A Shrinkage Approach to Model Uncertainty and Asset Allocation

Zhenyu Wang*

Columbia University
Graduate School of Business

July 1, 2002†

Abstract

Aversion to uncertainty about an asset-pricing model may lead investors to hold a portfolio that is inefficient for probability distributions estimated from any prior beliefs in the model. The shrinkage approach explicitly shows how prior beliefs affect the estimated distribution and helps to solve the optimal portfolio for uncertainty-averse investors. Using the shrinkage approach, U.S. investors' home bias is measured to be 0.9 on a scale from 0 to 1. Uncertainty about the world equilibrium model does not justify the strong home bias because the diversified world market portfolio is still optimal for uncertainty-averse investors.

JEL Classification: G11, G12, C11

Keywords: model uncertainty, portfolio choice, home bias, Bayesian analysis, robust analysis

*For helpful comments, I thank Torben Andersen, Andrew Ang, Geert Bekaert, Mark Carhart, Fangruo Chen, Silverio Foresi, Larry Glosten, Rick Green, Bob Hodrick, Michael Johannes, Bob Litterman, Lubos Pastor, David Robinson, Tano Santos, Rob Stambaugh, and seminar participants at Columbia University, Northwestern University, Lehman Brothers, Goldman Sachs, University of Minnesota, the 2001 American Finance Association Meeting, and the Federal Reserve Bank of Atlanta. I am especially grateful to Gur Huberman and Ravi Jagannathan for their encouragement and suggestions since the beginning of the project. Adam Kolasinski provided excellent research assistance. Corresponding address: 3022 Broadway, Uris Hall 421, New York, NY 10027. Email: zw24@columbia.edu

†Some of the earlier versions have been circulated since April 10, 1999 under the title "A shrinkage theory of asset allocation using asset-pricing models."

A Shrinkage Approach to Model Uncertainty and Asset Allocation

Abstract

Aversion to uncertainty about an asset-pricing model may lead investors to hold a portfolio that is inefficient for probability distributions estimated from any prior beliefs in the model. The shrinkage approach explicitly shows how prior beliefs affect the estimated distribution and helps to solve the optimal portfolio for uncertainty-averse investors. Using the shrinkage approach, U.S. investors' home bias is measured to be 0.9 on a scale from 0 to 1. Uncertainty about the world equilibrium model does not justify the strong home bias because the diversified world market portfolio is still optimal for uncertainty-averse investors.

JEL Classification: G11, G12, C11

Keywords: model uncertainty, portfolio choice, home bias, Bayesian analysis, robust analysis

In standard finance theory, investors optimally allocate assets using a given stochastic model of asset returns. It follows that the optimal asset allocation depends on the choice of the model. Uncertainty about the correct choice of a stochastic model, or simply *model uncertainty*, has recently become a research topic of greater interest. This issue has been well known in the academic literature and even appears recently in the popular press. In the national best seller, *When Genius Failed: The Rise and Fall of Long-Term Capital Management* by Lowenstein (2000), the issue of model uncertainty is described as follows: "There is a key difference between a share of IBM...and a pair of dice. With dice, there is *risk*—you could, after all, roll snake eyes—but there is no *uncertainty*, because you know (for certain) the chances of getting a 7 and every other result. Investing confronts us with both risk and uncertainty. There is a risk that the price of a share of IBM will fall, and there is uncertainty about how likely it is to do so." The author of the book believes that the partners of Long-Term Capital failed to consider the chances that their models, which prescribe their investment strategies, are wrong. When models are likely to be wrong and the correct model is unknown, investors are naturally averse to the uncertainty. This paper examines the implications of this aversion in asset allocation.

When specifying the distribution of asset returns, finance professionals often impose restrictions of equilibrium models. A major reason for doing so is that using sample estimates of the mean and variance to obtain optimal allocation over a large number of risky assets is well known to be problematic[1]. The restriction of an asset-pricing model such as the CAPM or Fama-French model reduces the dimension of the estimation problem and leads investors to allocate wealth among the factor portfolios in the model. Another major reason for imposing restrictions of asset-pricing models is that asset prices inconsistent with the market equilibrium are believed to be less likely to persist. Although restrictions of asset-pricing models help us obtain portfolios that are more intuitive and easier to implement, we face uncertainty regarding models' pricing ability because all models are rejected in some empirical tests[2]

Currently, a popular approach for analyzing model uncertainty is Bayesian inference, which combines the prior beliefs in models and the information in data.[3] Of course, different prior beliefs lead to different posterior and predictive probability distributions of asset returns. Thus, researchers have tried various prior distributions of the pricing errors to see how they affect asset allocation.

[1] Best and Grauer (1991), Green and Hollifield (1992), Michaud (1989), and Britten-Jones (1999).

[2] Hodrick and Zhang (2001) show that all the recently proposed multi-factor models can be statistically rejected. Although the Fama-French model generally fits the data best, its validity is still questionable (see, for example, Daniel and Titman (1997)).

[3] Black and Litterman (1991, 1992), He and Litterman (1999), Pastor (2000), and Pastor and Stambaugh (2000).

In practice, researchers often take the empirical Bayes approach, in which prior distributions are estimated from the observed samples using the maximum likelihood method. The Bayesian analysis in the current literature, however, does not incorporate investors' aversion to model uncertainty.

Since the most difficult task in Bayesian analysis is to choose the prior distribution, investors naturally dislike the fact that they do not know what the correct prior belief is. They are averse to model uncertainty because each choice of prior belief will lead to gain or loss in investment value or utility. This aversion should affect their asset allocation. The purpose of this paper is to investigate how investors, who are averse to model uncertainty, make domestic and international asset allocations when they are uncertain about the correct prior belief in asset-pricing models. It helps us gain more understanding of the empirically observed asset allocations. In the standard Bayesian analysis of asset allocation, optimal portfolio is obtained from the predictive distribution implied by a chosen prior. It is important to understand whether a certain choice of prior belief can lead to an optimal portfolio for uncertainty-averse investors.

The theoretical framework of aversion to model uncertainty is the robust analysis, in which a preference is represented by the minimum of the expected utility over the set of possible probability distributions. Investors solve the maxmin problem, in which they choose the portfolio that maximizes the minimum expected utility. Intuitively, the minimum expected utility over the choices of priors reflects their conservative attitude and thus their aversion to uncertainty. Using an axiomatic approach, Gilboa and Schmeidler (1989) demonstrates that the minimum expected utility represents the preference with aversion to uncertainty about the probability distributions. A solution to the maxmin problem gives the asset allocation optimal for investors who are averse to model uncertainty. Using robust analysis, this paper examines asset allocations associated with the models and data that attract the most attention in the empirical finance literature. The models are the CAPM and Fama-French model related to the mean-variance utility function. The data are the returns on the domestic and international equity.

An important observation made in this paper is that aversion to uncertainty about an asset-pricing model may lead investors hold a portfolio that is mean-variance inefficient for any prior belief in the model. Over the portfolios sorted by firm size and book-to-market, there exists no prior belief in the CAPM such that the corresponding tangency portfolio gives the optimal allocation for uncertainty-averse investors. This can be viewed as a failure of the CAPM in the context of asset allocation. Uncertainty-averse investors do not want to rely on either the estimate restricted by the CAPM or the unrestricted estimate because of considerable doubts about the estimates.

Moreover, the optimal allocation by uncertainty-averse investors is not the tangency portfolio for the mean and variance estimated from any prior belief in the model. It is commonly observed that investors do not appear to hold mean-variance efficient portfolios. Aversion to model uncertainty can potentially be an important reason. This issue has not been previously investigated in the literature.

For some models and data, however, there exist prior beliefs so that the tangency portfolios are optimal for uncertainty-averse investors. For example, this paper shows that the tangency portfolio corresponding to the dogmatic prior belief in the Fama-French three-factor model is optimal for investors who are averse to model uncertainty. The reason is that the Fama-French model is a summary of the empirical properties of the portfolios sorted by firm size and book-to-market and gives a more reliable estimate of the mean and variance than the unrestricted estimate. The analysis in this paper also demonstrates that the tangency portfolio corresponding to the prior belief obtained from the empirical Bayes approach is not optimal for investors with aversion to model uncertainty.

An important empirical observation in international finance is investor's home bias. Finance professionals sometimes contend that investors' distrust of the world equilibrium model might justify the home bias. Elton and Gruber (1995) argue that, since there is little evidence to support the world CAPM, investors with no ability to forecast expected returns might seek to minimize the variance of their portfolio rather than consider the trade-off between the risk and returns. When discussing the possible explanations for home bias, French and Poterba (1991) write, "The statistical uncertainty associated with estimating expected returns in equity markets makes it difficult for investors to learn that expected returns in domestic markets are not systematically higher than those abroad. Because it is difficult to estimate ex ante returns, investors may follow their own idiosyncratic investment rules with impunity." This interesting discussion invites questions. What can be the idiosyncratic investment rule? Is the idiosyncratic rule simply their home bias? Can the rule be explained by aversion to model uncertainty? The analysis in this paper demonstrates that even investors averse to uncertainty about the world CAPM should hold the world market portfolio. Therefore, the strong home bias of investors cannot be justified by aversion to model uncertainty. The optimal asset allocation with aversion to uncertainty about the world CAPM is also inconsistent with variance minimization.

In order to apply robust analysis to Bayesian analysis of asset allocation using asset-pricing models, a shrinkage approach is developed in this paper. It shows that Bayesian analysis of model uncertainty shrinks both the predictive mean and variance of asset returns from the unrestricted sample moments to the estimates restricted by the asset-pricing model. More importantly, the predictive mean and variance share the same shrinkage factor, which indicates the relationship between the mean and variance implied by the asset-pricing model. The shrinkage approach reveals that a Bayesian investor, facing uncertainty about an asset-pricing model under consideration, implicitly assigns a weight between the unrestricted estimate and the estimate restricted by the asset-pricing model. The weight is the shrinkage factor. For a given prior distribution, the weight on the estimate restricted by the asset-pricing model is larger if the frontier of the factor portfolios has a higher Sharpe ratio. The weight on the model is large if a long history of stationary data is not available. Investors who take the empirical Bayes approach, in which they use the observed data to estimate the prior, choose the shrinkage factor to be 1/2, which assigns equal weights to the restricted and unrestricted estimates. The shrinkage approach is very useful because it explicitly shows how a prior belief affects the estimates of the mean and variance. For the purpose of this paper, the shrinkage approach allows us to obtain the properties of the solutions to our maxmin problems in robust analysis.

The shrinkage approach also helps us gain insights into investors' asset allocation decisions in international financial markets. We can use the shrinkage factor to measure the degree of home bias on a scale from 0 to 1. For U.S. investors, the degree of home bias measured on this scale is 0.9, which should be considered to be rather high. In the literature of international finance, there are many studies of international diversification benefits for the home-biased U.S. investors. Some studies do not impose any asset-pricing models[4], while some others impose the world CAPM.[5] This paper shows that the use of the world CAPM substantially affects the perceived international diversification benefit for U.S. investors. The greatest lower bound of the diversification benefit is less than one-third of the benefit measured without the model restriction but still well above zero. Interestingly, the benefit measured under the model restriction is close to the lower bound. The analysis in this paper offers a unified understanding of the studies using and not using the world CAPM.

The rest of the paper is organized as follows. Section I gives an overview of the classic, Bayesian, and robust analysis in the context of asset allocation using asset-pricing models. Section II develops the shrinkage approach to the Bayesian and robust analysis. The next two sections are empirical

[4]Examples include Bekaert and Urias (1996) and Lewis (1999).
[5]Examples include De Santis and Gerard (1997).

investigations. Section III examines domestic asset allocations. Section IV investigates international asset allocation and the issue of home bias. Conclusions and future research are discussed in Section V. Mathematical derivations are provided in the Appendix.

