omega}\alpha + \frac{1}{2t^2}\mathrm{tr}\{\mathbf{\Omega J}_{\mathbf{qq}'}\} + \frac{1}{t}W\alpha'\mathbf{\Omega}J_{w\mathbf{q}} + J_W W \times [r + \alpha'(\mathbf{m}^*(\mathbf{q}, t) - r\mathbf{i})] + \Big(\frac{\mathbf{m}(\mathbf{q}, t) - \mathbf{q}}{t}\Big)' J_{\mathbf{q}} + J_t\Bigg) = 0. \quad (45)$$

For the isoelastic utility function of Equation (38), the investor's indirect utility function may be written as $J(W, \mathbf{q}, t) = \frac{W^{1-\gamma}}{1-\gamma}\phi(\mathbf{q}, t)$, where

$$\max_{\alpha}\Bigg(\frac{1}{2t^2}\mathrm{tr}\{\boldsymbol{\Omega}\phi_{\mathbf{q}\mathbf{q}'}\} + \frac{1}{t}\Big((1-\gamma)\alpha'\boldsymbol{\Omega} + (\mathbf{m}(\mathbf{q}, t) - \mathbf{q})\Big)\phi_{\mathbf{q}} + (1-\gamma)[r + \alpha'(\mathbf{m}^*(\mathbf{q}, t) - r\mathbf{i}) - \frac{1}{2}\gamma\alpha'\boldsymbol{\Omega}\alpha]\phi + \phi_t\Bigg) = 0. \quad (46)$$

The first-order conditions for Equation (46) are

$$\alpha^* = \frac{1}{\gamma}\Big[\boldsymbol{\Omega}^{-1}(\mathbf{m}^*(\mathbf{q}, t) - r\mathbf{i}) + \frac{1}{t}\xi(\mathbf{q}, t)\Big], \quad (47)$$

where $\xi(\mathbf{q}, t) = [\phi_{q1}/\phi, \phi_{q2}/\phi, \phi_{q3}/\phi]'$.

Both the drift and the diffusion of the stochastic process for $\mathbf{q}$, Equation (44), are time dependent. As $t \to \infty$, $\mathbf{q}, \mathbf{m} \to \mathbf{x}$; that is, the average realized return vector and the investor's assessment of the drift vector both approach the true drift vector, $\mathbf{x}$. However, $\mathbf{q}$ becomes ill-behaved for small t. Therefore, in our numerical analysis we define the state variable as $\hat{\mathbf{q}}_t \equiv t\mathbf{q}_t$ for $t \leq 1$. The expressions that are used for $\hat{\mu}_1(\mathbf{q}, t)$, $\widehat{\boldsymbol{\Sigma}}_1(t)$ and $\pi(\mathbf{q}, t)$ that enter $\mathbf{m}(\mathbf{q}, t)$ in the control problem [Equation (46)] depend on whether the prior is based on an *approximate* or an *exact* factor pricing model. In the latter case, $\hat{\mu}_{\eta,1} = 0$ and $A_1(\mathbf{q}, t)$ is given by Equation (27).

4. Assessing the Fama and French CAPM Anomalies

To illustrate the effects of both model and parameter uncertainty on optimal portfolio strategies, and to estimate the economic value of the CAPM anomaly represented by the Fama and French SMB and HML portfolios, the optimal portfolio strategy was computed for an investor with a 20-year horizon and an isoelastic utility function using parameter estimates derived from historical returns on the market portfolio and the Fama and French SMB and HML portfolios.[18] The data are drawn from the period July 1963–December 1991. Panel A of Table 1 reports the vector of (arithmetic) mean returns in excess of the Treasury-bill rate and associated standard deviations, as well as the correlation matrix for the three-factor returns, and the standard errors of the means. All estimates are annualized by multiplying the corresponding monthly figures by 12 or the square root of 12 as appropriate. The figures show the familiar high reward to risk ratio of the HML portfolio and, to a lesser degree, of the SML portfolio.

---

[18] We are grateful to Eugene Fama for making these available to us.

Table 1
Statistics for Fama and French factor portfolios

	Mean excess return	Standard deviation	Standard deviation of mean	Correlations		
				Market	SMB	HML
Panel A. July 1963–December 1991						
Market	5.21%	15.7	2.94	1.0	0.32	−0.38
SMB	3.25%	10.0	1.89	0.32	1.0	−0.08
HML	4.78%	8.8	1.65	−0.38	−0.08	1.0
Panel B. January 1959–December 1978						
Market	5.95%	14.5	3.27	1.0	0.39	−0.25
SMB	6.35%	10.5	2.30	0.39	1.0	−0.03
HML	9.45%	8.4	1.86	−0.25	−0.03	1.0

This table gives summary statistics for Fama and French (1996) portfolios from July 1963 to December 1991. The market return is the return in excess of the Treasury-bill rate. All returns are annualized by multiplying the monthly figure by 12.

4.1 Pure prior distributions

We consider first the case in which the investor places all his probability on either an *exact* version of the CAPM or assigns zero probability to the model and employs an unconstrained normal prior; in both cases we assume that the parameters of the investor's prior distribution of the mean returns are derived from the historical data, and that the real riskless interest rate is 3%. For example, for the CAPM prior we assume that the investor takes the volatility of the return on the market portfolio as 15.7% and known, while his distribution over the mean of the excess return on the market portfolio is normally distributed with mean 5.21% and standard deviation of 2.94%. The investor's information is analogous for the unconstrained prior except that the means and standard deviations are augmented by the correlation matrix which is also assumed to be known. The opportunity set looks much less risky to the investor with an unconstrained prior than to the CAPM investor because of the negative correlation of the return on the HML portfolio with the returns on the market and SMB portfolios: for example, using the data in panel A of Table 1 we can reject the null hypothesis that the true Sharpe ratio is zero at the 0.06% level for the unconstrained prior, but only at the 10.7% level for the CAPM prior.[19] These two cases of pure CAPM or pure unconstrained priors are examples of the normal prior analyzed in Section 2.3. Therefore the investor's optimal strategies in the two cases are given by Equation (43). The ordinary differential equations [Equations (40)–(42)] were solved by finite difference approximation and used to calculate the optimal portfolio strategy [Equation (43)]. The results for an investor with a 20-year horizon are shown in panel A of Table 2. In this table the results for the CAPM prior are shown in bold italics, while those for the unconstrained normal prior are shown in normal type.

[19] Under the null hypothesis of the CAPM with a zero market risk premium the squared Sharpe ratio for the period is distributed $F_{1,341}$; under the unconstrained normal prior the corresponding distribution is $F_{3,339}$. See, for example, Mackinlay (1995).

Table 2
Optimal portfolio strategies and certainty equivalent returns

γ	Security	Prior mean risk premium	α_M	α_H	α^* Unconstrained	α^* Constrained	Certainty equivalent return Unconstrained	Certainty equivalent return Constrained
Panel A. Investor with a 20-year horizon and two alternative priors over the mean return vector								
2.0	***Market***	***5.21%***	***1.06***	***−0.27***	***0.79***	***0.80***	***5.5%***	***5.4%***
	Market	5.21	1.78	−0.38	1.40	0.32	17.4%	7.7%
	SMB	3.25	1.03	−0.10	0.93	0.00		
	HML	4.78	4.41	−1.34	3.07	0.68		
3.0	***Market***	***5.21%***	***0.70***	***−0.21***	***0.49***	***0.49***	***4.5%***	***4.5%***
	Market	5.21	1.19	−0.32	0.87	0.25	11.7%	6.9%
	SMB	3.25	0.69	−0.04	0.65	0.00		
	HML	4.78	2.94	−1.12	1.82	0.00		
4.0	***Market***	***5.21%***	***0.53***	***−0.18***	***0.35***	***0.35***	***4.1%***	***4.1%***
	Market	5.21	0.89	−0.26	0.63	0.28	9.2%	5.9%
	SMB	3.25	0.52	−0.02	0.50	0.00		
	HML	4.78	2.20	−0.92	1.28	0.00		
5.0	***Market***	***5.21%***	***0.42***	***−0.14***	***0.28***	***0.28***	***3.9%***	***3.9%***
	Market	5.21	0.71	−0.22	0.49	0.19	7.8%	5.2%
	SMB	3.25	0.41	−0.00	0.41	0.00		
	HML	4.78	1.76	−0.77	0.99	0.00		
Panel B. Investor with a 10-year horizon and two alternative priors over the mean return vector								
2.0	***Market***	***5.21%***	***1.06***	***−0.16***	***0.90***	***0.86***	***5.7%***	***5.7%***
	Market	5.21	1.78	−0.21	1.57	0.34	19.3%	7.8%
	SMB	3.25	1.03	−0.10	0.93	0.00		
	HML	4.78	4.41	−0.74	3.67	0.66		
3.0	***Market***	***5.21%***	***0.70***	***−0.13***	***0.57***	***0.57***	***4.7%***	***4.7%***
	Market	5.21	1.19	−0.18	1.01	0.32	13.1%	7.3%
	SMB	3.25	0.69	−0.07	0.62	0.00		
	HML	4.78	2.94	−0.64	2.30	0.68		
4.0	***Market***	***5.21%***	***0.53***	***−0.11***	***0.42***	***0.42***	***4.3%***	***4.3%***
	Market	5.21	0.89	−0.15	0.74	0.29	10.3%	6.9%
	SMB	3.25	0.52	−0.06	0.46	0.00		
	HML	4.78	2.20	−0.52	1.68	0.71		
5.0	***Market***	***5.21%***	***0.42***	***−0.09***	***0.33***	***0.33***	***4.0%***	***4.0%***
	Market	5.21	0.71	−0.12	0.59	0.25	8.7%	6.2%
	SMB	3.25	0.41	−0.04	0.37	0.00		
	HML	4.78	1.76	−0.44	1.32	0.75		

