Model Misspecification and Under-Diversification*

Raman Uppal† and Tan Wang‡

January 2002

*We are extremely grateful for detailed discussions with Bernard Dumas and Larry Epstein. We would also like to acknowledge comments and suggestions from Suleyman Basak, Murray Carlson, Adlai Fisher, Francisco Gomes, Tim Johnson, Burton Hollifield, Kai Li, Pascal Maenhout, Jun Pan, Sinan Sarpca, Stephen Schaefer, Luis Viceira, Bill Ziemba and seminar participants at London Business School, MIT Sloan School of Management, University of Amsterdam, University of British Columbia, University of Warwick and the 2002 meetings of the AFA.

†London Business School, IFA, 6 Sussex Place Regent's Park, London, England NW1 4SA; Email: ruppal@london.edu.

‡University of British Columbia, 2053 Main Mall, Vancouver Canada V6T 1Z2, Canada; Email: tan.wang@commerce.ubc.ca.

Model Misspecification and Under-Diversification

Abstract

In this paper we develop a model of intertemporal portfolio choice where an investor accounts explicitly for the possibility of model misspecification. This work is motivated by the difficulty in estimating precisely the probability law for asset returns. Our contribution is to develop a framework that allows for ambiguity about the joint distribution of returns for all stocks being considered for the portfolio, and also for different levels of ambiguity for the marginal distribution of returns for *any* subset of these stocks. We then use this framework to derive in closed-form the optimal portfolio weights of an investor who accounts for model misspecification. We illustrate the model by calibrating it to data on international equity returns. The calibration shows that when the overall ambiguity about the joint distribution of returns is high, then small differences in ambiguity for the marginal return distribution will result in a portfolio that is significantly under-diversified relative to the standard mean-variance portfolio.

Keywords: Portfolio choice, uncertainty, ambiguity, robust control
JEL Classification: G11, D81

1 Introduction

Traditional rational expectations models of portfolio choice assume that an investor knows exactly the true probability law for the stochastic processes of asset returns. We consider a setting where an agent formulates a *reference model* of the probability law based on the data available, but recognizes that it is only an approximation to the true model and hence subject to misspecification.[1] The main contribution of our work is to develop a framework that allows for multiple state variables whose (marginal) probability laws are known to the investor with possibly different levels of ambiguity. We then use this framework to show how an investor would choose a portfolio that is optimal in the presence of model misspecification.

Several approaches have been developed in the literature to account for imprecise knowledge about the probability law. In one approach, the parameter uncertainty or estimation risk is modeled by a prior belief and a Bayesian methodology is adopted. A limitation of the Bayesian approach is that it makes the strong assumption that the prior belief can be described by a probability measure. However, this assumption is inconsistent with the evidence from experimental economics and psychology such as the Ellsberg paradox (Ellsberg (1961)), where uncertainty arises from two sources: uncertainty about the states of the world and uncertainty about the model itself. While there is agreement that the states of the world can be described by an objective probability law, evidence from the Ellsberg-type experiment raises the question whether model uncertainty can be described by a (subjective) probability prior.

Two classes of models have been developed in the literature as alternatives to the Bayesian approach. In one class, Dow and Werlang (1992) study the portfolio choice problem of an investor under Knightian uncertainty in a static setting using Choquet expected utility. Epstein and Wang (1994) in dynamic discrete time, and Chen and Epstein (2000) in continuous time, extend the Lucas (1978) model to incorporate the effect of Knightian uncertainty by allowing for multiple priors. An application of this approach to international portfolio choice in an equilibrium setting is presented in Epstein and Miao (2000). Epstein and Miao is closest to our paper in that it

[1] For instance, Merton (1980) discusses the difficulty in estimating the expected return of an asset; French and Poterba (1991) report that the standard error of the estimated mean annual return on the US stock market (based on 60 years of data) is 200 basis points, and Gorman and Jorgensen (1999) report similar evidence for several non-US equity markets.

addresses both the model uncertainty issue and the issue of different levels of ambiguity about the (marginal) probability laws governing the states. One feature of their minimization problem is that all priors in the set are treated indiscriminately, other things being equal. Consequently, agents exhibit extreme pessimism with respect to the given subset of priors and always pick the worst-case scenario. In contrast, while agents under our approach also have multiple priors, they do not exhibit extreme pessimism; instead, they use the *reference model* to differentiate among the priors. This is an important conceptual difference: knowledge of the data and the economic environment, although not perfect (otherwise there would be no model uncertainty), is used by economic agents in discriminating among candidate priors for the true model of the economy. This difference also leads to a formulation that has the differentiability needed for deriving the Bellman equation. Consequently, our characterization of the optimal portfolio is a transparent extension of the standard Merton (1971) portfolio model without ambiguity.

In the second class of models,[2] Hansen, Sargent, and Tallarini (1999), Anderson, Hansen, and Sargent (1999), and Hansen and Sargent (2001a) introduce model misspecification and preference for "robustness" into the Lucas model.[3] In their model, agents recognize the possibility of model misspecification, and account for it in their decisions. Maenhout (1999) applies this framework to study portfolio choice between a riskless and a single risky asset. In Anderson, Hansen, and Sargent, and Maenhout, model uncertainty is described by a *single* parameter that reflects the overall level of ambiguity. In contrast, we allow for *differences* in the degree of ambiguity about the (marginal) probability laws for the various elements of the state vector process, and in the portfolio context, for the returns of different assets.[4] Our formulation is sufficiently general to incorporate ambiguity about the joint distribution of returns for all stocks being considered, and different levels of ambiguity also for the return distribution of *any* subset of these stocks, with the subsets possibly overlapping.

[2]There is an on-going discussion about the exact relation between these two classes of models. The reader is referred to Epstein and Schneider (2001) and Hansen and Sargent (2001b,c) for this discussion.

[3]Hansen and Sargent (2000) provides an extensive discussion of the relation of the robust decision-making approach to Bayesian models, adaptive models, and models with filtering.

[4]As a by-product, we show that once one allows for differences in the level of ambiguity across assets, the investor's preferences are no longer observationally equivalent to recursive utility (Epstein and Zin (1989); Duffie and Epstein (1992)). Hence, the observational equivalence result in Anderson, Hansen, and Sargent (1999) and Maenhout (1999) holds only under the extreme case where the level of ambiguity is the same for all assets.

The rest of this paper is organized as follows. In Section 2, we develop a utility function for an agent who recognizes the possibility of model misspecification. In Section 3, we apply this utility function to study the problem of portfolio selection when there are multiple risky assets, and analyze some special cases that convey the intuition underlying our framework and its implications for portfolio selection. In Section 4, we calibrate the model to data on international equity returns to illustrate how one would apply our model, and also to show how one can gauge whether the parameters determining the level of ambiguity are reasonable. We conclude in Section 5. Proofs are presented in the appendix.

2 Preferences in the presence of model misspecification

In the first part of this section, we explain how the standard time-separable preferences have been extended in the recent literature to allow for decision making in the presence of model misspecification. Our main contribution to the existing literature is in the second part, where we extend this basic framework to allow for differences in the degree of ambiguity about the various elements of the state vector process. While we will be using a model set in continuous time, we start by motivating the analysis in discrete time.

2.1 The basic model with a single source of misspecification

In the standard model of portfolio choice and asset pricing, the investor is typically assumed to have intertemporally additive expected utility of the form:

$$V_t = u(c_t) + \beta E[V_{t+1}]. \tag{1}$$

A fundamental assumption behind this model of investor preferences is that the investor knows precisely the true probability law of asset returns when computing the expectation in the equation above. It has been argued in the literature that this assumption is too strong and that agents should be allowed to account for model misspecification in their decision process.[5]

[5] For references to this literature, see Epstein and Wang (1994), Chen and Epstein (2000), Hansen, Sargent, and Tallarini (1999), Anderson, Hansen, and Sargent (1999), and Maenhout (1999).

