

A Simple Theory of Asset Pricing under Model Uncertainty

Leonid Kogan and Tan Wang*

February, 2003

Abstract

The focus of our paper is on the implications of model uncertainty for the cross-sectional properties of returns. We perform our analysis in the context of a tractable single-period mean-variance framework. We show that there is an uncertainty premium in equilibrium expected returns on financial assets and study how the premium varies across the assets. In particular, the cross-sectional distribution of expected returns can be formally described by a two-factor model, where expected returns are derived as compensation for the asset's contribution to the equilibrium risk and uncertainty of the market portfolio. In light of the large empirical literature on the cross-sectional characteristics of asset returns, understanding the implications of model uncertainty even in such a simple setting would be of significant value. By characterizing the cross-section of returns we are also able to address some of the observational equivalence issues raised in the literature. That is, whether model uncertainty in financial markets can be distinguished from risk, and whether uncertainty aversion at an individual level can be distinguished from risk aversion.

*Leonid Kogan is with the Sloan School of Management of the Massachusetts Institute of Technology, Cambridge, MA 02142, USA. lkogan@mit.edu; Tan Wang is with the Faculty of Commerce and Business Administration, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z2. tan.wang@commerce.ubc.ca. The authors are grateful to Lorenzo Garlappi, Burton Hollifield, Raman Uppal, Stan Zin, and seminar participants at Carnegie Mellon University and MIT for helpful comments. The paper was presented at the 2002 Western Finance Association Meetings. Tan Wang thanks the Social Sciences and Humanities Research Council of Canada for financial support.

1 Introduction

The purpose of this paper is to study the implications of model uncertainty for the cross-sectional properties of asset prices in a simplest possible equilibrium setting.

The focus on model uncertainty is motivated by the difficulty of reconciling existing asset pricing theories with the empirical data. Limited success of the standard theories could be in part due to the commonly made assumption that economic agents possess perfect knowledge of the data generating process. For instance, the classical theories of Sharpe (1964), Lucas (1978), Breeden (1979) and Cox, Ingersoll and Ross (1985), assume that, while the payoffs of financial assets are random, agents know the underlying probability law exactly. In reality this is often not the case. Then the natural question is: how are the prices of financial assets affected by investors' lack of knowledge about the probability law, or their uncertainty about what the true model is.

The importance of model uncertainty has long been recognized in finance. While the literature appears under different names, such as parameter uncertainty, Knightian uncertainty, the defining characteristic of that literature is the recognition of the fact that the agents of the economy do not have a perfect knowledge of the probability law that governs the realization of the states of the world. Various issues have been studied. Dow and Werlang (1992) use the uncertainty averse preference model developed by Schmeidler (1989) to study a single period portfolio choice problem. Maenhout (1999) examines a similar problem in a continuous-time economy, but from the point of view of robust portfolio rules. Kandel and Stambaugh (1996), Brennan (1998), Barberis (2000), and Xia (2001) show that parameter uncertainty can affect significantly investors' portfolio choice. Frost and Savarino (1986), Genotte (1986), Balduzzi and Liu (1999), Pastor (2000) and Uppal and Wang (2001) examine the implication of model uncertainty for portfolio choices when there are multiple risky assets. Detemple (1986), Epstein and Wang (1994), Chen and Epstein (2001), Epstein and Miao (2001), and Brennan and Xia (2001) study the implications for equilibrium asset prices in the representative agent and heterogenous agent economies respectively. Routledge and Zin (2002) examine the connection between model uncertainty and liquidity. There is also a significant literature, for example Lewellen and Shanken (2001) and Brav and Heaton (2002), on the effect of learning about an unknown parameter on the equilibrium asset prices.

The focus of our paper is on the cross-sectional properties of returns. We perform our analysis in the context of a tractable single-period mean-variance framework. We show that there is an uncertainty premium in equilibrium expected returns on financial assets and study how the premium varies across the assets. In light of the large empirical literature on the cross-sectional characteristics of asset returns, understanding the implications of model uncertainty and uncertainty aversion even in such a simple setting would be of significant value. While prior research on model uncertainty has been concerned with its implications for the time-series of asset prices, by characterizing the cross-section of returns we are able to address some of the observational equivalence issues raised in the literature. That is, whether model uncertainty in financial markets can be distinguished from risk, and whether uncertainty aversion at an individual level can be distinguished from risk aversion (Anderson, Hansen and Sargent (1999)).

In the rest of this introduction, we will describe briefly our approach to formalizing model uncertainty and its relation to the the literature. The most common way of modelling imperfect knowledge of the model and parameters is in the Bayesian framework (Kandel and Stambaugh (1996), Lewellen and Shanken (2001), Barberis (2000) and Pástor (2000)). The key feature of this approach is that if a parameter of the model is unknown, a prior distribution of the parameter is introduced. The second approach, adopted by Dow and Werlang (1992), Epstein and Wang (1994, 1995), Chen and Epstein (2001), Epstein and Miao (2001), and the third approach, adopted by Anderson, Hansen and Sargent (1999), Maenhout (1999), Uppal and Wang (2001), follow the view of Knight (1921) that model uncertainty, or more precisely, the decision makers' view of model uncertainty, cannot be represented by a probability prior. There is a significant literature in psychology and experimental economics that documents the contrast between the Bayesian and Knightian approaches. The evidence documented there is that, when faced with uncertainty about the true probability law, people's behavior tend to be inconsistent with the prediction of the Bayesian approach (Ellsberg (1963)). In fact, the behavior is inconsistent with any probabilistically sophisticated preference (Machina and Schmeidler (1992)).

The second and third approaches differ in how uncertainty and uncertainty aversion are modelled. Maenhout (1999), Uppal and Wang (2001), use the preference first introduced by Anderson, Hansen and Sargent (1999) in their study of the implications of preference for

robustness for macroeconomic and general asset pricing issues.¹ This class of preferences has been extended in Uppal and Wang (2001), and axiomatized in a static setting in Wang (2001). For this class of preferences, uncertainty is described by a set of priors and the investor's aversion to it is introduced through a penalty function. Dow and Werlang (1992), Epstein and Wang (1994, 1995), Chen and Epstein (2001) and Epstein and Miao (2001) use the multi-prior expected utility developed by Gilboa and Schmeidler (1989).² Here both uncertainty and uncertainty aversion are introduced through a set of priors. This paper is based on the multi-prior expected utility preferences with a careful design of the set of priors to distinguish between the uncertainty and uncertainty aversion aspects of the set.

The rest of the paper is organized as follows. Section 2 describes the model. Section 4 presents the main result of this paper, the asset pricing implication of model uncertainty. Section 5 discusses several issues related to model uncertainty. Finally, Section 6 concludes.

2 The Model

2.1 The Setting

We assume a one-period representative agent economy. Consumption takes place only at the end of the period. The agent is endowed with an initial wealth W_0 . Without loss of generality, we assume $W_0 = 1$.

There are N risky assets and one risk-free asset in the economy, which is available in zero net supply. As indicated in the introduction, the investors do not have perfect knowledge of the distribution of the returns of the N risky assets. More specifically, they know that the returns $R = (R_1, \dots, R_N)^\top$ follow a joint normal distribution with density function

$$f(R) = (2\pi)^{-n/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2} (R - \mu)^\top \Omega^{-1} (R - \mu) \right\}$$

¹See Hansen and Sargent (2001) for more on this type of preferences.

²Dow and Werlang is based more directly on the Choquet expected utility developed by Schmeidler (1989). However, for the case they studied, Choquet expected utility coincides with multi-prior expected utility.

where

$$\mu = \mathbb{E}[R], \quad \Omega = \mathbb{E}[(R - \mu)(R - \mu)^\top].$$

The risk of returns is summarized by the non-degenerate variance-covariance matrix Ω . We assume that investors have precise knowledge of Ω . However, they do not know exactly the mean return vector μ . This is motivated by the fact that it is much easier to obtain accurate estimates of the variance and covariance of returns than their expected values, e.g., Merton (1992). The imperfect knowledge of the asset return distribution gives rise to model uncertainty.