The priors are the CAPM and an unconstrained normal prior where the mean vectors and covariance matrices are estimated from monthly data from July 1963 to December 1991 as shown in Table 1; the risk-free interest rate is 3%. Figures shown in bold italics are for the CAPM prior; other figures relate to the unconstrained normal prior. α_M is the fraction of wealth allocated to each security in the optimal myopic portfolio; α_H is the additional allocation to each security to hedge against changes in parameter estimates; α^* is the optimal total allocation to each security. The certainty equivalent rate of return is the sure rate of return up to the horizon that would leave the investor as well off as having \$1 of current wealth and 20 or 10 years, respectively, to invest with the currently assessed investment opportunity set. Columns labeled "constrained" impose the constraint that portfolio allocations (including riskless asset) be nonnegative.

The column α_M denotes the optimal portfolio for an investor who behaves myopically, selecting his portfolio simply on the basis of the instantaneous mean-variance trade-off.[20] The column α_H denotes the vector of asset allocations (as a proportion of the investor's wealth) to a hedge portfolio which

[20] As shown in Equation (43), $\alpha_M = \frac{1}{\gamma}\boldsymbol{\Omega}^{-1}(\mathbf{m}^* - r\mathbf{i})$.

is designed to hedge against changes in the investor's perceived investment opportunities.[21] $\alpha^* \equiv \alpha_M + \alpha_H$ denotes the investor's aggregate portfolio allocation vector.[22] The certainty equivalent rate of return is such that the investor with the stated value of γ would be indifferent between earning this rate of return for sure up to the horizon on the one hand, and on the other hand, receiving \$1 today with the opportunity to invest for the 20 years, given the perceived investment opportunity set and the prospect of learning more about it as time passes. In other words, the investor would be as well off with a riskless investment opportunity offering the certainty equivalent rate of return as he would be with \$1 of initial wealth and the perceived investment opportunity.

The results for $\gamma = 1$ (not shown) correspond to the case of log utility. It is well known that in this case the investor does not hedge against changes in the perceived investment opportunity set, so that the optimal portfolio is the myopic portfolio, and the investor allocates 2.11 times his wealth to the market portfolio.

For $\gamma = 2$ the CAPM investor not only halves his myopic allocation to stocks to 1.06 times his wealth, but takes a short position in the hedge portfolio equal to 27% of his wealth, so that his net allocation to stocks drops from 2.11 times wealth to only 0.79 times wealth. As Brennan (1998) has shown in a similar setting, parameter uncertainty has a large effect on the optimal allocation to risky assets for a (nonlog) investor with a long horizon. The certainty equivalent rate of return for this investor is 5.5%, or 2.5% above the riskless interest rate. However, for the same investor, the certainty equivalent rate of return is 17.4% if he assigns all the probability to the unconstrained hypothesis and is able to take unconstrained portfolio positions. The investor with an unconstrained prior also takes short positions in the three factor portfolios to form his hedge portfolio; however, it is noticeable that the ratio of the short positions in the hedge portfolio to the long positions in the myopic portfolio is much less for the unconstrained prior investor than for the CAPM investor. The fact that the positions in the hedge portfolio are short for both the CAPM investor and the unconstrained prior investor can be understood in terms of intertemporal diversification. If the early returns on, say, HMB are negative, the investor will revise down his estimate of $\mathbf{m}_3(t)$ and will tend to earn lower returns in the future; this means that his assessed future investment opportunities will deteriorate (improve) when a current low (high) return occurs; by taking a short position in the hedge portfolio today he hedges or diversifies his risk over time. The magnitude of the hedge position depends on the elasticity of the marginal utility of wealth with respect to the state variables, J_{Wm} in Equation (37) or $\xi(\mathbf{q}, t)$ in Equation (47).

[21] From Equation (43) the hedge portfolio is given by $\alpha_H = \frac{1}{\gamma}\Omega^{-1}G(B(t) + C(t)m)$.

[22] The columns labeled "constrained" report the results when the investor is prohibited from taking short positions or borrowing. Unless otherwise stated, our discussion relates to the unconstrained results.

Table 3
Optimal portfolio strategies and certainty equivalent returns

γ	Security	Prior mean risk premium	α_M	α_H	α^* Unconstrained	α^* Constrained	Certainty equivalent return Unconstrained	Certainty equivalent return Constrained
Panel A. Investor with a 20-year horizon and an *exact* model mixed (50:50) normal prior over the mean return vector								
2.0	Market	5.21%	1.42	−0.69	0.73	0.67	10.2%	6.8%
	SMB	2.16	0.52	−0.78	−0.26	0.00		
	HML	1.84	2.20	−2.40	−0.20	0.33		
3.0	Market	5.21%	0.94	−0.53	0.41	0.48	7.1%	5.8%
	SMB	2.16	0.34	−0.57	−0.23	0.00		
	HML	1.84	1.46	−1.77	−0.31	0.00		
4.0	Market	5.21%	0.71	−0.43	0.28	0.32	5.9%	5.1%
	SMB	2.16	0.26	−0.45	−0.19	0.00		
	HML	1.84	1.10	−1.38	−0.28	0.00		
5.0	Market	5.21%	0.57	−0.36	0.21	0.24	5.2%	4.7%
	SMB	2.16	0.21	−0.37	−0.16	0.00		
	HML	1.84	0.88	−1.13	−0.25	0.00		
Panel B. Investor with a 10-year horizon and an *exact* model mixed (50:50) normal prior over the mean return vector								
2.0	Market	5.21%	1.42	−0.40	1.02	0.68	10.8%	7.8%
	SMB	2.16	0.52	−0.48	0.04	0.00		
	HML	1.84	2.20	−1.46	0.74	0.32		
3.0	Market	5.21%	0.94	−0.32	0.62	0.57	7.6%	6.1%
	SMB	2.16	0.34	−0.36	−0.02	0.00		
	HML	1.84	1.46	−1.13	0.33	0.43		
4.0	Market	5.21%	0.71	−0.26	0.45	0.50	6.3%	5.5%
	SMB	2.16	0.26	−0.29	−0.03	0.00		
	HML	1.84	1.10	−0.90	0.20	0.50		
5.0	Market	5.21%	0.57	−0.22	0.35	0.32	5.5%	5.1%
	SMB	2.16	0.21	−0.24	−0.03	0.00		
	HML	1.84	0.88	−0.74	0.14	0.00		

The prior distribution over the mean vector is a mixture of normal distributions that correspond to the approximate CAPM prior ($\sigma_\eta = 0\%$) and an unconstrained normal prior where the mean vectors and covariance matrices are estimated from monthly data from July 1963 to December 1991 as shown in Table 1; the risk-free interest rate is 3%. Figures shown in bold italics are for the CAPM prior; other figures relate to the unconstrained normal prior. α_M is the fraction of wealth allocated to each security in the optimal myopic portfolio; α_H is the additional allocation to each security to hedge against changes in parameter estimates; α^* is the optimal total allocation to each security. The certainty equivalent rate of return is the sure rate of return up to the horizon that would leave the investor as well off as having \$1 of current wealth and 20 and 10 years, respectively, to invest with the currently assessed investment opportunity set. Columns labeled "constrained" impose the constraint that portfolio allocations (including riskless asset) be nonnegative.

As shown in Table 3, under some priors it is possible for the negative hedge positions in SMB and HML to exceed the positive positions in the myopic portfolio, making it optimal for a long-horizon investor to short these portfolios, despite their current high expected returns. As the risk aversion increases, the myopic portfolio allocation, the certainty equivalent return, and the absolute value of the hedge positions decline monotonically; however, the ratio of the hedge portfolio value to the myopic portfolio value increases.

Panel B of Table 2 reports the corresponding results when the investor's horizon is 10 years. The general effect of reducing the horizon is to reduce the size of the short positions in the hedge portfolio. The certainty equivalent rates of return are lower the longer is the horizon. Pastor and Stambaugh (1999), in comparing portfolio choices implied by different asset pricing models under parameter uncertainty, calculate annualized one-month certain equivalent returns. It is therefore of interest to note that the annualized one-month certainty equivalent return can differ significantly from the 10-year certainty equivalent return in this model; for example, when $\gamma = 3$ the annualized one-month return for the three-factor model is 15.3%, while the 10-year return is only 13.1%.