Our approach, where we formulate a recursive utility as in Epstein and Zin (1989) and use relative entropy for the certainty equivalent, can be described intuitively as follows. Let the knowledge of the investor about the uncertainty in the economy be described by a model of the relevant economic variables such as stock prices, the interest rate, etc. Since what matters to an investor is the distribution of the relevant economic variables, this model can equivalently be described by a probability measure P, called the reference probability or *reference model*. It is often the case that P is the result of some estimation process. If P could be estimated perfectly, there would be no ambiguity about it representing the true economic model. However, due to insufficient data, the estimation process is not capable of distinguishing all different models. Thus, P is subject to misspecification error. Because the investor is not sure if P is the right model, it is natural that he would consider alternative models. Let a possible alternative to the reference model P be described by a probability measure Q^{ξ} given by

$$dQ^{\xi} = \xi(X_{t+1})dP, \tag{2}$$

where X_t is the state variable and $\xi(x)$ is a density function. Of course, there can be many possible alternatives; knowing that the reference measure P is subject to misspecification error and that the possible alternatives are Q^{ξ}, the investor's problem is how to take into account the possible alternatives when making his decisions.

To evaluate the alternative models, the investor needs an index that tells him, given his information, how each alternative compares with the reference model. There can be many useful indices, one of which is the relative entropy index, $\phi L(\xi)$, given by

$$\phi L(\xi) = \phi E^{\xi} [\ln \xi] , \tag{3}$$

where E^{ξ} is the expectation under Q^{ξ} and $\phi \geq 0$ is a parameter whose role is explained below, after we provide an intuitive explanation for this index.

One interpretation of (3) is that it is an approximation to the empirical likelihood ratio adjusted for the level of ambiguity.[6] To elaborate, suppose that the data set available to the investor has T

[6]See Anderson, Hansen, and Sargent (1999) and Hansen and Sargent (2001a) for other interpretations of the index $L(\xi)$.

observations. For the alternative model, the log-likelihood of the data is

$$\frac{1}{T}\sum_{t=1}^{T}\ln \xi(X_t)f(X_t),$$

where $f(x)$ is the density function of the reference model P and $\xi(x)$ is the density function of the alternative model with respect to P. Thus the empirical log-likelihood ratio of the two models is

$$\frac{1}{T}\sum_{t=1}^{T}\ln \xi(X_t).$$

According to traditional likelihood ratio tests, if the above sum is large then the two models, Q^{ξ} and P, can be clearly distinguished. By the Law of Large Numbers, under the alternative model $\frac{1}{T}\sum_{t=1}^{T}\ln \xi(X_t)$ converges to $L(\xi) = E^{\xi}[\ln \xi(x)]$. Thus, if Q^{ξ} is the true probability law, $L(\xi)$ is a good approximation to the empirical log-likelihood when T is large. Given that in reality, data is limited and T is finite, ambiguity arises as to how close $L(\xi)$ is to the empirical log-likelihood. For instance, the same value of the empirical log-likelihood obtained from a longer data series is more precise than that from a shorter data series.[7]

More generally, the ambiguity can be caused by any factor that affects the evaluation of the true log-likelihood. The parameter ϕ is introduced to capture this ambiguity and a lower ϕ corresponds to a higher level of ambiguity. For example, a shorter data series is associated with a smaller ϕ, while a longer data series is associated with a larger ϕ', so that $\phi L(\xi)$ and $\phi' L(\xi)$ represent the empirical log-likelihood ratios adjusted for the level of ambiguity.[8]

To incorporate information about model misspecification into his decision process, the investor needs a function that combines this information with his future state-contingent utility. This is done through a certainty equivalent function, CE$(\cdot)$, such that the utility of the investor is given

[7] See Stambaugh (1997) for a discussion of portfolio choice using a sample where not all the data series are of equal length.

[8] It is worth emphasizing that large $\frac{1}{T}\sum_{t=1}^{T}\ln \xi(X_t)$ or $\phi L(\xi)$ should not be interpreted as evidence for *rejecting* the reference model P, as in the usual likelihood test: as explained in the introduction, the very fact that P is the reference model implies that the agent has already gone through the preliminary analysis and picked P over other models. The issue at this stage is only to find an index that summarizes the information available.

by[9]

$$V_t = u(c_t) + \beta\, \mathrm{CE}(V_{t+1}; \phi L). \tag{4}$$

The interpretation of (4) is that $\phi L(\xi)$ summarizes the investor's information about model misspecification, and the certainty equivalent, CE, captures how the investor uses this information to evaluate his future utility V_{t+1}.

One example of CE is relative entropy,

$$\inf_{\xi}[\psi(V_t)\phi L(\xi) + E_t^{\xi}[V_{t+1}]], \tag{5}$$

which, when substituted in (4), gives[10]

$$V_t = u(c_t) + \beta \inf_{\xi}[\psi(V_t)\phi L(\xi) + E_t^{\xi}[V_{t+1}]], \tag{6}$$

where $\psi(V_t)$ in (5) and (6) is a normalization factor that is introduced to convert the penalty to units of utility so that it is consistent with the units of $E^{\xi}[V_{t+1}]$; the particular functional form of $\psi(\cdot)$ is often chosen for analytical convenience. The minimization over ξ in (6) reflects the agent's aversion to ambiguity/model-misspecification.

The role of ϕ, as reflecting the ambiguity of the investor about the quality of the information/data, can be seen from (6). The investor ponders whether he should use model Q^{ξ} to evaluate his future utility. The term $\phi L(\xi)$ is used as a penalty function for rejecting the reference model P and accepting the alternative model Q^{ξ}. However, if the available data allows one to easily distinguish an alternative model Q^{ξ} from the reference model P, then accepting Q^{ξ} will incur a penalty. The magnitude of the penalty depends on the level of ambiguity in the reference model P. In the extreme case where $\phi \approx \infty$, i.e., the investor is extremely confident about P, any alternative model Q^{ξ} that deviates from the reference model will be penalized heavily. In this case, equation (6) reduces to the standard expected utility in equation (1). Thus the standard expected utility can be viewed as a special case of (6) where the investor knows the true model—rational expectations—and hence has no ambiguity about the reference model. On the other hand, for models that the

[9] The certainty equivalent $\mathrm{CE}(\cdot)$ has the following two properties: (a) $\mathrm{CE}(V) = V$ if V is a constant, that is, model misspecification is irrelevant when there is no uncertainty; and (b) $\mathrm{CE}(V) \geq \mathrm{CE}(V')$ if $V \geq V'$ almost surely, that is, if V first-order stochastically dominates V', model misspecification should not change the agent's preference for V over V'.

[10] An axiomatic foundation for these kind of preferences is available from the authors.

investor cannot clearly distinguish, considering them will result in only a small penalty. Among these models, due to his concern for model misspecification, the investor uses the one that gives the lowest expected utility. For the extreme case where $\phi \approx 0$, i.e., the investor has no knowledge about P, and equation (6) reduces to

$$V_t = u(c_t) + \beta \inf V_{t+1}.$$

In this case, the investor will consider the worst-case scenario as the only possible outcome.

In general, the investor balances his concern about model misspecification and the knowledge he has about the economy as represented by P. He does not wish to throw away information by setting $\phi \approx 0$ and only guarding against model misspecification, nor does he want to ignore his ambiguity about the information by setting $\phi \approx \infty$ and overlooking completely the possibility of model misspecification.

2.2 Extension: Different levels of ambiguity for each state variable

While the basic model in the previous section captures the investor's concern for model misspecification, it does not allow for different levels of ambiguity for different components of the state process. In this section, we extend the basic model to allow for such differences in ambiguity, which distinguishes our work from that of Anderson, Hansen, and Sargent (1999) and Maenhout (1999). The basic intuition underlying this extended model is the same as that elucidated in the preceding section for the model with a single state variable; the main change is the development of an appropriate penalty function.

2.2.1 Discrete time

Suppose that uncertainty is generated by more than a single state variable. Imagine an investor whose knowledge about the probability law for the state variables is limited, and this information comes from separate sources and the investor is more confident about some sources relative to others. For instance, in a universe with only two countries, each having one large firm and one small firm, we would like to allow for knowledge about the joint distribution of returns for all four stocks from an analyst who covers a broad spectrum of stocks, and also additional information from analysts specializing in a subset of these four stocks: the set consisting of only foreign stocks, only domestic stocks, only large stocks, only small stocks, and each of the individual stocks. We would like to develop a framework that is sufficiently general to allow for different levels of ambiguity for information from different sources and about different subsets of assets. In order to have a model capable of reflecting this feature of the investor's information, we extend the basic model described in the previous section by first generalizing the relative entropy index in (3) and then incorporating this more general index into the utility function in (6).