2.2 The Preferences

Each agent in the economy has a state-independent CARA utility function

$$u(W) = -\frac{1}{\gamma} \exp(-\gamma W).$$

Due to lack of perfect knowledge of the probability law of asset returns, however, the agent's preference is not represented by the standard expected utility, but instead by a multi-prior expected utility

$$U(W, \mathcal{P}(P, \phi)) = \min_{Q \in \mathcal{P}(P, \phi)} \{ \mathbb{E}^Q[u(W)] \}, \quad (1)$$

where \mathbb{E}^Q denotes the expectation under the probability measure Q , $\mathcal{P}(P, \phi)$ is a set of probability measures that depends on the probability measure P , called the reference prior, and the parameter $\phi \geq 0$, which is called the uncertainty aversion parameter. The set $\mathcal{P}(P, \phi)$ captures both the degree of model uncertainty and the agent's degree of uncertainty aversion. In particular, we assume that the larger the uncertainty aversion parameter ϕ , the larger the set $\mathcal{P}(P, \phi)$. The multi-prior expected utility preferences exhibits uncertainty aversion. The general nature and the axiomatic foundation of these preferences has been well studied in the literature (Gilboa and Schmeidler (1989)). What is specific to this paper is the structure of $\mathcal{P}(P, \phi)$, in particular, the use of ϕ as the uncertainty aversion parameter. Since the exact structure of $\mathcal{P}(P, \phi)$ is important to our analysis and to the understanding of our results, we now turn to the description of the exact dependence of $\mathcal{P}(P, \phi)$ on P and ϕ .

A Single Source of Information

We begin with the case of a single source of information about the distribution of stock returns,

$$\mathcal{P}(P, \phi) = \{Q : E[\xi \ln \xi] \leq \phi^2 \eta\},$$

where ξ is the density of Q with respect to P and η is a parameter to be described shortly.

The intuition behind this formulation of $\mathcal{P}(P, \phi)$ can be explained as follows. Since the investor lacks a perfect knowledge of the probability law of the returns, an econometrician is asked to estimate a model of the asset returns for the investor. After a slew of econometric analysis, typically including specification analysis and parameter estimation, the econometrician comes up with a model described by the probability measure P . However, the econometrician is not completely confident that this is the true model, due to not having enough data in the specification analysis and the parameter estimation, or due to simplifying assumptions made for tractability. On the other hand, the econometric analysis does provide more information than just the probability measure P . The true model can be narrowed down to a set \mathcal{P} of probability measures. At the end of the analysis, the investor is presented with a probability measure P , called the reference prior, and a set \mathcal{P} that summarize the precision of the econometric analysis.

Since both the econometrician and the investor are not completely sure of the reference prior P , each element in \mathcal{P} is a possible alternative to the reference prior P . Let Q be an element in \mathcal{P} and let its density be denoted by ξ , so that

$$dQ = \xi dP. \tag{2}$$

Knowing that the reference measure P is subject to misspecification and that the possible alternative is Q , the problem is how to evaluate the alternative. One way is to use the relative entropy index, $E[\xi \ln \xi]$. One interpretation of the index is that it is an approximation to the empirical log-likelihood ratio.³ To elaborate, suppose that the data set available to the investor has T observations. Then the empirical log-likelihood ratio of the two models is

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

³See Anderson, Hansen and Sargent (1999) and Hansen and Sargent (2000) for other interpretations of the index.

Now suppose that X_t , $t = 1, \dots, T$, takes finitely many values, x_1, \dots, x_k in the data series. Then

$$\frac{1}{T} \sum_{t=1}^T \ln \xi(X_t) = \frac{1}{T} \sum_{i=1}^k \sum_{X_t=x_i} \ln \xi(X_t) = \sum_{i=1}^k \frac{T_i}{T} \ln \xi(x_i),$$

where T_i is the number of t such that $X_t = x_i$. By the law of large numbers, under the alternative model Q , T_i/T converges to $Q(x) = \xi(x)P(x)$ and hence $\frac{1}{T} \sum_{t=1}^T \ln \xi(X_t)$ converges to $E[\xi \ln \xi]$. Thus, if Q is the true probability law, $E[\xi \ln \xi]$ is a good approximation to the empirical log-likelihood when T is large. According to the traditional likelihood ratio theory, if the above sum is large, then the two alternatives, Q and P , can be clearly distinguished.⁴ Therefore the set of possible alternative models according to the econometrician is given by

$$\mathcal{P}(P) = \{Q : E[\xi \ln \xi] \leq \eta\}$$

where η is the parameter the econometrician uses to describe how much uncertainty there is about the reference probability P . For example, η could be chosen to define a rejection region for a test of the reference model P with a 95% confidence level.

The investor's uncertainty aversion is introduced through the parameter ϕ , so that the set provided by the econometrician is scaled up or down by ϕ :

$$\mathcal{P}(P, \phi) = \{Q : E[\xi \ln \xi] \leq \phi^2 \eta\}.$$

Larger values of ϕ allow for a larger set of alternative models. Thus, more uncertainty averse agents are willing to entertain alternative models that are relatively far from the reference model P , as measured by their relative entropy. An investor more averse to uncertainty would require a higher level of confidence, say 99%, than the one used by the econometrician, and vice versa.

For analytical tractability, we assume that stock returns are jointly normally distributed under the alternative models. Furthermore, we assume that the variance-covariance matrix of the returns is the same under all measures in \mathcal{P} , reflecting the fact that the investor knows

⁴It is worth emphasizing that large $\frac{1}{T} \sum_{t=1}^T \xi(X_t) \ln \xi(X_t)$ should not be interpreted as evidence for *rejecting* the reference model P , as in the usual likelihood test: as explained above, the very fact that P is the reference prior implies that the econometrician has already gone through the preliminary analysis and picked P . The issue at this stage is only to find an index that summarizes the information available.

the variance-covariance matrix Ω precisely. Let Q be a measure in \mathcal{P} with the density given by

$$(2\pi)^{-n/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2} (R - \hat{\mu})^\top \Omega^{-1} (R - \hat{\mu}) \right\},$$

which can be written as

$$(2\pi)^{-n/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2} (R - \mu)^\top \Omega^{-1} (R - \mu) \right\} \\ \times \exp \left\{ -\frac{1}{2} (\mu - \hat{\mu})^\top \Omega^{-1} (\mu - \hat{\mu}) - (\mu - \hat{\mu})^\top \Omega^{-1} (R - \mu) \right\}.$$

Thus, the likelihood ratio of Q over P is given by

$$\xi(R) = \exp \left\{ \frac{1}{2} (\mu - \hat{\mu})^\top \Omega^{-1} (\mu - \hat{\mu}) - (\mu - \hat{\mu})^\top \Omega^{-1} (R - \hat{\mu}) \right\}. \quad (3)$$

Given this particular structure of the set $\mathcal{P}(P, \phi)$, the representative investor's utility function can be written as

$$\min_{v \in \mathcal{V}(\phi)} \mathbb{E} [\xi u(W)], \quad (4)$$

where ξ is given by (3), $v = \mu - \hat{\mu}$ and the set \mathcal{V} corresponds to \mathcal{P} :

$$\mathcal{V}(\phi) = \left\{ v : \mathbb{E}[\xi \ln \xi] = \frac{1}{2} v^\top \Omega^{-1} v \leq \phi^2 \eta \right\}.$$

