

Ambiguity, Learning, and Asset Returns*

Nengjiu Ju[†] and Jianjun Miao[‡]

September 2007

Abstract

We develop a consumption-based asset-pricing model in which the representative agent is ambiguous about the hidden state in consumption growth. He learns about the hidden state under ambiguity by observing past consumption data. His preferences are represented by the smooth ambiguity model axiomatized by Klbanoff et al. (2005, 2006). Unlike the standard Bayesian theory, this utility model implies that the posterior of the hidden state and the conditional distribution of the consumption process given a state cannot be reduced to a predictive distribution. By calibrating the ambiguity aversion parameter, the subjective discount factor, and the risk aversion parameter (with the latter two values between zero and one), our model can match the first moments of the equity premium and riskfree rate found in the data. In addition, our model can generate a variety of dynamic asset pricing phenomena, including the procyclical variation of price-dividend ratios, the countercyclical variation of equity premia and equity volatility, and the mean reversion and long horizon predictability of excess returns.

Keywords: Ambiguity aversion, learning, asset pricing puzzles, model uncertainty, robustness, pessimism

JEL Classification: D81, E44, G12

*We thank Massimo Marinacci and Tom Sargent for encouragement, Lars Hansen for helpful comments, and Costis Skiadas for helpful conversations. We are particularly grateful to Massimo Marinacci for his patient explanations of our questions.

[†]Department of Finance, the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: nengjiu@ust.hk. Tel: (+852) 2358 8318.

[‡]Department of Economics, Boston University, 270 Bay State Road, Boston MA 02215, USA, and Department of Finance, the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: miaoj@bu.edu. Tel: (+852) 2358 8298.

1. Introduction

Under the rational expectations hypothesis, there exists an objective probability law governing the state process, and economic agents know this law which coincides with their subjective beliefs. This rational expectations hypothesis has become the workhorse in macroeconomics and finance. However, it faces serious difficulties when confronting with asset markets data. Most prominently, Mehra and Prescott (1985) show that for a standard rational, representative-agent model to explain the high equity premium observed in the data, an implausibly high degree of risk aversion is needed, resulting in the equity premium puzzle. Weil (1989) shows that this high degree of risk aversion generates an implausibly high riskfree rate, resulting in the riskfree rate puzzle. In addition, a number of empirical studies document puzzling links between aggregate asset markets and macroeconomics: Price-dividend ratios move procyclically (Campbell and Shiller (1988a)) and conditional expected equity premia move countercyclically (Campbell and Shiller (1988a) and Fama and French (1989)). Excess returns are serially correlated and mean reverting (Fama and French (1988b) and Poterba and Summers (1988)). Excess returns are forecastable; in particular, the log dividend yield predicts long-horizon realized excess returns (Campbell and Shiller (1988b), Fama and French (1988a)). Conditional volatility of stock returns is persistent and moves countercyclically (Bollerslev et al. (1992)).

In this paper, we develop a representative-agent consumption-based asset-pricing model that helps explain the preceding puzzles simultaneously by departing from the rational expectations hypothesis. Our model has two main ingredients. First, we assume that aggregate consumption follows a hidden Markov regime-switching process. The agent learns about the hidden state based on past consumption data. The posterior state beliefs capture fluctuating economic uncertainty and drive asset return dynamics. Second, we assume that the agent is ambiguous about the hidden state and his preferences are represented by the smooth ambiguity model of Klibanoff et al. (2005, 2006). In order to derive quantitative implications, we study two tractable utility specifications. The log-exponential specification features a unit coefficient of relative risk aversion and a constant coefficient of absolute ambiguity aversion. This specification is equivalent to the multiplier preferences (Hansen and Sargent (2001)) and the risk-sensitive preferences (Tallarini (2000)), as pointed out by Hansen (2007). Ambiguity aversion is manifested through a pessimistic distortion of state beliefs. Under the distorted state beliefs, smaller values of continuation utilities receive relatively higher weight. We also consider the power-power specification in which the agent exhibits constant relative risk aversion and constant relative ambiguity aversion. In this case, ambiguity aversion is manifested

through a distortion of the standard pricing kernel. This distortion also features pessimism, but does not admit an interpretation based on the multiplier or risk-sensitive preferences. For both specifications, we can find reasonable parameter values (both the subjective discount factor and the risk aversion coefficient are between zero and one) to match the mean riskfree rate and the mean equity premium in the historical data. However, the log-exponential specification cannot deliver interesting aggregate stock return dynamics because the consumption-wealth ratio is constant. In this case, the price-dividend ratio is also constant when equilibrium aggregate consumption is equal to aggregate dividends. By contrast, the power-power specification can generate the dynamic asset pricing phenomena mentioned earlier.

We motivate our adoption of the smooth ambiguity model in two ways. First, the Ellsberg Paradox (Ellsberg (1961)) and related experimental evidence demonstrate that the distinction between risk and ambiguity is behaviorally meaningful. Roughly speaking, risk refers to the situation where there is a probability measure to guide choice, while ambiguity refers to the situation where the decision maker is uncertain about this probability measure due to cognitive or informational constraints. Knight (1921) and Keynes (1936) emphasize that ambiguity may be important for economic decision-making. We assume that the agent in our model is ambiguous about the hidden state in consumption growth. Our adopted smooth ambiguity model captures this ambiguity and attitude towards ambiguity. Our second motivation is related to the robustness theory developed by Hansen and Sargent (2001) and Hansen (2007). Specifically, the agent in our model may fear model misspecification in the consumption process. He is concerned about this model uncertainty, and thus, seeks robust decision-making. We may interpret the smooth ambiguity model as a model of robustness in the presence of model uncertainty.

Our modeling of learning echoes with Hansen's (2007) suggestion that one should put econometricians and economic agents on comparable footings in terms of statistical knowledge. When estimating the regime-switching consumption process, econometricians typically apply Hamilton's (1989) maximum likelihood method and assume that they do not observe the hidden state. However, the rational expectations hypothesis often requires economic agents to be endowed with more precise information than econometricians. A typical assumption is that agents know all parameter values underlying the consumption process (e.g., Cecchetti et al. (1990, 2000)). In this paper, we show that there are important quantitative implications when agents are concerned about statistical ambiguity by removing the information gap between them and econometricians, while the standard Bayesian learning has small quantitative effects.¹

¹There is a large literature on learning in asset pricing using the standard Bayesian framework. Notable works include Brandt et al. (2004), Brennan and Xia (2001), David (1997), Detemple (1986), Dothan and Feldman (1986), Timmermann (1993), Veronesi (1999, 2000), Wang (1993), and Weitzman (2006).

Learning is naturally embedded in the recursive smooth ambiguity model. In this model, the posterior of the hidden state and the conditional distribution of the consumption process given a state cannot be reduced to a compound predictive distribution, unlike in the standard Bayesian analysis. It is this irreducibility that captures ambiguity or model uncertainty. An important advantage of the smooth ambiguity model over other models of ambiguity such as the maxmin expected utility (or multiple-priors) model of Gilboa and Schmeidler (1989) is that it achieves a separation between ambiguity (beliefs) and ambiguity attitude (tastes). This feature allows us to do comparative statics with respect to the ambiguity aversion parameter holding ambiguity fixed, and to calibrate it for quantitative analysis. Another advantage is that we can apply the usual differential analysis for the smooth ambiguity model under standard regularity conditions. We can then derive the pricing kernel quite tractably. By contrast, the widely applied maxmin expected utility model lacks this smoothness property.

Our paper contributes to a growing body of literature that studies the implications of ambiguity and robustness for finance and macroeconomics.² Here we discuss closely related papers only. Epstein and Schneider (2007a) model learning under ambiguity using a set of priors and a set of likelihoods. Both sets are updated by Bayes' Rule in a suitable way. Applying this learning model, Epstein and Schneider (2007b) analyze asset pricing implications. Leippold et al. (2007) extend this model to a continuous-time environment. Hansen and Sargent (2006) formulate a learning model that allows for two forms of model misspecification: (i) misspecification in the underlying Markov law for the hidden states, and (ii) misspecification of the probabilities assigned to the hidden Markov states. Hansen and Sargent (2007) apply this learning model to study time-varying model uncertainty premia. Hansen (2007) surveys models of learning and robustness. He analyzes a continuous-time model similar to our log-exponential case. But he does not consider the power-power case and does not conduct a thorough quantitative analysis as in our paper. Our paper is also related to Abel (2002), Brandt et al. (2004), and Cecchetti et al. (2000) who model the agent's pessimism and doubt in specific ways and show that their modeling helps explain many asset pricing puzzles. Our adopted smooth ambiguity model captures pessimism and doubt with a decision theoretic foundation.

The remainder of the paper proceeds as follows. Section 2 presents the smooth ambiguity

²See Cao et al. (2005), Chen and Epstein (2002), Epstein and Miao (2003), Epstein and Wang (1994, 1995), Garlappi et al. (2007), Miao and Wang (2007), and Routledge and Zin (2001) for asset pricing applications of the multiple-prior utility model. See Anderson et al. (2002), Cagetti, et al. (2002), Hansen and Sargent (2001), Hansen et al. (1999), Hansen et al. (2006), Liu et al. (2005), Maenhout (2004), and Uppal and Wang (2003) for models of robustness and applications. Maccheroni et al. (2006) provide an axiomatic foundation for one of Hansen and Sargent's robustness formulations – the multiplier preferences. See Backus et al. (2005) and Hansen and Sargent (2008) for a survey.

model. Section 3 analyzes its asset pricing implications in a Lucas-style model. Section 4 calibrates the model and studies its quantitative implications. Section 5 concludes. Appendices contain proofs and an outline of the numerical method.

2. Smooth Ambiguity Preferences

In order to study its asset pricing implications, we first review the smooth ambiguity model developed by Klibanoff et al. (2005, 2006), and then derive its utility gradient. We also discuss related alternative approaches. We refer the reader to the preceding two papers for further details and for axiomatic foundations.

2.1. Static Smooth Ambiguity Model

We start with the static model of Klibanoff et al. (2005). Suppose uncertainty is represented by a measurable space (S, \mathcal{S}) . An agent ranks uncertain prospects or acts, maps from S into some outcome set. An example of acts is consumption. The agent's smooth ambiguity preferences over consumption are represented by the following utility function:³

$$\phi^{-1} \left(\int_{\Pi} \phi(\mathbb{E}_{\pi} u(C)) d\mu \right), \quad \forall C : S \rightarrow \mathbb{R}_+, \quad (1)$$

where \mathbb{E}_{π} is the expectation operator with respect to the probability distribution π on (S, \mathcal{S}) , u is a vN-M utility function, ϕ is an increasing function, and μ is a subjective prior over the set Π of probability measures π that the agent thinks possible.

A key feature of this model is that it achieves a separation between ambiguity, identified as a characteristic of the agent's subjective beliefs, and ambiguity attitude, identified as a characteristic of the agent's tastes.⁴ Specifically, ambiguity is characterized by properties of the subjective set of measures Π . Attitudes towards ambiguity are characterized by the shape of ϕ , while attitudes towards pure risk are characterized by the shape of u , as usual. In particular, the agent displays ambiguity aversion if and only if ϕ is concave. Intuitively, an ambiguity averse agent prefers consumption that is more robust to the possible variation in probabilities. That is, he is averse to mean-preserving spreads in the distribution μ_C induced by the prior μ and the consumption act C . This distribution represents the uncertainty about *ex ante* evaluation of C

³The utility function in (1) is ordinally equivalent to the utility function $\mathbb{E}_{\mu} \phi(\mathbb{E}_{\pi} u(c))$ given in Klibanoff et al. (2005).

⁴The behavioral foundation of ambiguity and ambiguity attitude is based on the theory developed by Ghirardato and Marinacci (2002). Epstein (1999) provides a different foundation. The main difference is that the benchmark ambiguity neutral preference is the expected utility preference according to Ghirardato and Marinacci (2002), while Epstein's (1999) benchmark is the probabilistic sophisticated preferences.

given π , $\mathbb{E}_\pi u(C)$. Note that there is no reduction between μ and π in general. It is possible when ϕ is linear. In this case, the agent is ambiguity neutral and the smooth ambiguity model in (1) reduces to the standard expected utility model. Segal (1987) first points out the idea of modeling ambiguity attitude by relaxing the axiom of reduction of compound lotteries. An alternative interpretation of the irreducibility of compound lotteries is related to the timing of resolution of information studied by Epstein and Zin (1989) and Kreps and Porteus (1978).

An important advantage of the smooth ambiguity model over other models of ambiguity, such as the widely adopted maxmin expected utility model of Gilboa and Schmeidler (1989), is that it is tractable and admits a clear-cut comparative statics analysis. Tractability is revealed by the fact that the well-developed machinery for dealing with risk attitudes can be applied to ambiguity attitudes. In addition, the indifference curve implied by (1) is smooth under regularity conditions, rather than kinked as in the case of the maxmin expected utility model. More importantly, comparative statics of ambiguity attitudes can be easily analyzed using the function ϕ only, holding ambiguity fixed. Such a comparative static analysis is not evident for the maxmin expected utility model since the set of priors in that model may characterize ambiguity as well as ambiguity attitudes.

Analogous to the standard risk theory, Klibanoff et al. (2005) define the coefficients of absolute and relative ambiguity aversion at x as $-\phi''(x)/\phi'(x)$ and $-x\phi''(x)/\phi'(x)$, respectively. We are particularly interested in the following two cases:

- constant absolute ambiguity aversion (CAAA) utility:

$$\phi(x) = -e^{-\frac{x}{\theta}}, \theta > 0, \quad (2)$$

where $1/\theta$ is the parameter of CAAA.

- constant relative ambiguity aversion (CRAA) utility:

$$\phi(x) = \frac{x^{1-\alpha}}{1-\alpha}, \alpha > 0, \neq 1 \quad (3)$$

where α is the parameter of CRAA. We identify the case $\alpha = 1$ as $\phi(x) = \log x$.

Klibanoff et al. (2005) show that when the coefficient of absolute ambiguity aversion goes to infinity (e.g., $\theta \rightarrow 0$),⁵ the smooth ambiguity model converges to the maxmin expected utility model:

$$\inf_{\pi \in \Pi} \mathbb{E}_\pi u(C).$$

⁵The coefficient of absolute ambiguity aversion need not be constant for this result to hold as long as a regularity condition in Klibanoff et al. (2006) is satisfied.

Thus, the maxmin expected utility model is the limiting case where the agent displays extreme ambiguity aversion. In the other extreme case where $\theta \rightarrow \infty$, the agent is ambiguity neutral and we obtain the standard expected utility model, $\mathbb{E}_\pi u(C)$.

When ϕ is given by (2), the smooth ambiguity model has an interesting connection to the robust control theory developed by Hansen and Sargent (2008). One can show that

$$\begin{aligned} \phi^{-1}(\mathbb{E}_\mu \phi(\mathbb{E}_\pi u(C))) &= \min_{m \geq 0, \mathbb{E}_\mu[m]=1} \mathbb{E}_\mu[m \mathbb{E}_\pi u(C)] + \theta \mathbb{E}_\mu[m \log m] \\ &= -\theta \log \mathbb{E}_\mu \exp\left(-\frac{\mathbb{E}_\pi u(C)}{\theta}\right). \end{aligned} \quad (4)$$

This equation shows that one can give two different alternative interpretations for the smooth ambiguity preferences. The expression in the second line of (4) gives the risk-sensitive formulation used in the control theory, while the expression in the first line of (4) gives the robust control formulation. Here m represents a Radon-Nikodym derivative used to distort the prior μ . The set of possible distortions is defined by a relative entropy criterion. The parameter θ can be interpreted as the Lagrange multiplier associated with the set of densities. Anderson et al. (2003) advocate to use model detection error probabilities to calibrate θ .

More generally, we may interpret the utility model defined in (1) as a model of robustness in which the agent is concerned about model misspecification, and thus, seeks robust decision making. Specifically, each distribution π in Π describes an economic model. The agent is ambiguous about which is the right model specification. He has a subjective prior μ over alternative models. He is averse to model uncertainty, and thus, evaluates different models using a concave expected utility function ϕ .