4.2 Mixed prior distributions

We consider next the case in which the investor's prior distribution over $\mathbf{x}$ is derived by assigning probability π_0 to the CAPM hypothesis and $(1-\pi_0)$ to the unconstrained normal hypothesis. The parameters of the prior distributions of the mean returns under both hypotheses are derived from the historical data for the period July 1963–December 1991 shown in panel A of Table 1. During this period the returns on the SMB and HML portfolios were anomalous with respect to the CAPM, so that the investor is implicitly allowing for the possibility that the anomaly is genuine. We consider two versions of the CAPM hypothesis; under both versions the expected returns on the HML and SMB portfolios are determined mechanically from the expected return on the market portfolio by the familiar CAPM relation, with the betas which are estimated from the historical data being treated as known; under the *exact* version of the CAPM hypothesis, the variance-covariance matrix of the prior means, Σ_1, is singular with typical element $\Sigma_{ij} = \beta_i \beta_j \Sigma_{11,1}$. The *approximate* version of the CAPM hypothesis allows for errors in the CAPM, due for example to mismeasurement of the market portfolio, so that the model may not hold exactly; in this case the diagonal elements of the variance-covariance matrix are augmented by $\sigma^2_{\eta_i}$ for $i > 1$.

Figure 1 shows the prior distributions over x_i $(i = 1, \ldots 3)$, for the three portfolios, for $\pi_0 = 0, 0.5$, and 1. $\pi_0 = 1$ corresponds to the pure *exact* CAPM hypothesis, $\pi_0 = 0$ corresponds to the unconstrained normal hypothesis, and $\pi_0 = 0.5$ corresponds to a situation in which the investor assigns a 50% probability to the validity of the (exact) CAPM. Note that while the marginal distribution for the market drift is normal, for SMB the prior distribution is clearly nonnormal and for HML it is actually bimodal for $\pi_0 = 0.5$.

We have distinguished between a prior based on an *exact* pricing model ("exact model prior") and a prior based on an *approximate* pricing model ("approximate model prior"). For the exact model prior, $\pi(\mathbf{q}, t)$ is the investor's assessment of the probability that the CAPM holds and his assessment is updated as described in part (2) of Theorem 2. For the approximate model prior, $\pi(\mathbf{q}, t)$ is simply a parameter of the posterior distribution, and

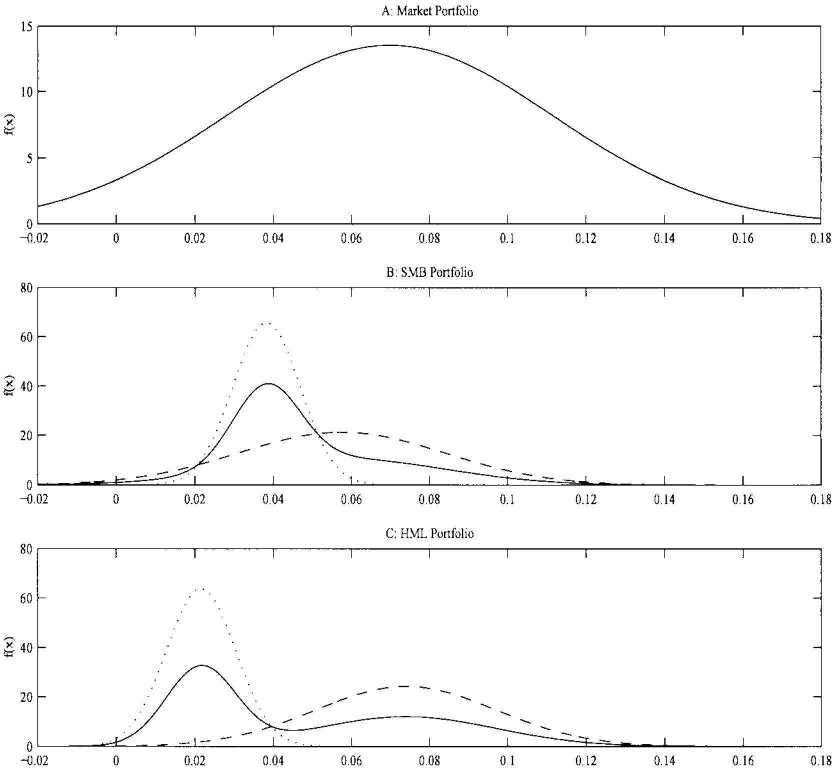

Figure 1
Plots of the marginal density functions of drift parameter for (A) market, (B) SMB, and (C) HML portfolios
The dotted lines $(\cdots)$ represents the pdf's for parameter estimates derived from historical data for the period July 1963–December 1991 under the exact CAPM hypothesis ($\pi_0 = 1$); the dashed lines (— — —) represent the pdf's for parameter estimates derived from the same historical data under the unconstrained normal model ($\pi_0 = 0$); the solid lines (__) represent the pdf's for parameter estimates derived from the same historical data under the mixed distribution ($\pi_0 = 0.5$).

the prior is updated as described in part (1) of Theorem 2. Thus the posterior mean vector, $\mathbf{m}(\mathbf{q}, t)$, is affected by whether or not the prior is based on an *exact* or an *approximate* model, and, as we shall see, this affects the portfolio policy.

Tables 4 and 5 reports results for an *approximate* mixed normal prior distribution; the prior distribution is a mixture of an *approximate* version of the CAPM and the unconstrained normal distribution. The mixing parameter π_0 is 0.5. σ_η is set equal to 1%, 2%, and 4%, $\text{cov}(\eta_i, \eta_j) = 0, i \neq j$, and horizons of 10 and 20 years are considered. Thus in forming his prior, the investor is assumed to place equal weight on parameter values from the (approximate) CAPM and the alternative hypothesis; under the approximate CAPM the model is assumed to predict the mean returns on the SMB and

Table 4
Optimal portfolio strategies and certainty equivalent returns

γ	Security	Prior mean risk premium	α_M	α_H	α^* Unconstrained	α^* Constrained	Certainty equivalent return Unconstrained	Certainty equivalent return Constrained
Panel A. Investor with a 20-year horizon and an *approximate* model ($\sigma_\eta = 2\%$) mixed (50:50) normal prior over the mean return vector								
2.0	Market	5.21%	1.42	−0.58	0.84	0.62	8.6%	6.3%
	SMB	2.16	0.52	−0.37	0.15	0.00		
	HML	1.84	2.20	−1.53	0.67	0.38		
3.0	Market	5.21%	0.94	−0.44	0.50	0.51	6.4%	5.4%
	SMB	2.16	0.34	−0.26	0.08	0.00		
	HML	1.84	1.46	−1.10	0.36	0.49		
4.0	Market	5.21%	0.71	−0.36	0.35	0.32	5.4%	4.7%
	SMB	2.16	0.26	−0.21	0.05	0.00		
	HML	1.84	1.10	−0.85	0.25	0.00		
5.0	Market	5.21%	0.57	−0.30	0.27	0.25	4.8%	4.3%
	SMB	2.16	0.21	−0.17	0.04	0.00		
	HML	1.84	0.88	−0.69	0.19	0.00		
Panel B. Investor with a 10-year horizon and an *approximate* ($\sigma_\eta = 2\%$) model mixed (50:50) normal prior over the mean return vector								
2.0	Market	5.21%	1.42	−0.35	1.07	0.66	9.4%	6.4%
	SMB	2.16	0.52	−0.25	0.27	0.00		
	HML	1.84	2.20	−1.04	1.16	0.34		
3.0	Market	5.21%	0.94	−0.28	0.66	0.54	6.9%	5.6%
	SMB	2.16	0.34	−0.18	0.16	0.00		
	HML	1.84	1.46	−0.79	0.67	0.46		
4.0	Market	5.21%	0.71	−0.23	0.48	0.48	5.8%	5.1%
	SMB	2.16	0.26	−0.15	0.11	0.00		
	HML	1.84	1.10	−0.63	0.47	0.52		
5.0	Market	5.21%	0.57	−0.19	0.38	0.44	5.1%	4.6%
	SMB	2.16	0.21	−0.12	0.08	0.00		
	HML	1.84	0.88	−0.52	0.36	0.54		

The prior distribution over the mean vector is a mixture of normal distributions that correspond to the approximate CAPM prior ($\sigma_\eta = 2\%$) and an unconstrained normal prior where the mean vectors and covariance matrices are estimated from monthly data from July 1963 to December 1991 as shown in Table 1; the risk-free interest rate is 3%. Figures shown in bold italics are for the CAPM prior; other figures relate to the unconstrained normal prior. α_M is the fraction of wealth allocated to each security in the optimal myopic portfolio; α_H is the additional allocation to each security to hedge against changes in parameter estimates; α^* is the optimal total allocation to each security. The certainty equivalent rate of return is the sure rate of return up to the horizon that would leave the investor as well off as having $1 of current wealth and 20 or 10 years, respectively, to invest with the currently assessed investment opportunity set. Columns labeled "constrained" impose the constraint that portfolio allocations (including riskless asset) be nonnegative.

