Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach

Lorenzo Garlappi
University of Texas at Austin

Raman Uppal
London Business School and CEPR

Tan Wang
University of British Columbia and CCFR

We develop a model for an investor with multiple priors and aversion to ambiguity. We characterize the multiple priors by a "confidence interval" around the estimated expected returns and we model ambiguity aversion via a minimization over the priors. Our model has several attractive features: (1) it has a solid axiomatic foundation; (2) it is flexible enough to allow for different degrees of uncertainty about expected returns for various subsets of assets and also about the return-generating model; and (3) it delivers closed-form expressions for the optimal portfolio. Our empirical analysis suggests that, compared with portfolios from classical and Bayesian models, ambiguity-averse portfolios are more stable over time and deliver a higher out-of sample Sharpe ratio. (*JEL* G11)

Expected returns, variances, and covariances are the key inputs of every portfolio selection model. These parameters are not known a priori and are usually estimated with error. The classical mean-variance approach to portfolio selection estimates the moments of asset returns via their sample counterparts and ignores the estimation error. The outcome of this process are portfolio weights that entail extreme positions in the assets, fluctuate substantially over time, and deliver abysmal out-of-sample performance.[1]

---

[1] For a discussion of the problems entailed in implementing mean-variance optimal portfolios, see Hodges and Brealey (1978), Michaud (1989), Best and Grauer (1991), and Litterman (2003).

The standard method adopted in the literature to deal with estimation error is to use a Bayesian approach, where the unknown parameters are treated as random variables. A Bayesian decision-maker combines a pre-specified prior over the parameters with observations from the data to construct a predictive distribution of returns. Bayesian optimal portfolios then maximize expected utility, where the expectation is taken with respect to the predictive distribution.

The Bayesian decision-maker, however, is assumed to have only a *single prior* or, equivalently, to be *neutral* to uncertainty in the sense of Knight (1921). Given the difficulty in estimating moments of asset returns, the sensitivity of portfolio weights to the choice of a particular prior, and the substantial evidence from experiments that agents are not neutral to ambiguity [Ellsberg (1961)], it is important to consider investors with *multiple priors* who are *averse* to this ambiguity[2] and hence desire *robust* portfolio rules that work well for a set of possible models.[3]

In this article, we examine the normative implications of parameter and model uncertainty for investment management, using a model that allows for multiple priors and where the decision-maker is averse to ambiguity. Our main contribution is to demonstrate how this model can be applied to the practical problem of portfolio selection if expected returns are estimated with error, and to compare explicitly the portfolio weights from this approach with those from the mean-variance and traditional Bayesian models.[4]

In our model, we show that the portfolio selection problem of an ambiguity-averse fund manager can be formulated by making two additions to the standard mean-variance model: (1) imposing an *additional constraint* on the mean-variance portfolio optimization program that restricts the expected return for each asset to lie within a specified confidence interval of its estimated value, and (2) introducing an *additional minimization* over the set of possible expected returns subject to the additional constraint. The additional constraint recognizes the possibility

---

[2] The aversion to ambiguity is particularly strong in cases where people feel that their competence in assessing the relevant probabilities is low [Heath and Tversky (1991)] and when subjects are told that there may be other people who are more qualified to evaluate a particular risky position [Fox and Tversky (1995)]. Gilboa and Schmeidler (1989), Epstein and Wang (1994), Anderson, Hansen, and Sargent (2000), Chen and Epstein (2002), and Uppal and Wang (2003) develop models of decision-making that allow for multiple priors where the decision-maker is not neutral to ambiguity.

[3] There are two terms used to describe aversion to Knightian uncertainty in the literature. Epstein and coauthors describe this as "ambiguity aversion," whereas Hansen and Sargent (and their coauthors) describe it as "wanting robustness." Both streams of the literature have as their origin the multi-priors model of Gilboa and Schmeidler (1989), and both streams agree that ambiguity aversion and wanting robustness are terms describing the same concept; see, for instance, Chen and Epstein (2002, p. 1405).

[4] We focus on the error in estimating expected returns of assets because as shown by Merton (1980) they are much harder to estimate than the variances and covariances. Moreover, Chopra and Ziemba (1993) estimate the cash-equivalent loss from the use of estimated rather than true parameters. They find that errors in estimating expected returns are over 10 times as costly as errors in estimating variances, and over 20 times as costly as errors in estimating covariances.

of estimation error; that is, the point estimate of the expected return is not the only possible value of the expected return considered by the investor. The additional minimization over the estimated expected returns reflects the investor's aversion to ambiguity; that is, in contrast to the standard mean-variance model or the Bayesian approach, in the model we consider, the investor is not neutral toward ambiguity.[5]

To understand the intuition underlying the multi-prior model, consider the case in which expected returns are estimated via their sample counterparts. Because of the constrained minimization over expected returns, if the confidence interval of the expected return of a particular asset is large (that is, the mean is estimated imprecisely), then the investor relies less on the estimated mean and hence reduces the weight invested in this asset. When this interval is small, the minimization is constrained more tightly, and hence, the portfolio weight is closer to the standard weight that one would get from a model that ignores estimation error. In the limit, if the confidence interval is zero for the expected returns on all the assets, the optimal weights are those from the classical mean-variance model.

Our formulation of the portfolio selection model with multiple priors and ambiguity aversion (AA) has several attractive features. First, just like the Bayesian model, the multi-prior model has solid axiomatic foundations—the max-min characterization of the objective function is consistent with the multi-prior approach advocated by Gilboa and Schmeidler (1989) and developed in a static setting by Dow and Werlang (1992) and Kogan and Wang (2002), in dynamic discrete-time by Epstein and Wang (1994), and in continuous time by Chen and Epstein (2002).

Second, in several economically interesting cases, we show that the model with ambiguity aversion can be simplified to a mean-variance model but where the expected return is adjusted to reflect the investor's ambiguity about its estimate. The analytic expressions we obtain for the optimal portfolio weights allow us to provide insights about the effects of parameter and model uncertainty if investors are ambiguity averse. In one special case, we show that the optimal portfolio weights can be interpreted as a weighted average of the classical mean-variance portfolio and the minimum-variance portfolio, with the weights depending on the precision with which expected returns are estimated *and* the investor's aversion to ambiguity. This special case is of particular importance because it allows us to compare the ambiguity-averse model of this article with the traditional Bayesian approach, where the decision-maker is neutral to ambiguity. The analytic solutions also indicate how the model with

---

[5] See Section 1 and Bewley (1988) for a discussion of how confidence intervals obtained from classical statistics are related to Knightian uncertainty and Section 2 for the relation of our model to Bayesian models of decision-making.

ambiguity aversion can be implemented as a simple maximization problem instead of a much more complicated saddle-point problem.

Third, the multi-prior model with ambiguity aversion is flexible enough to allow for the case where the expected returns on all assets are estimated jointly and also where the expected returns on assets are estimated in subsets. The estimation may be undertaken using classical methods such as maximum likelihood or using a Bayesian approach. Moreover, the framework can incorporate both parameter and model uncertainty; that is, it can be implemented if one is estimating expected returns using only sample observations of the realized returns or if one assumes a particular factor model for returns such as the Capital Asset Pricing Model (CAPM) or the Arbitrage Pricing Theory (APT) and is ambiguous about this being the true model.

Finally, the model with ambiguity aversion does not introduce ad-hoc short sale constraints on portfolio weights that rule out short positions even if these were optimal under the true parameter values. Instead, the constraints are imposed because of the investor's aversion to parameter and model uncertainty. At the same time, our formulation can accommodate real-world constraints on the size of trades or position limits.[6]

To demonstrate the differences between the portfolios selected by an ambiguity-averse investor and the ones selected by an ambiguity-neutral investor using either classical or Bayesian estimates of the moments, we apply the multi-prior model with ambiguity aversion to the problem of a fund manager who is allocating wealth across eight international equity indices. In the first application, there is uncertainty only about the expected returns on these indices, and in the second application, there is both parameter uncertainty and uncertainty about the return-generating model. For both applications, we characterize the properties of the portfolio weights under the ambiguity-averse model and compare them to the standard mean-variance portfolio that ignores estimation error and the Bayesian portfolios that allow for estimation error but are neutral to ambiguity. We find that the portfolio weights using the multi-prior model with ambiguity aversion are less unbalanced and vary much less over time compared with the mean-variance portfolio weights. Moreover, allowing for a small degree of ambiguity in the parameter estimates results in out-of-sample Sharpe ratios that are usually greater than those of the mean-variance and Bayesian portfolios.

It is important to point out that the portfolio choice model we propose is particularly appropriate for investors with a stable and significant degree of ambiguity aversion. For investors who are not significantly ambiguity averse, uncertainty about return distributions can be dealt with using a

---

[6] In addition to the features described above, the multi-prior approach with ambiguity aversion is consistent with any utility function, not just utility defined over mean and variance. Our focus on the mean-variance objective function is only because of our desire to compare our results with those in this literature.

more traditional Bayesian, ambiguity neutral, approach. A general property of the portfolios delivered by our model is that they are conservative, because they tend to over-weight the "safe" asset in the optimal allocation. The safe asset, however, does not have to be the traditional riskless asset. For instance, in the absence of a riskless asset, it is the minimum-variance portfolio that ignores estimates of expected returns, and hence, it is not subject to ambiguity about expected returns. More generally, the safe asset could be any benchmark portfolio relative to which performance is being measured. Thus, the portfolios recommended by our model are ones that would be used by conservative investors. Aversion to ambiguity is only one possible origin of this conservative behavior. More broadly, the desire for conservative behavior could also be driven by institutional reasons, arising, for instance, when compensation is based on performance relative to a benchmark, or because of the fiduciary responsibility of managers of pension funds who face serious concerns about underfunding, or, in circumstances where investors face large downside risk.

Our article is closely related to several papers in the literature on portfolio decisions that are robust to model uncertainty or incorporate aversion to ambiguity.[7] Goldfarb and Iyengar (2003), Halldórsson and Tütüncü (2000), and Tütüncü and Koenig (2004) develop algorithms for solving max-min saddle-point problems numerically and apply the algorithms to portfolio choice problem, whereas Wang (2005) shows how to obtain the optimal portfolio numerically in a Bayesian setting in the presence of aversion to model uncertainty. Our article differs from those of Goldfarb and Iyengar (2003), Halldórsson and Tütüncü (2000), and Tütüncü and Koenig (2004) in serval respects. First, we incorporate not only parameter uncertainty but also model uncertainty. Second, we introduce joint constraints on expected returns instead of only individual constraints. In contrast to Wang (2005), uncertainty in our model is characterized by "confidence intervals" around the estimate of expected returns instead of a set of priors with different precisions. This modeling device allows us to obtain, in several cases, not just numerical solutions but also closed-form expressions for the optimal portfolio weights, which enables us to provide an economic interpretation of the effect of aversion to ambiguity.

The rest of the article is organized as follows. In Section 1, we show how one can formulate the problem of portfolio selection for a fund manager who is averse to parameter and model uncertainty. In Section 2, we discuss the relation of the multi-prior model with ambiguity aversion to the traditional

---

[7] Other approaches for dealing with estimation error are to impose arbitrary portfolio constraints prohibiting short sales [Frost and Savarino (1988) and Chopra (1993)], which, as shown by Jagannathan and Ma (2003), can be interpreted as shrinking the extreme elements of the covariance matrix, and the use of resampling based on simulations advocated by Michaud (1998). Scherer (2002) and Harvey et al. (2003) describe the resampling approach in detail and discuss some of its limitations. Black and Litterman (1990, 1992) propose an approach that combines two sets of priors—one based on an equilibrium asset-pricing model and the other based on the subjective views of the investor.

ambiguity-neutral Bayesian approach for dealing with estimation error, and we compare analytically the portfolio weights under the two approaches. Then, in Section 3, we illustrate the out-of-sample properties of the model with ambiguity aversion by considering the case of an investor who allocates wealth across eight international equity-market indices. Our conclusions are presented in Section 4. Proofs for propositions are collected in the Appendix.

1. Portfolio Choice with Ambiguity Aversion

This section is divided into two parts. In the first part, Section 1.1, we summarize the standard mean-variance model of portfolio choice where estimation error is ignored. In the second part, Section 1.2, we show how this model can be extended to incorporate aversion to ambiguity about the estimated parameters and the return-generating model. Throughout the article, we will use the terms "uncertainty" and "ambiguity" equivalently.