Multiple Sources of Information

In reality, the investor's knowledge about the distribution of asset returns often comes from different sources and it is often about a subset of the assets, as opposed to the joint distribution of all assets as in the previous subsection. To accommodate this, let J_k , $k = 1, \dots, K$, be subsets of $\{1, \dots, N\}$, each set J_k having N_k elements. Sets J_k are not necessarily disjoint. But we will assume that $\cup_k J_k = \{1, \dots, N\}$, so that the investor has at least some information about each asset. Each k is an index of a source of information about assets in the set J_k . Let $J_k = \{j_1, \dots, j_{N_k}\}$, so that the information is about the distribution of $R_{J_k} = (R_{j_1}, \dots, R_{j_{N_k}})$. We assume that the reference probability distributions implied

by the various sources of information for the corresponding subsets of assets coincide with the marginal distributions of the reference model P . Consider the density function of the distribution of R_{J_k} ,

$$(2\pi)^{-1} |\Omega_{J_k}|^{-1/2} \exp \left\{ -\frac{1}{2} (R_{J_k} - \hat{\mu}_{J_k})^\top \Omega_{J_k}^{-1} (R_{J_k} - \hat{\mu}_{J_k}) \right\},$$

where $\hat{\mu}_{J_k} = (\hat{\mu}_{j_1}, \dots, \hat{\mu}_{j_{N_k}})$, and Ω_{J_k} is the variance-covariance matrix of R_{J_k} , which is a sub-matrix of Ω . This density function can be written as

$$\begin{aligned} & \exp \left\{ -\frac{1}{2} (\mu_{J_k} - \hat{\mu}_{J_k})^\top \Omega_{J_k}^{-1} (\mu_{J_k} - \hat{\mu}_{J_k}) - (\mu_{J_k} - \hat{\mu}_{J_k})^\top \Omega_{J_k}^{-1} (R_{J_k} - \mu_{J_k}) \right\} \\ & \times (2\pi)^{-1} |\Omega_{J_k}|^{-1/2} \exp \left\{ -\frac{1}{2} (R_{J_k} - \mu_{J_k})^\top \Omega_{J_k}^{-1} (R_{J_k} - \mu_{J_k}) \right\}. \end{aligned}$$

Thus, the likelihood ratio of the marginal distribution Q_{J_k} over P_{J_k} is

$$\xi_{J_k} = \exp \left\{ \frac{1}{2} (\mu_{J_k} - \hat{\mu}_{J_k})^\top \Omega_{J_k}^{-1} (\mu_{J_k} - \hat{\mu}_{J_k}) - (\mu_{J_k} - \hat{\mu}_{J_k})^\top \Omega_{J_k}^{-1} (R_{J_k} - \hat{\mu}_{J_k}) \right\}.$$

To relate to the probability measure Q , suppose its density function is

$$(2\pi)^{-n/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2} (R - \hat{\mu})^\top \Omega^{-1} (R - \hat{\mu}) \right\}.$$

Then

$$\begin{aligned} & (2\pi)^{-1} |\Omega_{J_k}|^{-1/2} \exp \left\{ -\frac{1}{2} (R_{J_k} - \hat{\mu}_{J_k})^\top \Omega_{J_k}^{-1} (R_{J_k} - \hat{\mu}_{J_k}) \right\} \\ & = \int (2\pi)^{-n/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2} (R - \hat{\mu})^\top \Omega^{-1} (R - \hat{\mu}) \right\} dR_{J_k-}, \end{aligned}$$

where $J_k- = \{1, \dots, N\} - J_k$. Thus, ξ_{J_k} is the likelihood ratio of the marginal distribution of Q over that of P .

For notational convenience, let $\hat{\Omega}_{J_k}^{-1}$ denote the $N \times N$ -matrix whose element in the j_m th row and j_n th column, for j_m and j_n in J_k , is equal to the element in the m th row and n th

column of the matrix $\Omega_{J_k}^{-1}$; otherwise it is zero. Then

$$(\mu_{J_k} - \hat{\mu}_{J_k})^\top \Omega_{J_k}^{-1} (\mu_{J_k} - \hat{\mu}_{J_k}) = (\mu - \hat{\mu})^\top \hat{\Omega}_{J_k}^{-1} (\mu - \hat{\mu}) = v^\top \hat{\Omega}_{J_k}^{-1} v$$

In the case where there are multiple sources of information, the representative investor's utility function is given by

$$\min_{v \in \mathcal{V}(\phi)} \mathbb{E} [\xi u(W)], \quad (5)$$

where ξ is given by (3), and similar to the single source information case,

$$\mathcal{V}(\phi) = \{v : \mathbb{E}[\xi_{J_k} \ln \xi_{J_k}] = \frac{1}{2} v^\top \hat{\Omega}_{J_k}^{-1} v \leq \phi^2 \eta_k, \quad k = 1, \dots, K\}. \quad (6)$$

2.3 A Measure of Uncertainty

To understand how the investor trades off uncertainty and expected return, it is useful to introduce a metric for uncertainty of various random variables. Let x be a random variable whose distribution is normal and whose variance is the same under P and all measures $Q \in \mathcal{P}$. An example of such a random variable would be the return on a portfolio of N risky assets. Define

$$\Delta(x) = \sup_{Q \in \mathcal{P}} \mathbb{E}^Q[\xi x]. \quad (7)$$

to be the uncertainty of a random variable x .

Applying the definition of uncertainty to the case of portfolio returns, $x = \theta R$, in the general case where the investor has multiple sources of information, for the portfolio θ ,

$$\Delta(\theta) = \sup_v \theta^\top v \quad (8)$$

subject to

$$\mathbb{E}[\xi_{J_k} \ln \xi_{J_k}] = \frac{1}{2} v^\top \hat{\Omega}_{J_k}^{-1} v \leq \eta_k, \quad k = 1, \dots, K. \quad (9)$$

Intuitively, if η_k , $k = 1, \dots, K$, represent the confidence level, then $[-\Delta(\theta), \Delta(\theta)]$ is the corresponding confidence interval that the true mean of the portfolio falls into.

The value function $\Delta(\theta)$ is independent of ϕ . Thus, our definition of uncertainty reflects the properties of the set \mathcal{P} of candidate probability measures, not the preferences of the decision maker. Moreover, only the “shape” of the set \mathcal{P} is important in determining the relative uncertainty of various portfolios. Scaling all constraints η_k by the same constant, thus preserving the shape of the set \mathcal{P} , has no effect on the measure of uncertainty.

We will denote a solution of (8) by $v(\theta)$. Note that the solution may not be unique in general, with multiple values of v corresponding to the same value of the objective function. The following lemma shows that when all portfolio weights are non-zero, which is the case for the market portfolio in equilibrium, the solution of (8) is indeed unique.⁵

Lemma 1 *For θ such that all of its components are non-zero, the solution of (8) is unique. There exists a set of nonnegative coefficients $\phi_k(\theta)$ depending on θ such that*

$$v(\theta) = \Omega_u(\theta)\theta, \tag{10}$$

where

$$\Omega_u(\theta) = \left(\sum_{k=1}^K \phi_k(\theta) \hat{\Omega}_{J_k}^{-1} \right)^{-1}.$$

The coefficient $\phi_k(\theta)$ is equal to zero if the k th constraint is not binding, but at least one of the coefficients is strictly positive.

2.4 Diversification of Uncertainty

In this section we summarize some of the properties of our measure of uncertainty, drawing a parallel with the variance as a measure of risk (return variance is the appropriate measure of risk in our model, since asset returns are jointly normally distributed).

⁵One of the typical features of the multi-prior expected utility model is that the solution of the maxmin problem is often not unique. The analytical feature of our formulation of the set $\mathcal{P}(P, \phi)$ is that, due to Lemma 1, the minimizer for the equilibrium situation we are considering is always unique. The crucial property of the set $\mathcal{P}(P, \phi)$ that gives rise to this uniqueness is the strict convexity of the relative entropy function, as can be seen in the proof of lemma 5.