HML portfolios with a standard error of 1, 2, or 4% per year.[23] In this case of an *approximate* prior, while the CAPM is used to construct the prior distribution, the posterior mean vector $\hat{\mu}_1(t)$ does not necessarily satisfy even the approximate CAPM relation for $t > 0$ since the posterior mean of η will in general be nonzero. Comparing Table 4A ($\sigma_\eta = 2\%$) with Table 2A, we see that the mixed prior ($\pi_0 = 0.5$) leads to certainty equivalent return

[23] Pastor (1998) considers prior distributions of possible mispricing of the SMB and HML portfolios within the CAPM which have standard deviations of up to 10% per year. However, Pastor implicitly restricts the mixing parameter π_0 to 1.0.

Table 5
Optimal portfolio strategies and certainty equivalent returns

γ	Security	Prior mean risk premium	α_M	α_H	α^* Unconstrained	α^* Constrained	Certainty equivalent return Unconstrained	Certainty equivalent return Constrained
Panel A. Investor with a 20-year horizon and an *approximate* ($\sigma_\eta = 1\%$) model mixed (50:50) normal prior over the mean return vector								
2.0	Market	5.21%	1.42	−0.58	0.84	0.62	8.1%	6.3%
	SMB	2.16	0.52	−0.38	0.14	0.00		
	HML	1.84	2.20	−1.58	0.62	0.38		
3.0	Market	5.21%	0.94	−0.45	0.49	0.51	6.4%	5.3%
	SMB	2.16	0.34	−0.27	0.07	0.00		
	HML	1.84	1.46	−1.14	0.32	0.49		
4.0	Market	5.21%	0.71	−0.36	0.35	0.44	5.1%	4.7%
	SMB	2.16	0.26	−0.21	0.05	0.00		
	HML	1.84	1.10	−0.88	0.22	0.56		
5.0	Market	5.21%	0.57	−0.30	0.27	0.24	4.6%	4.3%
	SMB	2.16	0.21	−0.18	0.03	0.00		
	HML	1.84	0.88	−0.72	0.16	0.00		
Panel B. Investor with a 20-year horizon and an *approximate* ($\sigma_\eta = 4\%$) model mixed (50:50) normal prior over the mean return vector								
2.0	Market	5.21%	1.42	−0.53	0.89	0.65	11.2%	6.5%
	SMB	2.16	0.52	−0.29	0.23	0.00		
	HML	1.84	2.20	−1.31	0.89	0.35		
3.0	Market	5.21%	0.94	−0.41	0.53	0.49	7.7%	5.5%
	SMB	2.16	0.34	−0.21	0.13	0.00		
	HML	1.84	1.46	−0.96	0.50	0.00		
4.0	Market	5.21%	0.71	−0.34	0.37	0.34	6.3%	4.8%
	SMB	2.16	0.26	−0.17	0.09	0.00		
	HML	1.84	1.10	−0.75	0.35	0.00		
5.0	Market	5.21%	0.57	−0.28	0.29	0.25	5.6%	4.4%
	SMB	2.16	0.21	−0.14	0.07	0.00		
	HML	1.84	0.88	−0.61	0.27	0.00		

The prior distribution over the mean vector is a mixture of normal distributions that correspond to the exact CAPM prior (σ_η = 1% or 4%), respectively and an unconstrained normal prior where the mean vectors and covariance matrices are estimated from monthly data from July 1963 to December 1991 as shown in Table 1; the risk-free interest rate is 3%. The mixing parameter is the investor's probability assessment that the approximate CAPM holds. Figures shown in bold italics are for the CAPM prior; other figures relate to the unconstrained normal prior. α_M is the fraction of wealth allocated to each security in the optimal myopic portfolio; α_H is the additional allocation to each security to hedge against changes in parameter estimates; α^* is the optimal total allocation to each security. The certainty equivalent rate of return is the sure rate of return up to the horizon that would leave the investor as well off as having \$1 of current wealth and 10 years to invest with the currently assessed investment opportunity set. Columns labeled "constrained" impose the constraint that portfolio allocations (including riskless asset) be nonnegative.

that is intermediate between the pure prior cases ($\pi_0 = 0, 1$). The mixed prior does not affect the prior mean market return, but significantly reduces the prior risk premia of SMB and HML relative to the pure unconstrained normal prior ($\pi_0 = 0$), since it shrinks them toward their CAPM values. This affects the composition of the myopic portfolio: as compared with the pure unconstrained normal prior, the allocation to the SMB and HML portfolios is roughly halved, while the allocation to the market portfolio also declines, but by only about 20%. The effect of the mixed prior on the hedge position is generally to raise the dollar size of the position, and in every case to raise the size of the position relative to holdings in the myopic portfolio. For example, for $\gamma = 3$, the hedge portfolio allocations to the market, SMB and HML rise from 22%, 7%, and 34% of the myopic allocations to 47%, 76%, and 75%, respectively.

The importance of taking account of learning, which is manifest in the size of the hedge portfolio, is striking. For example, for $\gamma = 3$, the myopic portfolio allocations are 0.94, 0.34, 1.46. These drop to 0.50, 0.08, and 0.36 when account is taken of learning and the horizon is 20 years. Even when the horizon is only 10 years (Table 4B), the corresponding figures are 0.66, 0.16, and 0.67, so that the optimal portfolio allocations are reduced relative to the myopic allocations by 30%, 53%, and 54%, respectively.

Table 5A and B shows the effect of changing the degree of the CAPM approximation parameter, σ_η. The smaller the value of σ_η, the tighter is the prior about the CAPM, the larger the size of the hedge positions, and the lower is the certainty equivalent return.

Table 3A and B reports the results when the *exact* version of the CAPM is used to construct the prior distribution: in this case the posterior mean vector $\hat{\mu}_1(t)$ satisfies the CAPM relation for $t > 0$. Note that whenever an approximate prior is used, $\hat{\mu}_1(t) \to \hat{\mu}_2(t) \to \mathbf{q}(t)$ for large t and $\pi(t) \to 0.5$. The effect of an exact CAPM prior distribution is to impose a restriction in updating $\hat{\mu}_1(t)$, which prevents it from converging to $\mathbf{q}(t)$ unless the CAPM is supported by the realized returns. This makes it possible for $\pi(t)$ to converge to 0, in which case the influence of the CAPM prior on the posterior distribution is eliminated. The effect of imposing the exact CAPM prior is dramatic. The certainty equivalent rate of return (for $\gamma = 3$, $T = 20$ years) rises from 6.4% when $\sigma_\eta = 1\%$ to 7.1% for the exact model prior ($\sigma_\eta = 0$). We conjecture that this is because resolution of parameter uncertainty is accelerated by the restrictions on parameter updating that are introduced by the exact model prior.

The sensitivity of the investor's expected utility to the instantaneous rate of return is greatly increased under the exact model prior. This is apparent in the much larger hedge positions, particularly in the (CAPM anomalous) SMB and HML portfolios [see Equation (47)]. For example, for $\gamma = 3$, the hedge positions in the three portfolios increase from $(0.45, 0.27, 1.14)$ when $\sigma_\eta = 1\%$ to $(0.53, 0.57, 1.77)$ for the exact model prior ($\sigma_\eta = 0$). This increased sensitivity of the investor's expected utility to the instantaneous asset returns is due to the more rapid resolution of parameter uncertainty under the exact model prior.

Note that the myopic allocation with the model (50:50) mixed normal prior shown in Tables 3 and 4 is a simple average of the allocations under the CAPM prior and the unconstrained normal prior shown in Table 2: this is because the myopic portfolio is determined by the prior mean, and the prior mean under the mixed prior is a simple average of the two pure priors. However, the hedge portfolio under the mixed prior is very different from the average of the hedge portfolios under the two pure priors; indeed the size of the hedge positions for the mixed prior is greater than that for either of the pure priors. This is because the mixture of normals prior is associated with another layer of uncertainty—model uncertainty—in addition to the parameter uncertainty considered under the pure priors. This implies much greater uncertainty about the mean vector than either of the pure priors, and the additional allocation to the hedge position may be roughly interpreted as the effect of model uncertainty. Indeed, the uncertainty about the mean vector for the mixed normal prior is so great that the optimal investments in the SMB and HML portfolios actually become negative for an investor with a 20-year horizon. The hedge positions are smaller for the approximate model (50:50) mixed normal prior shown in Tables 4 and 5 than for the exact model-based prior in Table 3; this is consistent with the result in Tables 4 and 5 that the size of the hedge position is decreasing in the approximation parameter. As seen in Equation (47), the size of the hedge position depends on the elasticity of the indirect utility function with respect to the realized returns; this is evidently greater for the exact model-based prior, because the rate of learning is greater the more precise are the two distributions underlying the mixed prior.

For the sake of completeness, the optimal portfolio strategies were also calculated subject to the constraints that there be no short positions or borrowing. When these constraints are imposed it is no longer possible to distinguish simply between the myopic and hedge components of the portfolios. Therefore in Tables 2–5 we report, under the columns headed "constrained," the optimal portfolios and the corresponding certainty equivalent rate of return. In almost all cases except for the pure CAPM prior ($\pi_0 = 1$), the constraints significantly reduce the certainty equivalent returns, and it is striking that for neither the 10-year nor the 20-year horizon is it optimal to hold any position in the SMB portfolio; and while it is generally optimal to take a long position in the HML portfolio, the general effect of imposing the constraints is to increase the relative importance of the market component of the optimal portfolio.