Let $X_t = (X_{1t}, \ldots, X_{nt})$ be the vector of all state variables. Let Q^ξ represents an alternative model as in the previous section, with

$$dQ^\xi = \xi dP,$$

where ξ is a scalar that perturbs P, the joint distribution of all the state variables. Let $J_i = \{j_1, \ldots, j_{n_i}\}$ be a subset of $\{1, \ldots, n\}$, and let $X_{J_i} = (X_{j_1}, \ldots, X_{j_{n_i}})$ be the corresponding sub-vector of X_t. Suppose that the investor has a separate source of information about the subset of state variables, X_{J_i}. Then, as in (3), we can use an index to describe this information. However, because the information is about the subset of state variables, the index is now calculated with respect to the marginal distribution of X_{J_i}:

$$\phi_i L(\xi_i) = \phi_i \int \left[\xi_i(X_{J_i,t+1}) \ln \xi_i(X_{J_i,t+1}) \right] \, dP_{J_i},$$

where P_{J_i} is the marginal distribution of the sub-vector X_{J_i} under the reference probability measure P, and $\xi_i = dQ^\xi_{J_i}/dP_{J_i}$. If there are K sources of information for the various subsets of state variables, then the overall index is taken to be the sum,

$$\sum_{i=1}^{K} \phi_i \, L(\xi_i). \tag{7}$$

The investor's utility function is now given by the following recursive equation, which is similar to (6), but with the index (7) that allows for multiple sources of information about the vector of state variables:

$$V_t = u(c_t) + \beta \inf_{\xi} \left\{ \psi(V_t) \sum_{i=1}^{K} \phi_i L(\xi_i) + E^\xi [V_{t+1}] \right\}, \tag{8}$$

where, as before, $\psi(V)$ is a normalization factor. The interpretation of (8) is essentially the same as in the previous section. The only difference is that if one source of information about a particular subset of the state variables is more reliable, the investor will assign a higher penalty for deviating from that information. For instance, if the investor has very reliable information about the return of a particular stock, he will put a high penalty for any alternative model whose marginal distribution for the return on this stock deviates from that of the reference model.

2.2.2 Continuous time

In this section, we extend the utility function formulated in (8) to continuous time. Suppose that the state variables $X_t = (X_{1t}, \dots, X_{nt})$ follow the process

$$dX_t = \mu_X(X_t, t)dt + \sigma_X(X_t, t)dw_t,$$

where w_t is a n-dimensional Brownian motion. Let

$$\mathcal{A}(f) = f_t + \mu_X f_X + \frac{1}{2}\text{tr}\left(f_{XX}\sigma_X\sigma_X^\top\right)$$

be the differential operator associated with the diffusion process X_t. Denote by $\left[\sigma_{J_i}\sigma_{J_i}^\top\right]_n$ the $n \times n$-matrix whose element in the j_k-th row and j_ℓ-th column, for j_k and j_ℓ in J_i, is equal to the element in the kth row and ℓth column of the matrix $[\sigma_{X_{J_i}}\sigma_{X_{J_i}}^\top]^{-1}$, which is the inverse of the variance-covariance matrix of X_{J_i}; otherwise it is zero.[11]

Theorem 1 *The continuous-time version of (8) is*

$$0 = \inf_v \left\{ u(c) - \rho V + \mathcal{A}(V) + v^\top V_X + \frac{\psi(V)}{2} v^\top \Phi v \right\}, \tag{9}$$

where $v = (v_1, \dots, v_n)^\top$ *and*

$$\Phi = \sum_i \phi_i \left[\sigma_{J_i}\sigma_{J_i}^\top\right]_n. \tag{10}$$

In equation (9), the first three terms correspond to the standard Hamilton-Jacobi-Bellman equation for the expected utility function under the reference probability P. The second-last term

[11]Specific examples of $\left[\sigma_{J_i}\sigma_{J_i}^\top\right]_n$ can be seen in equations (20) and (25).

arises from the change of probability measure from P to Q^ξ in (8). By Girsanov's Theorem, the change of probability measure is equivalent to a change in the drift term of the process of X_t. The drift change is given by $v = (v_1, \ldots, v_n)$. That is, under the probability Q^ξ, the process for X_t is

$$dX_t = [\mu_X(X_t, t) + v_t]\, dt + \sigma_X(X_t, t) dw_t^\xi,$$

where w_t^ξ is a Brownian motion under Q^ξ. Observe that the effect of the change from the reference model P to the alternative model Q^ξ is completely captured by this term, which will be useful for understanding the results in the portfolio choice problem that we will consider in the next two sections. The last term in (9) corresponds to the penalty function in (8). The fact that the utility function of the agent can be characterized by the Hamilton-Jacobi-Bellman equation (9) indicates that our formulation of the agent's preference is dynamically consistent.[12]

We conclude this section with the following remarks on the comparison between the Bayesian approach and our approach to model misspecification. In the Bayesian approach, model misspecification often comes in the form of parameter uncertainty. To be specific, suppose that a model of the probability law for asset returns is estimated in which a parameter cannot be estimated precisely. Let $P(X;\alpha)$ be the probability distribution function and α be the parameter about which one is uncertain. Given that the parameter α is unknown, the question for the investor is how to incorporate the parameter uncertainty into his decision process. The critical assumption of the Bayesian approach is that this parameter-uncertainty/model-misspecification can be represented by a prior distribution F, and that the investor's utility can be computed by[13]

$$V_t = u(c_t) + \beta E_F[E_{P(\alpha)}[V_{t+1}]].$$

In terms of the certainty equivalent, the equation can be written as

$$V_t = u(c_t) + \beta\, \mathrm{CE}(V_{t+1}; F), \qquad \text{where} \quad \mathrm{CE}(V_{t+1}; F) = E_F[E_{P(\alpha)}[V_{t+1}]].$$

In contrast, under our framework one need not restrict model misspecification to uncertainty regarding a particular parameter. More importantly, we do not assume that model misspecification, as a subjective matter, can be represented by a probability distribution. This difference between

[12]This implies that the criticism in Epstein and Schneider (2001) about the lack of dynamic consistency of the Hansen-Sargent formulation does not apply to our model.

[13]If learning/updating is to be incorporated, then F is the posterior.

the Bayesian approach and ours is exactly the same as that between the Savagian and Knightian approaches to decision making under uncertainty. For a more extensive discussion of this difference, see Ellsberg (1961).

3 Portfolio selection with multiple risky assets

In this section, we study the portfolio choice problem of an investor who is concerned about model misspecification. The portfolio choice model we use is standard (Merton, 1971, 1973) except for the preferences of the investor, which are the ones developed in the previous section.

3.1 Individual investor's portfolio choice

The investor can consume a single good, can invest in N risky stocks, and can also borrow and lend at an exogenously given riskless rate r_t. We use c to denote the consumption rate of the investor, W the wealth of the investor, and π_j the share of the investor's wealth invested in the j-th risky asset.

The return processes of the N stocks are given by

$$dR_t \equiv \mu_R(R_t, Y_t)dt + \sigma_R(R_t, Y_t)dw_t, \quad (11)$$

$$dY_t = \mu_Y(Y_t)dt + \sigma_Y(Y_t)dw_t. \quad (12)$$

These processes are viewed as the reference model. We assume that Y_t is a K-dimensional process and that the Brownian motion is $(N + K)$-dimensional.

The dynamics of the investor's wealth, for a given investment decision π and a consumption decision c, is:

$$dW_t = W_t \left[r_t + \pi_t (\mu_R - r_t) - \frac{c_t}{W_t} \right] dt + W_t \pi_t \sigma_R dw_t. \quad (13)$$

The investor wishes to maximize his intertemporal lifetime utility

$$E\left[\int_0^T e^{-\rho t} u(c_t) dt\right],$$

subject to the budget equation (13) while taking into account model misspecification when making his decisions.

