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 ¯nance. However, it faces serious di±culties 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. Shiller (1981) ¯nds that equity volatility is too high to be justi¯ed by changes in the fundamental. 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 consumption and dividends follow a hidden Markov regime-switching model. The agent learns about the hidden state based on past data. The posterior state beliefs capture °uctuating economic uncertainty and drive asset return dynamics. Second, and more importantly, we assume that the agent is ambiguous about the hidden state and his preferences are represented by a generalized recursive smooth ambiguity model that allows for a three-way separation among risk aversion, ambiguity aversion and intertemporal substitution. We propose novel tractable homothetic utility speci¯cations. These speci¯cations nest Epstein-Zin preferences (Epstein and Zin (1989)), smooth ambiguity preferences (Klibano® et al. (2005, 2008)), multiplier preferences (Hansen and Sargent (2001)), and risk-sensitive preferences (Tallarini (2000)) as special cases. Ambiguity aversion is manifested through a pessimistic distortion of the pricing kernel in the sense that the agent attaches more weight on low continuation values in recessions. It is this pessimistic behavior that allows our model to explain the asset pricing puzzles.

We motivate our adoption of the recursive 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 and dividend growth. Our adopted ambiguity model captures this ambiguity and attitude towards ambiguity. Our second motivation is related to the robustness theory developed by Hansen and Sargent (2001, 2008) and Hansen (2007). Speci¯cally, the agent in our model may fear model misspeci¯cation. He is concerned about model uncertainty, and thus, seeks robust decision-making. We may interpret our ambiguity model as a model of robustness in the presence of model uncertainty.

Our modelling 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 e®ects.[1]

Learning is naturally embedded in our recursive 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 ¯xed, and to calibrate it for quantitative analysis. Another advantage is that we can apply the usual di®erential 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

[1] There is a large literature on learning in asset pricing using the standard Bayesian framework. Notable works

expected utility model lacks this smoothness property.

Our paper is related to a growing body of literature that studies the implications of ambiguity and robustness for ¯nance and macroeconomics.[2] We contribute to this literature by (i) proposing a novel generalized recursive ambiguity model and tractable homothetic speci¯cations, and (ii) putting this utility model in quantitative work to address a variety of asset pricing puzzles.

We now discuss closely related papers. In the max-min framework, Epstein and Schneider (2007) 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 (2008) analyze asset pricing implications. Leippold et al. (2008) embed this model in a continuous-time environment. In contrast to our paper, there is no distinction between risk aversion and intertemporal substitution and no separation between ambiguity and ambiguity attitudes in the preceding three papers. Hansen and Sargent (2007a) formulate a learning model that allows for two forms of model misspeci¯cation: (i) misspeci¯cation in the underlying Markov law for the hidden states, and (ii) misspeci¯cation of the probabilities assigned to the hidden Markov states. Hansen and Sargent (2007b) 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 speci¯cation. But he does not consider the general homothetic form 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 speci¯c ways and show that their modelling 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 our generalized recursive smooth ambiguity 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. Appendix A contains proofs.

[2] See Cao et al. (2005), Chen and Epstein (2002), Epstein and Miao (2003), Epstein and Wang (1994), Garlappi et al. (2007), and Routledge and Zin (2001) for asset pricing applications of the multiple-priors utility model. See Anderson et al. (2003), Cagetti, et al. (2002), Hansen and Sargent (2001), Hansen et al. (1999), 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.

2. Generalized Recursive Ambiguity Preferences

In this section, we introduce the recursive ambiguity utility model adopted in our paper. In a static setting, this utility model delivers essentially the same functional form that has appeared in some other papers, e.g., Ergin and Gul (2009), Klibanoff et al. (2005), Nau (2006), and Seo (2008). These papers provide various different axiomatic foundations and interpretations. Our adopted dynamic model is axiomatized by Hayashi and Miao (2008) and closely related to Klibanoff et al. (2008). Here we focus on the utility representation and refer the reader to the preceding papers for axiomatic foundations.

We start with a static setting in which a decision maker's ambiguity preferences over consumption are represented by the following utility function:

$$v^{-1}\left(\int_{\Pi} v\left(u^{-1}\left(\int_{S} u(C)\,d\pi\right)\right)d\mu(\pi)\right),\ \forall C: S \to \mathbb{R}_{+}, \tag{1}$$

where u and v are increasing functions and μ is a subjective prior over the set Π of probability measures on S that the decision maker thinks possible. We have defined utility in (1) in terms of two certainty equivalents. When we define $\phi = v \circ u^{-1}$, it is ordinally equivalent to the smooth ambiguity model of Klibanoff et al. (2005):

$$\int_{\Pi} \phi\left(\int_{S} u(C)\,d\pi\right)d\mu(\pi) \equiv \mathbb{E}_{\mu}\phi\left(\mathbb{E}_{\pi}u(C)\right). \tag{2}$$

A key feature of this model is that it achieves a separation between ambiguity, identified as a characteristic of the decision maker's subjective beliefs, and ambiguity attitude, identified as a characteristic of the decision maker's tastes.[3] Specifically, ambiguity is characterized by properties of the subjective set of measures Π. Attitudes towards ambiguity are characterized by the shape of ϕ or v, while attitudes towards pure risk are characterized by the shape of u. In particular, the decision maker displays risk aversion if and only if u is concave, while he displays ambiguity aversion if and only if ϕ is concave or, equivalently, if and only if v is a concave transformation of u. Intuitively, an ambiguity averse decision maker 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 the *ex ante* utility evaluation of C, $\mathbb{E}_{\pi}u(C)$ for all $\pi \in \Pi$. Note that there is no reduction between μ and π in general. It is possible when

[3] The behavioral foundation of ambiguity and ambiguity attitude is based on the theory developed by Ghirardato and Marinacci (2002) and Klibanoff et al. (2005). 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.

ϕ is linear. In this case, the decision maker is ambiguity neutral and the smooth ambiguity model reduces to the standard expected utility model.

Klibanoff et al. (2005) show that the multiple-priors model of Gilboa and Schmeidler (1989), $\min_{\pi\in\Pi}\mathbb{E}_{\pi}u(C)$, is a limiting case of the smooth ambiguity model with infinite ambiguity aversion. An important advantage of the smooth ambiguity model over other models of ambiguity, such as the multiple-priors utility model, 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 (2) is smooth under regularity conditions, rather than kinked as in the case of the multiple-priors utility model. More importantly, comparative statics of ambiguity attitudes can be easily analyzed using the function ϕ or v only, holding ambiguity fixed. Such a comparative static analysis is not evident for the multiple-priors utility model since the set of priors Π in that model may characterize ambiguity as well as ambiguity attitudes.

We may alternatively interpret the utility model defined in (1) as a model of robustness in which the decision maker is concerned about model misspecification, and thus seeks robust decision making. Specifically, each distribution $\pi \in \Pi$ describes an economic model. The decision maker 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 function v.

We now embed the static model (1) in a dynamic setting. Time is denoted by $t = 0, 1, 2, ...$ The state space in each period is denoted by S. At time t, the decision maker's information consists of history $s^t = \{s_0, s_1, s_2, ..., s_t\}$ with $s_0 \in S$ given and $s_t \in S$. The decision maker ranks adapted consumption plans $C = (C_t)_{t\geq 0}$. That is, C_t is a measurable function of s^t. The decision maker is ambiguous about the probability distribution on the full state space S^∞. This uncertainty is described by an unobservable 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)).

