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

Lorenzo Garlappi† Raman Uppal‡ Tan Wang§

November 8, 2005

*We gratefully acknowledge financial support from INQUIRE UK; this article however represents the views of the authors and not of INQUIRE. We are very grateful to an anonymous referee and to Luboš Pástor for detailed suggestions on how to improve the paper. We would also like to acknowledge comments from Nicholas Barberis, Suleyman Basak, Ian Cooper, Victor DeMiguel, Francisco Gomes, Tim Johnson, Maureen O’Hara, Catalina Stefanescu, Yongjun Tang, Sheridan Titman, Roberto Wessels and participants at presentations given at Copenhagen Business School, Imperial College, Lancaster University, London Business School, University of Maryland, University of Minnesota, University of Texas at Austin, the INQUIRE Fall 2003 conference, the 2004 AGSM Accounting and Finance Summer Camp, the 2004 UBC summer conference, and the 2005 meetings of the European Finance Association.

†McCombs School of Business, The University of Texas at Austin, Austin TX, 78712; Email: lorenzo.garlappi@mccombs.utexas.edu.

‡London Business School and CEPR; IFA, 6 Sussex Place Regent’s Park, London, United Kingdom NW1 4SA; Email: rupal@london.edu.

§Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver Canada V6T 1Z2, Canada; Email: tan.wang@sauder.ubc.ca.
Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach

Abstract

In this paper, we show how an investor can incorporate uncertainty/ambiguity about expected returns when choosing a mean-variance optimal portfolio. In contrast to the Bayesian approach to estimation error, where there is typically a single prior and the investor is neutral to ambiguity, we consider the case where the investor has multiple priors and is averse to ambiguity. We characterize the set of multiple priors by a "confidence interval" around the estimated value of expected returns and we model aversion to ambiguity via a minimization over the set of priors. The multi-prior model with ambiguity aversion has several attractive features: One, just like the Bayesian model, it has a solid axiomatic foundation. Two, it is flexible enough to allow for different degrees of uncertainty about expected returns for different subsets of assets, and also about the underlying return-generating model. Three, for several formulations of the model we obtain closed-form expressions for the optimal portfolio, and in one special case we prove that the optimal portfolio is equivalent to a "shrinkage" portfolio based on the mean-variance and minimum-variance portfolios, which allows for a transparent comparison with Bayesian portfolios. We illustrate how to use the multi-prior model with ambiguity aversion by considering the portfolio problem of a fund manager allocating wealth across eight international equity indices; our empirical analysis suggests that portfolios that incorporate aversion to parameter and model uncertainty tend to over-weight the risk-free asset, are more stable over time, and deliver a higher out-of-sample Sharpe ratio than the portfolios from both classical and Bayesian models.

Keywords: Portfolio choice, asset allocation, estimation error, ambiguity, uncertainty, robustness.

JEL Classification: G11, D81
1 Introduction

Expected returns, variances, and covariances are key inputs of every portfolio selection model. These parameters are not known \textit{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.\footnote{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 \textit{single prior} or, equivalently, to be \textit{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 \textit{multiple-priors} who are \textit{averse} to this ambiguity,\footnote{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.} and hence, desire \textit{robust} portfolio rules that work well for a set of possible models.\footnote{There are two terms used to describe aversion to Knightian uncertainty in the literature. Epstein and coauthors describe this as “ambiguity aversion,” while Hansen and Sargent (and their coauthors) describe it as “wanting robustness”. Both streams of the literature have as their origin the multi-priors model in 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).}

In this paper, 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.\footnote{We focus on the error in estimating expected returns of assets because as shown in 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 ten times as costly as errors in estimating variances, and over twenty times as costly as errors in estimating covariances.}

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: (i) Imposing an \textit{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 (ii) Introducing an \textit{additional minimization} over the set of possible expected returns subject to the additional constraint. The additional constraint recognizes the possibility 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.\footnote{See Section 2 and Bewley (1988) for a discussion of how confidence intervals obtained from classical statistics are related to Knightian uncertainty and Section 3 for the relation of our model to Bayesian models of decision making.}

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 has several attractive features. One, 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.

Two, 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 paper with the traditional Bayesian approach, where the decision maker is neutral to ambiguity. The analytic solutions also indicate how the model with ambiguity aversion can be implemented as a simple maximization problem instead of a much more complicated saddle point problem.

Three, 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 CAPM or the 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.\footnote{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 to those in this literature.}

In order to demonstrate the differences between the portfolios selected by an ambiguity averse investor from 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 to 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 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, which ignores estimates of expected returns, and hence, 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 paper 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

---

\(^7\)Other approaches for dealing with estimation error are to impose arbitrary portfolio constraints prohibiting shortsales (Frost and Savarino (1988) and Chopra (1993)), which Jagannathan and Ma (2003) show 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, Liechty, Liechty, and Müller (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.
algorithms for solving max-min saddle-point problems numerically and apply the algorithms to portfolio choice problem, while Wang (2005) shows how to obtain the optimal portfolio numerically in a Bayesian setting in the presence of aversion to model uncertainty. Our paper differs from Goldfarb and Iyengar (2003), Halldórsson and Tütüncü (2000), and Tütüncü and Koenig (2004) in several 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 paper is organized as follows. In Section 2, 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 3, we discuss the relation of the multi-prior model with ambiguity aversion to the traditional ambiguity-neutral Bayesian approach for dealing with estimation error, and we compare analytically the portfolio weights under the two approaches. Then, in Section 4, 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 5. Proofs for propositions are collected in the Appendix.

