A Shrinkage Approach to Model Uncertainty and Asset Allocation

Zhenyu Wang*

Columbia University
Graduate School of Business

September 5, 2003†

Abstract

This paper takes a shrinkage approach to examine the empirical implications of aversion to model uncertainty. The shrinkage approach explicitly shows how predictive distributions incorporate data and prior beliefs. It enables us to solve the optimal portfolios for uncertainty-averse investors. Aversion to uncertainty about the CAPM leads investors to hold a portfolio that is not mean-variance efficient for any predictive distribution. However, mean-variance efficient portfolios corresponding to extremely strong beliefs in the Fama-French model are approximately optimal for uncertainty-averse investors. The empirical Bayes approach does not deliver optimal portfolios when investors are averse to uncertainty. Uncertainty aversion does not justify U.S. investors' home bias, and diversification benefit is robust to uncertainty about the world CAPM.

JEL Classification: G11, G12, G15, C11

Keywords: model uncertainty, portfolio choice, Bayesian analysis, maxmin analysis, international finance, home bias

---

*For helpful comments, I thank Torben Andersen, Andrew Ang, Geert Bekaert, Mark Carhart, Fangruo Chen, Wayne Ferson, Larry Epstein, Silverio Foresi, Larry Glosten, Rick Green, Bob Hodrick, Michael Johannes, Bob Litterman, Anthony Lynch, Lubos Pastor, David Robinson, Tano Santos, Rob Stambaugh, an anonymous referee, and seminar participants at Columbia University, Northwestern University, Lehman Brothers, Goldman Sachs, University of Minnesota, Federal Reserve Bank of Atlanta, University of North Carolina at Chapel Hill, University of Texas at Austin, University of Houston, American Finance Association Meeting in 2001, and National Bureau of Economic Research Summer Institute in 2002. 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

This paper takes a shrinkage approach to examine the empirical implications of aversion to model uncertainty. The shrinkage approach explicitly shows how predictive distributions incorporate data and prior beliefs. It enables us to solve the optimal portfolios for uncertainty-averse investors. Aversion to uncertainty about the CAPM leads investors to hold a portfolio that is not mean-variance efficient for any predictive distribution. However, mean-variance efficient portfolios corresponding to extremely strong beliefs in the Fama-French model are approximately optimal for uncertainty-averse investors. The empirical Bayes approach does not deliver optimal portfolios when investors are averse to uncertainty. Uncertainty aversion does not justify U.S. investors' home bias, and diversification benefit is robust to uncertainty about the world CAPM.

JEL Classification: G11, G12, G15, C11

Keywords: model uncertainty, portfolio choice, Bayesian analysis, maxmin analysis, international finance, home bias

Contents:

1. Introduction

In standard finance theory, investors optimally allocate their investment funds to 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 interest. This issue has been well known as Knightian uncertainty in the academic literature. It 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 are wrong. When models are likely to be wrong and the correct model is unknown, investors may be averse to the uncertainty. I examine the empirical implications of investor's aversion to model uncertainty in an asset allocation problem.

Finance professionals often impose restrictions of equilibrium models when specifying the distribution of asset returns. A major reason 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 (Best and Grauer (1991), Britten-Jones (1999), Green and Hollifield (1992), Michaud (1989), Jobson and Korkie (1980), and Polson and Tew (2000)). The resulting portfolio usually contains extremely long and short positions, which are difficult to implement. To solve this problem, portfolio weights are often constrained to be non-negative, but the optimal portfolios are usually corner solutions that assign zero weights to many assets and nonzero weights to a few highly risky assets. Such corner solutions conflict with investors' intuition about the benefit of diversification. More importantly, the optimal portfolio weights are extremely sensitive to variations in the estimated mean and covariance matrix. Another major reason for imposing restrictions of asset-pricing models is that asset prices inconsistent with market equilibrium are believed to be unlikely to persist. Black and Litterman (1991, 1992) discuss reasons for imposing asset-pricing models.

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. 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. Shanken (1987) argues that no model is correct and suggests that empirical Bayesian approach is a more appropriate way to study and use models.

Bayesian inference, which combines the prior beliefs in models and the information in data, is currently a popular approach in empirical studies of model uncertainty (Frost and Savarino (1986), Pastor and Stambaugh (2000), and Polson and Tew (2000)). Since different prior beliefs lead to different posterior and predictive probability distributions of asset returns, researchers have tried various prior distributions of the pricing errors to see how they affect asset allocation. In practice, researchers often take the empirical Bayes approach, in which prior distributions are estimated from the observed samples. In the current literature, Bayesian analysis of asset allocation does not incorporate investors' aversion to uncertainty in prior beliefs.

Investors may have multiple prior beliefs in models and may be averse to model uncertainty. 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 asset allocations when they are uncertain about their prior belief in asset-pricing models. This paper helps us gain more understanding of the empirically observed asset allocations. In the standard Bayesian analysis of asset allocation, the 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 leads to an optimal portfolio for uncertainty-averse investors.

The framework for aversion to model uncertainty is the maxmin analysis, in which preference is represented by the minimum of the expected utility over a 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. The maxmin principle has been a tradition in Bayesian decision theory. Berger (1985) offers a detailed overview on maxmin analysis of Bayesian decisions. Using an axiomatic approach, Gilboa and Schmeidler (1989) further 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 maxmin analysis, this paper examines asset allocations associated with the models and data that attract significant attention in the empirical finance literature. The models are the CAPM related to the mean-variance utility function and Fama-French model derived from empirical experience. The data are the returns on equity portfolios. Many recent papers apply maxmin analysis to finance[1] to examine theoretical issues. I bring uncertainty aversion into empirical studies of asset allocation.

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 ratio, there exists no prior belief in the CAPM such that the tangency portfolio with respect to the predictive distribution gives the optimal allocation for uncertainty-averse investors. This can be viewed as a failure of the CAPM in the context of asset allocation. With considerable doubts about both estimates, uncertainty-averse investors do not want to rely on either the estimate restricted by the CAPM or the unrestricted estimate. 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 is potentially an important reason.

For some models and data, however, there exist prior beliefs such 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 model is optimal or approximately 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 ratio and gives a more precise 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 investors' 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 risk and returns. When discussing the possible explanations for home bias, French and Poterba (1991) write, "The statistical

[1] For examples, see Epstein and Wang (1994), Chamberlain (2000), Hansen and Sargent (2001), Routledge and Zin (2001), Kogan and Wang (2002), Maenhout (2002) and Uppal and Wang (2003).