The decision maker has a prior μ_0 over the parameter z. Each parameter z gives a probability distribution π_z over the full state space. The posterior μ_t and the conditional likelihood $\pi_{z,t}$ can be obtained by Bayes' Rule. Inspired by Kreps and Porteus (1978) and Epstein and Zin (1989), we consider the following generalized recursive ambiguity utility function:

$$V_t(C) = W\left(C_t, \mathcal{R}_t\left(V_{t+1}(C)\right)\right), \quad \mathcal{R}_t(x) = v^{-1}\left(\mathbb{E}_{\mu_t}\left[v \circ u^{-1}\mathbb{E}_{\pi_{z,t}}\left[u(x)\right]\right]\right), \tag{3}$$

where W is a time aggregator, $\mathcal{R}$ is an uncertainty aggregator, and u and v admit the same interpretation as in the static setting. When $v \circ u^{-1}$ is linear, (3) reduces to the recursive utility model of Epstein and Zin (1989). In particular, the posterior μ_t and the likelihood $\pi_{z,t}$ can be reduced to a *predictive distribution*, which is the key idea underlying the Bayesian analysis. When $v \circ u^{-1}$ is nonlinear, the posterior μ_t and the likelihood $\pi_{z,t}$ cannot be reduced to a single distribution. It is this *irreducibility of compound distributions* that captures ambiguity, as pointed out by Hansen (2007), Klibanoff et al. (2005, 2008), and Segal (1987).

Our generalized recursive ambiguity utility model in (3) permits a three-way separation among risk aversion, ambiguity aversion and intertemporal substitution. In application, it proves tractable to consider the following homothetic specification:

$$W(c,y) = \left[(1-\beta)c^{1-\rho} + \beta y^{1-\rho}\right]^{\frac{1}{1-\rho}}, \ \rho > 0, \tag{4}$$

and u and v are given by:

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}, \ \gamma > 0, \neq 1, \tag{5}$$

$$v(x) = \frac{x^{1-\eta}}{1-\eta}, \ \eta > 0, \neq 1, \tag{6}$$

where $\beta \in (0,1)$ is the subjective discount factor, $1/\rho$ represents the elasticity of intertemporal substitution (EIS), γ is the risk aversion parameter, and η is the ambiguity aversion parameter. We then have

$$V_t(C) = \left[(1-\beta)C_t^{1-\rho} + \beta\left\{\mathcal{R}_t(V_{t+1}(C))\right\}^{1-\rho}\right]^{\frac{1}{1-\rho}}, \tag{7}$$

$$\mathcal{R}_t(V_{t+1}(C)) = \left\{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{\frac{1-\eta}{1-\gamma}}\right\}^{\frac{1}{1-\eta}}. \tag{8}$$

If $\eta = \gamma$, the decision maker is ambiguity neutral and (7) reduces to the recursive utility model of Epstein and Zin (1989) and Weil (1989). The decision maker displays ambiguity aversion if and only if $\eta > \gamma$. By the property of certainty equivalent, a more ambiguity averse agent with a higher value of η has a lower utility level.

In the limiting case with $\rho = 1$, (7) reduces to:

$$U_t = (1-\beta)\ln C_t + \frac{\beta}{1-\eta}\ln\left\{\mathbb{E}_{\mu_t}\exp\left(\frac{1-\eta}{1-\gamma}\ln\left[\mathbb{E}_{\pi_{z,t}}\exp\left((1-\gamma)U_{t+1}\right)\right]\right)\right\}, \tag{9}$$

where $U_t = \ln V_t$. This specification is closely related to the robust control model studied by Hansen (2007) and Hansen and Sargent (2001, 2007a, 2008). In particular, there are two risk-sensitivity adjustments in (9). The first risk-sensitivity adjustment for the distribution $\pi_{z,t}$ reflects the agent's concerns about the misspecification in the underlying Markov law given a hidden state z. The second risk-sensitivity adjustment for the distribution μ_t reflects the agent's concerns about the misspecification of the probabilities asiigned to the hidden states.

If we further take limit in (9) when $\gamma \to 1$, equation (9) becomes:

$$U_t = (1-\beta)\ln C_t + \frac{\beta}{1-\eta}\ln\left\{\mathbb{E}_{\mu_t}\exp\left((1-\eta)\,\mathbb{E}_{\pi_{z,t}}[U_{t+1}]\right)\right\}. \tag{10}$$

This is the log-exponential speci¯cation studied by Ju and Miao (2007). In this case, there is only one risk-sensitive adjustment for the state beliefs μ_t. Following Klibano® et al. (2005), we can show that when $\eta \to \infty$, (10) becomes:

$$U_t = (1-\beta)\ln C_t + \beta\min_z \mathbb{E}_{\pi_{z,t}}[U_{t+1}]. \tag{11}$$

This utility function belongs to the class of the recursive multiple-priors model of Epstein and Wang (1994) and Epstein and Schneider (2003, 2007). The agent is extremely ambiguity averse by choosing the worst continuation utility value each period.

Klibano® et al. (2008) propose the following closely related recursive smooth ambiguity model:

$$V_t(C) = u(C_t) + \beta\phi^{-1}\left(\mathbb{E}_{\mu_t}\phi\left(\mathbb{E}_{\pi_{z,t}}[V_{t+1}(C)]\right)\right), \tag{12}$$

where $\beta \in (0,1)$ is the discount factor, and u and ϕ admit the same interpretation as in the static model (2). In this model, risk aversion and intertemporal substitution is confounded. In addition, Ju and Miao (2007) ¯nd that when u is de¯ned in (5) and $\phi(x) = x^{1-\alpha}/(1-\alpha)$ for $x > 0$ and $1 \neq \alpha > 0$, the model (12) is not well de¯ned for $\gamma > 1$. Thus, they consider (7) with $\gamma = \rho$ and $\alpha \equiv 1-(1-\eta)/(1-\gamma)$, which is ordinally equivalent to (12) when $\gamma \in (0,1)$. The utility function in (12) is always well de¯ned for the speci¯cation $\phi(x) = -e^{-\frac{x}{\theta}}$ for $\theta > 0$. The nice feature of this speci¯cation is that it has a connection with risk-sensitive control and robustness, as studied by Hansen (2007) and Hansen and Sargent (2008). The disadvantage of this speci¯cation is that the utility function generally does not have the homogeneity property. Thus, the curse of dimensionality makes the numerical analysis of the decision maker's dynamic programming problem complicated, except for the special case where $u(c) = \ln(c)$ as in (10) (see Ju and Miao (2007) and Collard et al. (2009)). As a result, we will focus on the homothetic speci¯cation (7) in our analysis below.

3. Asset Pricing Implications

3.1. The Economy

We consider a representative-agent pure-exchange economy. There is only one consumption good with aggregate consumption given by C_t in period t. The agent trades multiple assets. Among these assets, we focus on the riskfree bond with zero net supply and equity that pays aggregate dividends D_t in period $t = 0, 1, 2, \ldots$. Let $R_{f,t+1}$ and $R_{e,t+1}$ denote their gross returns from period t to period $t+1$, respectively. We specify aggregate consumption by a regime-switching process as in Cecchetti et al. (1990, 1993, 2000) and Kandel and Stambaugh (1991):

$$\ln\left(\frac{C_{t+1}}{C_t}\right) = \kappa_{z_{t+1}} + \sigma\varepsilon_{t+1},\ \sigma > 0, \tag{13}$$

where ε_t is an independently and identically distributed (iid) standard normal random variable, and z_{t+1} follows a Markov chain which takes values 1 or 2 with transition matrix (λ_{ij}) where $\sum_j \lambda_{ij} = 1,\ i,j = 1,2$. We may identify state 1 as the boom state and state 2 as the recession state in that $\kappa_1 > \kappa_2$.