2 Portfolio choice with ambiguity aversion

This section is divided into two parts. In the first part, Section 2.1, we summarize the standard mean-variance model of portfolio choice where estimation error is ignored. In the second part, Section 2.2, we show how this model can be extended to incorporate aversion to ambiguity about the estimated parameters and the return-generating model. Throughout the paper, 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 is the “intertemporal hedging component,” which turns out to be quite small once the model is calibrated to realistic values for the parameters driving the processes for asset returns.\footnote{The case of dynamic portfolio choice with only a single risky asset in a robust-control setting is addressed in Maenhout (2004); the case of multiple risky assets if investors are averse to ambiguity is considered in Chen and Epstein (2002), Epstein and Miao (2003), and Uppal and Wang (2003).}

### 2.1 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,

$$
\max_w w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w,
$$

where $\mu$ is the $N$-vector of the true expected excess returns, $\Sigma$ is the $N \times N$ covariance matrix, and the scalar $\gamma$ is the risk aversion parameter. The solution to this problem is

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

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

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

where $\mu^0$ is the expected return on the zero-beta portfolio associated with the optimal portfolio $w$ and is given by

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

where $A = \mathbf{1}_N^\top \Sigma^{-1} \mathbf{1}_N$, $B = \mu^\top \Sigma^{-1} \mathbf{1}_N$ and $\mathbf{1}_N$ a $N$-vector of ones.
A fundamental assumption of the standard mean-variance portfolio selection model in (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

\[
\max_w \ w^\top \hat{\mu} - \frac{\gamma}{2} w^\top \Sigma w,
\]

subject to \( w^\top 1_N = 1 \). The problem in (5) coincides with (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 (5) tend to consist of extreme positions that swing dramatically over time. Moreover, these portfolios often perform poorly out of sample.

### 2.2 Extension of the standard model to incorporate ambiguity aversion

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. One, 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. Two, 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, \( \mu \), 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:

\[
\max_w \min_\mu \ w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w,
\]

subject to

\[
f(\mu, \hat{\mu}, \Sigma) \leq \epsilon,
\]

\[
w^\top 1_N = 1.
\]

As before, equation (8) constrains the weights to sum to unity in the absence of a riskfree asset; if a riskfree asset is available, this constraint can be dropped. And in the constraint
in (7), $f(\cdot)$ is a vector-valued function, and $\epsilon$ is a vector of constants that reflects both the investor’s ambiguity and his aversion to ambiguity. Specifically, the parameter $\epsilon$ should be understood as the product of ambiguity aversion (common across assets) and ambiguity (asset-specific). Hence, while ambiguity aversion is a general property of preferences, ambiguity is phenomenon-specific.\footnote{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.} 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 one.\footnote{See Ghirardato, Maccheroni, and Marinacci (2004) and Klibanoff, Marinacci, and Mukerji (2005) for models where ambiguity aversion is potentially separable from ambiguity; but, these models require the decision maker to specify subjective probabilities over the priors.} 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 $\epsilon$ 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.

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

### 2.2.1 Uncertainty about expected returns estimated asset-by-asset

We start by considering the case where $f(\mu, \hat{\mu}, \Sigma)$ has $N$ components,

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

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

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

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 $\epsilon_j$ in constraints (10) determine confidence intervals. When all the $N$ constraints in (10) are taken together, (10) is closely related to the probabilistic statement

$$P(\mu_1 \in I_1, \ldots, \mu_N \in I_N) = 1 - p, \quad (11)$$
where \( I_j, j = 1, \ldots, N \), 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 \( \epsilon_j \), then the probability that all the \( N \) true expected returns fall into the \( N \) intervals, respectively, is \( 1 - p = (1 - p_1)(1 - p_2) \cdots (1 - p_N) \).\(^{11}\)

While 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 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 (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

\[
\max_w \left\{ w^\top (\bar{\mu} - \mu^{\text{adj}}) - \frac{\gamma}{2} w^\top \Sigma w \right\},
\]

where \( \mu^{\text{adj}} \) is the \( N \)-vector of adjustments to be made to the estimated expected return:

\[
\mu^{\text{adj}} \equiv \left\{ \begin{array}{c}
sign(w_1) \frac{\sigma_1}{\sqrt{T} \sqrt{\epsilon_1}}, \ldots, sign(w_N) \frac{\sigma_N}{\sqrt{T} \sqrt{\epsilon_N}} \end{array} \right\}.
\]

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 (5), but where the mean has been adjusted to reflect the uncertainty about its estimated value. The term “\( \text{sign}(w_j) \)” in (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, while if it is negative (short position) the expected return on the asset is increased. We characterize the optimal solution for this problem in Section 2.2.3.

