Learning and Asset Prices under Ambiguous Information

Markus Leippold          Fabio Trojani
Paolo Vanini

First version: September 2003
Current version: July 2006

This research has been carried out within the NCCR FINRISK project on “New Methods in Theoretical and Empirical Asset Pricing”
Learning and Asset Prices Under Ambiguous Information*

Markus Leippold\textsuperscript{a}  Fabio Trojani\textsuperscript{b}  Paolo Vanini\textsuperscript{a,c}

\textsuperscript{a}Swiss Banking Institute, University of Zurich, Switzerland
\textsuperscript{b}Swiss Institute of Banking and Finance and Department of Economics, University of St. Gallen, Switzerland
\textsuperscript{c}Zurich Cantonal Bank, Switzerland

(First version: September 2003; This version: July 2, 2006)

* We are grateful to Andrew Abel and an anonymous referee for many valuable suggestions. We also thank Freddy Delbaen, David Feldman, Günter Franke, Rajna Gibson, Jens Jackwerth, Yvan Lengwiler, Marco LiCalzi, Abraham Lioui, Alessandro Sbuelz, Sandra Sizer, Pietro Veronesi, Yihong Xia, the participants of the 2004 European Summer Symposium in Financial Markets in Gerzensee, the 2005 International Finance Conference at the University of Copenhagen, the EFMA 2006 (Madrid), and of the finance seminars at the University of Basel, the University of Konstanz, the University of Frankfurt, the University of Venice, the University of Zurich, and the ETH Zurich. The authors gratefully acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK and grants 101312-103781/1 and 100012-105745/1) and the University Research Priority Program “Finance and Financial Markets” of the University of Zurich.

Correspondence:
Markus Leippold, Swiss Banking Institute, University of Zurich, Plattenstrasse 14, 8032 Zurich, Switzerland, \texttt{leippold@isb.unizh.ch}.
Fabio Trojani (corresponding author), Swiss Institute of Banking and Finance, University of St. Gallen, Rosenbergstr. 52, 9000 St. Gallen, Switzerland, e-mail: \texttt{fabio.trojani@unisg.ch}.
Paolo Vanini, Swiss Banking Institute, University of Zurich, Plattenstrasse 14, 8032 Zurich, Switzerland, e-mail: \texttt{paolo.vanini@zkb.ch}
Learning and Asset Prices under Ambiguous Information

Abstract

In a Lucas exchange economy with standard power utility, we study asset prices under learning and ambiguous information. If the representative investor is ambiguity averse, a discount for ambiguity arises only for a relative risk aversion below one or, equivalently, an intertemporal elasticity of substitution above one. The presence of both learning and ambiguity enforces large equity premia, and supports model predictions that are consistent with well-known asset pricing puzzles. For realistic amounts of ambiguity, the absence of learning or ambiguity aversion implies low volatilities or low equity premia.

Keywords: Financial Equilibria, Learning, Knightian Uncertainty, Ambiguity Aversion, Model Misspecification.

JEL Classification: C60, C61, G11.
The equity premium puzzle, the interest rate puzzle, and excess volatility are well-known empirical failures of the consumption capital asset pricing model pioneered by Lucas (1978) and Breeden (1979). In this paper, we present an extension of the Lucas exchange economy that helps explain these puzzles in a setting of learning and ambiguity (Knightian uncertainty) aversion.

In our economy, we equip the representative investor with a standard CRRA utility. The investor has only partial information on the dynamics of the aggregate endowment process, as in Veronesi (2000). However, in contrast to rational settings of Bayesian learning,\(^1\) we describe these dynamics with a set of multiple likelihoods that model the ambiguity of the underlying endowment process. Although there is evidence that to some extent, ambiguity aversion can explain the equity premium and the low interest rate puzzles in full information economies, we also know from Veronesi (2000) that in a learning setting, the equity premium is even more of a puzzle. Therefore, we ask if, in our simple exchange economy, the interaction of learning and ambiguity aversion can explain the equity premium puzzle.

We also know that pure learning settings can explain excess volatility and volatility clustering of asset returns. At the same time, models of ambiguity aversion with a constant opportunity set do not substantially affect the expected equity returns and equity volatility (see, e.g., Maenhout (2004) and Sbuelz and Trojani (2002)). Thus, we analyze the combination of learning and ambiguity aversion to see if that combination still generates excess volatility and volatility clustering.

First, we find that learning under ambiguity aversion implies an equilibrium discount for ambiguity if and only if the relative risk aversion is less than one, or, equivalently, the elasticity of intertemporal substitution (EIS) is above one. For low risk aversion, learning and ambiguity aversion increase conditional equity premia and volatilities. We also show that the part of the equity premium that is due to the interaction of learning and ambiguity aversion is the dominant part.

Second, learning and ambiguity aversion imply lower equilibrium interest rates, regardless of risk aversion. Thus, with low risk aversion we get both a higher equity premium and a lower interest rate. This finding is an important feature of our setting, because it explains both the equity premium and the interest rate puzzles for moderate amounts of ambiguity.

Third, the true theoretical equilibrium relation between excess returns and conditional variances is highly time-varying. This feature generates estimated relations between excess returns and conditional variances that have undetermined signs over time, and a time-varying bias in the naively estimated risk-return tradeoff using, e.g., regression methods.

The paper is organized as follows. The next section reviews the literature on ambiguity aversion. Section 2 introduces our model. In Section 3, we study the properties of the optimal learning dynamics. Section 4 characterizes and discusses conditional asset pricing relations. Section 5 concludes.

1. Background

From the Ellsberg (1961) paradox, we know that investors behave differently under ambiguity and risk aversion. Therefore, distinguishing between ambiguity aversion and risk aversion is economically and behaviorally important. Gilboa and Schmeidler (1989) suggest an atemporal axiomatic framework of ambiguity aversion in which preferences are represented by max-min expected utility over a set of multiple prior distributions. Inspired by their approach, Epstein and Wang (1994) study some asset pricing implications of max-min expected utility in a discrete-time infinite horizon economy. Epstein and Schneider (2003) provide a discrete-time axiomatic foundation for that model, and show that a dynamically consistent version of the Gilboa and Schmeidler (1989) preferences can be represented by a recursive max-min expected utility over a set of multiple distributions. Chen and Epstein (2002) extend the model of Epstein and Schneider (2003) to continuous time. Hansen, Sargent, and Tallarini (1999, in discrete time) and Anderson, Hansen, and Sargent (2003, in continuous time) propose another model of intertemporal ambiguity aversion based on an alternative form of max-min expected utility.

Various authors propose models of full information economies with ambiguity aversion to explain several important characteristics of asset prices. In these models, the investor observes perfectly the state variables that determine the opportunity set, but she is not fully aware of the probability distribution of the state variables. Consequently, some form of conservative worst-case optimization determines her optimal decision rules. Examples of such models include, among others, Gagliardini, Porchia, and Trojani (2004, term structure of interest rates), Epstein and Miao (2003, home bias), Liu, Pan, and Wang (2004, option pricing with rare events),
mium, excess volatility, and low interest rate puzzles simultaneously. Moreover, Maenhout (2004) and Sbuelz and Trojani (2002) show that it is difficult to fit the equity premium puzzle in a setting with standard power utility without imposing an unrealistically high degree of ambiguity.

Only more recently have a few authors addressed the issue of learning under ambiguity aversion. In a production economy driven by a two-state Markov chain, Cagetti, Hansen, Sargent, and Williams (2002) use robust filtering theory to study the impact of learning and ambiguity aversion on the aggregate capital stock, equity premia, and price-dividend ratios. Using numerical methods, they show that the precautionary saving increases in a way similar to the effect of a higher subjective time preference rate. The equity premium increases and price-dividend ratios turn out to be lower.

Our model differs from Cagetti, Hansen, Sargent, and Williams (2002) in several aspects. We work with a Lucas exchange economy and use a tractable homothetic setting of preferences under ambiguity aversion. Because our model is tractable, we can study and analyze in closed form all the relevant asset-pricing relations and their dependence on model parameters. For instance, we show that in the partial information Lucas economy, ambiguity aversion can fail to increase equity premia if standard risk aversion is too high or, equivalently, the EIS is too low. We also allow for external public signals and for heterogeneous ambiguity across the possible states of the economy. These extensions have important implications for asset prices. For example, we find that ambiguity premia caused by public signals can be quite large and that there is an ambiguity premium for information quality.

Epstein and Schneider (2002) use a discrete-time model to highlight that learning about an unknown parameter under multiple likelihoods can fail to resolve ambiguity asymptotically, even when the underlying state process is not subject to regime shifts. In a similar model, Epstein and Schneider (2004) illustrate the impact of an ambiguous signal precision on asset prices. Using numerical methods, they show that in a risk-neutral economy, ambiguous information quality, which they define by a set of possible values of the signal precision parameter, can generate skewed asset returns and excess volatility.

Our paper is different from Epstein and Schneider (2004). Their focus is on ambiguous signal precision and its impact on return volatility and skewness, but we focus on ambiguous signals about

the expected dividend growth and the resulting implications primarily for equity premia. In this context, we show that for low risk aversion and a realistic amount of ambiguity, the joint interaction of learning and ambiguity aversion is responsible for large equity premia. Additionally, the model predictions are consistent with the interest rate and excess volatility puzzles. Since we compute equilibrium quantities under standard assumptions on the utility function, we can also disentangle the impact of learning, ambiguity aversion, and risk aversion on equity premia.

2. The Model

We develop an equilibrium setting of learning under ambiguity aversion. Our model consists of three main components: a parametric reference model for the unobservable dividend drift, a set of multiple likelihoods for the dynamics of dividends, and a max-min expected utility optimization problem for the representative investor.\(^3\)

2.1. The Reference Model

We consider a Lucas (1978) economy populated by a CRRA investor with utility function

\[
u(C,t) = e^{-\delta t} \frac{C^{1-\gamma}}{1-\gamma},\]

where \(\gamma > 0\) is the coefficient of relative risk aversion.\(^4\) We give the representative investor a parametric reference model, which is an approximate description of the dynamics of dividends \(D\).

To focus on the genuine implications of learning and ambiguity aversion, we assume a reference model in which dividends follow a geometric Brownian motion, as in Veronesi (2000).