In a standard Lucas-style model (Lucas (1978)), dividends and consumption are identical in equilibrium. This assumption is clearly violated in reality. There are several ways to model dividends and consumption separately in the literature (Cecchetti, Lam, and Mark (1993)). Here, we follow Bansal and Yaron (2004) and assume:

$$\ln\left(\frac{D_{t+1}}{D_t}\right) = \phi\ln\left(\frac{C_{t+1}}{C_t}\right) + g_d + \sigma_d e_{t+1}, \tag{14}$$

where e_{t+1} is an iid standard normal random variable, and is independent of all other random variables. The parameter $\phi > 0$ can be interpreted as the leverage ratio on expected consumption growth as in Abel (1999). This parameter and the parameter σ_d allows us to calibrate volatility of dividends (which is significantly larger than consumption volatility) and their correlation with consumption. The parameter g_d helps match the expected growth rate of dividends. Our modelling of the dividend process is convenient because it does not introduce any new state variable in our model.

The model of consumption and dividends in (13) and (14) is a nonlinear counterpart of the long-run risk processes discussed in Campbell (1999) and Bansal and Yaron (2004) in that both consumption and dividends contain a common persistent component of Markov chain. Garcia et al. (2008) show that the processes in (13) and (14) can be obtained by discretizing the long-run risk model Case I in Bansal and Yaron (2004). Unlike Case II in Bansal and Yaron (2004), we assume that volatility σ is constant and independent of regimes. In the Bansal-Yaron model, fluctuating volatility of consumption growth is needed to generate time-varying expected equity premium. Our assumption of constant σ intends to generate this feature through endogenous learning rather than exogenous fluctuations in consumption volatility.

Unlike the long-run risks model, the regime-switching model can be easily estimated by the maximum likelihood method. Following Hansen (2007), we put economic agents and econometricians on equal footing by assuming that economic regimes are not observable. What is observable in period t is the history of consumption and dividends $s^t = \{C_0, D_0, C_1, D_1, ..., C_t, D_t\}$. The agent has ambiguous beliefs about the hidden states. His preferences are represented by the generalized ambiguity utility defined in (7). To apply this utility function, we need to derive the evolution of the posterior state beliefs. Let $\mu_t = \Pr\left(z_{t+1} = 1|s^t\right)$.[4] The prior belief μ_0 is given. By Bayes' Rule, we can derive:

$$\mu_{t+1} = \frac{\lambda_{11} f\left(\ln\left(C_{t+1}/C_t\right), 1\right)\mu_t + \lambda_{21} f\left(\ln\left(C_{t+1}/C_t\right), 2\right)\left(1-\mu_t\right)}{f\left(\ln\left(C_{t+1}/C_t\right), 1\right)\mu_t + f\left(\ln\left(C_{t+1}/C_t\right), 2\right)\left(1-\mu_t\right)}, \tag{15}$$

where $f(y, i) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\left(y-\kappa_i\right)^2 / \left(2\sigma^2\right)\right]$ is the density function of the normal distribution with mean κ_i and variance σ^2. By our modelling of dividends in (14), dividends do not provide any new information for belief updating and for the estimation of the hidden states.

3.2. Asset Pricing

As is standard in the literature, we derive the pricing kernel or the stochastic discount factor to understand asset prices. Following Duffie and Skiadas (1994) or Hansen et al. (2008), we use the homogeneity property of the generalized recursive ambiguity utility (7) to show that its pricing kernel is given by:

$$M_{z_{t+1},t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\rho} \left(\frac{V_{t+1}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{\rho-\gamma} \left(\frac{\left(\mathbb{E}_{z_{t+1},t}\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{-(\eta-\gamma)}, \quad z_{t+1} = 1, 2, \tag{16}$$

where $\mathbb{E}_{z_{t+1},t}$ denotes the expectation operator for the distribution of the consumption process conditioned on the history s^t and the period-$t+1$ state z_{t+1}. Given this pricing kernel, the return $R_{k,t+1}$ on any traded asset k satisfies the Euler equation:

$$\mathbb{E}_t\left[M_{z_{t+1},t+1} R_{k,t+1}\right] = 1, \tag{17}$$

where $\mathbb{E}_t$ is the expectation operator for the predictive distribution conditioned on history s^t. We distinguish between the unobservable price of aggregate consumption claims and the observable price of aggregate dividend claims. The return on the consumption claims is also the return on the wealth portfolio, which is unobservable, but can be solved using equation (17).

A challenge in estimating our model empirically is that the continuation value V_{t+1} in (16) is not observable. One possible approach is to use the following relation between continuation value and wealth proved in the appendix:

$$\frac{W_t}{C_t} = \frac{1}{1-\beta}\left(\frac{V_t}{C_t}\right)^{1-\rho}, \tag{18}$$

[4] We abuse notation here since we have used μ_t to denote the posterior distribution over the parameter space in Section 2.

where W_t is the wealth level at time t. We can then represent the pricing kernel (16) in terms of consumption growth and the return on the wealth portfolio, as in Epstein and Zin (1989, 1991). However, the return on the wealth portfolio is also unobservable, which makes empirical estimation of our model difficult.

We now turn to the interpretation of our pricing kernel in (16). The last multiplicative factor in (16) reflects the effect of ambiguity aversion. In the case of ambiguity neutrality (i.e., $\eta = \gamma$), this term vanishes and the pricing kernel reduces to that for the recursive utility model of Epstein and Zin (1989) and Weil (1989). When the agent is ambiguity averse with $\eta > \gamma$, a recession is associated with a high value of the pricing kernel. Intuitively, the agent has a lower continuation value V_{t+1} in a recession state, causing the adjustment factor in (16) to take a higher value in a recession than in a boom.

To explain asset pricing puzzles, a number of studies propose to adjust the standard pricing kernel. As Campbell and Cochrane (1999) argue, they have to answer the basic question: Why do people fear stocks so much? In the Campbell and Cochrane habit formation model, people fear stocks because stocks do poorly in recessions, times when consumption falls low relative to habits. Our model's answer is that people fear stocks because they are pessimistic and have low continuation values in recessions. This pessimistic behavior will reduce the stock price and raise the stock return. In addition, it will reduce the riskfree rate because the agent wants to save more for the future. More formally, using equation (17), we can derive:

$$\mathbb{E}_t\left[R_{e,t+1} - R_{f,t+1}\right] = \frac{-Cov_t\left(M_{z_{t+1},t+1}, R_{e,t+1}\right)}{\mathbb{E}_t\left[M_{z_{t+1},t+1}\right]}. \tag{19}$$

Because stocks do poorly in recessions when ambiguous averse people put more weight on the pricing kernel, ambiguity aversion helps generate high negative correlation between the pricing kernel and stock returns. This high negative correlation increases equity premium as shown in equation (19).[5]

To better understand an agent's pessimistic behavior, we consider the special case of the unitary EIS ($\rho = 1$). In this case, the recursive ambiguity utility function reduces to the Hansen and Sargent (2008) robust control model (9) and the pricing kernel becomes:

$$M_{z_{t+1},t+1} = \beta \frac{C_t}{C_{t+1}} \frac{V_{t+1}^{1-\gamma}\left(\mathbb{E}_{z_{t+1},t}\left[V_{t+1}^{1-\gamma}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}}{\left[\mathcal{R}_t\left(V_{t+1}\right)\right]^{1-\eta}}. \tag{20}$$