The definition of portfolio uncertainty $\Delta(\theta)$ given in (8) implies that it is a convex and symmetric function of the portfolio composition, $\Delta(-\theta) = \Delta(\theta)$, just as the variance of portfolio returns.

As with risk, one can draw a distinction between the total uncertainty of an asset (or a portfolio) and its systematic uncertainty. The systematic uncertainty of the asset i with respect to a portfolio θ is defined as its marginal contribution to the total portfolio uncertainty, in analogy with the definition of systematic risk:

$$\Delta_i(\theta) = \frac{\partial \Delta(\theta)}{\partial \theta_i}.$$

The following lemma shows that $\Delta_i(\theta)$ is well defined, as long as all components of the portfolio θ are non-zero and characterizes the sensitivity of the portfolio uncertainty to its composition.

Lemma 2 *Assuming that all components of the portfolio weights vector θ are non-zero, the sensitivity of the uncertainty of a portfolio to a change in its composition is given by*

$$\frac{\partial \ln \Delta(\theta)}{\partial \theta} = \frac{1}{\Delta(\theta)} v(\theta) = \frac{\Omega_u(\theta)\theta}{\theta^\top \Omega_u(\theta)\theta}. \quad (11)$$

This lemma implies in particular that systematic uncertainty of the market portfolio is equal to its total uncertainty. Also, since $v(\theta) \in \mathcal{V}(\phi)$, it is immediate that the total uncertainty of an asset exceeds its systematic uncertainty, i.e.,

$$\Delta(e_i) = \max_{v \in \mathcal{V}(\phi)} e_i^\top v \geq \frac{\partial \Delta(\theta)}{\partial \theta_i} = e_i^\top v(\theta),$$

where $e_i = (0, \dots, 0, 1, 0, \dots, 0)^\top$ with 1 being in the i th component of the vector.

In the above, we have considered the sensitivity of portfolio uncertainty to a change in the composition of the portfolio when the portfolio weights are non-zero. This corresponds to the case when the portfolio already has a loading of all the assets. The other interesting case is when an asset is not in the portfolio to begin with, but is to be added to the portfolio. As the following lemma shows, this case is not as simple as the other case and the reason is that $\Delta(\theta)$ is in general no longer differentiable.

Lemma 3 *Let θ be a portfolio with $\theta_j = 0$. Let $\mathcal{K} = \{k : j \in J_k\}$. If there exists a solution \bar{v} of (8) such that for all $k \in \mathcal{K}$,*

$$\frac{1}{2}\bar{v}^\top \hat{\Omega}_{J_k}^{-1} \bar{v} = \frac{1}{2}\bar{v}_{J_k}^\top \Omega_{J_k}^{-1} \bar{v}_{J_k} < \eta_k, \quad (12)$$

then $\Delta(\theta)$ is not differentiable in θ_j at $\theta_j = 0$. Otherwise $\Delta(\theta)$ is differentiable in θ_j at $\theta_j = 0$ and $\partial\Delta(\theta)/\partial\theta_j = \bar{v}_j$, where \bar{v} is any solution of (8).

The intuition of this lemma is best illustrated with the following example. There are two assets and two sources of information, one for each asset,

$$\frac{1}{2}v_j^2\sigma_j^2 \leq \eta_j, \quad j = 1, 2.$$

Let $\theta = (\theta_1, \theta_2)$ be a portfolio where $\theta_1 > 0$ and $\theta_2 = 0$. In this case,

$$\Delta(\theta) = \sqrt{2\eta_1}\sigma_1\theta_1.$$

and the solutions of (8) are of the form, $\bar{v} = (\sqrt{2\eta_1}\sigma_1, \bar{v}_2)$ where \bar{v}_2 is arbitrary as long as it satisfies the constraint above. According to the lemma, $\partial\Delta(\theta)/\partial\theta_2$ at $\theta_2 = 0$ does not exist.

While this example is special, the insight revealed applies more generally. Notice that, when $\theta_2 = 0$, the source of information about the second asset is irrelevant for the uncertainty of the portfolio. In other words, the source of information is not reflected in the portfolio uncertainty when $\theta_2 = 0$. The moment when θ_2 becomes positive, this source of information starts to contribute to the uncertainty of the portfolio. The rate at which it adds to the uncertainty of the portfolio is given by $\sqrt{2\eta_2}\sigma_2$. This rate is $-\sqrt{2\eta_2}\sigma_2$ when θ_2 becomes negative. As a result, $\Delta(\theta)$ is not differentiable in θ_2 . More generally, when the information about a particular asset has not been fully reflected, which is what (12) characterizes, the rates at which an asset contributes to the uncertainty of the portfolio differ, depending on whether the asset is added in a long or short position, and non-differentiability arises.

Interestingly, this potential non-differentiability has an equilibrium implication for the bid and ask spread of an asset price. See Routledge and Zin (2002). However, we will assume in the rest of the paper, except in section 5, that all assets are in positive supply.

3 Portfolio Choice

Using the utility function introduced above, the investor's utility maximization problem is

$$\sup_{\theta} \min_{v \in \mathcal{V}(\phi)} \mathbb{E} [\xi u(W)], \quad (13)$$

subject to the wealth constraint

$$W = W_0[\theta^\top (R - r\mathbf{1}) + 1 + r].$$

Without loss of generality we set $W_0 = 1$. The following proposition shows that the solution of the minimization is given by the solution $v(\theta)$ of (8).

Proposition 1 *Problem (13) is equivalent to*

$$\max_{\theta} \min_{|y| \leq \phi \Delta(\theta)} \{ \mathbb{E} [\xi(\theta, y) u(W)] \}, \quad (14)$$

where

$$\xi(\theta, y) = \exp \left\{ -\frac{y^2}{2\theta^\top \Omega \theta} - \frac{y(\theta^\top R - \theta^\top \mu + y)}{\theta^\top \Omega \theta} \right\}.$$

is the density of the return on the portfolio θ . Furthermore, if (θ, v) is the solution of (13) and θ is such that all of its components are non-zero, then

$$v = \phi v(\theta). \quad (15)$$

Moreover, the optimal portfolio policy θ satisfies

$$\mathbb{E} [u'(W - \phi \Delta(\theta)) (\theta^\top R - r\mathbf{1} - \phi v(\theta))] = 0.$$

The following figure illustrates, for the case where the individual is risk neutral, the trade-off between expected return and uncertainty implicit in the proposition above. In the figure, each dot corresponds to a portfolio. The horizontal coordinate of the point is the uncertainty of the portfolio, $\Delta(\theta)$; and the vertical coordinate of it is the expected return of

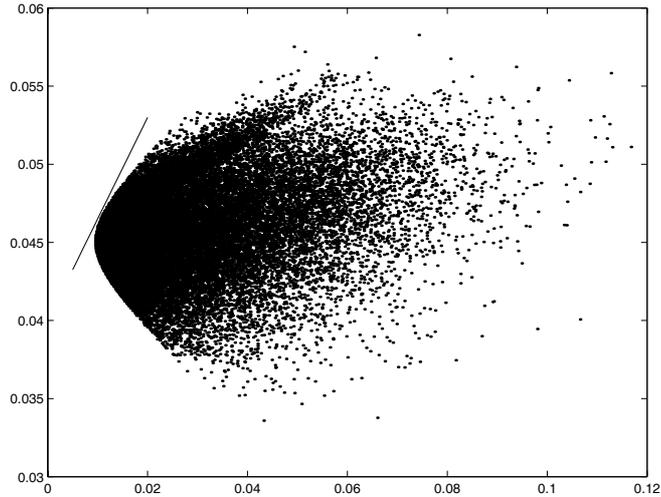


Figure 1: Mean uncertainty frontier.

the portfolio. In analogy with mean-sd diagram, we can define mean-uncertainty frontier as the collection of portfolios that solve

$$\min_{\theta} \Delta(\theta),$$

subject to $\theta^\top \mu = a$, for some a . The straight line is an indifference curve. The slope of the indifference curve is ϕ . Thus, analogous with the case of risk, an uncertainty averse investor will always choose portfolios on the upper boundary. The more uncertainty averse the investor, the steeper the indifference curve and hence the more compensation in expected return he requires for each unit increase in uncertainty.