**Definition 1** Under the reference model probability \(\mathbb{P}\), dividends have the dynamics

\[
\frac{dD}{D} = \theta dt + \sigma_D dB_D,
\]

where \(\theta \in \Theta := \{\theta_1, \theta_2, ..., \theta_n\}\), with \(\theta_1 < \theta_2 < ... < \theta_n\), \(\sigma_D > 0\) and \(B_D\) is a \(\mathbb{P}\)-Brownian motion.

\(^3\)See also Gilboa and Schmeidler (1989), Chen and Epstein (2002), Epstein and Schneider (2003), and Knox (2005).
The reference model is a rough approximation of reality. Therefore, the representative investor has some motivated doubts about the specification of the reference model. The representative investor observes dividends and a noisy unbiased signal $e$ for the dividend drift. Under the reference model probability, the signal dynamics are

$$de = \theta dt + \sigma_e dB_e$$

where $\sigma_e > 0$ and $B_e$ is a $\mathbb{P}$–Brownian motion independent of $B_D$.

In the standard Bayesian framework, Definition 1 implies a single likelihood for the dividend dynamics. The specific value of the parameter $\theta$ is unknown and the only relevant statistical uncertainty about the dividend dynamic (2) is parametric. Therefore, the standard filtering process leads to asymptotic learning of the unknown dividend drift $\theta$ in the class of candidate drift values $\Theta$.

We depart from the above setting by allowing for a misspecification in the reference model. Misspecifications can have a general nonparametric form, so that they cannot be consistently detected even if we use approaches based on parametric Bayesian model selection.

### 2.2. Multiple Likelihoods

Compared with the Bayesian single-likelihood hypothesis in Definition 1, an investor with multiple likelihood beliefs is less ambitious. She uses a set of multiple likelihoods to compute a set of relevant beliefs about the unknown dividend drift. This set of beliefs represents the investor’s ambiguity about the unobservable expected dividend growth rate. It can also be interpreted as a class of alternative specifications of the reference model.

**Assumption 1**

(i) The set of multiple likelihoods consists of absolutely continuous probabilities $\mathbb{P}^h(\theta)$, such that the dividend dynamics are

$$\frac{dD(t)}{D(t)} = (\theta + h(\theta, t)\sigma_D)dt + \sigma_D dB_D(t), \quad t \geq 0,$$
for some $\theta \in \Theta$ and some drift distortion $\theta + h(\theta) \sigma_D \in \Xi(\theta)$. (ii) Under any probability $\mathbb{P}^h(\theta)$ signals are unbiased:

$$de(t) = (\theta + h(\theta, t)\sigma_D)dt + \sigma_e dB_e(t), \quad t \geq 0 .$$

(iii) The representative investor has some beliefs $(\hat{\pi}_1, \ldots, \hat{\pi}_n)$ at time $t = 0$ about the a priori plausibility of the different sets $\Xi(\theta_1), \ldots, \Xi(\theta_n)$ of drift distortions $h$.

Under Assumption 1, the representative investor recognizes that statistically, the whole neighborhood $\Xi(\theta)$ of dividend drift changes is hardly distinguishable from a zero drift change, i.e., from the reference model dynamics with drift $\theta$ in Definition 1. If $\Xi(\theta) = \{0\}$ for all $\theta \in \Theta$, the Bayesian setup of Veronesi (2000) follows and the investor is concerned exclusively with the noisiness of the signal on the parameter value $\theta$.

The size of the neighborhood $\Xi(\theta)$ describes the degree of ambiguity associated with any possible reference model drift $\theta$. The broader $\Xi(\theta)$, the more ambiguous are the signals about a specific dividend drift $\theta + h(\theta) \sigma_D \in \Xi(\theta)$. This ambiguity reflects the fact that there are aspects of the unobservable dividend dynamics that are hardly possible, or even impossible, to ever know. Therefore, our representative investor is aware that empirically, identifying the exact functional form of a possible mean reversion in the dividend drift dynamics is a virtually infeasible task.\textsuperscript{4} Accordingly, the investor tries to understand only a limited number of features of the dividend dynamics.

We specify $\Xi(\theta)$ by ensuring that it contains likelihood specifications that are statistically close, in some appropriate statistical measure of model discrepancy, to the reference model. Since we constrain the relevant misspecifications to be small, they are hardly statistically detectable. This property defines a whole neighborhood of small, but otherwise arbitrary, misspecifications of the reference model.

\textbf{Assumption 2} For any $\theta \in \Theta$ we define $\Xi(\theta)$ by

$$\Xi(\theta) := \left\{ \theta + h(\theta) \sigma_D : \frac{1}{2} h^2(\theta, t) \leq \eta(\theta) \text{ for all } t \geq 0 \right\} ,$$

\textsuperscript{4}Shepard and Harvey (1990) show that in finite samples, it is very difficult to distinguish between a purely iid process and one which incorporates a small persistent component.
where $\eta(\theta_1), \ldots, \eta(\theta_n) \geq 0$. Moreover, for any $i \neq j$,

$$\Xi(\theta_i) \cap \Xi(\theta_j) = \emptyset . \quad (7)$$

Assumption 2 states that for any probability $P^h(\theta)$ implied by an admissible likelihood, the maximal drift distortion relative to the reference model is bounded by:

$$\theta - \sqrt{2\eta(\theta)\sigma_D} \leq \theta + h(\theta, t)\sigma_D \leq \theta + \sqrt{2\eta(\theta)\sigma_D} . \quad (8)$$

Therefore, by using the parameter $\eta$ we can control the size of the set of multiple likelihoods of the representative investor and parameterize in a parsimonious way her perception of the ambiguity of dividends. For $\eta = 0$ ambiguity vanishes and we obtain the standard Bayesian setting. For increasing $\eta$, we obtain settings in which ambiguity is more pronounced. Since equation (6) does not make specific assumptions on the parametric functional form of $h(\theta)$, the set of likelihoods implied by the neighborhood $\Xi(\theta)$ is nonparametric and there is no simple way we can put a prior on it.

Condition (7) means that the representative investor is confronted with ambiguity about candidate drifts within neighborhoods, but not between neighborhoods. So, for instance, we can map different macroeconomic conditions into disjoint sets of likely drift dynamics. We can then study a setting in which the reference model drifts describe the approximate average growth rate of the economy under structurally different conditions, such as booms or recessions. Conditional on a boom or a recession, there is a whole neighborhood of similar models for dividends that are very difficult to distinguish statistically. The size of this neighborhood can be a function of the relevant economic state. For example, we can use this additional degree of freedom to model a different degree of statistical uncertainty across the relevant structural states of the economy.

2.3. The Size of Ambiguity

Given a set of multiple likelihoods, we have to ask how we can impose realistic maximal drift distortions $\theta + h(\theta)\sigma_D \in \Xi(\theta)$ relative to the reference model.
Given the reference belief, condition (7) already imposes a constraint on the size of $\eta(\theta)$. Therefore, we can only select a moderate size for the ambiguity parameter $\eta(\theta)$, relative to the distances between reference model drifts $\theta_1, \ldots, \theta_n$. In addition, we note from condition (7) that the size of $\eta(\theta)$ determines the degree of pessimism implied by a worst-case dividend drift $\theta - \sqrt{2\eta(\theta)}\sigma_D$ in the neighborhood $\Xi(\theta)$. Given the model parameters, we can compute the quantity $\sqrt{2\eta(\theta)}\sigma_D$. In this way, we can study the difference between the expected consumption growth rate under the reference belief and under the worst-case belief. In all our subsequent model calibrations, we use this approach to constrain the size of $\eta(\theta)$ in a way that avoids worst-case beliefs that are too pessimistic. Table 1 reports the differences between the reference belief and the worst-case belief about the expected consumption growth rate. For all calculations, we use $\sigma_D = 0.0325$, which corresponds to the volatility of consumption growth as estimated in Campbell (1999) for an annual sample from 1891 to 1994. For the given choices of ambiguity $\eta(\theta)$ in Table 1, the differences between the reference belief and the worst-case belief of the expected consumption growth rate are always less than 0.46%.

Finally, from a statistical viewpoint, we can make use of an argument proposed by Anderson, Hansen, and Sargent (2003) to constrain the degree of pessimism in our model. They quantify the statistical closeness of two likelihood beliefs by calculating the average error probability in a Bayesian likelihood ratio test of two competing models, given a continuous record of data of length $N$. Intuitively, likelihoods that are statistically close must be associated with large error probabilities, but likelihoods that are easy to distinguish have to imply low error probabilities.

We use this idea of Anderson, Hansen, and Sargent (2003) to select ambiguity parameters that are consistent with nontrivial model detection error probabilities in a test of the implied worst-case likelihood against the likelihood of the reference model. Let $l_N(\theta)$ be the log-likelihood function of the worst-case likelihood relative to the likelihood of the reference model associated with neighborhood $\Xi(\theta)$. Treating these likelihoods symmetrically, the average probability of a model detection error in the corresponding likelihood ratio test is

$$
\epsilon_N(\theta) = 0.5 \cdot P(l_N(\theta) > 0) + 0.5 \cdot P^{\text{tr}}(l_N(\theta) < 0),
$$

(9)
where \( h^* = -\sqrt{2\eta(\theta)} \) is the standardized worst case drift distortion associated with neighborhood \( \Xi(\theta) \). Using the geometric Brownian motion structure of dividends under the reference model and the worst case likelihood, it follows that

\[
\epsilon_N(\theta) = \Phi \left( -\sqrt{\frac{\eta(\theta)N}{2}} \right),
\]

where \( \Phi \) is the cumulative distribution function of the standard normal distribution. We note that \( \epsilon_N(\theta) \) is decreasing in \( N \) and \( \eta(\theta) \). Hence, for a given sample size, we can identify large model discrepancies more easily and, for a given size of ambiguity, we can more frequently detect model errors as the sample size increases.

In Table 1, we report model detection error probabilities for different sizes of ambiguity and for realistic sample sizes. For the largest choice of \( \eta(\theta) \) and a century of annual observations \( (N = 100) \), the error probability in Table 1 is no less than 24% and for a century of quarterly observations \( (N = 400) \) no less than 8%. In both cases, the average error probability is nontrivial and above the significance level typically used to reject a null hypothesis in many empirical studies.