I. Aversion to Model Uncertainty

There are m risky assets. Let r_{1t} be the $m \times 1$ vector of excess returns (over the risk-free rate) on the assets during period t. An asset-pricing model is given and there are k factor portfolios in the model. Let r_{2t} be the $k \times 1$ vector of excess returns on the factor portfolios during period t. The time series of T observations, denoted by $R = \{r_t\}_{t=1,\cdots,T} = \{(r'_{1t}, r'_{2t})'\}_{t=1,\cdots,T}$, are assumed to follow a normal distribution with mean μ and variance Ω, independently across t. The mean and variance are decomposed into the following parts corresponding to the m assets and k factors:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \qquad \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}. \tag{1}$$

The mean and variance can be summarized by the parameters in a regression model:

$$r_{1t} = \alpha + \beta r_{2t} + u_t, \tag{2}$$

where α is the $m \times 1$ vector of Jensen's alpha, β is the $m \times k$ matrix of the betas, and u_t is the $m \times 1$ vector of the residual terms in the regression. The variance of u_t is assumed to be Σ. It follows that the mean and variance of the returns can be expressed as

$$\mu = \begin{pmatrix} \alpha + \beta\mu_2 \\ \mu_2 \end{pmatrix} \qquad \Omega = \begin{pmatrix} \beta\Omega_{22}\beta' + \Sigma & \beta\Omega_{22} \\ \Omega_{22}\beta' & \Omega_{22} \end{pmatrix}. \tag{3}$$

The asset-pricing model $\mu_1 = \beta\mu_2$ holds if and only if $\alpha = 0_{m\times 1}$, where $0_{m\times 1}$ is the $m \times 1$ vector of zeros.

In the classic framework of asset allocation using asset-pricing models, investors choose either to believe or not to believe the asset-pricing model. Those who do not believe the asset-pricing model estimate the parameters without restricting α to zero. Denote the maximum likelihood estimates of α, β and Σ by $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\Sigma}$ respectively. The maximum likelihood estimates of the mean and variance are

$$\hat{\mu} = \begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{pmatrix} = \begin{pmatrix} \hat{\alpha} + \hat{\beta}\hat{\mu}_2 \\ \hat{\mu}_2 \end{pmatrix} \quad \text{and} \quad \hat{\Omega} = \begin{pmatrix} \hat{\beta}\hat{\Omega}_{22}\hat{\beta}' + \hat{\Sigma} & \hat{\beta}\hat{\Omega}_{22} \\ \hat{\Omega}_{22}\hat{\beta}' & \hat{\Omega}_{22} \end{pmatrix}. \tag{4}$$

where $\hat{\mu}_2$ and $\hat{\Omega}_{22}$ are the sample mean and variance of r_{2t}. The above estimates of μ and Ω are in fact equivalent to the sample mean and variance of r_t. Those investors who do not believe the model make asset allocation based on the unrestricted estimates of the mean and variance in equation (4). Those who believe the asset-pricing model, however, impose the restriction $\alpha = 0_{m\times 1}$. Let $\bar{\beta}$ and $\bar{\Sigma}$ be the maximum likelihood estimates of β and Σ with the restriction of $\alpha = 0_{m\times 1}$. The restricted maximum likelihood estimates of the mean and variance are

$$\bar{\mu} = \begin{pmatrix} \bar{\beta}\hat{\mu}_2 \\ \hat{\mu}_2 \end{pmatrix} \quad \text{and} \quad \bar{\Omega} = \begin{pmatrix} \bar{\beta}\hat{\Omega}_{22}\bar{\beta}' + \bar{\Sigma} & \bar{\beta}\hat{\Omega}_{22} \\ \hat{\Omega}_{22}\bar{\beta}' & \hat{\Omega}_{22} \end{pmatrix} . \tag{5}$$

Those investors who believe the model make asset allocation decisions based on the restricted estimates of the mean and variance.

Discarding the dichotomy between disbelieving and believing the model, the Bayesian framework introduces an informative prior distribution of α to represent an investor's belief in the asset pricing model. The prior of α, conditional on Σ, is assumed to be a normal distribution with mean $0_{m\times 1}$ and variance $\theta\Sigma$, i.e.,

$$p(\alpha|\Sigma) \;=\; N(0_{m\times 1}, \theta\Sigma) . \tag{6}$$

The parameter θ is a positive number that controls the variance of the prior distribution of Jensen's alpha. Since an asset-pricing model does not impose any restrictions on β, Σ, μ_2, or Ω_{22}, prior distributions of these parameters are assumed to be independent and non-informative. Specifically, each of the prior probability density functions of β and μ_2 is proportional to a constant. The prior probability density function of Ω_{22} is proportional to $|\Omega_{22}|^{-(k+1)/2}$, and the prior probability density function of Σ is proportional to $|\Sigma|^{-(m+1)/2}$. These are the standard specifications for non-informative prior distributions.

In Bayesian framework, asset allocation is based on the predictive distribution, which is the distribution of r_{T+1} conditional on the observed data R. Let $E[r_{T+1}|R,\theta]$ and $V[r_{T+1}|R,\theta]$ be the mean and variance of the predictive distribution. They depend on the choice of the parameter θ in the prior distribution. For a given parameter θ and portfolio x, the mean-variance utility function of an investor is

$$U(x;\theta) = E[r_{T+1}|R,\theta]'x - \frac{1}{2}\gamma x'V[r_{T+1}|R,\theta]x , \tag{7}$$

where γ is the degree of risk aversion. Investors choose the following vector of portfolio positions on the risky assets:

$$x(\theta) = \gamma^{-1}\left(V[r_{T+1}|R,\theta]\right)^{-1} E[r_{T+1}|R,\theta] . \tag{8}$$

If we scale $x(\theta)$ to portfolio weights, it will be independent of the risk-avesion parameter γ. It is well known that an investor with a mean-variance utility function chooses the risky portfolio that maximizes the Sharpe ratio. This portfolio of risky assets is referred to as the tangency portfolio in the finance literature.

Since different values of θ lead to different predictive distributions and asset allocations, investors face the problem of making a choice among alternative values of θ. Investors are naturally averse to the uncertainty about the value of θ. Gilboa and Schmeidler (1989) demonstrate that an investor with aversion to model uncertainty chooses his portfolio by maximizing the minimum expected utility over the set of probability distributions. That is, he solves

$$\max_x \min_\theta U(x;\theta) \ . \tag{9}$$

It will be shown in the next section that the solution x^* to this maxmin problem exists and is in fact unique. The portfolio choice x^* is the optimal asset allocation for investors who are averse to model uncertainty. Fixing x^*, we can find the solution θ^* to $\min_\theta U(x^*;\theta)$. It is well known that that x^* is not necessarily the solution to $\max_x U(x;\theta^*)$. That is, x^* can be different from the asset allocation obtained by setting $\theta = \theta^*$ in (8). Notice that if we scale x^* to portfolio weights, it may still depend on the risk-aversion parameter γ.

The most popular way of solving a maxmin problem is to transform the maxmin problem to a minmax problem. The minmax problem is

$$\min_\theta \max_x U(x;\theta) \ . \tag{10}$$

In the minmax problem, investors are conservative and choose the probability distribution that gives the minimum of optimal utility. If $\theta^\dagger$ is a solution to the minmax problem, $x^\dagger$ is the solution to the utility maximization problem for given $\theta^\dagger$. That is, $x^\dagger = \arg\max_x U(x;\theta^\dagger) = x(\theta^\dagger)$. The asset allocation $x^\dagger$ can be understood as a conservative portfolio choice. If we substitute equation (8) into the utility function in (7), the minmax problem (10) becomes

$$\min_\theta \ E[r_{T+1}|R,\theta]' \left(V[r_{T+1}|R,\theta]\right)^{-1} E[r_{T+1}|R,\theta] \ . \tag{11}$$

Here, the parameter θ is chosen to minimize the highest Sharpe ratio of the predictive mean and variance. The resulting optimal Sharpe ratio is the greatest lower bound of the Sharpe ratios that an investor can obtain by varying parameter θ in the prior distribution.

Transformation of a maxmin problem to a minmax problem requires some assumptions on the utility function. If the utility function U is (weakly) concave in x and (weakly) convex in θ, the maxmin and the minmax problems are equivalent. This is the well known MiniMax Theorem. Sion (1958) and Bazaraa (1993) provide the general form of the theorem. When the MiniMax Theorem holds, the maxmin and minmax problem are equivalent and $(x^*;\theta^*) = (x^\dagger;\theta^\dagger)$. It implies $x^* = \arg\max_x U(x;\theta^*) = x(\theta^*)$, i.e., the tangency portfolio corresponding to the particular prior belief in the model with $\theta = \theta^*$ will be a portfolio that is even optimal for investors averse to model uncertainty. However, when the maxmin and minmax problems are not equivalent, there may not exist an prior so that the corresponding tangency portfolio is optimal for uncertaity-averse investors. This point will be demonstrated in Section III.

II. The Shrinkage Approach

In the Bayesian approach, different prior beliefs in asset-pricing models lead to different predictive distributions. However, for each given prior belief, it is not clear how much impact it has on the predictive distribution. In other words, we would like to know how a prior belief in an asset-pricing model affects the estimate of the mean and variance. This section develops the shrinkage approach that explicitly shows how a prior belief modifies the sample estimate of the mean and variance toward to the estimate restricted by the asset-pricing model. The shrinkage approach is then used to solve the maxmin problem.

The shrinkage formulation is summarized in the following theorem, which relates the predictive mean and variance to the maximum likelihood estimates:

THEOREM 1: *Let $\hat{S}$ be the highest Sharpe ratio of the efficient frontier spanned by the sample mean and variance of the factor portfolios, and let*

$$\omega = \frac{1}{1 + T\theta/(1+\hat{S}^2)} \ . \tag{12}$$

Then, the mean and variance of the predictive distribution are

$$E[r_{T+1}|R,\omega] = \omega \begin{pmatrix} \bar{\beta}\hat{\mu}_2 \\ \hat{\mu}_2 \end{pmatrix} + (1-\omega) \begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{pmatrix} \ , \tag{13}$$

$$V[r_{T+1}|R,\omega] = \begin{pmatrix} V_{11}(\omega) & V_{12}(\omega) \\ V_{12}(\omega)' & b\hat{\Omega}_{22} \end{pmatrix} \ , \tag{14}$$

where $V_{11}(\omega)$ and $V_{12}(\omega)$ are given by

$$\begin{aligned} V_{11}(\omega) &= b[\omega\bar{\beta} + (1-\omega)\hat{\beta}]\hat{\Omega}_{22}[\omega\bar{\beta} + (1-\omega)\hat{\beta}]' \\ &\qquad + h[\omega\bar{\delta} + (1-\omega)\hat{\delta}][\omega\bar{\Sigma} + (1-\omega)\hat{\Sigma}] \ , \end{aligned} \tag{15}$$

$$V_{12}(\omega) = b[\omega\bar{\beta} + (1-\omega)\hat{\beta}]\hat{\Omega}_{22} \ . \tag{16}$$

Here, $\bar{\delta}$, $\hat{\delta}$, b and h are scalers and defined as follows:

$$\bar{\delta} = \frac{T(T-2)+k}{T(T-k-2)} - \frac{k+3}{T(T-k-2)} \cdot \frac{\hat{S}^2}{1+\hat{S}^2}\,, \tag{17}$$

$$\hat{\delta} = \frac{(T-2)(T+1)}{T(T-k-2)}\,, \tag{18}$$

$$b = \frac{T+1}{T-k-2}\,, \tag{19}$$

$$h = \frac{T}{T-m-k-1}\,. \tag{20}$$

Equation (13) states that the predictive mean is a weighted average of the estimated means restricted and unrestricted by the asset-pricing model.[6] It is a shrinkage estimator. The shrinkage target is the maximum likelihood estimate of μ under the restriction of the asset-pricing model. According to Efron and Morris (1973), ω is referred to as *the shrinkage factor*. The predictive variance of the assets, $V_{11}(\omega)$, is a quadratic function of the shrinkage factor ω, rather than a linear weighted average. Although this is not a shrinkage estimator in the traditional sense, it can be understood as a nonlinear extension to the classic shrinkage estimator. If the shrinkage factor is one, $V_{11}(\omega)$ is proportional to the estimate restricted by the model. If the shrinkage factor is zero, $V_{11}(\omega)$ is proportional to the unrestricted sample estimate. The theorem also states that the predictive covariance, $V_{12}(\omega)$, between the assets and the factors is proportional to the weighted average of the estimates restricted and unrestricted by the model. The weight is the same shrinkage factor ω.