Table 6 shows the optimal portfolios as a function of the investor's prior probability that the exact version of the CAPM holds, π_0, for 10- and 20-year horizons. Surprisingly, even a 1% deviation from the dogmatic prior that the CAPM does not hold ($\pi_0 = 0$) can have a dramatic effect on the optimal portfolio composition. For a 20-year horizon the 1% possibility that the CAPM holds decreases the optimal holding in the market portfolio from 87% to 68%, in the SMB portfolio from 65% to 29%, and in the HML portfolio from 182% to 143%.

5. Simulated Portfolio Strategies

In order to illustrate the differences in portfolio strategies induced by the different priors, normal prior distributions for the asset drifts were calibrated using historical data on the monthly portfolio returns for the period

Table 6
Optimal portfolio strategies for an investor with an *exact* model mixed normal prior over the mean return vector under different prior probability that the CAPM holds

	Unconstrained				CAPM
π_0	0.00	0.01	0.50	0.99	1.00
Panel A. Horizon: 20 years					
Market					
Myopic	1.19	1.18	0.94	0.71	0.70
Hedge	−0.32	−0.50	−0.44	−0.29	−0.21
Optimal	**0.87**	**0.68**	**0.50**	**0.42**	**0.49**
SMB					
Myopic	0.69	0.68	0.34	0.01	0.00
Hedge	−0.04	−0.39	−0.27	−0.10	0.00
Optimal	**0.65**	**0.29**	**0.07**	**−0.10**	**0.00**
HML					
Myopic	2.94	2.91	1.46	0.03	0.00
Hedge	−1.12	−1.48	−1.13	−0.32	0.00
Optimal	**1.82**	**1.43**	**0.33**	**−0.29**	**0.00**
Panel B. Horizon: 10 years					
Market					
Myopic	1.19	1.18	0.94	0.71	0.70
Hedge	−0.18	−0.28	−0.27	−0.15	−0.09
Optimal	**1.01**	**0.91**	**0.67**	**0.55**	**0.57**
SMB					
Myopic	0.69	0.68	0.34	0.01	0.00
Hedge	−0.07	−0.21	−0.18	−0.04	0.00
Optimal	**0.62**	**0.47**	**0.16**	**−0.04**	**0.00**
HML					
Myopic	2.94	2.91	1.46	0.03	0.00
Hedge	−0.64	−0.79	−0.77	−0.14	0.00
Optimal	**2.30**	**2.12**	**0.69**	**−0.11**	**0.00**

The prior distribution over the mean vector is a mixture of normal distributions that correspond to the exact CAPM prior ($\sigma_\eta = 0\%$) and an unconstrained normal prior where the mean vectors and covariance matrices are estimated from monthly data from July 1963 to December 1991 as shown in Table 1; the risk-free interest rate is 3%. The mixing parameter is the investor's probability assessment that the approximate CAPM holds. π_0 is the prior probability that the approximate version of the CAPM holds. The investor's risk aversion parameter, γ, is 3.

1959–1978, both with and without imposing the CAPM. The parameters of these distributions are given in panel B of Table 1. An *exact* model prior distribution was constructed by combining the exact CAPM with the unconstrained normal distribution with a value of $\pi_0 = 0.5$. An *approximate* model prior distribution was constructed by combining an approximate CAPM ($\sigma_\eta = 2\%$) with the unconstrained normal distribution with $\pi_0 = 0.5$. Note that for the *exact* model prior, π_t corresponds to the investor's posterior probability assessment that the CAPM holds; it has no such interpretation in the *approximate* model prior.

Then, for each of these prior distributions, the control problem, [Equation (46)] was solved[24] for an initial horizon $T = 20$ years and $\gamma = 3$.

[24] The problems differ for the two priors in the expressions that are used for $\hat{\mu}_1(\mathbf{q}, t)$, $\widehat{\boldsymbol{\Sigma}}_1(t)$, and $\pi(\mathbf{q}, t)$ that enter $\mathbf{m}(\mathbf{q}, t)$ in the control problem [Equation (46)]. For the *approximate* model prior the expressions in

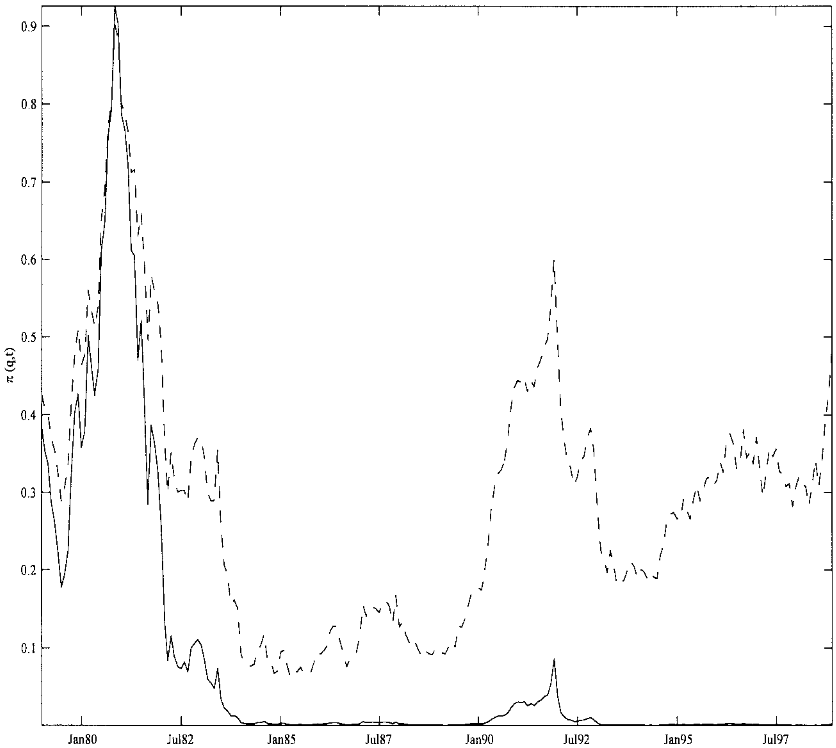

Figure 2
Plots of the time series of the mixing parameter of the posterior distribution
The figure shows the time series of $\pi(\mathbf{q}, t)$, the mixing parameter of the investor's posterior distribution which depends on the prior and the realized vector of asset returns, $\mathbf{q}$, over the period January 1979–December 1998. The prior distribution is obtained by combining the CAPM-based prior and the unconstrained normal prior based on the historical parameter estimates for the period 1959–1978 reported in panel B of Table 1 with a mixing parameter $\pi_0 = 0.5$. The dashed lines (— — —) correspond to the *approximate* ($\sigma_\eta = 2\%$) model prior in which $\pi(\mathbf{q}, t)$ is simply a mixing parameter. The solid lines (_____) correspond to the *exact* model prior in which $\pi(\mathbf{q}, t)$ is the investor's posterior probability that CAPM holds.

Next the realized average return vector $\mathbf{q}(t)$ was calculated for every month from January 1979 to December 1998, and was used to calculate both the posterior mean vector $\mathbf{m}(t)$ and the mixing parameter $\pi(\mathbf{q}, t)$ under both the *exact* model prior and the *approximate* model prior. Finally, values of $\mathbf{q}(t)$ were used to construct the optimal portfolio vectors $\alpha^*(\mathbf{q}, t)$ under the two different priors.

Figure 2 plots the realized values of $\pi(\mathbf{q}, t)$ under both the approximate and the exact model priors. It is interesting to note that for both priors $\pi(\mathbf{q}, t)$ initially rises rapidly to above 90%; this appears to be due to low returns

the first part of Theorem 2 apply; for the *exact* model prior the corresponding expressions from the second part of the theorem are used.

realized on the HML and SMB portfolios over the period 1979–1981 which bring their posterior means close to the CAPM predictions. However, from mid-1981 $\pi(\mathbf{q}, t)$ drops rapidly for both priors, and by 1984 for the investor using the *exact* model prior, the probability that the CAPM restriction holds is very close to zero, although it rebounds to around 6% briefly in 1991, once again reflecting low returns on HML. It is noticeable that π is much more volatile under the approximate model prior than it is under the exact model prior, where it sinks to zero. The reason for this is that under the *exact* model prior the posterior mean vector, $\hat{\mu}_1(t)$, always satisfies the CAPM restriction, whereas under the *approximate* model prior it evolves toward the realized sample mean return vector $\mathbf{q}(t)$; as a result the differences between $\hat{\mu}_1(\mathbf{q}, t)$ and $\hat{\mu}_2(\mathbf{q}, t)$ tend to disappear over time under the *approximate* model prior but not under the *exact* model prior.

Figure 3 plots the posterior means $m_i(\mathbf{q}, t)$ under the two priors for the three portfolios ($i = 1, \ldots 3$). Whether the prior is *approximate* or *exact* makes no difference for the market mean return, but at times relatively large differences for the SMB and HML portfolio: these differences are mainly due to differences in the mixing parameter $\pi(\mathbf{q}, t)$ under the two priors as seen in Figure 2.