To use Theorem 1 appropriately for deriving the Bellman equation corresponding to the investor's utility maximization problem, we need to distinguish between exogenous and endogenous state variables. As we know from Merton (1971), the investor's knowledge about the investment opportunities is described by the reference model given by (11) and (12). Thus, the state variables for the problem without model misspecification are R_t, Y_t and the investor's wealth W_t. However, the investor's wealth process (13), is derived from the stock returns. This can be seen by expressing the evolution of wealth in terms of stock returns:

$$dW_t = W_t(1 - \pi_t \mathbf{1}) r_t dt - c_t dt + W_t \pi_t dR_t. \tag{14}$$

Thus, W_t itself is not a source of the investor's concern of model misspecification. It only inherits the model misspecification through R_t. Because of this difference, R_t and Y_t are called *exogenous* state variables, while W_t is called an *endogenous* state variable. These two set of state variables need to be treated differently: in particular, the drift adjustment for W is $v_W = W_t \pi_t v_R$, because the amount of wealth invested in the risky asset is $W_t \pi_t$ and, under probability Q^{ξ}, the drift adjustment for R is v_R.

We write the investor's indirect utility function as $V(W_t, R_t, Y_t, t)$. Applying Theorem 1, and using the appropriate drift adjustment for W as discussed above, the Hamilton-Jacobi-Bellman equation for the investor's utility maximization problem is:

$$\begin{aligned} 0 \;=\; \sup_{c,\pi} \inf_{v_Y, v_R} \Bigg\{ & u(c) - \rho V + V_t + W V_W \left[r + \pi \left(\mu_R - r \right) - \frac{c}{W} \right] + \frac{W^2}{2} V_{WW} \pi^{\top} \sigma_R \sigma_R^{\top} \pi + V_R \mu_R \\ & + V_Y \mu_Y + \frac{1}{2} \mathrm{tr} \left[\begin{pmatrix} V_{RR} & V_{RY} \\ V_{YR} & V_{YY} \end{pmatrix} \begin{pmatrix} \sigma_R \\ \sigma_Y \end{pmatrix} \begin{pmatrix} \sigma_R \\ \sigma_Y \end{pmatrix}^{\top} \right] + W \pi \sigma_R \sigma_Y^{\top} V_{WY} + W \pi \sigma_R \sigma_R^{\top} V_{WR} \\ & + V_W W \pi v_R + V_Y v_Y + V_R v_R + \frac{\psi(V)}{2} v^{\top} \Phi v \Bigg\}, \end{aligned} \tag{15}$$

where Φ is defined by the expression (10). The terms in the first two lines of this equation are the same as the ones that would appear in the standard Bellman equation. The next three terms,

$V_W W \pi v_R + V_Y v_Y + V_R v_R$, reflect the drift adjustment due to the change of probability measure, while the last term, $(1/2)\psi(V)v^\top \Phi v$, is the penalty function.

Let I_R to be the identity matrix of the same dimension as R, I the identity matrix of the dimension of R plus that of Y, and π^{Merton} the optimal portfolio when there is no model misspecification. Then, substituting the first-order condition for the minimization problem in (15) into the first-order conditions for the maximization problem, and solving for π gives the following result.

Theorem 2 *The optimal portfolio of an investor is given by*

$$\pi = -\frac{1}{WV_{WW}}\left[\sigma_R\sigma_R^\top\right]^{-1}\left[V_W(\mu_R - r + v_R^*) + \sigma_R\sigma_Y^\top V_{WY} + \sigma_R\sigma_R^\top V_{WR}\right] \tag{16}$$

where

$$\left[\begin{array}{c} v_R^* \\ v_Y^* \end{array}\right] = -\frac{1}{\psi(V)}\Phi^{-1}\left[\begin{array}{c} V_W W\pi_t + V_R \\ V_Y \end{array}\right].$$

Or, in closed-form,

$$\begin{aligned}\pi &= \frac{-1}{WV_{WW}}B\left[\sigma_R\sigma_R^\top\right]^{-1}\left[V_W(\mu_R - r) + \sigma_R\sigma_Y^\top V_{WY} + \sigma_R\sigma_R^\top V_{WR}\right] \\ &\quad + \frac{V_W}{WV_{WW}}B\left[\sigma_R\sigma_R^\top\right]^{-1}\left[\begin{array}{cc} I_R & 0 \\ 0 & 0 \end{array}\right]\Phi^{-1}\left[\begin{array}{c} V_R/\psi(V) \\ V_Y/\psi(V) \end{array}\right] \\ &= B\pi^{\text{Merton}} + \frac{V_W}{WV_{WW}}B\left[\sigma_R\sigma_R^\top\right]^{-1}\left[\begin{array}{cc} I_R & 0 \\ 0 & 0 \end{array}\right]\Phi^{-1}\left[\begin{array}{c} V_R/\psi(V) \\ V_Y/\psi(V) \end{array}\right], \end{aligned} \tag{17}$$

with

$$B = \left(I - \frac{(V_W)^2}{\psi(V)V_{WW}}\left[\sigma_R\sigma_R^\top\right]^{-1}\left[\begin{array}{cc} I_R & 0 \\ 0 & 0 \end{array}\right]\Phi^{-1}\left[\begin{array}{c} I_R \\ 0 \end{array}\right]\right)^{-1}. \tag{18}$$

Without the term v_R^*, equation (16) reduces to the standard Merton formula. As noted above, this term corresponds to the drift adjustment to the wealth process due to R. Hence, (16) can be viewed as the Merton formula with μ_R adjusted by v_R^*. Similarly, when the ϕs tend to infinity, $\Phi^{-1} \to 0$, and equation (17) reduces to the familiar Merton (1971) result.

3.2 Understanding the portfolio model

To gain some insight into the expression for the portfolio weight in (16), we consider an economy where the stock price processes are given by geometric Brownian motions, the riskless rate is constant and $r < \mu_R$, the investor has power utility of the form $U(c) = c^{1-\gamma}/(1-\gamma)$ and is long lived ($T = \infty$). Following Maenhout (1999), we also set $\psi(V_t) = \frac{1-\gamma}{\gamma} V$. Under these assumptions, $V(W) = \kappa_0 W^{1-\gamma}/(1-\gamma)$, where κ_0 is a constant that depends on the parameters of the economy. We first look at the case where there is a single risky asset (this is the case considered in Maenhout (1999)) and then the case where there are two risky assets.

3.2.1 Example with one risky asset

In the case where there is only one risky asset, the explicit expression for the optimal portfolio simplifies to

$$\pi = \left(\frac{\phi}{1+\phi}\right) \underbrace{\frac{1}{\gamma}\frac{\mu_R - r}{\sigma_R^2}}_{\text{Merton weight}} . \tag{19}$$

The implications of model misspecification for portfolio choice can be observed from equation (19). For the case of $\phi = \infty$, the expression for the portfolio is simply the optimal Merton weight. However, for values of $\phi < \infty$ the investment in the risky asset is less than what it would be in the absence of model misspecification. In the limit, as ϕ approaches zero, investment in the risky asset drops to zero and the investor holds only the riskless asset. Thus, in the context of the portfolio choice problem, the consequence of model misspecification is a reduction in the investment in the asset about whose process the investor is ambiguous. However, the adjustment is limited by the penalty for being too far away from the reference model; this concern keeps the investor from choosing the most pessimistic scenario.

Also observe that the portfolio weight is exactly the same as that in the Merton model when the investor's risk aversion is given by $\gamma(1 + 1/\phi)$. In other words, observationally the ambiguity parameter ϕ is not separable from risk aversion. This is the observational-equivalence result noted in Anderson, Hansen, and Sargent (1999) and Maenhout (1999). The intuition behind this result is that when there is the possibility of model misspecification, this adds another source of uncertainty to the riskiness of the consumption process. If the investor is averse to model misspecification and pessimistic in choosing alternative models, he appears more risk averse in evaluating the given consumption process.

3.2.2 Example with two risky assets

In this section, we illustrate how concern for model misspecification affects the optimal portfolio by deriving the weights for the setting where there are two risky assets whose returns are given by Brownian motions. For expositional convenience, we assume that the two assets have the same expected return (μ) and volatility (σ), and that their returns are uncorrelated. The more general case, where the expected return and volatility is different for each asset and the returns are correlated, is considered in the next section.