2.2. Learning and Recursive Smooth Ambiguity Model

We now turn to the dynamic model of Klibanoff et al. (2006). Consider an infinite horizon environment. Time is denoted by $t = 0, 1, 2, \dots$. The state space in each period is denoted by S . Thus the full state space is given by S^∞ . At time t , the agent's information consists of history $s^t = \{s_0, s_1, s_2, \dots, s_t\}$ with $s_0 \in S$ given and $s_t \in S$. The agent ranks adapted consumption plans $C = (C_t)_{t \geq 0}$. That is, C_t is a measurable function of s^t . The agent is ambiguous about the probability distribution on the full state space. This uncertainty is described by a parameter z in the space Z . The parameter z can be interpreted in several different ways. It could be an unknown model parameter, a discrete indicator of alternative models, or a hidden state that evolves over time in a regime-switching process (Hamilton (1989)).

Each parameter z gives a probability distribution π_z . This distribution is updated by Bayes' Rule to deliver $\pi_z(\cdot | s^t)$ conditioned on information s^t . The agent has a prior μ over

the parameter z . The posterior $\mu(\cdot|s^t)$ is updated by Bayes' Rule. At any time t , conditioned on information s^t , the agent's ambiguity preferences are represented by the following utility function:

$$V_t(C; s^t) = u(C_t) + \beta \phi^{-1} \left(\int_Z \phi \left(\int_S V_{t+1}(C; s^t, s_{t+1}) d\pi_z(s_{t+1}|s^t) \right) d\mu(z|s^t) \right), \quad (5)$$

where $\beta \in (0, 1)$ is the discount factor, and u and ϕ admit the same interpretation as in the static model. Note that the utility process in (5) is defined recursively, as in Kreps and Porteus (1978) and Epstein and Zin (1989). Thus, it satisfies dynamic consistency. Dynamic consistency is a tractable feature to analyze dynamic problems because the standard dynamic programming technique can be applied.

As in the static model discussed in the previous subsection, we can still derive the equivalence of the robustness, risk-sensitive control, and smooth ambiguity models when ϕ takes the CAAA specification in (2). We can also derive a limiting result as the coefficient of absolute ambiguity aversion goes to infinity. Formally, we can show that the limit satisfies:

$$V_t(C; s^t) = u(C_t) + \beta \inf_{z \in Z} \int_S V_{t+1}(C; s^t, s_{t+1}) d\pi_z(s_{t+1}|s^t). \quad (6)$$

This utility process is similar to the Epstein and Wang (1994) and Epstein and Schneider (2003, 2007a) recursive multiple-priors utility model. In Epstein and Schneider's (2007a) learning model, there exist both a set of priors and a set of likelihoods given a parameter. Both sets are updated by Bayes' Rule in a suitable way. By contrast, for the utility function defined in (6), the set of priors is not updated. The agent always chooses the Dirac measure over the parameter space that minimizes the continuation utility at each date.

We now turn to the following question: Does ambiguity persist in the long run? Klibanoff et al. (2006) show that if ϕ^{-1} is Lipschitz and if the parameter space Z is finite, then the smooth ambiguity utility model converges to the standard expected utility model with the true parameter z^* :

$$V_t(C; s^t) = u(C_t) + \beta \int_S V_{t+1}(C; s^t, s_{t+1}) d\pi_{z^*}(s_{t+1}|s^t). \quad (7)$$

However, if the parameter space is infinite, then such convergence fails. In our application below, we will consider a hidden Markov switching process. In this case, the parameter space is infinite, and thus, ambiguity persists even in the long run.

2.3. Utility Gradient and Pricing Kernel

To study asset pricing implications, it is useful to introduce utility gradient (Duffie and Skiadas (1994)). A utility gradient of the utility process (V_t) at the consumption plan C is an adapted

process g_z such that for every adapted process h ,

$$\lim_{\delta \downarrow 0} \frac{V_0(C + \delta h) - V_0(C)}{\delta} = \mathbb{E} \left[\sum_{t=0}^{\infty} g_{t,z} h_t \right].$$

Note that we use the notation g_z to indicate that the utility gradient depends on an known parameter z because the agent has partial information. In Appendix A, we show the following:

Proposition 1 *Suppose that ϕ and u are differentiable.⁶ Then the utility gradient is given by*

$$g_{t,z} = \times_{\tau=1}^t \frac{\phi'(\mathbb{E}_{\pi_{\tau-1,z}}[V_{\tau}(C)])}{\phi'(\phi^{-1}(\mathbb{E}_{\mu_{\tau-1}}[\phi(\mathbb{E}_{\pi_{\tau-1,z}}[V_{\tau}(C)])]))} \beta^t u'(C_t), \quad g_{0,z} = u'(C_0), \quad (8)$$

where $\mathbb{E}_{\pi_{\tau-1,z}}$ and $\mathbb{E}_{\mu_{\tau-1}}$ denote the conditional expectation operators with respect to the distributions $\pi_z(\cdot|s^{\tau-1})$ and $\mu(\cdot|s^{\tau-1})$, respectively.

As is standard in the literature, we call the intertemporal marginal rate of substitution $g_{t+1,z}/g_{t,z}$ at consumption plan C the pricing kernel at C or the stochastic discount factor at C . Given the preceding utility gradient, we can derive the following pricing kernel for the recursive smooth ambiguity model:

$$M_{t+1,z} = \frac{\phi'(\mathbb{E}_{\pi_{t,z}}[V_{t+1}(C)])}{\phi'(\phi^{-1}(\mathbb{E}_{\mu_t}[\phi(\mathbb{E}_{\pi_{t,z}}[V_{t+1}(C)])]))} \frac{\beta u'(C_{t+1})}{u'(C_t)}. \quad (9)$$

The last term in (9) gives the pricing kernel for the standard time-additive expected utility model. The first term reveals the effect of ambiguity aversion. It is this term that generates interesting asset pricing implications.

3. Asset Pricing Implications

In this section, we study the asset pricing implications of the smooth ambiguity model by analyzing a Lucas-style pure-exchange economy (Lucas (1978)). Our model is based on the fully rational model of Cecchetti (2000, Section II) with two departures: (i) we introduce learning, and (ii) we incorporate ambiguity.

3.1. The Economy

There is a representative agent in the economy. The agent trades a risky and ambiguous stock with unit supply and a riskfree bond with zero supply.⁷ The stock pays dividends D_t in period

⁶A transversality condition given in Appendix A must also be satisfied.

⁷We can easily generalize this model to incorporate multiple risky assets.

$t = 0, 1, 2, \dots$. The dividend process is governed by a Markov regime-switching process,

$$\log \left(\frac{D_{t+1}}{D_t} \right) = \kappa_{z_{t+1}} + \sigma \varepsilon_{t+1}, \quad D_0 \text{ given}, \quad (10)$$

where ε_t is iid standard normal and the state $z_t \in \{1, 2, \dots, N\}$ follows a N state Markov chain with transition matrix (λ_{ij}) where $\sum_j \lambda_{ij} = 1$. Here $\kappa_{z_{t+1}}$ denotes the expected growth rate of dividends when the economy in period $t + 1$ is in the state z_{t+1} . Assume $\kappa_1 > \kappa_2 > \dots > \kappa_N$.

Let $R_{e,t+1}$ and $R_{f,t+1}$ denote the gross returns on the stock and the bond between periods t and $t + 1$, respectively. Let W_t denote the period t financial wealth and let ψ_t be the proportion of wealth after consumption invested in the stock. Then the agent's budget constraint is given by

$$W_{t+1} = (W_t - C_t) R_{m,t+1}, \quad (11)$$

where the market return $R_{m,t+1}$ is given by

$$R_{m,t+1} = \psi_t R_{e,t+1} + (1 - \psi_t) R_{f,t+1}. \quad (12)$$

We assume that the agent does not observe the state of the economy. He learns about it given his information about the history of dividends $s^t = \{s_0, s_1, \dots, s_t\}$, where $s_t = D_t$. To model learning within the standard expected utility model, one must specify a subjective prior over the hidden state and a conditional distribution of data given a state. The prior is updated by Bayes' Rule to deliver a posterior. The posterior and the conditional distribution can be reduced to a predictive distribution over the observable data. Thus, the model with learning is observationally equivalent to a complete information model without learning.⁸ This Bayesian theory of learning precludes ambiguity about hidden states. We will show in Section 4 that it has only modest quantitative implications for asset returns. To incorporate ambiguity about hidden states, we assume that the agent's preferences are represented by the recursive smooth ambiguity utility function defined in (5), where the parameter $z = (z_t)_{t \geq 1}$ describes the state of the economy and the parameter space is given by $Z = \{1, 2, \dots, N\}^\infty$. An important feature of this model is that the preceding compound distribution cannot be reduced. We will show in Section 4 that this irreducibility and ambiguity aversion have significant quantitative implications for asset returns.

We are ready to define equilibrium. A competitive equilibrium of this economy consists of processes of consumption (C_t) , trading strategies (ψ_t) , and returns $(R_{e,t+1})$, and $(R_{f,t+1})$ such

⁸Brandt et al. (2004), David (1997), and Veronesi (1999, 2000) use this Bayesian theory to study asset pricing models with hidden Markov switching processes. Brandt et al. (2004) also analyze some alternative learning rules to Bayes' Rule.

that: (i) (C_t) and (ψ_t) maximize the agent's utility (5) subject to the budget constraint (11), and (ii) markets clear in that $\psi_t = 1$ and $C_t = D_t$ for all t . Because in equilibrium consumption is equal to dividends, we may directly refer to the dividend process in (10) as the aggregate consumption process.

3.2. State Beliefs

We now describe the evolution of the posterior state beliefs. Let $\mu_t(j) = \Pr(z_{t+1} = j | s^t)$ and $\mu_t = (\mu_t(1), \mu_t(2), \dots, \mu_t(N))$. That is, $\mu_t(j)$ is the conditional probability that the economy at date $t + 1$ is in state j given the history of dividends $s^t = \{D_0, D_1, D_2, \dots, D_t\}$. The prior belief μ_0 is given. We need to derive the updating process of the posterior beliefs μ_t . To this end, we let $\mu_t^t(j) = \Pr(z_t = j | s^t)$. We then have

$$\mu_{t+1}(j) = \sum_{i=1}^N \lambda_{ij} \mu_{t+1}^{t+1}(i). \quad (13)$$

By Bayes' Rule,

$$\mu_{t+1}^{t+1}(i) = \frac{f(\log(D_{t+1}/D_t), i) \mu_t(i)}{\sum_j f(\log(D_{t+1}/D_t), j) \mu_t(j)} \quad (14)$$

where

$$f(y, i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y - \kappa_i)^2}{2\sigma^2}\right] \quad (15)$$

is the density function of the normal distribution with mean κ_i and variance σ^2 .

Combining (13) and (14), we obtain the belief updating process:

$$\mu_{t+1}(j) = B_j(\log(D_{t+1}/D_t), \mu_t),$$

where the belief updating function is given by

$$B_j(y, \mu_t) = \frac{\sum_{i=1}^N \lambda_{ij} f(y, i) \mu_t(i)}{\sum_{i=1}^N f(y, i) \mu_t(i)}.$$

We denote the vector of these functions by $B = (B_1, B_2, \dots, B_N)$. We can then write the belief updating equation as

$$\mu_{t+1} = B(\log(D_{t+1}/D_t), \mu_t). \quad (16)$$

3.3. Optimality and Equilibrium

As is well known from the recursive utility models (Epstein and Zin (1989)), one needs to specify functional forms for primitives in order to derive sharp characterizations of asset pricing implications. In what follows, we consider two tractable specifications.

3.3.1. Log-Exponential Specification

Under the log-exponential specification, we assume $u(C) = \log(C)$, and ϕ is CAAA given by (2). We can use the utility gradient and the pricing kernel derived in Section 2.3 to derive equilibrium restrictions on asset returns. Instead of using this method, we directly solve the agent's optimization problem by dynamic programming theory, and then impose market-clearing conditions. This method is of independent interest because it gives the solution to the agent's optimal consumption and portfolio choice problem.

We choose wealth and beliefs (W_t, μ_t) as state variables. We solve an equilibrium in which the stock return and the riskfree rate are functions of the dividend growth D_{t+1}/D_t and beliefs μ_t . Let $J(W_t, \mu_t)$ be the value function (or indirect utility function) associated with the utility function (5). By a standard dynamic programming argument, we can show that J satisfies the following Bellman equation:

$$J(W_t, \mu_t) = \max_{C_t, \psi_t} \log(C_t) - \beta\theta \log \left(\sum_j \mu_t(j) \exp \left(-\frac{1}{\theta} \mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})] \right) \right), \quad (17)$$

subject to the budget constraint (11) and the belief updating equation (16). Here $\mathbb{E}_{t,j}$ denotes the expectation for the distribution of the dividend growth given in (10) conditional on μ_t and the state $z_{t+1} = j$. This expectation is actually taken with respect to the normally distributed random variable ε_{t+1} given $z_{t+1} = j$. The following proposition characterizes the equilibrium and optimality.

Proposition 2 (i) *The equilibrium stock price and return are given by*

$$P_t = \frac{\beta}{1-\beta} D_t, \quad R_{e,t+1} = R_{m,t+1} = \frac{1}{\beta} \frac{D_{t+1}}{D_t}. \quad (18)$$

(ii) *The equilibrium bond return is given by*

$$\frac{1}{R_{f,t+1}} = \sum_j \mu_t(j) \mathbb{E}_{t,j} [M_{t+1,j}], \quad (19)$$

where the pricing kernel is given by

$$M_{t+1,j} = \beta \frac{C_t}{C_{t+1}} \frac{\exp \left(-\frac{1}{\theta} \mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})] \right)}{\sum_j \mu_t(j) \exp \left(-\frac{1}{\theta} \mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})] \right)}. \quad (20)$$

(iii) *Given the returns $R_{e,t+1}$ and $R_{f,t+1}$ in parts (i) and (ii), the value function is given by*

$$J(W_t, \mu_t) = \frac{1}{1-\beta} \log(W_t) + G(\mu_t), \quad (21)$$

where the function G satisfies

$$G(\mu_t) = \log(1 - \beta) - \beta\theta \log \left(\sum_j \mu_t(j) \exp \left(-\frac{1}{\theta} \mathbb{E}_{t,j} \left[\frac{\log(\beta R_{m,t+1})}{1 - \beta} + G(\mu_{t+1}) \right] \right) \right). \quad (22)$$

The optimal consumption rule is given by

$$C_t = (1 - \beta) W_t, \quad (23)$$

and the optimal trading strategy ψ_t satisfies

$$0 = \sum_j \mu_t(j) \exp \left(-\frac{1}{\theta} \mathbb{E}_{t,j} \left[\frac{1}{1 - \beta} \log(R_{m,t+1}) + G(\mu_{t+1}) \right] \right) \mathbb{E}_{t,j} \left[\frac{R_{e,t+1} - R_{f,t+1}}{R_{m,t+1}} \right], \quad (24)$$

where $R_{m,t+1} = \psi_t R_{e,t+1} + (1 - \psi_t) R_{f,t+1}$.

An important feature of the log-exponential specification is that the optimal consumption-to-wealth ratio is constant as in the standard logarithmic expected utility model. Given this consumption rule, we use the market-clearing condition to derive

$$C_t = D_t = (1 - \beta) W_t = (1 - \beta) (P_t + D_t).$$

This equation delivers a closed-form solution to the stock price and return given in part (i). This closed-form solution implies that with log-exponential specification, learning and ambiguity aversion do not affect the stock price and return.