More generally, the trade-off implicit in Proposition 1 is a three-way trade-off among expected return, risk and uncertainty of the portfolio. In the case of CARA utility, the indifference surface is given by

$$-\gamma[\theta^\top \mu - \phi \Delta(\theta) + r(1 - \theta \mathbf{1}) + 1] + \frac{\gamma^2}{2} \theta^\top \Omega \theta.$$

4 The Equilibrium

The definition of equilibrium for our economy is that of the standard rational expectations equilibrium extended to account for the fact that the objective probability law is not known. Specifically, the econometricians provide an estimate of the probability law of the (exogenous) dividend vector and a set of possible alternatives (at certain confidence level). Through the equation

$$R_j = \frac{D_j}{p_j} - 1, \quad j = 1, \dots, N,$$

this translates, for a fixed price vector $p = (p_1, \dots, p_N)$, to an estimated law for the returns and a set of possible alternatives. Taking these as given,⁶ the investors determine their asset demands. The equilibrium arises if the price vector $p = (p_1, \dots, p_N)$ is such that the markets for all assets clear.

4.1 Risk Premium and Uncertainty Premium

Let θ_m denote the market portfolio and $\Delta_m = \Delta(\theta_m)$. Define

$$\zeta = \frac{u'(W - \phi\Delta_m)}{\mathbb{E}[u'(W - \phi\Delta_m)]}$$

Then, according to Proposition 1,

$$\mathbb{E}[\zeta R] = r1 + \phi v(\theta_m), \quad (16)$$

and hence

$$\mathbb{E}[\zeta R_m] = r + \phi\Delta_m. \quad (17)$$

By applying Stein's Lemma to (16) and (17), we find that the expected return premia on the individual stocks and on the market are given by

$$\mu - r = \gamma \text{cov}(R_m, R) + \phi v(\theta) \quad (18)$$

$$\mu_m - r = \underbrace{\gamma \sigma_m^2}_{\text{risk}} + \phi \underbrace{\Delta_m}_{\text{uncertainty}} \quad (19)$$

⁶Earlier, for expositional convenience, we expressed everything in terms of returns.

The first term in (19) may be viewed as the market risk premium, being proportional to the variance of the market portfolio. The proportionality coefficient depends on the preferences of the representative agent, i.e., the absolute risk aversion coefficient of the agent, γ . We will denote the first term by λ . The second term, $\phi\Delta(\theta)$, has a natural interpretation of the market uncertainty premium, given by the product of the uncertainty aversion parameter and the degree of uncertainty of the market portfolio. We will denote it by λ_u .

Equations (16) and (17) imply a relation between expected excess returns on individual assets, which we state as the following theorem.

Theorem 2 *The equilibrium vector of expected excess returns is given by*

$$\mu - r\mathbf{1} = \lambda\beta + \lambda_u\beta_u, \quad (20)$$

where λ and λ_u are the market risk and uncertainty premia and β and β_u are the risk and uncertainty betas with respect to the market:

$$\begin{aligned} \beta &= \frac{\partial \ln \sigma^2(\theta_m)}{\partial \theta} = \frac{1}{\sigma_m^2} \Omega \theta, \\ \beta_u &= \frac{\partial \ln \Delta(\theta_m)}{\partial \theta_m} = \frac{1}{\Delta_m} \Omega_u(\theta_m) \theta_m. \end{aligned}$$

In the theorem β defines the vector of market risk betas of stocks, i.e., their betas with respect to the market portfolio. As stated in the theorem, an equivalent definition of the market risk beta is as sensitivity of the total risk of the market portfolio to a change in its composition, i.e., $\beta = \partial \ln \sigma^2(\theta_m) / \partial \theta$. The definition of the market uncertainty betas β_u is analogous. According to Lemma 2, β_u defines the sensitivity of the uncertainty of the market portfolio to a change in its composition. Note that the uncertainty betas depend only on the relative uncertainty of various portfolios and not on the uncertainty aversion of the representative agent. Scaling the constraint set \mathcal{P} by multiplying η_k 's by the same constant has no effect on β_u . We also find that, like risk, uncertainty is partially “diversifiable” in a sense that for a particular asset only its contribution the total market uncertainty is compensated in equilibrium by higher expected return.

In equilibrium, the investor is compensated for bearing both risk and uncertainty. Thus, two assets with the same beta with respect to the market risk can have different equilibrium expected returns. This not only sets our model apart conceptually from the standard CAPM, but also points to the empirical relevance of our model. To elaborate, consider first the case where there is a single source of information. In this case,⁷

$$\Omega_u = \frac{\sqrt{2\eta}}{\sigma_\theta} \Omega,$$

where θ is the equilibrium market portfolio, and hence

$$\mu - r1 = \gamma\Omega\theta + \phi\frac{\sqrt{2\eta}}{\sigma_\theta}\Omega\theta = \left(\gamma + \phi\frac{\sqrt{2\eta}}{\sigma_m}\right)\sigma_m^2\beta.$$

Since the risk aversion coefficient γ is not observable, the cross-sectional distribution of expected asset returns in a world with a single source of information will be observationally indistinguishable from that in a world where there is no model uncertainty. In fact, it follows from the equation above that

$$\mu - r1 = \beta(\mu_m - r).$$

Thus the standard CAPM holds.

Note that in the case of a single source of information, the reason that the uncertainty premium is observationally indistinguishable from the risk premium is that the two are proportional to each other in the cross-section. When there is more than one source of information, this is no longer the case (Section 4.3 contains an example) and hence the observational equivalence no longer holds. Therefore, by studying the cross-section of asset returns, one can potentially test for the existence of uncertainty premia.

4.2 Two Sources of Information

To help illustrate the implications of theorem (2), we consider an important special case where in addition to information about the joint distribution of all N assets, there is an additional source of information about the joint distribution of the first J assets. For instance,

⁷See Section 4.3.

it may be that the historical sample of returns on the first J assets is longer than the overall sample and hence there is less uncertainty about the expected returns on these assets. In this case, the uncertainty of a portfolio θ is defined as

$$\Delta(\theta) = \sup_v \theta^\top v$$

subject to

$$\frac{1}{2}v^\top \hat{\Omega}^{-1}v \leq \eta \quad (21)$$

$$\frac{1}{2}v^\top \hat{\Omega}_J^{-1}v \leq \eta_J \quad (22)$$

Letting ϕ and ϕ_J denote the Lagrange multipliers on the constraints (21) and (22) respectively, the uncertainty matrix is given by

$$\Omega_u = \frac{1}{\phi} \left(\hat{\Omega}^{-1} + \frac{\phi_J}{\phi} \hat{\Omega}_J^{-1} \right)^{-1}$$

It is straightforward to verify that

$$\Omega_u = \frac{1}{\phi} \left(\Omega - \frac{\phi_J}{\phi + \phi_J} \Omega \hat{\Omega}_J^{-1} \Omega \right)$$

Given the explicit form of the uncertainty matrix, the uncertainty beta β_u in Theorem 2 is given by

$$\beta_u = \frac{1}{\Delta_m} \left(\Omega \theta_m - \frac{\phi_J}{\phi + \phi_J} \Omega \hat{\Omega}_J^{-1} \Omega \theta_m \right)$$

The first term, $\Omega \theta_m$ is a vector of covariances of returns with the market portfolio and is proportional to the vector of standard risk betas. The second term is proportional to the vector of covariances of returns with a portfolio with weights $\pi = (1^\top \hat{\Omega}_J^{-1} \Omega \theta_m)^{-1} \hat{\Omega}_J^{-1} \Omega \theta_m$. Such a portfolio has a simple intuitive interpretation. The return on the portfolio with weights π is a linear projection of the market return on the space of the first J assets.