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, aversion to model uncertainty does not justify home bias.

In order to apply maxmin analysis to asset allocation using asset-pricing models, a shrinkage approach is developed in this paper. I show 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. The predictive mean and variance share the same shrinkage factor, which indicates the relation 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, 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 mean and variance. The shrinkage approach allows me to solve the maxmin problem and to obtain the properties of the solutions in the maxmin analysis.

The rest of the paper is organized as follows. Section 2 gives an overview of the classic, Bayesian, and maxmin analysis in the context of asset allocation using asset-pricing models. Section 3 develops the shrinkage approach to the Bayesian and maxmin analysis. Section 4 compares the maxmin analysis with the empirical approach. Section 5 examines domestic asset allocations. Section 6 investigates international asset allocation and the issue of home bias. Conclusions and future research are discussed in Section 7. Mathematical derivations are provided in the Appendix.

2. Aversion to Model Uncertainty

In this section, we review the classic, Bayesian, and maxmin analysis in the context of asset allocation using asset-pricing models. 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 (Berger (1985) and Bernardo and Smith (1994)).

In Bayesian framework, asset allocation without uncertainty aversion 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 parameter θ. 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:

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

Notice that $x(\theta)$ is not a vector of portfolio weights because the elements of the vector do not sum up to 1. In order to examine the return of the optimal portfolio, we need to scale $x(\theta)$ to portfolio weights. That is, the vector of portfolio weights is $\tilde{x}(\theta) = x(\theta)/(1'_{m+k}x(\theta))$, where 1_{m+k} is the $m+k$ dimension vector of ones. It is obvious that the vector of portfolio weights is independent of the risk aversion parameter γ because γ is canceled out. It is well known that the risky portfolio $\tilde{x}(\theta)$ maximizes the Sharpe ratio. This portfolio of risky assets is referred to as the tangency portfolio in the finance literature.

Suppose that the investor has uncertainty about the correct model, and is averse to such uncertainty. Such a case can be modeled as the investor having a set of priors parameterized in terms of $\theta \in [0, +\infty)$, and the utility function $\min_{\theta \in [0,+\infty)} U(x, \theta)$ (see Gilboa and Schmeidler (1989)). The optimal portfolio choice for such an investor is to solve

$$\max_{x} \min_{\theta \in [0,+\infty)} U(x; \theta) \ . \tag{9}$$

In order to show that the solution x^* to this maxmin problem exists and is unique, we employ the shrinkage approach developed in the next section. The solution x^* is sometimes referred to as robust portfolio in the literature. The asset allocation of x^* is optimal 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 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). The solution x^* and its portfolio weights depend on the risk aversion parameter γ.

The set of predictive distributions obtained by changing θ is not a convex set, i.e., a weighted average of two predictive probability distribution functions is not necessarily a predictive distribution. However, Gilboa and Schmeidler (1989) require the set of probability distributions in consideration to be convex in order to show that uncertainty aversion implies a maxmin problem. We can consider the convex hull of the predictive distributions and then apply the implication from uncertainty aversion to maxmin problems established by Gilboa and Schmeidler. Since a maxmin problem implies behavior of uncertain aversion over any set of probability distributions, we can then use the maxmin problem over the set of predictive distributions that is not convex. In fact, the maxmin principle had been used in Bayesian decisions to express aversion to uncertainty (see Berger (1985)) before Gilboa and Schmeidler (1989) established the equivalence axiomatically under the convexity assumption.

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

$$\min_{\theta \in [0,+\infty)} \max_{x} U(x; \theta) \ . \tag{10}$$

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

$$\min_{\theta\in[0,+\infty)} E[r_{T+1}|R,\theta]' \left(V[r_{T+1}|R,\theta]\right)^{-1} E[r_{T+1}|R,\theta] \ . \tag{11}$$

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 choosing parameter θ.

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 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 problems 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 optimal for investors averse to model uncertainty. However, when the maxmin and minmax problems are not equivalent, there is no prior such that the tangency portfolio given the prior is optimal for uncertainty-averse investors.

3. The Shrinkage Approach

The shrinkage approach to be developed in this section is new. 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 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. 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*. This result is consistent with the well-known observation that posterior mean is a Bayes-Stein estimator. 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 frameworks for 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 a prior belief influences 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, 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 same Sharpe ratio as 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 (a 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. For example, if the investor thinks that the model and the data are equally important, then they should choose $\omega = 1/2$. If the investor thinks that the data are twice as important as the model, they should choose $\omega = 1/3$. This is consistent with industry practice, in which people often take average estimates obtained from different models. For example, the asset management group at Goldman Sachs averages the estimate from equilibrium models and estimate from the views of investors (He and Litterman (1999)).

A major distinction of Theorem 1 from other shrinkage estimation in the literature is that the shrinkage factor for the variance is the same as that for the mean. The work by Polson, Johannes and Stroud (2002), Polson, Gron and Jorgensen (2001), Polson, Jacquier and Rossi (2002), and MacKinlay and Pastor (2000) demonstrate that it is important to estimate variance precisely and consider the relation between mean and variance implied by asset pricing models. Theorem 1 shows that the predictive mean and variance both depend on the same shrinkage factor. It exhibits the importance of the predictive variance through the relation between the predictive mean and variance implied by an asset-pricing model. In the next section, it will be further demonstrated that shrinkage of 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.[2] 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 variance is closely related to risk. Bayesian inference of variance and its shrinkage formulation are considered in a few studies,[3] 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 that do not reflect any relation between the mean and variance.

Theorem 1 allows us to obtain useful properties of the maxmin problem and to 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

$$\max_{x} \min_{\omega\in[0,1]} U(x;\omega) \ . \tag{21}$$

Since $U(x;\omega)$ is continuous in $(x;\omega)$ and the set of possible ω is compact, the function $\tilde{U}(x) \equiv \min_\omega U(x;\omega)$ is continuous in x. Since $U(x;\omega)$ is strictly concave in x, the function $\tilde{U}(x)$ 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 exist but may not be unique.