To understand the effect of ambiguity aversion on the riskfree rate, we consider the pricing kernel given in equation (20), which can also be derived using equation (9). This equation admits an intuitive interpretation. The first term $\beta C_t / C_{t+1}$ is the pricing kernel for the standard logarithmic expected utility function. This case is specialized when θ goes to infinity. Let $m_{t,j}^*$ denote the second term in equation (20). This term can be interpreted as a Radon-Nikodym derivative with respect to μ_t since $\sum_j \mu_t(j) m_{t,j}^* = 1$. We can then rewrite equation (19) as

$$\frac{1}{R_{f,t+1}} = \sum_j \hat{\mu}_t(j) \mathbb{E}_{t,j} \left[\beta \frac{C_t}{C_{t+1}} \right],$$

where $\hat{\mu}_t(j) = \mu_t(j) m_{t,j}^*$ is the distorted posterior probability of state j . Thus, the effect of ambiguity aversion on the riskfree rate under the log-exponential specification is manifested through distorting the posterior beliefs about the hidden states. Importantly, the expression of $m_{t,j}^*$ given in (20) reveals that the agent puts relatively more weight on smaller continuation values than larger values in the distorted probability distribution. Thus, an increase in the

degree of ambiguity aversion implies a first-order stochastic dominated shift of state beliefs. This pessimism induces the agent to save more for future consumption. Thus, it lowers the riskfree rate and raises the equity premium.

The preceding interpretation is related to the robustness and risk-sensitive control theory developed by Hansen and Sargent discussed in Section 2.1. Formally, we can show that the Bellman equation in (17) is equivalent to

$$J(W_t, \mu_t) = \max_{C_t, \psi_t} \log(C_t) + \beta \left(\min_{m_t \geq 0, \mathbb{E}_{\mu_t}[m_t] = 1} \mathbb{E}_{\mu_t} [m_t \mathbb{E}_{t,j} \{J(W_{t+1}, \mu_{t+1})\}] + \theta \mathbb{E}_{\mu_t} [m_t \log(m_t)] \right),$$

where m_t is a Radon-Nikodym derivative with respect to μ_t defined on the set of states $\{1, 2, \dots, N\}$. This derivative m_t distorts the state beliefs μ_t . One can show that the minimizing derivative is given by $m_t(j) = m_{t,j}^*$ defined earlier.

It is interesting to consider the limiting case when $\theta \rightarrow 0$. As discussed in Section 2.2, the model in this case reduces to a version of the recursive multiple-priors utility:

$$J(W_t, \mu_t) = \max_{C_t, \psi_t} \log(C_t) + \beta \min_j \mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})]. \quad (25)$$

That is, the agent exhibits extreme ambiguity aversion by choosing the worst continuation utility value. In terms of the distorted beliefs interpretation, the agent views the state with the lowest continuation value has probability 1. Taking limit in (20) shows that the pricing kernel for this model is equal to $\beta C_t / [C_{t+1} \mu_t(N)]$ for state N with the lowest continuation values, and to zero, otherwise. Thus, it follows from (19) that the limiting riskfree rate is given by

$$r_{f,t+1} = R_{f,t+1} - 1 = \frac{1}{\beta} \exp(\kappa_N - 0.5\sigma^2) - 1, \quad (26)$$

where κ_N denotes the lowest expected growth rate of consumption. This equation gives the lower bound of the mean riskfree rate. This lower bound corresponds to the value obtained from Abel's (2002) formulation of uniform pessimism in which the agent's perceived distribution of consumption growth is a first-order stochastically dominated shift of the objective distribution. Thus, our model provides a foundation of Abel's modeling of pessimism. Our model, however, cannot capture Abel's (2002) modeling of doubt which is characterized by a mean-preserving spread of the objective distribution of consumption growth.

We now take the asset returns as exogenously given and consider the effect of ambiguity aversion on the agent's portfolio choice decision. We use $R_{m,t+1} = W_{t+1} / (W_t - C_t)$ and (21) to rewrite (24) as

$$0 = \sum_j \mu_t(j) \exp\left(-\frac{1}{\theta} \mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})]\right) \mathbb{E}_{t,j} \left[\frac{R_{e,t+1} - R_{f,t+1}}{R_{m,t+1}} \right]. \quad (27)$$

Dividing this equation by $\sum_j \mu_t(j) \exp\left(-\frac{1}{\theta} \mathbb{E}_{t,j}[J(W_{t+1}, \mu_{t+1})]\right)$, we obtain

$$0 = \sum_j \hat{\mu}_t(j) \mathbb{E}_{t,j} \left[\frac{R_{e,t+1} - R_{f,t+1}}{\psi_t(R_{e,t+1} - R_{f,t+1}) + R_{f,t+1}} \right]. \quad (28)$$

This equation reveals that an ambiguous averse agent behaves as an expected utility agent with distorted state beliefs $\hat{\mu}_t$. Suppose the bond and excess returns are positive and the excess returns are higher for higher growth state. Because an increase in the ambiguity aversion parameter from $1/\theta_1$ to $1/\theta_2$ implies a first-order stochastic dominated shift of the state beliefs from $\hat{\mu}_t^1$ to $\hat{\mu}_t^2$, we can show that

$$0 = \sum_j \hat{\mu}_t^1(j) \mathbb{E}_{t,j} \left[\frac{R_{e,t+1} - R_{f,t+1}}{\psi_t^1(R_{e,t+1} - R_{f,t+1}) + R_{f,t+1}} \right] \geq \sum_j \hat{\mu}_t^2(j) \mathbb{E}_{t,j} \left[\frac{R_{e,t+1} - R_{f,t+1}}{\psi_t^1(R_{e,t+1} - R_{f,t+1}) + R_{f,t+1}} \right].$$

Thus, we deduce that a more ambiguity averse agent demand less stocks in that $\psi_t^2 \leq \psi_t^1$. As shown earlier, in general equilibrium this reduction in demand increases equity premium and this increase is due to a decrease in the riskfree rate only, with the stock return unchanged.

3.3.2. Power-Power Specification

We now turn to the power-power specification in which $u(C) = C^{1-\gamma}/(1-\gamma)$, and ϕ is CRAA given by (3). Here $1 \neq \gamma > 0$ is the relative risk aversion parameter.⁹ The following proposition characterizes equilibrium returns.

Proposition 3 (i) *The equilibrium stock price and return are given by*¹⁰

$$P_t = \varphi(\mu_t) D_t, \\ R_{e,t+1} = R_{m,t+1} = \frac{D_{t+1}}{D_t} \frac{1 + \varphi(\mu_{t+1})}{\varphi(\mu_t)}, \quad (29)$$

where the function φ satisfies

$$1 = \sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[\frac{1 + \varphi(\mu_{t+1})}{\varphi(\mu_t)} \beta \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \right] \right)^{1-\alpha}. \quad (30)$$

⁹Note that the case of $u(C) = \log(C)$ is not nested here. When $\gamma > 1$, the utility is not well defined for all values of α , e.g., $\alpha = 0.5$. To overcome this problem, we may follow Epstein and Zin (1989) and define an ordinally equivalent utility function as

$$V_t(C) = \left[C_t^{1-\gamma} + \beta \left(\sum_j \mu_t(j) (\mathbb{E}_{t,j}[V_{t+1}^{1-\gamma}(C)])^{1-\alpha} \right)^{\frac{1}{1-\alpha}} \right]^{\frac{1}{1-\gamma}}.$$

This formulation does not change our asset pricing results.

¹⁰A transversality condition must be satisfied, which insures that the value function is finite and the price-dividend ratio is positive. This condition is satisfied for all numerical solutions in Section 4.

(ii) The equilibrium bond return is given by

$$\frac{1}{R_{f,t+1}} = \sum_j \mu_t(j) \mathbb{E}_{t,j} [M_{t+1,j}], \quad (31)$$

where the pricing kernel is given by

$$M_{t+1,j} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\mathbb{E}_{t,j} \left[R_{m,t+1} \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right] \right)^{-\alpha}. \quad (32)$$

This proposition demonstrates that unlike the log-exponential specification, both learning and ambiguity aversion affect the stock return. This effect is manifested through the price-dividend ratio φ which is a function of state beliefs μ_t . This function varies with the risk aversion parameter γ and the ambiguity aversion parameter α as revealed by equation (30).

Turn to the pricing kernel given in equation (32). The first term on its right-hand side gives the pricing kernel for the standard power expected utility model with $\alpha = 0$. The second term captures the effect of ambiguity aversion. Unlike (20) under the log-exponential specification, this term cannot be interpreted as a Radon-Nikodym derivative. As a result, the equivalence to the robustness and risk-sensitive formulations discussed in Hansen (2007) does not hold here. We also observe from (30) and (32) that the effect of ambiguity aversion depends crucially on the risk aversion parameter γ . We will show numerically in Section 4 that ambiguity aversion has different effects for the $\gamma > 1$ case and for the $\gamma < 1$ case.

We can derive the pricing kernel in (32) using the utility gradient approach discussed in Section 2.3. To this end, we need the following:

Proposition 4 *Let the returns $R_{e,t+1}$ and $R_{f,t+1}$ be given in Proposition 3 parts (i) and (ii). Then the value function is given by*

$$J(W_t, \mu_t) = \frac{W_t^{1-\gamma}}{1-\gamma} G(\mu_t). \quad (33)$$

The optimal consumption rule is given by

$$C_t = W_t [G(\mu_t)]^{-1/\gamma}, \quad (34)$$

and the optimal trading strategy ψ_t satisfies

$$0 = \sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} G(\mu_{t+1}) \right] \right)^{-\alpha} \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{-\gamma} G(\mu_{t+1}) (R_{e,t+1} - R_{f,t+1}) \right] \right), \quad (35)$$

where $R_{m,t+1} = \psi_t R_{e,t+1} + (1 - \psi_t) R_{f,t+1}$ and G satisfies

$$G(\mu_t) = [G(\mu_t)]^{\frac{\gamma-1}{\gamma}} + \beta \left(1 - [G(\mu_t)]^{-\frac{1}{\gamma}}\right)^{1-\gamma} \left(\sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} G(\mu_{t+1}) \right] \right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}. \quad (36)$$

Finally, the pricing kernel in (32) is equal to

$$\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \frac{(\mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})])^{-\alpha}}{\left(\sum_j \mu_t(j) \left[(\mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})])^{1-\alpha} \right] \right)^{\frac{-\alpha}{1-\alpha}}}. \quad (37)$$

Taking the stock and bond returns as exogenously given, Proposition 4 characterizes the agent's optimal consumption and portfolio rules. To understand the effect of ambiguity aversion on the optimal portfolio, we follow a similar procedure described in the log-exponential case to rewrite (35) as

$$0 = \sum_j \hat{\mu}_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{-\gamma} G(\mu_{t+1}) (R_{e,t+1} - R_{f,t+1}) \right] \right), \quad (38)$$

where

$$\hat{\mu}_t(j) = \mu_t(j) \frac{(\mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})])^{-\alpha}}{\sum_j \mu_t(j) (\mathbb{E}_{t,j} [J(W_{t+1}, \mu_{t+1})])^{-\alpha}}.$$

Equation (38) reveals that the ambiguous averse agent behaves as an expected utility agent with the distorted beliefs $\hat{\mu}_t$. As in the log-exponential case, a more ambiguity averse agent attaches more weight to the smaller continuation values. Because we know from the standard risk analysis that a first-order stochastic dominated shift in the distribution of the stock returns does not necessarily reduce the demand for the stock depending on the risk aversion coefficient γ , a more ambiguity averse agent does not necessarily demand less stocks. Gollier (2006) finds a similar result in a static model and gives conditions for a comparative statics result. Our dynamic model with learning is more complicated and does not permit such a characterization.

4. Quantitative Results

We first calibrate our model and describe stylized facts. We then study properties of unconditional and conditional moments of returns generated by our model. Our model does not admit an explicit analytical solution. We thus solve the model numerically using the projection method (Judd (1998)) and run Monte Carlo simulations to compute model moments. For comparison, we also solve two benchmark models. Benchmark model I is the fully rational model with complete information studied by Cecchetti et al. (2000). Benchmark model II incorporates learning and is otherwise the same as benchmark model I. This model is similar to Veronesi (1999, 2000) and its solution is specialized by setting the ambiguity aversion parameter to zero.

4.1. Calibration and Stylized Facts

We calibrate our model at the annual frequency. Because our model is based on Cecchetti et al. (2000), we use their estimates for the consumption process. Cecchetti et al. (2000) apply Hamilton's maximum likelihood method to estimate parameters of the two-state regime-switching process given in (10) using the annual per capita US consumption data covering the period 1890-1994. Table 1 reproduces their estimates. This table reveals that the high-growth state is highly persistent, with consumption growth in this state being 2.251 percent. The economy spends most of the time in this state with the unconditional probability of being in this state given by $(1 - \lambda_{22}) / (2 - \lambda_{11} - \lambda_{22}) = 0.96$. The low-growth state is moderately persistent, but very bad, with consumption growth in this state being -6.785 percent.

[Insert Table 1 Here]

We reproduce data moments estimated by Cecchetti et al. (2000) in Table 2. Panel A of this table reveals that the mean values of equity premium and riskfree rate are given by 5.75 and 2.66 percent, respectively.¹¹ In addition, the volatility of equity premium is 19.02 percent. These values are hard to match in a standard asset-pricing model under reasonable calibration. This fact is often referred to as the equity premium, riskfree rate and equity volatility puzzles (see Campbell (1999) for a survey). Panel A of Table 2 also reports that the equity premium and the riskfree rate are negatively correlated with the correlation coefficient -0.24 . Panel B of Table 2 reports that the log dividend yield predicts long-horizon realized excess returns. It also shows that the regression slope and R^2 increase with the return horizon. This return predictability puzzle is first documented by Campbell and Shiller (1988b) and Fama and French (1988a). Panel B of Table 2 also reports variance ratio statistics for the equity premium. These ratios are generally less than 1 and fall with the horizon. This evidence suggests that excess returns are negatively serially correlated, or asset prices are mean reverting (Fama and French (1988b) and Poterba and Summers (1988)).

In addition to the preceding stylized facts reported in Table 2, we will use our model to explain three other stylized facts: (i) persistent and countercyclical variation in conditional volatility of stock returns (Bollerslev et al. (1992)), (ii) procyclical variation in price-dividend ratios (Campbell and Shiller (1988a)), and (iii) countercyclical variation in conditional expected equity premia (Campbell and Shiller (1988a,b) and Fama and French (1989)).

[Insert Table 2 Here]

¹¹We follow Cecchetti et al. (2000) and report arithmetic average returns in both data and model solutions. Mehra and Prescott (1985) also report arithmetic averages.

To explain the above facts, we need to calibrate baseline preference parameters. As argued by Mehra and Prescott (1985) and Kocherlakota (1996), we require β to be between zero and one and γ to be between zero and ten. However, we do not have any information about the magnitude of the degree of ambiguity aversion. For the log-exponential specification, we set $\beta = 0.940$ and $1/\theta = 1.292$ to match the first moments of the equity premium and the riskfree rate reported in Table 2. For the power-power specification, we set $\beta = 0.944$, $\gamma = 0.647$, and $\alpha = 48.367$ to match the means of the equity premium and the riskfree rate and their correlation coefficient reported in Table 2. Based on the preceding calibrated baseline parameter values, we will study the cyclical behavior of returns and examine the comparative statics effects of risk aversion and ambiguity aversion on unconditional and conditional moments of returns. We will also investigate the role of learning quantitatively.

4.2. Unconditional Moments of Returns

We start by discussing the unconditional moments of returns generated from our model. We consider two different utility specifications studied in Section 3.

4.2.1. Log-Exponential Specification

Table 3 reports the results for the log-exponential case. Panel A reports the results for the baseline parameter values. To examine the comparative static properties of β and $1/\theta$, Panels B-D report results for different values of β and $1/\theta$. These panels reveal that both the mean riskfree rate and the mean stock return decrease with β for a fixed value of $1/\theta$. They also reveal that the riskfree rate decreases with $1/\theta$ for fixed β . The intuition is that a more ambiguity averse agent attaches more weight to lower values of continuation utilities. Thus, he saves more for future consumption, resulting in a lower riskfree rate. As shown in Proposition 2 and Table 3, ambiguity aversion has no effect on the stock return because the consumption-wealth ratio is constant as in the standard logarithmic utility model. Consequently, we can choose a low value for β to match the high stock return and then choose a high value for $1/\theta$ to match the low riskfree rate.