### 2.4. Ambiguity Aversion and Intertemporal Max-Min Expected Utility

We use \( \mathcal{F}(t) \) to denote the information available at time \( t \). This information contains all possible realizations of dividends and signals. \( P \) is the price of the risky asset in the economy, \( r \) the instantaneous interest rate and \( \eta(\theta) \) the function that describes ambiguity. The representative investor determines consumption and investment plans \( C(t) \) and \( w(t) \) by solving the intertemporal max-min expected utility optimization problem

\[
\max_{C,w} \inf_{h(\theta)} \left\{ \int_0^\infty u(C,s) \, ds \mid \mathcal{F}(0) \right\},
\]

subject to the dividend and wealth dynamics

\[
dD = (\theta + h(\theta)\sigma_D)Ddt + \sigma_DDdB_D, \\
dW = W \left[ w \left( \frac{dP + Ddt}{P} \right) + (1 - w)r dt \right] - Cdt,
\]
where for any $\theta \in \Theta$ the standardized drift distortion is such that $\theta + h(\theta) \sigma_D \in \Xi(\theta)$ and Assumption 2 holds.

The max-min problem (11) models the representative investor’s optimal behavior, given her attitudes to risk and ambiguity. In addition to an optimal consumption/investment policy, the representative investor must select an optimal worst-case belief $\theta + h(\theta)\sigma_D$ out of the admissible class $\Xi(\theta)$. The fact that such an optimal belief is determined endogenously as a function of preferences differs sharply from the standard Bayesian setting, in which beliefs are fixed by a parametric assumption on the unobservable dynamics of the dividend drift process.

3. Asset Prices

We first study the representative investor’s optimal learning behavior under ambiguity. With the endogenous set of optimal dividend-drift prediction processes at hand, we solve the equilibrium optimization problem and compute asset prices.

3.1. Learning Under Ambiguity Aversion

To derive the equilibrium asset prices, we first need to determine the endogenous set of optimal dividend-drift prediction processes. Learning under ambiguity requires a set of Bayesian ahead beliefs as functions of drift distortions $\theta + h(\theta) \sigma_D \in \Xi(\theta)$. For a given likelihood $h(\theta)$, let $\pi_i(t)$ be the investor’s belief that the drift rate is $\theta_i + h(\theta_i) \sigma_D$, conditional on past dividends and signals, i.e.,

$$\pi_i(t) = \mathbb{P}(\theta + h(\theta)\sigma_D = \theta_i + h(\theta_i)\sigma_D | \mathcal{F}(t)) \quad .$$

(12)

The distribution $\Pi(t) := (\pi_1(t), \ldots, \pi_n(t))$ summarizes the investor’s beliefs at time $t$, under the likelihood $h(\theta)$. Given such beliefs, the investor can compute the expected dividend drift at time $t$ as

$$m_{\theta,h} := m_{\theta} + m_{h(\theta)} := \sum_{i=1}^{n} (\theta_i + h(\theta_i)\sigma_D) \pi_i(t) \quad ,$$

(13)

where

$$m_{\theta} = \sum_{i=1}^{n} \theta_i \pi_i(t) \quad , \quad m_{h(\theta)} = \sum_{i=1}^{n} h(\theta_i) \pi_i(t) \sigma_D \quad .$$

(14)
The filtering equations implied by any likelihood $h(\theta)$ are standard (see Liptser and Shiryaev (2001)).

**Lemma 1** Suppose that at time zero, the investor’s beliefs are represented by the prior probabilities $\pi_1, \ldots, \pi_n$. Under a likelihood $h(\theta)$, the dynamics of the optimal filtering probabilities vector $\pi_1, \ldots, \pi_n$ are given by

$$d\pi_i = \pi_i (\theta_i + h(\theta_i) \sigma_D - m_{\theta,h}) \left( k_D d\tilde{B}^h_D + k_e d\tilde{B}^h_e \right) ; \quad i = 1, \ldots, n,$$

(15)

where

$$d\tilde{B}^h_D = k_D (dD/D - m_{\theta,h} dt) , \quad d\tilde{B}^h_e = k_e (de - m_{\theta,h} dt) , \quad k_D = 1/\sigma_D , \quad k_e = 1/\sigma_e .$$

In this equation, $(\tilde{B}^h_D, \tilde{B}^h_e)$ is a $\mathbb{P}^h(\theta)$-standard Brownian motion in $\mathbb{R}^2$ with respect to the filtration $\{\mathcal{F}(t)\}$.

Which learning behavior should the investor adopt in an ambiguous environment? Since the investor is not particularly comfortable with a specific element of $\Xi(\theta)$, she will base her beliefs on the whole set of likelihoods implied by $\Xi(\theta)$. This approach generates a whole class $\mathcal{P}$ of dividend-drift prediction processes, which are given by

$$\mathcal{P} = \{ m_{\theta,h} : \theta + h(\theta) \sigma_D \in \Xi(\theta) \} ,$$

(16)

where the dynamics of the posterior probabilities $\pi_1, \ldots, \pi_n$ under the likelihood $h(\theta)$ are given in (15). The set $\mathcal{P}$ represents investor’s ambiguity about the true dividend-drift process, conditional on the available information $\mathcal{F}(t)$ generated by dividends and signals. The larger the size of $\Xi(\theta)$, i.e., the ambiguity about the dividend dynamics, the larger the set $\mathcal{P}$ of dividend-drift prediction processes.

Using the set $\mathcal{P}$, we write the continuous-time optimization problem (11) as a full information problem in which we define the relevant dynamics in terms of the filtration $\{\mathcal{F}(t)\}$. Since all beliefs implied by likelihoods $\mathbb{P}^h(\theta)$ are absolutely continuous, all relevant processes $(\tilde{B}^h_D, \tilde{B}^h_e)'$ generate the same filtration $\{\mathcal{F}(t)\}$, and the dynamic budget constraint associated with the optimization problem (11) can be equivalently formulated in terms of $m_{\theta,h}$ and $(\tilde{B}^h_D, \tilde{B}^h_e)'$. (See Miao (2001) for a related discussion).
3.2. Equilibrium with Learning and Ambiguity Aversion

In our economy, an equilibrium is a vector of processes \((C(t), w(t), P(t), r(t), h(\theta, t))\) such that the optimization problem (11) is solved and markets clear, i.e., \(w(t) = 1\) and \(C(t) = D(t)\). In equilibrium, the relevant problem reads

\[
J(\Pi, D) = \inf_{h(\theta)} E \left[ \int_0^{\infty} e^{-\delta t} \left( \frac{D(t)^{1-\gamma}}{1-\gamma} \right) dt \right], \tag{17}
\]

subject to the dynamics

\[
dD = m_{\theta,h} D dt + \sigma_D d\tilde{B}^D_h, \tag{18}
\]

\[
d\pi_i = \pi_i (\theta_i + h(\theta_i) \sigma_D - m_{\theta,h}) \left( k_D d\tilde{B}_D^h + k_e d\tilde{B}_e^h \right), \tag{19}
\]

where for any \(\theta \in \Theta\) we have \(\theta + h(\theta) \sigma_D \in \Xi(\theta)\) and Assumption 2 holds.

The key difference to the standard Bayesian setting is that in equation (17) the investor must select the optimal worst-case forecast procedure for the unknown dividend drift. This worst-case belief selection generates an endogenous systematic discrepancy between the reference model belief and the belief applied by the investor to value financial assets.

In Proposition 1, we solve problem (17) and compute the impact of ambiguity aversion on the price of equity and the equilibrium interest rate.

**Proposition 1** Let \(\hat{\theta}_i := \delta + (\gamma - 1) \theta_i + \gamma (1 - \gamma) \frac{\sigma^2}{2}\) and assume that

\[
\hat{\theta}_i + (1 - \gamma) \sqrt{2\eta(\theta_i)} \sigma_D > 0, \quad i = 1, \ldots, n. \tag{20}
\]

It then follows:

a) The normalized misspecification \(h^*(\theta)\) that solves problem (17) is given by:

\[
h^*(\theta_i) = -\sqrt{2\eta(\theta_i)}, \quad i = 1, \ldots, n. \tag{21}
\]
b) The equilibrium price function $P(\Pi, D)$ for the risky asset is given by

$$P(\Pi, D) = D \sum_{i=1}^{n} \pi_i C_i ,$$

(22)

where

$$C_i = 1/\left(\hat{\theta}_i + (1 - \gamma) \sqrt{2\eta(\theta_i)\sigma_D}\right) , \quad i = 1, \ldots, n .$$

(23)

c) The equilibrium interest rate $r$ is

$$r = \delta + \gamma m_{\theta,h^*} - \frac{1}{2} \gamma (\gamma + 1) \sigma_D^2 ,$$

(24)

where $m_{\theta,h^*} = m_{\theta} + m_{h^*}(\theta)$ is such that

$$m_{h^*}(\theta) = \sum_{i=1}^{n} h^*(\theta_i) \pi_i \sigma_D = - \sum_{i=1}^{n} \sqrt{2\eta(\theta_i)\pi_i \sigma_D} .$$

(25)

The constant $C_i$ in equation (23) is proportional to the representative investor’s expectation of discounted lifetime dividends, conditional on the worst-case drift $\theta_i - \sqrt{2\eta(\theta_i)\sigma_D}$ selected from the neighborhood $\Xi(\theta_i)$:

$$C_i = E^{h^*(\theta_i)} \left[ \int_{s}^{\infty} e^{-\delta(t-s)} \left( \frac{D(t)}{D(s)} \right)^{1-\gamma} \, dt \right] = \frac{1}{D(s)} E^{h^*(\theta_i)} \left[ \int_{s}^{\infty} \frac{u_c(D(t), t)}{u_c(D(s), s)} D(t) \, dt \right] ,$$

(26)

where $E^{h^*(\theta_i)} [\cdot]$ denotes the expectation under a geometric Brownian motion process for $D$ having drift $\theta_i - \sqrt{2\eta(\theta_i)\sigma_D}$.

A high $C_i$ implies that the representative investor is willing to pay a high price for the ambiguous state $\Xi(\theta_i)$. Since she cannot exactly identify a particular state, the price of the risky asset is a weighted sum, in which $C_i$ is weighted by the posterior probability of the neighborhood $\Xi(\theta_i)$. $C_i$ is a function of both ambiguity aversion, via the parameter $\eta(\theta_i)$, and the relative risk aversion $\gamma$.