\(^{11}\)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. The fact that the data for different assets may be of different lengths does not present a serious problem for the multivariate Normal distribution setting, as shown by Stambaugh (1997).
2.2.2 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

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

(14)

has a $\chi^2$ distribution with $N$ degrees of freedom.\footnote{If $\Sigma$ is not known, then the expression in (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 4, we will use an $F$ distribution.}

Let $f = \frac{T(T-N)}{(T-1)N} (\hat{\mu} - \mu)^\top \Sigma^{-1} (\hat{\mu} - \mu)$ and $\epsilon$ be a chosen quantile for the $\chi^2$-distribution. Then the constraint (7) can be expressed as

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

(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 (8) and (15) can be simplified into a maximization problem which 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 (8) and (15) is equivalent to the following maximization problem

$$\max_w \; w^\top \hat{\mu} - \frac{\gamma}{2} w^\top \Sigma w - \sqrt{\epsilon} w^\top \Sigma w,$$

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

$$w^* = \frac{1}{\gamma} \Sigma^{-1} \left( \frac{1}{1 + \sqrt{\frac{\epsilon}{D(P)}}} \right) \left( \hat{\mu} - \frac{B - \gamma \left( \mu \left(1 + \frac{\sqrt{\epsilon}}{\sigma^*_P} \right) \right)}{A} 1_N \right),$$

(17)

where $A = 1_N^\top \Sigma^{-1} 1_N$, $B = \hat{\mu}^\top \Sigma^{-1} 1_N$, and $\sigma^*_P$ is the variance of the optimal portfolio that can be obtained from solving the polynomial equation (A11) in the appendix.
We can now use the expression in (17) for the optimal weights to interpret the effect of aversion to parameter uncertainty. Note that as \( \varepsilon \to 0 \), that is either \( \epsilon \to 0 \) or \( T \to \infty \), the optimal weight \( w^* \) converges to the mean-variance portfolio

\[
 w^* = \frac{1}{\gamma} \Sigma^{-1} \left( \hat{\mu} - \frac{B - \gamma}{A} \mathbf{1}_N \right) 
\]

\[
 = \frac{1}{\gamma} \Sigma^{-1} \left( \hat{\mu} - \mu^0 \mathbf{1}_N \right),
\]

(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 other hand, as \( \varepsilon \to \infty \) the optimal portfolio converges to

\[
 w^* = \frac{1}{A} \Sigma^{-1} \mathbf{1}_N,
\]

(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.

### 2.2.3 Uncertainty about expected returns estimated for subsets of assets

In Section 2.2.1 we described the case where there was uncertainty about expected returns that were estimated individually asset-by-asset, and in Section 2.2.2 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 a \( M \)-valued function with

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

(20)

Then (15) becomes

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

(21)

\[13\]In taking these limits, it is important to realize that \( \sigma^*_P \) also depends on the weights. In order to obtain the correct limits, it is useful to look at equation (A11) that characterizes \( \sigma^*_P \).
Just as in the earlier specifications, these constraints correspond to the probabilistic statement
\[ P(X_1 \in I_1, \ldots, X_M \in I_M) = 1 - p, \]
where \( X_m, m = 1, \ldots, M, \) are sample statistics defined by the left hand side of the inequalities in (21).

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

**Proposition 3** Consider the case of \( M \) non-overlapping subsets of assets and assume \( f \) in (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:

\[
w_m = \max \left[ 1 - \frac{\sqrt{\varepsilon_m}}{\sqrt{g(w_m)^\top \Sigma_m^{-1} g(w_m)}} \cdot 0 \right],
\]

for \( m = 1, \ldots, M, \) where \( \varepsilon_m \equiv \epsilon_m \left( \frac{(T_m - 1)N_m}{T_m(T_m - N_m)} \right), \) \( w_m \) represents the weights in the assets not in subclass \( m, \) \( \Sigma_m \) is the variance-covariance matrix of the asset in subclass \( m, \) and

\[
g(w_m) = \hat{\mu}_m - \gamma \Sigma_{m,-m} w_m, \quad m = 1, \ldots, M \tag{23}
\]

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

### 2.2.4 Uncertainty about the return-generating model and expected returns

In this section, we explain how the general model developed in Section 2.2.3 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 with \( r_{at} \) the \( N \times 1 \) vector of excess

---

\(^{14}\text{We can characterize analytically the portfolio weights also for the case where the investor does not have access to the riskfree asset; but, the characterization of the weights for this case is less transparent because this problem involves an extra constraint that the weights sum to unity. Therefore, for expositional reasons, we focus on the case where a riskfree asset is available.}\)
returns of the non-benchmark 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:

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

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

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

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

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

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

Let $w \equiv (w_a, w_b)^\top$ to 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

$$
(\hat{\mu}_a - \mu_a)^\top \Sigma_{aa}^{-1} (\hat{\mu}_a - \mu_a) \leq \epsilon_a, \quad (\hat{\mu}_b - \mu_b)^\top \Sigma_{bb}^{-1} (\hat{\mu}_b - \mu_b) \leq \epsilon_b.
$$  \tag{28}

Equations (28) and (29) capture 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.$^{15}$

$^{15}$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 $\mu_b = \mu_b$ (i.e., unless $\epsilon_b = 0$) the difference $\hat{\mu}_a - \mu_a$ does not represent Jensen’s $\alpha$. 