Theorem 1 also makes it easy to solve for the optimal portfolio x^* in the maxmin problem. It follows that

$$E[r_{T+1}|R,\omega] \quad = \quad \begin{pmatrix} \hat{\mu}_1 + \omega(\bar{\beta}\hat{\mu}_2 - \hat{\mu}_1) \\ \hat{\mu}_2 \end{pmatrix} \tag{22}$$

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

[3] See Brown (1976), Frost and Savarino (1986) and Jagannathan and Ma (2002).

$$V[r_{T+1}|R,\omega] = \begin{pmatrix} \Psi_0 + \Psi_1\omega + \Psi_2\omega^2 & b[\hat{\beta} + \omega(\bar{\beta} - \hat{\beta})]\hat{\Omega}_{22} \\ b\hat{\Omega}_{22}[\hat{\beta} + \omega(\bar{\beta} - \hat{\beta})]' & b\hat{\Omega}_{22} \end{pmatrix} \tag{23}$$

where

$$\Psi_0 = b\hat{\beta}\hat{\Omega}_{22}\hat{\beta}' + h\hat{\delta}\hat{\Sigma} \tag{24}$$

$$\Psi_1 = b(\bar{\beta} - \hat{\beta})\hat{\Omega}_{22}\hat{\beta}' + b\hat{\beta}\hat{\Omega}_{22}(\bar{\beta} - \hat{\beta})' + h(\bar{\delta} - \hat{\delta})\hat{\Sigma} + h\hat{\delta}(\bar{\Sigma} - \hat{\Sigma}) \tag{25}$$

$$\Psi_2 = b(\bar{\beta} - \hat{\beta})\hat{\Omega}_{22}(\bar{\beta} - \hat{\beta})' + h(\bar{\delta} - \hat{\delta})(\bar{\Sigma} - \hat{\Sigma}) \tag{26}$$

Then, the predictive mean and variance can be written as

$$E[r_{T+1}|R,\omega] = A_0 + \omega A_1 \tag{27}$$

$$V[r_{T+1}|R,\omega] = B_0 + \omega B_1 + \omega^2 B_2 \tag{28}$$

where

$$A_0 = \begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{pmatrix} \tag{29}$$

$$A_1 = \begin{pmatrix} \bar{\beta}\hat{\mu}_2 - \hat{\mu}_1 \\ 0_{k\times 1} \end{pmatrix} \tag{30}$$

$$B_0 = \begin{pmatrix} \Psi_0 & b\hat{\beta}\hat{\Omega}_{22} \\ b\hat{\Omega}_{22}\hat{\beta}' & b\hat{\Omega}_{22} \end{pmatrix} \tag{31}$$

$$B_1 = \begin{pmatrix} \Psi_1 & b(\bar{\beta} - \hat{\beta})\hat{\Omega}_{22} \\ b\hat{\Omega}_{22}(\bar{\beta} - \hat{\beta})' & 0_{k\times k} \end{pmatrix} \tag{32}$$

$$B_2 = \begin{pmatrix} \Psi_2 & 0_{m\times k} \\ 0_{k\times m} & 0_{k\times k} \end{pmatrix} \tag{33}$$

Substituting equations (27) and (28) into the utility function, we obtain

$$U(x,\omega) = U_0(x) + \omega U_1(x) + \omega^2 U_2(x) \tag{34}$$

where

$$U_0(x) = x'A_0 - \frac{1}{2}\gamma\, x'B_0 x \tag{35}$$

$$U_1(x) = x'A_1 - \frac{1}{2}\gamma\, x'B_1 x \tag{36}$$

$$U_2(x) = -\frac{1}{2}\gamma\, x'B_2 x \tag{37}$$

Basic algebra for quadratic functions implies that $\tilde{U}(x) = \min_\omega U(x,\omega)$ has the following analytical expression:

$$\tilde{U}(x) = \begin{cases} U_0(x) - U_1^2(x)/(4U_2(x)) & \text{if} \quad x \in \mathcal{X} \\ \min\{U_0(x), U_0(x) + U_1(x) + U_2(x)\} & \text{if} \quad x \notin \mathcal{X} \end{cases} \tag{38}$$

where the set $\mathcal{X}$ is defined as

$$\mathcal{X} = \{x : \; U_2(x) > 0, \; 0 < -0.5 U_1(x)/U_2(x) < 1\} \tag{39}$$

Since $\tilde{U}(x)$ is strictly concave, multivariate optimization programs work well in finding x^* on computer.

The shrinkage is critical for obtaining the properties and solutions of the maxmin problem. It also allows us to check if the utility function $U(x, \omega)$ is convex in ω so that we know if we can apply the MiniMax Theorem. The function is convex in ω for any x if and only if $U_2(x)$ is non-negative for any x. This is true if and only if the matrix B_2, or equivalently Ψ_2, is negative semi-definite. Since the nature of Ψ_2 depends on the observed data, convexity of the function $U(x, \omega)$ in ω depends on the observed data. We cannot impose assumptions on the shape of $U(x, \omega)$ because such assumptions can be violated by the observed data. In empirical applications, we should examine eigenvalues of Ψ_2 because Ψ_2 is negative semi-definite if and only if all eigenvalues are non-positive.

It is important to keep in mind that the Minimax Theorem is only a sufficient but not necessary condition for the maxmin and minmax problems to be equivalent. Therefore, if the function $U(x, \omega)$ is not convex in ω for every x, we need to solve the minmax problem to find out whether the maxmin and minmax problems are equivalent. We are interested in the equivalent of the two problems because it tells us whether an optimal portfolio for uncertainty averse investors can be a tangency portfolio given a particular value of θ in the prior. Using the shrinkage factor ω, the minmax problem becomes

$$\min_{\omega \in [0,1]} \max_{x} U(x; \omega) \; . \tag{40}$$

A solution $\omega^{\dagger}$, referred to as a conservative shrinkage factor in this paper, solves

$$\min_{\omega} \; E[r_{T+1}|R, \omega]' \left(V[r_{T+1}|R, \omega] \right)^{-1} E[r_{T+1}|R, \omega] \; . \tag{41}$$

The solution $\omega^{\dagger}$ exists because the objective function in (41) is continuous and the set of ω is compact. For any given ω, a tangent portfolio determined by the predictive mean and variance is

$$x(\omega) = \gamma^{-1} \left(V[r_{T+1}|R, \omega] \right)^{-1} E[r_{T+1}|R, \omega] \; . \tag{42}$$

The portfolio choice corresponding to the minmax solution is $x^{\dagger} = x(\omega^{\dagger})$.