Remark 1 We can also describe equation (26) as

$$C_i = E \left[ \int_{s}^{\infty} e^{-\left[\delta+(1-\gamma)\sqrt{2\eta(\theta_i)\sigma_D}\right](t-s)} \left( \frac{D(t)}{D(s)} \right)^{1-\gamma} \, dt | \theta = \theta_i \right] ,$$

(27)
where $E [ \cdot | \theta = \theta_i ]$ denotes the reference model expectation conditional on a constant drift $\theta = \theta_i$. Therefore, the impact of ambiguity aversion on the price of equity in the ambiguous state $\Xi(\theta_i)$ is equivalent to the impact of an adjusted time preference rate

$$\delta \rightarrow \delta + (1 - \gamma) \sqrt{2\eta(\theta_i)}\sigma_D. \quad (28)$$

The adjustment depends on the amount of ambiguity of neighborhood $\Xi(\theta_i)$, risk aversion $\gamma$ and dividend-growth volatility $\sigma_D$. Cagetti, Hansen, Sargent, and Williams (2002) give numerical evidence that in their production economy ambiguity aversion decreases the aggregate capital stock in a way that is similar, but not identical under general power utility, to the effect of an increased subjective discount rate. In our model, when ambiguity is homogeneous across neighborhoods, its effect on the price of equity is identical to the impact of an increased time preference rate. In the general case of a heterogeneous degree of ambiguity, the final effect on asset prices cannot be mapped into an adjustment of the time preference rate.

In Corollary 1 we summarize the dependence of the price $C_i$ of an ambiguous neighborhood $\Xi(\theta_i)$ on the ambiguity parameter $\eta(\theta_i)$.

**Corollary 1** The price of any ambiguous state $\Xi(\theta_i)$ is a decreasing function of the degree of ambiguity $\eta(\theta_i)$ if and only if $\gamma < 1$. In such a case, $C_i$ is a convex function of $\eta(\theta_i)$, which is more convex for smaller risk aversion $\gamma$.

From Corollary 1, the marginal relative price of ambiguity is negative if and only if the relative risk aversion $\gamma$ is less than one. If $\gamma > 1$, then we obtain the somewhat counterintuitive implication that the price of an ambiguous state is higher than the one of an unambiguous state. This implication contradicts the basic intuition provided by the Ellsberg (1961) paradox. To clarify this finding, we note that $C_i$ is determined by the representative investor’s worst-case marginal indirect utility of future dividends. In equilibrium, a lower dividend growth rate implies both a lower consumption growth and a higher stochastic discount factor. Since the last effect dominates for high risk aversion, a lower expected dividend growth implied by the worst-case belief under ambiguity aversion yields a higher price for ambiguous states.
For $\gamma > 1$, settings of learning and ambiguity aversion deliver both low equity premia and interest rates that are high and highly variable. Therefore, we focus on settings with low risk aversion.\footnote{There is some experimental evidence favoring low risk aversion under ambiguity, collected by Wakker and Deneffe (1996), who estimate a virtually linear utility function when using a utility elicitation procedure robust to the presence of ambiguity. In these experiments, utility functions estimated by procedures that are not robust to the presence of ambiguity are clearly concave.}

**Assumption 3** The representative investor has a relative risk aversion parameter $\gamma < 1$.

Since we adopt a setting with power utility, Assumption 3 is equivalent to assuming an elasticity of intertemporal substitution $1/\gamma > 1$. This assumption is consistent with the intuition that in our model excess returns mainly reflect some premium for ambiguity, rather than a premium for risk. The empirical evidence on the size of the EIS is mixed. Vissing-Jorgensen (2002) estimates an EIS above one when using Euler equations for Treasury bills in a model with time-additive preferences. For the UK, Attanasio, Banks, and Tanner (2002) estimate an EIS between 0.81 and 2.54 ‘likely’ shareholders. Vissing-Jorgensen and Attanasio (2003) estimate a model with Epstein-Zin preferences using after-tax returns and find an EIS parameter between 1.03 and 1.43 for U.S. stockholders. In a model with long run risk and fluctuating economic uncertainty, Bansal and Yaron (2004) use a calibrated EIS parameter of 1.5 to justify several empirical features of asset returns.

### 3.3. Price-Dividend Ratios and Interest Rates

Under Assumption 3, we obtain from Proposition 1 a few direct implications for the behavior of the price-dividend ratio.

**Corollary 2** Let Assumption 3 be satisfied. It then follows:

a) The price-dividend ratio $P/D$ is a decreasing convex function of the amount of ambiguity $(\eta(\theta_1), \ldots, \eta(\theta_n))$. Moreover, $P/D$ is a more convex function for lower risk aversion $\gamma$.

b) A mean-preserving spread $\hat{\Pi}$ of $\Pi$ implies $\hat{P}/D > P/D$. Therefore, the price-dividend ratio is increasing in the amount of uncertainty.
Finding a) is a direct implication of (22) and (27). Finding b) follows from the convexity of $C_i$ in (22) as a function of $\theta_i - \sqrt{2\eta(\theta_i)}\sigma_D$. From Corollary 2, we see that the impact of a higher ambiguity on $P/D$ ratios and of a higher uncertainty in the economy differ in signs, implying a distinct prediction of ambiguity aversion for the behavior of $P/D$.

Table 2 illustrates the impact of ambiguity aversion on $P/D$ ratios in settings of no learning (column NL) and learning (column L), for risk aversion parameters between 0.6 and 0.9. These levels of risk aversion correspond to EIS parameters between 1.1 and 1.7. These EIS parameters are consistent with the empirical findings in the literature cited above.

Insert Table 2 about here

In Table 2, we compute $P/D$ ratios from equation (22) in a setting with three possible reference model drifts $\theta_1 = 0.005$, $\theta_2 = 0.0175$, $\theta_3 = 0.03$, and a discretized normal distributed prior $\Pi$ centered at $\theta_2$. To fix the prior dispersion, we set $\theta_1$, and $\theta_3 = \theta_2 \pm \sigma_e$, and use a signal standard deviation $\sigma_e = 0.0125$. Under these conditions, the posterior mean for expected consumption growth is $m_\theta = 0.0175$ and coincides with the U.S. average consumption growth estimated by Campbell (1999) for an annual sample spanning the period from 1891 to 1994. Again, we fix the volatility of consumption growth at $\sigma_D = 0.0325$.

With this prior and this parameter structure, we compute the model implied $P/D$ ratios by using the explicit formula (22) in Proposition 1. The first row ($\eta = 0$) in Table 2 presents the quantities that prevail in the absence of ambiguity. The entries in the following rows for $\eta$ ranging from 0.0001 to 0.01 report the $P/D$ ratios for an increasing (homogeneous) ambiguity parameter. We see that the $P/D$ ratios decrease with risk aversion. Due to the convexity of $C_i$ as a function of $\theta_i - \sqrt{2\eta(\theta_i)}\sigma_D$, these ratios are higher in a setting of learning. Ambiguity aversion lowers $P/D$ ratios monotonically, and the size of this decrease depends on the investor’s risk aversion. For the model with learning and for $\gamma = 0.6$, the $P/D$ ratios decrease from 36.27 to 33.96 for $\eta = 0$ and $\eta = 0.01$. For $\gamma = 0.9$, the corresponding $P/D$ ratio decreases very little from 30.06 to 29.65.
Since $r$ is a decreasing convex function of $(\eta(\theta_1), \ldots, \eta(\theta_n))$, learning and ambiguity aversion reduce the equilibrium interest rate. We obtain the interest rate in the Bayesian setting of Veronesi (2000) for $\eta(\theta) = 0$:

$$r = \delta + \gamma m_\theta - \frac{1}{2} \gamma (\gamma + 1) \sigma_D^2.$$  \hspace{1cm} (29)

The interest rate with ambiguity but no learning arises for a degenerate distribution $\Pi$. In this case, we have $m_\theta + m_{h^*}(\theta) = \theta_l - \sqrt{2\eta(\theta_l)\sigma_D}$ for some $\theta_l \in \Theta$ and the interest rate is equal to

$$r = \delta + \gamma \left( \theta_l - \sqrt{2\eta(\theta_l)\sigma_D} \right) - \frac{1}{2} \gamma (\gamma + 1) \sigma_D^2.$$  \hspace{1cm} (30)

When $\Pi$ is nondegenerate, the contribution of ambiguity aversion to the interest rate is summarized by the term $m_{h^*}(\theta)$, which is a weighted sum of the contributions in the single-model neighborhoods $\Xi(\theta_1), \ldots, \Xi(\theta_n)$. The weights in $m_{h^*}(\theta)$ are the posterior probabilities $\Pi$.

4. Conditional Asset Returns

Given the worst-case dividend drift $\theta - \sqrt{2\eta(\theta)\sigma_D}$ conditional on the ambiguous state $\Xi(\theta)$, we obtain the dynamics of the equilibrium equity excess return $R$, which is defined by

$$dR = \frac{dP + Ddt}{P} - rdt.$$  \hspace{1cm} (31)

**Proposition 2** (i) Under the investor’s subjective optimal worst-case belief $h^*(\theta)$ in Proposition 1, the equilibrium excess return satisfies

$$dR = \mu_R^{wc} dt + \sigma_D d\tilde{B}_D^{h^*} + V_{\theta,h^*} \left( k_D d\tilde{B}_D^{h^*} + k_e d\tilde{B}_e^{h^*} \right),$$  \hspace{1cm} (32)

where

$$\mu_R^{wc} = \gamma \left( \sigma_D^2 + V_{\theta,h^*} \right), \quad V_{\theta,h^*} = \sum_{i=1}^n \frac{\pi_i C_i \left( \theta_i - \sqrt{2\eta(\theta_i)\sigma_D} \right)}{\sum_{i=1}^n \pi_i C_i} - m_{\theta,h^*},$$  \hspace{1cm} (33)

with Brownian motion increments with respect to the filtration $\{\mathcal{F}(t)\}$ given by

$$d\tilde{B}_D^{h^*} = k_D \left( \frac{dD}{D} - m_{\theta,h^*} dt \right), \quad d\tilde{B}_e^{h^*} = k_e \left( de - m_{\theta,h^*} dt \right).$$

17
(ii) Under the reference model belief, the equilibrium excess return satisfies

\[ dR = \mu_R dt + \sigma_D d\tilde{B}_D + V_{\theta,h^*} \left( k_D d\tilde{B}_D + k_e d\tilde{B}_e \right), \tag{34} \]

where

\[ \mu_R = \mu_R^{wc} - m_h^*(\theta) \left( 1 + kV_{\theta,h^*} \right), \quad k = k_D^2 + k_e^2, \tag{35} \]

with Brownian motion increments with respect to the filtration \( \mathcal{F}(t) \) given by

\[ d\tilde{B}_D = k_D \left( \frac{dD}{D} - m_\theta dt \right), \quad d\tilde{B}_e = k_e (de - m_\theta dt). \]

By using the first part of Proposition 2, we can study the first effect of learning and ambiguity aversion on excess returns. The description of \( R \) with respect to the Brownian motions \( \tilde{B}_D^{h^*} \) and \( \tilde{B}_e^{h^*} \) yields the dynamics under the worst-case scenario \( h^*(\theta) \in \Xi(\theta) \). Thus, we can interpret the resulting expected excess return on equity as the worst-case equity premium.