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

\[
\begin{align*}
  w_a &= \max \left[ 1 - \frac{\sqrt{\epsilon_a}}{\sqrt{g(w_b)\Sigma^{-1}_{aa} g(w_b)}}, 0 \right] \frac{1}{\gamma} \Sigma^{-1}_{aa} g(w_b), \\
  w_b &= \max \left[ 1 - \frac{\sqrt{\epsilon_b}}{\sqrt{h(w_a)\Sigma^{-1}_{bb} h(w_a)}}, 0 \right] \frac{1}{\gamma} \Sigma^{-1}_{bb} h(w_a),
\end{align*}
\]

(30)

(31)

where

\[
\begin{align*}
  g(w_b) &= \hat{\mu}_a - \gamma \Sigma_{ab} w_b, \\
  h(w_a) &= \hat{\mu}_b - \gamma \Sigma_{ba} w_a.
\end{align*}
\]

(32)

(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 non-benchmark assets decrease and the holdings of the factor portfolios increase. Third, given a certain level of uncertainty about the assets (i.e., keeping fixed \( \epsilon_a \)), as \( \epsilon_b \) increases, the holdings of risky non-benchmark 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 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 \( \alpha \). The prior on \( \alpha \), conditional on \( \Omega \), is assumed to have a Normal distribution with mean zero and variance \( \tau \Omega \), with \( \tau \) determining the precision of the prior belief over the validity of the asset pricing model. Moreover, the priors on the factor loadings \( \beta \), the variance-covariance of the residuals \( \Omega \), as well as on the expected returns and variance-covariance matrix of the factors are assumed to be non informative 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 towards the model moments. More specifically, denoting by \( \hat{\beta}_a \), \( \hat{\mu}_b \), 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:

\[
\mu_{DM} = \omega \left( \beta \hat{\mu}_a \right) + (1 - \omega) \left( \hat{\mu}_a \right). \tag{34}
\]

In the above expression, \( \beta \) is the estimator of the factor loading \( \beta \) obtained by estimating the model (25) with the restriction that \( \alpha = 0_{N \times 1} \) and \( \omega \) 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 \( \omega \) 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 \( \Sigma_{DM} \). 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 4, we will compare the ambiguity-averse portfolio (30)-(31) with the ambiguity-neutral Bayesian Data-and-Model portfolio (35) by implementing both approaches for data on international equity indices.

3 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: (i) the standard mean-variance portfolio that ignores estimation error, (ii) the minimum-variance portfolio, (iii) the portfolio based on Bayes-diffuse-prior estimates as in Bawa, Brown, and Klein (1979), and (iv) the portfolio based on the empirical Bayes-Stein estimator, as described in 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, while in Section 4 this comparison is undertaken empirically and the comparison set includes also the weights obtained by using the “Data-and-Model” approach of Pástor (2000).

### 3.1 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 \( \Theta \). In the setting of this paper, \( \Theta \) 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 \( \Theta \) 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 \( \Theta \) is often unknown and needs to be estimated from data, i.e., there is parameter uncertainty. In the presence of such parameter uncertainty, Savage’s expected utility approach assumes the investor has a prior over the unknown parameter \( \Theta \) 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

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

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

\[
g(R|X) = \int g(R|\Theta) p(\Theta|X) d\Theta. \tag{38}
\]

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 non-informative diffuse prior or a predictive distribution obtained by integrating over the unknown parameter. In a second implementation, Jobson and Korkie (1980), Jorion (1985,\(^{16}\)The use of the predictive distributions was pioneered by Zellner and Chetty (1965).
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.

3.2 Comparison of the ambiguity-averse approach with 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 returns 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, 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 $\Sigma$ which appears in equation (6). The uncertainty about the unknown mean return vector, $\mu$, is captured by the constraint (7). The two are further separated by the minimization over $\mu$, subject to the constraint in (7), which reflects that the investor is averse, rather than neutral, to ambiguity.
3.3 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 to those obtained if using traditional Bayesian methods to deal with estimation error. We start by describing the portfolio obtained if using the empirical Bayes-Stein estimator. The Bayes-diffuse-prior portfolio is then obtained as a special case of this portfolio, while 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 risk by “shrinking” the sample mean towards 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 \(\bar{\mu}\) to be the mean of the minimum variance portfolio, \(\mu_{MIN}\). More specifically, following Jorion (1986), we use the following shrinkage estimator for the expected return and covariance matrix

\[
\begin{align*}
\mu_{BS} &= (1 - \phi_{BS}) \hat{\mu} + \phi_{BS} \mu_{MIN} 1_N, \quad (39) \\
\Sigma_{BS} &= \Sigma \left( 1 + \frac{1}{T + \nu_\mu} \right) + \frac{\nu_\mu}{T(T + 1 + \nu_\mu)} \frac{1_N 1_N^\top}{1_N \Sigma^{-1} 1_N}. \quad (40)
\end{align*}
\]

where \(\hat{\mu}\) is the sample mean, \(\mu_{MIN}\) is the minimum-variance portfolio, \(\phi_{BS}\) is the shrinkage factor for the mean and \(\nu_\mu\) is the precision (or “tightness”) of the prior on \(\mu\):

\[
\begin{align*}
\phi_{BS} &= \frac{\nu_\mu}{T + \nu_\mu}, \quad (41) \\
\nu_\mu &= \frac{N + 2}{(\hat{\mu} - \mu_{MIN})^\top \Sigma^{-1} (\hat{\mu} - \mu_{MIN})}. \quad (42)
\end{align*}
\]

Note that the case of zero precision (\(\nu_\mu = 0\)) corresponds to the Bayes-diffuse-prior case considered in Bawa, Brown, and Klein (1979) in which 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

\(^{17}\)Stein (1955) and Berger (1974) developed the idea of shrinking the sample mean towards 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.
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 (3). Substituting the empirical Bayes-Stein (BS) estimator $\mu_{BS}$ in (3), one can show that the optimal weights can be written as

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

(43)

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

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

(44)

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

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

(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-aversion (AA) 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

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

(46)

where

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

(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 $\epsilon$ 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 Jorion (1985, 1986), or both the data and an asset pricing model, as in 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: one, 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 two, 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 (43). But, one could just as well have a third case that takes estimation uncertainty into account and allows for ambiguity aversion. In this case, the optimal portfolio weights are given by the following expression.