When there are two risky assets, there are three return distributions about which an investor may have knowledge: the joint distribution for the returns on assets 1 and 2, the marginal distribution for asset 1, and the marginal distribution for asset 2. We use ϕ_0 to denote the investor's knowledge of the joint distribution of returns, and ϕ_j, $j = \{1, 2\}$, for the ambiguity about the marginal distribution for the individual asset j. Thus,

$$\Phi = \phi_0 \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}^{-1} + \phi_1 \begin{pmatrix} \sigma_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} + \phi_2 \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{22}^{-1} \end{pmatrix}. \tag{20}$$

Case 1: Equal ambiguity about both return processes

We first consider the case where an investor has knowledge only about the joint process for the returns on assets 1 and 2 implying that $\phi_0 > 0$, and that there is no additional knowledge about the marginal distribution for returns on asset 1 or 2 ($\phi_1 = \phi_2 = 0$). Then, the optimal portfolio weight is:

$$\pi = \underbrace{\frac{1}{\gamma}\frac{\mu - r}{\sigma^2}}_{\text{Merton weights}} \begin{bmatrix} \frac{\phi_0}{(1+\phi_0)} \\ \frac{\phi_0}{(1+\phi_0)} \end{bmatrix}. \tag{21}$$

This is the expression we would get if we used the Maenhout (1999) formulation with multiple risky assets. As one can see, the adjustment factor for model misspecification to the Merton portfolio weights is the same for both assets 1 and 2 , $\phi_0/(1+\phi_0)$. Thus, under such a specification ambiguity about the return distributions would not bias the portfolio toward a particular asset. In this setting where the agent has knowledge only of the joint process for the returns on the two risky assets, the adjustment to the Merton portfolio weights can be interpreted either as a change in risk aversion from γ to $\gamma(1+1/\phi_0)$ or as a change in the expected return from μ to $\left[\mu - \frac{(\mu-r)}{(1+\phi_0)}\right]$.

Case 2: Unequal ambiguity about the returns processes

In order to focus on the effect of differences in ambiguity about the returns processes for the two assets, we now assume that $\phi_j \neq 0, j = \{1,2\}$, while ϕ_0 is set equal to zero in order to obtain a more transparent expression for the portfolio weights.

Under this specification, the optimal portfolio weights are:

$$\pi = \underbrace{\frac{1}{\gamma}\frac{\mu-r}{\sigma^2}}_{\text{Merton weights}} \begin{bmatrix} \frac{\phi_1}{(1+\phi_1)} \\ \frac{\phi_2}{(1+\phi_2)} \end{bmatrix}. \tag{22}$$

We can interpret these weights as the Merton weights, adjusted by the factor $\phi_j/(1+\phi_j)$, with $j = \{1,2\}$. In the limit, as $\phi_j \to 0$, $\pi_j \to 0$; on the other hand, as $\phi_j \to \infty$, π_j approaches the Merton weight.

In this setting, where the agent has knowledge only of the marginal distributions for the returns on the two risky assets but no knowledge in the joint distribution ($\phi_0 = 0$), the adjustment to the Merton portfolio weights can no longer be interpreted in terms of a change in the agent's risk aversion, and the appropriate interpretation is that of an *asset-specific* change in the expected return from μ to $\left[\mu - \frac{(\mu-r)}{(1+\phi_j)}\right], j = \{1,2\}$. Thus, the the observational-equivalence result noted in Anderson, Hansen, and Sargent (1999) and Maenhout (1999) between ambiguity aversion and Stochastic Differential Utility is valid only if the agent is equally ambiguous about the distribution of returns for all assets.

Finally, we note that the ratio of the portfolio weight for asset 1 to asset 2 is given by:

$$\frac{\phi_1\ (1+\phi_2)}{\phi_2\ (1+\phi_1)},$$

which is greater than unity only if $\phi_1 > \phi_2$. Thus, if $\phi_1 > \phi_2$ then the portfolio that accounts for model misspecification will appear biased toward asset 1 relative to the Merton (1971) portfolio that ignores model misspecification and also relative to the Maenhout (1999) model where there is a single parameter governing ambiguity toward all risky assets. In the next section, we examine the magnitude of this bias in the context of international portfolio choice.

4 Calibration to international equity returns

In this section, we illustrate how one can apply the model developed above by exploring its implications for underdiversification. Motivated by the evidence in Tversky and Heath (1991) that individuals behave as though unfamiliar gambles are riskier than familiar ones (even though they assign identical probability distributions to the two gambles), we calibrate the portfolio model to data on domestic and foreign stock returns, and explore how the portfolio weights change as agents exhibit a greater ambiguity about the return distribution for foreign stocks relative to domestic stocks. We would like to emphasize that the goal of this exercise is not to reproduce the weights documented in the literature on the "home-bias" puzzle,[14] but rather: (i) to illustrate how one can apply the model, (ii) to understand the conditions under which the model will yield a portfolio that is under-diversified, and (iii) to show how one can evaluate whether the parameter values chosen are reasonable.

In Section 4.1 we describe the choice of parameter values, in Section 4.2 we report the portfolio weights for a range of ambiguity levels, and in Section 4.3 we explain how one can assess whether the values chosen for the parameters determining the level of ambiguity are reasonable.

4.1 Choice of parameter values

We examine the problem from the perspective of a US investor under the assumption that asset prices are geometric Brownian motions and the investment opportunity set is constant. We use the same data on quarterly stock returns (from MSCI) as that used in French and Poterba (1990) and French and Poterba (1991). This data is for the period 1975-89 and consists of CPI-adjusted real

[14] A survey of the papers attempting to provide various explanations for this puzzle can be found in Stulz (1995) and Lewis (1999).

returns where the investor is assumed to use 3-month forward contracts to fully hedge the amount of the initial investment. We look at a universe with three "countries": US, Japan and Europe (consisting of France, West Germany, Switzerland, and the United Kingdom). The three countries are indexed by $i = \{1, 2, 3\}$. The volatilities of the real rates of return for the US, Japan and Europe are (0.1650, 0.1825, 0.2000), and the US-JP, US-EU, and EU-JP correlations are (0.53, 0.55, 0.40). These estimates of volatilities and correlations are taken from French and Poterba (1990), with the numbers for Europe being averages of the reported estimates for the France, West Germany, Switzerland, and the United Kingdom. The expected real rate of returns on US, Japanese and European equities, computed from the estimates in Table 2 of French and Poterba (1991) under the assumption that the investor's degree of risk aversion is 3, are (0.0464, 0.0430, 0.0460).

In the calibration exercise, we compute the portfolio weights for our model under the assumption that the investor has knowledge of two return distributions. The first is the joint distribution of the stock returns of *all* three countries; the investor's ambiguity about the joint distribution is represented by the parameter $\phi_0 = \phi$. The investor is also assumed to have some knowledge about the marginal distribution of US stock returns and the ambiguity about this is denoted by ϕ_1; in order to constrain the additional knowledge that the investor has about the marginal distribution of US returns, we specify that $\phi_1 = m\,\phi_0 = m\phi$. Finally, we assume that the US agent has no additional knowledge of European and Japanese stock returns over and above what is known about the joint distribution.

From Theorem 2, the portfolio weights for this specification are:

$$\pi = B\,\frac{1}{\gamma}\left[\sigma_R\sigma_R^\top\right]^{-1}(\mu_R - r), \tag{23}$$

$$B = \left(I - \left[\sigma_R\sigma_R^\top\right]^{-1}\Phi^{-1}\right)^{-1}, \tag{24}$$

$$\Phi = \phi\begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}^{-1} + m\,\phi\begin{pmatrix} \sigma_{11}^{-1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{25}$$

where σ_{ii} is the variance of the stock returns for country i, and $\sigma_{ij}, i \neq j$ is the covariance between the stock returns of i and j. Thus, for the calibration we need to specify values for two additional parameters: ϕ and m. We first report the portfolio weights and the total investment in risky assets based on (23) for a range of values for these these two parameters, ϕ and m, and then explain how one can assess whether the values chosen for these parameters are reasonable.

4.2 Portfolio weights

The portfolio weights from the model are compared to those from the Merton (1971) model, where agents are assumed to have no ambiguity about the returns process ($\phi = \infty$ so that $B = I$), and also to the Maenhout (1999) model, where the agent has some ambiguity about the joint process for the returns of *all* risky assets in the portfolio ($0 < \phi < \infty$), but no additional knowledge about the returns process of any individual asset ($m = 0$).