The expression $\beta C_t/C_{t+1}$ is the pricing kernel for the standard log utility. It is straightforward to show that the adjustment factor in (20) is the density with respect to the predictive distribution

[5]Using a static smooth ambiguity model, Gollier (2006) analyzes the effect of ambiguity aversion on the pricing kernel. He shows that ambiguity aversion may not generally reinforce risk aversion.

because we can use the law of iterated expectations to show that:

$$\mathbb{E}_t \left\{ \frac{V_{t+1}^{1-\gamma} \left(\mathbb{E}_{z_{t+1},t} \left[V_{t+1}^{1-\gamma} \right] \right)^{-\frac{\eta-\gamma}{1-\gamma}}}{\left[\mathcal{R}_t \left(V_{t+1} \right) \right]^{1-\eta}} \right\} = 1.$$

As a result, we can write the Euler equation (17) as $\hat{\mathbb{E}}_t \left[\beta C_t / C_{t+1} R_{k,t+1} \right] = 1,$ where $\hat{\mathbb{E}}_t$ is the conditional expectation operator for the slanted predictive distribution. In this case, the model is observational equivalent to an expected utility model with distorted beliefs. The distorted beliefs attach more weight to the recession state. A similar observation equivalence result also appears in the multiple-priors model. (see Epstein and Miao (2003) for a discussion.) An undesirable feature of the unitary EIS case is that the consumption-wealth ratio is constant in that $C_t = (1-\beta) W_t$ by (18), which is inconsistent with empirical evidence.

Ju and Miao (2007) consider further the special case (10) with $\rho = \gamma = 1$. In this log-exponential case, the pricing kernel becomes:

$$M_{z_{t+1},t+1} = \beta \frac{C_t}{C_{t+1}} \frac{\exp \left((1-\eta) \mathbb{E}_{z_{t+1},t} \left[\ln V_{t+1} \right] \right)}{\mu_t \exp \left((1-\eta) \mathbb{E}_{1,t} \left[\ln V_{t+1} \right] \right) + (1-\mu_t) \exp \left((1-\eta) \mathbb{E}_{2,t} \left[\ln V_{t+1} \right] \right)}.$$

The agent slants his state beliefs towards the state with the lower continuation value or the recession state. Ju and Miao (2007) also show that the return on equity satisfies $R_{e,t+1} = \frac{1}{\beta} C_{t+1}/C_t$ if dividends are equal to aggregate consumption, $C_t = D_t$. Consequently, this case cannot generate interesting stock returns dynamics.

We now turn to the general homothetic specification with $\rho \neq 1$.[6] In this case, the effect of ambiguity aversion is not distorting beliefs because the multiplicative adjustment factor in (16) is not a probability density. Thus, unlike in the case of $\rho = 1$, our model with $\rho \neq 1$ is not observational equivalent to an expected utility model because one cannot find a change in beliefs of an expected utility maximizer that can account for the ambiguity aversion behavior in our model. However, our interpretation of the ambiguity aversion behavior as attaching more weight (the preceding adjustment factor) to the recession state with worse continuation utility is still valid, but the weight may not be mixture linear in state beliefs.

Let $P_{e,t}$ denote the date t price of dividend claims. Using equations (16) and (17) and the homogeneity property of V_t, we can show that the price-dividend ratio $P_{e,t}/D_t$ is a function of

[6] Ju and Miao (2007) study the power-power case with $\rho = \gamma \neq 1,$ in which risk aversion and intertemporal substitution are confounded. They require $\gamma < 1$ to explain the asset pricing puzzles. Embedding the multiple-priors model of Epstein and Schneider (2007) in a continuous-time framework, Leippold et al. (2008) also assume $\rho = \gamma < 1$ as in Ju and Miao (2007). Unlike the present paper, they assume that (i) dividends are equal to consumption, (ii) dividend growth takes finitely many unknown values without regime shifts, and (iii) the agent receives an additional signal about dividends. In addition, they do not discuss their calibration procedure.

the state beliefs, denoted by $\varphi(\mu_t)$. By de¯nition, we can write the equity return as:

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

This equation implies that the state beliefs drive changes in the price-dividend ratio, and hence dynamics of equity returns. In the next section, we will show that ambiguity aversion and learning under ambiguity help amplify consumption growth uncertainty, while Bayesian learning has a modest quantitative e®ect.

4. Quantitative Results

We ¯rst describe stylized facts and calibrate our model. 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.[7] For comparison, we also solve two benchmark models. Benchmark model I is the fully rational model with Epstein-Zin preferences under complete information similar to that studied by Bansal and Yaron (2004). Benchmark model II incorporates learning and is otherwise the same as benchmark model I. This model is a special case of our ambiguity model when $\eta = \gamma$. A special case of benchmark model II with time-additive expected utility ($\eta = \gamma = \rho$) is similar to the continuous-time model of Veronesi (1999, 2000).

4.1. Stylized Facts and Calibration

We start by summarizing some asset pricing puzzles documented in the empirical literature. Using annual US data from 1891-1993, Cecchetti et al. (2000) ¯nd that the mean values of equity premium and riskfree rate are given by 5.75 and 2.66 percent, respectively, as reported in Panel A of Table 1.[8] 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 B of Table 1 reports that the log dividend yield predicts long-horizon realized excess returns. It also shows that the regression slopes and R^2's increase with the return horizon. This return predictability puzzle is ¯rst documented by Campbell and Shiller (1988b) and Fama and French (1988a). Panel B of Table 1 also reports variance

[7] The Fortran codes and a technical appendix detailing our numerical method are available upon request.

[8] Campbell (1999) and Campbell and Cochrane (1999) ¯nd similar estimates using log returns. We follow Cecchetti et al. (2000) and report arithmetic average returns in both data and model solutions. Mehra and Prescott (1985) also report arithmetic averages. Cecchetti et al. (2000) measure the riskfree rate using returns on six-month commercial paper.

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)).

[Insert Table 1 Here.]

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

To explain the above asset pricing phenomena, we calibrate our model at the annual frequency. We ¯rst calibrate parameters in consumption and dividends processes. Cecchetti et al. (2000) apply Hamilton's maximum likelihood method to estimate parameters in (13) using the annual per capita US consumption data covering the period 1890-1994. Table 2 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. The long-run average rate of consumption growth is 1.86 percent.

[Insert Table 2 Here.]

We next calibrate parameters in the dividend process (14). We follow Abel (1999) and set the leverage parameter $\phi = 2.74$. We then follow Bansal and Yaron (2004) and choose $g_d = -0.0323$ so that the average rate of dividend growth is equal to that of consumption growth. We choose σ_d to match the volatility of dividend growth in the data. Using di®erent century-long annual samples, this volatility is equal to 0.136 and 0.142, according to the estimates given by Cecchetti et al. (1990) and Campbell (1999), respectively. Here, we take 0.13 and ¯nd $\sigma_d = 0.084$. Our calibrated values of σ_d and ϕ imply that the correlation between consumption growth and dividend growth is about 0.76. This value may seem high. However, Campbell and Cochrane (1999) argue that the correlation is di±cult to measure and it may approach 1.0 in the very long run since dividends and consumption should share the same long-run trends.