Figure 4 plots the allocations to the three portfolios under the two priors. Note first that, although the expected market return is independent of the prior, the allocation to the market portfolio does depend on the prior slightly. The steep increase in the allocation to the market portfolio through the sample period for both priors is due in part to the increasing posterior mean, which is visible in Figure 3; but it is also due to the reduced size of the short position in the hedge portfolio as the horizon is approached. In January 1979 the allocation to the market portfolio under the *exact* prior is equal to 28.6% of wealth. This consists of a long position in the myopic portfolio of 61.1% of wealth, offset by a short position in the hedge portfolio of 32.5% of wealth; this hedge position goes to zero as the horizon approaches.

The choice of prior makes a large difference to the allocation to the SMB portfolio at the start of the sample period. This is entirely due to the differences in the hedge portfolio; as the uncertainty is resolved more quickly with the *exact* model prior, the absolute size of the (short) hedge position (76.1% of wealth) is larger than under the *approximate* model prior (37.3% of wealth). A similar phenomenon is apparent for the HML portfolio.

6. Conclusion

In this article we have shown how to calculate the optimal dynamic investment strategy when asset returns follow diffusion processes with constant coefficients but the drift coefficient is unknown to the investor and must be inferred from the observed returns. When the prior distribution over the asset price drift is normal, the posterior is also normal, and we show that for an

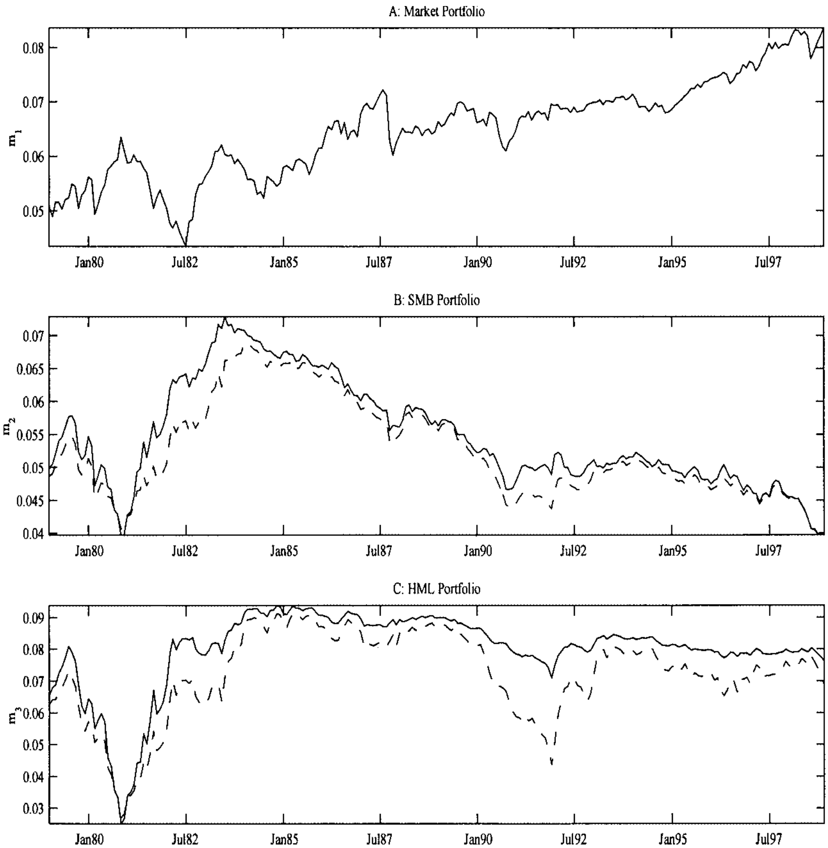

Figure 3

Plots of the time series of the posterior asset drift parameters under the two different priors for (A) market, (B) SMB, and (C) HML portfolios

The figure shows the time series of $\mathbf{m}(\mathbf{q}, t)$, the mean of the investor's posterior distribution which depends on the prior and the realized vector of asset returns, $\mathbf{q}$, over the period January 1979–December 1998. The prior distribution is obtained by combining the CAPM-based prior and the unconstrained normal prior based on the historical parameter estimates for the period 1959–1978 reported in panel B of Table 1 with a mixing parameter $\pi_0 = 0.5$. The dashed lines (— — —) correspond to the *approximate* model prior in which an approximate version of the CAPM ($\sigma_\eta = 2\%$) is used. The solid lines (_____) correspond to the *exact* model prior.

investor with power utility the expected utility and optimal portfolio strategy can be determined as the solution to a recursive set of ordinary differential equations. More interesting is the case in which the prior is a mixture of normal distributions; in this case, following earlier results of Benes and Karatzas (1983) and Detemple (1991) we show that the investor's perceived investment opportunities can be summarized in terms of the realized asset returns, and solve the resulting control problem numerically. We argue that the mixture of normal priors is a natural one for an investor who places some weight on

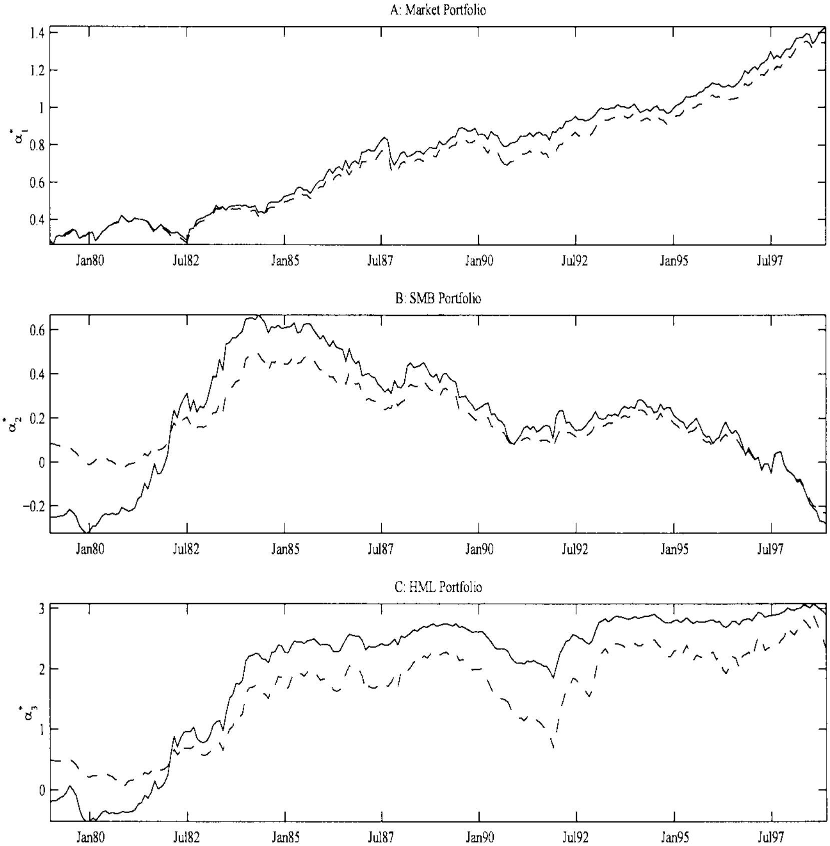

Figure 4
Plots of the time series of optimal portfolio holdings under the two different priors for (A) Market, (B) SMB, and (C) HML portfolios
The figure shows the time series of $\boldsymbol{\alpha}^*(\mathbf{q}, t)$, the investor's optimal portfolio holdings, which depend on the prior and the realized vector of asset returns, $\mathbf{q}$, over the period January 1979–December 1998. The prior distribution is obtained by combining the CAPM-based prior and the unconstrained three-factor prior based on the historical parameter estimates for the period 1959–1978 reported in Panel B of Table 1 with a mixing parameter $\pi_0 = 0.5$. The dashed lines (— — —) correspond to the *approximate* model prior in which an approximate version of the CAPM ($\sigma_\eta = 2\%$) is used. The solid lines (_____) correspond to the *exact* model prior.

a particular asset pricing model which constrains relative asset returns, but also allows the possibility that an empirical finding about asset returns may represent a genuine anomaly. Such an investor must take account of the fact that he will learn more about the anomaly over his investment horizon. We solve the investment problem of an investor who places some weight on the CAPM and some weight on the empirical findings of Fama and French that the HML and SMB portfolios have returns that are anomalous with respect to the CAPM.

Our findings are striking. Consider, for example, an investor with a coefficient of relative risk aversion of 3 who would place 70% of his wealth in the market portfolio if he knew that the CAPM was valid and that the market risk premium was equal to the average market excess return over the period 1963–1991. This investor, taking account of the fact that the market risk premium is only estimated rather than known, will reduce his investment in the market portfolio to 49% of his wealth if his horizon is 20 years, and 57% if his horizon is 10 years. For the 10-year investor, the investment opportunity set with its estimated market risk premium offers a certainty equivalent rate of return of 4.7%, which is 1.7% above the riskless interest rate; that is, the investor would be as well off with a single riskless asset offering 4.7% return as with the actual investment opportunity set. The same investor, if he is sure that the CAPM does not hold, and uses the historical data to estimate the returns on the three portfolios, will conclude that the capital market offers a certainty equivalent rate of return of 13.1%, *even after taking account of the fact that the mean returns on the portfolio are estimated.* In other words, he will conclude that the dollar bills lying on the sidewalks of Wall Street are of pretty high denomination!