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

(48)

where the minimum-variance-portfolio, \( w_{MIN} \), and the Bayesian portfolio, \( w_{BS} \), are defined in (44) and (43), respectively. Observe that the expression in (48) is similar to that in (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 (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 (43) and (46), which suggests that the Bayes-Stein approach in Jorion (1985, 1986) and the ambiguity-averse approach (with a single joint constraint, as described in Section 2.2.2) 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 2.2.3, where the estimation is done separately for different subsets of assets, a special case of which is discussed in Section 2.2.1, where the expected return on each asset is estimated individually. Once the constraint in (7) is imposed separately for different

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

4 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 to other procedures commonly used, with a particular emphasis on the Bayes-Stein approach in Jorion (1985, 1986) and the Bayesian “Data-and-Model” approach in Pástor (2000).

The experiment considered in this section of the paper is similar in spirit to the one in Jorion (1985, 1986). In these papers, Jorion, considers a 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. But, 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 4.1 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 2.2.2. Then, in Section 4.2 we illustrate the case with both uncertainty about expected returns and the factor model generating these returns; this is the model described in Section 2.2.4.

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, United Kingdom, United States. For the second application, we consider also the return on the world market portfolio. Data are from MSCI (Morgan Stanley Capital International). 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 United States 30 day T-bill as reported in the 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 121\textsuperscript{st} 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 $\gamma = 1$.

4.1 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 2.2.2. Using this model, we compute the portfolios that account for different degrees of uncertainty in the statistical estimate ($\epsilon$) 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 in 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

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

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.\textsuperscript{19} Notice from Panel A that the case of $\epsilon = 0$ corresponds to the mean-variance portfolio, while the case of $\epsilon \to \infty$ corresponds to the minimum-variance portfolio, as discussed in the previous section. From

\textsuperscript{19}The number in parenthesis appearing in the table refer to the percentage-confidence interval implied by different value of $\epsilon$ and computed from a $F_{8,112}$ distribution.
the table, we see that compared to the mean-variance strategy in which historical mean returns $\hat{\mu}$ are taken to be the estimator of expected returns $\mu$, 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 ratios than any of the portfolios that account for ambiguity aversion.\footnote{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)) while 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.}

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 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, while this factor for the model with aversion to parameter uncertainty is 0.8302 if $\epsilon = 1$ and, as shown in Proposition 2, this factor increases with $\epsilon$. 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 56% 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 $\epsilon$; moreover, the shrinkage factor fluctuates much less for higher $\epsilon$.

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 (BS) 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 ($\epsilon = 1$) which corresponds to a degree of uncertainty expressed roughly by the 56% confidence interval for an $F_{8,112}$ centered around the sample mean, while 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 has less extreme positions and the portfolio weights vary much less over time compared to 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 $\epsilon$ means a higher confidence interval and, consequently, more uncertainty in the estimates. As a consequence, the larger is $\epsilon$, 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 2.2.2, but with the additional constraint that short sales are not allowed: $w \geq 0_N$. 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.\footnote{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.}

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. 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). Though 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 \textit{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 (BS) 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.

### 4.2 Parameter and model uncertainty

In this section, we implement the model discussed in Section 2.2.4, 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 2.2.4) 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 \( \epsilon_a \). 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 \( \hat{\mu}_a = \beta \hat{\mu}_b \), where \( \beta \) is the \( 8 \times 1 \) vector of betas obtained by estimating the model (25) with the restriction that \( \alpha = 0_{8 \times 1} \). In this case too, the investor allows for uncertainty about the estimate from the model, captured by the parameter \( \epsilon_a \). 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 \( \epsilon_b \). 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 (\( \omega = 0 \)), while 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 in 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 in 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.

On the other hand, 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 (\( \epsilon_b \)) has a greater impact on out-of-sample Sharpe ratios than uncertainty about the expected returns of the assets (\( \epsilon_a \)). In fact, the Sharpe ratios hardly change as we allow for more uncertainty over the expected returns of assets (going down rows) while they are significantly affected by changes in uncertainty about the expected returns of the benchmark, \( \epsilon_b \) (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 dataset 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 \( \epsilon_b \), the performance of the ambiguity-averse portfolio declines. On the other hand, 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 \( \epsilon_a = 3 \)).
5 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 paper, 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 on 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 paper 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 fac-
tor 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 to the standard mean-variance portfolio weights and also portfolios from the Bayesian models developed in 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,\textsuperscript{22} we do not allow for learning in the formal sense. One could defend the decision to ignore the effect of learning on several grounds. For instance, some papers find the effect of learning to be small (see Hansen, Sargent, and Wang (2002, p. 4)), while 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 ambiguity have been developed in 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.