Now, we select baseline preference parameters. We follow Bansal and Yaron (2004) and set EIS to 1.5, implying $\rho = 1/1.5$. An EIS greater than 1 is critical to generate procyclical variation of the price-consumption ratio. Researchers in macroeconomics and ¯nance generally believe that the risk aversion parameter is around 2. We thus set $\gamma = 2$, in order to demonstrate that the main force of our model comes from ambiguity aversion, but not risk aversion. We next select the discount factor β and ambiguity aversion parameter η to match the mean riskfree rate of 0.0266 and the mean equity premium of 0.0575 from the data reported in Table 1. We find $\beta = 0.975$ and $\eta = 8.864$.

There is no independent study of the magnitude of ambiguity aversion in the literature. To judge whether our calibrated value is reasonable, we conduct a thought experiment related to the Ellsberg Paradox (Ellsberg (1961)) in a static setting. Suppose there are two urns. Subjects are told that there are 50 black and 50 white balls in urn 1. Urn 2 also contains 100 balls, but may contain either 100 black balls or 100 white balls. If a subject picks a black ball from an urn, he wins a prize, otherwise he does not win or lose anything. Experimental evidence reveals that most subjects prefer to bet on urn 1 rather than urn 2 (Camerer (1999) and Halevy (2007)). Paradoxically, if the subject is asked to pick a white ball, he still prefers to bet on urn 1. The standard expected utility model with any beliefs or any risk aversion level cannot explain this paradox. Our adopted smooth ambiguity model in the static setting (1) can explain this paradox whenever subjects display ambiguity aversion (i.e., v is more concave than u). Thus, ambiguity aversion and risk aversion have distinct behavioral meanings.

Formally, Let w be a subject's wealth level and d be the prize money. Because the subject knows that the distribution of black and white balls in urn 1 is $(1/2, 1/2)$, when he evaluates a bet on urn 1, his utility level in terms of certainty equivalent is equal to:

$$u^{-1}\left(\frac{1}{2}u(w+d)+\frac{1}{2}u(w)\right). \tag{21}$$

The subject believes that there are two possible equally likely distributions $(0,1)$ and $(1,0)$ in urn 2, and thus $\Pi = \{(0,1),(1,0)\}$ and $\mu = (1/2, 1/2)$. But he is not sure which one is the true distribution and is averse to this uncertainty. When he evaluates a bet on urn 2, his utility level in terms of certainty equivalent is equal to:

$$v^{-1}\left(\int_{\Pi} v\left(u^{-1}\left(\int_{S} u(c)\,d\pi\right)\right)d\mu(\pi)\right), \tag{22}$$

where $c = w + d$ or w and $S = \{\text{black, white}\}$. The expression in (21) is larger than that in (22) if v is more concave than u, causing the subject to bet on urn 1 rather than urn 2. The difference between the certainty equivalents in (21) and (22) is a measure of ambiguity premium. Given power functions of u and v and fixing the risk aversion parameter, we can use the size of the ambiguity premium to gauge the magnitude of ambiguity aversion.[9] It is straightforward to compute that the ambiguity premium is equal to 1.7 percent of the expected prize value for our

[9] See Chen, Ju and Miao (2009) for a more extensive discussion and an application of our generalized recursive ambiguity model to a portfolio choice problem.

calibrated ambiguity aversion parameter η = 8.864, when we set γ = 2 and the prize-wealth ratio of 1 percent. Increasing the prize-wealth ratio raises the ambiguity premium. Camerer (1999) reports that the ambiguity premium is typically in the order of 10-20 percent of the expected value of a bet in the Ellsberg-Paradox type experiments. Given this evidence, our calibrated ambiguity aversion parameter seems small and reasonable.

4.2. Unconditional Moments of Returns

As a first check of the performance of our calibrated model, we compare the model prediction of the volatility of the equity premium and the volatility of the riskfree rate with the data. Panel A of Table 3 reports model results. This table reveals that our model can match the volatility of the equity premium in the data quite closely (0.1826 versus 0.1902). However, our model generated volatility of the riskfree rate is lower than the data (0.0116 versus 0.0513). Campbell (1999) argues that the high volatility of the riskfree rate in the century-long annual data could be due to large swings in inflation in the interwar period, particularly in 1919-21. Much of this volatility is probably due to unanticipated inflation and does not reflect the volatility in the ex ante real interest rate. Campbell (1999) reports that the annualized volatility of the real return on Treasury Bills is 1.8 percent using the US postwar quarterly data.

[Insert Table 3 Here.]

To understand why our model is successful in matching unconditional moments of returns, we conduct a comparative statics analysis in Panels B-E of Table 3. The first row of each of these panels gives the result of benchmark model II with Epstein-Zin preferences under Bayesian learning. We first consider the effects of the three standard parameters (β, ρ, γ) familiar from the Epstein-Zin model. Equation (17) implies that the riskfree rate $R_{f,t+1} = 1/\mathbb{E}_t\left[M_{z_{t+1},t+1}\right]$. Because the pricing kernel $M_{z_{t+1},t+1}$ increases with the subjective discount factor β, a high value of β helps match the low riskfree rate. Table 3 reveals that an increase in EIS (or $1/\rho$) from 1.5 to 2.0 generally lowers the riskfree rate and stock returns due to the high intertemporal substitution effect. In addition, an increase in γ from 2.0 to 5.0 also lowers the riskfree rate and raises stock returns. These results follow from the usual intuition in the Epstein-Zin model.

Next, consider the role of ambiguity aversion, which is unique in our model. Table 3 reveals that an increase in the ambiguity aversion parameter η lowers the riskfree rate and raises stock returns. The intuition follows from the discussion in Section 3.2. An ambiguity averse agent displays pessimistic behavior by attaching more weight to the worst state with low continuation utilities. Thus, he saves more for the future and invests less in the stock. In addition, as more weight is attached to the low-growth state, there is less variation of $\mathbb{E}_t\left[M_{z_{t+1},t+1}\right]$, and hence the riskfree rate $R_{f,t+1}$ is less volatile. By contrast, ambiguity aversion makes the pricing kernel $M_{z_{t+1},t+1}$ more volatile as revealed by the last term in (16), leading to high and volatile equity premium. It also generates a high market price of uncertainty defined by the ratio of the volatility of the pricing kernel and the mean of the pricing kernel (Hansen and Jagannathan (1991)). For our calibrated baseline parameter values, the market price of uncertainty is equal to 0.60, as reported in Panel A Table 3. It is equal to 0.09 in benchmark model II with $\eta = \gamma$.

Finally, we analyze the role of learning under ambiguity. We decompose the riskfree r_f in our model into three components:

$$r_f = r_f^* + \left(r_f^L - r_f^*\right) + \left(r_f - r_f^L\right), \tag{23}$$

where r_f^*, r_f^L, and r_f are the means of the riskfree rate delivered by benchmark model I, benchmark model II, and our ambiguity model, respectively. We do a similar decomposition for the mean stock returns and the volatility of the equity premium.[10] Table 4 presents this decomposition.

[Insert Table 4 Here.]

Panel A of this table shows that under the baseline parameter values, benchmark model I with full information predicts that the mean riskfree rate $r_f^* = 0.0363$, the mean equity returns $r_e^* = 0.046$, and the volatility of equity premium $\sigma_{eq}^* = 0.1448$. For benchmark model II with Epstein-Zin preferences, the standard Bayesian learning lowers the riskfree rate and raises the equity return and equity volatility, but by a negligible amount. By contrast, the component $(r_f - r_f^L)$ due to learning under ambiguity accounts for most of the decrease in the riskfree rate and the increase in the equity return and the volatility of the equity premium. In addition, the magnitude of this component is larger for a larger degree of ambiguity aversion. We find the same result also for various values of the risk aversion parameter as presented in Panels B-C. In particular, when the risk aversion parameter $\gamma = 2$ and 5, the corresponding effects of Bayesian learning are to lower the mean riskfree rate by 0.01 and 0.03 percent, to raise the mean stock return by 0.01 and 0.03 percent, and to raise the equity premium volatility by 0.02 and 0.05 percent. These effects are quantitatively negligible. Increasing EIS from 1.5 to 2.0 does not change this result much as revealed by Panels D-E.