The second part of Proposition 2 allows us to study the second impact of learning and ambiguity aversion on excess returns. This impact arises because of the discrepancy between the likelihood belief implied by the reference model and the optimal worst-case belief that the representative investor adopts to compute asset prices. This discrepancy generates an additional premium for ambiguity. We can analyze this second effect of learning and ambiguity aversion by describing excess returns with respect to the filtered Brownian motions \( \tilde{B}_D, \tilde{B}_e \), which are the optimal ones implied by the reference model belief. This description yields the correct excess return from the perspective of an outside observer such as an econometrician, who believes in the reference model as an approximate description of dividends and who knows that the representative investor is ambiguity averse.

To analyze excess returns in more detail, we must study the sign and the comparative statics of the quantities \( m_{h^*}(\theta) \) and \( V_{\theta,h^*} \) in equations (32) and (34). The term

\[ m_{h^*}(\theta) = m_{\theta,h^*} - m_\theta = -\sum_{i=1}^{n} \sqrt{2\eta(\theta_i)} \pi_i \sigma_D \tag{36} \]

is a correction to the reference model’s posterior mean \( m_\theta \). This correction is negative and incorporates misspecification doubts in the posterior mean of the economy’s growth rate. \( V_{\theta,h^*} \) reflects
the difference between the worst-case expected growth rate of the economy, $m_{\theta,h^*}$, and its value-adjusted counterpart. This quantity is larger when the representative investor has more diffuse beliefs about $\Xi(\theta_1), \ldots, \Xi(\theta_n)$ or when she values the asset very differently across states. Differences in valuation across states depend on the heterogeneity of the worst-case growth rate $\theta - \sqrt{2\eta(\theta)}\sigma_D$.

Under Assumption 2, it follows that:

$$\theta_1 - \sqrt{2\eta(\theta_1)}\sigma_D < \theta_2 - \sqrt{2\eta(\theta_2)}\sigma_D < \ldots < \theta_n - \sqrt{2\eta(\theta_n)}\sigma_D .$$

Therefore, we can apply the arguments in the proof of Lemma 3 in Veronesi (2000) to characterize the behavior of $V_{\theta,h^*}$.

**Lemma 2** Let Assumption 2 be satisfied. It then follows:

1. $V_{\theta,h^*}$ is a decreasing function of $\gamma$.

2. The following statements are equivalent:
   (a) Assumption 3 holds.
   (b) $V_{\theta,h^*} > 0$.
   (c) For any mean-preserving spread $\tilde{\Pi}$ of $\Pi$ it follows that

   $$\tilde{V}_{\theta,h^*} > V_{\theta,h^*} ,$$

   where "\∽" denotes quantities under $\tilde{\Pi}$.

In particular, $V_{\theta,h^*}$ is positive and increasing for mean-preserving spreads of $\Pi$ if and only if $\gamma < 1$. The positivity of $V_{\theta,h^*}$ is crucial for avoiding asset prices that are inconsistent with the equity premium puzzle predictions.

### 4.1. Equity Premia

From equation (33), see that the equity premium is the sum of three conceptually different components:

$$\mu_R = \gamma\left(\sigma_D^2 + V_{\theta,h^*}\right) - \underbrace{m_{h^*}}_{(A)} - \underbrace{m_{h^*}kV_{\theta,h^*}}_{(B)} .$$

(37)
Component (A) is both the equity premium for the standard risk exposure, i.e., the risk premium, and also the worst-case equity premium in our economy. The sum (B)+(C) is the equity premium caused by ambiguity, i.e., the ambiguity premium. (B) is the ambiguity premium caused by misspecification doubts about the dynamics of dividends. (C) is the ambiguity premium caused by misspecification doubts about the dynamics of the optimal posterior probabilities II.

**Insert Figure 1 about here**

In Figure 1, we compute the risk and equity premia in a setting with the same reference model drifts, prior structure, and parameter values as in Table 2. We plot these drifts and priors in Panel A of Figure 1. In Panel B, we also plot the sizes of the ambiguous neighborhoods $\Xi(\theta_1), \ldots, \Xi(\theta_n)$ as ellipses centered at $\theta_1, \ldots, \theta_n$. In Panel D of Figure 1, we present a typical pattern of the equity premium $\mu_R$ implied by equation (35) for a risk aversion parameter between 0.7 and 1 and for a homogeneous degree of ambiguity $\eta = 0.001$. The equity premium is a monotonically decreasing function of risk aversion. For some risk aversions strictly smaller than one, this feature turns out to be consistent with the predictions of the equity premium puzzle. For instance, for a risk aversion between 0.7 and 0.8 the equity premium ranges from 12.5% to 4.8%. This effect arises despite the very small size of the ambiguity parameter used. Following our discussion in Section 2.3, the difference between the reference and the worst-case belief is only 0.16%, which implies a worst-case expected consumption growth of

$$ m_{\theta, h^*} = m_{\theta} - \sqrt{2\eta} \sigma_D = 1.75\% - 0.16\% = 1.59\%. $$

Furthermore, for the given parameter structure with $\eta = 0.001$, the model detection error probability for the sample size used by Campbell (1999) is $\epsilon_N(\theta) = 0.41$.

### 4.1.1. Premia for Risk

Component (A) in equation (37) is the equity premium perceived by an investor under the optimal worst-case likelihood $h^* (\theta)$ in Proposition 1. From Proposition 2, we get

$$ \mu_{R}^{wc} = \gamma (\sigma_D^2 + V_{\theta, h^*}) = \gamma \text{Cov}_{t} h^* (dR, dD/D) = \gamma \text{Cov}_{t} (dR, dD/D) \quad , $$

(38)
where \( \text{Cov}^{h^*}_{t} \) and \( \text{Cov}_{t} \) denote conditional covariances under the worst-case likelihood \( h^* \) and under the reference model likelihood, respectively. The last equality in (38) arises because worst-case and reference model likelihoods are absolutely continuous. Hence, the covariance \( \text{Cov}_{t}(dR, dD/D) \) between returns and aggregate consumption growth is linked only to the fraction \( (A) \) of the whole equity premium \( \mu_R \).

In our economy, the risk premium \( (A) \) is quite small. \( \gamma \sigma_D^2 \) is the risk premium under ambiguity aversion but no learning.\(^6\) For realistic risk aversion parameters it is typically small. Moreover, Lemma 2 implies that \( V_{\theta, h^*} \) is decreasing in risk aversion. Therefore, the risk premium \( (A) \) is bounded as a function of risk aversion and is negligible for practical purposes, as in Veronesi (2000).

In Panel C of Figure 1, we plot a typical profile of the risk premium \( (A) \) (solid line) together with the risk premium implied by the standard Bayesian learning setting (dashed line). The two risk premia almost coincide. In general, the risk premium under learning and ambiguity aversion is different, because the term \( V_{\theta, h^*} \) in equation (38) is different from the one in a setting of pure learning. But for realistic structures of the function \( \eta(\theta) \) we always find that the two risk premia are numerically similar and small. For \( \gamma = 0.6 \), the risk premia in the models with and without ambiguity aversion are 0.57% and 0.59%, respectively. As emphasized by Veronesi (2000), the Mehra and Prescott (1985) equity premium puzzle is even more puzzling in a Bayesian setting, because equity premia cannot be matched by risk premia, even for very high risk aversion. This finding on risk premia is also true for our model. However, in our model, equity premia consist of premia for both risk and ambiguity.

4.1.2. Premia for Ambiguity

The equity premium in equation (37) depends on the ambiguity premia \( (B) \) and \( (C) \), both of which are positive under Assumption 3. The premium \( (B) + (C) \) is due to the discrepancy between the reference model belief and the worst-case belief used by the representative investor to update

her subjective posterior mean for the growth rate of dividends. The resulting worst-case return
dynamics in (32) depend on the two filtered random shocks,

\[ d\tilde{B}^h_D = d\tilde{B}_D - k_D m_{h^*(\theta)} dt, \quad d\tilde{B}^h_e = d\tilde{B}_e - k_e m_{h^*(\theta)} dt, \]  

(39)

which define a \( \{\mathcal{F}(t)\} \)-Brownian motion \((\tilde{B}^h_D, \tilde{B}^h_e)\) under the worst-case belief \( h^*(\theta) \), and define a \( \{\mathcal{F}(t)\} \)-Brownian motion with drift under the reference model belief. We can interpret the discrepancies

\[ \lambda^A_D := \frac{(d\tilde{B}^h_D - d\tilde{B}_D)/dt}{-k_D m_{h^*(\theta)}}, \]  

(40)

and

\[ \lambda^A_e := \frac{(d\tilde{B}^h_e - d\tilde{B}_e)/dt}{-k_e m_{h^*(\theta)}}, \]  