\textsuperscript{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 (Epstein and Miao, 2003; Routledge and Zin, 2001). Similarly, existing robust control models (Hansen, Sargent, and Tallarini, 1999; Cagetti, Hansen, Sargent, and Williams, 2002) do not allow the ‘concern for robustness’ to change in response to new observations. Neither is learning modeled in Uppal and Wang (2003) that pursues a third approach to accommodating ambiguity or robustness.”
Appendix: Proofs of all propositions

Proof of Proposition 1

The solution to the inner minimization problem is

\[ \mu_j = \hat{\mu}_j - \text{sign}(w_j)\sigma_j \sqrt{\epsilon_j} / \sqrt{T}, \quad j = 1, \ldots, N, \]

which, if substituted back into the original problem, gives

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

subject to \( w^\top 1_N = 1 \).

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

\[
\min_{\mu} w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w,
\]

subject to

\[
(\hat{\mu} - \mu)^\top \Sigma^{-1} (\hat{\mu} - \mu) \leq \epsilon.
\]

The Lagrangian is

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

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

\[
\min_{\mu} \max_{\lambda} \mathcal{L}(\mu, \lambda).
\]

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

\[
\mu^* = \hat{\mu} - \frac{1}{2\lambda} \Sigma w.
\]

Note that the objective function is not differentiable at \( w_j = 0, j = 1, \ldots, N \).
Substituting this in the Lagrangian (A4), we get
\[ L(\mu^*, \lambda) = w^\top \hat{\mu} - \left( \frac{1}{4\lambda} + \frac{\gamma}{2} \right) w^\top \Sigma w - \lambda \varepsilon. \]  
(A7)

Hence, the original max-min problem in (6) subject to (8) and (15) is equivalent to the following maximization problem
\[ \max_{w, \lambda} w^\top \hat{\mu} - \left( \frac{1}{4\lambda} + \frac{\gamma}{2} \right) w^\top \Sigma w - \lambda \varepsilon, \]  
subject to \( w^\top \mathbf{1}_N = 1 \). Solving for \( \lambda \) we obtain \( \lambda = \frac{1}{2} \sqrt{\frac{w^\top \Sigma w}{\varepsilon}} > 0 \) from which, upon substitution in (A8), we obtain (16).

The maximization in (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
\[ L(w, \lambda) = w^\top \hat{\mu} - \frac{\gamma}{2} w^\top \Omega(w) w + \lambda(1 - w^\top \mathbf{1}_N). \]

The first-order conditions with respect to \( w \) gives
\[ \hat{\mu} - \left( \frac{\sqrt{\varepsilon} + \gamma \sqrt{w^\top \Sigma w}}{\sqrt{w^\top \Sigma w}} \right) \Sigma^{-1} w - \lambda_1 \mathbf{1}_N = 0. \]

Let \( \sigma_P \equiv \sqrt{w^\top \Sigma w} \). From the last equation
\[ w = \left( \frac{\sigma_P}{\sqrt{\varepsilon} + \gamma \sigma_P} \right) \Sigma^{-1} (\hat{\mu} - \lambda \mathbf{1}_N). \]  
(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 (A9) we arrive at

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

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 \( \sigma_P^* \) of the following polynomial equation

\[
A_\gamma^2 \sigma_P^4 + 2A_\gamma \sqrt{\varepsilon} \sigma_P^3 + (A \varepsilon - AC + B^2 - \gamma^2) \sigma_P^2 - 2\gamma \sqrt{\varepsilon} \sigma_P - \varepsilon = 0, \tag{A11}
\]

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, since \( \Sigma \) is definite positive, the above polynomial equation always has at least one positive real root. Let \( \sigma_P^* \) be the unique positive real root of this equation.\(^{24}\) Then, the optimal portfolio \( w^* \) is given by

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

which simplifies to (17).

**Proof of Proposition 3**

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

\[
\max \min_{w, \mu_a, \mu_b} w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w, \tag{A13}
\]

subject to

\[
(\hat{\mu}_a - \mu_a)^\top \Sigma_a^{-1} (\hat{\mu}_a - \mu_a) \leq \epsilon_a, \tag{A14}
\]

\[
(\hat{\mu}_b - \mu_b)^\top \Sigma_b^{-1} (\hat{\mu}_b - \mu_b) \leq \epsilon_b. \tag{A15}
\]

\(^{24}\)It is possible to show that the fourth degree polynomial in (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 (A11) is negative at \( \sigma_P = 0 \) and has at least a negative local maximum, implying that the positive real root of (A11) is unique.
The Lagrangian of the inner minimization is
\[
L(\mu_a, \mu_b, \lambda_a, \lambda_b) = w_a^\top \mu_a + w_b^\top \mu_b - \frac{\gamma}{2} w^\top \Sigma w \\
- \lambda_a \left( \epsilon_a - (\hat{\mu}_a - \mu_a)^\top \Sigma_{aa}^{-1} (\hat{\mu}_a - \mu_a) \right) \\
- \lambda_b \left( \epsilon_b - (\hat{\mu}_b - \mu_b)^\top \Sigma_{bb}^{-1} (\hat{\mu}_b - \mu_b) \right).
\] (A16)

Solving for \(\mu_a\) and \(\mu_b\) in the inner minimization yields
\[
\mu_a^* = \hat{\mu}_a - \frac{1}{2\lambda_a} \Sigma_{aa} w_a, \quad \mu_b^* = \hat{\mu}_b - \frac{1}{2\lambda_b} \Sigma_{bb} \\
\lambda_a = \frac{1}{2} \sqrt{\frac{w_a^\top \Sigma_{aa} w_a}{\epsilon_a}}, \quad \lambda_b = \frac{1}{2} \sqrt{\frac{w_b^\top \Sigma_{bb} w_b}{\epsilon_b}},
\] (A17)