A surprising feature of benchmark model II with Bayesian learning is that equity premium can become negative when risk aversion γ is sufficiently large in the special case of time-additive utility $\gamma = \rho$. Increasing risk aversion may worsen the equity premium puzzle. In a similar

[10] In a continuous-time multiple-priors model without learning, Chen and Epstein (2002) provide a similar decomposition and show that equity premium reflects a premium for risk and a premium for ambiguity.

continuous-time model, Veronesi (2000) proves this result analytically. The intuition is that an increase in risk aversion raises the agent's hedging demand for the stock after bad news in dividends, thereby counterbalancing the negative pressure on prices due to the bad news in dividends. The former e®ect may dominate so that the pricing kernel and stock returns are positively correlated, resulting in negative equity premia (see equation (19)). By contrast, in our model, an ambiguity averse agent invests less in the stock, thereby counterbalancing the preceding hedging e®ect. In contrast to risk aversion, an increase in the degree of ambiguity aversion helps increase equity premium.

4.3. Price-Consumption and Price-Dividend Ratios

Panel A of Figure 1 presents the price-consumption ratio as a function of the posterior probabilities μ_t of the high-growth state for three values of η, holding other parameters ¯xed at the baseline values. It reveals two properties. First, the price-consumption ratio is increasing and convex. The intuition is similar to that described by Veronesi (1999) who analyzes time-additive expected exponential utility. When times are good (μ_t is close to 1), a bad piece of news decreases μ_t, and hence decreases future expected consumption. But it also increases the agent's uncertainty about consumption growth since μ_t 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. Since the agent wants to be compensated for bearing more risk, they will require an additional discount on the price of consumption claims. Thus, the price reduction due to a bad piece of news in good times is higher than the reduction in expected future consumption. By contrast, suppose the agent believes times are bad and hence μ_t is close to zero. A good piece of news increases the expected future consumption, but also raises the agent's perceived uncertainty since it moves μ_t closer to 0.5. Thus, the price-consumption 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-consumption ratio because it induces the agent to invest less in the asset. In addition, an increase in the degree of ambiguity aversion raises the curvature of the price-consumption ratio function, thereby helping increase the asset price volatility. In the special case of benchmark model II with $\eta = \gamma$, the price-consumption ratio is close to be a linear function of the state beliefs.[11] Thus, this model cannot generate high asset price volatility.

[Insert Figure 1 Here]

[11] We can follow Veronesi (1999) to prove analytically that both the price-consumption and price-dividend ratios are linear in state beliefs for time-additive expected utility.

Panel B of Figure 1 presents the price-consumption ratio function for various values of ρ, holding other parameters ¯xed at the baseline values. It reveals that the price-consumption ratio is an increasing function of μ_t when $\rho < 1$, while it is a decreasing function when $\rho > 1$. When $\rho = 1$, it is equal to $\beta/(1-\beta)$ by (18) because wealth is equal to consumption plus the price of consumption claims. This result follows from the usual intuition in the Epstein-Zin model (see Bansal and Yaron (2004)). When $\rho < 1$, EIS is greater than 1 and hence the intertemporal substitution e®ect dominates the wealth e®ect. In response to good news of consumption growth, the agent buys more assets and hence the price-consumption ratio rises. The opposite result holds true when $\rho > 1$.

Panels C and D of Figure 1 present similar ¯gures for the price-dividend ratio. We ¯nd that the e®ects of η and ρ are similar. One di®erence is that the price-dividend ratio is not constant when $\rho = 1$ because dividends and aggregate consumption are not identical in our model. Due to leverage, we need a su±ciently small EIS (or a large ρ) to make the price-dividend ratio decrease with μ_t. Bansal and Yaron (2004) ¯nd a similar result in a full information model with Epstein-Zin preferences.

In summary, ambiguity aversion helps generate the variation in the price-consumption and price-dividend ratios. An EIS greater than 1 is important for generating procyclical price-consumption and price-dividend ratios.

4.4. Time-Varying Equity Premia and Equity Volatility

Panels A of Figure 2 plots the conditional expected equity premium as a function of the posterior probability μ_t of the high-growth state for various values of η. We ¯nd that this function is hump-shaped and peaks when μ_t is around 0.6. This shape seems to suggest that a negative consumption shock can lead to either higher or lower equity premium, depending on whether μ_t is close to 0 or to 1. However, since the economy spends most of the time in the high-growth state, the steady-state distribution of the posterior is highly skewed. This implies that μ_t is close to 1 in most of the time, leading to the pattern that equity premium rises following negative consumption shocks. As a result, our model can generate the countercyclical variation in equity premium observed in the data.

What is the role of ambiguity aversion? Panel A of Figure 2 shows that the curvature of the conditional expected equity premium function increases with η, implying that ambiguity aversion helps amplify the variation in equity premium. In benchmark model II with Bayesian learning ($\eta = \gamma = 2.0$), the conditional expected equity premium is almost °at. Consequently, it cannot generate highly time-varying expected equity premia. By contrast, when η is increased from 2 to 8.864, conditional equity premium can rise from about 3 percent to 28 percent.

[Insert Figure 2 Here.]

Panel B of Figure 2 plots the conditional volatility of equity premium as a function of μ_t for various values of η. This function is also hump-shaped, with the maximum attained at a value of μ_t close to 0.6. Following similar intuition discussed above, our model generates countercyclical variation in conditional volatility of equity premium observed in the data. In addition, ambiguity aversion helps amplify this variation.

Our model can also generate persistent changes in conditional volatility of equity premium, documented by Bollerslev et al. (1992). The intuition is 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. The persistence of beliefs drives the persistence of the volatility of equity premium.

[Insert Figure 3 Here.]

Figure 3 illustrates the time-varying properties of the expected equity premium and the volatility of equity premium by a Monte Carlo simulation. Panel A plots a time series of consumption growth simulated using (13). Panel B plots the time series of the posterior probability of the high-growth state μ_t, computed using (15). It reveals that in most of the time the agent believes that the economy is in the high-growth state in that μ_t is close to 1. After a few negative shocks to consumption growth, the agent believes the low-growth state is more likely in that μ_t 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, dividend growth, and the posterior probabilities, 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 3, respectively. From these panels, we can see that both the conditional volatility of equity premium and conditional expected equity premium are time-varying and move with business cycles countercyclically.

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 5 reports the model implied values of the variance ratios, the regression slopes and the R^2's, at horizons of 1, 2, 3, 5, and 8 years based on the baseline parameter values given in Table 3. 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 5, we observe that all three models can generate the pattern that variance ratios are less than 1 and decrease with the horizon, suggesting that excess returns are negatively serially correlated. In terms of predictive regressions, benchmark models I and II deliver very small R^2's, implying weak predictability.[12] One may expect that learning should help generate return predictability. The intuition is that the change of state beliefs is persistent, and hence the price-dividend ratio is also persistent and positively serially correlated. However, Table 5 reports that benchmark model II with Bayesian learning helps little quantitatively. In a related model, Brandt et al. (2004) ¯nd a similar result.