[Insert Table 3 Here]

An alternative way to understand the equity premium puzzle is to study the Hansen-Jagannathan bound or the market price of uncertainty (Hansen and Jagannathan (1991)) defined as the ratio of the standard deviation to the mean of the pricing kernel. This bound is close to zero in models without ambiguity as revealed by the first row in each of panels B-D. This

bound rises significantly as we increase the ambiguity aversion parameter $1/\theta$. In particular, it is equal to 3.792 for our baseline parameter values, while it is equal to 0.037 for benchmark model II without ambiguity.

Column 6 of Table 3 reveals that ambiguity may raise volatility of equity premium, but by a very small amount. For the baseline parameter values, the model implied volatility 3.853% of the equity premium is too low, compared to the data value 19.02% reported in Table 2. The intuition follows from the closed-form solution to the stock return in equation (18). This equation shows that the volatility of the stock return is determined by the volatility of consumption growth since the price-dividend ratio is constant. The latter volatility is extremely low in the data as reported in Table 1.

To study the role of learning, we compare our model with benchmark models I and II. Because the stock return is the same for these three models as shown in Proposition 2, we focus on the riskfree rate. Consistent with the findings reported by Cecchetti. (2000), Columns 8-9 of Table 3 show that the equity premium μ_{eq}^* is too low and the riskfree rate r_f^* is too high in benchmark model I. We decompose the riskfree r_f in our model into three components:

$$r_f = r_f^* + (r_f^L - r_f^*) + (r_f - r_f^L), \quad (39)$$

where r_f^L is the mean riskfree rate in benchmark model II. Column 10 of Table 3 reports the second component $\Delta r_f^L = r_f^L - r_f^*$ which measures the effect of the standard Bayesian learning without ambiguity. This column reveals that learning lowers the riskfree rate, but by a negligible amount. Column 11 reports the third component $\Delta r_f = r_f - r_f^L$. This component accounts for the effect of ambiguity aversion and learning under ambiguity. It reveals that the reduction of the riskfree rate is attributed almost exclusively to ambiguity.

It is interesting to consider the limiting case where $1/\theta$ converges to infinity. In this case, our model reduces to a version of the recursive multiple-priors model. In our simulations reported in Table 3, we find that this limit is approached very quickly. For example, when $1/\theta = 2$ and $\beta = 0.98$, we can verify that the implied riskfree rate is extremely close to the analytical solution given in (26). This analytical solution is obtained when the agent pessimistically believes that consumption grows according to the rate in the low-growth state. This extreme pessimism cannot match the first moments of the equity premium and riskfree rate simultaneously by choosing one parameter β only.

4.2.2. Power-Power Specification

Since the log-exponential specification implies that the price-dividend ratio is constant and the stock return is equal to consumption growth discounted by the subjective discount rate, this

specification cannot deliver interesting dynamics of stock returns and equity premium. We now turn to the power-power specification and restrict attention to this case in the remainder of Section 4.

Panel A of Table 4 reports results for the baseline parameter values $\beta = 0.944$, $\gamma = 0.647$, and $\alpha = 48.367$. Panels B-F reports comparative static results. First, as in the log-exponential case, both the riskfree rate and the stock return decrease with the subjective discount factor β . Second, the first rows of Panels B-F of Table 4 reveal that an increase in the risk aversion parameter γ from 0 to 3.0 raises both the riskfree rate and the equity premium for benchmark model II with $\alpha = 0$. We also experiment with many other parameter values for $\beta \in (0, 1)$, $\gamma \in (0, 10)$ and $\alpha = 0$. But we are unable to match both the low riskfree rate and the high equity premium reported in Table 2 simultaneously. This result shows that benchmark model II with Bayesian learning cannot resolve the equity premium and riskfree rate puzzles. We will return to this point below.

We now consider the role of ambiguity aversion with $\alpha > 0$. Table 4 reveals that the effects of ambiguity aversion are quite different for the cases with $\gamma < 1$ and $\gamma > 1$.¹² For the $\gamma < 1$ case, an increase in α lowers the riskfree rate and raises the stock return, and hence raises the equity premium. The intuition is that a more ambiguity averse agent saves more for future consumption and invests less in the stock. It is this property that permits us to find parameter values to match the first moments of the riskfree rate and the equity premium.

[Insert Table 4 Here]

More formally, when $\gamma < 1$, a more ambiguous averse agent puts a higher weight on the low-growth state in the pricing kernel distortion given in (32) than on the high-growth state. To help understand the intuition, we consider the extreme case where α is very large. In this case, the agent puts the maximal weight $1/\mu_t(2)$ on the low-growth state and zero weight on the high-growth state in order for equation (30) to be satisfied. That is, the expression

$$\left(\mathbb{E}_{t,j} \left[\frac{1 + \varphi(\mu_{t+1})}{\varphi(\mu_t)} \beta \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \right] \right)^{1-\alpha}$$

increases to $1/\mu_t(2)$ for $j = 2$ and decreases to 0 for $j = 1$. Note that φ varies with α and the expectation in the preceding expression does not have a limit. We can then use equations (29),

¹²Applying Epstein and Schneider's (2007b) recursive multiple-priors model with learning, Leippold et al. (2005) find a similar result analytically in a continuous-time framework.

(31), and (32) to show that when we increase α , the riskfree rate decreases to the value

$$\begin{aligned}
r_{f,t+1} &= R_{f,t+1} - 1 \simeq \frac{\mathbb{E}_{t,2} \left[\frac{1 + \varphi(\mu_{t+1})}{\varphi(\mu_t)} \beta \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \right]}{\mathbb{E}_{t,2} \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right]} - 1 \\
&= \frac{1}{\beta} \exp(\gamma\kappa_2 - 0.5\gamma^2\sigma^2) \mathbb{E}_{t,2} \left[\frac{1 + \varphi(\mu_{t+1})}{\varphi(\mu_t)} \beta \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \right] - 1.
\end{aligned} \tag{40}$$

We numerically verify that the expectation term is always less than 1, which allows us to find a value of α to match the low riskfree rate observed in the data.

By contrast, for the $\gamma > 1$ case, an increase in α raises both the mean riskfree rate and the mean stock return. To understand the opposite effect on the riskfree rate, we examine equation (30). This equation reveals that when $\gamma > 1$, the marginal utility discount on consumption growth $(C_{t+1}/C_t)^{-\gamma}$ dominates so that a more ambiguity averse agent with $\gamma > 1$ believes that the high-growth state is less valuable. Thus, he puts a relatively higher weight on the high-growth state in the pricing kernel distortion than on the low-growth state. Consequently, he saves less, resulting in a higher riskfree rate. We verify numerically that as we increase α , the riskfree rate increases to the value given by equation (40) with the low-growth state 2 replaced with the high-growth state 1. By extensive numerical simulations, we are unable to find parameter values for $\alpha > 0$, $\beta \in (0, 1)$, and $\gamma > 1$ to match the first moments of the riskfree rate and the equity premium simultaneously. In addition, we also find that the mean equity premium may decrease with the degree of ambiguity aversion. In summary, we find that ambiguity aversion plays a very different role than risk aversion and may not reinforce risk aversion. Gollier (2006) also makes this point in a simple static model without learning.

We next consider the role of learning. Table 5 reports the decomposition of the riskfree rate, the stock return and the equity premium similar to equation (39). This table reveals that for benchmark model II without ambiguity, the standard Bayesian learning lowers the riskfree rate, but by a negligible amount. In addition, learning raises the stock return for $\gamma < 1$ and lowers it for $\gamma > 1$, both by a negligible amount. In addition, the standard risk premium component and the learning component account for a negligible amount of the mean equity premium. The latter component may be negative for a sufficiently high degree of risk aversion. Thus, a standard Bayesian learning model may worsen the equity premium puzzle. In a related continuous-time model without ambiguity, Veronesi (2000) finds a similar result. The intuition follows from the following equation with $\alpha = 0$:

$$\mathbb{E}_t [R_{e,t+1} - R_{f,t+1}] = \frac{-Cov_t(M_{t+1}, R_{e,t+1})}{\mathbb{E}_t [M_{t+1}]}. \tag{41}$$

Specifically, following a negative innovation in consumption growth, the agent revises his beliefs about the future consumption growth downward. He then increases his hedging demand for the stock to avoid low levels of consumption in the future. This effect tends to increase the stock price. When the agent is sufficiently risk averse, this effect dominates the opposing effect on the stock price due to the initial negative shock to consumption growth, leading to a negative correlation between consumption growth and the stock return. This negative correlation results in a positive covariance between the pricing kernel $M_{t+1} = \beta (C_{t+1}/C_t)^{-\gamma}$ and the stock return, and hence a negative equity premium. Veronesi (2000) also shows that there is an upper bound of the equity premium and this bound is independent of the degree of risk aversion. Thus, the high equity premium observed in the data cannot be matched by choosing a high value of the risk aversion parameter. By contrast, in our learning model ambiguity aversion can help explain both the equity premium and riskfree rate puzzles because the ambiguity averse agent tends to invest less in the stock, thereby counteracting the preceding hedging demand effect. Table 5 shows that ambiguity aversion reduces the riskfree rate for $\gamma < 1$ and increases the mean stock return. In addition, the ambiguity premium component $\Delta\mu_{eq}$ reported in Column 10 accounts for a significant amount of the equity premium.

[Insert Table 5 Here]

We now turn to the second moments. Table 4 reveals that for the baseline parameter values, the implied volatility of the riskfree rate and the stock return is too low compared with the data reported in Table 2. Thus, for both the log-exponential and power-power cases, our model with learning under ambiguity cannot resolve the equity volatility puzzle (Shiller (1981)).¹³ To understand the intuition, we observe from equation (29) that the stock return volatility comes from two components: (i) volatility of consumption growth, and (ii) variation in the price-dividend ratio φ due to learning about the hidden state. Clearly the first component is too small as reported in Table 2. To examine the second component, we plot the price-dividend ratio as a function of the posterior beliefs about the high-growth state $\mu_t(1)$ in Figure 1. This figure reveals that the price-dividend ratio does not vary much as $\mu_t(1)$ moves from 0 to 1. Thus, learning under ambiguity does not contribute much to the unconditional stock return volatility.

[Insert Figure 1 Here]

¹³In a model similar to ours, but without ambiguity and with Epstein-Zin utility, Brandt et al. (2004) show that the model implied stock volatility is too low compared to data for several alternative learning rules.

4.3. Price-Dividend Ratio

We now analyze the properties of the price-dividend ratio function φ under the power-power specification by varying the ambiguity and risk aversion parameters, respectively, around the baseline parameter values. Panel a of Figure 1 presents this function for different values of α , holding other parameters fixed at the baseline values. It reveals two properties. First, the price-dividend ratio is an increasing and convex function of the posterior probability of the high-growth state $\mu_t(1)$. The intuition is similar to that described by Veronesi (1999) who assumes expected exponential utility. When times are good ($\mu_t(1)$ is close to 1), a bad piece of news decreases $\mu_t(1)$, and hence decreases future expected dividends. But it also increases the agent's uncertainty about consumption growth since $\mu_t(1)$ is now closer to 0.5, which gives approximately the maximal conditional volatility of the posterior probability of the high-growth state in the next period, as shown in Panel a of Figure 2. Since the agent wants to be compensated for bearing more risk, they will require an additional discount on the stock price. Thus, the price reduction due to a bad piece of news in good times is higher than the reduction in expected future dividends. By contrast, suppose the agent believes times are bad and hence $\mu_t(1)$ is close to zero. A good piece of news increases the expected future dividends, but also raises the agent's perceived uncertainty since it moves $\mu_t(1)$ closer to 0.5. Thus, the price-dividend ratio increases, but not as much as it would in a present-value model. The second property of Panel a of Figure 1 is that an increase in the degree of ambiguity aversion lowers the price-dividend ratio because it induces the agent to invest less in the stock. In addition, the increase in the degree of ambiguity aversion raises the curvature of the price-dividend ratio function, thereby helping increase the equity volatility. In the special case of benchmark model II with $\alpha = 0$, we can prove analytically that the price-dividend ratio is a linear function of the state beliefs. Note that the curvature is not very sensitive to the ambiguity aversion parameter. As a result, ambiguity aversion cannot have a large impact on equity volatility, and hence cannot resolve the equity volatility puzzle.

[Insert Figure 2 Here]

Panel b of Figure 1 presents the price-dividend ratio function for different values of γ , holding other parameters fixed at the baseline values. It reveals that the price-dividend ratio is an increasing function of $\mu_t(1)$ for the $\gamma < 1$ case, as in Panel a, while it is a decreasing function for the $\gamma > 1$ case. This property implies that the price-dividend ratio is countercyclical when $\gamma > 1$, which is inconsistent with empirical evidence. The intuition is that the cyclicity of the price-dividend ratio depends on the relative importance of two offsetting wealth and

substitution effects. Suppose the economy is in a boom and the agent believes that the high-growth state is more likely. Anticipating future high consumption growth, the wealth effect means that the agent wants to consume more and sell assets today. This effect lowers the stock price. On the other hand, future high consumption growth, implies that the relative price of future goods is lower. Thus, the intertemporal substitution effect implies that the agent wants to consume less and buy more assets. This effect raises the stock price. When the elasticity of intertemporal substitution $1/\gamma > 1$, the substitution effect dominates so that the price-dividend ratio is procyclical and increases with the posterior probability $\mu_t(1)$ of the high-growth state. When $\gamma > 1$, the opposite result holds. A similar finding appears in Bansal and Yaron (2004), Brandt et al. (2004), and Cecchetti et al. (1990).

Panel b of Figure 1 also reveals that the price-dividend ratio rises as we increase γ from 0.2 to 0.8, and then falls as we increase γ further to 1.5 and 3.0. The intuition is that an increase in γ when $\gamma < 1$ reinforces the substitution effect which induces the agent to buy more assets. By contrast, an increase in γ when $\gamma > 1$ reinforces the wealth effect which induces the agent to sell more assets.

4.4. Time-Varying Equity Premia and Equity Volatility

Although our model cannot generate high unconditional volatility of the stock return, in this subsection we show that our model under the power-power specification can generate several interesting patterns of conditional moments of returns observed in the data such as countercyclicality of conditional expected equity premia and conditional equity volatility.

Panels a and b of Figure 3 plot the conditional expected equity premium as a function of the posterior probability of the high-growth state for different values of α and γ , respectively. Several properties emerge. First, this function is hump-shaped. This shape follows from the convexity of the price-dividend ratio function presented in Figure 1 and the shape of the belief function presented in Figure 2. We observe from Figure 2a that the conditional volatility of the posterior probability of the high-growth state in the next period takes the maximum value when the current posterior probability of the high-growth state is around 0.5. Since during recessions, the agent's perceived uncertainty about the hidden state is high, it follows from Figure 3 that the conditional expected equity premium is high during recessions. As a result, our model can generate the countercyclical variation in equity premia observed in the data. Second, the curvature of this function increases with α , implying that ambiguity aversion helps explain the time-varying equity premium. By contrast, in benchmark model II with $\alpha = 0$, the conditional expected equity premium is almost flat with the posterior probability of the high-

growth state. Consequently, it cannot generate time-varying expected equity premia. Third, the conditional expected equity premium may decrease with the risk aversion parameter γ when γ rises from 0.2 to 3.0 for a wide range of the posterior probabilities of the high-growth state. It can even take negative values for $\gamma > 1$. The intuition follows from equation (41) and is similar to that discussed in Section 4.2 for benchmark model II with $\alpha = 0$. That is, when γ is high enough, the agent's hedging demand is so high that consumption growth and the stock return are negatively correlated. Consequently, the pricing kernel and the stock return is positively correlated, leading to negative conditional equity premia by (41).