We have made a conscious decision to use as a starting point of our analysis the static mean-variance portfolio model of Markowitz (1952) rather than the dynamic portfolio selection model of Merton (1971). There are three reasons for this choice. First, our motivation is to relate the model with ambiguity aversion to the ambiguity-neutral Bayesian models of decision-making that have been considered in the investments literature, which are typically set in a static setting [see, for instance, Jorion (1985, 1986, 1991, 1992), Pástor (2000), and Pástor and Stambaugh (2000)]. Second, considering the static portfolio model allows us to derive explicit expressions for the optimal portfolio weights, and hence, show more clearly how to implement the idea of ambiguity aversion and the benefits from doing so. Finally, in many cases (but not all), the optimal portfolio in the dynamic model is very similar to the portfolio in the static model. The reason for this is that the difference between the portfolios from the static and dynamic models, that is, the "intertemporal hedging component," turns out to be quite small once the model is calibrated to realistic values for the parameters driving the processes for asset returns.[8]

1.1 The classical mean-variance portfolio model
According to the classical mean-variance model [Markowitz (1952, 1959), Sharpe (1970)], the optimal portfolio of N risky assets, w, is given by the solution of the following optimization problem:

$$\max_{w} w^{\top}\mu - \frac{\gamma}{2}w^{\top}\Sigma w, \tag{1}$$

---

[8] The case of dynamic portfolio choice with only a single risky asset in a robust control setting is addressed by Maenhout (2004); the case of multiple risky assets if investors are averse to ambiguity is considered by Chen and Epstein (2002), Epstein and Miao (2003), and Uppal and Wang (2003).

where μ is the N-vector of the true expected *excess* returns over the risk-free asset, Σ is the $N \times N$ covariance matrix, and the scalar γ is the risk aversion parameter. The solution to this problem is

$$w = \frac{1}{\gamma} \Sigma^{-1} \mu. \tag{2}$$

In the absence of a risk-free asset, the problem faced by the investor has the same form as Equation (1), but now μ represents the vector of true expected returns (instead of excess returns) and the portfolio weights have to sum to 1. The solution in this case is

$$w = \frac{1}{\gamma} \Sigma^{-1} \left(\mu - \mu^0 \, \mathbf{1}_N \right), \tag{3}$$

where μ^0 is the expected return on the "zero-beta" portfolio associated with the optimal portfolio w and is given by

$$\mu^0 = \frac{B - \gamma}{A}, \tag{4}$$

where $A = \mathbf{1}_N^\top \Sigma^{-1} \mathbf{1}_N$, $B = \mu^\top \Sigma^{-1} \mathbf{1}_N$, and $\mathbf{1}_N$ is a N-vector of 1's.

A fundamental assumption of the standard mean-variance portfolio selection model in Equation (1) is that the investor knows the true expected returns. In practice, however, the investor needs to estimate expected returns. Denoting the estimate of expected returns by $\hat{\mu}$, the actual problem that the investor solves is

$$\max_{w} w^\top \hat{\mu} - \frac{\gamma}{2} w^\top \Sigma w, \tag{5}$$

subject to $w^\top \mathbf{1}_N = 1$. The problem in Equation (5) coincides with Equation (1) only if expected returns are estimated with infinite precision, that is, $\hat{\mu} = \mu$. In reality, however, expected returns are notoriously difficult to estimate. As a result, portfolio weights obtained from solving Equation (5) tend to consist of extreme positions that swing dramatically over time. Moreover, these portfolios often perform poorly out of sample.

1.2 Extension of the standard model to incorporate AA
To explicitly take into account that expected returns are estimated imprecisely, we introduce two new components into the standard mean-variance portfolio selection problem described above. First, we impose an *additional constraint* on the mean-variance optimization program that restricts the expected return for each asset to lie within a specified

confidence interval of its estimated value. This constraint implies that the investor recognizes explicitly the possibility of estimation error; that is, the point estimate of the expected return is not the only possible value considered by the investor. Second, we introduce an *additional optimization*—the investor *minimizes* over the choice of expected returns and/or models subject to the additional constraint. This minimization over expected returns, μ, reflects the investor's aversion to ambiguity [Gilboa and Schmeidler (1989)].

With the two changes to the standard mean-variance model described above, the model takes the following general form:

$$\max_{w} \min_{\mu} w^{\top}\mu - \frac{\gamma}{2} w^{\top}\Sigma w, \tag{6}$$

subject to

$$f(\mu,\hat{\mu},\Sigma) \leq \epsilon, \tag{7}$$

$$w^{\top}\mathbf{1}_{N} = 1. \tag{8}$$

As before, Equation (8) constrains the weights to sum to unity in the absence of a risk-free asset; if a risk-free asset is available, this constraint can be dropped. In Equation (7), $f(\cdot)$ is a vector-valued function that characterizes the constraint, and ϵ is a vector of constants that reflects both the investor's ambiguity and his aversion to ambiguity. Specifically, the parameter ϵ should be understood as the *product* of ambiguity aversion (common across assets) and ambiguity (asset-specific). Hence, although ambiguity aversion is a general property of preferences, ambiguity is phenomenon-specific.[9] Because the ambiguity-aversion parameter of the investor is not observationally separable from the level of ambiguity in our model, we choose a parsimonious characterization of preferences by normalizing the degree of ambiguity-aversion to 1.[10] In our model, the investor does not choose the degree of ambiguity aversion but only the asset-specific level of ambiguity. As we will see in the next section, the parameter ϵ may then be *interpreted* in a classical statistics sense as the size of a confidence interval, provided we restrict the set of priors to be Gaussian.

---

[9] This distinction is conceptually similar to the distinction between risk aversion, which is a general property of preferences, and risk, which is specific for each asset.

[10] See Ghirardato, Maccheroni, and Marinacci (2004) and Klibanoff, Marinacci, and Mukerji (2005) for models where ambiguity aversion is potentially separable from ambiguity; these models, however, require the decision-maker to specify subjective probabilities over the priors.

In the rest of this section, we illustrate several possible specifications of the constraint given in Equation (7) and their implications for portfolio selection.

1.2.1 Uncertainty about expected returns estimated asset by asset.
We start by considering the case where $f(\mu,\hat{\mu},\Sigma)$ has N components,

$$f_j(\mu,\hat{\mu},\Sigma) = \frac{(\mu_j - \hat{\mu}_j)^2}{\sigma_j^2/T_j}, \quad j = 1,...,N, \tag{9}$$

where T_j is the number of observations in the sample for asset j. In this case, the constraint in Equation (7) becomes

$$\frac{(\hat{\mu}_j - \mu_j)^2}{\sigma_j^2/T_j} \le \epsilon_j, \quad j = 1,...,N. \tag{10}$$

The constraints (10) have an immediate interpretation as confidence intervals. For instance, it is well known that if returns are Normally distributed, then $\frac{\hat{\mu}_j - \mu_j}{\sigma_j/\sqrt{T_j}}$ follows a Normal distribution. Thus, the ϵ_j in constraints (10) determine confidence intervals. When all the N constraints in Equation (10) are taken together, Equation (10) is closely related to the probabilistic statement

$$P(\mu_1 \in I_1,...,\mu_N \in I_N) = 1 - p, \tag{11}$$

where I_j, $j = 1,...,N$, are intervals in the real line and p is a significance level. For instance, if the returns are independent of each other and if p_j is the significance level associated with ϵ_j, then the probability that all the N true expected returns fall into the N intervals, respectively, is $1 - p = (1 - p_1)(1 - p_2) \cdots (1 - p_N)$.[11]

Although confidence intervals or significance levels are often associated with hypothesis testing in statistics, Bewley (1988) shows that they can be interpreted also as a measure of the level of uncertainty associated with the parameters estimated. An intuitive way to see this is to envision an econometrician who estimates the expected returns for an investor. He can report to the investor his best estimates of the expected returns. He can, at the same time, report the uncertainty of his estimates

---

[11] When the asset returns are not independent, the calculation of the confidence level of the event involves multiple integrals. In general, it is difficult to obtain a closed-form expression for the confidence level.

by stating, say, that the confidence level of $\mu_j \in I_j$ for all $j = 1,\ldots, j = N$ is 95%.

To gain some intuition regarding the effect of uncertainty about the estimated mean on the optimal portfolio weight, one can simplify the max-min portfolio problem, subject to the constraint in Equation (10), as follows.

Proposition 1. *The max-min problem (6) where the constraint (7) takes the form (10) is equivalent to the following maximization problem*

$$\max_w \left\{ w^\top (\hat{\mu} - \mu^{adj}) - \frac{\gamma}{2} w^\top \Sigma w \right\}, \tag{12}$$

where μ^{adj} is the N-vector of adjustments to be made to the estimated expected return:

$$\mu^{adj} \equiv \left\{ \text{sign}(w_1) \frac{\sigma_1}{\sqrt{T}} \sqrt{\epsilon_1}, \ldots, \text{sign}(w_N) \frac{\sigma_N}{\sqrt{T}} \sqrt{\epsilon_N} \right\}. \tag{13}$$

The proposition above shows that the multi-prior model, which is expressed in terms of a max-min optimization, can be interpreted as the mean-variance optimization problem in Equation (5) but where the mean has been adjusted to reflect the uncertainty about its estimated value. The term "sign(w_j)" in Equation (13) ensures that the adjustment leads to a "shrinkage" of the portfolio weights; that is, if a particular portfolio weight is positive (long position), then the expected return on this asset is reduced, whereas if it is negative (short position), the expected return on the asset is increased. We characterize the optimal solution for this problem in Section 1.2.3.

1.2.2 Uncertainty about expected returns estimated jointly for all assets. Instead of stating the confidence intervals for the expected returns of the assets individually as described in the previous section, one could do this jointly for all assets. Stambaugh (1997) provides motivation for why one may wish to do this. Suppose that expected returns are estimated by their sample mean $\hat{\mu}$. If returns are drawn from a Normal distribution, then the quantity

$$\frac{T(T-N)}{(T-1)N} (\hat{\mu} - \mu)^\top \Sigma^{-1} (\hat{\mu} - \mu) \tag{14}$$

has a χ^2 distribution with N degrees of freedom.[12] Let $f = \frac{T(T-N)}{(T-1)N}(\hat{\mu} - \mu)^\top \Sigma^{-1}(\hat{\mu} - \mu)$ and ϵ be a chosen quantile for the χ^2 distribution. Then the constraint (7) can be expressed as

$$\frac{T(T-N)}{(T-1)N}(\hat{\mu} - \mu)^\top \Sigma^{-1}(\hat{\mu} - \mu) \leq \epsilon. \tag{15}$$

In other words, this constraint corresponds to the probabilistic statement

$$P\left[\frac{T(T-N)}{(T-1)N}(\hat{\mu} - \mu)^\top \Sigma^{-1}(\hat{\mu} - \mu) \leq \epsilon\right] = 1 - p,$$

for some appropriate level p.

The following proposition shows how the max-min problem (6) subject to Equations (8) and (15) can be simplified into a maximization problem that is easier to solve and how one can obtain an intuitive characterization of the optimal portfolio weights.

Proposition 2. *The max-min problem (6) subject to Equations (8) and (15) is equivalent to the following maximization problem*

$$\max_w w^\top \hat{\mu} - \frac{\gamma}{2} w^\top \Sigma w - \sqrt{\varepsilon w^\top \Sigma w}, \tag{16}$$

subject to $w^\top \mathbf{1}_N = 1$, where $\varepsilon \equiv \epsilon \frac{(T-1)N}{T(T-N)}$. Moreover, the expression for the optimal portfolio weights can be written as:

$$w^* = \frac{1}{\gamma}\Sigma^{-1}\left(\frac{1}{1+\frac{\sqrt{\varepsilon}}{\gamma\sigma_P^*}}\right)\left[\hat{\mu} - \frac{B - \gamma\left(1+\frac{\sqrt{\varepsilon}}{\gamma\sigma_P^*}\right)}{A}\mathbf{1}_N\right], \tag{17}$$

where $A = \mathbf{1}_N^\top \Sigma^{-1} \mathbf{1}_N$, $B = \hat{\mu}^\top \Sigma^{-1} \mathbf{1}_N$, and σ_P^* are the variances of the optimal portfolio that can be obtained from solving the polynomial Equation (A11) in the Appendix.

We can now use the expression in Equation (17) for the optimal weights to interpret the effect of aversion to parameter uncertainty. Note that as

---

[12] If Σ is not known, then the expression in Equation (14) follows an F distribution with N and $T - N$ degrees of freedom (Johnson and Wichern, 1992, p. 188). Hence, for the empirical applications in Section 3, we will use an F distribution.

$\varepsilon \to 0$, that is either $\epsilon \to 0$ or $T \to \infty$, the optimal weight w^* converges to the mean-variance portfolio[13]

$$w^* = \frac{1}{\gamma} \Sigma^{-1} \left(\hat{\mu} - \frac{B-\gamma}{A} \mathbf{1}_N \right)$$
$$= \frac{1}{\gamma} \Sigma^{-1} \left(\hat{\mu} - \mu^0 \mathbf{1}_N \right), \tag{18}$$

where $\frac{B-\gamma}{A} = \mu^0$ is the expected return on the zero-beta portfolio associated with w^* defined in Equation (3). Thus, in the absence of parameter uncertainty, the optimal portfolio reduces to the mean-variance weights. On the contrary, as $\varepsilon \to \infty$ the optimal portfolio converges to

$$w^* = \frac{1}{A} \Sigma^{-1} \mathbf{1}_N, \tag{19}$$

which is the minimum-variance portfolio. These results suggest that, in the presence of ambiguity aversion, parameter uncertainty shifts the optimal portfolio away from the mean-variance weights toward the minimum-variance weights.

1.2.3 Uncertainty about expected returns estimated for subsets of assets. In Section 1.2.1, we described the case where there was uncertainty about expected returns that were estimated individually asset by asset, and in Section 1.2.2, we described the case where the expected returns were estimated jointly for all assets. In this section, we present a generalization that allows the estimation to be done separately for different subclasses of assets, and we show that this generalization unifies the two specifications described above.