Thus, in presence of an additional source of information about the first J risky assets, the equilibrium vector of expected excess returns can be described by a two-factor model,

in which the first factor is the market portfolio and the second factor is a projection of the market return on the subset of the first J assets.

The special case studied here can be naturally related to the size effect documented in the literature (Banz (1981)). Anecdotal evidence suggests that larger companies tend to be followed by larger numbers of analysts than smaller companies. It is reasonable to assume that, when a company is followed by a larger number of stock analysts, there is less uncertainty about its expected stock return than when it is followed by fewer stock analysts. Think of the stocks in the group J as those followed by large numbers of analysts so that there is an additional source of information about the uncertainty of the expected returns of these stocks. Then the equilibrium expected returns of these stocks are given by, for $j \in J$,

$$\mu_j - r = \lambda\beta + \lambda_u \frac{1}{\Delta_m} \Omega \theta_m - \lambda_u \frac{\phi_J}{\phi + \phi_J} \Omega \hat{\Omega}_J^{-1} \Omega \theta_m.$$

The equilibrium expected returns of the stocks not in J are given by, for $j \notin J$,

$$\mu_j - r = \lambda\beta + \lambda_u \frac{1}{\Delta_m} \Omega \theta_m.$$

Thus small stocks enjoys higher premia than large stocks with the same market risk.

The analysis above can be readily extended to the case of multiple nested sources of information. Suppose that $J_1 \subset \dots \subset J_N$ with J_1 corresponding to the stocks with the largest number of analysts and those not in $\cup_k J_k$ with the smallest number of analysts. It is readily verified that in this case,

$$\Omega_{\mathbf{u}} = \frac{1}{\phi} \left(\Omega - \sum_{k=1}^N \epsilon_k \Omega \hat{\Omega}_{J_k}^{-1} \Omega \right)$$

where $\epsilon_k > 0$, $k = 1, \dots, N$, are constants, and

$$\beta_{\mathbf{u}} = \frac{1}{\Delta_m} \left(\Omega \theta_m - \sum_{k=1}^N \epsilon_k \Omega \hat{\Omega}_{J_k}^{-1} \Omega \theta_m \right).$$

Therefore, for the same market risk, stocks in J_1 enjoy the lowest premia and hence have the lowest cost of capital, whereas those other stocks enjoy higher equity premium with those not in $\cup_k J_k$ having the highest premia.

4.3 Independent Sources of Information

To gain further understanding of the implications of theorem (2), we consider the special case where the N risky assets can be divided into K groups, with investor having a separate source of information about each group. Without loss of generality, assume that the first N_1 assets are in the first group, the next N_2 in the second group, and so on.

Lemma 4 *If the K sources of information are independent, then*

$$\phi_k(\theta) = \frac{1}{\sqrt{2\eta_k}}\sigma_k, \quad k = 1, \dots, K, \quad (23)$$

where σ_k is the standard deviation of returns on the portfolio P_k of assets in group k combined with their market portfolio weights.

As a result of this lemma, the model uncertainty matrix simplifies to a block diagonal matrix,

$$\Omega_{\mathbf{u}} = \begin{bmatrix} \frac{\sqrt{2\eta_1}}{\sigma_1}\Omega_{J_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\sqrt{2\eta_K}}{\sigma_K}\Omega_{J_K} \end{bmatrix} \quad (24)$$

Intuitively, the block-diagonal form of $\Omega_{\mathbf{u}}$ could arise if the agent had separate models for returns on each group of assets, e.g., different models for returns on fixed income securities, stocks, and commodities or a different model of returns on equity in the United States, Japan and Europe. After all, it is common practice in academic research to specify, estimate, and test the models of individual classes of assets independently of each other. This would imply that if the uncertainty faced by the agent about the model of returns on the group of assets k were to change, it would have no effect on the amount of uncertainty remaining about the model of returns on any other group of assets. Then Theorem 2 implies that the market uncertainty beta of an asset j from the asset group k is given by

$$\beta_{\mathbf{u},j} = \frac{\sqrt{2\eta_k}\sigma_k}{\sum_{n=1}^K \sqrt{2\eta_n}\sigma_n} \beta_{P_k,j},$$

where σ_k^2 is the variance of returns on the portfolio P_k of assets in group k combined with their market portfolio weights and $\beta_{P_k,j}$ is the beta of returns on asset j with respect to such portfolio. The market price of uncertainty is given by

$$\lambda_u = \phi^2 \sum_{k=1}^K \sqrt{2\eta_k} \sigma_k$$

As a result, we have the following.

Corollary 3 *If the uncertainty matrix has a block-diagonal form (24), the cross-section of expected returns on the assets in group k is characterized by*

$$\mu_j = r + \lambda\beta_j + \phi^2 \sqrt{2\eta_k} \sigma_k \beta_{P_k,j}. \quad (25)$$

Thus, the cross-sectional differences in returns within each group of assets can be described by the assets' loadings on two "factors" – the aggregate market portfolio and the value-weighted portfolio of assets within the corresponding group.

The relation (25) could be tested empirically using the standard cross-sectional methodology. Note that (25) implies that the second factor used in addition to the market is specific to the group of assets under consideration. The presence of the second factor distinguishes our model from the standard static CAPM. The pricing relation (25) is also distinct from dynamic, multi-factor models, in which *all* assets earn risk premium as compensation for their covariation with the systematic risk factors. Under model uncertainty, the asset's expected return is affected by its correlation with the portfolio P_k only if such asset is subject to the same source of model uncertainty as other assets in that portfolio. Moreover, factors in the standard intertemporal pricing model earn excess return because they could be used to hedge against changes in the investment opportunity set. This does not have to be the case under model uncertainty. In our static model, the investment opportunity set cannot change by assumption, yet portfolios P_k appear to serve as pricing factors within the corresponding group of assets.

Using (25), one can also identify a number of restrictions across the asset groups. For instance, a within-the-group cross-sectional regression of returns on the market betas and the

group-portfolio betas should recover the two coefficients: λ and $\phi^2\sqrt{2\eta_k}\sigma_k$. One can then test whether the estimates of the market risk premium λ are identical across the groups. Moreover, since $\sum_{k=1}^K \phi^2\sqrt{2\eta_k}\sigma_k = \lambda_u$, one could compare the resulting estimate of $\lambda + \lambda_u$ with the direct estimate of the expected return on the market portfolio.

5 The Effect of Changes in Uncertainty

Another way to highlight the effect of uncertainty on asset prices is by performing a comparative statics experiment of increasing the degree of uncertainty in the model.

All zero net supply derivative securities written on the primitive assets can be priced using the risk-neutral probability density.⁸ The risk-neutral density is given by

$$\frac{f(R)\xi(R)u'(W)}{\int f(R)\xi(R)u'(W) dR},$$

where R is the vector of equilibrium returns on the primitive securities, W is the end-of-period wealth, and $\xi(R)$ is the probability density corresponding to the equilibrium value of $v(\theta)$. Then the price of any security with payoff $X(R)$ is given by

$$\frac{1}{1+r} \frac{\int f(R)\xi(R)u'(W)X(R) dR}{\int f(R)\xi(R)u'(W) dR}.$$

Let Ω_D denote the variance-covariance matrix of dividends. Similarly, define $\Omega_{uD} = \left(\sum_{k=1}^K \phi_k \hat{\Omega}_{D,J_k}^{-1}\right)^{-1}$. Then the risk-neutral probability density can be expressed as

$$(2\pi)^{-n/2} |\Omega_D|^{-1/2} \exp \left\{ -\frac{1}{2} (D + v_D - E[D])^\top \Omega_D^{-1} (D + v_D - E[D]) \right\} \\ \times \exp \left\{ -\gamma \left(D + v_D - E[D] - \frac{1}{2} \gamma^2 \mathbf{1}^\top \Omega_D \mathbf{1} \right) \right\},$$

where

$$v_D = \exp \left\{ -\gamma(1+r) - \frac{\gamma^2}{2} \mathbf{1}^\top \Omega_D \mathbf{1} \right\} \Omega_{uD} \mathbf{1}$$

⁸This is because the introduction of these derivative securities does not introduce additional uncertainty.