When the maxmin and minmax problems have different solutions, the optimal portfolio for uncertainty-averse investors is not in the set of tangency portfolios, $\mathcal{P}$, which is defined as $\mathcal{P} = \{x :$

$x = x(\omega)$ for some $\omega \in [0,1]\}$. If investors are restricted to the set of tangency portfolios, there will be a loss to them in terms of utility $\min_\omega U(x,\omega)$. The loss can be calculated as

$$L = \max_x \min_\omega U(x,\omega) - \max_{x\in\mathcal{P}} \min_\omega U(x,\omega) \ . \tag{43}$$

4. A Comparison to Empirical Bayes Approach

Bayesian analyses of asset allocation in the current literature do not use maxmin analysis to incorporate aversion to model uncertainty. They often use the empirical Bayes approach to specify prior beliefs. In the standard 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 to estimate 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 same set of data used for computing the posterior distribution. Berger (1985) and Bernardo and Smith (1994) provide overviews of the compound decision problems. Since the empirical Bayes approach is often applied in asset allocation, 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,

$$\hat{\alpha} \,|\, \Sigma \quad \sim \quad N\left(0_{m\times 1}, \frac{1}{T}(1+\hat{S}^2)\Sigma\right) \ . \tag{44}$$

This sampling distribution is the multi-factor version of formula (5.3.16) in Campbell, Lo and MacKinley (1997), which has only one factor. It follows from equation (6) that the parameter in the prior distribution of α should be specified as

$$\theta = T^{-1}\left(1+\hat{S}^2\right) \ . \tag{45}$$

Substituting this choice of θ into equation (12) gives $\omega = 1/2$. This important result is stated in the following theorem:

THEOREM 2: *If the prior distribution of* α *is chosen to satisfy* $V[\alpha|\Sigma] = V[\hat{\alpha}|\Sigma]$*, the shrinkage factor is* $\omega = 1/2$*.*

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. Most applications in asset allocation deviate from the standard compound decision problem by choosing θ different from equation (45). For example, one can estimate a prior variance for alpha using an earlier or different set of data. Alternatively, one can estimate a prior variance of alpha using the cross-sectional variance of estimated alphas in a large cross-section of assets. Even when some people choose θ without estimation, they tend to make it more or less consistent with our empirical experience with the data. No matter how a prior variance is inferred from data, it implies a shrinkage factor. If a prior variance of alpha is close to empirical estimate, the implied shrinkage factor is close to 1/2.

It seems that the empirical Bayes approach is biased toward models — the shrinkage factor is always 1/2, no matter how many data are observed and used. This is caused by setting the prior mean of Jensen's alpha to zero. One may let $\alpha \,|\, \Sigma \sim N\,(\lambda\ ,\ \theta\Sigma)$, where λ is an $m \times 1$ vector, in the prior distribution and take the empirical Bayes approach to determine both the hyperparameters λ and θ. The resulting predictive mean and variances will be close to the sample estimates. In this case, there is no faith in the asset-pricing model, and the model is in fact not used in asset allocation. Using classical statistical inference to determine both λ and θ can be criticized as non-Bayesian because it relies solely on data. We are only interested in the case of $\lambda = 0$, in which the informative prior distribution of Jensen's alpha imposes a prior belief in the model. Choosing a prior distribution for Jensen's alpha with a zero mean is equivalent to choosing the model as reference. The setup of this kind of prior distribution allows the model to exert influence on the posterior and predictive distributions. The built-in influence of the model is the purpose of this paper's exercise — asset allocation *using asset-pricing models*.

In the empirical Bayes approach, researchers sometimes use different sets of data to estimate the prior and posterior distributions. For example, a researcher may use earlier observations to estimate the prior distribution and use later observations to estimate the posterior distribution. Or, a researcher may construct her prior distribution using a different set of assets. Alternatively, a researcher might form her prior opinion by reading other people's research that uses either different set of assets or different period of observations. This kind of empirical Bayes approach can be different from the compound decision problem.

Suppose the prior distribution is estimated by a set of observed returns $R_o = (Y_o, X_o)$ on some assets and factors. This set of data can be entirely different from, overlap with or coincide with the data $R = (Y, X)$, which are used for estimating the posterior and predictive distributions. Although the factors in the two sets of data must be the same, the actual observed data X_o and X can be different if the two sets of data cover different periods. Suppose there are T_o observations and m_o assets in the set of data R_o. Let $\hat{\Sigma}_o$ be the estimated variance of the residual in the regression of Y_o on X_o, and $\hat{S}_o$ the highest sample Sharpe ratio of the factor portfolios estimated from X_o. It follows from (44) that

$$\hat{\alpha} \,|\, \Sigma \;\sim\; N\left(0_{m_o \times 1}, \frac{1}{T_o}(1 + \hat{S}_o^2)\Sigma\right) \; . \tag{46}$$

Therefore, we should choose $\theta = T_o^{-1}(1 + \hat{S}_o^2)$.

With the above choice of θ, the shrinkage factor will depend on the sample size and the factor Sharpe ratio. Substituting the above θ into equation (12), we obtain

$$\omega = \left(1 + \frac{T}{T_o} \cdot \frac{1 + \hat{S}_o^2}{1 + \hat{S}^2}\right)^{-1} \tag{47}$$

Since the factor Sharpe ratios are usually much smaller than 1, the ratio $(1 + \hat{S}_0^2)/(1 + \hat{S}^2)$ is not very different from 1. The shrinkage factor is then mainly determined by the relative size of the two sets of samples, i.e., $\omega \approx (1 + T/T_o)^{-1}$. If more observations are used for estimating the prior variance of Jensen's alpha, more weight will be assigned to the estimate restricted by the asset-pricing model. When researchers arbitrarily determine which data are used for the prior and which are used for the posterior, the resulting shrinkage factor is arbitrary. One should be cautious about the empirical Bayes approach that is not a compound decision problem.