We ¯nally consider our model in which we introduce ambiguity aversion into benchmark model II. Table 5 reveals that while all three models can generate the pattern that the regression slopes increase with the horizon, our model with learning under ambiguity produces much more signi¯cant quantitative e®ects. In particular, compared to benchmark models I and II, our model implied values of the regression slopes and R^2's are much higher. However, our model still cannot replicate the same numbers estimated from the data reported in Panel B of Table 2. In addition, all three models cannot generate the pattern that R^2's increase with the horizon. The model predicted R^2's ¯rst increase with the horizon and then decrease with it after period 3. This could be due to the fact that the model generated price-dividend ratios are not persistent enough.[13] We should recognize that the predictability results in the empirical literature are quite sensitive to data sets, changing samples, and estimation techniques (Welch and Goyal (2008)). Thus, one should be cautious in interpreting empirical evidence on predictability.

[Insert Table 5 Here]

5. Conclusion

In this paper, we have proposed a novel generalized recursive smooth ambiguity model which allows a three-way separation among risk aversion, ambiguity aversion and intertemporal substitution. This model nests some utility models commonly adopted in the literature as special cases. We also propose a tractable homothetic speci¯cation and apply this model to asset pricing. When modelling consumption growth and dividend growth as regime-switching processes (nonlinear counterpart of the long-run risk processes as in Bansal and Yaron (2004)), our asset pricing model can help explain a variety of asset pricing puzzles. Our calibrated model can match the mean equity premium, the mean riskfree rate, and the volatility of equity premium

[12] Beeler and Campbell (2009) and Garcia et al. (2008) re-examine the Bansal-Yaron model and ¯nd that it cannot match the predictability in the data, contrary to the ¯nding of Bansal and Yaron (2004).

[13] See Campbell et al. (1997, pp. 271-273) for a theoretical analysis of why R^2's may ¯rst increase and then decrease with the horizon.

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 of excess returns.

We show that ambiguity aversion and learning under ambiguity play a key role in explaining asset pricing puzzles. An ambiguity averse agent displays pessimistic behavior in that he attaches more weight to the pricing kernel in bad times when his continuation values are low. This pessimistic behavior helps propagate and amplify shocks to consumption growth, and generates time-varying equity premium. We also ¯nd that Bayesian learning in the expected utility framework has a modest quantitative e®ect on asset returns, while learning under ambiguity is important to explain dynamic asset pricing phenomena. One limitation of our model is that it cannot reproduce the predictability pattern in the data.

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.[14] As a result, the agent's risk aversion is time varying. Bansal and Yaron (2004) apply the Epstein-Zin preferences, and incorporate °uctuating volatility and a persistent component in consumption growth.[15] Their calibrated risk aversion parameter is 10. Our model of consumption and dividend processes is similar to Bansal and Yaron (2004), but is much easier to estimate. We shut down exogenous °uctuations in consumption growth volatility and analyze how endogenous learning under ambiguity can generate time-varying equity premium.

We view our model as a ¯rst 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 empirically estimate parameters of ambiguity aversion, risk aversion, and intertemporal substitution would be important future research topics. In addition, our proposed novel generalized recursive ambiguity model can be applied to many other problems in ¯nance and macroeconomics.

[14] Ljungqvist and Uhlig (2009) show that government interventions that occasionally destroy part of endowment can be welfare improving when endogenizing aggregate consumption choices in the Campbell-Cochrane habit formation model.

[15] See Beeler and Campbell (2009) for a critique of the Bansal-Yaron model.

Appendix

A Proofs of Results in Section 3.2

We follow the method of Hansen et al. (2008) to derive the marginal utility of consumption and continuation value as:

$$MC_t = \frac{\partial V_t(C)}{\partial C_t} = (1-\beta) V_t^{\rho} C_t^{-\rho},$$

$$MV_{z_{t+1},t+1} = \frac{\partial V_t(C)}{\partial V_{z_{t+1},t+1}} = \beta V_t^{\rho} \left[\mathcal{R}_t(V_{t+1})\right]^{\eta-\rho} \left(\mathbb{E}_{z_{t+1},t}\left[V_{t+1}^{1-\gamma}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}} V_{t+1}^{-\gamma},$$

where $V_{z_{t+1},t+1}$ denotes the continuation value $V_{t+1}(C)$ conditioned on the period $t+1$ state being z_{t+1}. The pricing kernel is given by $M_{z_{t+1},t+1} = \left(MV_{z_{t+1},t+1}\right)(MC_{t+1})/MC_t$, which delivers (16).

We next use the dynamic programming method of Epstein and Zin (1989) to derive other results in Section 3.2. Suppose the agent trades N assets. The budget constraint is $W_{t+1} = (W_t - C_t) R_{w,t+1}$, where the return on the wealth portfolio $R_{w,t+1}$ is equal to $\sum_{k=1}^{N} \psi_{kt} R_{k,t+1}$, ψ_{kt} is the portfolio weight on asset k, and $R_{k,t+1}$ denotes its return. The value function $J(W_t,\mu_t)$ satisfies the Bellman equation:

$$\begin{aligned} J(W_t,\mu_t) \;=\; \max \Bigg[(1-\beta) C_t^{1-\rho} + \beta \Bigg\{ \Bigg(& \mu_t \left(\mathbb{E}_{1,t}\left[J^{1-\gamma}(W_{t+1},\mu_{t+1})\right]\right)^{\frac{1-\eta}{1-\gamma}} \\ & + (1-\mu_t)\left(\mathbb{E}_{2,t}\left[J^{1-\gamma}(W_{t+1},\mu_{t+1})\right]\right)^{\frac{1-\eta}{1-\gamma}} \Bigg)\Bigg\}^{\frac{1-\rho}{1-\eta}} \Bigg]^{\frac{1}{1-\rho}}. \end{aligned} \tag{A.1}$$

Conjecture

$$J(W_t,\mu_t) = A_t W_t, \text{ and } C_t = a_t W_t, \tag{A.2}$$

where A_t and a_t are to be determined. Substituting (A.2) and the budget constraint into (A.1), we can then rewrite the Bellman equation as:

$$A_t = \max_{a_t,\{\psi_{kt}\}} \left[(1-\beta) a_t^{1-\rho} + (1-a_t)^{1-\rho} \beta \left(\mathbb{E}_{\mu_t} \left(\mathbb{E}_{z_{t+1},t}\left[(A_{t+1}R_{w,t+1})^{1-\gamma}\right]\right)^{\frac{1-\eta}{1-\gamma}} \right)^{\frac{1-\rho}{1-\eta}} \right]^{\frac{1}{1-\rho}}$$

Use the first-order condition for consumption to derive:

$$\left(\frac{a_t}{1-a_t}\right)^{-\rho} = \frac{\beta}{1-\beta} \left(\mathbb{E}_{\mu_t} \left(\mathbb{E}_{z_{t+1},t}\left[(A_{t+1}R_{w,t+1})^{1-\gamma}\right]\right)^{\frac{1-\eta}{1-\gamma}} \right)^{\frac{1-\rho}{1-\eta}}. \tag{A.3}$$

From the above two equations, we have:

$$A_t = (1-\beta)^{1/(1-\rho)} a_t^{-\rho/(1-\rho)} = (1-\beta)^{1/(1-\rho)} \left(\frac{C_t}{W_t}\right)^{-\rho/(1-\rho)}. \tag{A.4}$$

Substituting equation (A.4) into (A.2) yields equation (18).

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 E®ects 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 Semi-groups for Model Speci¯cation, 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 Rogo® (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.

Beeler, Jason and John Y. Campbell, 2009, The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment, working paper, Harvard University.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner, 1992, ARCH Modelling 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.