This expression for the risk-neutral density implies that an increase in uncertainty, i.e., an increase in Ω_{uD} , leads to a shift in the mean of the risk-neutral distribution. The effect is particularly easy to visualize in a one risky asset case. An increase in uncertainty results in a downward shift in the mean of the risk-neutral distribution, as illustrated in Figure 2.

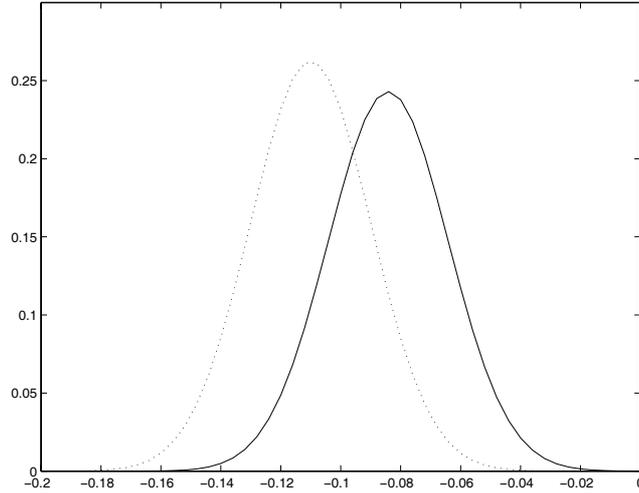


Figure 2: Risk-neutral distributions. The dotted line corresponds to the economy with a higher degree of uncertainty.

It is instructive to compare this behavior with an increase in prior uncertainty in the standard Bayesian framework. The Bayesian approach assumes uncertainty neutrality, whereas our approach assumes uncertainty aversion. This is best illustrated using Ellsberg experiment and its following variant: in the second urn the number of red ball is between 10 and 90 so that the probability of drawing red is between 0.1 and 0.9. In this case, the Bayesian approach would still assign 0.5 to drawing red and hence be indifferent between a bet on the second urn and that on the first urn, just as in the original Ellsberg experiment. Thus even though the amount of uncertainty is different in these two experiment, the Bayesian approach makes no distinction. Note that in this experiment the mean of the posterior distribution remains unchanged in the Bayesian approach.

Increased prior variance in the Bayesian framework results in an increase in the variance of the risk-neutral distribution, while in our model an increase in uncertainty would shift the mean of the risk-neutral distribution downward, as illustrated in Figure 3. Thus, an increase in model uncertainty would have different implications for the prices of derivative securities

relative to an increase in prior variance in a Bayesian model. Prices of out-of-the-money call options fall as model uncertainty increases. However, this may not be the case in the Bayesian framework.

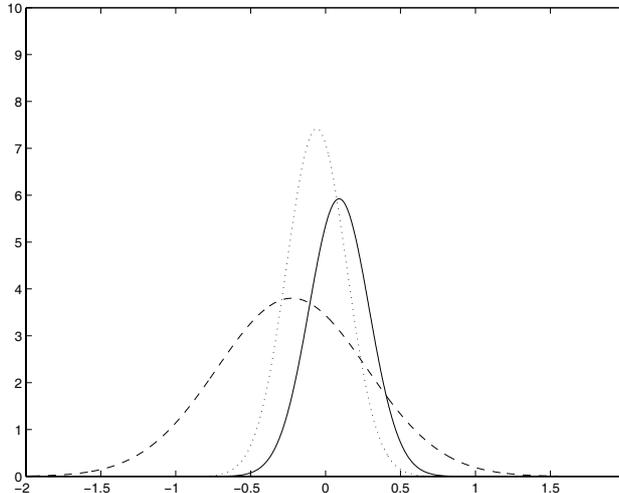


Figure 3: Risk-neutral distributions. The dotted line corresponds to the economy with uncertainty aversion and a higher degree of uncertainty. The dashed line corresponds to the Bayesian economy with increased prior uncertainty.

6 Conclusion

We have developed a single-period equilibrium model incorporating, not only risk, but also uncertainty and uncertainty aversion. We have shown that there is an uncertainty premium in equilibrium expected returns on financial assets. In particular, the cross-sectional distribution of expected returns can be formally described by a two-factor model, where expected returns are derived as compensation for the asset's contribution to the equilibrium risk and uncertainty of the portfolio held by the agent. We were able to derive several empirically testable implications of this result. While prior research on model uncertainty has been concerned with its implications for the time-series of asset prices, by characterizing the cross-section of returns we were able to address some of the observational equivalence issues raised in the literature. In particular, we demonstrated that the effect of model uncertainty in our framework is distinct from risk aversion and cannot be captured by any specification of the risk aversion parameter.

Appendix

Proof of Lemma 1

Suppose to the contrary that \bar{v} and \underline{v} are two distinct solutions. Let $v(a) = a\bar{v} + (1-a)\underline{v}$. The strict convexity of all the functions defining the choice set implies that for $a \in (0, 1)$,

$$\frac{1}{2}v(a)^\top \hat{\Omega}_{J_k}^{-1}v(a) \leq \eta_k, \quad k = 1, \dots, K.$$

Now let k , if exists, be such that

$$\frac{1}{2}v(a)^\top \hat{\Omega}_{J_k}^{-1}v(a) = \eta_k$$

holds for $a = 0$, $a = 1$, and for some $a \in (0, 1)$. Then it must be the case that $\bar{v}_{J_k} = \underline{v}_{J_k}$.

Denote by A the set of such k . If

$$J_A = \cup_{k \in A} J_k = \{1, \dots, n\},$$

then $\bar{v} = \underline{v}$, a contradiction to assumption. So, $J_A \neq \{1, \dots, n\}$. Without loss of generality, we assume that $J_A = \{2, \dots, n\}$. Then for all v of the form $v = (v_1, \bar{v}_2, \dots, \bar{v}_n)$ with $v_1 \in R$,

$$\frac{1}{2}v^\top \hat{\Omega}_{J_k}^{-1}v = \eta_k, \quad k \in A.$$

Note that $v(a)$ is of the form $(a\bar{v}_1 + (1-a)\underline{v}_1, \bar{v}_2, \dots, \bar{v}_n)$. Thus for $v = (0.5\bar{v}_1 + 0.5\underline{v}_1, \bar{v}_2, \dots, \bar{v}_n)$,

$$\frac{1}{2}v^\top \hat{\Omega}_{J_k}^{-1}v < \eta_k, \quad k \notin A.$$

Combining the two cases, $k \in A$ and $k \notin A$, together, by continuity, there is a $\epsilon > 0$ such that for all $v = (v_1, \bar{v}_2, \dots, \bar{v}_n)$ with $v_1 \in (0.5\bar{v}_1 + 0.5\underline{v}_1 - \epsilon, 0.5\bar{v}_1 + 0.5\underline{v}_1 + \epsilon)$,

$$\frac{1}{2}v^\top \hat{\Omega}_{J_k}^{-1}v \leq \eta_k, \quad k = 1, \dots, K.$$

But, given the linearity of the objective function, this means \bar{v} and \underline{v} cannot be the solution of (8). This is a contradiction.

The second statement of the lemma is a straightforward application of the Lagrangian duality approach.