The first three columns of numbers in Table 1 give the weights allocated to US, Japanese and European equities in a portfolio consisting of *only* these three risky assets; in the next two columns, we report the total proportion of wealth invested in the three risky assets and the proportion invested in the riskless asset; and the last three columns of the table give the adjustment to the expected return. The first row (from Table 1 of French and Poterba (1991)) reports the "Observed weight" for the US economy, which exhibit a strong bias toward US equity. The second row (also from Table 1 in French and Poterba) gives the value-weighted market weights. The third row gives the weights determined from the Merton model.[15]

Panels A, B and C of Table 1 give the portfolio weights of an investor who accounts for model misspecification. The three panels correspond to different levels of overall ambiguity, indexed by ϕ; recall that a lower ϕ corresponds to a higher level of ambiguity. Within each panel, the portfolio weights are reported for m ranging from 0 to 5, where 0 corresponds to the case where the investor has no additional knowledge about the marginal distribution of US stock returns. Studying the effect of the parameter m, we see that for the case where $m = 0$ the weights in the risky-asset portfolio are the same as the Merton portfolio weights. This is true across all panels. Thus, a model with only a single parameter controlling the concern for model misspecification, as is the case in Anderson, Hansen, and Sargent (1999) and Maenhout (1999), cannot generate the limited diversification we observe in the data. As m increases, the investor's portfolio is increasingly biased

[15]The closeness of the market portfolio and the Merton portfolio is not typical, and the reader should not infer that the two sets of weights will always be close.

toward US equity relative to the Merton portfolio. This effect is smaller in Panel A, where the overall level of ambiguity is given by $\phi = 1$, and increases as ϕ decreases. For instance, relative to the Merton weight in US equities of 0.504, the weight is 0.647 for the case where $\phi = 1$ and $m = 2$ in Panel A, and it increases to 0.699 when $\phi = 0.5$ and $m = 2$ in Panel B, and to 0.743 when $\phi = 0.25$ and $m = 2$ in Panel C.

The first plot in Figure 1 shows the share allocated to US equities in the portfolio of only risky assets for a broader range of values of ϕ than displayed in the table. The horizontal solid line shows that for the case where $m = 0$ the share allocated to US equities does not change with ϕ. However, as m increases, the bias toward US equities increases. The figure shows that the bias is highest when ϕ is low and m is high.

In addition to studying the home bias, we also evaluate the implications of the model for the *total* share of net wealth invested in equities. Based on the Survey of Consumer Finances, Heaton and Lucas (2000) provide estimates of the proportion of net worth that US households invest in stocks. Their estimates, which depend on how one defines net worth, range from 15% to 34%.[16] From Panel A of Table 1, we see that with an overall level of $\phi = 1$, the concern for model misspecification leads to a decrease in the proportion of wealth invested in the risky assets (see the column titled "Weight in risky assets"): this drops from 0.700 to 0.350 for the "$m = 0$" case, and it increases with m, suggesting that $\phi = 1$ may be too high relative to the estimates reported in Heaton and Lucas. In Panel B, where $\phi = 0.5$, the total investment in risky assets is 0.233, which increases to 0.328 for $m = 1$; for $m \leq 2$, the total investment in the risky assets is within the range documented by Heaton and Lucas. In Panel C, where $\phi = 0.25$, we find that the weight in risky assets is within the reported range for $m \leq 3$. The total proportion of wealth invested in the three risky assets, for a broader range of values of ϕ, can also be seen in the middle plot of Figure 1.

[16] Heaton and Lucas (2000) consider three definitions of net worth: (i) liquid net worth, defined as the sum of cash, bonds, stocks and mutual funds less mortgages and consumer loans; (ii) financial net worth, which is liquid net worth plus real estate holdings, proprietary businesses, pensions and trusts; and, (iii) total net worth, which is equal to financial net worth plus capitalized labor and social security. In Panel A of their Table 1, the mean investment in stocks as a proportion of liquid net worth based on the data for 1989 is 0.155 for households with income between \$10,000 and \$100,000 and 0.340 for households with income greater than \$1 million, with the rest invested in cash and bonds. As a proportion of financial wealth (their Table 2), the proportion invested in stocks is about 0.150 in 1989 and for 1995 the proportion increases to about 0.250. Finally, as a proportion of total net worth (their Table 5), investment in stocks ranges from 0.040 for individuals with a net worth between \$10,000 than \$100,000, to 0.257 for individuals with a net worth exceeding \$1 million.

The table and the plot suggest that to match the overall investment in risky assets the value for ϕ needs to be low, which is also the condition under which the bias toward US equities is high.

4.3 Appropriate choice of ϕ and m

Ideally, one would like have a priori information about the appropriate range for ϕ and for m. Since we do not have direct information about ϕ and m, we infer this indirectly by examining the adjustment to the drift of the returns process implied by the different levels of ϕ and m; this is useful because it is easier to interpret an adjustment to the expected return—for instance, by comparing it to the standard error in estimating expected returns–than to assess whether m and ϕ are reasonable.

The adjustments to the US, Japanese and European expected returns arising from model misspecification are given in the last three column of Table 1. Again focussing on the case where $m = 2$, we see from the first row of Panel A that the adjustment to US expected returns is -0.0155, which is less than the 200 basis points standard error reported in French and Poterba (1991). This is also illustrated in the bottom plot of Figure 1. Note also that the adjustment to the mean return *decreases* as m increases; this is important because an increase in m corresponds to an increase in the bias toward US equities. Thus, a small absolute adjustment to the expected returns is sufficient to generate a large bias in portfolio holdings.

We conclude this section by summarizing the main observation: differences in the level of ambiguity about the returns distributions leads to portfolios that are under-diversified relative to the Merton (1971) model where there is no ambiguity, and also to the Maenhout (1999) model where there is a single parameter measuring the agent's ambiguity toward the distribution of all the asset returns. The under-diversification effect is strongest when the overall level of ambiguity is high.

5 Conclusion

In this paper, we have developed a model which formalizes the problem of investors who are concerned about model misspecification because they understand that the distributions of assets returns are not estimated with perfect precision. Our model allows agents to have different levels of ambiguity for the distribution of returns for each of the stocks in the portfolio. The model shows that when the overall degree of ambiguity is high, then small differences in ambiguity about the marginal distribution of asset returns will lead to a strong bias in portfolio holdings.

Traditional models of portfolio choice predict that investors should hold diversified portfolios. However, there is substantial evidence of a bias toward familiar assets in both international and domestic portfolios of institutions and individual households. International equity portfolios are strongly biased toward domestic stocks (Cooper and Kaplanis, 1994; French and Poterba, 1991); and, of the limited foreign investments by US and Canadian investors, a disproportionate share is invested across the border, even though the correlation between US and Canadian returns are higher than the correlations with Japanese and European equity returns (Tesar and Werner, 1995). Evidence on domestic portfolios reveals a similar lack of diversification: US households are more likely to invest in their local US Regional Bell Operating Companies rather than some other Regional Bell Operating Company (Huberman, 2001); workers tend to hold their own company's stock in their retirement accounts (Schultz, 1996); and, Grinblatt and Keloharju (1999) report that Finnish households are more likely to invest in firms that are located close to them and that communicate in the investor's native language (Swedish vs. Finnish). At the institutional level, US mutual fund managers exhibit a preference for local companies (Coval and Moskovitz, 1999). The model we develop can be viewed as offering at least a partial explanation for the observed under-diversification and bias toward familiar securities

Appendix: Proofs for theorems

Proof of Theorem 1

Let Q be an alternative model. According to Girsanov's Theorem, $dQ/dP = \xi_T$ is given by,

$$\xi_t = \exp\left\{-\int_0^t a_s^\top dw_s - \frac{1}{2}\int_0^t |a_s|^2\, ds\right\},$$

for some appropriate adapted process a_t. The result of this change of probability is a drift adjustment to the process of X_t, given by $-\sigma_X a_t$. In other words, a_t can be chosen of the following form

$$a_t^\top = -v_t^\top [\sigma_X \sigma_X^\top]^{-1} \sigma_X.$$

Then, applying Girsanov's Theorem to X_{J_i}, $dQ^\xi_{J_i}/dP_{J_i}$ is given by

$$\xi_{it} = \exp\left\{-\int_0^t a_{J_i s}^\top dw_s - \frac{1}{2}\int_0^t |a_{J_i s}|^2\, ds\right\},$$

where

$$a_{J_i t}^\top = -v_{J_i t}^\top [\sigma_{X_{J_i}} \sigma_{X_{J_i}}^\top]^{-1} \sigma_{J_i X}.$$