The above theorem establishes the link between the classic and Bayesian framework of asset allocation using asset-pricing models. The link is the shrinkage factor ω, which indicates the relative weight between the estimates restricted and unrestricted by the model. More importantly, the predictive mean and variance share exactly the same shrinkage factor. The theorem explicitly shows how prior beliefs influence the estimated mean and variance. As expected, the shrinkage factor ω is a decreasing function of θ. When θ approaches 0, the weight assigned to the estimate restricted by the model converges to 1. When θ becomes infinitely large, the weight assigned to the unrestricted estimate converges to 1. Therefore, choosing θ is equivalent to choosing the weight between the estimates restricted and unrestricted by the model.

Equation (12) captures another new feature, stating that the shrinkage factor ω is an increasing function of the Sharpe ratio of the frontier spanned by the factor portfolios. Given T and θ,

[6] Pastor and Stambaugh (1999) have noticed that if there is one asset and one factor, the posterior mean of Jensen's alpha is approximately a weighted average of the estimated alpha restricted and unrestricted by the model.

more weight is assigned to the estimate restricted by the model if the Sharpe ratio of the factor portfolios is higher. This is reasonable. If the asset-pricing model holds, some combination of the factor portfolios should give the highest Sharpe ratio of the efficient frontier spanned by all the assets. If the Sharpe ratio of the frontier spanned by the factors is much lower than the Sharpe ratio of the frontier spanned by all the assets, the asset-pricing model is likely to be wrong. If the Sharpe ratio of the factors is higher, it is closer to the highest Sharpe ratio of all the assets, and the model is more likely to be correct. More weight (higher shrinkage factor) is therefore assigned to the estimate restricted by the model.

Theorem 1 points to a simple way to implement the Bayesian analysis. First, obtain the maximum likelihood estimate subject to the model restriction and the estimate without the restriction; Then, choose ω and use the formula in the theorem to calculate the predictive mean and variance. This simple implementation is more accessible to practitioners than the framework of Bayesian analysis, especially to those not trained in Bayesian statistics. It is also more intuitive to practitioners because they can choose ω to reflect their belief in the relative importance between the model and data. If they think the model and data are equally important, they should choose ω as 1/2. If they think the data is twice as important as the model, the choice of ω should be 1/3. Each choice of ω implies a choice of θ and thus a choice of prior distribution.

A major distinction of Theorem 1 from other shrinkage estimation in the literature is to shrink both the mean and variance with the same shrinkage factor. In fact, MacKinlay and Pastor (2000) argue that it is very important to consider the relationship between the mean and variance implied by asset-pricing models. Using Monte Carlo simulation, they demonstrate that portfolios constructed using some relationship between the mean and variance have higher Sharpe ratios than those constructed without using any relationship. This idea is formalized in Theorem 1, which shows that the predictive mean and variance both depend on the same shrinkage factor. It reflects the relationship between the predictive mean and variance implied by an asset-pricing model. In the next section, it will be further demonstrated that shrinkage of the variance is important for optimal asset allocation among a large number of assets. Most existing shrinkage theorems in the literature shrink only the mean of asset returns but assume the variance to be given.[7] Since the main issue in finance is the trade-off between return and risk, the estimation of variance should be at least as important as the mean, because the variance is closely related to risk. Bayesian inference

[7]For example, see Andersen (1971), Black and Litterman (1991), Efron and Morris (1973), Jobson and Korkie, Ratti (1979), Jorion (1986, 1991).

of variance and its shrinkage formulation are considered in a few studies,[8] but their prior distributions do not consider the restrictions of asset-pricing models. Those studies shrink the predictive mean and variance separately with different shrinkage factors. It does not reflect any relationship between the mean and variance.

Most importantly, Theorem 1 allows us to obtain useful properties of the maxmin problem and solve it easily because the predictive mean and variance is a simple function of ω, whose range is from 0 to 1 (rather than from 0 to infinity). The mean-variance utility function of an investor is a function of x and ω, denoted by $U(x;\omega)$. The maxmin problem becomes

$$\max_x \min_\omega U(x;\omega) \ . \tag{21}$$

Since $U(x;\omega)$ is continuous in $(x;\omega)$ and the set of possible ω is compact, the function $\min_\omega U(x;\omega)$ is continuous in x. Since $U(x;\omega)$ is strictly concave in x, the function $\min_\omega U(x;\omega)$ is also strictly concave in x. It follows that there is a unique x^* that solves the maxmin problem (21). With the fixed x^*, solutions ω^* to $\min_\omega U(x^*;\omega)$ always exit but may not be unique.

Using the shrinkage factor ω, the minmax problem becomes

$$\min_\omega \max_x U(x;\omega) \ . \tag{22}$$

A solution $\omega^\dagger$, referred to as a conservative shrinkage factor, solves

$$\min_\omega \ E[r_{T+1}|R,\omega]' \left(V[r_{T+1}|R,\omega]\right)^{-1} E[r_{T+1}|R,\omega] \ . \tag{23}$$

The solution ω exists because the objective function in (23) is continuous and the set of ω is compact. The corresponding conservative portfolio choice is

$$x^\dagger = \frac{1}{\gamma} \left(V[r_{T+1}|R,\omega^\dagger]\right)^{-1} E[r_{T+1}|R,\omega^\dagger] \ . \tag{24}$$

The Bayesian analyses of asset allocation in the current literature do not incorporate aversion to model uncertainty. They often use the empirical Bayes approach to specify prior beliefs. In the basic parametric empirical Bayes approach, researchers specify a functional form of the prior distribution, which sometimes depends on hyperparameters. Researchers estimate the hyperparameters using the marginal distribution of data. Either maximum likelihood or the method of moments can be used for estimating the hyperparameters. This approach is also called the *compound decision problem* in the statistics literature. In this approach, the set of data used for estimating hyperparameters is the

[8] See Brown (1976) and Frost and Savarino (1986).

same set of data used for computing the posterior distribution.[9] Since the empirical Bayes approach is often applied in asset allocation,[10] it is important to understand what shrinkage factor it implies and whether the resulting asset allocation is optimal for investors averse to model uncertainty.

For the issues examined in this paper, the hyperparameter is θ. In order to determine θ, classic statistical inference is used to estimate the sampling distribution of Jensen's alpha. The estimated sampling distribution is then used as the informative prior distribution of Jensen's alpha. In the Appendix, it is shown that, under the null hypothesis of $\alpha = 0_{m\times 1}$, the sampling distribution of the maximum likelihood estimator of α, conditional on the factor and the parameters $(\alpha, \beta, \Sigma, \mu_2, \Omega_{22})$, is a normal distribution with mean $0_{m\times 1}$ and variance $T^{-1}(1+\hat{S}^2)\Sigma$. That is,[11]

$$\hat{\alpha} \,|\, \Sigma \quad \sim \quad N\left(0_{m\times 1}, \frac{1}{T}(1+\hat{S}^2)\Sigma\right) \,. \tag{25}$$

It follows from equation (6) that the parameter in the prior distribution of α should be specified as $\theta = T^{-1}(1+\hat{S}^2)$. Substituting this choice of θ into equation (12) gives $\omega = 1/2$. A researcher taking the empirical Bayes approach, therefore, assigns equal weights to the estimates restricted and unrestricted by the asset-pricing model, implying that the model and data have equal importance. This makes it convenient to compare the asset allocation obtained using the empirical Bayes approach with the allocation obtained using the maxmin approach.

III. Domestic Asset Allocation

Let us first look at allocations over a set of assets widely used in empirical studies when the CAPM is used to affect the estimates. The factor in the asset-pricing model is the monthly excess return on the value-weighted market index portfolio of NYSE, AMEX and Nasdaq during the period from July 1963 to December 1998. The excess returns on the assets are the monthly excess returns on the portfolios used by Fama and French (1993) and updated to the end of 1998. In order to compare the CAPM with the Fama-French model, the returns on the two Fama-French factors, SMB and HML, are also included as part of the asset returns.[12] The last five of the Fama-French portfolios that contain the largest firms are excluded because the market, SMB and HML factors are almost a linear combination of the 25 Fama-French portfolios. Table I presents the tangency portfolio weights for various values of the shrinkage factor ω. As expected, most of the weight is

[9] Berger (1985), Bernardo and Smith (1994) provide overviews of the compound decision problems.

[10] For example, see Frost and Savarino (1986) and Pastor and Stambaugh (2000).

[11] This sampling distribution has been derived by Gibbons, Ross and Shanken (1989) for the CAPM, which has only one factor. The distribution in (25) is a multi-factor extension.

[12] The data are made available by Kenneth French on his Web page.

on the market portfolio when ω is close to 1.

Since the Fama-French model is empirically successful in fitting the data, it is natural to apply it to asset allocation. The returns on the factors in this model include the returns on the SMB and HML portfolios in addition to the monthly excess returns on the value-weighted market index portfolio of NYSE, AMEX and Nasdaq. The excess returns on the assets are the monthly excess returns on the Fama-French portfolios considered in Table I. The tangency portfolio weights for various values of ω are reported in Table II. When ω is close to 1, most of the portfolio weight is assigned to the three Fama-French factors. When the value of ω is .1 or less, the portfolio weights in Tables I and II are very similar. In this case, the tangency portfolios are mainly determined by the unrestricted sample estimates.

Now, let us examine the optimal allocation with aversion to uncertainty about the CAPM. The solution x^* to the maxmin problem corresponding to each selected degree of risk aversion[13] is scaled into a vector of portfolio weights and reported in columns 2–5 of Table III. It is interesting that the solution to the maxmin problem assigns large positive or negative weights to either SMB or the HML portfolios, even when the CAPM is used to shrink the estimate. The optimal portfolio for an investor with aversion to model uncertainty is very different from the market portfolio prescribed by the CAPM. It is also very different from the portfolio based on the unrestricted sample estimate. It indicates that the utility loss caused by the uncertainty about the CAPM is probably as much as the loss caused by the uncertainty about the unrestricted estimate. More importantly, the optimal allocations by uncertainty-averse investors in no way resemble the tangency portfolios obtained in Table I. In fact, there is no prior belief in the CAPM that yields the optimal allocation for uncertainty-averse investors. That is, there is no ω such that $x^* = \arg\max_x U(x;\omega)$. Thus, the optimal allocation for uncertainty-averse investors is not a tangency portfolio corresponding to any prior beliefs in the CAPM. In particular, the tangency portfolio obtained using the empirical Bayes approach ($\omega = 0.5$) is not optimal for uncertainty-averse investors.

Naturally, we should also look at the optimal allocation with aversion to uncertainty about the Fama-French model. The solutions x^* to the maxmin problems are reported in columns 6–9 of Table III. When the Fama-French model is used, the portfolio weights in the solution to the maxmin problem is very similar to the tangency portfolio weights implied by the model with ω being close to 1 in Table II. Therefore, when the Fama-French model is under consideration, an investor

[13]The risk aversion γ is set to be 3, 5, 7, or 9 in this paper. Bodie et al. (1999, see page 191–193 for details) calibrated that $\gamma = 2.96$ for U.S. investors. Similarly, Pastor and Stambaugh (2000) calibrated that $\gamma = 2.84$.

with aversion to model uncertainty can approximate his/her desired asset allocation by setting the shrinkage factor ω close to 1 or simply assuming the model holds. It appears puzzling why aversion to the uncertainty about the Fama-French model leads to asset allocations similar to the tangency portfolio obtained from the model. In fact, in the set-up of the maxmin problem, the investor is making a choice between two stochastic models—the predictive distribution restricted by the Fama-French model and the unrestricted predictive distribution. The unrestricted predictive distribution may be more likely to be incorrect than the Fama-French model because the estimate depends too much on noisy observations. In this case, aversion to the uncertainty about the unrestricted predictive distribution forces the investor to rely more on the Fama-French model. Notice that the optimal asset allocation with uncertainty-aversion is very different from the tangency portfolio obtained from the empirical Bayes approach ($\omega = 0.5$) when investors use the Fama-French model.