(41)

as the market prices of ambiguity for \( d\tilde{B}^h_D \) and \( d\tilde{B}^h_e \) shocks. The filtered shocks \( d\tilde{B}^h_D \) and \( d\tilde{B}^h_e \) influence the worst-case dynamics in equation (32) in two ways: through the impact of \( d\tilde{B}^h_D \) on the filtered dynamics for dividends and through the joint impact of \( d\tilde{B}^h_D \) and \( d\tilde{B}^h_e \) on the filtered dynamics for \( \Pi \). Let \( h = h^* \) in equations (18) and (19). It then follows:

\[ dD/D = m_{\theta,h^*} dt + \sigma_D d\tilde{B}^h_D, \]  

(42)

\[ d\pi_i/\pi_i = (\theta_i + h(\theta_i) \sigma_D - m_{\theta,h^*}) \left( k_D d\tilde{B}^h_D + k_e d\tilde{B}^h_e \right), \quad i = 1, \ldots, n. \]  

(43)

The ambiguity premium deriving from a unit shock in the filtered dynamics in (43) is

\[ k_D (d\tilde{B}^h_D - d\tilde{B}_D)/dt + k_e (d\tilde{B}^h_e - d\tilde{B}_e)/dt = k_D \lambda^A_D + k_e \lambda^A_e = -k m_{h^*(\theta)} , \]  

(44)

where \( k = k_D^2 + k_e^2 \). According to the worst-case dynamics in (32), \( V_{\theta,h^*} \) measures the reaction of equilibrium equity returns to shocks in the dynamics in (43). Therefore, component (C) in equation (37) is the ambiguity premium for the misspecification of these shocks. The ambiguity premium for a shock in the filtered dividend dynamics is given in (40). Therefore, component (B) in equation (37) is the ambiguity premium for misspecified dividend shocks.
Component (B) is positive under a degenerate $\Pi$-distribution, i.e., in the absence of learning. We can interpret it as a pure premium for ambiguity. In contrast, the ambiguity premium (C) is nonzero if and only if $\Pi$ is nondegenerate, i.e., if the representative investor has only partial information about the underlying model neighborhood $\Xi(\theta)$. Therefore, we can interpret (C) as an ambiguity premium component caused by the interaction of learning and ambiguity aversion.

To illustrate the contribution of learning and ambiguity to equity premia, we compute these quantities in Table 3 for a setting of no learning (column NL) and a setting of learning (column L). The first row ($\eta = 0$) in Table 3 presents the quantities that prevail in the absence of ambiguity. The remaining rows show the equilibrium risk and equity premia (column RP and EP) for an increasing (homogeneous) ambiguity parameter $\eta$ and for risk aversion parameters $\gamma = 0.8$ (Panel A), $\gamma = 0.85$ (Panel B), and $\gamma = 0.9$ (Panel C). The reference belief dividend drifts, prior structure, and all other model parameters are the same as those used for Figure 1.

Insert Table 3 about here

In columns RP of Table 3, we obtain small risk premia for all levels of risk aversion and for a setting both with and without learning. The risk premia $\mu^{SC}_R$ range between 0.20% and 0.33%.

In column EP, we present the equity premia. In the absence of ambiguity ($\eta = 0$), the risk and equity premia are identical. Introducing ambiguity into the model increases equity premia for both the setting of no learning (column NL) and of learning (column L). In column NL of Panel B with $\gamma = 0.85$, we observe a premium for pure ambiguity that increases the equity premia from 0.09% ($\eta = 0$) to 0.55% ($\eta = 0.01$). We can attribute this increase to the pure ambiguity premium (B), which amounts to 0.44% at $\eta = 0.01$. In relative terms, this increase is substantial. However, the resulting equity premium is still too small for practical purposes.

In contrast, in Column L of Panel B we present large equity premia that increase from 0.26% ($\eta = 0$) to 7.44% ($\eta = 0.01$) due to a large ambiguity premium (B)+(C). Since the ambiguity premium (B) is only 0.44%, the large increase is due to premium component (C) arising from the joint presence of learning and ambiguity aversion. In Panel B for $\eta = 0.01$, this component amounts to 6.89%.
For $\eta = 0.0075$ and $\eta = 0.01$, equilibrium interest rates are close to 2\% for all levels of risk aversion. The low interest is due to both ambiguity aversion and the low risk aversion. The latter induces a sufficiently strong desire to substitute consumption intertemporally, which additionally helps to obtain the low interest rate.

Insert Table 4 about here

Table 4 shows that the equity premium and the interest rates obtained in Table 3 have sizes comparable to those of the interest rate and equity premium estimated in Campbell (1999). The historical average interest rate in Table 4 is about 2\% (see column IR), and the equity premium is about 6.26\% (see column EP).

As noted above, component (B) is a pure premium for ambiguity. In the absence of learning, we could make this component arbitrarily large by increasing the parameter $\eta$ to fit the stylized facts of asset prices. However, to obtain an empirically reasonable equity premium, we would need to impose an excessive degree of pessimism by making the uncertainty parameter $\eta$ unreasonably large. In a model with ambiguity aversion only, we would have to set $\eta = 1.8$ to obtain an equity premium around 6.28\%. This ambiguity level implies a rather pessimistic worst-case expected consumption growth $m_{\theta,h^*} = -4.42\%$ and yields a model detection error probability equal to zero. Therefore, only by simultaneously modeling both learning and ambiguity aversion, we can generate a substantial equity premium with a more reasonable amount of ambiguity.

4.1.3. Is There an Ambiguity Premium for Imprecise Signals?

Consistent with the findings of Veronesi (2000), in our Lucas economy there is no quantitatively relevant risk premium for imprecise signals. However, under ambiguity aversion the key equity premium component is the ambiguity premium (C). Therefore, we ask under which conditions an ambiguity premium for imprecise signals exists.

**Corollary 3** Let Assumption 3 be satisfied and suppose that function $\sqrt{\eta(\theta)}$ is a convex function of $\theta$. Then, for any mean-preserving spread $\tilde{\Pi}$ of $\Pi$ it follows that:

$$-\tilde{m}_{h^*}(1 + k\tilde{V}_{\theta,h^*}) > -m_{h^*}(1 + kV_{\theta,h^*}), \quad (45)$$

24
where $\sim$ denotes quantities under $\tilde{\Pi}$.

Corollary 3 states that under a convex function $\sqrt{\eta(\theta)}$, there is always an ambiguity premium for information noisiness. When the realized signal precision is low, the posterior probabilities $\Pi$ are more diffuse and the quantity $V_{\theta,h^*}$ in equation (45) increases (see again Lemma 2). We can write $V_{\theta,h^*}$ as the covariance between equity excess returns $R$ and signals $e$. Therefore, the increased ambiguity premium under a less precise signal follows from the higher covariance between equity returns and public signals. This covariance amplifies the impact of the ambiguity premium (C) on the total equity premium. The higher covariance between returns and signals arises because a less precise dividend drift prediction $m_{\theta,h^*}$ implies a lower sensitivity of the investor’s hedging demand to signals. In that case, positive (negative) signals are more directly linked to a positive (negative) excess demand for equity and imply a higher covariance between equilibrium equity returns and signals.

4.2. Equity Volatility and Risk/Return Relations

From Proposition 2, the variance of stock returns is

$$\sigma^2_R = \sigma^2_D + V_{\theta,h^*} (2 + kV_{\theta,h^*}) \cdot$$

(46)

The variance is increasing in the absolute value of $V_{\theta,h^*}$ and for $V_{\theta,h^*} = 0$ attains the minimum $\sigma^2_D$. As in Veronesi (2000), to obtain empirically relevant volatilities it is crucial to introduce learning into the model. However, we know from Section 4.1 that the only parameter choice consistent with sizable equity premia is $\gamma < 1$.

Column Vol of Table 3 illustrates quantitatively the contribution of learning and ambiguity to equity return volatilities, for both the setting of no learning (column NL) and of learning (column L). In the absence of learning, we obtain small equity volatilities of 3.25%. In the presence of learning and for $\gamma = 0.85$ (Panel B), volatility ranges from 18.49%, without ambiguity ($\eta = 0$), to 18.47%, for an ambiguity parameter $\eta = 0.01$. These volatility levels correspond to the 18.5% unconditional volatility of returns estimated in Table 4 (see column Vol).
From Proposition 2 and equation (46), we see that the relations between risk or equity premia and the conditional variance of returns are given by

\[ \mu^\text{rc}_R = \gamma \sigma^2_R - \gamma V_{\theta, h^*} (1 + k V_{\theta, h^*}) \]  

and

\[ \mu_R = \left( \gamma - \frac{m h^*(\theta)}{V_{\theta, h^*}} \right) \sigma^2_R - \gamma V_{\theta, h^*} (1 + k V_{\theta, h^*}) + m h^*(\theta) \left( 1 + \frac{\sigma_D}{V_{\theta, h^*}} \right) . \]  

Equation (48) implies a truly positive, but time-varying, theoretical relation between the equity premium \( \mu_R \) and the conditional variance \( \sigma^2_R \). This time-varying relation is due to the ambiguity premium and results from the interaction of learning and ambiguity aversion. The true relation between the risk premium \( \mu^\text{rc}_R \) and the conditional variance \( \sigma^2_R \) is linear and constant. Both relations (47) and (48) are biased by a heteroscedastic error term that has a nonzero conditional mean.

Figure 2 illustrates the theoretical relation between risk premia, equity premia, and conditional variances of equity returns. Using the same initial prior and the model parameters of Figure 1, we simulate a time series of posterior probabilities \( \Pi(t) \) and compute the resulting theoretical (time-varying) coefficient

\[ b_t = \gamma - m h^*(\theta)/V_{\theta, h^*}, \]  

from the expected excess return and variance relation in equation (48).

Insert Figure 2 about here

The theoretical (time-varying) equity premium “sensitivity” \( b_t \) to changes in \( \sigma^2_R \) is large compared to the sensitivity of the risk premium, which is just the risk aversion coefficient \( \gamma \). By definition, ambiguity premia derive from model misspecification, not from covariances between asset returns and aggregate consumption growth. Therefore, we can expect them to be difficult to identify by, for instance, regression methods. Figure 3 highlights this point by plotting the time series of estimated parameters for a sequence of rolling regressions of \( R \) on the theoretical value of \( \sigma^2_R \) defined by equation (46). To compute the time series of \( R \) and \( \sigma^2_R \) and the corresponding linear regression estimates, we use the same simulated trajectories of \( \Pi(t) \) as in Figure 2. To estimate the rolling regression coefficients, we use samples of 180 observations in each regression.
In Figure 3, highly time-varying regression estimates arise. These estimates indicate that the estimated relation between \( \mu_R \) and \( \sigma^2_R \) over different time periods can even switch signs. The estimated coefficients clearly fail to identify the theoretical coefficient \( b_t \) in equation (49). For instance, the estimated parameters for \( \gamma = 0.9 \) are never above 0.1, but the theoretical coefficient \( b_t \) in Figure 2 stays above two for all ambiguity aversion parameters.