We also demonstrate that model uncertainty, uncertainty about the genuineness of the anomaly, can have a major effect on portfolio choice. Consider, for example, an investor with an horizon of 20 years and a coefficient of relative risk aversion of 3 who would place 49% of his wealth in the market portfolio if he knew that the CAPM was valid (and take zero positions in the SMB and HML portfolios). The same investor, if he were only 99% sure that the CAPM holds, would reduce his investment in the market portfolio to 42% of his wealth and take short positions in the SMB and HML portfolios of 10% and 29%, respectively.

Appendix

Proof of Theorem 1. Consider the mixture of normal prior distributions over the $(n \times 1)$ mean vector, $\mathbf{x}$:

$$f_0(\mathbf{x}) = \frac{\pi_0}{2\pi^{\frac{n}{2}}|\Sigma_1|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu_1)'\Sigma_1^{-1}(\mathbf{x}-\mu_1)\right\} + \frac{1-\pi_0}{2\pi^{\frac{n}{2}}|\Sigma_2|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu_2)'\Sigma_2^{-1}(\mathbf{x}-\mu_2)\right\}. \tag{A.1}$$

Then it follows from Lemma 1, Equation (2) that the posterior distribution over the mean vector at time t, $f_t(\mathbf{x})$, is given by

$$f_t(\mathbf{x}) \equiv f(\mathbf{x}; q, t) = \frac{\exp\left\{-\frac{1}{2}t\mathbf{x}'\Omega^{-1}(\mathbf{x}-2q)\right\} f_0(\mathbf{x})}{\int \exp\left\{-\frac{1}{2}(t\mathbf{x}'\Omega^{-1}(\mathbf{x}-2q)\right\} f_0(\mathbf{x})\,\mathbf{dx}}, \tag{A.2}$$

where $q_t = \frac{1}{t}[\ln(P_t) - \ln(P_0)]$.

Then, substituting for $f_0(\mathbf{x})$ from Equation (A.1) in Equation (A.2), $f_t(\mathbf{x})$ may be written as

$$f_t(\mathbf{x}) \equiv \frac{\xi(t, \mathbf{x}, q)}{\int \xi(t, \mathbf{x}, q)\, \mathbf{dx}}, \tag{A.3}$$

where

$$\xi(\mathbf{x}, q, t) \equiv \frac{\pi_0}{(2\pi)^{\frac{n}{2}}|\Sigma_1|^{\frac{1}{2}}} \exp^{\frac{-1}{2}[\mathbf{x}'\Sigma_1^{-1}\mathbf{x}-2\mathbf{x}'\Sigma_1^{-1}\mu_1+\mu_1'\Sigma_1^{-1}\mu_1+t\mathbf{x}'\Omega^{-1}(\mathbf{x}-2tq)]} + \frac{1-\pi_0}{(2\pi)^{\frac{n}{2}}|\Sigma_2|^{\frac{1}{2}}} \exp^{\frac{-1}{2}[\mathbf{x}'\Sigma_2^{-1}\mathbf{x}-2\mathbf{x}'\Sigma_2^{-1}\mu_2+\mu_2'\Sigma_2^{-1}\mu_2+t\mathbf{x}'\Omega^{-1}(\mathbf{x}-2tq)]}. \tag{A.4}$$

Simplifying Equations (A.3) and (A.4), we obtain

$$f_t(\mathbf{x}) = \pi(t, q)\frac{1}{(2\pi)^{\frac{n}{2}}|\widehat{\Sigma}_1(t, q)|^{\frac{1}{2}}} \exp^{-\frac{1}{2}\{(\mathbf{x}-\hat{\mu}_1(t,q))'\widehat{\Sigma}_1^{-1}(\mathbf{x}-\hat{\mu}_1(t,q))\}} + (1-\pi(t, q))\frac{1}{(2\pi)^{\frac{n}{2}}|\widehat{\Sigma}_2(t, q)|^{\frac{1}{2}}} \exp^{-\frac{1}{2}\{(\mathbf{x}-\hat{\mu}_2(t,q))'\widehat{\Sigma}_2^{-1}(\mathbf{x}-\hat{\mu}_2(t,q))\}}, \tag{A.5}$$

where $\hat{\mu}_i(t, q) = (\Sigma_i^{-1} + t\Omega^{-1})^{-1}(\Sigma_i^{-1}\mu_i + t\Omega^{-1}q)$, and

$$\pi(t, q) = \frac{\pi_0|\Sigma_1^{-1} + t\Omega^{-1}|^{\frac{1}{2}}|\Sigma_1|^{\frac{-1}{2}} \exp^{\{\frac{-1}{2}(\mu_1'\Sigma_1^{-1}\mu_1-\hat{\mu}_1'\widehat{\Sigma}_1^{-1}\hat{\mu}_1)\}}}{\left(\begin{array}{c}\pi_0|\Sigma_1^{-1} + t\Omega^{-1}|^{\frac{1}{2}}|\Sigma_1|^{\frac{-1}{2}} \exp^{\{\frac{-1}{2}(\mu_1'\Sigma_1^{-1}\mu_1-\hat{\mu}_1'\widehat{\Sigma}_1^{-1}\hat{\mu}_1)\}} + \\ (1-\pi_0)|\Sigma_2^{-1} + t\Omega^{-1}|^{\frac{1}{2}}|\Sigma_2|^{\frac{-1}{2}} \exp^{\{\frac{-1}{2}(\mu_2'\Sigma_2^{-1}\mu_2-\hat{\mu}_2'\widehat{\Sigma}_2^{-1}\hat{\mu}_2)\}}\end{array}\right)}. \tag{A.6}$$

Then Equation (17) for $\mathbf{G}(\mathbf{q}, t)$ follows from substituting for $f_0(\mathbf{x})$ from Equation (7) in Equation (5).

Proof of Theorem 2. It is convenient to recognize explicitly that under distribution one, the approximate factor pricing hypothesis, the investor is updating $\mathbf{x}_1$, the drift term of the factors, and η, the factor model deviations.

Define the observable signals of the model deviations by

$$d\hat{S}_i \equiv dS_i - \sum_{k=1}^{K}\beta_{ik}dS_k - c_i\, dt = \eta_i\, dt + \left(\sigma_i - \sum_{k=1}^{K}\beta_{ik}\sigma_k\right) dz \tag{A.7}$$
$$i = K+1, \ldots, N.$$

Then partition μ_1, the vector of mean drifts under the approximate factor pricing hypothesis, into its first K elements, $\mu_{\mathbf{x}_1,1}$, and its remaining $N-K$ elements, $\mu_{\mathbf{x}_2,1}$, and partition $\mathbf{q}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{\Omega}$ conformably.

The vector of factor returns and signals, $[dS_1, \ldots, dS_K, d\hat{S}_{K+1}, \ldots, d\hat{S}_N]'$, has a variance-covariance matrix

$$\widehat{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22} - \boldsymbol{\Omega}_{21}\boldsymbol{\Omega}_{11}^{-1}\boldsymbol{\Omega}_{12}, \end{bmatrix},$$

where Ω_{11} is the variance-covariance matrix of the K factor returns; and Ω_{12} and Ω_{22} are defined similarly.

The prior under the approximate factor pricing hypothesis may be written as

$$[\mathbf{x}_1, \eta]' \sim N\left(\begin{bmatrix} \mu_{\mathbf{x}_1,1} \\ \mu_{\eta,1} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{x}_1\mathbf{x}_1,1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\eta,1} \end{bmatrix}\right),$$

where $\mu_{\eta,1} = \mathbf{0}$ and $\boldsymbol{\Sigma}_{\eta,1}$ is a diagonal matrix with $\sigma_{\eta_i}^2$ as the (i, i) element.