One should not simply conclude that the CAPM is a useless model. First, the CAPM is a static model applied to 30 years of asset returns. It is likely that this static model cannot capture the time-variation of risk. Second, the market portfolio in the CAPM contains only equity, which is a small part of the aggregate wealth of the economy. Investments on capital other than equity could be partially responsible for the empirical failure of the CAPM.[14] These might be the reasons why investors with uncertainty-aversion do not choose the static CAPM. In contrast, the Fama-French model is constructed to fit the data. The SMB and HML factors serve as the correction portfolios, in the sense of Hansen and Jagannathan (1997), for the pricing errors of the CAPM. The two correction portfolios proxy the changing risk in the economy and the premium on the missing assets in the market portfolio. This might be why the Fama-French model is chosen when investors compare it with the unrestricted sample estimate.

In the above application of Gilboa and Schmeidler's (1989) preference theory, the maxmin problem is solved directly, unlike in other papers that apply the MiniMax theorem to transform the maxmin problem to a minmax problem.[15] The MiniMax theorem cannot be applied to the problems discussed in this paper. The trouble is that $U(x;\omega)$ can be shown to be a weakly concave (rather than convex) function in ω for the problems and data discussed in this paper. In the issue addressed in this paper, the unequivalence of the maxmin and minmax problems is an important implication of the data to the model. If we assume that the maxmin and minmax are equivalent,

[14] Jagannathan and Wang (1996) argue that human capital can be important.

[15] For example, Chamberlain (2001) applies the maxmin utility theory to autoregressive models for panel data. Hansen and Sargent (2001) apply the maxmin utility theory to robust control problems facing model uncertainty. Both papers transform their maxmin problems to minmax problems.

we will lose some important information contained in the data.

Since the conditions in the MiniMax Theorem is sufficient but not necessary, the maxmin and minmax problems may happen to be equivalent in our case. Therefore, we should also solve the minmax problem to compare with the solution to the maxmin problem. In Figure 1, when the CAPM is used, the optimal Sharpe ratio is plotted as a function of the shrinkage factor (the solid curve). Since the function is strictly decreasing, the conservative shrinkage factor is $\omega^{\dagger} = 1$. A similar function can be plotted for the Fama-French model (the dotted curve). This function is also decreasing and thus the conservative shrinkage factor is also $\omega^{\dagger} = 1$. For both the CAPM and Fama-French model, the solutions to the minmax problems assign all the weights to the models, and the corresponding portfolios have positions only on the factor portfolios. Since the asset allocation obtained from the Fama-French model is similar to the allocation obtained from the maxmin problem, the minmax problem and maxmin problem give similar (but not always identical) asset allocations when the Fama-French model is in consideration. This implies that investors who are averse to the uncertainty about the Fama-French model can obtain an approximate optimal asset allocation by choosing the conservative shrinkage factor. However, this is not true when the CAPM is used.

Besides helping to solve the maxmin problem, the shrinkage approach developed in this paper offers additional insights into the Bayesian analysis of asset allocation under model uncertainty. To examine the impact of model uncertainty on asset allocation, Pastor (2000) and Pastor and Stambaugh (1999, 2000) specify the prior distribution of Jensen's alpha as $N(0\,,\,(\sigma^2/s^2)\Sigma)$, where s^2 is defined as $s^2 = \text{trace}(\hat{\Sigma})/m$. We can view s^2 as the average sample variance of the residual terms in the regression equation (2). They show that σ^2 can be interpreted as the average prior variance of Jensen's α, unconditional on Σ. Therefore, σ is viewed as the measure of uncertainty about the model's pricing ability. They experiment with different values of σ and look at the resulting asset allocations. Clearly, $\theta = \sigma^2/s^2$. Thus, each value of σ corresponds to a value of θ and a shrinkage factor ω, which determines the tangency portfolio weights. Since the shrinkage factor ω is a decreasing function of θ, it is also a decreasing function of σ. Like θ, the range of σ is from 0 to infinity.

Pastor and Stambaugh (2000) observe that the difference in tangency portfolios using different models "are substantially reduced by modest uncertainty about the models' pricing abilities." In Figure 2, the shrinkage factor ω is plotted as a function of σ. The shrinkage factor drops quickly as σ increases. The sharp reduction of the shrinkage factor is consistent with Pastor and Stambaugh's observation and explains why modest value of σ substantially reduces models' impact on portfolio choice. The main reason for the fast drop of the shrinkage factors is the large number of observations. It is easy to see this from equation (12). Since the factor Sharpe ratio is .1219 for the CAPM and 0.2551 for the Fama-French model, the magnitude of the second term in the denominator is of the order of $T\theta$, which equals $T(\sigma^2/s^2)$. When T is large, the shrinkage factor ω drops fast as σ increases. This is reasonable. When more than 30 years of monthly observations are used in the inference of model parameters, it is implicitly assumed that the parameters are constant and the stochastic process of the returns is stationary over 30 years. With such a long period, we should be able to estimate the constant parameters very well with the observed data. Thus, little weight will be assigned to the model when the model's pricing ability is quite uncertain. If the model parameters are varying over time, we are not able to use such a long period of observation in the inference. It will then be more important to apply the restrictions of asset-pricing models.

Theorem 1 can be used to examine whether the shrinkage estimate of the variance is important to optimal portfolio choice. If the predictive variance is not affected by the shrinkage factor, the vector of portfolio weights is the weighted average of the tangency portfolios with and without the restriction of the asset-pricing model. In this case, the tangency portfolio based on the unrestricted sample estimate linearly shrinks to the tangency portfolio of the factors. This observation offers a way to check whether the shrinkage of variance is important. The weights in Tables I and II are, however, nonlinear with respect to the shrinkage factor. In Table I, the portfolio weight on the market index increases from 1.00 to 2.97 when the shrinkage factor decreases from 1.00 to 0.90. The weight then decreases to 0.57 when the shrinkage factor decreases to 0.75. As the shrinkage factor drops further, the portfolio weight on the market index rises again. In Table II, we also observe complicated changes of the portfolio weight on the market index. These indicate that the shrinkage estimate of the variance has substantial influence on asset allocation. The shrinkage estimate of the variance could be less important for some small number of assets, but it is more likely to be important when the number of risky assets is large. It is the large number of risky assets that motivates us to impose the restrictions of asset-pricing models, which offers better estimates by reducing the dimension of our problems.

IV. International Asset Allocation

Perhaps the most puzzling empirical observation in international asset allocation is home bias—people invest heavily in their home country. About 92 percent of U.S. investors' equity holding is domestic. Investors in many other countries also have high domestic equity ownership.[16] The shrinkage factor developed in this paper offers a measure of investors' bias on a scale from 0 to 1. As French and Poterba (1991) point out, investors "may impute extra 'risk' to foreign investments because they know less about foreign markets, institutions, and firms." The "extra risk" relative to the home equity market can be naturally expressed as the uncertainty about the prior distribution of α in the following factor model:

$$r_{\text{foreign}} = \alpha + \beta r_{\text{US}} + \epsilon \qquad \epsilon \sim N(0, \Sigma) \tag{26}$$

An "asset-pricing model" can be introduced by setting $\alpha = 0$ in equation (26). An investor who is completely U.S.–biased allocates all his/her wealth to the U.S. market. A model that justifies this allocation to be optimal is the asset-pricing model that specifies the U.S. market index as the only factor. Therefore, this model is referred to as the U.S.–bias model.[17] Since a shrinkage factor is the weight assigned to a model when investors combine the information from the model and data, it is a measure from 0 to 1 for the bias toward such a model. For example, if an investor always uses the model despite the observed data, he chooses a shrinkage factor of 1. An investor who takes the empirical Bayes approach to determine the uncertainty about Jensen's alpha will be half-way biased toward the model. Following this idea, we can use the shrinkage factor to measure the home bias of U.S. investors.

Economists are often interested in the incremental diversification benefit obtained by adding emerging markets to the portfolio of industrial countries. The shrinkage factor can be used to model the bias toward industrial countries. For this purpose, we consider a model that has the index portfolios in seven industrial countries as its factors. This model can be referred to as the G7-bias model. The index portfolios in emerging markets are treated as the assets. If the portfolios of the industrial countries span the efficient frontier of all the markets, the G7-bias model should correctly price the portfolios of the emerging markets, in which case investors only allocate their wealth to the industrial countries. The shrinkage factor for the G7-bias model therefore measures the degree of bias toward the seven industrial countries.

Consider the asset allocation over a set of major international markets, which include seven developed markets in industrial countries (the United States, the United Kingdom, Canada, France, Germany, Italy, and Japan) and eight emerging markets (Argentina, Brazil, Chile, Mexico, Hong Kong, South Korea, Singapore, and Thailand). The historical data of the dollar-denominated

[16] See Bohn and Tesar (1996), Cooper and Kaplanis (1994), and French and Poterba (1991).
[17] Pastor (2000) refers to it as the domestic CAPM.

monthly returns on the equity indices in the emerging markets, with the exception of Hong Kong and Singapore, are obtained from the International Finance Corporation (IFC). The data for other markets are obtained from Morgan Stanley Capital International (MSCI). The data cover the period from January 1976 to December 1998. All returns are converted to excess returns by subtracting the monthly returns on the U.S. Treasury Bills, which are obtained from CRSP. Table IV reports the optimal portfolio weights on the U.S. for various degrees of U.S. bias. It also reports the sum of the optimal portfolio weights on G7 countries for various degrees of G7 bias.

We can extract investors' degree of bias by comparing Table IV with the actual asset allocation of investors. If portfolios are unconstrained, the tangency portfolio weight on the U.S. is nearly 90 percent, even when the degree of bias is as low as 0.01. This is mainly due to the large short position on Canada (not reported) because of the low average return on the Canadian market and its high correlation with the U.S. market. A large short position on a country is especially unrealistic for long-term asset allocation. If short sales are not allowed, the unbiased tangency portfolio weight on the U.S. is about 40 percent. Since U.S. investors' actual foreign equity holdings account for only about 8 percent of their total equity holdings, Table IV indicates that the degree of U.S. bias is 0.9, if short sales are not allowed. Therefore, the degree of home bias of U.S. investors is measured as 0.9, which is rather high on a scale from 0 to 1. In particular, U.S. investors do not seem to use the empirical Bayes approach ($\omega = 0.5$) to determine their beliefs in the U.S.–biased model.

Given the strong home bias of U.S. investors, there are many studies on the existence and magnitude of the international diversification benefits. However, different studies often take different approaches separately. Some studies examine the efficient frontier spanned by the international markets. Other studies assume the world version of the CAPM that specifies the world portfolio as the efficient portfolio. Still other studies consider the effects of short-sale constraints on international diversification.[18] Each study addresses an important perspective separately. This makes it difficult to obtain a unified understanding of those studies. The world CAPM is rejected in several empirical tests.[19] It is also well known to be difficult for investors to take some short positions. In addition, all the studies assume that U.S. investors are completely biased and they invest 100 percent of their wealth in the U.S. The international diversification benefits obtained in the empirical analysis depend on whether the world CAPM is used, how biased the investors are, and whether short positions are constrained.

[18] Li, Sarkar and Wang (2001) offer an overview and Bayesian analysis on this issue.

[19] Example of such tests are reported by De Santis and Gerard (1997) and Fama and French (1998).