5. Conclusion

We study asset prices in a continuous-time partial-information Lucas economy with ambiguity aversion. We find that for reasonable parameter values, the joint presence of learning and ambiguity generates a high equity premium, a low interest rate, and excess volatility. Model settings that do not take learning or ambiguity aversion into account need an unrealistically large amount of pessimism to generate sizable equity premia. At the same time, model settings without learning imply small equity volatilities. A further implication of our model is a highly time-varying relation between excess returns and their conditional variances. This feature leads to estimated relations between excess returns and conditional variances with an undetermined sign and implies large time-varying biases in the naively estimated risk-return tradeoff.

By using a time-additive power utility function in our model, we can obtain easily interpretable closed-form solutions for the equilibrium at the cost of constraining the relation between risk aversion and elasticity of intertemporal substitution (EIS). Disentangling risk aversion and EIS could be useful to introduce an additional degree of freedom in the choice of the risk aversion parameter, which would make the model more flexible in the fit of, e.g., higher worst-case equity premia. Such extensions are interesting directions for future research.
Appendix: Proofs

Proof of Proposition 1. We have for any likelihood \( h(\theta) \in \Xi(\theta) \),

\[
V^{h(\theta)}(\Pi, D) = E^{h(\theta)} \left[ \int_{s}^{\infty} e^{-\delta(t-s)} \frac{D^{1-\gamma}}{1-\gamma} dt \bigg| F(s) \right] = E^{h(\theta)} \left[ \int_{s}^{\infty} e^{-\delta(t-s)} \frac{D(t)}{1-\gamma} dt \bigg| \pi_{1}(s) = \pi_{1}, \ldots, \pi_{n}(s) = \pi_{n}, D(s) = D \right] = \frac{D^{1-\gamma}}{1-\gamma} \sum_{i=1}^{n} \pi_{i} E^{h(\theta_{i})} \left[ \int_{s}^{\infty} e^{-\delta(t-s)} \left( \frac{D(t)}{D(s)} \right)^{1-\gamma} dt \bigg| \tilde{\theta} = \tilde{\theta}_{i} \right], \quad (A1)
\]

where \( \tilde{\theta} = \theta + h(\theta) \sigma_{D}, \tilde{\theta}_{i} = \theta_{i} + h(\theta_{i}) \sigma_{D} \). Therefore, for any vector \( \Pi \):

\[
J(\Pi, D) = \inf_{h(\theta)} \left\{ V^{h(\theta)}(\Pi, D) \right\} \geq \frac{D^{1-\gamma}}{1-\gamma} \sum_{i=1}^{n} \pi_{i} \inf_{h(\theta_{i})} E^{h(\theta_{i})} \left[ \int_{s}^{\infty} e^{-\delta(t-s)} \left( \frac{D(t)}{D(s)} \right)^{1-\gamma} dt \bigg| \tilde{\theta} = \tilde{\theta}_{i} \right] . \quad (A2)
\]

Conditional on \( \tilde{\theta}_{i} \), the \( h(\theta) \)–drift misspecified dynamics are

\[
dD = (\theta_{i} + h(\theta_{i}) \sigma_{D}) Ddt + \sigma_{D} DdB \quad . \quad (A3)
\]

Therefore, Assumption 2 implies that we can focus on solving the problem

\[
V^{i}(D) = \inf_{h(\theta_{i})} E \left( \int_{s}^{\infty} e^{-\delta(t-s)} \frac{D(t)^{1-\gamma}}{1-\gamma} dt \bigg| D(s) = D \right) ,
\]

subject to

\[
\frac{1}{2} h(\theta_{i})^{2} \leq \eta(\theta_{i})
\]

and the dividend dynamics

\[
dD = (\theta_{i} + h(\theta_{i}) \sigma_{D}) Ddt + \sigma_{D} DdB \quad . \quad (A4)
\]

The Hamilton Jacobi Bellman equation for this problem reads

\[
0 = \inf_{h(\theta_{i})} \left\{ -\delta V^{i} + \frac{D^{1-\gamma}}{1-\gamma} + (\theta_{i} + h(\theta_{i}) \sigma_{D}) D V^{i}_{D} + \frac{1}{2} \sigma_{D}^{2} D^{2} V^{i}_{DD} + \lambda \left( \frac{1}{2} h(\theta_{i})^{2} - \eta(\theta_{i}) \right) \right\} . \quad (A5)
\]

28
where $\lambda \geq 0$ is a Lagrange multiplier for the constraint $\frac{1}{2}h(\theta_i)^2 \leq \eta(\theta_i)$. Equation (A5) implies the optimality condition
\[ h(\theta_i) = -\frac{\sigma_D D}{\lambda} V_i^j . \] (A6)

Slackness then gives
\[ \frac{\sigma_D^2 D^2}{\lambda^2} (V_i^j)^2 = 2\eta(\theta_i) , \] (A7)

implying
\[ h(\theta_i) = -\sqrt{2\eta(\theta_i)} \text{sign}[\sigma_D D V_i^j] = -\sqrt{2\eta(\theta_i)} . \] (A8)

This result proves the first statement. To prove the second statement, we note that
\[ V_i^j(D) = D_{1-\gamma} \mathbb{E} \left[ \int_s^\infty e^{-\delta(t-s)} \left( \frac{D(t)}{D(s)} \right)^{1-\gamma} dt \bigg| \tilde{\theta} = \theta_i - \sqrt{2\eta(\theta_i)} \sigma_D \right] . \] (A9)

Conditional on $\tilde{\theta} = \theta_i - \sqrt{2\eta(\theta_i)} \sigma_D$, the solution of the dividend dynamics gives
\[ \left( \frac{D(t)}{D(s)} \right)^{1-\gamma} = \exp \left\{ (1-\gamma) \left( \theta_i - \sqrt{2\eta(\theta_i)} \sigma_D - \frac{\sigma_D^2}{2} \right) (t-s) + (1-\gamma) \sigma_D \left( B_D(t) - B_D(s) \right) \right\} , \]

implying, under the given assumptions,
\[ \mathbb{E} \left[ \int_s^\infty e^{-\delta(t-s)} \left( \frac{D(t)}{D(s)} \right)^{1-\gamma} dt \bigg| \tilde{\theta} = \theta_i - \sqrt{2\eta(\theta_i)} \sigma_D \right] = \frac{1}{\tilde{\theta}_i + (1-\gamma) \sqrt{2\eta(\theta_i)} \sigma_D} , \]

where
\[ \tilde{\theta}_i = \delta - (1-\gamma) \theta_i + \gamma (1-\gamma) \frac{\sigma_D^2}{2} > 0 . \] (A10)

We thus obtain for the equilibrium price of the risky asset with dividend process $D$
\[ \frac{P(t)}{D(t)} = \sum_{i=1}^n \pi_i \mathbb{E} \left[ \int_t^\infty e^{-\delta(s-t)} \left( \frac{D(s)}{D(t)} \right)^{1-\gamma} ds \bigg| \tilde{\theta} = \theta_i - \sqrt{2\eta(\theta_i)} \sigma_D \right] , \] (A11)

or equivalently
\[ P(t) \rho(t) = \sum_{i=1}^n \pi_i \mathbb{E} \left[ \int_t^\infty \rho(s) D(s) \, ds \bigg| \tilde{\theta} = \theta_i - \sqrt{2\eta(\theta_i)} \sigma_D \right] \] (A12)
where \( \rho(t) = u_c(D(t), t) = e^{-\delta t} D(t)^{-\gamma} \). This result proves the second statement of the proposition. Writing equation (A12) in differential form and applying it to the risky asset paying a “dividend” \( D = r \) we obtain

\[
rdt = -\sum_{i=1}^{n} \pi_i E_t \left[ \frac{d\rho}{\rho} \right] \bar{\theta} = \theta_i - \sqrt{2 \eta(\theta_i) \sigma_D} = \left( \delta + \gamma (m_\theta + m_{h(\theta)} \sigma_D) - \frac{1}{2} \gamma (\gamma + 1) \sigma_D^2 \right) dt,
\]

i.e., the third statement of the proposition, concluding the proof. ■

Proof of Proposition 2. Statement (i) follows by applying the proof of Proposition 2 in Veronesi (2000) to the \( D \)-dynamics in (18) under the worst-case likelihood \( h^*(\theta) = -\sqrt{2 \eta(\theta)} \). Statement (ii) follows by expressing the excess return dynamics obtained in (i) with respect to the filtered Brownian motions \( \tilde{B}_D, \tilde{B}_e \) under the reference model. ■

Proof of Corollary 3. From Lemma 2, \( V_{\theta, h^*} \) is increasing in mean-preserving spreads \( \Pi \), under the given assumptions. Moreover,

\[
-m_{h^*} = \sum_{i=1}^{n} \pi_i \sqrt{\eta(\theta_i)} \sigma_D \quad (A13)
\]

is also increasing in mean-preserving spreads, because of the assumed convexity of \( \sqrt{2 \eta(\theta)} \) as a function of \( \theta \). ■
References


Figures and Tables

**Fig. 1.** Risk premium and ambiguity premium. Panel A plots the set of probabilities $\Pi$ relevant for the figure. Panel B plots the different relevant reference model states $\theta_1, \ldots, \theta_n$. The true dividend drift state of the reference model is marked with a square and is equal to the posterior expected value $\sum \pi_i \theta_i$. Calculations are based on a model with three reference model drift states $\Theta = \{0.005, 0.0175, 0.03\}$ with prior probabilities $\Pi = \{0.266, 0.468, 0.266\}$. We use a small amount of homogeneous ambiguity $\eta = 0.001$. In Panel B, the size of the ambiguous neighborhoods $\Xi(\theta_1), \ldots, \Xi(\theta_n)$ is highlighted by the ellipses centered at $\theta_1, \ldots, \theta_n$. Further, we set $\delta = 0.01$, $\sigma_D = 0.0325$ and $\sigma_e = 0.0125$. With this prior structure and these parameters, Panels C and D plot the resulting risk premium $\mu_{wc}^R$ and the equity premium $\mu_R$ as functions of $\gamma$ for ambiguity aversion (solid line) and no ambiguity aversion (dashed line).
Fig. 2. Theoretical time-varying return/variance tradeoff. For different parameters $\gamma = 0.8, 0.9$, we simulate the theoretical (time-varying) coefficient $b_t$ in equation (49). In all panels, we plot $b_t$ as a function of time for homogeneous ambiguity structures $\eta = 0.001$ (solid line) and 0.005 (dash-dotted line). The dashed line represents $b_t = \gamma$. The three reference model drift states are $\Theta = \{0.005, 0.0175, 0.03\}$ with initial prior probabilities $\Pi = \{0.266, 0.468, 0.266\}$. The true dividend drift is $\theta = 0.0175$. The remaining parameters are $\delta = 0.01, \sigma_D = 0.0325, \sigma_e = 0.0125$. 