[Insert Figure 3 Here]

Panels a and b of Figure 4 plot the conditional volatility of stock returns as a function of the posterior probability of the high-growth state for different values of α and γ , respectively. The shape of this function follows from the shape of the price-dividend ratio function presented in Figure 1 and the shape of the belief function presented in Figure 2. As a result, the conditional stock volatility function is hump-shaped, with the maximum attained at a value of posterior close to 0.5. Thus, our model is consistent with the empirical evidence that uncertainty and conditional volatility of stock returns are higher during recessions. We next observe from Panel b of Figure 2 that the agent's beliefs are persistent in the sense that if he believes the high-growth state today has a high probability, then he expects the high-growth state tomorrow also has a high probability on average. It follows from Figure 4 that changes in return volatility tend to be persistent, giving rise to the volatility clustering phenomenon documented by Bollerslev et al. (1992). Finally, Figure 4a reveals that ambiguity aversion ($\alpha > 0$) raises conditional volatility of stock returns, compared to benchmark model II with $\alpha = 0$. But there is no monotonic relationship between conditional stock return volatility and the degree of ambiguity aversion. Figure 4b shows that conditional volatility of stock returns first decreases with the risk aversion parameter γ , and then increases with γ when γ increases from 0.2 to 3.0. Consequently, ambiguity aversion amplifies the stock return volatility, while risk aversion may dampen it.

[Insert Figure 4 Here]

Figure 5 illustrates the time-varying properties of equity premia and stock return volatility by a Monte Carlo simulation. Panel a plots a time series of dividend growth simulated using (10). Panel b plots the time series of the posterior probability of the high-growth state $\mu_t(1)$, computed using (16). It reveals that in most of the time the agent believes that the economy is in the high-growth state in that $\mu_t(1)$ is close to 1. After a few negative innovations in consumption

growth, the agent believes the low-growth state is more likely in that $\mu_t(1)$ decreases and is close to 0.5. At this value, the agent's perceived uncertainty about the high-growth state in the next period is the highest. Using the simulated series of consumption growth and the posterior probability, we can compute the series of conditional volatility of stock returns and conditional expected equity premium. We plot these series in Panels c and d of Figure 5, respectively. From these panels, we can see that both the conditional volatility of stock returns and conditional expected equity premium are time-varying and move with business cycles countercyclically.

[Insert Figure 5 Here]

4.5. Serial Correlation and Predictability of Returns

To examine the ability of our model to generate the serial correlation and predictability of returns reported in Table 2, we compare our model with benchmark models I and II. Table 6 reports the model implied values of the variance ratios, the regression slope and the R^2 's, at horizons of 1, 2, 3, 5, and 8 years based on the baseline parameter values given in Table 4. To account for the small sample bias in these statistics, we generate them using 10,000 Monte Carlo experiments as described in Cecchetti et al. (2000).

From Table 6, we observe that benchmark model I produces variance ratios close to 1 and R^2 's close to zero, and thus, it cannot generate the mean reversion and predictability of excess returns reported in Table 2. Cecchetti et al. (2000) find the same result and show that a model with distorted beliefs can help explain this stylized fact. We now consider benchmark model II with Bayesian learning. Because of learning, the agent's posterior state beliefs are a state variable that drives asset returns. These posterior state beliefs respond to consumption innovations and evolve over time by Bayes updating. Because the change of state beliefs is persistent, the price-dividend ratio is also persistent and positively serially correlated. Intuition suggests that learning should help explain the mean reverting and predictability pattern. However, Table 6 reports that benchmark model II with Bayesian learning helps little quantitatively. Brandt et al. (2004) find a similar result. We finally consider our model in which we introduce ambiguity into benchmark model II. Table 6 reveals that while all three models can generate the pattern that the regression slope increases with the horizon and the variance ratio decreases with the horizon, our model with learning under ambiguity produces much more significant quantitative effects. In particular, compared to benchmark model II, our model implied values of the regression slope and R^2 are higher, while our model implied values of variance ratio are smaller. However, our model still cannot replicate the same numbers estimated from the data reported in Panel B of Table 2.

[Insert Table 6 Here]

5. Conclusion

In this paper, we have proposed a consumption-based asset pricing model that can match the first moments of the equity premium and riskfree rate observed in the data. In addition, our model can generate a variety of dynamic asset pricing phenomena, including the procyclical variation of price-dividend ratios, the countercyclical variation of equity premia and equity volatility, and the mean reversion and long horizon predictability of excess returns. There are two main ingredients of our model. First, we assume that consumption growth is governed by a hidden Markov regime-switching process. The representative agent formulates posterior probabilities on the hidden states by observing past consumption data. This posterior distribution is a state variable that drives asset return dynamics. Second, and most importantly, the agent is ambiguous about the hidden state in consumption growth. His preferences are represented by the smooth ambiguity model proposed by Klibanoff et al. (2005, 2006). The agent's degree of ambiguity aversion plays a key role in determining asset returns. It helps propagate and amplify shocks to the dynamics of asset returns. Without ambiguity aversion, our model cannot match the first moments of the equity premium and riskfree rate observed in the data, and cannot generate significant time-varying equity premia. One limitation of our model is that it cannot match equity volatility observed in the data. One potential way to resolve this puzzle is to separate consumption from dividends since dividends are much more volatile than consumption or to introduce leverage (see Abel (1999) and Cecchetti et al. (1993)). We leave this extension for future research.

Other models can also simultaneously generate the unconditional moments and dynamics of asset returns observed in the data. For example, Campbell and Cochrane (1999) introduce a slow moving habit or time-varying subsistence level into a standard power utility function. As a result, the agent's risk aversion is time varying. Bansal and Yaron (2004) apply the Epstein-Zin recursive utility function, and incorporate fluctuating volatility and a persistent component in consumption growth. Nevertheless, we view our model as a first step toward understanding the quantitative implications of learning under ambiguity for asset returns. We have shown that our model can go a long way to explain many asset pricing puzzles. Much work still remains to be done. For example, how to distinguish between risk aversion and intertemporal substitution in our model and how to empirically estimate parameters of ambiguity aversion, risk aversion, and intertemporal substitution would be important future research topics.

Appendix

A Proofs

Proof of Proposition 1: Define the function $\Phi_t(\delta) = V_t(C + \delta h)$. Then $\Phi'_t(0)$ is equal to

$$\lim_{\delta \downarrow 0} \frac{V_t(C + \delta h) - V_t(C)}{\delta}.$$

Using (5), we can derive

$$\begin{aligned} \Phi'_t(0) &= u'(C_t) h_t + \beta \frac{\mathbb{E}_{\mu_t} \phi'(\mathbb{E}_{\pi_{t,z}}[V_{t+1}])}{\phi'(\phi^{-1}(\mathbb{E}_{\mu_t}[\phi(\mathbb{E}_{\pi_{t,z}}[V_{t+1}])]))} \mathbb{E}_{\pi_{t,z}}[\Phi'_{t+1}(0)] \\ &= u'(C_t) h_t + \beta \mathbb{E}_t \left[\frac{\phi'(\mathbb{E}_{\pi_{t,z}}[V_{t+1}])}{\phi'(\phi^{-1}(\mathbb{E}_{\mu_t}[\phi(\mathbb{E}_{\pi_{t,z}}[V_{t+1}])]))} \Phi'_{t+1}(0) \right], \end{aligned}$$

where \mathbb{E}_t denotes the conditional expectation given history s^t . Solving the preceding equation forward recursively starting from $t = 0$, we can derive the utility gradient g_z given in (8). We also need a transversality condition, $\lim_{T \rightarrow \infty} \mathbb{E}[g_{T,z} \Phi'_T(0)] = 0$. The expression for the pricing kernel (9) follows from this equation. Q.E.D.

Proof of Proposition 2: We conjecture that the value function takes the form in (21). Substituting this conjectured value function and the budget constraint (11) into the Bellman equation (17) yields

$$\begin{aligned} J(W_t, \mu_t) &= \max_{C_t, \psi_t} \log(C_t) + \frac{\beta}{1 - \beta} \log(W_t - C_t) \\ &\quad - \beta \theta \log \left(\sum_j \mu_t(j) \exp \left(-\frac{1}{\theta} \mathbb{E}_{t,j} [A \log(R_{m,t+1}) + G(\mu_{t+1})] \right) \right). \end{aligned} \tag{A.1}$$

The first-order condition with respect to C_t delivers the consumption rule (23). Substituting this consumption rule back into the Bellman equation (A.1) and using the conjectured value function, we obtain equation (22).

Using the equilibrium market-clearing condition, we can derive

$$C_t = D_t = (1 - \beta) W_t = (1 - \beta) (P_t + D_t).$$

This equation gives the equilibrium stock price and hence the stock return given in part (i).

We turn to the riskfree rate. Substituting (12) for $R_{m,t+1}$ into (A.1) and taking first-order condition with respect to the trading strategy ψ_t , we obtain equation (24). Imposing the market

clearing condition $\psi_t = 1$ so that $R_{m,t+1} = R_{e,t+1}$, we can solve for $R_{f,t+1}$ to obtain (19) with the pricing kernel given by

$$M_{t+1,j} = \frac{\beta C_t}{C_{t+1}} \frac{\exp\left(-\frac{1}{\theta} \mathbb{E}_{t,j} \left[\frac{1}{1-\beta} \log(R_{m,t+1}) + G(\mu_{t+1}) \right]\right)}{\sum_j \mu_t(j) \exp\left(-\frac{1}{\theta} \mathbb{E}_{t,j} \left[\frac{1}{1-\beta} \log(R_{m,t+1}) + G(\mu_{t+1}) \right]\right)}. \quad (\text{A.2})$$

We finally use the value function (21) and $R_{m,t+1} = W_{t+1}/(W_t - C_t)$ from (12) to rewrite (A.2) as equation (20). Q.E.D.

Proof of Propositions 3 and 4: We conjecture that the value function takes the form in (33) and optimal consumption is linear in wealth $C_t = a_t W_t$, where a_t and G are to be determined. By a standard dynamic programming argument, the value function J satisfies the following Bellman equation

$$\frac{W_t^{1-\gamma}}{1-\gamma} G(\mu_t) = \max_{C_t, \psi_t} \frac{C_t^{1-\gamma}}{1-\gamma} + \beta \left(\sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[\frac{W_{t+1}^{1-\gamma}}{1-\gamma} G(\mu_{t+1}) \right] \right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}.$$

Substituting the budget constraint into the preceding equation yields:

$$\begin{aligned} & \frac{W_t^{1-\gamma}}{1-\gamma} G(\mu_t) \\ &= \max_{C_t, \psi_t} \frac{C_t^{1-\gamma}}{1-\gamma} + \beta \frac{(W_t - C_t)^{1-\gamma}}{1-\gamma} \left(\sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} G(\mu_{t+1}) \right] \right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}. \end{aligned} \quad (\text{A.3})$$

Taking first-order condition with respect to consumption delivers:

$$C_t^{-\gamma} = \beta (W_t - C_t)^{-\gamma} \left(\sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} G(\mu_{t+1}) \right] \right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}.$$

Substituting the conjectured consumption rule $C_t = a_t W_t$ into the preceding two equation, we have

$$G(\mu_t) = a_t^{1-\gamma} + \beta (1 - a_t)^{1-\gamma} \left(\sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} G(\mu_{t+1}) \right] \right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}, \quad (\text{A.4})$$

and

$$a_t^{-\gamma} = \beta (1 - a_t)^{-\gamma} \left(\sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} G(\mu_{t+1}) \right] \right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}. \quad (\text{A.5})$$

It follows from these two equations that

$$G(\mu_t) = a_t^{-\gamma} = \left(\frac{C_t}{W_t} \right)^{-\gamma}. \quad (\text{A.6})$$

Substituting (A.6) for a_t into (A.4) yields (36).

Substitute (A.6) into (A.5) to deduce that

$$a_t^{-\gamma} = \beta (1 - a_t)^{-\gamma} \left(\sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} \left(\frac{C_{t+1}}{W_{t+1}} \right)^{-\gamma} \right] \right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}. \quad (\text{A.7})$$

Use the budget constraint and the consumption rule $C_t = a_t W_t$ to derive:

$$W_{t+1} = R_{m,t+1} (W_t - C_t) = R_{m,t+1} C_t \frac{1 - a_t}{a_t}. \quad (\text{A.8})$$

Substitute this equation into (A.7) to derive

$$a_t^{-\gamma} = \beta (1 - a_t)^{-\gamma} \left(\sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} \left(\frac{a_t}{1 - a_t} \right)^{-\gamma} \left(\frac{C_{t+1}}{C_t R_{m,t+1}} \right)^{-\gamma} \right] \right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}. \quad (\text{A.9})$$

Simplifying yields

$$1 = \sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1} \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right] \right)^{1-\alpha}. \quad (\text{A.10})$$

We turn to the riskfree rate. Substituting (12) for $R_{m,t+1}$ into (A.3) and taking first-order condition with respect to the trading strategy ψ_t , we obtain equation (35). Using (A.6) to substitute $G(\mu_{t+1})$ into this equation yields:

$$0 = \sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{1-\gamma} \left(\frac{C_{t+1}}{W_{t+1}} \right)^{-\gamma} \right] \right)^{-\alpha} \left(\mathbb{E}_{t,j} \left[R_{m,t+1}^{-\gamma} (R_{e,t+1} - R_{f,t+1}) \left(\frac{C_{t+1}}{W_{t+1}} \right)^{-\gamma} \right] \right).$$

Substituting W_{t+1} given in (A.8) into the preceding equation yields:

$$0 = \sum_j \mu_t(j) \left(\mathbb{E}_{t,j} \left[R_{m,t+1} \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right] \right)^{-\alpha} \left(\mathbb{E}_{t,j} \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (R_{e,t+1} - R_{f,t+1}) \right] \right). \quad (\text{A.11})$$

Equations (A.10) and (A.11) imply that the pricing kernel is given by (32), and that the stock return and the riskfree rate satisfy the Euler equations (29) and (31).

Conjecture that the price dividend ratio is given by

$$P_t = \varphi(\mu_t) D_t,$$

where φ is a function to be determined. Then in equilibrium,

$$R_{e,t+1} = R_{m,t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{D_{t+1}}{D_t} \frac{1 + \varphi(\mu_{t+1})}{\varphi(\mu_t)}.$$

Substituting the preceding equation into (A.10) yields equation (30).

From equations (33)-(34), we can rewrite the value functions at date t as $J(W_t, \mu_t) = W_t C_t^{-\gamma} / (1 - \gamma)$. We now use this expression and (9) to rewrite (37) as

$$\frac{\left(\mathbb{E}_{t,j} \left[W_{t+1} C_{t+1}^{-\gamma} \right]\right)^{-\alpha}}{\left(\mathbb{E}_{\mu_t} \left[\left(\mathbb{E}_{t,j} \left[W_{t+1} C_{t+1}^{-\gamma} \right]\right)^{1-\alpha} \right]\right)^{\frac{-\alpha}{1-\alpha}}} \beta \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma}.$$

Using equation (12), we rewrite the first term on the right-hand side as

$$\begin{aligned} & \frac{\left(\mathbb{E}_{t,j} \left[(W_t - C_t) R_{m,t+1} C_{t+1}^{-\gamma} \right]\right)^{-\alpha}}{\left(\mathbb{E}_{\mu_t} \left[\left(\mathbb{E}_{t,j} \left[(W_t - C_t) R_{m,t+1} C_{t+1}^{-\gamma} \right]\right)^{1-\alpha} \right]\right)^{\frac{-\alpha}{1-\alpha}}} \\ &= \frac{\left(\mathbb{E}_{t,j} \left[R_{m,t+1} \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right]\right)^{-\alpha}}{\left(\mathbb{E}_{\mu_t} \left[\left(\mathbb{E}_{t,j} \left[R_{m,t+1} \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right]\right)^{1-\alpha} \right]\right)^{\frac{-\alpha}{1-\alpha}}}, \end{aligned}$$

where we cancel out $(W_t - C_t)$ and multiply the expressions in the numerator and the denominator by $(\beta C_t^\gamma)^{-\alpha}$ to obtain the equality. Finally, we use equations (29) and (30) to deduce that the denominator in the last expression is equal to one. We thus obtain the pricing kernel given in (32). Q.E.D.