The shrinkage approach can provide a synthesis of these studies. To appreciate the economic importance of international diversification, the diversification benefit is often measured by the increase in utility, a measure that depends on the degree of risk aversion. For an investor who is U.S.–biased/G7-biased (ω_h), the vector of his/her portfolio positions is $x_h = \gamma^{-1}\Omega_h^{-1}\mu_h$, where μ_h and Ω_h are the predictive mean and variance affected by the U.S.–bias/G7-bias model with the shrinkage factor ω_h. For an analyst who assigns weight ω_a to the world CAPM, the vector of tangency portfolio positions should be $x_a = \gamma^{-1}\Omega_a^{-1}\mu_a$, where μ_a and Ω_a are the predictive mean and variance affected by the world CAPM with the shrinkage factor ω_a. The international diversification benefit is calculated as the difference in the utility of the analyst's tangency portfolio and the utility of the investor's biased portfolio. That is, the diversification benefit is measured by

$$\Delta U = \left(\mu_a' x_a - \frac{1}{2}\gamma x_a' \Omega_a x_a\right) - \left(\mu_a' x_h - \frac{1}{2}\gamma x_h' \Omega_a x_h\right) \ . \tag{27}$$

The utility difference can also be interpreted as the change in certainty equivalent return. Substituting out x_h and x_a, we obtain

$$\Delta U = \gamma^{-1}\left(\mu_a'(\Omega_a^{-1}\mu_a - \Omega_h^{-1}\mu_h) - \frac{1}{2}(\mu_a'\Omega_a^{-1}\mu_a - \mu_h'\Omega_h^{-1}\Omega_a\Omega_h^{-1}\mu_h)\right). \tag{28}$$

Since ΔU is proportional to γ, the relative comparison of the diversification benefits should not be affected by the degree of risk aversion. Following French and Poterba (1991), the degree of risk aversion γ is assumed to be 3 in the calculation of ΔU.

To see how the diversification benefit over international markets depend on investors' bias, the world CAPM, and the short-sale constraints, Table V reports ΔU for various values of ω_a and ω_h. Let us start with the case where portfolios are unconstrained. For an investor who is strongly U.S.–biased ($\omega_h \geq .90$), the increase in the utility due to diversification is well above 1 percent in annual certainty equivalent return, if the analyst assigns a very small weight ($\omega_a = 0.05$) to the world CAPM. If the analyst assumes that the world CAPM holds exactly, he/she concludes that the diversification benefit is only .26 percent in annual certainty equivalent return, which is less than one-fourth of the diversification benefit measured by the analyst who does not use the world CAPM. If the analyst uses the empirical Bayes approach ($\omega_a = 0.50$) to construct his/her prior belief in the world CAPM, he/she concludes that the increase in the annual certainty equivalent return for an investor with U.S. bias is between .34 and .47 percent. This is about one-third of the benefit perceived by the analyst who does not use the world CAPM but higher than the benefit perceived by the analyst who completely relies on the world CAPM.

Since the international diversification benefit reported in economic analysis depends on the shrinkage factor for the world CAPM, it is important to know the range of the diversification benefit as the shrinkage factor changes. Especially, we are interested to know if the range is above zero. Corresponding to the degree of U.S. bias measured as $\omega_h = 0.9$, the greatest lower bound of the diversification benefit measured by the certainty equivalent return is about 0.2 percent per annum, which is above zero.

In summary, three important observations can be made from Table V. First, the lower bound of the international diversification benefit is positive, indicating the existence of diversification benefit for U.S. investors. Second, the lower bound is less than one-third of the benefit measured without the model restriction, indicating the sensitivity of the estimated diversification benefit to the prior belief in the model. Third, the benefit measured under the restriction of the world CAPM is close to the lower bound, indicating that the world CAPM offers a rather conservative estimate of the diversification benefit.

The short-sale constraints reduce the magnitude of the diversification benefit, especially when the shrinkage factor for the World CAPM is small. However, the effects of the world CAPM on the diversification benefits are the same as in the case of unconstrained portfolios. Moreover, short-sale constraints do not eliminate diversification benefits for U.S. investors. Table V also presents the international diversification benefit for G7-biased investors. The results are qualitatively similar to those for U.S.–biased investors.

Although the lower bound of the diversification benefit is positive, U.S. investors do not diversify over the international markets. One possible reason is investors' aversion to the uncertainty about the world equilibrium model. To examine this issue, the maxmin problem for the uncertainty about the world CAPM is solved directly by numerical search.[20] The solution suggests investing 100 percent in the world market portfolio, as shown in Table VI for various degrees of risk aversion γ. It turns out that the solution to each maxmin problem is exactly the same as the solution to the corresponding minmax problem. In Figure 1, for the world CAPM, the shrinkage factor of $\omega = 1$ gives the lowest Sharpe ratio. It shows that uncertainty about the world CAPM does not justify putting most wealth in the U.S. In the solution to the maxmin problem, the weight on the world market portfolio is 100 percent. It indicates that the utility loss caused by the uncertainty about the world CAPM is probably much smaller than the loss caused by the uncertainty about

[20]Again, the Minimax Theorem cannot be applied because the mean-variance utility function in the problem is not convex with respect to the shrinkage factor.

the unrestricted estimate.[21] Since the actual weight of the U.S. in the world market portfolio is only around 50 percent, aversion to the uncertainty about this world equilibrium model does not justify the strong home bias of U.S. investors. Uncertainty-averse investors should impose the world CAPM and hold the world market portfolio rather than use the empirical Bayes approach or variance minimization. This also implies that the world CAPM should be imposed when we measure the international diversification benefit for uncertainty-averse investors.

Since there is also uncertainty about the home-bias model, it is natural to question whether aversion to this uncertainty pushes U.S. investors' asset allocation away from the U.S. Table VI presents the asset allocations obtained from the maxmin problem when the U.S.–bias model is used to affect the predictive mean and variance. The maxmin problems are solved for various degrees of risk aversion γ. As shown in the table, even if investors are unsure about the U.S.–bias model and are averse to this uncertainty, they still allocate all the wealth to the U.S. Uncertainty about the U.S.–bias model does not deter them from making home biased asset allocation. In this case, the solution to the maxmin problem is also the solution to the minmax problem because the lowest Sharpe ratio is at $\omega = 1$, as shown in Figure 1. The heavy portfolio weight on the U.S. assets thus indicates that U.S. investors are using the U.S.–bias model rather than the world CAPM.

V. Conclusion

Standard investment theory seeks to equate owning a share of stock with rolling a pair of fair dice. In practice, this linkage is tenuous. Because the true stochastic process of asset returns is uncertain, owning stock is like rolling dice without knowing whether the dice are fair or even how many faces they have. The distinction between owning a stock and participating in a lottery causes investors to allocate their assets differently in reality than in a model because of aversion to model uncertainty. The analysis in this paper maintains the distinction as well as the link between investments and lotteries. It is shown that investors with aversion to model uncertainty may choose an asset allocation that is not mean-variance efficient for the probability distribution estimated from any particular prior belief in the model. In most cases, the portfolio obtained from the empirical Bayes approach is suboptimal for uncertainty-averse investors.

This paper demonstrates that a Bayesian investor implicitly assigns a weight (shrinkage factor) between the restricted asset-pricing model and the data (or the unrestricted statistical model). For

[21] In fact, Hodrick, Ng and Sengmueller (1999) show that the world CAPM usually prices the country indices well. The model is strongly rejected only when test assets include international book-to-market portfolios.

a given prior distribution, the weight on the asset-pricing model is larger if the frontier of the factor portfolios has higher Sharpe ratio, i.e., history lends stronger credibility to the factor-based pricing model. The connection between Bayesian inference and shrinkage estimation is used to understand various issues on asset allocation in the presence of model uncertainty. Not surprisingly, the paper shows that the weight on the asset-pricing model is large if a long history of stationary data is not available. The shrinkage approach helps to understand the empirical Bayes approach in this context. For example, the popular empirical Bayes approach is shown to assign equal weights to the asset-pricing model and the data.

This paper illustrates how to use the shrinkage approach to understand investors' international asset allocation decisions. The shrinkage factor measures the degree of home bias on a scale from 0 to 1, and it is found to be 0.9 for U.S. investors. Most researchers probably agree that U.S. investors are home-biased, but the perceived international diversification benefit can be quite different, depending on the extent to which they trust the historical data and the model for prediction. It is shown that the greatest lower bound of the benefit is positive, although it is less than one-third of the benefit measured without the restriction of any asset-pricing model. Interestingly, the benefit measured under the restriction of the world CAPM is close to the lower bound. More importantly, the paper shows that the strong home bias of U.S. investors cannot be justified by aversion to uncertainty about the world market equilibrium model.[22]

The approach to home bias indicates that shrinkage factors for models are actually very flexible for analyzing portfolio advice. Most investment opinions can be expressed as some suggested portfolios. One can specify an "asset-pricing model" that contains the suggested portfolios as the factors. The shrinkage factor for this "asset-pricing model" gives the mix of the prior opinion and the observed data in asset allocation. One can also use the same methodology to measure the degree of home bias in other countries or regions.

Although the paper focuses on the choice between the estimated probability distributions restricted and unrestricted by an asset-pricing model, economists usually face the choice of several asset-pricing models. It is an interesting but difficult task to develop a Bayesian framework of asset allocation using multiple asset-pricing models. However, it might be possible and useful to implement the shrinkage approach to the choice of multiple asset-pricing models. The shrinkage

[22] Although many other possible reasons for home bias have been explored in the literature, it is still unclear what causes the bias. Most recently, Huberman (2001) suggests that familiarity is the potential explanation. He argues that similar bias is observed even in the domestic arena, for instance, in employee's strong tendency to hold company stock in their 401(k) retirement plans.

estimate of the mean can be a weighted average of the estimated means restricted by various models. The shrinkage estimate of the variance might be constructed as a quadratic function of the weights and the estimated variances restricted by the models. The extension of the analysis of model uncertainty to multiple asset-pricing models deserves more research.

In this paper, it is assumed that the observations of returns are drawn from identical and independent normal distributions with constant parameters. The asset allocation problem is considered for an investor with mean-variance utility function during a single period. Although this is the simplest problem of asset allocation over multiple risky assets, it offers a starting point for understanding the complicated problems of dynamic asset allocation with more general stochastic processes. Numerous recent papers empirically examine optimal asset allocation problems in multiple periods when returns are predictable.[23] When returns are not identically and independently distributed over time, it would be useful to incorporate dynamic equilibrium models into asset allocation decisions during multiple periods. Dynamic optimal allocation over a large number of risky assets is more difficult and sensitive to the estimates of the conditional probability distribution. In this area, it may be even more interesting and fruitful to address the issue of model uncertainty using the shrinkage approach.

[23]For example, Ang and Bekaert (2000) and Lynch (2000) examine the dynamic optimal portfolios of multiple risky assets when asset returns are predictable.