Here $v_{J_i t} = (v_{j_1 t}, \ldots, v_{j_{n_i} t})$, $[\sigma_{X_{J_i}} \sigma_{X_{J_i}}^\top]$ is the instantaneous variance-covariance matrix of X_{J_i}, and $\sigma_{J_i X}$ is the matrix whose rows are those of σ_X that correspond to X_{J_i}. Furthermore, by Girsanov's Theorem,

$$dw_t^\xi = dw_t + a_t dt$$

is a Brownian motion under Q^ξ, and

$$dX_t = [\mu_X(X_t, t) + v_t]\, dt + \sigma_X(X_t, t) dw_t^\xi.$$

Let

$$V_t = \inf_\xi \left\{ u(c_t)\Delta + e^{-\rho\Delta}\left[\sum_i^K \psi(V_t)\phi_i L(\xi_{it+\Delta}) + E_t^\xi[V_{t+\Delta}]\right]\right\}.$$

Then

$$[V_t - u(c_t)\Delta]\, e^{\rho\Delta} = \inf_\xi \left\{ \psi(V_t) \sum_{i=1}^K \phi_i L(\xi_{it+\Delta}) + E_t^\xi[V_{t+\Delta}]\right\},$$

and thus,

$$0 = \inf_{\xi} \left\{ \psi(V_t) \sum_{i=1}^{K} \phi_i L(\xi_{it+\Delta}) + E_t^{\xi}[V_{t+\Delta}] - V_t + V_t - [V_t - u(c_t)\Delta]\, e^{\rho\Delta} \right\}.$$

Letting $\Delta \to 0$,

$$\frac{[E_t^{\xi}[V_{t+\Delta}] - V_t]}{\Delta} \to \mathcal{A}[V_t] + v_t^{\top} V_X,$$

and

$$-\frac{\Big[[V_t - u(c_t)\Delta]\, e^{\rho\Delta} - V_t\Big]}{\Delta} \to [u(c_t) - \rho V_t].$$

Since $d\xi_{it} = -\xi_{it} a_{J_i t}^{\top} dw_t$, we have,

$$\begin{aligned} L(\xi_{it+\Delta}) &= \frac{E_t[\xi_{it+\Delta} \ln \xi_{it+\Delta}] - E[\xi_{it+\Delta}] \ln E_t[\xi_{it+\Delta}]}{E_t[\xi_{it+\Delta}]} \\ &= \frac{E_t[\xi_{it+\Delta} \ln \xi_{it+\Delta}] - \xi_{it} \ln \xi_{it}}{\xi_{it}}. \end{aligned}$$

After a straightforward calculation using Itô's Lemma, we have,

$$L(\xi_{it+\Delta})/\Delta \to \frac{1}{2} a_{J_i t}^{\top} a_{I_1 t} = \frac{1}{2} v_{J_i t}^{\top} [\sigma_{X_{J_i}}\, \sigma_{X_{J_i}}^{\top}]^{-1}\, v_{J_i t}.$$

Substituting yields the desired result. ∎

Proof of Theorem 2

The first-order condition for the minimization problem in (15) is:

$$0 = \begin{bmatrix} V_W W \pi_t + V_R \\ V_Y \end{bmatrix}^{\top} + \psi(V_t) \begin{bmatrix} v_R \\ v_Y \end{bmatrix}^{\top} \Phi.$$

The first-order conditions for the maximization problem are

$$\begin{aligned} 0 &= u'(c) - V_W, \\ 0 &= V_W[(\mu_R - r_t) + v_R] + \sigma_R \sigma_Y^{\top} V_{WY} + \sigma_R \sigma_R^{\top} V_{WR} + W_t V_{WW} \sigma_R \sigma_R^{\top} \pi_t. \end{aligned}$$

The first part of the theorem follows directly from these two first order conditions. For the second part, re-write the first part of the theorem as,

$$\pi = -\frac{1}{WV_{WW}}[\sigma_R\sigma_R^\top]^{-1}\left(V_W(\mu_R - r) + \sigma_R\sigma_Y^\top V_{WY} + \sigma_R\sigma_R^\top V_{WR} + V_W\begin{bmatrix} I_R & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_R^* \\ v_Y^* \end{bmatrix}\right)$$

and

$$\begin{bmatrix} v_R^* \\ v_Y^* \end{bmatrix} = -\frac{V_W W}{\psi(V)}\Phi^{-1}\begin{bmatrix} I_R \\ 0 \end{bmatrix}\pi - \frac{1}{\psi(V)}\Phi^{-1}\begin{bmatrix} V_R \\ V_Y \end{bmatrix}.$$

Solving for π yields the closed-form solution. ∎

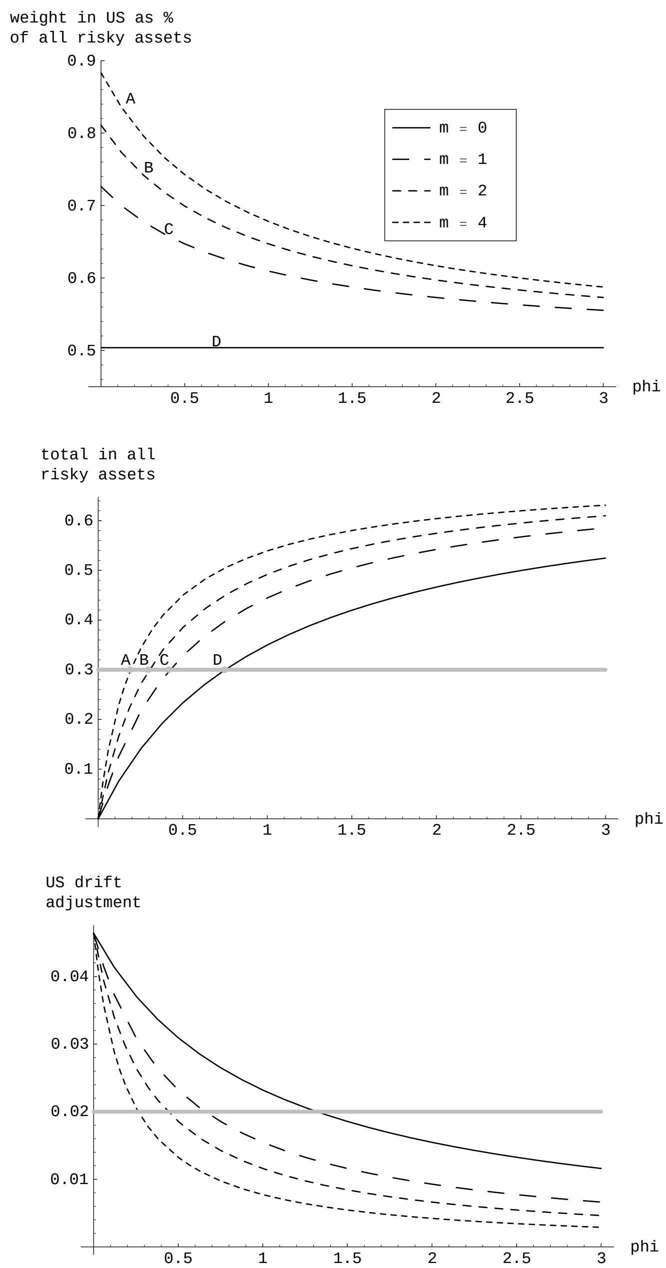
Table 1: Portfolio weights and adjustment to expected returns

This table compares the portfolio weights of a US investor who accounts for model misspecification to those from the Merton (1971) and Maenhout (1999) models. The first three columns of numbers report the weight allocated to the US, Japanese (JP) and European (EU) equities in the risky-asset portfolio. The next two columns give the proportion of total wealth invested in the three risky assets and the weight in the riskfree asset. The last three columns report the adjustment to expected returns implied by the choice of ϕ and m. As in French and Poterba (1991), the investor is assumed to have a risk aversion of 3, while the vector of expected real rates of return on US, Japanese and US equities is $\{.0464, .0430, .0460\}$, the volatility vector for US, Japanese and European markets is $\{.1650, .1825, .2000\}$, and the US-JP, US-EU, and EU-JP correlations are $\{.53, .55, .40\}$. The three panels correspond to different levels of overall ambiguity, indexed by ϕ. The first row in *each* panel, "$m = 0$" corresponds to the case in equation (25), where one has knowledge about the joint distribution of asset returns for the three indexes (US, JP and EU) but no additional knowledge about any of the marginal distributions; this matches the model in Maenhout (1999)). The rows titled "$m = 1$" to "$m = 5$" correspond to the case where the investor is less ambiguous about the marginal distribution of US returns. The table shows that the bias toward US equities increases as ambiguity about the marginal distribution for US stock returns decreases (measured by an increase in m); this effect is larger when the overall level of ambiguity is high (low ϕ).