B Numerical Method: Two-State Case

For the log-exponential case, we solve the functional equation (22) for G . For the power-power specification, we solve the functional equation (30) for φ . In both cases, we use the projection method described in Judd (1998). Here, we only outline the algorithm for the power-power specification with two states $N = 2$. The algorithm for the log-exponential case is similar.

Let $z \in \{1, 2\}$. Let $p_t = \mu_t(1)$. We approximate the function $\varphi(p_t, 1 - p_t)$ by the function

$$\Phi(p_t) = \sum_{i=0}^n c_i T_i(p_t),$$

where $T_i(p)$ is an n -order Chebyshev function (with the domain adjusted to $[-1, 1]$) and c_0, c_1, \dots, c_n are coefficients to be determined. Let $c = (c_0, c_1, \dots, c_n)$.

We define the function

$$\begin{aligned} H(p_t, j) &= \mathbb{E}_{t,j} \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \left(1 + \sum_{i=0}^n c_i T_i(p_{t+1}) \right) \right] \\ &= \beta \int \exp((1-\gamma)y) \left(1 + \sum_{i=0}^n c_i T_i(B_1(y, p_t)) \right) f(y, j) dy, \end{aligned}$$

where we define the belief transition function

$$B_1(y, x) = \frac{\lambda_{11} f(y, 1) x + (1 - \lambda_{22}) f(y, 2) (1 - x)}{f(y, 1) x + f(y, 2) (1 - x)}.$$

We then obtain the approximated residual function

$$R(p_t; c) = p_t (H(p_t, 1))^{1-\alpha} + (1 - p_t) (H(p_t, 2))^{1-\alpha} - \left(\sum_{i=0}^n c_i T_i(p_t) \right)^{1-\alpha}.$$

Our objective is to make $R(p_t; c)$ close to zero. To this end, we use the collocation method. Let x_j be the $n+1$ roots of the Chebyshev function $T_{n+1}(x)$. We then solve the system of $n+1$ equations

$$R(x_j; c) = 0, \quad j = 1, \dots, n+1,$$

for $n+1$ unknowns (c_0, c_1, \dots, c_n) .

References

- Abel, Andrew B., 1999, Risk Premia and Term Premia in General Equilibrium, *Journal of Monetary Economics* 43, 3-33.
- Abel, Andrew B., 2002, An Exploration of the Effects of Pessimism and Doubt on Asset Returns, *Journal of Economic Dynamics and Control* 26, 1075-1092.
- Anderson, Evan W., Lars Peter Hansen and Thomas J. Sargent, 2003, A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection, *Journal of the European Economic Association*, 1, 68-123.
- Backus, David, Bryan Routledge and Stanley Zin, 2005, Exotic Preferences for Macroeconomists, in Mark Gertler and Kenneth Rogoff (Eds.), *NBER Macroeconomics Annual 2004*, 319-390. Cambridge: MIT Press.
- Bansal, Ravi and Amir Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 59, 1481-1509.
- Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner, 1992, ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics* 52, 5-59.
- Brandt, Michael W., Qi Zeng, and Lu Zhang, 2004, Equilibrium Stock Return Dynamics under Alternative Rules of Learning about Hidden States, *Journal of Economic Dynamics and Control* 28, 1925-1954.
- Brennan, Michael J. and Yihong Xia, 2001, Stock Price Volatility and the Equity Premium, *Journal of Monetary Economics* 47, 249-283.
- Cagetti, Marco, Lars Peter Hansen, Thomas Sargent, and Noah Williams, 2002, Robustness and Pricing with Uncertain Growth, *Review of Financial Studies* 15, 363-404.
- Campbell, John Y., 1999, Asset Prices, Consumption and the Business Cycle, in John B. Taylor, and Michael Woodford, eds.: *Handbook of Macroeconomics*, vol. 1 (Elsevier Science, North-Holland, Amsterdam).
- Campbell, John Y., and John H. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, *Journal of Political Economy* 107, 205-251.
- Campbell, John Y., and Robert J. Shiller, 1988a, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195-227.
- Campbell, John Y. and Robert J. Shiller, 1988b, Stock Prices, Earnings, and Expected Dividends, *Journal of Finance* 43, 661-676.
- Cao Henry, Tan Wang, and Harold Zhang, 2005, Model Uncertainty, Limited Market Participation, and Asset Prices, *Review of Financial Studies* 18, 1219-1251.
- Cecchetti, Stephen G., Pok-sang Lam, and Nelson C. Mark, 1990, Mean Reversion in Equilibrium Asset Prices, *American Economic Review* 80, 398-418.

- Cecchetti, Stephen G., Pok-sang Lam, and Nelson C. Mark, 1993, The Equity Premium and the Riskfree Rate: Matching the Moments, *Journal of Monetary Economics*, 31, 21-45.
- Cecchetti, Stephen G., Pok-sang Lam, and Nelson C. Mark, 2000, Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True, *American Economic Review* 90, 787-805.
- Chen, Zengjin and Larry G. Epstein, 2002, Ambiguity, Risk and Asset Returns in Continuous Time, *Econometrica* 4, 1403-1445.
- David, Alexander, 1997, Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility, *Journal of Financial and Quantitative Analysis* 32, 427-462.
- Detemple, Jerome, 1986, Asset Pricing in a Production Economy with Incomplete Information, *Journal of Finance* 41, 383-391.
- Dothan, Michael U. and David Feldman, 1986, Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy, *Journal of Finance* 41, 369-382.
- Duffie, Darrell and Costis Skiadas, 1994, Continuous-Time Security Pricing: A Utility Gradient Approach, *Journal of Mathematical Economics* 23, 107-132.
- Ellsberg, Daniel, 1961, Risk, Ambiguity and the Savage Axiom, *Quarterly Journal of Economics* 75, 643-669.
- Epstein, Larry G. , 1999, A Definition of Uncertainty Aversion, *Review of Economic Studies* 66, 579-608.
- Epstein, Larry G. and Jianjun Miao, 2003, A Two-Person Dynamic Equilibrium under Ambiguity, *Journal of Economic Dynamics and Control* 27, 1253-1288.
- Epstein, Larry G. and Martin Schneider, 2003, Recursive multiple-priors, *Journal of Economic Theory* 113, 1-31.
- Epstein, Larry G. and Martin Schneider, 2007a, Learning under ambiguity, forthcoming in *Review of Economic Studies*.
- Epstein, Larry G. and Martin Schneider, 2007b, Ambiguity, Information Quality and Asset Pricing, forthcoming in *Journal of Finance*.
- Epstein, Larry G. and Tan Wang, 1994, Intertemporal Asset Pricing under Knightian Uncertainty, *Econometrica* 62, 283-322.
- Epstein, Larry G. and Tan Wang, 1995, Uncertainty, Risk-Neutral Measures and Security Price Booms and Crashes, *Journal of Economic Theory* 67, 40-82.
- Epstein, Larry G. and Stanley Zin, 1989, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* 57, 937-969.

- Fama, Eugene F., and Kenneth R. French, 1988a, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics* 22, 3-25.
- Fama, Eugene F., and Kenneth R. French, 1988b, Permanent and Temporary Components of Stock Prices, *Journal of Political Economy* 96, 246-273.
- Fama, Eugene F., and Kenneth R. French, 1989, Business Conditions and Expected Returns on Stocks and Bonds, *Journal of Financial Economics* 25, 23-49.
- Garlappi, Lorenzo, Raman Uppal and Tan Wang, 2007, Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach, *Review of Financial Studies* 20, 41-81.
- Ghirardato, Paolo and Massimo Marinacci, 2002, Ambiguity Made Precise: A Comparative Foundation, *Journal of Economic Theory* 102, 251-289.
- Gilboa, Itzhak and David Schmeidler, 1989, Maxmin Expected Utility with Non-unique Priors, *Journal of Mathematical Economics* 18, 141-153.
- Gollier, Christian, 2006, Does Ambiguity Aversion Reinforce Risk Aversion? Applications to Portfolio Choices and Asset Prices, working paper, University of Toulouse.
- Hamilton, James D., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica* 57, 357-384.
- Hansen, Lars Peter, 2007, Beliefs, Doubts and Learning: The Valuation of Macroeconomic Risk, Forthcoming in *American Economic Review, Papers and Proceedings*.
- Hansen, Lars Peter and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225-262.
- Hansen, Lars Peter and Thomas J. Sargent, 2001, Robust Control and Model Uncertainty, *American Economic Review* 91, 60-66.
- Hansen, Lars Peter and Thomas J. Sargent, 2006, Recursive Robust Estimation and Control without Commitment, Forthcoming in *Journal of Economic Theory*.
- Hansen, Lars Peter and Thomas J. Sargent, 2007, Fragile Beliefs and the Price of Model Uncertainty, working paper, New York University.
- Hansen, Lars Peter and Thomas J. Sargent, 2008, *Robustness*, Forthcoming, Princeton University Press.
- Hansen, Lars Peter, Thomas J. Sargent, and Thomas Tallarini, 1999, Robust Permanent Income and Pricing, *Review of Economic Studies* 66, 873-907.
- Hansen, Lars Peter, Thomas J. Sargent, G. A. Turmuhambetova, and N. Williams, 2006, Robust Control, Min-Max Expected Utility, and Model Misspecification, *Journal of Economic Theory* 128, 45-90.
- Judd, Kenneth, 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, MA.

- Keynes, Maynard, 1936, *The General Theory of Employment, Interest, and Money*, London Macmillan.
- Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2005, A Smooth Model of Decision Making under Ambiguity, *Econometrica* 73, 1849-1892.
- Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2006, Recursive Smooth Ambiguity Preferences, working paper, Northwestern University and Collegio Carlo Alberto.
- Knight, Frank, 1921, *Risk, Uncertainty, and Profit*, Boston: Houghton Mifflin.
- Kocherlakota, Narayana R., 1996, The Equity Premium: It's Still a Puzzle, *Journal of Economic Literature* 34, 42-71.
- Kreps, David, M. and Evan L. Porteus, 1978, Temporal Resolution of Uncertainty and Dynamic Choice, *Econometrica* 46, 185-200.
- Leippold, Markus, Fabio Trojani, and Paolo Vanini, 2005, Learning and Asset Prices under Ambiguous Information, Forthcoming in *Review of Financial Studies*.
- Liu, Jun, Jun Pan and Tan Wang, 2005, An Equilibrium Model of Rare Event Premia and Its Implication for Option Smirks, *Review of Financial Studies* 18, 131-164.
- Lucas, Robert E., 1978, Asset Prices in an Exchange Economy, *Econometrica* 46, 1429-1445.
- Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, 2006, Ambiguity Aversion, Robustness and the Variational Representation of Preferences, *Econometrica* 74, 1447-1498.
- Maenhout, Pascal J., 2004, Robust Portfolio Rules and Asset Pricing, *Review of Financial Studies* 17, 951-983.
- Mehra, Rajnish and Edward C. Prescott, 1985, The Equity Premium: A Puzzle, *Journal of Monetary Economics* 15, 145-161.
- Miao, Jianjun and Neng Wang, 2007, Risk, Uncertainty, and Option Exercise, working paper, Boston University.
- Poterba, James M. and Summers, Lawrence H., 1988, Mean Reversion in Stock Prices: Evidence and Implications, *Journal of Financial Economics* 22, 27-59.
- Routledge, Bryan and Stanley Zin, 2001, Model Uncertainty and Liquidity, NBER working paper #8683.
- Segal, Uzi, 1987, The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach, *International Economic Review* 28, 175-202.
- Shiller, Robert J., 1981, Do Stock Prices Move too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review* 71, 421-436.
- Tallarini, Thomas D. Jr., 2000, Risk-Sensitive Real Business Cycles, *Journal of Monetary Economics* 45, 507-532.

- Timmerman, Alan G., 1993, How Learning in Financial Markets Generates Excess Volatility and Predictability of Stock Returns, *Quarterly Journal of Economics* 108, 1135-1145.
- Uppal, Raman and Tan Wang, 2003, Model Misspecification and Underdiversification, *Journal of Finance* 58, 2465-2486.
- Veronesi, Pietra, 1999, Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model, *Review of Financial Studies* 12, 975-1007.
- Veronesi, Pietro, 2000, How Does Information Quality Affect Stock Returns? *Journal of Finance* 55, 807-839.
- Wang, Jiang, 1993, A Model of Intertemporal Asset Prices under Asymmetric Information, *Review of Economic Studies* 60, 249-282.
- Weil, Philippe, 1989, The Equity Premium Puzzle and the Risk-Free Rate Puzzle, *Journal of Monetary Economics* 24, 401-421.
- Weitzman, Martin L., 2006, Subjective Expectations and Asset-Return Puzzles, Forthcoming in *American Economic Review*.

Table 1. Maximum likelihood estimates of the consumption process

λ_{11}	λ_{22}	κ_1	κ_2	σ
0.978	0.516	2.251	-6.785	3.127

Notes: The numbers in the last three columns are expressed in percentage. This table is taken from Table 2 in Cecchetti et al. (2000).

Table 2. Stylized facts of equity and short-term bond returns using annual observations from 1871-1993

A. First and second moments as a percentage			
Mean equity premium		μ_{eq}	5.75
Mean risk-free rate		r_f	2.66
Standard deviation			
Equity premium		$\sigma(\mu_{eq})$	19.02
Risk-free rate		$\sigma(r_f)$	5.13
Correlation		$\rho_{eq,f}$	-0.24
B. Predictability and persistence of excess returns			
Horizon	Regression slope	R^2	Variance ratio
1	0.148	0.043	1.000
2	0.295	0.081	1.038
3	0.370	0.096	0.921
5	0.662	0.191	0.879
8	0.945	0.278	0.766

Notes: This table is taken from Table 1 in Cecchetti et al. (2000).