Appendix. Mathematical Derivations

A. Two Lemmas

It is convenient to prove two lemmas before deriving Theorem 1. The set of unknown parameters is denoted by $\Theta = (\Gamma, \Sigma, \mu_2, \Omega_{22})$, where $\Gamma = (\alpha, \beta)'$. The distribution assumption for asset returns in Section II and the prior distribution (6) imply that the posterior distribution of (Γ, Σ) is independent of the posterior distribution of (μ_2, Ω_{22}). Let $X = (r_{21}, \cdots, r_{2T})'$ and $Z = (\iota, X)$, where ι is the $T \times 1$ vector of ones. Also, define $D = \theta^{-1} J J'$, where J is a $(k+1) \times 1$ vector in which the first element is one and all other elements are zero. The posterior distributions imply the following posterior moments of the parameters:

$$
\begin{aligned}
E[\Gamma|\Sigma, R] &= (D + Z'Z)^{-1} Z'Z\hat{\Gamma} = \tilde{\Gamma} \equiv (\tilde{\alpha}, \tilde{\beta})' && \text{(A1)}\\
E[\Sigma|R] &= (T\hat{\Sigma} + \hat{\Gamma}' Q \hat{\Gamma})/(T - m - k - 1) \equiv \tilde{\Sigma} && \text{(A2)}\\
\mathrm{var}(\mathrm{vec}(\Gamma)|R) &= \tilde{\Sigma} \otimes (D + Z'Z)^{-1} && \text{(A3)}\\
E[\mu_2|\Omega_{22}, R] &= \hat{\mu}_2 && \text{(A4)}\\
E[\Omega_{22}|R] &= \hat{\Omega}_{22} T/(T - k - 2) \equiv \tilde{\Omega}_{22} && \text{(A5)}\\
\mathrm{var}(\mu_2|R) &= \hat{\Omega}_{22}/(T - k - 2) \ , && \text{(A6)}
\end{aligned}
$$

where $Q = Z'(I_T - Z(D + Z'Z)^{-1} Z')Z$, and the notation "$\equiv$" represents the expression "which is denoted by." The above formulas can be obtained by following the standard derivations in the textbooks of Bayesian statistics.

LEMMA 1: *The posterior means of the parameters are*

$$E[\Gamma \,|\, \Sigma, R] = \omega \bar{\Gamma} + (1 - \omega)\hat{\Gamma} \tag{A7}$$

$$E[\Sigma \,|\, R] = h \left[\omega \bar{\Sigma} + (1 - \omega)\hat{\Sigma}\right] \tag{A8}$$

$$E[\mu_2 \,|\, \Omega_{22}, R] = \hat{\mu}_2 \tag{A9}$$

$$E[\Omega_{22} \,|\, R] = a\,\hat{\Omega}_{22} \ , \tag{A10}$$

where $a = T/(T - k - 2)$ and h is defined as in Theorem 1.

Proof of Lemma 1: It follows from the properties of the inverse of partitioned matrix that

$$(D + Z'Z)^{-1} = \begin{pmatrix} \theta^{-1} + T & \iota' X \\ X'\iota & X'X \end{pmatrix}^{-1} = \omega G + (1 - \omega)(Z'Z)^{-1} \ , \tag{A11}$$

where ω and G are defined as follows:

$$\omega \quad = \quad \frac{\theta^{-1}}{\theta^{-1} + T - \iota' X (X'X)^{-1} X'\iota} \ , \tag{A12}$$

$$G = \begin{pmatrix} 0 & 0_{1\times k} \\ 0_{k\times 1} & (X'X)^{-1} \end{pmatrix} . \tag{A13}$$

It is easy to see that $\iota' X(X'X)^{-1}X'\iota = T\hat{\mu}_2'(\hat{\Omega}_{22} + \hat{\mu}_2\hat{\mu}_2')^{-1}\hat{\mu}_2$. The inverse matrix of $\hat{\Omega}_{22} + \hat{\mu}_2\hat{\mu}_2'$ is $\hat{\Omega}_{22}^{-1} - \hat{\Omega}_{22}^{-1}\hat{\mu}_2\hat{\mu}_2'\hat{\Omega}_{22}^{-1}(1 + \hat{\mu}_2'\hat{\Omega}_{22}^{-1}\hat{\mu}_2)^{-1}$, which the reader may verify by multiplying the two together. It follows that $\iota' X(X'X)^{-1}X'\iota = T(\hat{\mu}_2'\hat{\Omega}_{22}^{-1}\hat{\mu}_2)/(1 + \hat{\mu}_2'\hat{\Omega}_{22}^{-1}\hat{\mu}_2)$. It is well known that $\hat{\mu}_2'\hat{\Omega}_{22}^{-1}\hat{\mu}_2$ is the square of the highest Sharpe ratio of the frontier spanned by the sample mean $\hat{\mu}_2$ and variance $\hat{\Omega}_{22}$. Denoting the Sharpe ratio as $\hat{S}$, we have

$$\iota' X(X'X)^{-1}X'\iota = T\hat{S}^2/(1 + \hat{S}^2) \tag{A14}$$

and thus $T - \iota' X(X'X)^{-1}X'\iota = T/(1 + \hat{S}^2)$, which gives equation (12). Equation (A7) in Lemma 1 follows immediately from substituting equation (A11) into equation (A1). Since substitution of (A11) into Q gives

$$Q = \omega(T - \iota' X(X'X)^{-1}X'\iota)JJ' , \tag{A15}$$

we should have

$$\hat{\Gamma}'Q\hat{\Gamma} = \omega(T - \iota' X(X'X)^{-1}X'\iota)\hat{\alpha}\hat{\alpha}' . \tag{A16}$$

Noticing $T - \iota' X(X'X)^{-1}X'\iota = \iota'[I_T - X(X'X)^{-1}X']\iota$ and letting $M = I_T - \iota\iota'/T$, one can then show that

$$\begin{aligned} \hat{\Gamma}'Q\hat{\Gamma} &= \omega\hat{\alpha}\iota'[I_T - X(X'X)^{-1}X']\iota\hat{\alpha}' \\ &= \omega Y'\{[I - X(X'X)^{-1}X'] - [M - MX(X'MX)^{-1}X'M]\}Y \\ &= \omega(T\bar{\Sigma} - T\hat{\Sigma}) . \end{aligned} \tag{A17}$$

Substituting this into (A2) gives equation (A8), which completes the proof of Lemma 1.

LEMMA 2: *The posterior mean of μ and Ω are*

$$E[\mu|\Omega, R] = \omega \begin{pmatrix} \bar{\beta}\hat{\mu}_2 \\ \hat{\mu}_2 \end{pmatrix} + (1-\omega)\begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{pmatrix} , \tag{A18}$$

$$E[\Omega|R] = \begin{pmatrix} \Phi(\omega) & a[\omega\bar{\beta} + (1-\omega)\hat{\beta}]\hat{\Omega}_{22} \\ a\hat{\Omega}_{22}[\omega\bar{\beta} + (1-\omega)\hat{\beta}]' & a\hat{\Omega}_{22} \end{pmatrix} , \tag{A19}$$

where

$$\Phi(\omega) = a[\omega\bar{\beta} + (1-\omega)\hat{\beta}]\hat{\Omega}_{22}[\omega\bar{\beta} + (1-\omega)\hat{\beta}]' \tag{A20}$$

$$+ h[\omega\bar{\lambda} + (1-\omega)\hat{\lambda}][\omega\bar{\Sigma} + (1-\omega)\hat{\Sigma}] , \tag{A21}$$

$$\bar{\lambda} = [(T-2) - \hat{S}^2/(1 + \hat{S}^2)]/(T-k-2) , \tag{A22}$$

$$\hat{\lambda} = (T-2)/(T-k-2) , \tag{A23}$$

and ω, $\hat{S}$, and h are defined as in Theorem 1 and a is defined as in Lemma 1.

Proof of Lemma 2: Since the posterior distributions of (Γ, Σ) and (μ_2, Ω_{22}) are independent, we have

$$
\begin{aligned}
E[\mu|\Omega, R] &= \begin{pmatrix} \tilde{\alpha} + \tilde{\beta}\hat{\mu}_2 \\ \hat{\mu}_2 \end{pmatrix} && \text{(A24)} \\
E[\Omega|R] &= \begin{pmatrix} E[\beta\Omega_{22}\beta'|R] + \tilde{\Sigma} & \tilde{\beta}\tilde{\Omega}_{22} \\ \tilde{\Omega}_{22}\tilde{\beta}' & \tilde{\Omega}_{22} \end{pmatrix} . && \text{(A25)}
\end{aligned}
$$

It follows from equation (A3) and the law of iterated expectations that

$$
E[\beta\Omega_{22}\beta'|R] = \tilde{\beta}\tilde{\Omega}_{22}\tilde{\beta}' + \mathrm{tr}[F\tilde{\Omega}_{22}]\tilde{\Sigma} \ , \tag{A26}
$$

where F is the $k \times k$ submatrix in the lower-right corner of $(D + Z'Z)^{-1}$. Using equations (A5), (A11), (A14) and $\hat{S}^2 = \hat{\mu}_2'\hat{\Omega}_{22}^{-1}\hat{\mu}_2$, we obtain

$$
\mathrm{tr}(F\tilde{\Omega}_{22}) = \frac{1}{T-k-2}\left[\omega\left(k - \frac{\hat{S}^2}{1+\hat{S}^2}\right) + (1-\omega)k\right] . \tag{A27}
$$

Equation (A18) in the theorem follows easily from (A7) and (A24). Equation (A19) in the theorem follows from (A7), (A8), (A25), (A26) and (A27). This completes the proof of Lemma 2.

B. Proof of Theorem 1

It follows from the law of iterated expectations that

$$
E[r_{T+1}|R] = E[\mu|R] \tag{A28}
$$

$$
\mathrm{var}(r_{T+1}|R) = E[\Omega|R] + \mathrm{var}(\mu|R) \ . \tag{A29}
$$

Equation (13) holds because of equations (A28) and (A18). Given equations (A29) and (A19), we only need to figure out $\mathrm{var}(\mu|R)$. This can be written as

$$
\begin{aligned}
\mathrm{var}(\mu|R) &= \mathrm{var}\left(\begin{pmatrix} \alpha + \beta\mu_2 \\ \mu_2 \end{pmatrix} |R\right) \\
&= \frac{1}{T-k-2}\begin{pmatrix} \tilde{\beta}\hat{\Omega}_{22}\tilde{\beta}' & \tilde{\beta}\hat{\Omega}_{22} \\ \hat{\Omega}_{22}\tilde{\beta}' & \hat{\Omega}_{22} \end{pmatrix} + \begin{pmatrix} E[\mathrm{var}(\alpha + \beta\mu_2|\mu_2, R)|R] & 0_{m\times k} \\ 0_{k\times m} & 0_{k\times k} \end{pmatrix} . && \text{(A30)}
\end{aligned}
$$

The second equality in the above expression follows from

$$
\mathrm{var}(\alpha + \beta\mu_2|R) = \mathrm{var}(E[\alpha + \beta\mu_2|\mu_2, R]|R) + E[\mathrm{var}(\alpha + \beta\mu_2|\mu_2, R)|R] \tag{A31}
$$

as well as equations (A1), (A4), (A5), and (A6). Since $\alpha + \beta\mu_2 = (I_m \otimes (\, 1 \quad \mu_2' \,))\mathrm{vec}(\Gamma)$, it follows from equation (A3) that

$$
\begin{aligned}
\mathrm{var}(\alpha + \beta\mu_2|\mu_2, R) &= (I_m \otimes (\, 1 \quad \mu_2' \,))(\tilde{\Sigma} \otimes (D + Z'Z)^{-1})(I_m \otimes (\, 1 \quad \mu_2' \,))' \\
&= \tilde{\Sigma} \otimes [(\, 1 \quad \mu_2' \,)\, (D + Z'Z)^{-1}\, (\, 1 \quad \mu_2' \,)'] \\
&= \rho\tilde{\Sigma} \ , && \text{(A32)}
\end{aligned}
$$

where $\rho = (1 \quad \mu_2')\,(D + Z'Z)^{-1}\,(1 \quad \mu_2')'$. With equations (A4) and (A6), it is straightforward to show that the posterior mean of ρ is $E[\rho|R] = \tilde{\rho}$, where

$$\tilde{\rho} = \operatorname{tr}\left((D + Z'Z)^{-1}\begin{pmatrix} 1 & \hat{\mu}_2' \\ \hat{\mu}_2 & (T-k-2)^{-1}\hat{\Omega}_{22} + \hat{\mu}_2\hat{\mu}_2' \end{pmatrix}\right) . \tag{A33}$$

One can use equation (A11) and (A14) to show that

$$\tilde{\rho} = \omega\bar{\rho} + (1-\omega)\hat{\rho} \ , \tag{A34}$$

where

$$\bar{\rho} = \frac{k}{T(T-k-2)} + \frac{T-k-3}{T(T-k-2)} \cdot \frac{\hat{S}^2}{1+\hat{S}^2} \ , \tag{A35}$$

$$\hat{\rho} = \frac{T-2}{T(T-k-2)} \ . \tag{A36}$$

Therefore, we have

$$E[\operatorname{var}(\alpha + \beta\mu_2|\mu_2, R)|R] = [\omega\bar{\rho} + (1-\omega)\hat{\rho}]\tilde{\Sigma} \ . \tag{A37}$$

Substituting this into equation (A30), we obtain

$$\begin{aligned} \operatorname{var}(\mu|R) &= \operatorname{var}\left(\begin{pmatrix} \alpha + \beta\mu_2 \\ \mu_2 \end{pmatrix}|R\right) \\ &= \frac{1}{T-k-2}\begin{pmatrix} \tilde{\beta}\hat{\Omega}_{22}\tilde{\beta}' & \tilde{\beta}\hat{\Omega}_{22} \\ \hat{\Omega}_{22}\tilde{\beta}' & \hat{\Omega}_{22} \end{pmatrix} + \begin{pmatrix} [\omega\bar{\rho} + (1-\omega)\hat{\rho}]\tilde{\Sigma} & 0_{m\times k} \\ 0_{k\times m} & 0_{k\times k} \end{pmatrix} \ . \end{aligned} \tag{A38}$$

We then combine equation (A38), (A29) and Lemma 1 to get $\operatorname{var}(R_{T+1}|R)$. Finally, equation (14) is obtained by letting $\bar{\delta} = \bar{\lambda} + \bar{\rho}$ and $\hat{\delta} = \hat{\lambda} + \hat{\rho}$. This completes the proof of Theorem 1.