	Risky-asset portfolio			Weight in	Weight in	Drift adjustment		
	US	JP	EU	risky assets	riskless asset	US	JP	EU
Observed weights	0.938	0.031	0.031					
Market weights	0.496	0.276	0.228					
Merton-weights	0.504	0.278	0.218	0.700	0.300			
Panel A: $\phi = 1.00$								
$m = 0$	0.504	0.278	0.218	0.350	0.650	-0.0232	-0.0215	-0.0230
$m = 1$	0.610	0.219	0.171	0.444	0.556	-0.0155	-0.0170	-0.0178
$m = 2$	0.647	0.198	0.155	0.492	0.508	-0.0116	-0.0147	-0.0153
$m = 3$	0.666	0.187	0.146	0.520	0.480	-0.0093	-0.0133	-0.0137
$m = 4$	0.678	0.181	0.141	0.539	0.461	-0.0077	-0.0124	-0.0127
$m = 5$	0.686	0.176	0.138	0.553	0.447	-0.0066	-0.0118	-0.0120
Panel B: $\phi = 0.50$								
$m = 0$	0.504	0.278	0.218	0.233	0.767	-0.0309	-0.0287	-0.0307
$m = 1$	0.647	0.198	0.155	0.328	0.672	-0.0232	-0.0241	-0.0255
$m = 2$	0.699	0.169	0.132	0.385	0.615	-0.0186	-0.0214	-0.0224
$m = 3$	0.726	0.154	0.120	0.423	0.577	-0.0155	-0.0196	-0.0204
$m = 4$	0.743	0.144	0.113	0.450	0.550	-0.0133	-0.0183	-0.0189
$m = 5$	0.754	0.138	0.108	0.470	0.530	-0.0116	-0.0173	-0.0178
Panel C: $\phi = 0.25$								
$m = 0$	0.504	0.278	0.218	0.140	0.860	-0.0371	-0.0344	-0.0368
$m = 1$	0.678	0.181	0.141	0.216	0.784	-0.0309	-0.0308	-0.0327
$m = 2$	0.743	0.144	0.113	0.270	0.730	-0.0265	-0.0282	-0.0297
$m = 3$	0.776	0.125	0.098	0.310	0.690	-0.0232	-0.0262	-0.0275
$m = 4$	0.797	0.114	0.089	0.342	0.658	-0.0206	-0.0247	-0.0258
$m = 5$	0.811	0.106	0.083	0.367	0.633	-0.0186	-0.0235	-0.0244

Figure 1: Portfolio weights and adjustment to expected returns

The three panels plot as a function of the agent's ambiguity (ϕ): (i) the portfolio weight allocated to US equity relative to the total investment in risky assets, $\pi_{US}/(\pi_{US}+\pi_{JP}+\pi_{EU})$, (ii) the total investment in risky assets $(\pi_{US}+\pi_{JP}+\pi_{EU})$, and (iii) the adjustment to the US expected returns (drift). The figure is obtained using the same parameter values as the ones described in the legend of Table 1. In each panel, four cases are plotted. In the first case, an investor's knowledge of the joint distribution for the returns on US, Japanese and European equities is given by ϕ, but there is no additional knowledge about the marginal distribution of US returns ($m = 0$). The three other cases plotted correspond to $m = 1$, $m = 2$, and $m = 4$, where $m\,\phi$ measures the additional knowledge that the investor has about the marginal distribution for US equity returns. From the first panel, we see that the holding of the US assets increases with m, and is particularly pronounced for low values of ϕ, while the second panel shows that the total investment in risky assets increases with ϕ and with m. The third panel shows the adjustment to expected returns implied by different combinations of ϕ and m.



References

Anderson, E., L. Hansen, and T. Sargent, 1999, "Robustness detection and the price of risk," Working paper, University of Chicago.

Chen, Z., and L. Epstein, 2000, "Ambiguity, risk and asset returns in continuous time," Working paper, University of Rochester.

Cooper, I., and E. Kaplanis, 1994, "Home bias in equity portfolios, inflation hedging and international capital market equilibrium," *Review of Financial Studies*, 7, 45–60.

Coval, J., and T. Moskovitz, 1999, "Home bias at home: Local equity preference in domestic portfolios," *Journal of Finance*, 54, 2045–2073.

Dow, J., and S. Werlang, 1992, "Uncertainty aversion, risk aversion, and the optimal choice of portfolio," *Econometrica*, 60, 197–204.

Duffie, D., and L. Epstein, 1992, "Stochastic Differential Utility," *Econometrica*, 60, 353–394.

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

Epstein, L., and J. Miao, 2000, "A two-person dynamic equilibrium under ambiguity," Working paper, University of Rochester.

Epstein, L., and M. Schneider, 2001, "Recursive multiple-priors," Working paper, University of Rochester.

Epstein, L., and T. Wang, 1994, "Intertemporal asset pricing under Knightian uncertainty," *Econometrica*, 62, 283–322.

Epstein, L., and S. Zin, 1989, "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

French, K., and J. Poterba, 1990, "Japanese and U.S. cross-border common stock investments," *Journal of the Japanese and International Economics*, 4, 476–493.

French, K., and J. Poterba, 1991, "Investor diversification and international equity markets," *American Economic Review*, 81, 222–226.

Gorman, L., and B. Jorgensen, 1999, "Domestic versus international portfolio selection: A statistical examination of the home bias," Working paper, Harvard University.

Grinblatt, M., and Keloharju, 1999, "Distance bias, language bias, and investor sophistication: Results from Finland," Working paper, UCLA.

Hansen, L., and T. Sargent, 2000, "Robustness and Uncertainty Aversion," Working paper, University of Chicago.

Hansen, L., and T. Sargent, 2001a, "Robust Control and Model Uncertainty," *American Economic Review*, 91, 60–66.

Hansen, L., and T. Sargent, 2001b, "Wanting robustness in macroeconomics," Working paper, University of Chicago.

Hansen, L., and T. J. Sargent, 2001c, "Elements of robust control and filtering for macroeconomics," Working paper, University of Chicago.

Hansen, L. P., T. J. Sargent, and T. D. Tallarini, 1999, "Robust permanent income and pricing," *Review of Economic Studies*, 66, 873–907.

Heaton, J., and D. Lucas, 2000, "Portfolio choice and asset prices: The importance of entrepreneurial risk," *Journal of Finance*, 55, 1163–1198.

Huberman, G., 2001, "Familiarity breeds investment," *Review of Financial Studies*, 14, 659–680.

Lewis, K., 1999, "Trying to explain home bias in equities and consumption," *Journal of Economic Literature*, 37, 571–608.

Lucas, R., 1978, "Asset prices in an exchange economy," *Econometrica*, 46, 1429–1445.

Maenhout, P., 1999, "Robust portfolio rules and asset pricing," Working paper, Harvard University.

Merton, R., 1971, "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 3, 373–413.

Merton, R., 1973, "An intertemporal asset pricing model," *Econometrica*, 41, 867–888.

Merton, R., 1980, "On estimating the expected return on the market: An exploratory investigation," *Journal of Financial Economics*, 8, 323–361.

Schultz, E., 1996, "Workers put too much in their employer's stock," *Wall Street Journal*, September 13.

Stambaugh, R., 1997, "Analyzing investments whose histories differ in length," *Journal of Financial Economics*, 45, 285–331.

Stulz, R., 1995, "International portfolio choice and asset pricing: An integrative survey," in R. Jarrow, V. Maksimovic, and W. Ziemba (ed.), *Handbook in Operations Research and Management Science, Finance* . , vol. 9, North-Holland Amsterdam.

Tesar, L., and I. Werner, 1995, "Home bias and high turnover," *Journal of International Money and Finance*, 4, 467–492.

Tversky, A., and C. Heath, 1991, "Preferences and beliefs: Ambiguity and competence in choice under uncertainty," *Journal of Risk and Uncertainty*, 4, 5–28.