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

where
\[
\hat{\Sigma}_{aa}(w_a, \epsilon_a) = \left( 1 + \frac{2}{\gamma} \frac{\sqrt{\epsilon_a}}{\sqrt{w_a^\top \Sigma_{aa} w_a}} \right), \\
\hat{\Sigma}_{bb}(w_b, \epsilon_b) = \left( 1 + \frac{2}{\gamma} \frac{\sqrt{\epsilon_b}}{\sqrt{w_b^\top \Sigma_{bb} w_b}} \right).
\] (A20)

Let us define
\[
\sigma_a \equiv \sqrt{w_a^\top \Sigma_{aa} w_a}, \\
\sigma_b \equiv \sqrt{w_b^\top \Sigma_{bb} w_b}.
\] (A22)

Since \(\Sigma_{aa}\) and \(\Sigma_{bb}\) are positive definite, \(\sigma_a > 0\) unless \(w_a = 0_{M_a \times 1}\), and similarly, \(\sigma_b > 0\) unless \(w_b = 0_{M_b \times 1}\). Now the first-order conditions with respect to \(w_a\) and \(w_b\) yield
\[
\left( \frac{\sqrt{\epsilon_a} + \gamma \sigma_a}{\sigma_a} \right) \Sigma_{aa} w_a = \hat{\mu}_a - \gamma \Sigma_{ab} w_b, \\
\left( \frac{\sqrt{\epsilon_b} + \gamma \sigma_b}{\sigma_b} \right) \Sigma_{bb} w_b = \hat{\mu}_b - \gamma \Sigma_{ba} w_a.
\] (A24)

Given \(w_b\), we can solve for \(\sigma_a\) in (A24) and get
\[
\sigma_a = \frac{1}{\gamma} \max \left[ \sqrt{g(w_b)^\top \Sigma_{aa}^{-1} g(w_b)} - \sqrt{\epsilon_a}, 0 \right],
\] (A26)
where \( g(w_b) = \mu_a - \gamma \Sigma_{ab} w_b \). Substituting this back into (A24), we obtain (22) if \( m = a \).

Similarly, given \( w_a \), we can solve for \( \sigma_b \) in (A25) and get

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

(A27)

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

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

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

(A28)

Since \( \Upsilon \) is continuous and the “max” in (22) are bounded between zero and one, \( \Upsilon(\Gamma) \subseteq \Gamma \).

By the Brouwer Fixed-Point Theorem (see, for example, Stokey and Lucas (1989, p. 517)), \( \Upsilon \) has a fixed point in \( \Gamma \).
Table 1: Out-of-sample performance with only parameter uncertainty
This table reports the out-of-sample Mean, Standard Deviation and Mean-to-Standard Deviation ratio for the returns on different portfolio strategies, including the one from the ambiguity-averse model that allows for parameter uncertainty. Means and Standard Deviations are expressed as percentage per month. The data are obtained from MSCI (Morgan Stanley Capital International) and consist of monthly returns on eight international equity indices (Canada, France, Germany, Italy, Japan, Switzerland, United Kingdom, 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 $\epsilon$. The investor is assumed to have a risk aversion of $\gamma = 1$.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Mean/Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Short sales allowed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Variance</td>
<td>0.0049</td>
<td>0.2557</td>
<td>0.0190</td>
</tr>
<tr>
<td>Minimum-Variance</td>
<td>0.0118</td>
<td>0.0419</td>
<td>0.2827</td>
</tr>
<tr>
<td>Bayes-Stein</td>
<td>0.0071</td>
<td>0.1058</td>
<td>0.0671</td>
</tr>
<tr>
<td>Ambiguity-averse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\epsilon = 0.00$</td>
<td>0.0049</td>
<td>0.2557</td>
<td>0.0190</td>
</tr>
<tr>
<td>$\epsilon = 0.25$</td>
<td>0.0065</td>
<td>0.1307</td>
<td>0.0495</td>
</tr>
<tr>
<td>$\epsilon = 0.50$</td>
<td>0.0080</td>
<td>0.0953</td>
<td>0.0841</td>
</tr>
<tr>
<td>$\epsilon = 0.75$</td>
<td>0.0091</td>
<td>0.0765</td>
<td>0.1190</td>
</tr>
<tr>
<td>$\epsilon = 1.00$</td>
<td>0.0098</td>
<td>0.0654</td>
<td>0.1491</td>
</tr>
<tr>
<td>$\epsilon = 1.50$</td>
<td>0.0104</td>
<td>0.0545</td>
<td>0.1909</td>
</tr>
<tr>
<td>$\epsilon = 2.00$</td>
<td>0.0107</td>
<td>0.0501</td>
<td>0.2144</td>
</tr>
<tr>
<td>$\epsilon = 2.50$</td>
<td>0.0109</td>
<td>0.0480</td>
<td>0.2281</td>
</tr>
<tr>
<td>$\epsilon = 3.00$</td>
<td>0.0111</td>
<td>0.0467</td>
<td>0.2369</td>
</tr>
<tr>
<td>$\epsilon \rightarrow \infty$</td>
<td>0.0117</td>
<td>0.0419</td>
<td>0.2827</td>
</tr>
<tr>
<td><strong>Panel B: Short sales not allowed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Variance</td>
<td>0.0104</td>
<td>0.0587</td>
<td>0.1774</td>
</tr>
<tr>
<td>Minimum-Variance</td>
<td>0.0117</td>
<td>0.0412</td>
<td>0.2831</td>
</tr>
<tr>
<td>Bayes-Stein</td>
<td>0.0106</td>
<td>0.0511</td>
<td>0.2074</td>
</tr>
<tr>
<td>Ambiguity-averse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\epsilon = 0.00$</td>
<td>0.0104</td>
<td>0.0587</td>
<td>0.1774</td>
</tr>
<tr>
<td>$\epsilon = 0.25$</td>
<td>0.0113</td>
<td>0.0511</td>
<td>0.2214</td>
</tr>
<tr>
<td>$\epsilon = 0.50$</td>
<td>0.0115</td>
<td>0.0482</td>
<td>0.2391</td>
</tr>
<tr>
<td>$\epsilon = 0.75$</td>
<td>0.0115</td>
<td>0.0467</td>
<td>0.2472</td>
</tr>
<tr>
<td>$\epsilon = 1.00$</td>
<td>0.0116</td>
<td>0.0457</td>
<td>0.2533</td>
</tr>
<tr>
<td>$\epsilon = 1.50$</td>
<td>0.0116</td>
<td>0.0446</td>
<td>0.2607</td>
</tr>
<tr>
<td>$\epsilon = 2.00$</td>
<td>0.0117</td>
<td>0.0440</td>
<td>0.2647</td>
</tr>
<tr>
<td>$\epsilon = 2.50$</td>
<td>0.0117</td>
<td>0.0436</td>
<td>0.2671</td>
</tr>
<tr>
<td>$\epsilon = 3.00$</td>
<td>0.0117</td>
<td>0.0433</td>
<td>0.2695</td>
</tr>
<tr>
<td>$\epsilon \rightarrow \infty$</td>
<td>0.0117</td>
<td>0.0412</td>
<td>0.2831</td>
</tr>
</tbody>
</table>
Table 2: Out-of-sample performance with parameter and model uncertainty