5. Domestic Asset Allocation

Let us first look at allocations when 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. The data are obtained from Kenneth

French's Web page. 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 1 presents the tangency portfolio weights for various values of the shrinkage factor ω. As expected, most of the weight is 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. The excess returns on the assets are the monthly excess returns on the Fama-French portfolios considered in Table 1. The tangency portfolio weights for various values of ω are reported in Table 2. 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 1 and 2 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 is scaled into a vector of portfolio weights and reported in columns 2–5 of Table 3. The risk aversion γ is set to be 3, 5, 7, or 9 in this paper.[4] 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 important as the loss caused by the uncertainty about the unrestricted estimate. The optimal allocations by uncertainty-averse investors in no way resemble the tangency portfolios obtained in Table 1. For the CAPM, there is no prior belief that gives the allocation optimal 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 effects of aversion to uncertainty about the Fama-French

[4] Bodie et al. (1999, page 191–193) calibrated that $\gamma = 2.96$ for U.S. investors. Similarly, Pastor and Stambaugh (2000) calibrated that $\gamma = 2.84$.

model. The solutions to the maxmin problems are reported in columns 6–9 of Table 3. 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 2. Therefore, when the Fama-French model is under consideration, an investor 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 that aversion to 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 uncertainty about the unrestricted predictive distribution forces the investor to rely more on the Fama-French model. The optimal asset allocation with uncertainty aversion is also very different from the tangency portfolio obtained from the empirical Bayes approach ($\omega = 0.5$) when investors use the Fama-French model.

For uncertainty-averse investors, I calculate the utility loss L, defined in equation (43), due to the restriction of tangency portfolios. When the CAPM is used for asset-allocation, the utility loss L is 1.38%, 2.24%, 0.49%, and 5.88%, respectively, for $\gamma = 3$, 5, 7, and 9. In comparison, when the Fama-French model is used for asset allocation, the utility loss L is zero for $\gamma = 3$ and 5. For $\gamma = 7$ and 9, the utility loss L is not zero but are only 0.02% and 0.05% respectively. This further demonstrates that the tangency portfolio obtained by imposing the Fama-French model is optimal or approximately optimal for uncertainty-averse investors when they consider asset allocation over the portfolios sorted by firm size and book-to-market ratio.

Since the Fama-French model is constructed to fit mainly the data during 1970s and 1980s while the CAPM is motivated from theory, it is natural to examine the models using the data of 1990s. Table 4 presents the results using the data of 1990s. When the CAPM is the reference model, the optimal portfolios for uncertainty-averse investors constructed from the data of 1990s are not drastically different from those constructed from the whole sample. When Fama-French model is used, the optimal portfolios for uncertainty-averse investors using the data of 1990s are very different from the portfolios using the whole sample. The main change is in the weights among the three factor portfolios. However, all the optimal portfolios are heavily concentrated on the factor portfolios if the degree of risk aversion is high. Therefore, the tangency portfolio of Fama-French model still offers a good approximation to the optimal portfolio for uncertain averse investors with high degree of risk aversion.

In the above analysis, the maxmin problem is solved directly. The MiniMax theorem cannot be applied to the problems discussed in this paper because $U(x;\omega)$ is not a convex function in ω for the problems and data discussed in this paper. As pointed out in section 3, the function is convex in ω if and only if all the eigen values of Ψ_2 are non-positive. Numerical calculation shows that the largest eigenvalue for Ψ_2 is positive for both the CAPM and Fama-French model. Since the conditions in the MiniMax Theorem are sufficient but not necessary, the maxmin and minmax problems may happen to be equivalent. 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 curve is strictly decreasing, the conservative shrinkage factor is $\omega^{\dagger} = 1$. A similar curve is plotted for the Fama-French model (the dotted curve). This curve is also decreasing and thus the conservative shrinkage factor is $\omega^{\dagger} = 1$. 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 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.

Results presented in this section shows that, when investors are restricted in the set of available predictive distributions, the maxmin and minmax problems are not equivalent for the case of CAPM, and the optimal portfolio solved from the maxmin problem is not a tangency portfolio. This fact is true even if we consider the convex hull of predictive distributions in the maxmin problem. If investors are able to choose probability distributions from the convex hull,[5] the maxmin and minmax problems are still not equivalent in general because the function $U(x,\omega)$ is not always convex in ω. If the maxmin and minmax problems turn out to be equivalent in the convex hull of predictive distributions, the optimal portfolio must be a tangency portfolio for some probability distribution that is not in the set of predictive distributions. This still implies that none of the

[5] In practice, however, there is no obvious way to figure out what prior beliefs generate the additional probability distributions in the convex hull.

tangency portfolios obtained from the available predictive distributions are optimal for uncertainty-averse investors.

Besides enlarging the set of predictive distributions to its convex hull, we can also reduce the set of probability distributions. A smaller set of possible probability distributions represents less uncertainty. Investors may focus on a smaller set of prior beliefs in a model. Particularly interesting sets of available predictive distributions are those obtained by restricting ω between some lower and upper bounds, say $\underline{\omega}$ and $\overline{\omega}$. If we set $\overline{\omega} = 1$, higher $\underline{\omega}$ represents less uncertainty about the belief in a model. Table 5 presents the portfolio weights in the solutions to

$$\max_{x} \min_{\omega \in [\underline{\omega}, \overline{\omega}]} U(x, \omega)$$

where $\overline{\omega} = 1$ and $\underline{\omega}$ =0.25, 0.50, and 0.75. The degree of risk aversion is set to be 3. When the CAPM is used, the change of the set of ω substantially affects uncertainty-averse investor's optimal portfolio. When Fama-French model is used, the optimal portfolio weights are similar to those in Table 3 as long as the set contains $\omega = 1$. Since one might want to center their choice of predictive distributions on the empirical approach, Table 5 also presents the optimal portfolio weights with $\underline{\omega} = 0.25$ and $\overline{\omega} = 0.75$. These portfolios are very different from the optimal portfolios with ω ranging between 0 and 1.

Besides helping to solve the maxmin problem, the shrinkage approach developed in this paper offers additional insights into some Bayesian analysis in the literature on asset allocation. To examine the impact of model uncertainty on asset allocation, Pastor and Stambaugh (2000) represent θ in terms of another parameter σ by $\theta = \sigma^2/s^2$, where s^2 is defined as $s^2 = \text{trace}(\hat{\Sigma})/m$. They interpret σ^2 as the average prior variance of Jensen's α and 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 explains why they observe that modest value of σ substantially reduces models' impact on portfolio choice. The predictive distribution assigns a small weight to the model when σ increases modestly. The reason for the fast drop of the shrinkage factor 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 in 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 shrinkage of variance is important. The weights in Tables 1 and 2 are, however, nonlinear with respect to the shrinkage factor. In Table 1, 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 2, 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 (Polson and Tew (2000)).