Let $J_m = \{i_1, \ldots, i_{N_m}\}$, $m = 1, \ldots, M$, be M subsets of $\{1, \ldots, N\}$, each representing a subset of assets. Let f be an M-valued function with

$$f_m(\mu, \hat{\mu}, \Sigma) = \frac{T_m(T_m - N_m)}{(T_m - 1)N_m} (\hat{\mu}_{J_m} - \mu_{J_m})^{\top} \Sigma_{J_m}^{-1} (\hat{\mu}_{J_m} - \mu_{J_m}). \tag{20}$$

Then Equation (15) becomes

$$\frac{T_m(T_m - N_m)}{(T_m - 1)N_m} (\hat{\mu}_{J_m} - \mu_{J_m})^{\top} \Sigma_{J_m}^{-1} (\hat{\mu}_{J_m} - \mu_{J_m}) \leq \epsilon_m, \quad m = 1, \ldots, M. \tag{21}$$

Just as in the earlier specifications, these constraints correspond to the probabilistic statement

---

[13] In taking these limits, it is important to realize that σ_P^* also depends on the weights. In order to obtain the correct limits, it is useful to look at Equation (A11) that characterizes σ_P^*.

$$P(X_1 \in I_1,...,X_M \in I_M) = 1 - p,$$

where X_m, $m = 1,...,M$, are sample statistics defined by the left hand side of the inequalities in Equation (21).

The case where J_m, $m = 1,...,M$, do not overlap with each other and investors have access to a risk-free asset is of particular interest because we can obtain an analytic characterization of the portfolio weights, as shown in the following proposition.[14]

Proposition 3. *Consider the case of M nonoverlapping subsets of assets and assume f in Equation (7) is an M-valued function expressing the uncertainty of the investor for each subset of assets. Then, if the investor has access to a risk-free asset, the optimal portfolio is given by the solution to the following system of equations:*

$$w_m = \max\left[1 - \frac{\sqrt{\varepsilon_m}}{\sqrt{g(w_{-m})^\top \Sigma_m^{-1} g(w_{-m})}}, 0\right] \frac{1}{\gamma} \Sigma_m^{-1} g(w_{-m}), \qquad (22)$$

for $m = 1,...,M$, where $\varepsilon_m \equiv \epsilon_m \frac{(T_m-1)N_m}{T_m(T_m-N_m)}$, w_{-m} represents the weights in the assets not in subclass m, Σ_m is the variance-covariance matrix of the asset in subclass m, and

$$g(w_{-m}) = \hat{\mu}_m - \gamma\Sigma_{m,-m}w_{-m}, \ m = 1,...,M \qquad (23)$$

with $\Sigma_{m,-m}$ the matrix of covariances between assets in class m and assets outside class m.

1.2.4 Uncertainty about the return-generating model and expected returns. In this section, we explain how the general model developed in Section 1.2.3 where there are M subsets of assets can be used to analyze situations where investors rely on a factor model to generate estimates of expected return and are averse to ambiguity about *both* the estimated expected returns on the factor portfolios and the model generating the expected returns on investable assets.

To illustrate this situation, consider the case of a market with N risky assets in which an asset-pricing model with K factors is given. Denote

---

[14] We can characterize analytically the portfolio weights also for the case where the investor does not have access to the risk-free asset; but the characterization of the weights for this case is less transparent because this problem involves an extra constraint that the weights sum to unity. Therefore, for expositional reasons, we focus on the case where a risk-free asset is available.

with r_{at} the $N \times 1$ vector of *excess* returns of the nonbenchmark assets over the risk-free rate in period t. Similarly, denote by r_{bt} the excess return of the benchmark assets. The mean and variance of the assets and factors are

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}. \tag{24}$$

We can always summarize the mean and variance of the assets by the parameters of the following regression model

$$r_{at} = \alpha + \beta r_{bt} + u_t, \ \mathrm{cov}(u_t, u_t^\top) = \Omega, \tag{25}$$

where α is a $N \times 1$ vector, β is a $N \times K$ matrix of factor loadings, and u_t is a $N \times 1$ vectors of residuals with covariance Ω. Hence, the mean and variance of the returns can always be expressed as follows:

$$\mu = \begin{pmatrix} \alpha + \beta\mu_b \\ \mu_b \end{pmatrix}, \ \Sigma = \begin{pmatrix} \beta\Sigma_{bb}\beta^\top + \Omega & \beta\Sigma_{bb} \\ \Sigma_{bb}\beta^\top & \Sigma_{bb} \end{pmatrix}. \tag{26}$$

If the asset-pricing model is true, then $\mu_a = \beta\mu_b$, and the mispricing term, α, is zero.

Let $w \equiv (w_a, w_b)^\top$ be the $(N + K) \times 1$ vector of portfolio weights. An investor who is averse to ambiguity about both the expected returns on the factors and the model generating the returns on the assets will solve the following problem:

$$\max_w \min_{\mu_a, \mu_b} w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w, \tag{27}$$

subject to

$$(\hat{\mu}_a - \mu_a)^\top \Sigma_{aa}^{-1} (\hat{\mu}_a - \mu_a) \leq \epsilon_a, \tag{28}$$

$$(\hat{\mu}_b - \mu_b)^\top \Sigma_{bb}^{-1} (\hat{\mu}_b - \mu_b) \leq \epsilon_b. \tag{29}$$

Equations (28) and (29) capture the uncertainty about the estimates of the expected returns for the assets and the factors, respectively. If investors use the asset-pricing model to determine the estimate of $\hat{\mu}_a$, then $\hat{\mu}_a = \beta\mu_b$ and Equation (28) can be interpreted as a characterization of *model uncertainty* in the presence of ambiguity aversion. Setting

$\epsilon_a = 0$ corresponds to imposing that the investor believes dogmatically in the model.[15]

From Proposition 3, the solution to this problem is given by the following system of equations

$$w_a = \max\left[1 - \frac{\sqrt{\epsilon_a}}{\sqrt{g(w_b)^\top \Sigma_{aa}^{-1} g(w_b)}}, 0\right] \frac{1}{\gamma} \Sigma_{aa}^{-1} g(w_b), \qquad (30)$$

$$w_b = \max\left[1 - \frac{\sqrt{\epsilon_b}}{\sqrt{h(w_a)^\top \Sigma_{bb}^{-1} h(w_a)}}, 0\right] \frac{1}{\gamma} \Sigma_{bb}^{-1} h(w_a) \qquad (31)$$

where

$$g(w_b) = \hat{\mu}_a - \gamma \Sigma_{ab} w_b, \qquad (32)$$

$$h(w_a) = \hat{\mu}_b - \gamma \Sigma_{ba} w_a. \qquad (33)$$

In general, the solution to this problem will have the following properties. First, if $\epsilon_b = 0$, then for all values of $\epsilon_a > 0$, the investor is more uncertain about the assets than about the factors and hence will hold 100% of her wealth in the factor portfolios. Second, given a certain level of uncertainty about the factors (i.e., *keeping fixed* $\epsilon_b > 0$), as uncertainty about the estimates of expected returns for the assets increases, the holdings of the risky nonbenchmark assets decrease and the holdings of the factor portfolios increase. Third, given a certain level of uncertainty about the assets (i.e., *keeping fixed* ϵ_a), as ϵ_b increases, the holdings of risky nonbenchmark assets increase and the holdings of the factor portfolios decrease. These results are intuitive and suggest that the more uncertain is the estimate of the expected return of an asset, the less an investor is willing to invest in that asset. Obviously, the uncertainty in the assets and the factors are interrelated, and it is ultimately the *relative level of uncertainty* between the two classes of assets that determines the final portfolio.

We conclude this section by briefly summarizing the Bayesian approach to model uncertainty pioneered by Pástor (2000) and Pástor and Stambaugh (2000). We refer to this approach as the "Data-and-Model" (DM) approach because the decision-maker relies on both

---

[15] To be precise, the interpretation of Equation (28) as a characterization of model uncertainty is true only if $\epsilon_b = 0$. To see this, note that if $\hat{\mu}_a = \beta\mu_b$, $\hat{\mu}_a - \mu_a = \beta(\hat{\mu}_b - \mu_b) - \alpha$. Therefore, unless $\mu_b = \hat{\mu}_b$ (i.e., unless $\epsilon_b = 0$), the difference $\hat{\mu}_a - \mu_a$ does not represent Jensen's α.

the data and an asset-pricing model. A Bayesian, *ambiguity neutral* investor expresses his belief about the validity of the asset-pricing model by postulating a *prior* belief on the mispricing term α. The prior on α, conditional on Ω, is assumed to have a Normal distribution with mean zero and variance $\tau\Omega$, with τ determining the precision of the prior belief over the validity of the asset-pricing model. Moreover, the priors on the factor loadings β, the variance-covariance of the residuals Ω, as well as on the expected returns and variance-covariance matrix of the factors are assumed to be noninformative because the asset-pricing model does not impose any restrictions on these parameters. Under these assumptions, it is possible to show [see Wang (2005, Theorem 1)] that the estimators for the expected return and variance-covariance matrix that account for the belief of a Bayesian investor over the validity of a particular asset-pricing model are obtained by "shrinking" the sample moments toward the model moments. More specifically, denoting by $\hat{\mu}_a$, $\hat{\mu}_b$, and $\hat{\Sigma}_{bb}$, respectively, the sample mean of the assets' return and the sample mean and variance of the factors' returns, the estimator used for the expected return has the following structure:

$$\mu_{DM} = \omega \begin{pmatrix} \overline{\beta}\hat{\mu}_a \\ \hat{\mu}_b \end{pmatrix} + (1-\omega) \begin{pmatrix} \hat{\mu}_a \\ \hat{\mu}_b \end{pmatrix}. \tag{34}$$

In the above expression, $\overline{\beta}$ is the estimator of the factor loading β obtained by estimating the model (25) with the restriction that $\alpha = \mathbf{0}_{N\times 1}$, and ω represents the degree of confidence a Bayesian investor places in the asset-pricing model ($\omega = 1$ implies dogmatic belief in the model). Notice from Equation (34) that the parameter ω acts as a *linear* shrinkage factor for the means. Wang (2005) also shows an equivalent (quadratic) shrinkage results generating the estimator for the variance-covariance matrix Σ_{DM}. Combining these results, the portfolio of a Bayesian investor who relies on an asset-pricing model to inform his prior will be given by

$$w_{DM} = \frac{1}{\gamma}(\Sigma_{DM})^{-1}\mu_{DM}. \tag{35}$$

In Section 3, we will compare the ambiguity-averse portfolio (30)–(31) with the ambiguity-neutral Bayesian DM portfolio (35) by implementing both approaches for data on international equity indices.

2. Comparison with Other Approaches to Estimation Error

In this section, we relate the framework for portfolio choice in the presence of parameter and model uncertainty and aversion to ambiguity to other approaches considered in the literature and, in particular, to portfolios that use the traditional Bayesian approach, which is neutral to ambiguity. We compare the portfolio weights from the model with ambiguity aversion to the following: (1) the standard mean-variance portfolio that ignores estimation error, (2) the minimum-variance portfolio, (3) the portfolio based on Bayes-diffuse prior estimates as in the work of Bawa, Brown, and Klein (1979), and (4) the portfolio based on the empirical Bayes–Stein (BS) estimator, as described by Jorion (1985, 1986). In this section, the comparison is done in terms of the theoretical foundations of the models and their implications for portfolio weights, whereas in Section 3, this comparison is undertaken empirically and the comparison set includes also the weights obtained by using the Data-and-Model approach of Pástor (2000).

2.1 A summary of the traditional Bayesian approach

It is useful to begin with a brief summary of the traditional Bayesian approach. Let $U(R)$ be the utility function, where R is the return from the investment and $g(R|\Theta)$ the conditional density (likelihood) of asset returns given the parameter Θ. In the setting of this article, Θ is the vector of the expected returns of the risky assets. More generally, it can include the covariances of the asset returns as well as other relevant parameters. If the parameter Θ is known, then the conditional expected utility of the investor is

$$E[U(R)|\Theta] = \int U(R)g(R|\Theta)dR. \tag{36}$$

In practice, however, the parameter Θ is often unknown and needs to be estimated from data; that is, there is parameter uncertainty. In the presence of such parameter uncertainty, Savage's expected utility approach assumes that the investor has a prior over the unknown parameter Θ and derives a conditional prior (or *posterior*) $p(\Theta|X)$, after observing the history $X = (r_1,\ldots,r_T)$ of realized returns. The expected utility is hence given by

$$
\begin{aligned}
E[U(R)|X] &= \int \int U(R)g(R|\Theta)p(\Theta|X)dRd\Theta \\
&= \int U(R)\left(\int g(R|\Theta)p(\Theta|X)d\Theta \right)dR,
\end{aligned}
\tag{37}
$$

where the inner integral in Equation (37) represents the *predictive density*, $g(R|X)$:

$$g(R|X) = \int g(R|\Theta)p(\Theta|X)d\Theta. \qquad (38)$$

Thus, the key to the Bayesian approach is the incorporation of prior information and the information from data in the calculation of the posterior and predictive distributions. The effect of information on the investor's decision comes through its effect on the predictive distribution.[16]

The foundation for the Bayesian approach was provided by Savage (1954). In the portfolio choice literature, the Bayesian approach has been implemented in different ways. Barry (1974), Klein and Bawa (1976), and Bawa, Brown, and Klein (1979) use either a noninformative diffuse prior or a predictive distribution obtained by integrating over the unknown parameter. In a second implementation, Jobson and Korkie (1980), Jorion (1985, 1986), Frost and Savarino (1986), and Dumas and Jacquillat (1990) use empirical Bayes estimators. In a third implementation, Pástor (2000) and Pástor and Stambaugh (2000) use the equilibrium implications of an asset-pricing model to establish a prior and consider, in addition to parameter uncertainty, also model uncertainty.