Table 3. Unconditional Moments for the Log-Exponential Case

1	2	3	4	5	6	7	8	9	10	11
$1/\theta$	r_f	$\sigma(r_f)$	r_e	μ_{eq}	$\sigma(\mu_{eq})$	$\frac{\sigma(M)}{E[M]}$	μ_{eq}^*	$r_f = r_f^* + \Delta r_f^L + \Delta r_f$		
								r_f^*	Δr_f^L	Δr_f
Panel A: Baseline parameter values: $\beta = 0.940$, $1/\theta = 1.292$										
	2.660	0.914	8.410	5.750	3.853	3.792	0.133	8.276	-0.002	-5.614
Panel B: $\beta = 0.940$										
0.00	8.274	0.817	8.410	0.136	3.816	0.037	0.133	8.276	-0.002	0.000
0.25	7.990	1.107	8.410	0.420	3.831	0.185	0.133	8.276	-0.002	-0.283
0.50	7.369	1.399	8.410	1.041	3.874	0.560	0.133	8.276	-0.002	-0.904
0.75	6.027	1.477	8.410	2.383	3.905	1.436	0.133	8.276	-0.002	-2.246
1.00	4.318	1.260	8.410	4.092	3.883	2.614	0.133	8.276	-0.002	-3.955
1.25	2.867	0.962	8.410	5.543	3.856	3.644	0.133	8.276	-0.002	-5.407
1.50	1.791	0.699	8.410	6.619	3.846	4.420	0.133	8.276	-0.002	-6.483
1.75	1.028	0.496	8.410	7.382	3.849	4.976	0.133	8.276	-0.002	-7.246
2.00	0.495	0.347	8.410	7.915	3.857	5.367	0.133	8.276	-0.002	-7.778
Panel C: $\beta = 0.98$										
0.00	3.892	0.784	4.023	0.131	3.662	0.037	0.128	3.895	-0.002	0.000
0.25	1.398	1.396	4.023	2.625	3.747	1.675	0.128	3.895	-0.002	-2.494
0.50	-2.582	0.606	4.023	6.605	3.690	4.612	0.128	3.895	-0.002	-6.474
0.75	-4.011	0.206	4.023	8.034	3.712	5.705	0.128	3.895	-0.002	-7.903
1.00	-4.477	0.068	4.023	8.500	3.728	6.064	0.128	3.895	-0.002	-8.369
1.25	-4.628	0.022	4.023	8.651	3.734	6.181	0.128	3.895	-0.002	-8.520
1.50	-4.676	0.007	4.023	8.699	3.737	6.219	0.128	3.895	-0.002	-8.569
1.75	-4.692	0.002	4.023	8.715	3.737	6.231	0.128	3.895	-0.002	-8.584
2.00	-4.697	0.001	4.023	8.720	3.738	6.235	0.128	3.895	-0.002	-8.589
Panel D: $\beta = 0.92$										
0.00	10.668	0.835	10.807	0.139	3.901	0.037	0.136	10.670	-0.003	0.000
0.25	10.476	1.047	10.807	0.331	3.909	0.131	0.136	10.670	-0.003	-0.192
0.50	10.154	1.275	10.807	0.653	3.933	0.314	0.136	10.670	-0.003	-0.514
0.75	9.565	1.465	10.807	1.242	3.968	0.670	0.136	10.670	-0.003	-1.103
1.00	8.567	1.516	10.807	2.240	3.990	1.308	0.136	10.670	-0.003	-2.101
1.25	7.311	1.396	10.807	3.496	3.981	2.146	0.136	10.670	-0.003	-3.357
1.50	6.096	1.186	10.807	4.711	3.958	2.979	0.136	10.670	-0.003	-4.572
1.75	5.064	0.964	10.807	5.743	3.940	3.698	0.136	10.670	-0.003	-5.604
2.00	4.235	0.763	10.807	6.572	3.931	4.283	0.136	10.670	-0.003	-6.433

Notes: Except for the numbers in Columns 1 and 7, all numbers are in percentage. The variables in the first row and columns 2-6 are defined as in Table 2. $\sigma(M)/E[M]$ is the ratio of the standard deviation to the mean of the pricing kernel. r_f^* and μ_{eq}^* are the mean riskfree rate and the mean equity premium for benchmark model I. r_f^L is the mean riskfree rate for benchmark model II. $\Delta r_f^L = r_f^L - r_f^*$ denotes the change of the mean riskfree rate due to learning only. $\Delta r_f = r_f - r_f^L$ denotes change of the riskfree rate due to ambiguity.

Table 4. Comparative Statistics for the Power-Power Case

1	2	3	4	5	6	7	8	9
α	r_f	$\sigma(r_f)$	r_e	$\sigma(r_e)$	μ_{eq}	$\sigma(\mu_{eq})$	$\frac{\mu_{eq}}{\sigma(\mu_{eq})}$	$\frac{\sigma(M)}{E[M]}$
Panel A: Baseline parameter values: $\beta = 0.944$, $\gamma = 0.647$, $\alpha = 48.367$								
	2.660	0.952	8.410	4.581	5.750	4.715	1.219	2.640
Panel B: $\beta = 0.944$, $\gamma = 0.0$								
0.0	5.953	0.000	5.953	4.456	0.000	4.456	0.000	0.000
3.0	5.813	0.189	6.077	4.660	0.265	4.671	0.057	0.096
10.0	4.568	0.845	6.888	5.640	2.320	5.843	0.397	0.711
25.0	-1.736	0.932	10.466	6.037	12.202	6.011	2.030	4.215
60.0	0.004	1.026	13.212	4.927	13.208	4.996	2.643	6.152
Panel C: $\beta = 0.944$, $\gamma = 0.2$								
0.0	6.345	0.160	6.375	4.324	0.030	4.322	0.007	0.007
3.0	6.234	0.308	6.453	4.444	0.220	4.452	0.049	0.078
10.0	5.552	0.786	6.855	4.956	1.303	5.087	0.256	0.447
25.0	0.121	0.674	9.295	5.698	9.175	5.793	1.584	3.265
60.0	-0.970	0.912	11.793	4.865	12.763	4.925	2.592	5.975
Panel D: $\beta = 0.944$, $\gamma = 0.8$								
0.0	7.499	0.648	7.611	3.975	0.112	3.928	0.029	0.029
3.0	7.470	0.686	7.615	3.980	0.145	3.935	0.037	0.040
10.0	7.392	0.778	7.628	3.995	0.236	3.955	0.060	0.082
25.0	7.144	0.999	7.669	4.036	0.526	4.028	0.131	0.225
60.0	5.374	1.348	7.946	4.200	2.572	4.320	0.595	1.265
Panel E: $\beta = 0.944$, $\gamma = 1.5$								
0.0	8.800	1.235	8.986	3.717	0.186	3.516	0.053	0.055
3.0	8.877	1.137	9.013	3.733	0.137	3.546	0.039	0.050
10.0	9.015	0.932	9.059	3.763	0.044	3.604	0.012	0.095
25.0	9.195	0.606	9.114	3.808	-0.081	3.695	-0.022	0.171
60.0	9.360	0.255	9.164	3.858	-0.195	3.807	-0.051	0.233
Panel F: $\beta = 0.944$, $\gamma = 3.0$								
0.0	11.416	2.542	11.647	4.161	0.230	3.213	0.072	0.113
3.0	11.807	2.071	12.069	3.863	0.262	3.248	0.081	0.127
10.0	12.297	1.224	12.416	3.806	0.119	3.507	0.034	0.225
25.0	12.623	0.529	12.611	3.903	-0.012	3.784	-0.003	0.262
60.0	12.776	0.222	12.701	3.983	-0.075	3.936	-0.019	0.263

Notes: Except for numbers in Columns 1, 8 and 9, all numbers are in percentage. The variables in the first row and columns 2-6 are defined as in Table 2. $\sigma(M)/E[M]$ is the ratio of the standard deviation to the mean of the pricing kernel.

Table 5. Decomposition of r_f , r_e and μ_{eq} for the Power-Power Case

1	2	3	4	5	6	7	8	9	10
	$r_f = r_f^* + \Delta r_f^L + \Delta r_f$			$r_e = r_e^* + \Delta r_e^L + \Delta r_e$			$\mu_{eq} = \mu_{eq}^* + \Delta \mu_{eq}^L + \Delta \mu_{eq}$		
α	r_f^*	Δr_f^L	Δr_f	r_e^*	Δr_e^L	Δr_e	μ_{eq}^*	$\Delta \mu_{eq}^L$	$\Delta \mu_{eq}$
Panel A: Baseline parameter values: $\beta = 0.944$, $\gamma = 0.647$, $\alpha = 48.367$									
	7.208	-0.001	-4.547	7.299	0.001	1.111	0.090	0.002	5.657
Panel B: $\beta = 0.944$, $\gamma = 0.0$									
0.0	5.953	0.000	0.000	5.953	0.000	0.000	0.000	0.000	0.000
3.0	5.953	0.000	-0.140	5.953	0.000	0.124	0.000	0.000	0.265
10.0	5.953	0.000	-1.385	5.953	0.000	0.935	0.000	0.000	2.320
25.0	5.953	0.000	-7.689	5.953	0.000	4.513	0.000	0.000	12.202
60.0	5.953	0.000	-5.949	5.953	0.000	7.259	0.000	0.000	13.208
Panel C: $\beta = 0.944$, $\gamma = 0.2$									
0.0	6.345	0.000	0.000	6.375	0.001	0.000	0.030	0.001	0.000
3.0	6.345	0.000	-0.111	6.375	0.001	0.078	0.030	0.001	0.189
10.0	6.345	0.000	-0.793	6.375	0.001	0.479	0.030	0.001	1.273
25.0	6.345	0.000	-6.224	6.375	0.001	2.920	0.030	0.001	9.144
60.0	6.345	0.000	-7.315	6.375	0.001	5.417	0.030	0.001	12.733
Panel D: $\beta = 0.944$, $\gamma = 0.8$									
0.0	7.500	-0.001	0.000	7.610	0.001	0.000	0.110	0.003	0.000
3.0	7.500	-0.001	-0.028	7.610	0.001	0.005	0.110	0.003	0.033
10.0	7.500	-0.001	-0.106	7.610	0.001	0.017	0.110	0.003	0.123
25.0	7.500	-0.001	-0.355	7.610	0.001	0.059	0.110	0.003	0.414
60.0	7.500	-0.001	-2.125	7.610	0.001	0.335	0.110	0.003	2.460
Panel E: $\beta = 0.944$, $\gamma = 1.5$									
0.0	8.806	-0.006	0.000	8.989	-0.003	0.000	0.183	0.003	0.000
3.0	8.806	-0.006	0.077	8.989	-0.003	0.028	0.183	0.003	-0.049
10.0	8.806	-0.006	0.215	8.989	-0.003	0.074	0.183	0.003	-0.141
25.0	8.806	-0.006	0.395	8.989	-0.003	0.128	0.183	0.003	-0.267
60.0	8.806	-0.006	0.559	8.989	-0.003	0.179	0.183	0.003	-0.381
Panel F: $\beta = 0.944$, $\gamma = 3.0$									
0.0	11.444	-0.027	0.000	11.685	-0.038	0.000	0.241	-0.011	0.000
3.0	11.444	-0.027	0.390	11.685	-0.038	0.422	0.241	-0.011	0.032
10.0	11.444	-0.027	0.881	11.685	-0.038	0.769	0.241	-0.011	-0.111
25.0	11.444	-0.027	1.207	11.685	-0.038	0.964	0.241	-0.011	-0.243
60.0	11.444	-0.027	1.359	11.685	-0.038	1.055	0.241	-0.011	-0.305

Notes: Except for the numbers in Column 1, all numbers are in percentage. The variables r_f^* , r_e^* , and μ_{eq}^* are the mean riskfree rate, stock return, and the mean equity premium, respectively, for benchmark model I. The variables r_f^L , r_e^L , and μ_{eq}^L are the mean riskfree rate, stock return, and the mean equity premium, respectively, for benchmark model II. $\Delta r_f^L = r_f^L - r_f^*$ denotes the change of the mean riskfree rate due to learning only. $\Delta r_f = r_f - r_f^L$ denotes change of the riskfree rate due to ambiguity. The other variables Δr_e^L , Δr_e , $\Delta \mu_{eq}^L$, $\Delta \mu_{eq}$, are defined similarly.

Table 6. Predictability and persistence of excess returns

1	2	3	4	5	6	7	8	9	10
Horizon	Baseline parameter values			Benchmark model I			Benchmark model II		
	Slope	R^2	Variance ratio	Slope	R^2	Variance ratio	Slope	R^2	Variance ratio
1	1.040	0.064	1.000	0.443	0.022	1.000	0.776	0.020	1.000
2	1.372	0.074	0.848	0.599	0.025	0.989	1.028	0.022	0.987
3	1.511	0.071	0.769	0.685	0.024	0.979	1.137	0.022	0.975
5	1.635	0.061	0.683	0.780	0.022	0.960	1.346	0.021	0.955
8	1.741	0.053	0.619	0.907	0.020	0.933	1.649	0.020	0.927

Notes: The slope and R^2 are obtained from an OLS regression of the excess returns on the log dividend yield at different horizons. The variance ratio is computed in the same way as Cecchetti (1990, 2000). The reported numbers are the mean values of 10,000 Monte Carlo simulations, each consisting of 123 excess returns and dividend yields.

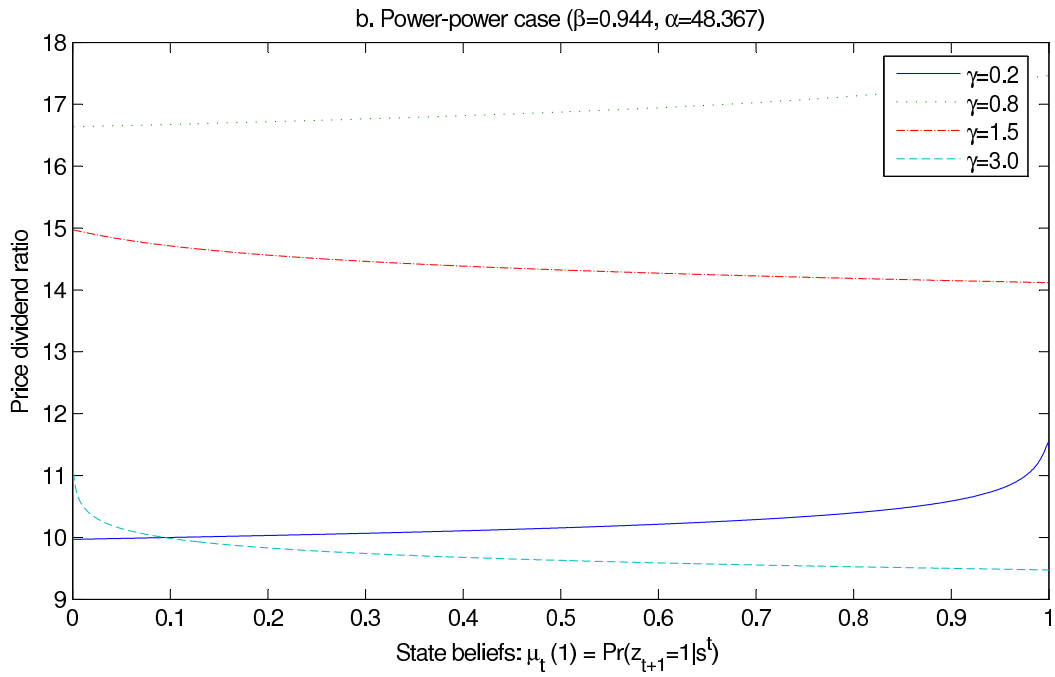
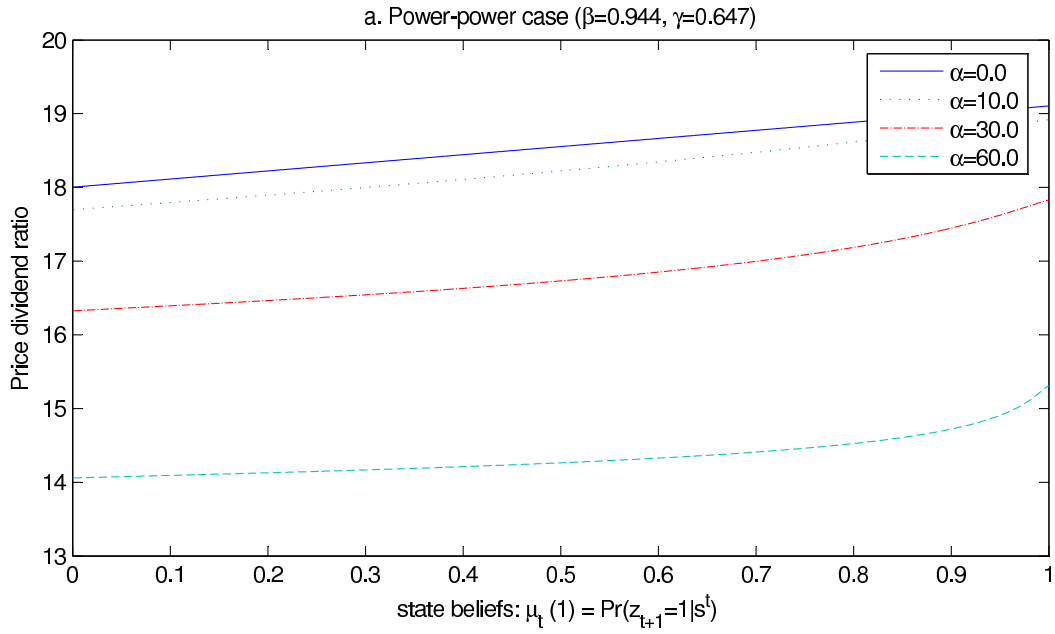


Figure 1: Price dividend ratio as a function of the posterior probability of the high-growth state.