The prior distribution over the mean vector $\mathbf{x} \equiv [\mathbf{x}_1, \mathbf{x}_2]'$ under the alternative hypothesis can be written as a distribution over $[\mathbf{x}_1, \eta]'$, which is

$$N\left(\begin{bmatrix}\mu_{\mathbf{x}_1,2}\\ \mu_{\eta,2}\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{x}_1\mathbf{x}_1,2} & \boldsymbol{\Sigma}_{\mathbf{x}_1\eta,2}\\ \boldsymbol{\Sigma}_{\eta\mathbf{x}_1,2} & \boldsymbol{\Sigma}_{\eta,2}\end{bmatrix}\right),$$

where

$$\mu_{\eta,2} \equiv \mu_{\mathbf{x}_2,2} - \beta\mu_{\mathbf{x}_1,2} - \mathbf{c}, \tag{A.8}$$

$$\boldsymbol{\Sigma}_{\eta,2} \equiv \boldsymbol{\Sigma}_{\mathbf{x}_2\mathbf{x}_2,2} + \beta\boldsymbol{\Sigma}_{\mathbf{x}_1\mathbf{x}_1,2}\beta' - 2\beta\boldsymbol{\Sigma}_{\mathbf{x}_1\mathbf{x}_2,2}, \tag{A.9}$$

$$\boldsymbol{\Sigma}_{\eta\mathbf{x}_1,2} \equiv \boldsymbol{\Sigma}_{\mathbf{x}_2\mathbf{x}_1,2} - \beta\Sigma_{\mathbf{x}_1\mathbf{x}_1,2}, \tag{A.10}$$

$$\boldsymbol{\Sigma}_{\mathbf{x}_1\eta,2} = \boldsymbol{\Sigma}'_{\eta\mathbf{x}_1,2}. \tag{A.11}$$

Therefore Theorem 1 can be applied to derive the posterior for $[\mathbf{x}_1, \eta]'$. Part (1) of Theorem 2 follows from the relation between $\mathbf{x}_2$ and η:

$$\eta = \mathbf{x}_2 - \beta\mathbf{x}_1 - \mathbf{c}. \tag{A.12}$$

Part (2) of Theorem 2 follows by letting $\max_i(\sigma_{\eta_i}) \to 0$, and noting that $\hat{\mu}_{\eta,1} \to 0$ at the limit.

The interpretation of $\pi(\mathbf{q}, t)$ as the probability that the factor pricing model holds can be derived from the posterior distribution of $\mathbf{x}$:

$$\begin{aligned} f(\mathbf{x}|\mathbf{q}, t) &\propto L(\mathbf{q}|\mathbf{x})f_0(\mathbf{x}) \\ &= \pi_0 L(\mathbf{q}|\mathbf{x})f_0(\mathbf{x}|H1) + (1-\pi_0)L(\mathbf{q}|\mathbf{x})f_0(\mathbf{x}|H2) \\ &= \pi_0 p(\mathbf{q}, \mathbf{x}|H1) + (1-\pi_0)p(\mathbf{q}, \mathbf{x}|H2) \\ &= \pi_0 p(\mathbf{q}|H1)f(\mathbf{x}|\mathbf{q}, H1) + (1-\pi_0)p(\mathbf{q}|H2)f(\mathbf{x}|\mathbf{q}, H2) \\ &\propto p(H1|\mathbf{q})f(\mathbf{x}|\mathbf{q}, H1) + p(H2|\mathbf{q})p(\mathbf{q}|H2)f(\mathbf{x}|\mathbf{q}, H2). \end{aligned} \tag{A.13}$$

Substituting in the prior distributions under hypotheses of $H1$ and $H2$ and the likelihood function, we find that $p(H1|\mathbf{q}) \propto A_1(\mathbf{q}, t) \propto \pi(\mathbf{q}, t)$ and $p(H2|\mathbf{q}) \propto A_2(\mathbf{q}, t) \propto (1 - \pi(\mathbf{q}, t))$. Therefore $\pi(\mathbf{q}, t)$ is the posterior probability that hypothesis 1 (the factor pricing model) holds.

References

Ball, R., and P. Brown, 1968, "An Empirical Examination of Accounting Income Numbers," *Journal of Accounting Research*, 6, 159–178.

Banz, R., 1981, "The Relationship Between Mean Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9, 3–18.

Barber, B. M., and J. D. Lyon, 1997, "Firm Size, Book-to-Market Ratio, and Security Returns: A Holdout Sample of Financial Firms," *Journal of Finance*, 52, 875–883.

Benes, V. E., and I. Karatzas, 1983, "Estimation and Control for Linear, Partially Observable Systems with Non-Gaussian Initial Distribution," *Stochastic Processes and Their Applications*, 14, 233–248.

Berk, J., 1995, "A Critique of Size-Related Anomalies," *Review of Financial Studies*, 8, 275–286.

Bernard, V. L., and J. K. Thomas, 1990, "Evidence That Stock Prices Do Not Fully Reflect the Implications of Current Earnings for Future Earnings," *Journal of Accounting and Economics*, 13, 305–340.

Black, F., 1993, "Beta and Return," *Journal of Political Economy*, 20, 8–18.

Breen, W. J., and R. A. Korajczyk, 1995, "On Selection Biases in Book-to-Market Based Tests of Asset Pricing Models," working paper, Northwestern University.

Brennan, M. J., 1998, "The Role of Learning in Dynamic Portfolio Decisions," *European Finance Review*, 1, 295–306.

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

Chan, L. K. C., N. Jegadeesh, and J. Lakonishok, 1995, "Evaluating the Performance of Value versus Glamor Stocks: The Impact of Selection Bias," *Journal of Financial Economics*, 38, 269–296.

Daniel, K., D. Hirshleifer and A. Subrahmanyam, 1998, "Investor Psychology and Security Market Under- and Overreactions," *Journal of Finance*, 53, 1839–1885.

Detemple, J. B., 1986, "Asset Pricing in an Economy with Incomplete Information," *Journal of Finance*, 61, 383–392.

Detemple, J. B., 1991, "Further Results on Asset Pricing with Incomplete Information," *Journal of Economic Dynamics and Control*, 15, 425–453.

Dimson, E., and P. Marsh, 1999, "Murphy's Law and Market Anomalies," *Journal of Portfolio Management*, 23, 53–69.

Fama, E. F., and K. R. French, 1992, "The Cross-Section of Expected Returns," *Journal of Finance*, 47, 427–465.

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

Fama, E. F., and K. R. French, 1995, "Size and Book-to-Market Factors in Earnings and Returns," *Journal of Finance*, 50, 131–155.

Fama, E. F., and K. R. French, 1996, "Multi-factor Explanations of Asset Pricing Anomalies," *Journal of Finance*, 51, 55–84.

Fama, E. F., and K. R. French, 1998, "Values versus Growth: The International Evidence," *Journal of Finance*, 53, 1975–2000.

Feldman, D., 1992, "Logarithmic Preferences, Myopic Decisions and Incomplete Information," *Journal of Financial and Quantitative Analysis*, 27, 619–630.

Gennotte, G., 1986, "Optimal Portfolio Choice under Incomplete Information," *Journal of Finance*, 61, 733–749.

Haugen, R., 1995, *The New Finance: The Case against Efficient Markets*, Prentice-Hall, Englewood Cliffs, NJ.

Hawawini, G., and D. Keim, 1999, "The Cross-Section of Common Stock Returns: A Synthesis of the Evidence and the Explanations," in Donald Keim and William Ziemba (eds.), *Security Market Imperfections in World Wide Equity Markets*, Cambridge: Cambridge University Press.

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

Jensen, M. C., 1978, "Some Anomalous Evidence Regarding Market Efficiency," *Journal of Financial Economics*, 6, 95–101.

Knez, P. J., and M. J. Ready, 1997, "On the Robustness of Size and Book-to-Market in Cross-Sectional Regressions," *Journal of Finance*, 52, 1355–1382.

Kothari, S. P., J. Shanken, and R. G. Sloan, 1995, "Another Look at the Cross-Section of Expected Stock Returns," *Journal of Finance*, 50, 185–224.

Lakonishok, J., A. Shleifer, and R. Vishny, 1994, “Contrarian Investment, Extrapolation, and Risk,” *Journal of Finance*, 49, 1541–1578.

La Porta, R., J. Lakonishok, A. Shleifer, and R. Vishny, 1997, “Good News for Value Stocks: Further Evidence on Market Efficiency,” *Journal of Finance*, 52, 859–874.

Lo, A. W., and A. C. MacKinlay, 1990, “Data Snooping Biases in Tests of Financial Asset Pricing Models,” *Review of Financial Studies*, 3, 431–467.

Loughran, T., 1997, “Book-to-Market Across Firm Size, Exchange, and Seasonality: Is There an Effect?,” *Journal of Financial and Quantitative Analysis*, 32, 249–268.

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

Pastor, L., 1999, “Portfolio Selection and Asset Pricing Models,” working paper, University of Pennsylvania.

Pastor, L., and R. F. Stambaugh, 1999, “Comparing Asset Pricing Models: An Investment Perspective,” working paper, University of Pennsylvania.

Reinganum, M. R., 1981, “Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings’ Yields and Market Values,” *Journal of Financial Economics*, 9, 19–46.

Roll, R., 1994, “What Every CFO Should Know about Scientific Progress in Financial Economics: What is Known and What Remains to be Resolved,” *Financial Management*, 23, 69–75.

Ross, S. A., 1973, “The Arbitrage Theory of Capital Asset Pricing,” *Journal of Economic Theory*, 13, 341–360.

Rosenberg, B., Reid, K., and R. Lanstein, 1984, “Persuasive Evidence of Market Inefficiency,” *Journal of Portfolio Management*, 11, 9–17.

Williams, J. T., 1977, “Capital Asset Prices with Heterogeneous Beliefs,” *Journal of Financial Economics*, 5, 219–239.

Xia, Y., 2001, “Learning about Predictability: The Effect of Parameter Uncertainty on Dynamic Asset Allocation,” *Journal of Finance*, 56, 205–246.