2.2 Comparison of the Ambiguity-Averse approach with the Bayesian approach

The decision-theoretic foundation of the model with ambiguity aversion is laid by Gilboa and Schmeidler (1989). Equally well-founded axiomatically as the Bayesian approach, the most important difference is that in the Bayesian approach the investor is implicitly assumed to be neutral rather than averse to parameter and/or model uncertainty.

That in the Bayesian approach the investor is ambiguity neutral is best seen through Equation (37). The last expression in the equation suggests that parameter and/or model uncertainty enters the investor's utility through the posterior $p(\Theta|X)$, which can affect the investor's utility only through its effect on the predictive density $g(R|X)$. In other words, as far as the investor's utility maximization decision is concerned, it does not matter whether the overall uncertainty comes from the conditional distribution $g(R|\Theta)$ of the asset return or from the uncertainty about the parameter/model, $p(\Theta|X)$, as long as the predictive distribution $g(R|X)$ is the same. That is, if the investor were in a situation where there is no parameter/model uncertainty, say, because the past data X could be used to identify the true parameter perfectly,

---

[16] The use of the predictive distributions was pioneered by Zellner and Chetty (1965).

and the distribution of asset returns were characterized by $g(R|X)$, then the investor would feel no different. In particular, there is no meaningful separation of risk aversion and ambiguity aversion. In this sense, we say that the investor is ambiguity *neutral*.

In the framework with ambiguity aversion, the *risk* (the conditional distribution $g(R|\Theta)$ of the asset returns) is treated differently from the *uncertainty* about the parameter/model of the data-generating process. For example, in the portfolio choice problem described by Equations (6)–(8), the risk of the asset returns is captured by Σ, which appears in Equation (6). The uncertainty about the unknown mean return vector, μ, is captured by the constraint (7). The two are further separated by the minimization over μ, subject to the constraint in Equation (7), which reflects that the investor is averse, rather than neutral, to ambiguity.

2.3 Analytic comparison of the portfolio weights from the various models

In this section, we compare analytically the portfolio weights from the model with ambiguity aversion with those obtained if using traditional Bayesian methods to deal with estimation error. We start by describing the portfolio obtained by using the empirical Bayes-Stein estimator. The Bayes-diffuse prior portfolio is then obtained as a special case of this portfolio, whereas the mean-variance portfolio and the minimum-variance portfolio are discussed as limit cases of the traditional Bayesian models and also the model with ambiguity aversion.

The intuition underlying the Bayes-Stein approach to asset allocation is to minimize the impact of estimation error by "shrinking" the sample mean toward a *common* value or, as it is usually called, a *grand mean*.[17] In our implementation of the Bayes-Stein approach, we take the grand mean $\overline{\mu}$ to be the mean of the minimum-variance portfolio, μ_{MIN}. More specifically, following Jorion (1986), we use the following shrinkage estimator for the expected return and covariance matrix:

$$\mu_{BS} = (1 - \phi_{BS})\hat{\mu} + \phi_{BS}\,\mu_{MIN}\,\mathbf{1}_N, \qquad (39)$$

$$\Sigma_{BS} = \Sigma\left(1 + \frac{1}{T + \nu_\mu}\right) + \frac{\nu_\mu}{T(T + 1 + \nu_\mu)}\frac{\mathbf{1}_N \mathbf{1}_N^\top}{\mathbf{1}_N^\top \Sigma^{-1}\mathbf{1}_N}, \qquad (40)$$

---

[17] Stein (1955) and Berger (1974) developed the idea of shrinking the sample mean toward a common value and showed that these kind of estimators achieve uniformly lower risk than the MLE estimator (here risk is defined as the expected loss, over repeated samples, incurred by using an estimator instead of the true parameter). The results from Stein and Berger can be interpreted in a Bayesian sense where the decision-maker assumes an *informative prior* over the unknown expected returns. This is what defines a Bayes-Stein estimator. An *empirical Bayes* estimator is a Bayes estimator where the grand mean and the precision are inferred from the data.

where $\hat{\mu}$ is the sample mean, μ_{MIN} is the mean of the minimum-variance portfolio, ϕ_{BS} is the shrinkage factor for the mean, and ν_{μ} is the precision (or "tightness") of the prior on μ:

$$\phi_{BS} = \frac{\nu_{\mu}}{T + \nu_{\mu}}, \tag{41}$$

$$\nu_{\mu} = \frac{N + 2}{(\hat{\mu} - \mu_{MIN})^{\top} \Sigma^{-1} (\hat{\mu} - \mu_{MIN})}. \tag{42}$$

Note that the case of zero precision ($\nu_{\mu} = 0$), corresponds to the Bayes-diffuse prior case considered by Bawa, Brown, and Klein (1979), in whose work the sample mean is the predictive mean but the covariance matrix is inflated by the factor $(1 + 1/T)$. Finally, observe that for $\nu_{\mu} \to \infty$, the predictive mean is the common mean represented by the mean of the minimum-variance portfolio.

We are now ready to determine the optimal portfolio weights using the Bayes-Stein estimators. Let us assume that we know the variance-covariance matrix and that only the expected returns are unknown. In the case where a risk-free asset is not available, we know that the classical mean-variance portfolio is given by Equation (3). Substituting the empirical Bayes-Stein estimator μ_{BS} in Equation (3), one can show that the optimal weights can be written as

$$w_{BS} = \phi_{BS}\, w_{MIN} + (1 - \phi_{BS})w_{MV}, \tag{43}$$

where the minimum-variance portfolio weights, which ignore expected returns altogether, are

$$w_{MIN} = \frac{1}{A} \Sigma^{-1} \mathbf{1}_N, \tag{44}$$

and the mean-variance portfolio weights formed using the maximum-likelihood estimates of the expected return are

$$w_{MV} = \frac{1}{\gamma} \Sigma^{-1} (\hat{\mu} - \hat{\mu}_0 \mathbf{1}_N). \tag{45}$$

We now compare the portfolio obtained from the Bayes-Stein approach, in Equation (43), with the optimal portfolio derived from the multi-prior ambiguity-averse approach that incorporates aversion to

parameter uncertainty and where the estimation of expected returns is done jointly for all assets—that is, the expression given in Equation (17). After some manipulation, the optimal portfolio for an investor who is averse to parameter uncertainty can be written as

$$w_{AA}(\epsilon) = \phi_{AA}(\epsilon)w_{MIN} + [1 - \phi_{AA}(\epsilon)]w_{MV}, \qquad (46)$$

where

$$\phi_{AA}(\epsilon) = \frac{\sqrt{\epsilon\frac{(T-1)N}{T(T-N)}}}{\gamma\sigma_P^* + \sqrt{\epsilon\frac{(T-1)N}{T(T-N)}}}, \qquad (47)$$

and w_{MIN} and w_{MV} are defined in (44) and (45), respectively.

Comparing the weights in Equation (43) that are obtained using a Bayes-Stein estimator to the weights in Equation (46) obtained from the model with ambiguity aversion, we notice that both methods shrink the mean-variance portfolio toward the minimum-variance portfolio, which is the portfolio that essentially ignores all information about expected returns. However, the magnitude of the shrinkage is different, that is, $\phi_{BS} \neq \phi_{AA}(\epsilon)$. In the next section, where we implement these different portfolio strategies using real-world data, we will find that the shrinkage factor from the ambiguity-averse approach is much greater than that for the empirical Bayes-Stein portfolio; that is, for reasonable values of $\epsilon, \phi_{AA}(\epsilon) > \phi_{BS}$. Note that, as mentioned earlier, the parameter ϵ represents the *product* of the agent's degree of ambiguity-aversion and the level of ambiguity. If there is no aversion to ambiguity and also no ambiguity about returns, then the optimal portfolio of the investor is indeed the classical mean-variance portfolio. However, if the agent is neutral to ambiguity but there is ambiguity about returns, then the investor deals with the ambiguity by using Bayesian shrinkage estimators for the asset returns, where these Bayesian estimators can be based on either just the data, as in the works of Jorion (1985, 1986), or both the data and an asset-pricing model, as in the works of Pástor (2000) and Pástor and Stambaugh (2000), so that the resulting portfolios are the ones that a Bayesian investor would choose.

So far, we have considered the following two cases: (i) where the investor uses classical maximum-likelihood estimators to estimate expected returns and then accounts for ambiguity aversion in obtaining the weights in Equation (46) and (ii) where the investor takes into account that the estimates are uncertain but does not allow for ambiguity aversion, which leads to the portfolio weights in Equation (43). But, one could just as well have a third case that takes estimation uncertainty into account *and* allows

for ambiguity aversion.[18] In this case, the optimal portfolio weights are given by the following expression:

$$w_{AA}^{BS}(\epsilon) = \phi_{AA}(\epsilon)w_{MIN} + [1 - \phi_{AA}(\epsilon)]w_{BS}, \qquad (48)$$

where the minimum-variance portfolio, w_{MIN}, and the Bayesian portfolio, w_{BS}, are defined in Equations (44) and (43), respectively. Observe that the expression in Equation (48) is similar to that in Equation (46) but where w_{BS} replaces w_{MV}; that is, the effect of uncertainty is to shrink the portfolio that is now obtained using Bayesian estimation methods, w_{BS}, toward w_{MIN}. Notice that in the limiting case where the investor is *neutral* toward uncertainty, setting $\epsilon = 0$ in Equation (48), which leads to $\phi_{AA}(\epsilon) = 0$, implies that the optimal portfolio reduces to w_{BS}. This is the sense in which the portfolio obtained using Bayesian estimation methods is nested in the ambiguity-averse model.

We conclude this section by noting the similarity between Equations (43) and (46), which suggests that the Bayes-Stein approach by Jorion (1985, 1986) and the ambiguity-averse approach (with a single joint constraint, as described in Section 1.2.2) are observationally equivalent. That is, the Bayes-Stein approach is a special case of the approach with ambiguity aversion. However, this observational equivalence is not true for the general case discussed in Section 1.2.3, where the estimation is done separately for different subsets of assets, a special case of which is discussed in Section 1.2.1, where the expected return on each asset is estimated individually. Once the constraint in Equation (7) is imposed separately for different subsets of assets, the "shrinkage factor" will differ across these subsets. Consequently, it will no longer be possible to express the optimal portfolio as a weighted average of w_{MIN} and w_{MV}; in fact, it is possible that the optimal portfolio will have zero investment in some of the subsets of assets.

3. Empirical Applications of the Ambiguity-Aversion Approach

Our goal in this section is (i) to demonstrate how the multi-prior approach with ambiguity aversion can be implemented in practice and (ii) to compare the portfolio recommendations from this approach with other procedures commonly used, with a particular emphasis on the Bayes-Stein approach by Jorion (1985, 1986) and the Bayesian Data-and-Model approach by Pástor (2000).

The experiment considered in this section of the article is similar in spirit to the one by Jorion (1985, 1986). In these articles, Jorion considers a

---

[18] In this comparison, the Bayesian approach is interpreted narrowly as an estimation technique rather than a decision-theoretic approach.

mean-variance investor and evaluates alternative estimators of expected returns, based on Bayesian methods. Ex ante (in sample) it is, of course, clear that each investor would attain the highest utility by adopting the portfolio policy obtained from maximizing her own objective function. However, the way such estimators are assessed by Jorion is by looking at their ex post or *out-of-sample* performance, typically measured by their Sharpe ratio. Similarly, in our empirical exercise we evaluate whether *even* the mean-variance investor could attain a higher out-of-sample utility (or equivalently, Sharpe ratio) by using the portfolio policy that incorporates ambiguity aversion. That is, we examine if the portfolio policies that incorporate ambiguity aversion are sufficiently *robust* to estimation uncertainty, so that out-of-sample they deliver a higher Sharpe ratio than that provided by the mean-variance policies, which ignore estimation error, or the Bayesian policies, which account for estimation error but are ambiguity neutral.

We consider two applications of our model. First, in Section 3.1 we illustrate the case with only parameter uncertainty when the expected returns on all assets are estimated jointly; this is the model that is described in Section 1.2.2. Then, in Section 3.2 we illustrate the case with both uncertainty about expected returns and the factor model generating these returns; this is the model described in Section 1.2.4.

For both applications, we use returns on eight international equity indices. These data are similar to the data considered by Jorion (1985) and De Santis and Gerard (1997) but span for a longer time period. The equity indices are for Canada, France, Germany, Italy, Japan, Switzerland, the United Kingdom, and the United States. For the second application, we consider also the return on the world market portfolio. Data are from Morgan Stanley Capital International (MSCI). The returns on each index are computed based on the month-end US dollar value of the index. Excess returns are obtained by subtracting the month-end return for the US 30–day T-bill as reported in the Center for Research in Security Prices (CRSP) data files. The data span from January 1970 to July 2001 (379 observations). The portfolio weights for each strategy are determined each month using moments estimated from a rolling-window of 120 months, and these portfolio weights are then used to calculate the returns in the 121st month. The resulting out-of-sample period spans from January 1980 to July 2001 (259 observations). The investor is assumed to have a risk aversion of $\gamma = 1$.