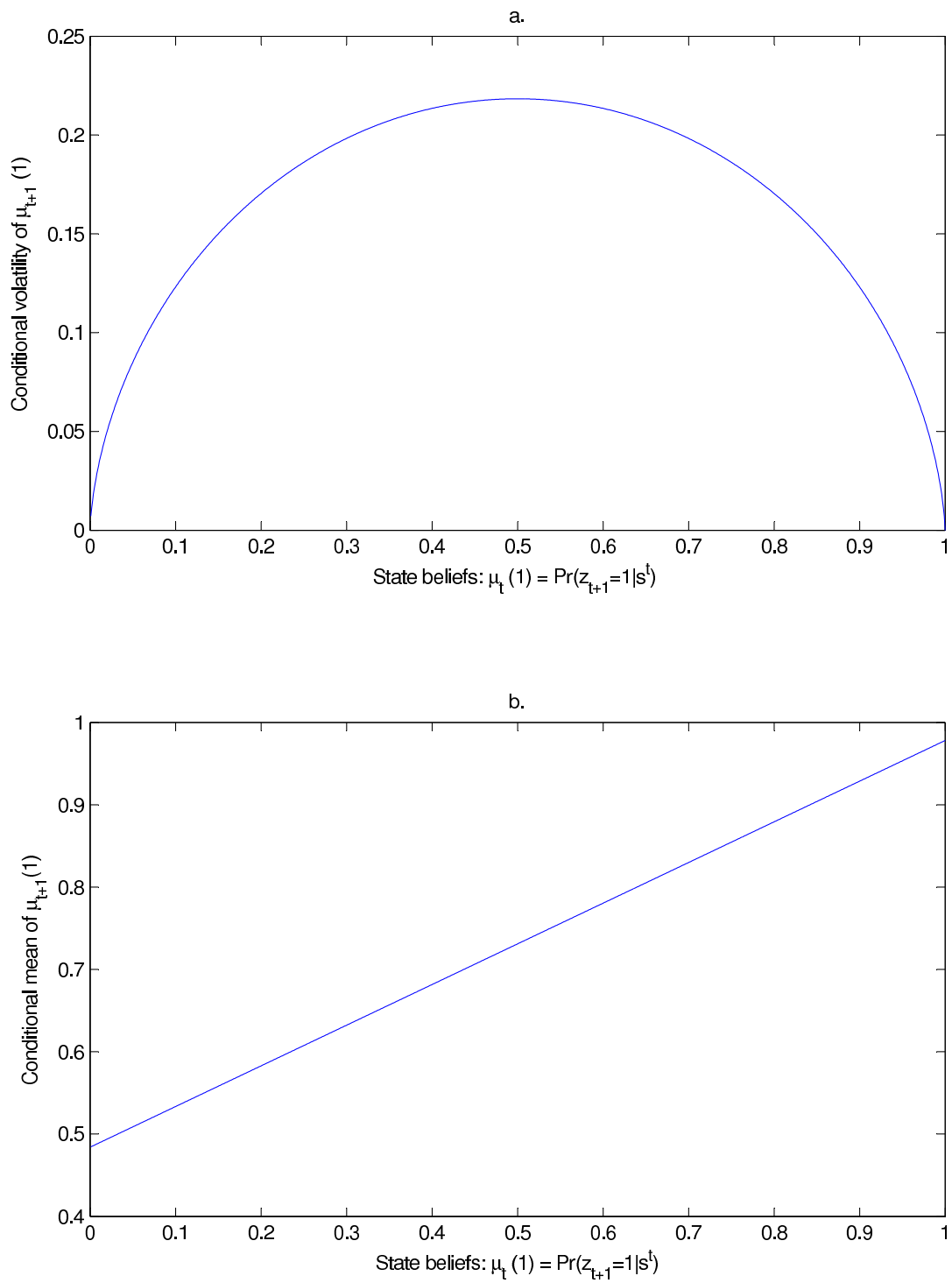


Figure 2: **Conditional mean and volatility of the probability of the high-growth state in the next period as functions of the current state beliefs.**

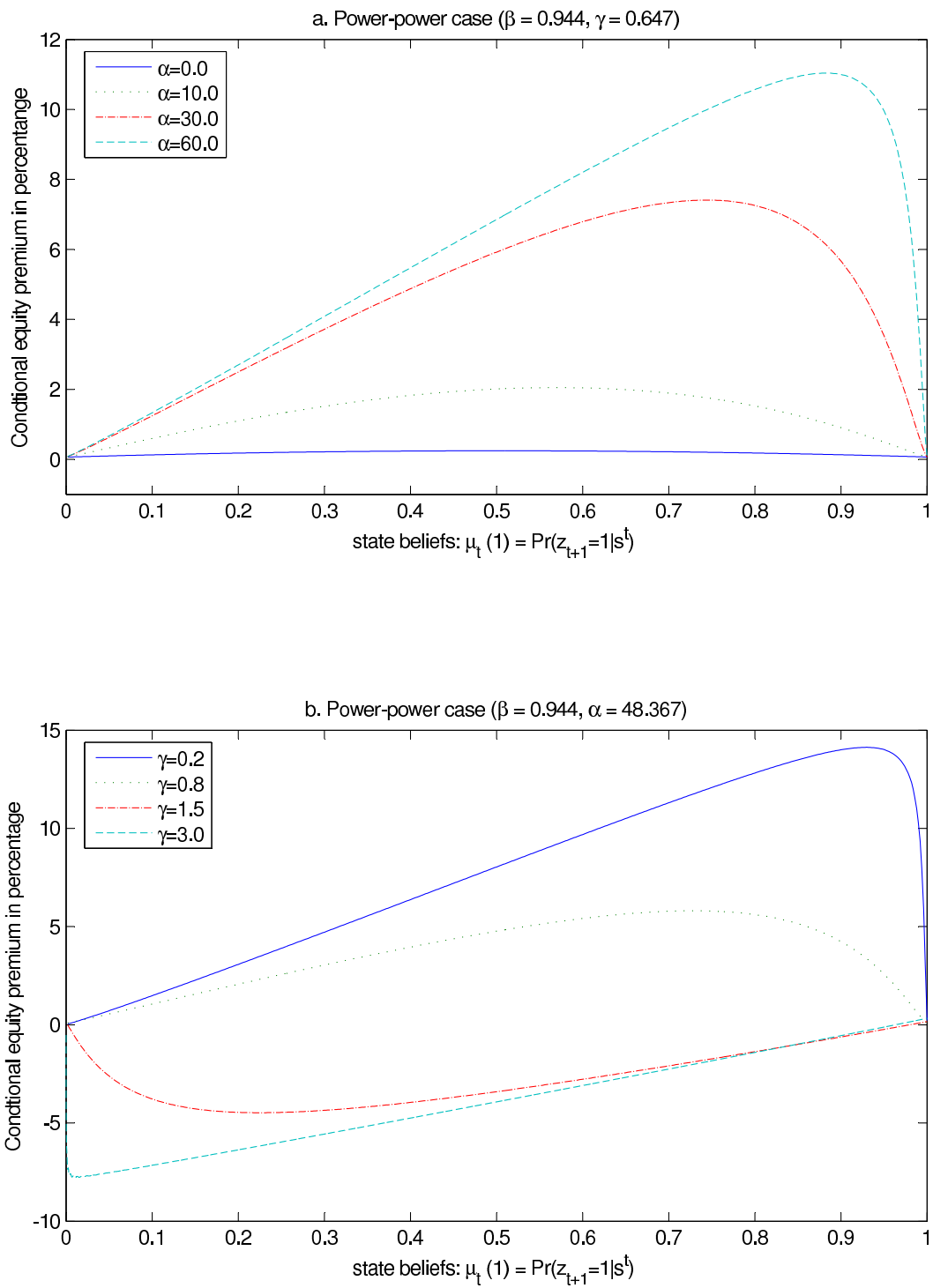


Figure 3: **Conditional expected equity premium as a function of the beliefs about the high-growth state.** Panel a plots this function for different values of the ambiguity aversion parameter α . Panel b plots this function for different values of the risk aversion parameter γ .

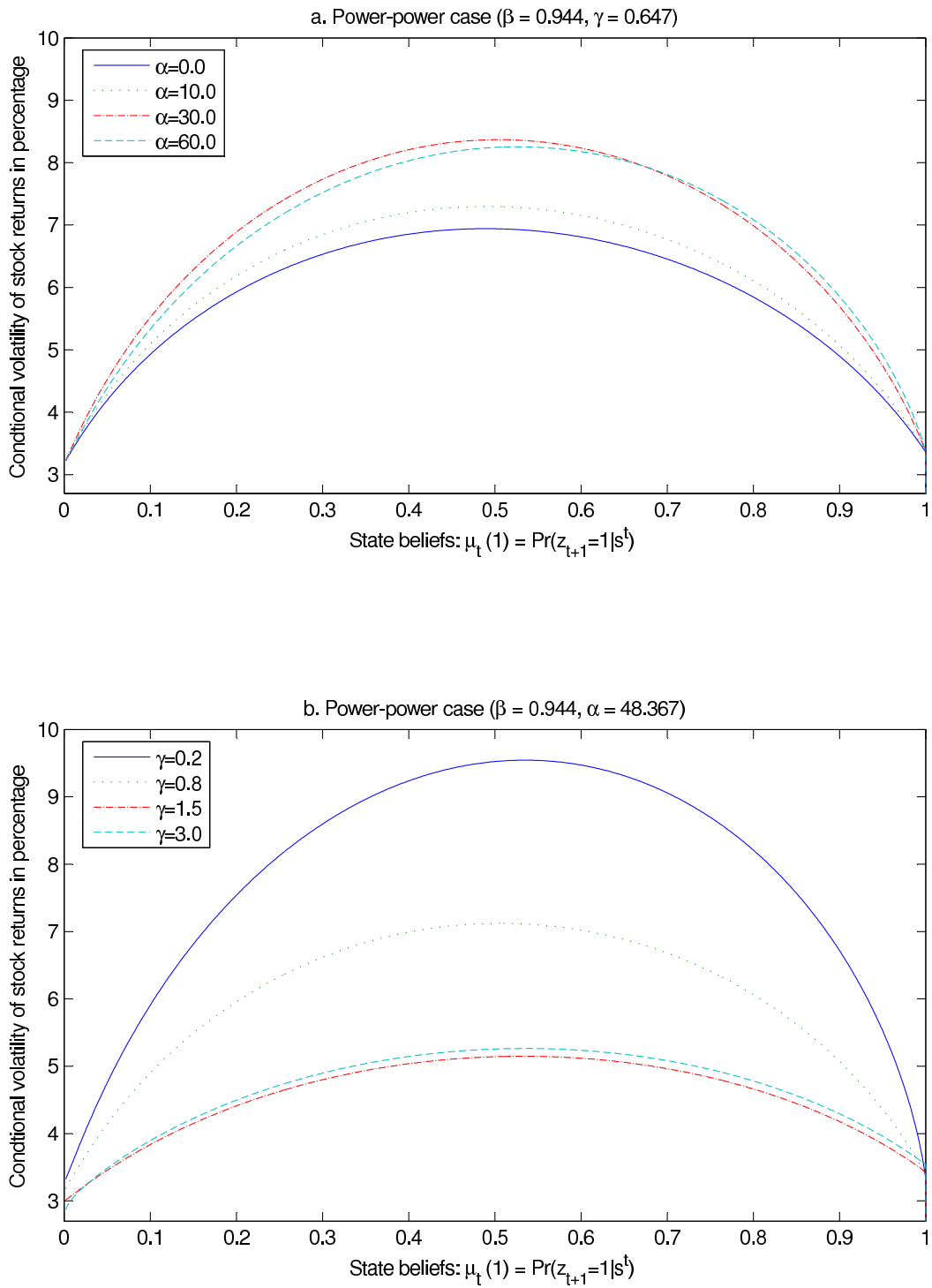


Figure 4: **Conditional volatility of stock returns as a function of the beliefs about the high-growth state.** Panel a plots this function for different values of the ambiguity aversion parameter α . Panel b plots this function for different values of the risk aversion parameter γ .

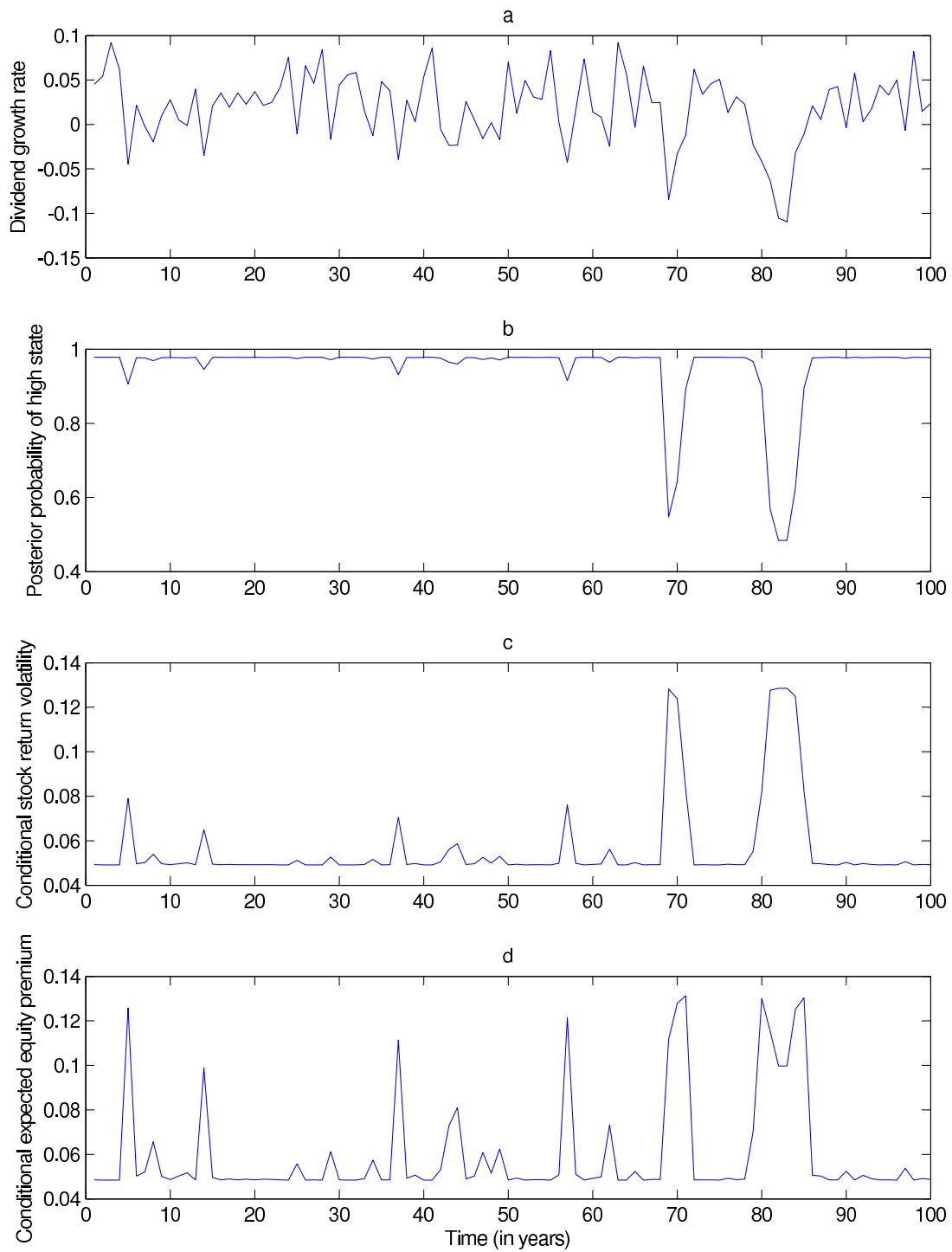


Figure 5: **Simulated time series of dividend (consumption) growth, posterior probability of the high-growth state, conditional volatility of stock returns, and conditional expected equity premium.** Parameter values are set as the baseline values given in Table 4.