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 MSCI (Morgan Stanley Capital International) and consist of monthly excess returns on eight international equity indices (Canada, France, Germany, Italy, Japan, Switzerland, United Kingdom, United States) in addition to the world market portfolio. Excess return are obtained by subtracting the month-end return on the United States 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 $\epsilon_a$ 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 $\epsilon_b$. The Sharpe ratio for the minimum-variance portfolio, which is not nested by any of the models considered in this table, is 0.1490.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Sharpe ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-Variance</td>
<td>-0.0719</td>
</tr>
<tr>
<td>Bayes-Stein</td>
<td>-0.0528</td>
</tr>
<tr>
<td>Bayesian Data-and-Model with $\omega = 0$</td>
<td>-0.0853</td>
</tr>
<tr>
<td>with $\omega = 1$</td>
<td>0.1239</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ambiguity-averse</th>
<th>$\epsilon_a$%</th>
<th>$\epsilon_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>with $\omega = 0$</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>0.00 (0.00)</td>
<td>-0.0853</td>
<td>0.1219</td>
</tr>
<tr>
<td>0.25 (2.01)</td>
<td>-0.0774</td>
<td>0.1032</td>
</tr>
<tr>
<td>0.50 (14.60)</td>
<td>-0.0475</td>
<td>0.0824</td>
</tr>
<tr>
<td>0.75 (35.27)</td>
<td>-0.0113</td>
<td>0.0604</td>
</tr>
<tr>
<td>1.00 (55.98)</td>
<td>0.0930</td>
<td>0.0655</td>
</tr>
<tr>
<td>1.50 (83.49)</td>
<td>0.1219</td>
<td>0.1218</td>
</tr>
<tr>
<td>2.00 (94.73)</td>
<td>0.1239</td>
<td>0.1252</td>
</tr>
<tr>
<td>2.50 (98.45)</td>
<td>0.1239</td>
<td>0.1275</td>
</tr>
<tr>
<td>3.00 (99.56)</td>
<td>0.1239</td>
<td>0.1284</td>
</tr>
</tbody>
</table>

| with $\omega = 1$              | (0.00)         | (0.00)       |
| 0.00 (0.00)                     | 0.1239         | 0.1202       |
| 0.25 (2.01)                     | 0.1239         | 0.1284       |
| 0.50 (14.60)                    | 0.1239         | 0.1284       |
| 0.75 (35.27)                    | 0.1239         | 0.1284       |
| 1.00 (55.98)                    | 0.1239         | 0.1284       |
| 1.50 (83.49)                    | 0.1239         | 0.1284       |
| 2.00 (94.73)                    | 0.1239         | 0.1284       |
| 2.50 (98.45)                    | 0.1239         | 0.1284       |
| 3.00 (99.56)                    | 0.1239         | 0.1284       |
Figure 1: Shrinkage factors $\phi_{AA}(\epsilon)$ and $\phi_{BS}$ 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 shrinkage approach, as explained in Section 3.3. 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, $\phi_{BS}$ 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.
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 ($\epsilon = 1$) which corresponds to uncertainty expressed roughly by the 56% confidence interval for an $F_{8,112}$ centered around the sample mean, while 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.

Panel A: Shortselling allowed

Panel B: Shortselling not allowed
References