3.1 Parameter uncertainty only

In this section, we consider the case where the investor is uncertain about expected returns. We assume that there is no risk-free asset and that the investor estimates the expected returns jointly by expressing uncertainty over the entire set of assets, as described in Section 1.2.2. Using this model, we compute the portfolios that account for different degrees of uncertainty in the statistical estimate (ϵ) about the expected returns. We also compute

(i) the standard mean-variance portfolio that ignores estimation error, (ii) the minimum-variance portfolio, and (iii) the portfolio based on Bayes-Stein estimators, as described by Jorion (1985, 1986), which is a combination of the minimum-variance portfolio and the mean-variance portfolio. For each of the portfolio models, we consider two cases: one where short selling is allowed and the other where short selling is not allowed.

In our analysis, we set $T = 120$ because the estimation is done using a rolling-window of 120 months and we set $N = 8$ because there are eight country indexes. Under the assumption that the returns are Normally distributed, if $\hat{\mu}$ is taken to be the sample average of the returns and $\hat{\Sigma}$ is the sample variance-covariance matrix, then the quantity

$$\frac{T(T-N)}{(T-1)N} (\hat{\mu} - \mu)^{\top} \hat{\Sigma}^{-1} (\hat{\mu} - \mu) \tag{49}$$

has an F-distribution with 8 and 112 degrees of freedom ($F_{8,112}$).

The results of our analysis are reported in Panel A of Table 1.[19] Notice from Panel A that the case of $\epsilon = 0$ corresponds to the mean-variance portfolio, whereas the case of $\epsilon \to \infty$ corresponds to the minimum-variance portfolio, as discussed in the previous section. From the table, we see that compared with the mean-variance strategy in which historical mean returns $\hat{\mu}$ are taken to be the estimator of expected returns μ, the portfolios constructed using the model with ambiguity aversion that allows for parameter uncertainty exhibit uniformly higher means, lower volatility, and, consequently, substantially higher Sharpe ratios. The same is true for the comparison with the empirical Bayes-Stein portfolio, which also has a lower mean, higher variance, and lower Sharpe ratio than any of the portfolios that account for ambiguity aversion.[20]

To understand the relatively poor performance of the empirical Bayes-Stein portfolio, recall from Equations (43) and (46) that the optimal portfolio of the investor can also be interpreted as one that is a weighted-average of the standard mean-variance portfolio and the minimum-variance portfolio, with the weight on the minimum-variance portfolio increasing as ambiguity aversion increases. The Bayes-Stein model performs poorly because it puts too much weight on the estimated expected returns and, consequently, does not shrink the portfolio weights sufficiently toward the minimum-variance

---

[19] The number in parenthesis appearing in the table refer to the percentage-confidence interval implied by different value of ϵ and computed from an $F_{8,112}$ distribution.

[20] We do not report the performance of the Bayes-diffuse-prior portfolio because it is virtually indistinguishable from the mean-variance case. To understand the reason for this, observe that for the case of the Bayesian-diffuse prior portfolio, parameter uncertainty is incorporated by inflating the variance-covariance matrix by the factor $1 + \frac{1}{T}$ [see Bawa, Brown, and Klein (1979)] although still using the historical mean as a predictor of expected returns. For large enough T (120 in our case), this correction to the variance-covariance matrix has only a negligible effect on performance.

Table 1
Out-of-sample performance with only parameter uncertainty

Strategy	Mean	SD	Mean/SD
Panel A: Short sales allowed			
Mean-variance	0.0049	0.2557	0.0190
Minimum-variance	0.0118	0.0419	0.2827
Bayes–Stein	0.0071	0.1058	0.0671
Ambiguity-averse			
$\epsilon=0.00$ (0.00%)	0.0049	0.2557	0.0190
$\epsilon=0.25$ (2.01%)	0.0065	0.1307	0.0495
$\epsilon=0.50$ (14.60%)	0.0080	0.0953	0.0841
$\epsilon=0.75$ (35.27%)	0.0091	0.0765	0.1190
$\epsilon=1.00$ (55.98%)	0.0098	0.0654	0.1491
$\epsilon=1.50$ (83.49%)	0.0104	0.0545	0.1909
$\epsilon=2.00$ (94.73%)	0.0107	0.0501	0.2144
$\epsilon=2.50$ (98.45%)	0.0109	0.0480	0.2281
$\epsilon=3.00$ (99.56%)	0.0111	0.0467	0.2369
$\epsilon\rightarrow\infty$ (100.00%)	0.0118	0.0419	0.2827
Panel B: Short sales not allowed			
Mean-variance	0.0104	0.0587	0.1774
Minimum-variance	0.0117	0.0412	0.2831
Bayes–Stein	0.0106	0.0511	0.2074
Ambiguity-averse			
$\epsilon=0.00$ (0.00%)	0.0104	0.0587	0.1774
$\epsilon=0.25$ (2.01%)	0.0113	0.0511	0.2214
$\epsilon=0.50$ (14.60%)	0.0115	0.0482	0.2391
$\epsilon=0.75$ (35.27%)	0.0115	0.0467	0.2472
$\epsilon=1.00$ (55.98%)	0.0116	0.0457	0.2533
$\epsilon=1.50$ (83.49%)	0.0116	0.0446	0.2607
$\epsilon=2.00$ (94.73%)	0.0117	0.0440	0.2647
$\epsilon=2.50$ (98.45%)	0.0117	0.0436	0.2671
$\epsilon=3.00$ (99.56%)	0.0117	0.0433	0.2695
$\epsilon\rightarrow\infty$ (100.00%)	0.0117	0.0412	0.2831

This table reports the out-of-sample mean, Standard Deviation (SD), and Mean-to-SD ratio for the returns on different portfolio strategies, including the one from the ambiguity-averse model that allows for parameter uncertainty. Means and SDs are expressed as percentage per month. The data are obtained from Morgan Stanley Capital International (MSCI) and consist of monthly returns on eight international equity indices (Canada, France, Germany, Italy, Japan, Switzerland, the United Kingdom, and the United States) from January 1970 to July 2001 (379 observations). The portfolio weights for each strategy are determined each month using moments estimated from a rolling-window of 120 months, and these portfolio weights are then used to calculate the returns in the 121st month. The resulting out-of-sample period spans from January 1980 to July 2001 (259 observations). In parenthesis, we report the percentage size of the confidence interval for a $F_{8,112}$ implied by the values of ϵ. The investor is assumed to have a risk aversion of $\gamma = 1$.

portfolio relative to the portfolio that incorporates aversion to ambiguity. The weighting factor assigned by the empirical Bayes-Stein model to the minimum-variance portfolio over the out-of-sample period averages to 0.6930, whereas this factor for the model with aversion to parameter uncertainty is 0.8302 if $\epsilon = 1$ and, as mentioned in Proposition 2, this factor increases with ϵ. In Figure 1, we report the evolution of the shrinkage factor over time for the ambiguity-averse and Bayes-Stein portfolios. The top line in the figure, starting at around 0.9, represents the shrinkage weight $\phi_{AA}(3)$ assigned by the investor to the minimum-variance portfolio if $\epsilon = 3$, corresponding approximately to a 95% confidence around the estimate expected return. The case for $\epsilon = 1$ (corresponding approximately to a 56%
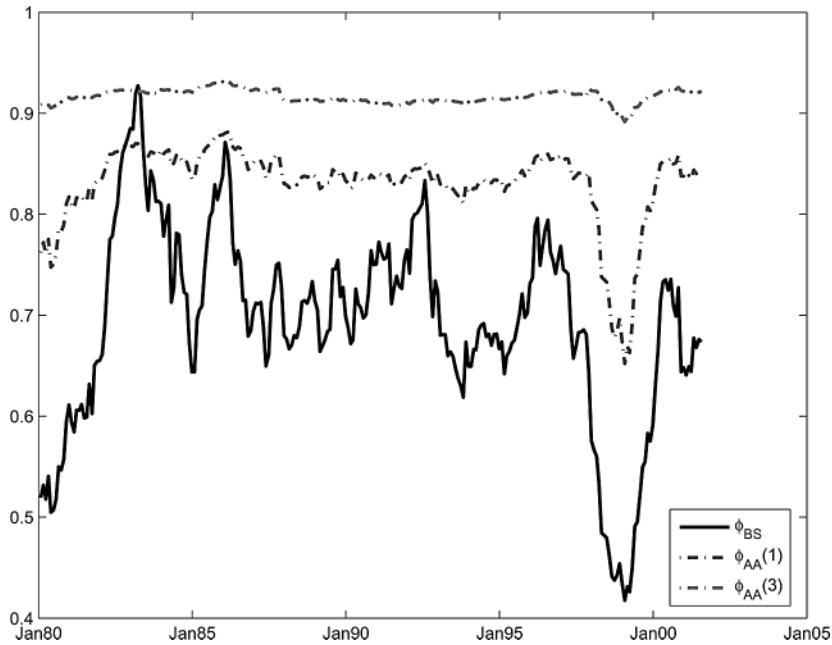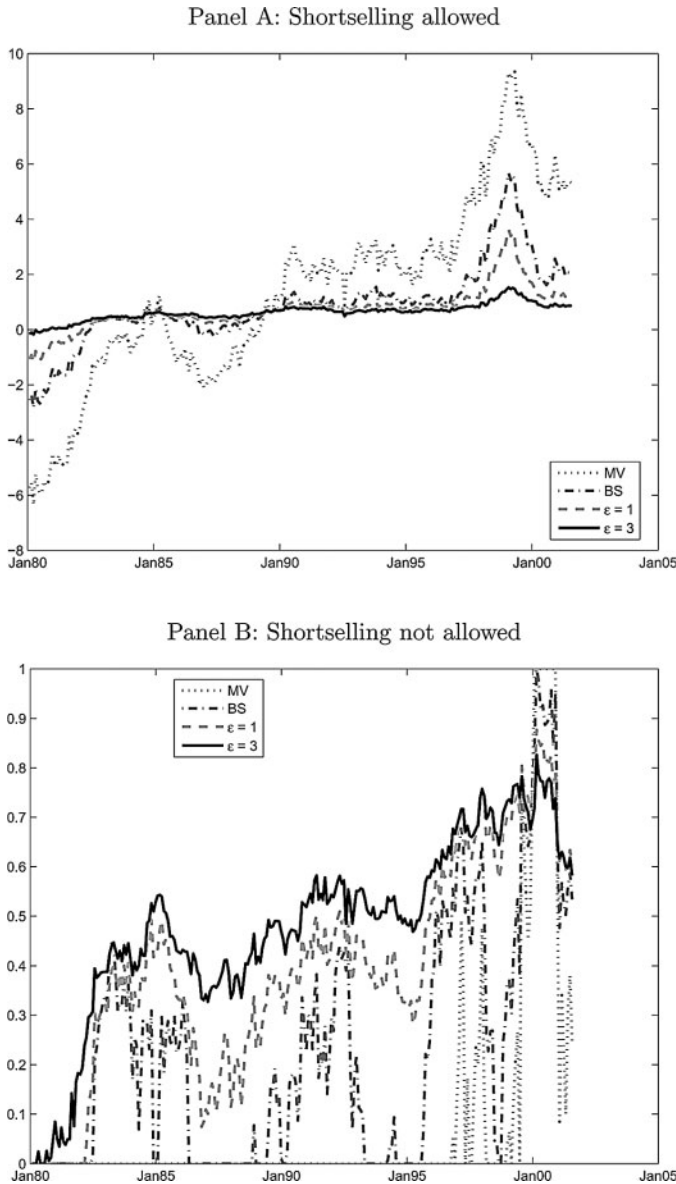
Figure 1
Shrinkage factors $\phi_{AA}(\epsilon)$ and ϕ_{BS} over time
The figure reports the weight put on the minimum-variance portfolio by a ambiguity-averse investor and by an investor following the Bayes-Stein (BS) shrinkage approach, as explained in Section 2.3. The plot $\phi_{AA}(3)$ (the very top line in the figure starting at 0.9) gives the weight on the minimum-variance portfolio if $\epsilon = 3$, and the plot $\phi_{AA}(1)$ (the middle line in the figure) gives the weight on the minimum-variance portfolio if $\epsilon = 1$. These quantities are defined in Equation (46). The solid line, ϕ_{BS} gives the weight on the minimum-variance portfolio suggested by the Bayes-Stein approach [see Equation (41)]. Details of the data are contained in the description of Table 1.

confidence interval) is represented by the middle line in the figure, labeled $\phi_{AA}(1)$. The solid line starting just about 0.5 represents the shrinkage factor for the Bayes-Stein approach. From the figure, we see that the shrinkage toward the minimum-variance portfolio increases with ϵ; moreover, the shrinkage factor fluctuates much less for higher ϵ.

To analyze the effect of ambiguity aversion on the individual weights in the risky portfolio, we report in Panel A of Figure 2 the percentage weight allocated to the US index from January 1980 to July 2001 for four different portfolio strategies. The dotted line (MV) refers to the percentage of wealth allocated to the US index under the mean-variance portfolio strategy, which is implemented using historical estimates of the moments of asset returns. The dash-dotted line refers to the Bayes-Stein portfolio. The other two lines refer to portfolios obtained by incorporating aversion to parameter uncertainty. Two levels of uncertainty are considered. The dashed line gives the weight from the ambiguity averse model ($\epsilon = 1$)

Panel A: Shortselling allowed

Panel B: Shortselling not allowed

Figure 2
Portfolio weight in the US index over time
This figure reports the portfolio weight in the US index from January 1980 to July 2001. Panel A shows the case where short selling is allowed and Panel B gives the case where short selling is not allowed. The dotted line (MV) refers to the mean-variance portfolio. The dash-dotted line refers to the Bayes-Stein (BS) portfolio. The dashed line gives the weight from the ambiguity-averse model ($\epsilon = 1$) that corresponds to uncertainty expressed roughly by the 56% confidence interval for an $F_{8,112}$ centered around the sample mean, whereas the solid line ($\epsilon = 3$) plots the weight for the case where uncertainty about expected returns is given by a 99% confidence interval. Details of the data are contained in the description of Table 1.