Proof of Lemma 2

Since the constraint set \mathcal{P} is convex and compact, $\Delta(\theta)$ is a convex function. Optimality conditions imply that $\phi^{-1}v(\theta)$ is a subgradient of the value function $\Delta(\theta)$ at θ . The solution v of is unique, according to lemma 1. Thus, the function $\Delta(\theta)$ has a unique subgradient, therefore it is in fact differentiable, and $\phi^{-1}v$ is equal to the gradient of $\Delta(\theta)$. This establishes the statement of the lemma.

Proof of Lemma 3

Observe that $\partial\Delta(\theta)/\partial\theta_j$ exists if and only if all solutions of (8) have the same j th component. For the first claim of the lemma, assume without loss of generality that $j = 1$. If the condition of the lemma is satisfied, there exists a $\epsilon > 0$ such that for any $|x| < \epsilon$, $v_x = \bar{v} + (x, 0, \dots, 0)$ satisfies all constraints of (8). Since $\theta_1 = 0$, v_x is also a solution of (8). The claim follows.

For the second part, let \bar{v} be a solution of (8). If it is the unique solution of (8), then $\partial\Delta(\theta)/\partial\theta_j$ exists. Suppose \bar{v} and \underline{v} are two distinct solutions of (8). Let $v(a) = a\bar{v} + (1-a)\underline{v}$. We claim that there exists a $k \in \mathcal{K}$ such that

$$\frac{1}{2}v(a)_{J_k}^\top \Omega_{J_k}^{-1} v(a)_{J_k} = \eta_k$$

holds for $a = 0$, $a = 1$ and some $a \in (0, 1)$. Suppose the contrary. By strict convexity,

$$\frac{1}{2}v(a)_{J_k}^\top \Omega_{J_k}^{-1} v(a)_{J_k} < \eta_k, \quad k \in \mathcal{K}$$

for $a \in (0, 1)$. Also the convexity of all the functions defining the choice set implies that for $a \in (0, 1)$,

$$\frac{1}{2}v(a)^\top \hat{\Omega}_{J_k}^{-1} v(a) = \frac{1}{2}v(a)_{J_k}^\top \Omega_{J_k}^{-1} v(a)_{J_k} \leq \eta_k, \quad k = 1, \dots, K.$$

Since the objective function of (8) is linear, $v(a)$ is a solution of (8) for all $a \in (0, 1)$. But this is a contradiction to assumption of the lemma. Thus the claim is shown. It then follows from the claim that $\bar{v}_{J_k} = \underline{v}_{J_k}$ and hence $\bar{v}_j = \underline{v}_j$. Since \bar{v} and \underline{v} are arbitrary, we have $\bar{v}_j = \underline{v}_j$ for all solutions of (8). The differentiability follows. \blacksquare

Proof of Lemma 1

Since the distribution of W depends only on the distribution of $\theta^\top R$, for each fixed θ , $E[\xi u(W)]$ depends only on $y = \theta^\top v$, and it is given by

$$E[\xi u(W)] = E[\xi(\theta, y)u(W)],$$

where

$$\xi(\theta, y) = \exp \left\{ -\frac{y^2}{2\theta^\top \Omega \theta} - \frac{y(R_\theta - \theta^\top \mu + y)}{\theta^\top \Omega \theta} \right\}.$$

Thus the original utility function can be written as

$$\max_{\theta} \min_{|y| \leq \phi \Delta(\theta)} (E[\xi(\theta, y)u(W)]) \tag{A1}$$

which is (14). The characterization for v follows immediately.

References

- [1] Anderson, E., L. Hansen and T. Sargent (1999): "Robustness, Detection and the Price of Risk," working paper, University of Chicago.
- [2] Banz, R. (1981): "The Relation between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9, 3-18.
- [3] Barberis, N. (2000): "Investing for the long run when returns are predictable," *Journal of Finance* 55, 225-264.
- [4] Balduzzi, P., and L. Liu (1999): "Estimation risk, learning, and international investment," Working Paper, Boston College.
- [5] Brav, A. and J.B. Heaton (2002): "Competing Theories of Financial Anomalies," *Review of Financial Studies*, 15, 575-606.
- [6] Breeden, D. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265-296.
- [7] Brennan, M. (1998): "The role of learning in dynamic portfolio decisions," *European Finance Review* 1, 295 -306.
- [8] Brennan, M. and Y. Xia (2001): "Stock Return Volatility and Equity Premium," *Journal of Monetary Economics*, 47, 249-283.
- [9] Chen, Z. and L. Epstein (2001): "Ambiguity, risk and asset returns in continuous time," *Econometrica*, 70, 1403-1443.
- [10] Cox, J., J. Ingersoll, and S. Ross (1985): "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, 363-384.
- [11] Detemple, J. (1986): "Asset pricing in an economy with incomplete information," *Journal of Finance* 61, 383-392.
- [12] Dow, J. and S. Werlang (1992): "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio", *Econometrica*, 60, pp.197-204.
- [13] Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, pp.643-669.
- [14] Epstein, L. and J. Miao (2001): "A two-person dynamic equilibrium under ambiguity," forthcoming in *Journal of Economic Dynamics and Control*.
- [15] Epstein, L. and T. Wang (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62, pp.283-322.
- [16] Epstein, L. and T. Wang (1995): "Uncertainty, Risk-Neutral Measures and Security Price Booms and Crashes," *Journal of Economic Theory*, 67, pp.40-80.

- [17] Frost, P. and J. Savarino, (1986): “An empirical Bayes approach to efficient portfolio selection,” *Journal of Financial and Quantitative Analysis* 21, 293 -305.
- [18] Gennotte, G. (1986): “Optimal portfolio choice under incomplete information,” *Journal of Finance* 41, 733-746.
- [19] Gilboa, I. and D. Schmeidler (1989): “Maxmin Expected Utility Theory with Non-Unique Prior,” *Journal of Mathematical Economics*, 18, pp.141-153.
- [20] Hansen, L. and T. Sargent (2001): “Robust Control and Model Uncertainty,” *American Economic Review*, Vol. 91, pp.60-66.
- [21] Kandel, S., and R. Stambaugh (1996): “On the predictability of stock returns: An asset allocation perspective,” *Journal of Finance* 51, 385-424.
- [22] Knight, F. (1921): *Risk, Uncertainty and Profit*, Houghton, Mifflin, Boston.
- [23] Lowellen, J. and J. Shanken (2001): “Learning, Asset-Pricing Tests, and Market Efficiency,” *Journal of Finance*, 57, 1113-1145.
- [24] Lucas, R. (1978): “Asset prices in an exchange economy,” *Econometrica* 46, 1429-1445.
- [25] Machina, M. and D. Schmeidler (1992): “A More Robust Definition of Subjective Probability,” *Econometrica*, 60, pp.745-780.
- [26] Maenhout, P. (1999): “Robust portfolio rules and asset pricing,” Working paper, Harvard University.
- [27] Merton, R. (1992): *Continuous-Time Finance*. Blackwell Publishers, Cambridge MA.
- [28] Pástor, L. (2000): “Portfolio selection and asset pricing models,” *Journal of Finance* 55, 179-223.
- [29] Routledge, B. and S. Zin (2002): “Model Uncertainty and Liquidity,” working paper, Carnegie Mellon University.
- [30] Schmeidler, D. (1989): “Subjective Probability and Expected Utility without Additivity,” *Econometrica*, 57, pp.571-587.
- [31] Sharpe, W. (1964): “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance*, 19, 425-442.
- [32] Uppal, R. and T. Wang (2001): “Model Misspecification and Under-Diversification,” forthcoming in *Journal of Finance*.
- [33] Xia, Y. (2001): “Learning about Predictability: the Effect of Parameter Uncertainty on Dynamic Asset Allocation,” *Journal of Finance* 56, 205-246.