C. Derivation of the Distribution in (25)

It follows from the definition of $\hat{\Gamma}$ and equation (2) that

$$\hat{\Gamma} = \Gamma + (Z'Z)^{-1}Z'U \ . \tag{A39}$$

Notice that $\alpha = \Gamma'J$ and $\hat{\alpha} = \hat{\Gamma}'J$. We have

$$\hat{\alpha} = \alpha + U'Z(Z'Z)^{-1}J \ , \tag{A40}$$

which implies that $\hat{\alpha}$, conditional on X and the parameters, has a normal distribution with mean α and variance

$$\operatorname{var}(\hat{\alpha}) = E\left[U'Z(Z'Z)^{-1}JJ'(Z'Z)^{-1}Z'U\right] \ . \tag{A41}$$

Let U_i be the i^{th} column of U and σ_{ij} be the element of Σ at i^{th} row and j^{th} column. The covariance between $\hat{\alpha}_i$ and $\hat{\alpha}_j$ can be calculated as

$$\begin{aligned} \mathrm{var}(\hat{\alpha}_i, \hat{\alpha}_j) &= E\left[U_i' Z(Z'Z)^{-1} J J' (Z'Z)^{-1} Z' U_j\right] \\ &= J'(Z'Z)^{-1} J \sigma_{ij} \ . \end{aligned} \tag{A42}$$

Using the formula of the inverse of partitioned matrix, one can show that

$$J'(Z'Z)^{-1} J = (T - \iota' X (X'X)^{-1} X' \iota)^{-1} \ . \tag{A43}$$

It then follows from equation (A14) that

$$\mathrm{var}(\hat{\alpha}_i, \hat{\alpha}_j) = \frac{1}{T}(1 + \hat{S}^2)\sigma_{ij} \ , \tag{A44}$$

which gives

$$\mathrm{var}(\hat{\alpha}) = \frac{1}{T}(1 + \hat{S}^2)\Sigma \ . \tag{A45}$$

Therefore, under the null hypothesis of $\alpha = 0$, we have

$$\hat{\alpha} \,|\, \Sigma \ \sim \ N\left(0_{m\times 1} \ , \ \frac{1}{T}(1 + \hat{S}^2)\Sigma\right) \ . \tag{A46}$$

References

Anderson, Theodore W., 1984, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons.

Ang, Andrew., and Geert Bekaert, 2001, International asset allocation with regime shifts. *Review of Financial Studies*, forthcoming.

Bazaraa, Mokhtar S., 1993, *Nonlinear Programming: Theory and Algorithms.* 2nd edition, John Wiley & Sons.

Bekaert, Geert, and Michael Urias, 1996, Diversification, integration, and emerging market closed-end funds. *Journal of Finance* 51, 835–870.

Berger, James O., 1985, *Statistical Decision Theory and Bayesian Analysis.* 2nd edition, Springer-Verlag.

Bernardo, Jose M., and Adrian F. M. Smith, 1994. *Bayesian Theory.* John Wiley & Sons.

Best, Michael J., and Robert R. Grauer, 1991, On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *Review of Financial Studies* 4, 315–342.

Black, Fischer, and Robert Litterman, 1991, Global asset allocation with equities, bonds, and currencies. *Fixed Income Research*, Goldman Sachs.

Black, Fischer, and Robert Litterman, 1992, Global portfolio optimization. *Financial Analysts Journal*, September–October 1992.

Bodie, Zvi., Alex Kane, and Alan J. Marcus, 1999, *Investments*, 4th edition. Irwin McGraw-Hill.

Bohn, Henning, and Linda L. Tesar, 1996. U.S. equity investment in foreign markets: Portfolio rebalancing or return chasing? *American Economic Review* 86, 77–81.

Brown, Stephen J., 1976, Optimal portfolio choice under uncertainty: a Bayesian approach. Unpublished Ph.D. dissertation, University of Chicago.

Britten-Jones, Mark, 1999, The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance* 54, 655–671.

Chamberlain, Gary, 2001, Minimax estimation and forecasting in a stationary autoregression model. *American Economic Review* 91, 55–59.

Cooper, Ian, and Evi Kaplanis, 1994, Home Bias in Equity Portfolios, Inflation Hedging, and International Capital Market Equilibrium, *Review of Financial Studies* 7, 45-60.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

De Santis, Giorgio, and Bruno Gerard, 1997, International asset-pricing and portfolio diversification with time-varying risk. *Journal of Finance* 52, 1881-1912.

Efron, Bradley, and Carl Morris, 1973, Stein's estimation rule and its competitors — an empirical Bayes approach. *Journal of the American Statistical Association* 68, 117–130.

Elton, Edwin, and Martin Gruber, 1995, *Modern Portfolio Theory and Investment Analysis*, John Wiley & Sons, New York.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.

Fama, Eugen F., and Kenneth R. French, 1997, Industry costs of equity. *Journal of Financial Economics* 43, 153–193.

Fama, Eugen F., and Kenneth R. French, 1998, Value versus growth: the international evidence. *Journal of Finance* 53, 1975–1999.

French, Kenneth R., and James M. Poterba, 1991, Investor Diversification and International Equity Markets, *American Economic Review* 81, 222-26.

Frost, Peter A., and James E. Savarino, 1986, An empirical Bayes approach to efficient portfolio selection. *Journal of Financial and Quantitative Analysis* 21, 293–305.

Gibbons, Michael R., Stephen A. Ross and Jay Shanken, 1989, A test of the efficiency of a given portfolio. *Econometrica* 57, 1121–1152.

Gilboa, Itzhak, and David Schmeidler (1989), Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18, 141–153.

Green, Richard C., and Burton Hollifield, 1992, When will mean-variance efficient portfolios be well diversified? *Journal of Finance* 47, 1785–1809.

Hansen, Lars, and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance* 52, 557–590.

Hansen, Lars Peter, and Thomas J. Sargent, 2001, Robust control and model uncertainty. *American Economic Review*, 91, 60–66.

He, Guanglian, and Robert Litterman, 1999, The intuition behind Black-Litterman model portfolios, *Investment Management Research*, Goldman Sachs.

Hodrick, Robert, David Ng, Paul Sengmueller, 1999. An international dynamic asset pricing model. *International Taxation and Public Finance* 6, 597-620

Hodrick, Robert, and Xiaoyan Zhang, 2001, Assessing specification errors in stochastic discount factor models. *Journal of Financial Economics* 62, 327–376.

Huberman, Gur, 2001. Familiarity breeds investment, *Review of Financial Studies* 14, 659–80

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jobson, J. D., B. Korkie, and V. Ratti, 1979, Improved estimation for Markowitz Portfolios using James-Stein type estimators, *Proceedings of the American Statistical Association, Business and Economics Statistics Section* 41, 279–284.

Jorion, Philippe, 1986, Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21, 279–292.

Jorion, Philippe, 1991, Bayesian and CAPM estimators of the means: implications for portfolio selection. *Journal of Banking and Finance* 15, 717–727.

Lewis, Karen K., 2000, Why do stocks and consumption imply such different gains from international risk sharing? *Journal of International Economics* 52, 1–35.

Li, Kai, Asani Sarkar, and Zhenyu Wang, 2001, Diversification benefits of emerging markets subject to portfolio constraints. *Journal of Empirical Finance*, forthcoming.

Lowenstein, Roger, 2000, *When Genius Failed.* Random House.

Lynch, Anthony W. 2001, Portfolio choice and equity characteristics: Characterizing the hedging demands induced by return predictability. *Journal of Financial Economics* 62, 67–130.

MacKinlay, A Craig., and Lubos, Pastor, 2000, Asset-pricing models: implications for expected returns and portfolio selection. *Review of Financial Studies* 13, 883–916.

Michaud, Richard, 1989, The Markowitz Optimization Enigma: Is Optimized Optimal? *Financial Analyst Journal* 45, 31–42.

Pastor, Lubos, 2000, Portfolio selection and pricing models. *Journal of Finance* 55, 179–223.

Pastor, Lubos, and Robert F. Stambaugh, 1999, Costs of equity capital and model mispricing. *Journal of Finance* 54, 67–121.

Pastor, Lubos, and Robert F. Stambaugh, 2000, Comparing asset-pricing models: an investment perspective. *Journal of Financial Economics* 56, 335–381.

Sion, Maurice, 1958, On general minimax theorems. *Pacific Journal of Mathematics* 8, 171–176.

Table I

Asset Allocation Using the CAPM

This table reports the tangency portfolio weights on assets and factors for various values of the shrinkage factor ω. The assets are the Fama-French portfolios, excluding the five portfolios that contain the largest firms, and the SMB and HML portfolios constructed by Fama and French (1993) according to firms' market capitalization and book-to-market ratio. There is one factor, which is the value-weighted market index return of the NYSE, Amex and Nasdaq. The monthly excess returns on these portfolios during the period from July 1963 to December 1998 are used for estimating the mean and variance of the predictive distribution. The sample Sharpe ratio ($\hat{S}$) of the factor portfolios and the number of time-series observations (T) are provided.