67

that corresponds to a degree of uncertainty expressed roughly by the 56% confidence interval for an $F_{8,112}$ centered around the sample mean, whereas the solid line ($\epsilon = 3$) plots the weight for the case where uncertainty about expected returns is given by a 99% confidence interval. We find that portfolio weights from the optimization incorporating parameter uncertainty have less extreme positions and the portfolio weights vary much less over time compared with the weights for the classical mean-variance portfolio and the Bayes-Stein portfolio. In particular, the figure shows that the position in the US asset is less extreme for the Bayes-Stein portfolio than it is for the mean-variance portfolio, but the ambiguity-averse portfolios for $\epsilon = 1$ and $\epsilon = 3$ are even more conservative than the Bayes-Stein portfolio. A larger ϵ means a higher confidence interval and, consequently, more uncertainty in the estimates. As a consequence, the larger is ϵ, the less extreme are the portfolio weights.

In the results described above, investors were permitted to hold short positions. We now repeat the analysis but prohibit short sales. Formally, the problem we now solve is the same as the one in Section 1.2.2 but with the additional constraint that short sales are not allowed: $w \geq \mathbf{0}_N$. The results of this analysis are reported in Panel B of Table 1. As in Panel A, this panel compares the out-of-sample mean return, volatility, and Sharpe ratio obtained from the model with ambiguity aversion to alternative portfolio strategies. Again, we find that the portfolio strategies that incorporate aversion to parameter uncertainty achieve a higher mean and lower volatility than the mean-variance portfolio and the Bayes-diffuse-prior portfolio. Just as before, the relatively poor performance of the empirical Bayes-Stein portfolio is due to the relatively low weight this approach assigns to the minimum-variance portfolio, as discussed above.[21]

It is well known [Frost and Savarino (1988), Jagannathan and Ma (2003)] that imposing a short selling constraint improves the performance of the mean-variance portfolio. This result can be confirmed by comparing Panel B of Table 1 with Panel A. Both the mean-variance portfolio and the Bayesian portfolios show a higher Sharpe ratio in the case in which short selling is not allowed. It is also interesting to note that the out-of-sample performance of the portfolio constructed by incorporating aversion to parameter uncertainty is less sensitive to the introduction of a short sale constraint. For these portfolios, the difference in Sharpe ratios between Panels A and B is much less dramatic than for the case of the mean-variance portfolio or the Bayesian portfolios.

---

[21] Note however that in Panels A and B of Table 1 the portfolio with the highest mean and lowest volatility is the minimum-variance portfolio or, equivalently, the portfolio for a very high level uncertainty, $\epsilon = \infty$. The reason for this is that in the particular data that we are using, returns are so noisy that expected returns are estimated very imprecisely, and hence, one is best off ignoring them all together. However, simulations reveal that if data is less noisy, then it is no longer optimal to hold only the minimum-variance portfolio.

This is because the effect of parameter uncertainty, as we saw previously for the case in which short sales were allowed, is to reduce extreme positions, producing a similar effect on the portfolio as a constraint on short selling. This intuition is confirmed by noting, for example, that for $\epsilon \geq 1.5$ (83.49-percentile of an $F_{8,112}$) the Sharpe ratios in Panel A of Table 1 for the portfolios that account for aversion to parameter uncertainty are larger than the Sharpe ratio for the constrained mean-variance portfolio in Panel B (0.1774). Although the effect of incorporating parameter uncertainty is similar to the effect of constraining short sales, there is one important difference: the "constraints" imposed by incorporating parameter uncertainty are endogenous rather than exogenous, and consequently, if it is optimal to have short positions in some assets, these are not ruled out a priori.

We report also the portfolio weights over time if short sales are prohibited, just as we did for the case without shortsale constraints. In Panel B of Figure 2, we report the percentage weight allocated to the US index from January 1980 to July 2001. As in Panel A of Figure 2, the dotted line (MV) refers to the mean-variance portfolio weight, the dash-dotted line refers to the Bayes-Stein portfolio, and the other two lines refer to weights obtained by incorporating aversion to ambiguity with $\epsilon = 1$ (dashed line, low uncertainty) and $\epsilon = 3$ (solid line, higher uncertainty). Note how the introduction of parameter uncertainty reduces the "bang-bang" nature displayed by the mean-variance and Bayes-Stein portfolio weights, and thus, incorporating parameter uncertainty reduces turnover along with improving out-of-sample performance.

3.2 Parameter and model uncertainty
In this section, we implement the model discussed in Section 1.2.4, where assets are assumed to follow a factor structure and investors are uncertain about the validity of the return-generating model in addition to being uncertain about expected returns. The return-generating model is assumed to be a single-factor CAPM, where the factor is the world market portfolio. In order for the CAPM to be valid, in this application we allow for the existence of a risk-free asset. However, for evaluating the performance of the different portfolio strategies, we consider the portfolio of only risky assets.

We consider two possible cases of model uncertainty: in the first, the investor does not believe in the model (that is, $\omega = 0$ in terms of the notation introduced in Section 1.2.4) and estimates the expected return on the asset, $\hat{\mu}_a$, using its sample average but allowing for uncertainty around this as indicated by the parameter ϵ_a. In the second case, the investor forms his expectation about $\hat{\mu}_a$ by believing completely in a single-factor model ($\omega = 1$) and therefore, in each month, forms estimates of the expected returns on the assets by using $\hat{\mu}_a = \overline{\beta}\hat{\mu}_b$, where $\overline{\beta}$ is

the 8×1 vector of betas obtained by estimating the model (25) with the restriction that $\alpha = \mathbf{0}_{8 \times 1}$. In this case too, the investor allows for uncertainty about the estimate from the model, captured by the parameter ϵ_a. In both cases ($\omega = 0$ and $\omega = 1$), the expected return on the benchmark, $\hat{\mu}_b$ is the sample average. However, the investor also allows for uncertainty about this estimate, as reflected by the parameter ϵ_b. When $\epsilon_a = 0$ and $\epsilon_b = 0$, the investor's portfolio will be the mean-variance portfolio if the reference estimator for $\hat{\mu}_a$ is the sample mean ($\omega = 0$), whereas it will be the market portfolio if the reference estimator for $\hat{\mu}_a$ is obtained through the CAPM ($\omega = 1$).

In Table 2, we report the Sharpe ratios of various portfolio strategies if one can invest in the world market (benchmark) portfolio in addition to the eight country indices. The strategies considered are the mean-variance portfolio, the Bayes-Stein portfolio, the single-prior Bayesian Data-and-Model approach by Pástor (2000), and the portfolio that now allows for aversion to both parameter and model uncertainty.

From Table 2, we see that the out-of-sample Sharpe ratio for the mean-variance strategy is -0.0719. This improves slightly for the Bayes-Stein strategy to -0.0528. For the Bayesian Data-and-Model approach, if $\omega = 0$, that is, the agent does not believe in the single-factor model, the Sharpe ratio is -0.0853; as explained by Wang (2005), the only reason that this is not the same as that for the mean-variance portfolio is because of an adjustment for the degrees of freedom in computing the variance of the portfolio. For the case where $\omega = 1$, that is, the agent believes completely in the model, the Sharpe ratio is 0.1239, which is the Sharpe ratio for holding the world market portfolio.

For the portfolios incorporating ambiguity aversion, we first note from Table 2 that if we set $\epsilon_a = \epsilon_b = 0$ implying that there is neither parameter nor model uncertainty, then we get the same Sharpe ratios for $\omega = 0$ and $\omega = 1$ as the corresponding values under the Data-and-Model approach.

Now, if we set $\epsilon_b = 0$ but let $\epsilon_a > 0$ implying that the investor is confident about expected returns on the benchmark but not about expected returns on the individual country indices, then for the case where the investor dogmatically believes in the model ($\omega = 1$ given in the lower panel in the table), the investor holds only the world market portfolio, which has a Sharpe ratio of 0.1239. And, for the case where the investor does not believe in the model at all ($\omega = 0$, upper panel) for $\epsilon_a \geq 1.50$, the Sharpe ratio is again 0.1239 implying that the investor holds only the benchmark portfolio (the world market) and does not invest at all in any of the individual country indices. The reason for this is that the uncertainty about the expected returns on individual country indices is sufficiently large, relative to zero uncertainty about the expected returns on the benchmark portfolio, that the investor finds it optimal to hold just the benchmark portfolio.

Table 2
Out-of-sample performance with parameter and model uncertainties

Strategy			Sharpe ratio						
Mean-variance						−0.0719			
Bayes–Stein						−0.0528			
Bayesian Data-and-Model									
With $\omega = 0$						−0.0853			
With $\omega = 1$						0.1239			
Ambiguity-averse						ϵ_b			
			0.00	0.50	1.00	1.50	2.00	2.50	3.00
			% (0.00)	(38.20)	(68.07)	(86.37)	(95.22)	(98.62)	(99.67)
	ϵ_a	%							
With $\omega = 0$	0.00	(0.00)	−0.0853	0.1127	0.1127	0.1127	0.1127	0.1127	0.1127
	0.25	(2.01)	−0.0774	0.1032	0.1043	0.1043	0.1043	0.1043	0.1043
	0.50	(14.60)	−0.0475	0.0824	0.0815	0.0824	0.0839	0.0842	0.0842
	0.75	(35.27)	−0.0113	0.0604	0.0568	0.0524	0.0477	0.0489	0.0481
	1.00	(55.98)	0.0930	0.0655	0.0601	0.0566	0.0562	0.0556	0.0447
	1.50	(83.49)	0.1219	0.1218	0.0814	0.0503	0.0391	0.0319	−0.0435
	2.00	(94.73)	0.1223	0.1252	0.0874	0.0643	0.0487	0.0341	−0.0492
	2.50	(98.45)	0.1239	0.1275	0.0909	0.0465	0.0126	0.0091	−0.0870
	3.00	(99.56)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0364	−0.0647
With $\omega = 1$	0.00	(0.00)	0.1239	0.1202	0.1202	0.1202	0.1202	0.1202	0.1202
	0.25	(2.01)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0383	0.0383
	0.50	(14.60)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0364	−0.0647
	0.75	(35.27)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0364	−0.0647
	1.00	(55.98)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0364	−0.0647
	1.50	(83.49)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0364	−0.0647
	2.00	(94.73)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0364	−0.0647
	2.50	(98.45)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0364	−0.0647
	3.00	(99.56)	0.1239	0.1284	0.0943	0.0551	0.0197	0.0364	−0.0647

This table reports the out-of-sample Sharpe ratios for the returns on different portfolio strategies, including the one from the ambiguity-averse model that allows for both parameter and model uncertainty. Sharpe ratios are expressed as percentage per month. The data are obtained from Morgan Stanley Capital International (MSCI) and consist of monthly excess returns on eight international equity indices (Canada, France, Germany, Italy, Japan, Switzerland, the United Kingdom, and the United States) in addition to the world market portfolio. Excess return are obtained by subtracting the month-end return on the US 30 day T-bill as reported by the CRSP data files and the sample span from January 1970 to July 2001 (379 observations). The portfolio weights for each strategy are determined each month using moments estimated from a rolling-window of 120 months, and these portfolio weights are then used to calculate the returns in the 121st month. The resulting out-of-sample period spans from January 1980 to July 2001 (259 observations). In parenthesis, we report the percentage size of the confidence interval for a $F_{8,112}$ implied by the values of ϵ_a and the percentage size of the confidence interval for a t_{119} (which is the limiting case of the F distribution if there is only one factor) implied by the values of ϵ_b. The Sharpe ratio for the minimum-variance portfolio, which is not nested by any of the models considered in this table, is 0.1490.

On the contrary, if we set $\epsilon_a = 0$ but let $\epsilon_b > 0$, then for the case where the investor does not believe in the model ($\omega = 0$), the investor holds only the individual country indexes and not the benchmark, and the Sharpe ratio in this case is 0.1127. In general, we can see that for the case where the investor believes dogmatically in the model ($\omega = 1$), uncertainty about the expected return of the benchmark (ϵ_b) has a greater impact on out-of-sample Sharpe ratios than uncertainty about the expected returns of the assets (ϵ_a). In fact, the Sharpe ratios hardly change as we allow for more uncertainty over the expected returns of assets (going down rows) but they

are significantly affected by changes in uncertainty about the expected returns of the benchmark, ϵ_b (going across columns). This makes sense because, as noted above, the Data-and-Model approach with $\omega = 1$ performs significantly better than the traditional mean-variance model. Allowing for uncertainty over the estimates of the benchmark will, therefore, cause a deviation from the portfolio that invests only in the world index. In general, whether this increases or decreases the out-of sample performance depends on the data set being considered.