36
Fig. 3. Rolling regression analysis of equity returns on equilibrium variances. For different parameters $\gamma = 0.8, 0.9$, we plot the estimated parameter $b$ in a series of rolling regressions of $R$ on $\sigma_R^2$. All panels present the results of one single simulation run. The rolling regressions are based on sample sizes of 180 observations, simulated from a model with three reference model drift states $\Theta = \{0.005, 0.0175, 0.03\}$ with initial probabilities $\Pi = \{0.266, 0.468, 0.266\}$ and under a homogeneous degree of ambiguity $\eta = 0.001$. The true dividend is $\theta = 0.0175$. The remaining parameters are $\delta = 0.01$, $\sigma_D = 0.0325$, $\sigma_e = 0.0125$. 

The entries report the degree of pessimism and the average model detection error probability $\epsilon_N(\theta)$ implied by different sizes for the ambiguity parameter $\eta(\theta)$ and for sample sizes $N = 100, 200, 300, 400$. We calculate the degree of pessimism as the difference between the expected consumption growth rate under the reference belief and under the worst-case belief. For the calculations, we assume $\sigma_D = 0.0325$. 

<table>
<thead>
<tr>
<th>$\eta(\theta)$</th>
<th>Degree of Pessimism</th>
<th>$\epsilon_N(\theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>N=100</td>
</tr>
<tr>
<td>0.0001</td>
<td>0.05 %</td>
<td>0.47</td>
</tr>
<tr>
<td>0.0010</td>
<td>0.15 %</td>
<td>0.41</td>
</tr>
<tr>
<td>0.0025</td>
<td>0.23 %</td>
<td>0.36</td>
</tr>
<tr>
<td>0.0050</td>
<td>0.33 %</td>
<td>0.31</td>
</tr>
<tr>
<td>0.0075</td>
<td>0.40 %</td>
<td>0.27</td>
</tr>
<tr>
<td>0.0100</td>
<td>0.46 %</td>
<td>0.24</td>
</tr>
</tbody>
</table>
The entries report \( P/D \) ratios for different risk aversion coefficients \( \gamma \) and different homogeneous levels of ambiguity \( \eta \), in model settings without learning (NL) and with learning (L). Calculations are based on a model with three reference model drift states \( \Theta = \{0.005, 0.0175, 0.03\} \) with prior probabilities \( \Pi = \{0.266, 0.468, 0.266\} \). The true dividend drift is \( \theta = 0.0175 \). Further parameters are \( \delta = 0.035, \sigma_D = 0.0325, \sigma_e = 0.0125 \).

<table>
<thead>
<tr>
<th>( \eta )</th>
<th>( \gamma = 0.6 )</th>
<th>( \gamma = 0.7 )</th>
<th>( \gamma = 0.8 )</th>
<th>( \gamma = 0.9 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NL</td>
<td>L</td>
<td>NL</td>
<td>L</td>
</tr>
<tr>
<td>0.0000</td>
<td>35.55</td>
<td>36.27</td>
<td>33.49</td>
<td>33.82</td>
</tr>
<tr>
<td>0.0001</td>
<td>35.32</td>
<td>36.02</td>
<td>33.33</td>
<td>33.66</td>
</tr>
<tr>
<td>0.0010</td>
<td>34.83</td>
<td>35.51</td>
<td>33.01</td>
<td>33.32</td>
</tr>
<tr>
<td>0.0025</td>
<td>34.43</td>
<td>35.08</td>
<td>32.73</td>
<td>33.04</td>
</tr>
<tr>
<td>0.0050</td>
<td>33.98</td>
<td>34.61</td>
<td>32.43</td>
<td>32.73</td>
</tr>
<tr>
<td>0.0075</td>
<td>33.65</td>
<td>34.25</td>
<td>32.20</td>
<td>32.49</td>
</tr>
<tr>
<td>0.0100</td>
<td>33.37</td>
<td>33.96</td>
<td>32.01</td>
<td>32.30</td>
</tr>
</tbody>
</table>

**Table 2**

Price-Dividend Ratios.
Table 3

Equilibrium Interest Rates, Risk Premia, Equity Premia, and Volatilities.

The entries report the equilibrium interest rates (IR), risk premia (RP), equity premia (EP), and volatilities (Vol) for risk aversions $\gamma = 0.8$ (Panel A), $\gamma = 0.85$ (Panel B) and $\gamma = 0.9$ (Panel C) and for different homogeneous levels of ambiguity $\eta$, under models without learning (NL) and with learning (L). Calculations are based on a model with three reference model drift states $\Theta = \{0.005, 0.0175, 0.03\}$ with prior probabilities $\Pi = \{0.266, 0.468, 0.266\}$. The true dividend is $\theta = 0.0175$. Further parameters are $\delta = 0.01$, $\sigma_D = 0.0325$, $\sigma_e = 0.0125$. 

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>IR (%)</th>
<th>RP (%)</th>
<th>EP (%)</th>
<th>Vol (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NL L</td>
<td>NL L</td>
<td>NL L</td>
</tr>
<tr>
<td>0.0000</td>
<td>2.32</td>
<td>0.08</td>
<td>0.33</td>
<td>0.08</td>
</tr>
<tr>
<td>0.0001</td>
<td>2.29</td>
<td>0.08</td>
<td>0.33</td>
<td>0.13</td>
</tr>
<tr>
<td>0.0010</td>
<td>2.21</td>
<td>0.08</td>
<td>0.33</td>
<td>0.13</td>
</tr>
<tr>
<td>0.0025</td>
<td>2.14</td>
<td>0.08</td>
<td>0.33</td>
<td>0.31</td>
</tr>
<tr>
<td>0.0050</td>
<td>2.06</td>
<td>0.08</td>
<td>0.33</td>
<td>0.41</td>
</tr>
<tr>
<td>0.0075</td>
<td>2.01</td>
<td>0.08</td>
<td>0.33</td>
<td>0.32</td>
</tr>
<tr>
<td>0.0100</td>
<td>1.96</td>
<td>0.08</td>
<td>0.33</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Panel A: $\gamma = 0.8$

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>IR (%)</th>
<th>RP (%)</th>
<th>EP (%)</th>
<th>Vol (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NL L</td>
<td>NL L</td>
<td>NL L</td>
</tr>
<tr>
<td>0.0000</td>
<td>2.40</td>
<td>0.09</td>
<td>0.26</td>
<td>0.09</td>
</tr>
<tr>
<td>0.0001</td>
<td>2.37</td>
<td>0.09</td>
<td>0.26</td>
<td>0.14</td>
</tr>
<tr>
<td>0.0010</td>
<td>2.28</td>
<td>0.09</td>
<td>0.26</td>
<td>0.24</td>
</tr>
<tr>
<td>0.0025</td>
<td>2.21</td>
<td>0.09</td>
<td>0.26</td>
<td>0.32</td>
</tr>
<tr>
<td>0.0050</td>
<td>2.13</td>
<td>0.09</td>
<td>0.26</td>
<td>0.41</td>
</tr>
<tr>
<td>0.0075</td>
<td>2.07</td>
<td>0.09</td>
<td>0.26</td>
<td>0.49</td>
</tr>
<tr>
<td>0.0100</td>
<td>2.01</td>
<td>0.09</td>
<td>0.26</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Panel B: $\gamma = 0.85$

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>IR (%)</th>
<th>RP (%)</th>
<th>EP (%)</th>
<th>Vol (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NL L</td>
<td>NL L</td>
<td>NL L</td>
</tr>
<tr>
<td>0.0000</td>
<td>2.48</td>
<td>0.10</td>
<td>0.20</td>
<td>0.10</td>
</tr>
<tr>
<td>0.0001</td>
<td>2.44</td>
<td>0.10</td>
<td>0.20</td>
<td>0.14</td>
</tr>
<tr>
<td>0.0010</td>
<td>2.35</td>
<td>0.10</td>
<td>0.20</td>
<td>0.24</td>
</tr>
<tr>
<td>0.0025</td>
<td>2.28</td>
<td>0.10</td>
<td>0.20</td>
<td>0.32</td>
</tr>
<tr>
<td>0.0050</td>
<td>2.19</td>
<td>0.10</td>
<td>0.20</td>
<td>0.42</td>
</tr>
<tr>
<td>0.0075</td>
<td>2.13</td>
<td>0.10</td>
<td>0.20</td>
<td>0.49</td>
</tr>
<tr>
<td>0.0100</td>
<td>2.07</td>
<td>0.10</td>
<td>0.20</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Panel C: $\gamma = 0.9$
The table reports the interest rate, equity premium and volatility parameters, estimated in Campbell (1999) for an annual sample spanning the period from 1891 to 1994.

Table 4
Estimated Interest Rate and Returns Parameters.

<table>
<thead>
<tr>
<th>IR (%)</th>
<th>EP (%)</th>
<th>Vol (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.95%</td>
<td>6.26%</td>
<td>18.5%</td>
</tr>
</tbody>
</table>

The table reports the interest rate, equity premium and volatility parameters, estimated in Campbell (1999) for an annual sample spanning the period from 1891 to 1994.