6. 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 (Bohn and Tesar (1996), Cooper and Kaplanis (1994), and French and Poterba (1991)). 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:

$$r_{\text{foreign}} = \alpha + \beta r_{\text{US}} + \epsilon \qquad \epsilon \sim N(0, \Sigma) \ . \tag{48}$$

An "asset-pricing model" can be introduced by setting $\alpha = 0$ in equation (48). 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. Pastor (2000a) has applied this model to analyze home bias and referred to it as domestic CAPM. 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 halfway 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 G7-bias model. The index portfolios in emerging markets are treated as the assets. If the portfolios of industrial countries span the efficient frontier of all the markets, the G7-bias model should correctly price the portfolios of emerging markets, in which case investors only allocate their wealth to 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 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 6 reports the tangency portfolio weights on the U.S. for various degrees of U.S. bias. It also reports the sum of the tangency portfolio weights on G7 countries for various degrees of G7 bias.

We can extract investors' degree of bias by comparing Table 6 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 tangency portfolio weight on the U.S. is about 40 percent when $\omega = 0.01$. Since U.S. investors' actual foreign equity holdings account for only about 8 percent of their total equity holdings, Table 6 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 $\omega = 0.9$, 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.–bias model.

Given the strong home bias of U.S. investors, there are many studies on the existence and magnitude of international diversification benefits. However, different studies often take different approaches separately. Some studies examine the efficient frontier spanned by the international markets.[6] Other studies assume the world version of the CAPM that specifies the world portfolio as the efficient portfolio.[7] Still other studies consider the effects of short-sale constraints on international diversification.[8] 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.[9] It is also well known to be difficult for investors to take 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.

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

[6] See Bekaert and Urias (1996), Harvey (1995) and Lewis (1999).

[7] See Harvey (1991) and De Santis and Gerard (1997).

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

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

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

$$\Delta U = \left(\mu_a' x_a - \frac{1}{2}\gamma x_a'\Omega_a x_a\right) - \left(\mu_a' x_h - \frac{1}{2}\gamma x_h'\Omega_a x_h\right) . \tag{49}$$

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

$$\Delta U = \gamma^{-1}\left(\mu_a'(\Omega_a^{-1}\mu_a - \Omega_h^{-1}\mu_h) - \frac{1}{2}(\mu_a'\Omega_a^{-1}\mu_a - \mu_h'\Omega_h^{-1}\Omega_a\Omega_h^{-1}\mu_h)\right). \tag{50}$$

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 7 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 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 ($\omega_a = 0.5$) to construct his/her prior belief in the world CAPM, he/she concludes that the increase in 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 in knowing 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 diversification benefit measured in certainty equivalent return is about 20 basis points per annum.

In summary, three important observations can be made from Table 7. 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.

Short-sale constraints reduce the magnitude of diversification benefits, 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 7 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 diversification benefit is positive, U.S. investors do not diversify over international markets. One possible reason is investors' aversion to uncertainty about the world equilibrium model. To examine this issue, the maxmin problem for uncertainty about the world CAPM is solved directly by numerical search.[10] The solution suggests investing 100 percent in the world market portfolio, as shown in Table 8 for various degrees of risk aversion γ. It turns out that the solution to each maxmin problem is the same as the solution to the corresponding minmax problem. In Figure 3, for the world CAPM, the shrinkage factor of $\omega = 1$ gives the lowest optimal 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 less important than the loss caused by the uncertainty about the unrestricted estimate. According to Hodrick, Ng and Sengmueller (1999), the world CAPM usually explains the country indices well. The model is strongly rejected only when test assets include international book-to-market portfolios. Since the actual weight of the U.S. in the world market portfolio is only around 50 percent, aversion to uncertainty about this world equilibrium

[10] 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.

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 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 8 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 optimal Sharpe ratio is at $\omega = 1$, as shown in Figure 3. 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.

7. 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 if investors are averse 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 a given prior distribution, the weight on the asset-pricing model is larger if the frontier of the factor portfolios has a 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 shrinkage factor derived in this 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 us understand the empirical Bayes approach in this context. For example, the standard 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 equilibrium model of the world markets. 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.

The approach to home bias indicates that shrinkage factors for models are flexible for analyzing portfolio advices. 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 bias toward portfolios other than the U.S. and G7 country indices.

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 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 multi-period settings when returns are predictable. For example, Ang and Bekaert (2000) and Lynch (2000) examine the dynamic optimal portfolios of multiple risky assets when asset returns are predictable. When returns are not distributed identically and independently over time, it would be useful to incorporate dynamic equilibrium models into asset allocation decisions for multi-period settings. 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.

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 3 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 $D = \theta^{-1}JJ'$, 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:

$$E[\Gamma|\Sigma, R] = (D + Z'Z)^{-1}Z'Z\hat{\Gamma} = \tilde{\Gamma} \equiv (\tilde{\alpha}, \tilde{\beta})' \quad \text{(A1)}$$

$$E[\Sigma|R] = (T\hat{\Sigma} + \hat{\Gamma}'Q\hat{\Gamma})/(T - m - k - 1) \equiv \tilde{\Sigma} \quad \text{(A2)}$$

$$\text{var}(\text{vec}(\Gamma)|R) = \tilde{\Sigma} \otimes (D + Z'Z)^{-1} \quad \text{(A3)}$$

$$E[\mu_2|\Omega_{22}, R] = \hat{\mu}_2 \quad \text{(A4)}$$

$$E[\Omega_{22}|R] = \hat{\Omega}_{22}T/(T - k - 2) \equiv \tilde{\Omega}_{22} \quad \text{(A5)}$$

$$\text{var}(\mu_2|R) = \hat{\Omega}_{22}/(T - k - 2)\ , \quad \text{(A6)}$$

where $Q = Z'(I_T - Z(D + Z'Z)^{-1}Z')Z$, and the notation "≡" 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:

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

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

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

$$E[\mu|\Omega, R] = \begin{pmatrix} \tilde{\alpha} + \tilde{\beta}\hat{\mu}_2 \\ \hat{\mu}_2 \end{pmatrix} \tag{A24}$$

$$E[\Omega|R] = \begin{pmatrix} E[\beta\Omega_{22}\beta'|R] + \tilde{\Sigma} & \tilde{\beta}\tilde{\Omega}_{22} \\ \tilde{\Omega}_{22}\tilde{\beta}' & \tilde{\Omega}_{22} \end{pmatrix} . \tag{A25}$$

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

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

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

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

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

$$\tilde\rho = \mathrm{tr}\left((D+Z'Z)^{-1}\begin{pmatrix}1 & \hat\mu_2'\\ \hat\mu_2 & (T-k-2)^{-1}\hat\Omega_{22}+\hat\mu_2\hat\mu_2'\end{pmatrix}\right)\ . \tag{A33}$$

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

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

Substituting this into equation (A30), we obtain

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

We then combine equation (A38), (A29) and Lemma 1 to get $\text{var}(R_{T+1}|R)$. 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 (44)

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

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

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

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

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

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

which gives

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

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.