In our example, for the general case where $\epsilon_a > 0$ and also $\epsilon_b > 0$, the investor allocates wealth to the world market portfolio and also to the individual country indices. For this particular data set, the factor portfolio is useful in describing returns. Thus, if one believes in the factor model ($\omega = 1$), then for small values of $\epsilon_a > 0$ and $\epsilon_b > 0$, such as $\epsilon_a = 0.25$ and $\epsilon_b = 0.50$, the ambiguity-averse portfolio has a Sharpe ratio of 0.1284, which is greater than that for the mean-variance portfolio and the Bayesian portfolios. However, for larger values of ϵ_b, the performance of the ambiguity-averse portfolio declines. On the contrary, an investor who does not believe in the factor model at all ($\omega = 0$) but is simply more confident about the estimate of the benchmark expected returns than about the estimates of the expected returns on each individual asset, will form a portfolio that has a larger investment in the factor portfolio, and, consequently, will achieve an out-of-sample performance similar to the portfolio chosen by an investor who dogmatically believes in the model (see, for example, the case in which $\omega = 0$, $\epsilon_b = 0.50$, and $\epsilon_a = 3$).

4. Conclusion

Traditional mean-variance portfolio optimization assumes that the expected returns used as inputs to the model are estimated with infinite precision. In practice, however, it is extremely difficult to estimate expected returns precisely. And, portfolios that ignore estimation error have very poor properties: the portfolio weights have extreme values that fluctuate dramatically over time and deliver very low Sharpe ratios over time. The Bayesian approach that is traditionally used to deal with estimation error assumes, however, that investors are neutral to ambiguity.

In this article, we have shown how to allow for the possibility of multiple priors and aversion to ambiguity about both the estimated expected returns and the underlying return-generating model. The model with ambiguity aversion relies on imposing constraints on the mean-variance portfolio optimization program, which restrict each parameter to lie within a specified confidence interval of its estimated value. This constraint reflects the possibility of estimation error. And, in addition to the standard maximization of the mean-variance objective function over the choice of weights, one also minimizes over the choice of

parameter values subject to this constraint. This minimization reflects ambiguity aversion, that is, the desire of the investor to guard against estimation error by making choices that are conservative.

We show analytically that the max-min problem faced by an investor who is concerned about estimation uncertainty can be reduced to a maximization-only problem but where the estimated expected returns are adjusted to reflect the parameter uncertainty. The adjustment depends on the precision with which parameters are estimated, the length of the data series, and the investor's aversion to ambiguity. For the case without a riskless asset and where estimation of expected returns for all assets is done jointly, we show that the optimal portfolio can be characterized as a weighted average of the standard mean-variance portfolio, which is the portfolio where the investor ignores the possibility of error in estimating expected returns, and the minimum-variance portfolio, which is the portfolio formed by completely ignoring expected returns. We also explain that the portfolio formed using Bayesian estimation methods is nested in the model with ambiguity aversion.

In a simple setting that is closely related to the familiar mean-variance asset-allocation setup, our article illustrates the ambiguity-aversion approach using data on returns for international equity indices. First, we consider the case where there is only parameter uncertainty about expected returns and then the case with uncertainty both about the factor model generating returns and also about expected returns. We find that the portfolio weights from the model with ambiguity aversion are less unbalanced and fluctuate much less over time compared with the standard mean-variance portfolio weights and also portfolios from the Bayesian models developed by Jorion (1985) and Pástor (2000). We find that allowing for a small amount of aversion to ambiguity about factor and asset returns leads to an out-of-sample Sharpe ratio that is greater than that of the mean-variance and Bayesian portfolios.

One limitation of our analysis is that, like other models in this literature, we do not allow for learning in the formal sense.[22] One could defend the decision to ignore the effect of learning on several grounds. For instance, some articles find the effect of learning to be small [see Hansen, Sargent, and Wang (2002, p. 4)], whereas other papers have argued that after a certain point, not much can be learned [see Anderson, Hansen, and Sargent (2000, p. 2) and Chen and Epstein (2002, p. 1406)]. It is clear, however, that the issue of learning in a world of ambiguity is a complex one. Some theories on how learning would work in the presence of

---

[22] Epstein and Schneider (2004, p. 29) state: "There exist a number of applications of multiple-priors utility or the related robust-control model to portfolio choice or asset pricing. None of these is concerned with learning. Multiple-priors applications typically employ a constant set of one-step-ahead probabilities [Routledge and Zin (2001), Epstein and Miao (2003)]. Similarly, existing robust control models [Hansen, Sargent, and Tallarini (1999), Cagetti et al. (2002)] do not allow the 'concern for robustness' to change in response to new observations. Neither is learning modeled by Uppal and Wang (2003), who pursue a third approach to accommodating ambiguity or robustness."

ambiguity have been developed by Epstein and Schneider (2003) and Wang (2003). The impact of learning on portfolio selection with multiple priors and ambiguity aversion is left as a topic for future research.

Appendix

Proof of Proposition 1

The solution to the inner minimization problem is

$$\mu_j = \hat{\mu}_j - \text{sign}(w_j)\sigma_j \frac{\sqrt{\epsilon_j}}{\sqrt{T}}, \quad j = 1,\dots,N,$$

which, if substituted back into the original problem, gives

$$\max_{w} \left\{ w^{\top}\hat{\mu} - \sum_{j}^{N} \text{sign}(w_j)w_j \frac{\sigma_j}{\sqrt{T}}\sqrt{\epsilon_j} - \frac{\gamma}{2}w^{\top}\Sigma w \right\}, \tag{A1}$$

subject to $w^{\top}\mathbf{1}_N = 1$.[23] Collecting the first two terms in the curly brackets gives the result in the proposition. ∎

Proof of Proposition 2

Let us focus first on the inner minimization

$$\min_{\mu} w^{\top}\mu - \frac{\gamma}{2}w^{\top}\Sigma w, \tag{A2}$$

subject to

$$(\hat{\mu} - \mu)^{\top}\Sigma^{-1}(\hat{\mu} - \mu) \leq \epsilon. \tag{A3}$$

The Lagrangian is

$$\mathcal{L}(\mu,\lambda) = w^{\top}\mu - \frac{\gamma}{2}w^{\top}\Sigma w - \lambda \left[\varepsilon - (\hat{\mu} - \mu)^{\top}\Sigma^{-1}(\hat{\mu} - \mu) \right]. \tag{A4}$$

It is well known that μ^* is a solution of the constrained problem (A2)–(A3) if and only if there exist a scalar $\lambda^* \geq 0$, such that (μ^*,λ^*) is a solution of the following unconstrained problem

$$\min_{\mu} \max_{\lambda} \mathcal{L}(\mu,\lambda). \tag{A5}$$

From the first-order conditions with respect to μ in Equation (A4), we obtain

---

[23] Note that the objective function is not differentiable at $w_j = 0, j = 1, \dots, N$.

74

$$\mu^* = \hat{\mu} - \frac{1}{2\lambda}\Sigma w. \tag{A6}$$

Substituting this in the Lagrangian (A4), we get

$$\mathcal{L}(\mu^*,\lambda) = w^\top\hat{\mu} - \left(\frac{1}{4\lambda} + \frac{\gamma}{2}\right)w^\top\Sigma w - \lambda\varepsilon. \tag{A7}$$

Hence, the original max-min problem in Equation (6) subject to Equation (8) and (15) is equivalent to the following *maximization* problem

$$\max_{w,\lambda} w^\top\hat{\mu} - \left(\frac{1}{4\lambda} + \frac{\gamma}{2}\right)w^\top\Sigma w - \lambda\varepsilon, \tag{A8}$$

subject to $w^\top\mathbf{1}_N = 1$. Solving for λ, we obtain $\lambda = \frac{1}{2}\sqrt{\frac{w^\top\Sigma w}{\varepsilon}} > 0$ from which, upon substitution in Equation (A8), we obtain Equation (16).

The maximization in Equation (16) can be rewritten as follows

$$\max_w w^\top\hat{\mu} - \frac{\gamma}{2}w^\top\Sigma w\left(1 + \frac{2\sqrt{\varepsilon}}{\gamma\sqrt{w^\top\Sigma w}}\right),$$

subject to $w^\top\mathbf{1}_N = 1$. Letting $\Omega(w) \equiv \left(1 + \frac{2\sqrt{\varepsilon}}{\gamma\sqrt{w^\top\Sigma w}}\right)$, we can write the above maximization as

$$\max_w w^\top\hat{\mu} - \frac{\gamma}{2}w^\top\Omega(w)w,$$

subject to $w^\top\mathbf{1}_N = 1$. The Lagrangian is

$$\mathcal{L}(w,\lambda) = w^\top\hat{\mu} - \frac{\gamma}{2}w^\top\Omega(w)w + \lambda(1 - w^\top\mathbf{1}_N).$$

The first-order conditions with respect to w gives

$$\hat{\mu} - \left(\frac{\sqrt{\varepsilon} + \gamma\sqrt{w^\top\Sigma w}}{\sqrt{w^\top\Sigma w}}\right)\Sigma w - \lambda\mathbf{1}_N = 0.$$

Let $\sigma_P \equiv \sqrt{w^\top\Sigma w}$. From the last equation

$$w = \left(\frac{\sigma_P}{\sqrt{\varepsilon} + \gamma\sigma_P}\right)\Sigma^{-1}(\hat{\mu} - \lambda\mathbf{1}_N). \tag{A9}$$

Using $w^\top\mathbf{1}_N = 1$, we can write

$$1 = w^\top \mathbf{1}_N = \left(\frac{\sigma_P}{\sqrt{\varepsilon} + \gamma \sigma_P} \right) (\hat{\mu}^\top - \lambda \mathbf{1}_N^\top) \Sigma^{-1} \mathbf{1}_N$$

$$= \left(\frac{\sigma_P}{\sqrt{\varepsilon} + \gamma \sigma_P} \right) (B - \lambda A),$$

where $A = \mathbf{1}_N^\top \Sigma^{-1} \mathbf{1}_N$ and $B = \hat{\mu}^\top \Sigma^{-1} \mathbf{1}_N$. From the last equality, we obtain

$$\lambda = \frac{1}{A} \left(B - \frac{\sqrt{\varepsilon} + \gamma \sigma_P}{\sigma_P} \right).$$

Substituting this in the expression for the weights in Equation (A9), we arrive at

$$w = \frac{\sigma_P}{\sqrt{\varepsilon} + \gamma \sigma_P} \Sigma^{-1} \left[\hat{\mu} - \frac{1}{A} \left(B - \frac{\sqrt{\varepsilon} + \gamma \sigma_P}{\sigma_P} \right) \mathbf{1}_N \right]. \qquad (A10)$$

We obtain, after some manipulation, that the variance of the optimal portfolio w^* subject to $w^\top \mathbf{1}_N = 1$ is given by the (unique) positive real solution σ_P^* of the following polynomial equation

$$A\gamma^2 \sigma_P^4 + 2A\gamma \sqrt{\varepsilon} \sigma_P^3 + (A\varepsilon - AC + B^2 - \gamma^2) \sigma_P^2 - 2\gamma \sqrt{\varepsilon} \sigma_P - \varepsilon = 0, \qquad (A11)$$

where $A = \mathbf{1}_N^\top \Sigma^{-1} \mathbf{1}_N$, $B = \hat{\mu}^\top \Sigma^{-1} \mathbf{1}_N$, and $C = \hat{\mu}^\top \Sigma^{-1} \hat{\mu}$. Note that because Σ is definite positive, the above polynomial equation always has at least one positive real root. Let σ_P^* be the unique positive real root of this equation.[24] Then, the optimal portfolio w^* is given by

$$w^* = \frac{\sigma_P^*}{\sqrt{\varepsilon} + \gamma \sigma_P^*} \Sigma^{-1} \left[\hat{\mu} - \frac{1}{A} \left(B - \frac{\sqrt{\varepsilon} + \gamma \sigma_P^*}{\sigma_P^*} \right) \mathbf{1}_N \right], \qquad (A12)$$

which simplifies to Equation (17). ∎

Proof of Proposition 3

Without loss of generality, we consider the case of $M = 2$ nonoverlapping subsets. The two subsets are labeled a containing M_a assets and f containing M_b factors (or assets), with $M_a + M_b = N$. Because there are only two subclasses, if we label by a subclass m, subclass $-m$ will be labeled by f and vice-versa. The investor faces the following problem:

$$\max_w \min_{\mu_a, \mu_b} w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w, \qquad (A13)$$

subject to

---

[24] It is possible to show that the fourth degree polynomial in Equation (A11) has at least two real roots, one of which is positive. This is because the polynomial is equal to $-\varepsilon$ at $\sigma_P = 0$ and tends to $+\infty$ for $\sigma_P \to \pm\infty$. Moreover, the first derivative of Equation (A11) is negative at $\sigma_P = 0$ and has at least a negative local maximum, implying that the positive real root of Equation (A11) is unique.

$$(\hat{\mu}_a - \mu_a)^\top \Sigma_{aa}^{-1}(\hat{\mu}_a - \mu_a) \leq \epsilon_a, \tag{A14}$$

$$(\hat{\mu}_b - \mu_b)^\top \Sigma_{bb}^{-1}(\hat{\mu}_b - \mu_b) \leq \epsilon_b. \tag{A15}$$

The Lagrangian of the inner minimization is

$$\begin{aligned}
\mathcal{L}(\mu_a, \mu_b, \lambda_b, \lambda_a) = \; & w_a^\top \mu_a + w_b^\top \mu_b - \frac{\gamma}{2} w^\top \Sigma w \\
& - \lambda_a \Big[\epsilon_a - (\hat{\mu}_a - \mu_a)^\top \Sigma_{aa}^{-1}(\hat{\mu}_a - \mu_a) \Big] \\
& - \lambda_b \Big[\epsilon_b - (\hat{\mu}_b - \mu_b)^\top \Sigma_{bb}^{-1}(\hat{\mu}_b - \mu_b) \Big].
\end{aligned} \tag{A16}$$