$\hat{S} = .1219, \quad T = 426$								
ω	1.00	.95	.90	.75	.50	.30	.10	.01
Optimal portfolio weights								
MKT	1.00	1.19	2.97	.57	.69	.71	.72	.73
SMB	.00	−.59	−6.03	1.33	.95	.88	.85	.84
HML	.00	−.32	−3.27	.72	.52	.48	.46	.45
S-L	.00	−.36	−3.65	.81	.58	.53	.51	.51
S-2	.00	.13	1.32	−.29	−.21	−.19	−.19	−.18
S-3	.00	−.04	−.36	.08	.06	.05	.05	.05
S-4	.00	.35	3.54	−.78	−.56	−.52	−.50	−.49
S-H	.00	.15	1.57	−.35	−.25	−.23	−.22	−.22
2-L	.00	.02	.16	−.04	−.03	−.02	−.02	−.02
2-2	.00	.07	.75	−.17	−.12	−.11	−.11	−.10
2-3	.00	.19	1.94	−.43	−.31	−.28	−.27	−.27
2-4	.00	.19	1.94	−.43	−.31	−.28	−.27	−.27
2-H	.00	.08	.81	−.18	−.13	−.12	−.11	−.11
3-L	.00	−.21	−2.12	.47	.33	.31	.30	.29
3-2	.00	.07	.72	−.16	−.11	−.10	−.10	−.10
3-3	.00	−.02	−.21	.05	.03	.03	.03	.03
3-4	.00	.12	1.26	−.28	−.20	−.18	−.18	−.18
3-H	.00	.12	1.27	−.28	−.20	−.19	−.18	−.18
4-L	.00	.30	3.08	−.68	−.49	−.45	−.43	−.43
4-2	.00	−.35	−3.55	.78	.56	.52	.50	.49
4-3	.00	−.11	−1.17	.26	.18	.17	.16	.16
4-4	.00	.05	.56	−.12	−.09	−.08	−.08	−.08
4-H	.00	−.05	−.54	.12	.08	.08	.08	.07

Table II
Asset Allocation Using the Fama-French Model

This table reports the tangency portfolio weights on assets and factors for various values of the corresponding shrinkage factor ω. The assets are the Fama-French portfolios constructed by Fama and French (1993) according to firms' market capitalization and book-to-market ratio, excluding the five portfolios that contain the largest firms. There are three factors, which are the SMB and HML risk factors constructed by Fama and French (1993) and the value-weighted market index return of the NYSE, Amex and Nasdaq. The monthly excess returns on these portfolios during the period from July 1963 to December 1998 are used for estimating the mean and variance of the predicative distribution. The sample Sharpe ratio ($\hat{S}$) of the factor portfolios and the number of time-series observations (T) are provided.

$\hat{S} = .2551, \quad T = 426$								
ω	1.00	.95	.90	.75	.50	.30	.10	.01
Optimal portfolio weights								
MKT	.32	.29	.24	−.12	1.46	.89	.77	.74
SMB	.04	−.03	−.13	−.82	2.25	1.15	.92	.86
HML	.64	.66	.68	.84	.12	.38	.43	.45
S-L	.00	−.04	−.10	−.55	1.41	.71	.56	.52
S-2	.00	.02	.04	.20	−.51	−.26	−.20	−.19
S-3	.00	.00	−.01	−.05	.14	.07	.06	.05
S-4	.00	.04	.10	.53	−1.36	−.68	−.54	−.51
S-H	.00	.02	.05	.24	−.61	−.30	−.24	−.22
2-L	.00	.00	.00	.02	−.06	−.03	−.02	−.02
2-2	.00	.01	.02	.11	−.29	−.14	−.11	−.11
2-3	.00	.02	.06	.29	−.75	−.38	−.30	−.28
2-4	.00	.02	.06	.29	−.75	−.38	−.30	−.28
2-H	.00	.01	.02	.12	−.31	−.16	−.12	−.12
3-L	.00	−.03	−.06	−.32	.82	.41	.32	.30
3-2	.00	.01	.02	.11	−.28	−.14	−.11	−.10
3-3	.00	.00	−.01	−.03	.08	.04	.03	.03
3-4	.00	.02	.04	.19	−.49	−.24	−.19	−.18
3-H	.00	.02	.04	.19	−.49	−.25	−.19	−.18
4-L	.00	.04	.09	.46	−1.19	−.60	−.47	−.44
4-2	.00	−.04	−.10	−.53	1.37	.69	.54	.51
4-3	.00	−.01	−.03	−.18	.45	.23	.18	.17
4-4	.00	.01	.02	.08	−.22	−.11	−.09	−.08
4-H	.00	−.01	−.02	−.08	.21	.10	.08	.08

Table III
Asset Allocation with Aversion to Model Uncertainty

For the assets and data considered in Tables I and II, this table presents the portfolio weights in the solutions to the maxmin problems, for various degree of risk aversion γ, when the CAPM or the Fama-French model are used in asset allocation.

	The CAPM				Fama-French Model			
γ	3.0	5.0	7.0	9.0	3.0	5.0	7.0	9.0
MKT	1.13	.96	.91	.86	.32	.32	.32	.31
SMB	−.49	.22	.33	.43	.04	.04	.03	.03
HML	−.20	.06	.08	.14	.64	.64	.64	.65
S-L	.03	−.01	−.05	−.04	.00	.00	.00	.00
S-2	.01	−.01	.01	−.05	.00	.00	.00	.00
S-3	.06	−.06	−.06	.02	.00	.00	.00	.00
S-4	−.01	.07	.03	−.05	.00	.00	.00	.01
S-H	.09	−.07	−.04	−.07	.00	.00	.00	.00
2-L	.03	−.03	−.05	−.02	.00	.00	.00	.00
2-2	.02	.00	.01	.00	.00	.00	.00	.00
2-3	.02	−.02	.02	−.03	.00	.00	.00	−.01
2-4	.08	−.06	−.09	−.05	.00	.00	.00	.01
2-H	.07	−.03	−.04	−.02	.00	.00	.00	.00
3-L	.00	−.02	−.01	−.02	.00	.00	.00	.00
3-2	.02	.01	−.06	−.02	.00	.00	.00	.01
3-3	.06	.05	−.05	−.07	.00	.00	.00	.00
3-4	.04	−.03	.01	−.01	.00	.00	.00	.00
3-H	.06	.00	−.02	−.03	.00	.00	.00	.00
4-L	.01	.00	.02	.00	.00	.00	.00	.00
4-2	.02	−.02	.02	−.01	.00	.00	.00	.01
4-3	−.05	−.03	−.01	.00	.00	.00	.00	.00
4-4	.01	.03	.03	.01	.00	.00	.00	.00
4-H	−.01	.00	.00	.02	.00	.00	.00	.00

Table IV

International Asset Allocation with Bias

This table reports the tangency portfolio weights on the U.S. or G7 countries for various degrees of bias toward the U.S. or G7 countries. The international asset allocation problem considers the tangency portfolio weights on the seven industrial countries (the U.S., Canada, Japan, France, Germany, Italy, and U.K.) and eight emerging markets (Argentina, Brazil, Chile, Mexico, Hong Kong, Singapore, Thailand, and Korea). Portfolio weights are either unconstrained or constrained to be non-negative (no short sales).

Degree of bias to U.S.	1.00	.95	.90	.75	.50	.30	.10	.01
	The weight on the U.S.							
if uncontrained	1.00	.99	.99	.97	.94	.92	.90	.89
if no short sales	1.00	.96	.92	.80	.64	.52	.43	.39
Degree of bias to G7	1.00	.95	.90	.75	.50	.30	.10	.01
	Sum of the weights on G7 countries							
if uncontrained	1.00	.98	.96	.91	.83	.77	.72	.70
if no short sales	1.00	.99	.98	.92	.82	.74	.66	.63

Table V

Diversification Benefits for U.S.– and G7-Biased Investors

For various values of the shrinkage factor (ω_a) assigned to the world CAPM, this table reports the diversification benefit for an investor with various degrees of U.S. and G7 bias (ω_h). The diversification benefit is reported as annual certainty equivalent returns (percent).

ω_a	1.00	.95	.90	.75	.50	.25	.10	.05
ω_h	Benefit for U.S.–biased investors							
	if portfolios are unconstrained							
1.000	.26	.24	.24	.26	.47	.89	1.25	1.39
.975	.25	.24	.23	.25	.44	.84	1.19	1.33
.950	.25	.23	.22	.23	.40	.79	1.13	1.26
.925	.26	.23	.21	.21	.37	.74	1.07	1.20
.900	.26	.23	.21	.20	.34	.69	1.01	1.13
	if no short sales are allowed							
1.000	.26	.24	.23	.24	.36	.59	.79	.86
.975	.25	.23	.22	.23	.33	.56	.75	.82
.950	.25	.23	.22	.21	.31	.52	.71	.78
.925	.24	.22	.21	.20	.29	.49	.67	.74
.900	.24	.22	.20	.19	.27	.46	.63	.70
ω_h	Benefit for G7-biased investors							
	if portfolios are unconstrained							
1.000	.41	.37	.34	.29	.36	.65	.93	1.05
.975	.41	.37	.34	.28	.34	.61	.89	1.00
.950	.42	.37	.33	.27	.31	.57	.84	.95
.925	.42	.38	.33	.26	.29	.54	.79	.90
.900	.43	.38	.34	.25	.27	.50	.75	.85
	if no short sales are allowed							
1.000	.13	.11	.10	.11	.21	.43	.62	.69
.975	.13	.11	.10	.11	.21	.43	.62	.69
.950	.13	.11	.10	.10	.21	.43	.62	.69
.925	.13	.11	.10	.10	.21	.43	.61	.69
.900	.13	.11	.10	.10	.20	.41	.59	.66

Table VI

International Asset Allocation with Aversion to Model Uncertainty

For the assets and data considered in Table IV, this table presents the portfolio weights in the solutions to the maxmin problems, for various degrees of risk aversion γ, when the U.S.–bias model or the world CAPM are used in asset allocation.

	The U.S.–bias Model				The World CAPM			
γ	3.0	5.0	7.0	9.0	3.0	5.0	7.0	9.0
World					1.00	1.00	1.01	1.00
United States	1.00	1.00	1.00	1.00	.00	.00	.00	.00
Canada	.00	.00	.00	.00	.00	.00	.00	.00
Japan	.00	.00	.00	.00	.00	.00	.00	.00
France	.00	.00	.00	.00	.00	.00	.00	.00
Germany	.00	.00	.00	.00	.00	.00	.00	.00
Italy	.00	.00	.00	.00	.00	.00	.00	.00
United Kingdom	.00	.00	.00	.00	.00	.00	.00	.00
Argentina	.00	.00	.00	.00	.00	.00	.00	.00
Brazil	.00	.00	.00	.00	.00	.00	.00	.00
Chile	.00	.00	.00	.00	.00	.00	.00	.00
Mexico	.00	.00	.00	.00	.00	.00	.00	.00
Hong Kong	.00	.00	.00	.00	.00	.00	.00	.00
Singapore	.00	.00	.00	.00	.00	.00	.00	.00
Korea	.00	.00	.00	.00	.00	.00	.00	.00
Thailand	.00	.00	.00	.00	.00	.00	.00	.00

Figure 1

Sharpe Ratio as a Function of the Shrinkage Factor

The highest Sharpe ratio of the efficient frontier spanned by the assets and factors is plotted as a function of the shrinkage factor ω. The solid line is for the CAPM with the assets considered in Table I. The dotted line is for the Fama-French model with the assets considered in Table II. The dash line is for the U.S.–bias model with the assets considered in Table VI. The dot-dash line is for the world CAPM with the assets considered in Table VI.

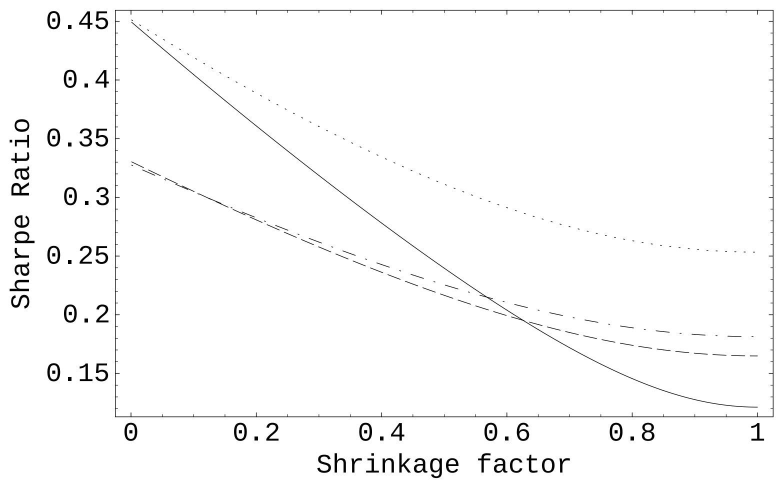

Figure 2
The Shrinkage Factor as a Function of σ

The shrinkage factor ω is plotted as a function of σ (in annual percentage returns) in the prior distribution. The solid line is for the CAPM with the assets considered in Table I. The dotted line is for the Fama-French model with the assets considered in Table II.