Campbell, J., A. Lo, and A. MacKinlay, 1997, The Econometrics of Financial Markets, Princeton University Press.

Chamberlain, Gary, 2000, Econometric applications of maxmin expected utility. *Journal of Applied Econometrics* 15, 625–644.

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.

Epstein, Larry G., and Tan Wang, 1994, International asset pricing under Knightian uncertainty. *Econometrica* 62, 283-322.

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.

Ferson, Wayne E., and Campbell R. Harvey, 1999, Conditioning Variables and the Cross Section of Stock Returns, *Journal of Finance* 54, 1325-1360.

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, Edward, Jose A. Lopez, and Zhenyu Wang, 2000, The Federal Reserve Banks' Imputed Cost of Equity Capital. Federal Research Bank of San Francisco Working Paper.

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

Gron, A., B. Jorgensen, and Nicholas G. Polson, 2001, Portfolio Choice and Stochastic Volatility.

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.

Harvey, Campbell R., 1991, The World Price of Covariance Risk. *Journal of Finance* 46, 111–158.

Harvey, Campbell R., 1995, Predictable Risk and Returns in Emerging Markets. *Review of Financial Studies* 8, 773–816.

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 Tongshu Ma, 2002. Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance*, forthcoming.

Jobson, J. D., B. Korkie, 1980, Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association* 75, 544–554.

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.

Kogan, Leonid, and Tan Wang, 2002, A simple theory of asset pricing under model uncertainty. Working paper, MIT.

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.

Maenhout, Pascal J., 2002, Robust portfolio rules, hedging and asset pricing, INSEAD, Working paper.

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.

Polson, Nicholas G., and Bernard V. Tew, 2000, Bayesian Portfolio Selection: An Empirical Analysis of the S&P 500 Index 1970–1996. *Journal of Business and Economic Statistics* 18, 164–173.

Polson, Nicholas G., Eric Jacquier and Peter E. Rossi, 2002, Bayesian Analysis of Stochastic Volatility Models. *Journal of Business and Economic Statistics* 20, 69–87.

Polson, Nicholas G., 2002, A Family of Distributions for a Covariance Matrix. Working paper, University of Chicago.

Polson, Nicholas G., Michael Johannes and J. Stroud, 2002, Sequintial Optimal Portfolio Allocation: Market and Volatility Timing. Working paper, University of Chicago.

Routledge, Bryan R., and Stanley E. Zin, 2000, Model uncertainty and liquidity. Working paper, Carnegie Mellon University.

Shanken, Jay, 1987, A Bayesian Approach to Testing Portfolio Efficiency. *Journal of Financial Economics* 19, 195-215

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

Uppal, R. and T. Wang (2002), Model Misspecification and Under Diversification, *Journal of Finance*, forthcoming.

Table 1
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.

$\hat{S} = .1219$, $T = 426$								
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 2

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 3
Asset Allocation with Aversion to Model Uncertainty

For the assets and data considered in Tables 1 and 2, 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 is 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	.97	.98	.89	.32	.32	.32	.31
SMB	−.42	.20	.08	.38	.04	.04	.04	.03
HML	−.20	.05	.02	.10	.64	.64	.64	.64
S-L	.03	.00	−.02	−.02	.00	.00	.00	.00
S-2	.00	−.01	−.01	−.04	.00	.00	.00	.00
S-3	.03	−.07	.01	.01	.00	.00	.00	.00
S-4	.02	.05	−.02	−.11	.00	.00	.00	.00
S-H	.08	−.04	.01	.00	.00	.00	.00	.00
2-L	.02	−.03	.00	−.05	.00	.00	.00	.00
2-2	.02	−.02	.00	−.01	.00	.00	.00	.00
2-3	.02	−.02	−.01	.00	.00	.00	.00	.00
2-4	.05	−.05	−.04	−.02	.00	.00	.00	.00
2-H	.06	−.01	.00	−.05	.00	.00	.00	.00
3-L	.01	.00	−.01	−.01	.00	.00	.00	.00
3-2	.02	.02	.00	−.04	.00	.00	.00	.00
3-3	.05	.01	−.03	−.06	.00	.00	.00	.00
3-4	.03	−.04	.00	.00	.00	.00	.00	.00
3-H	.04	.00	.01	−.04	.00	.00	.00	.00
4-L	.00	−.02	.02	.02	.00	.00	.00	.00
4-2	−.01	.00	.00	.00	.00	.00	.00	.00
4-3	.02	.00	.02	−.01	.00	.00	.00	.00
4-4	.01	.01	−.01	.03	.00	.00	.00	.00
4-H	.00	.00	.00	.03	.00	.00	.00	.00