Solving for μ_a and μ_b in the inner minimization yields

$$\mu_a^* = \hat{\mu}_a - \frac{1}{2\lambda_a} \Sigma_{aa} w_a, \quad \mu_b^* = \hat{\mu}_b - \frac{1}{2\lambda_b} \Sigma_{bb} \tag{A17}$$

$$\lambda_a = \frac{1}{2} \frac{\sqrt{w_a^\top \Sigma_{aa} w_a}}{\sqrt{\epsilon_a}}, \quad \lambda_b = \frac{1}{2} \frac{\sqrt{w_b^\top \Sigma_{bb} w_b}}{\sqrt{\epsilon_b}}, \tag{A18}$$

where $\lambda_a \geq 0$ and $\lambda_b \geq 0$ are the Lagrange multipliers for the constraints (A14) and (A15). Substituting these back in the Lagrangian, we can rewrite the original maxmin problem as

$$\max_{w_a, w_b} w_a^\top \hat{\mu}_a + w_b^\top \hat{\mu}_b - \frac{\gamma}{2} \big[w_a^\top \hat{\Sigma}_{aa}(w_a, \epsilon_a) w_a - w_a^\top \Sigma_{ab} w_b - w_b^\top \Sigma_{ba} w_a - w_b^\top \hat{\Sigma}_{bb}(w_b, \epsilon_b) w_b \big], \tag{A19}$$

where

$$\hat{\Sigma}_{aa}(w_a, \epsilon_a) = \left(1 + \frac{2}{\gamma} \frac{\sqrt{\epsilon_a}}{\sqrt{w_a \Sigma_{aa} w_a}} \right), \tag{A20}$$

$$\hat{\Sigma}_{bb}(w_b, \epsilon_b) = \left(1 + \frac{2}{\gamma} \frac{\sqrt{\epsilon_b}}{\sqrt{w_b \Sigma_{bb} w_b}} \right). \tag{A21}$$

Let us define

$$\sigma_a \equiv \sqrt{w_a \Sigma_{aa} w_a}, \tag{A22}$$

$$\sigma_b \equiv \sqrt{w_b \Sigma_{bb} w_b}. \tag{A23}$$

Because Σ_{aa} and Σ_{bb} are positive definite, $\sigma_a > 0$ unless $w_a = \mathbf{0}_{M_a \times 1}$, and similarly, $\sigma_b > 0$ unless $w_b = \mathbf{0}_{M_b \times 1}$. Now the first-order conditions with respect to w_a and w_b yield

$$\left(\frac{\sqrt{\epsilon_a} + \gamma \sigma_a}{\sigma_a} \right) \Sigma_{aa} w_a = \hat{\mu}_a - \gamma \Sigma_{ab} w_b, \tag{A24}$$

$$\left(\frac{\sqrt{\epsilon_b} + \gamma \sigma_b}{\sigma_b} \right) \Sigma_{bb} w_b = \hat{\mu}_b - \gamma \Sigma_{ba} w_a. \tag{A25}$$

Given w_b, we can solve for σ_a in Equation (A24) and get

$$\sigma_a = \frac{1}{\gamma} \max \left[\sqrt{g(w_b)^\top \Sigma_{aa}^{-1} g(w_b)} - \sqrt{\epsilon_a},\, 0 \right], \tag{A26}$$

where $g(w_b) = \hat{\mu}_a - \gamma \Sigma_{ab} w_b$. Substituting this back into Equation (A24), we obtain Equation (22) if $m = a$. Similarly, given w_a, we can solve for σ_b in Equation (A25) and get

$$\sigma_b = \frac{1}{\gamma} \max \left[\sqrt{h(w_a)^\top \Sigma_{bb}^{-1} h(w_a)} - \sqrt{\epsilon_a},\, 0 \right], \tag{A27}$$

where $h(w_a) = \hat{\mu}_b - \gamma \Sigma_{ba} w_a$. Substituting this back into Equation (A25) we obtain Equation (22) if $m = b$.

Note finally that it is always possible to find a set in which the mapping $\Upsilon : \mathbb{R}^{M_a} \times \mathbb{R}^{M_b} \to \mathbb{R}^{M_a} \times \mathbb{R}^{M_b}$ defined by Equation (22) admits a solution. To see this, let $W_a \subset \mathbb{R}^{M_a}$ and $W_b \subset \mathbb{R}^{M_b}$ be nonempty, closed, bounded, and convex subsets containing the origin. Define the set

$$\Gamma = \left[\frac{1}{\gamma} \Sigma_{aa}^{-1} g(W_b) \right] \times \left[\frac{1}{\gamma} \Sigma_{bb}^{-1} h(W_a) \right] \subset \mathbb{R}^{M_a} \times \mathbb{R}^{M_b}. \tag{A28}$$

Because Υ is continuous and the "max" in Equation (22) are bounded between 0 and 1, $\Upsilon(\Gamma) \subseteq \Gamma$.

By the Brouwer Fixed-Point Theorem [see, for example, Stokey and Lucas (1989, p. 517)], Υ has a fixed point in Γ. ∎

References

Anderson, E., L. P. Hansen, and T. J. Sargent, 2000, "Robustness Detection and the Price of Risk," working paper, University of Chicago.

Barry, C. B., 1974, "Portfolio Analysis under Uncertain Means, Variances, and Covariances," *Journal of Finance*, 29, 515–522.

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

Berger, J., 1974, "Minimax Estimator of a Multivariate Normal Mean Under Polynomial Loss," *Journal of Multivariate Analysis*, 8, 173–180.

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

Bewley, T., 1988, "Knightian Decision Theory and Econometric Inference," working paper, Yale University.

Black, F., and R. Litterman, 1990, "Asset Allocation: Combining Investor Views with Market Equilibrium," Discussion paper, Goldman, Sachs & Co.

Black, F., and R. Litterman, 1992, "Global Portfolio Optimization," *Financial Analysts Journal*, 48, 28–43.

Cagetti, M., L. P. Hansen, T. J. Sargent, and N. Williams, 2002, "Robustness and Pricing with Uncertain Growth," *Review of Financial Studies*, 15, 363–404.

Chen, Z., and L. Epstein, 2002, "Ambiguity, Risk, and Asset Returns in Continuous Time," *Econometrica*, 70, 1403–1443.

Chopra, V. K., 1993, "Improving Optimization," *Journal of Investing*, 8, 51–59.

Chopra, V. K., and W. T. Ziemba, 1993, "The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice," *Journal of Portfolio Management*, 19, 6–11.

De Santis, G., and B. Gerard, 1997, "International Asset Pricing and Portfolio Diversification with Time-Varying Risk," *Journal of Finance*, 52, 1881–1912.

Dow, J., and S. R. d. C. Werlang, 1992, "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio," *Econometrica*, 60, 197–204.

Dumas, B., and B. Jacquillat, 1990, "Performance of Currency Portfolios Chosen by a Bayesian Technique: 1967–1985," *Journal of Banking and Finance*, 14, 539–-558.

Ellsberg, D., 1961, "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

Epstein, L., and J. Miao, 2003, "A Two-Person Dynamic Equilibrium Under Ambiguity," *Journal of Economic Dynamics and Control*, 27, 1253–1288, working paper, University of Rochester.

Epstein, L., and M. Schneider, 2003, "Recursive Multiple-Priors," *Journal of Economic Theory*, 113, 1–31.

Epstein, L., and M. Schneider, 2004, "Learning Under Ambiguity," working paper, University of Rochester.

Epstein, L. G., and T. Wang, 1994, "Intertemporal Asset Pricing Under Knightian Uncertainty," *Econometrica*, 62, 283–322.

Fox, C., and A. Tversky, 1995, "Ambiguity Aversion and Comparative Ignorance," *Quarterly Journal of Economics*, 110, 585–603.

Frost, P. A., and J. E. Savarino, 1986, "An Empirical Bayes Approach to Efficient Portfolio Selection," *Journal of Financial and Quantitative Analysis*, 21, 293–305.

Frost, P., and J. Savarino, 1988, "For Better Performance Constrain Portfolio Weights," *Journal of Portfolio Management*, 15, 29–34.

Ghirardato, P., F. Maccheroni, and M. Marinacci, 2004, "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory*, 118, 133–173.

Gilboa, I., and D. Schmeidler, 1989, "Maxmin Expected Utility Theory with Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

Goldfarb, D., and G. Iyengar, 2003, "Robust Portfolio Selection Problems," *Mathematica of Operations Research*, 28, 1–38.

Halldórsson, B., and R. Tüutüncü, 2000, "An Interior-Point Method for a Class of Saddle Point Problems," working paper, Carnegie Mellon University.

Hansen, L. P., T. J. Sargent, and T. D. Tallarini, 1999, "Robust Permanent Income and Pricing," *Review of Economic Studies*, 66, 873–907.

Hansen, L. P., T. J. Sargent, and N. Wang, 2002, "Robust Permanent Income and Pricing with Filtering," *Macroeconomic Dynamics*, 6, 40–84.

Harvey, C. R., J. Liechty, M. Liechty, and P. Müller, 2003, "Portfolio Selection with Higher Moments," working paper, Duke University.

Heath, C., and A. Tversky, 1991, "Preferences and Beliefs: Ambiguity and Competence in Choice under Uncertainty," *Journal of Risk and Uncertainty*, 4, 5–28.

Hodges, S. D., and R. A. Brealey, 1978, "Portfolio Selection in a Dynamic and Uncertin World," in J. H. Lorie and R. A. Brealey (eds.), *Modern Developments in Investment Management* (2d ed.), Dryden Press, Hinsdale, Illinois.

Jagannathan, R., and T. Ma, 2003, "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance*, 58, 1651–1684.

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

Johnson, R., and D. W. Wichern, 1992, *Applied Multivariate Statistical Analysis*, Prentice Hall, Englewood Cliffs, NJ.

Jorion, P., 1985, "International Portfolio Diversification with Estimation Risk," *Journal of Business*, 58, 259–278.

Jorion, P., 1986, "Bayes–Stein Estimation for Portfolio Analysis," *Journal of Financial and Quantitative Analysis*, 21, 279–292.

Jorion, P., 1991, "Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection," *Journal of Banking and Finance*, 15, 717–727.

Jorion, P., 1992, "Portfolio Optimization in Practice," *Financial Analysts Journal*, 48, 68–74.

Klein, R. W., and V. S. Bawa, 1976, "The Effect of Estimation Risk on Optimal Portfolio Choice," *Journal of Financial Economics*, 3, 215–231.

Klibanoff, P., M. Marinacci, and S. Mukerji, 2005, "A Smooth Model of Decision Making Under Ambiguity," *Econometrica*, 73, 1849–1892.

Knight, F., 1921, *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston.

Kogan, L., and T. Wang, 2002, "A Simple Theory of Asset Pricing under Model Uncertainty," working paper, University of British Columbia.

Litterman, R., 2003, *Modern Investment Management: An Equilibrium Approach*, Wiley, New York.

Maenhout, P. J., 2004, "Robust Portfolio Rules and Asset Pricing," *Review of Financial Studies*, 17, 951–983.

Markowitz, H. M., 1952, "Mean-Variance Analysis in Portfolio Choice and Capital Markets," *Journal of Finance*, 7, 77–91.

Markowitz, H. M., 1959, *Portfolio Selection*, Wiley, New York.

Merton, R., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373–413.

Merton, R., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

Michaud, R. O., 1989, "The Markowitz Optimization Enigma: Is Optimized Optimal?" *Financial Analysts Journal*, 45, 31–42.

Michaud, R.O., 1998, *Efficient Asset Management*, Harvard Business School Press, Boston.

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

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

Routledge, B., and S. Zin, 2001, "Model Uncertainty and Liquidity," NBER Working Paper No. 8683.

Scherer, B., 2002, "Portfolio Resampling: Review and Critique," *Financial Analysts Journal*, 58, 98–109.

Sharpe, W. F., 1970, *Portfolio Theory and Capital Markets*, McGraw-Hill, New York.

Stambaugh, R. F., 1997, "Analyzing Investments whose Histories Differ in Length," *Journal of Financial Economics*, 45, 285–331.

Stein, C., 1955, "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," in *3rd Berkely Symposium on Probability and Statistics*, vol. 1, pp. 197–206, University of California Press, Berkeley.

Stokey, N. L., and J. R. E. Lucas, 1989, *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, MA.

Tütüncü, R., and M. Koenig, 2004, "Robust Asset Allocation," *Annals of Operations Research*, 132, 157–187.

Uppal, R., and T. Wang, 2003, "Model Misspecification and Underdiversification," *Journal of Finance*, 58, 2465–2486.

Wang, T., 2003, "Conditional Preferences and Updating," *Journal of Economic Theory*, 108, 286–321.

Wang, Z., 2005, "A Shrinkage Approach to Model Uncertainty and Asset Allocation," *Review of Financial Studies*, 18, 673–705.

Zellner, A., and V. K. Chetty, 1965, "Prediction and Decision Problems in Regression Models from the Bayesian Point of View," *Journal of the American Statistical Association*, 58, 608–616.