Table 4

Asset Allocation with Uncertainty Aversion Using the 1990s' Data

For the assets considered in Tables 1 and 2 and using the data from January 1990 to December 1998, 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 is 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	1.09	1.06	1.03	1.05	.65	.71	.73	.72
SMB	−.29	−.28	−.05	−.22	−1.30	−.45	−.27	−.29
HML	−.14	−.10	−.06	−.12	.87	.63	.58	.59
S-L	.01	.01	.01	.01	.04	.01	.00	.00
S-2	.02	.03	−.03	−.01	.03	.01	.00	.00
S-3	.02	.00	.01	.02	.04	.01	.00	.00
S-4	.02	.01	−.02	.01	.02	.00	.00	.01
S-H	.03	.04	.04	.05	.09	.02	−.01	−.01
2-L	.02	.02	.01	.00	.03	.00	.00	.00
2-2	.04	.02	.01	.01	.06	.01	.00	.00
2-3	−.02	.00	.00	.01	.01	.00	.00	.00
2-4	.04	.05	.02	.03	.13	.02	−.01	.00
2-H	.06	.04	.00	.03	.13	.01	−.01	.00
3-L	.00	.01	−.01	.01	.00	.00	.00	.00
3-2	.02	.02	.00	.02	.04	.00	.00	.00
3-3	.02	.00	.01	.01	.02	.00	.00	.00
3-4	.04	.04	.01	.01	.09	.01	.00	.00
3-H	.03	.02	.03	.03	.07	.01	.00	.00
4-L	−.02	.00	−.01	.00	−.01	.00	.00	.00
4-2	.00	.02	−.01	.01	.02	.00	.00	.00
4-3	−.01	.01	.00	.02	.01	.00	.00	.00
4-4	.02	.00	.01	.01	.02	.00	.00	.00
4-H	−.02	−.02	−.01	.00	−.04	.00	.00	.00

Table 5
Uncertainty Averse Asset Allocation with Subsets of Priors

For the assets and data considered in Tables 1 and 2, this table presents the portfolio weights in the solutions to the maxmin problems, for various bounds ($\underline{\omega}$ and $\overline{\omega}$) of shrinkage factors, when the CAPM or the Fama-French model is used in asset allocation. The degree of risk aversion γ is set to be 3.

	The CAPM				Fama-French Model			
$\overline{\omega}$	1.00	1.00	1.00	.75	1.00	1.00	1.00	.75
$\underline{\omega}$	.25	.50	.75	.25	.25	.50	.75	.25
MKT	1.07	1.03	.99	.57	.30	.29	.32	−.14
SMB	−.25	−.11	.03	1.34	.00	−.01	.02	−.86
HML	−.10	−.04	.01	.72	.66	.67	.65	.86
S-L	.02	.01	.00	.81	.00	.01	.00	−.56
S-2	.00	.00	−.01	−.29	.00	.01	.00	.20
S-3	.01	.01	.00	.08	.01	.00	.00	−.05
S-4	.00	.01	.01	−.79	.00	.01	.00	.54
S-H	.05	.01	−.01	−.35	.01	.00	.01	.25
2-L	.01	.00	.00	−.04	.01	.00	.00	.03
2-2	.02	.01	.01	−.16	.00	.01	.00	.11
2-3	.01	.01	.00	−.43	−.01	.00	.00	.30
2-4	.04	.02	.00	−.43	.01	.00	−.01	.31
2-H	.03	.02	.00	−.18	.01	.01	.00	.13
3-L	.01	.01	.00	.47	.01	.01	.01	−.33
3-2	.00	.01	−.01	−.16	.01	.00	.01	.11
3-3	.02	.01	−.01	.04	.00	.01	.00	−.03
3-4	.02	.01	.00	−.28	.01	.02	.00	.19
3-H	.02	.01	.00	−.28	.00	.00	.00	.20
4-L	−.01	.00	.00	−.68	.00	−.01	−.01	.48
4-2	.02	.00	.00	.79	.00	−.01	−.01	−.55
4-3	−.01	.00	.00	.26	.00	.00	.00	−.19
4-4	.00	−.01	.00	−.12	−.01	.00	.00	.09
4-H	.00	.00	−.01	.12	.00	.00	.00	−.08

Table 6
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 7
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 8
International Asset Allocation with Aversion to Model Uncertainty

For the assets and data considered in Table 6, 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

Optimal 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 1. The dotted line is for the Fama-French model with the assets considered in Table 2.

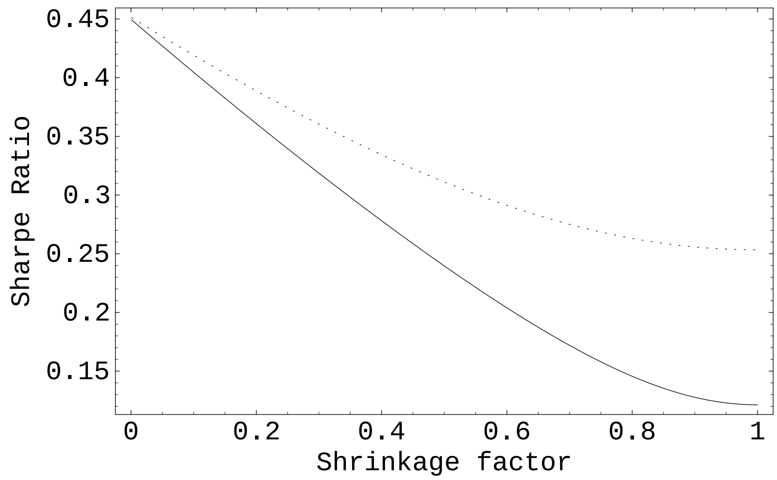

Figure 2

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 1. The dotted line is for the Fama-French model with the assets considered in Table 2.

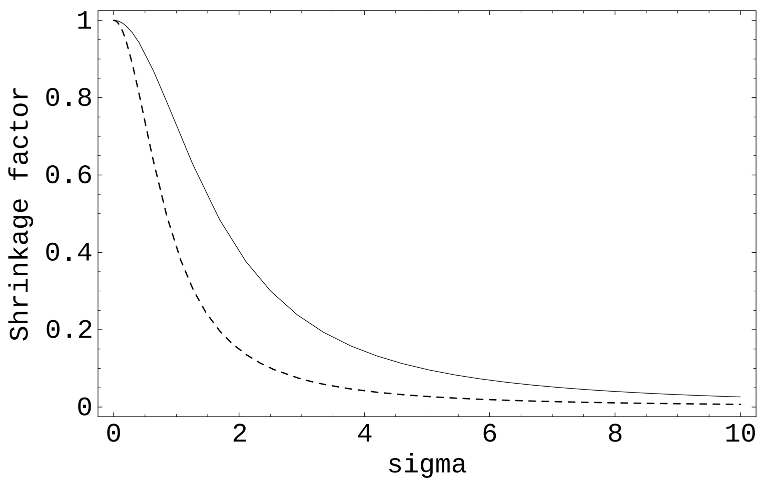

Figure 3

Optimal 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 dotted line is for the U.S.–bias model with the assets considered in Table 8. The solid line is for the world CAPM with the assets considered in Table 8.