Camerer, Colin F., 1999, Ambiguity-Aversion and Non-Additive Probability: Experimental Evidence, Models and Applications, *Uncertain Decisions: Bridging Theory and Experiments*, ed. by L. Luini, pp. 53-80. Kluwer Academic Publishers.

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., Andrew W. Lo, and A. Craig Mackinlay, 1997, *The Econometrics of Financial Markets,* Princeton University Press, Princeton.

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, Hui, Nengjiu Ju, and Jianjun Miao, 2009, Dynamic Asset Allocation with Ambiguous Return Predictability, working paper, Boston University.

Chen, Zengjin and Larry G. Epstein, 2002, Ambiguity, Risk and Asset Returns in Continuous Time, *Econometrica* 4, 1403-1445.

Collard, Fabrice, Sujoy Mukerji, Kevin Sheppard, and Jean-Marc Tallon, 2009, Ambiguity and Historical Equity Premium, working paper, Oxford University.

David, Alexander, 1997, Fluctuating Con¯dence 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.

Du±e, 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 De¯nition 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, 2007, Learning under ambiguity, *Review of Economic Studies* 74, 1275-1303.

Epstein, Larry G. and Martin Schneider, 2008, Ambiguity, Information Quality and Asset Pricing, *Journal of Finance* 63, 197-228.

Epstein, Larry G. and Tan Wang, 1994, Intertemporal Asset Pricing under Knightian Uncertainty, *Econometrica* 62, 283-322.

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.

Epstein, Larry G. and Stanley Zin, 1991, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis, *Journal of Political Economy* 99, 263-286.

Ergin, Huluk I. and Faruk Gul, 2009, A Theory of Subjective Compound Lotteries, *Journal of Economic Theory* 144, 899-929.

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.

Garcia, Rene, Nour Meddahi, and Romeo Tedongap, 2008, An Analytical Framework for Assessing Asset Pricing Models and Predictability, working paper, Edhec Business School.

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, Itzach 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.

Halevy, Yoran, 2007, Ellsberg Revisited: An Experimental Study, *Econometrica* 75, 503-536.

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, *American Economic Review* 97, 1-30.

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, John C. Heaton, and Nan Li, 2008, Consumption Strikes Back?: Measuring Long-Run Risk, *Journal of Political Economy* 116, 260-302.

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, 2007a, Recursive Robust Estimation and Control without Commitment, *Journal of Economic Theory* 136, 1-27.

Hansen, Lars Peter and Thomas J. Sargent, 2007b, Fragile Beliefs and the Price of Model Uncertainty, working paper, New York University.

Hansen, Lars Peter and Thomas J. Sargent, 2008, *Robustness*, Princeton University Press, Princeton and Oxford.

Hansen, Lars Peter, Thomas J. Sargent, and Thomas Tallarini, 1999, Robust Permanent Income and Pricing, *Review of Economic Studies* 66, 873-907.

Hayashi, Takashi and Jianjun Miao, 2009, Generalized Recursive Ambiguity Preferences, work in progress, Boston University.

Ju, Nengjiu and Jianjun Miao, 2007, Ambiguity, Learning, and Asset Returns, working paper, Boston University.

Judd, Kenneth, 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, MA.

Kandel, Shmuel and Robert F. Stambaugh, 1991, Asset Returns and Intertemporal Preferences, *Journal of Monetary Economics* 27, 39{71.

Keynes, Maynard, 1936, *The General Theory of Employment, Interest, and Money*, London Macmillan.

Klibano®, Peter, Massimo Marinacci, and Sujoy Mukerji, 2005, A Smooth Model of Decision Making under Ambiguity, *Econometrica* 73, 1849-1892.

Klibano®, Peter, Massimo Marinacci, and Sujoy Mukerji, 2008, Recursive Smooth Ambiguity Preferences, forthcoming in *Journal of Economic Theory*.

Knight, Frank, 1921, *Risk, Uncertainty, and Pro¯t*, Boston: Houghton Mi²in.

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, 2008, Learning and Asset Prices under Ambiguous Information, *Review of Financial Studies* 21, 2565-2597.

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.

Ljunqvist, Lars and Harold Uhlig, 2009, Optimal Endowment Destruction under Campbell-Cochrane Habit Formation, NBER working paper #14772.

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.

Nau, Robert F., 2006, Uncertainty Aversion with Second-Order Utilities and Probabilities, *Management Science* 52, 136-145.

Poterba, James M. and Lawrence H. Summers, 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.

Seo, Kyounwon, 2008, Ambiguity and Second-Order Belief, forthcoming in *Econometrica*.

Shiller, Robert J., 1981, Do Stock Prices Move too Much to be Justi¯ed 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.

Uppal, Raman and Tan Wang, 2003, Model Misspeci¯cation and Underdiversi¯cation, *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 A®ect Stock Returns? *Journal of Finance* 55, 807-839.

Weil, Philippe, 1989, The Equity Premium Puzzle and the Risk-Free Rate Puzzle, *Journal of Monetary Economics* 24, 401-421.

Weitzman, Martin L., 2007, Subjective Expectations and Asset-Return Puzzles, *American Economic Review* 97, 1102-1130.

Welch, Ivo and Amit Goyal, 2008, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* 21, 1455-1508.

Table 1. 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

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: The regression slope and R^2 are for regressions of the $k-$year ($k = 1, 2, 3, 5, 8$) ahead equity premium on the current log dividend-price ratio. The variance ratio is the variance of the $k-$year equity premium divided by k times the variance of the one-year equity premium. This table is taken from Table 1 in Cecchetti et al. (2000).

Table 2. 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 3. Unconditional Moments and Comparative Statistics

η	r_f	$\sigma(r_f)$	r_e	$\sigma(r_e)$	μ_{eq}	$\sigma(\mu_{eq})$	¾(M) / E[M]
Panel A (Baseline): γ = 2.0, ρ = 1/1.5							
8.864	2.66	1.16	8.41	17.98	5.75	18.26	0.60
Panel B: ρ = 1/1.5, γ = 2.0							
2.0	3.61	0.63	4.60	14.49	0.99	14.50	0.09
3.0	3.55	0.70	4.81	14.72	1.26	14.75	0.12
8.0	2.88	1.10	7.44	17.28	4.56	17.51	0.47
15.0	-0.18	1.06	16.44	18.75	16.62	19.01	2.20
Panel C: ρ = 1/1.5, γ = 5.0							
5.0	3.02	0.89	6.48	15.73	3.46	15.83	0.30
8.0	2.40	1.12	9.02	17.72	6.62	17.97	0.59
15.0	-0.87	0.97	17.88	18.48	18.75	18.68	2.43
Panel D: ρ = 1/2, γ = 2.0							
2.0	3.30	0.51	4.31	14.57	1.02	14.58	0.09
3.0	3.23	0.58	4.53	14.82	1.30	14.85	0.12
8.0	2.52	1.00	7.30	17.59	4.79	17.82	0.48
15.0	-0.82	0.96	16.42	19.10	17.23	19.34	2.22
Panel E: ρ = 1/2, γ = 5.0							
5.0	2.71	0.77	6.29	15.92	3.58	16.01	0.30
8.0	2.04	1.02	8.93	18.04	6.89	18.29	0.60
15.0	-1.52	0.88	17.88	18.80	19.39	19.00	2.45

Notes: Except for the numbers in the ¯rst and the last columns, all other numbers are in percentage. Columns 2-7 present the means and standard deviations of the riskfree rate, the equity return, and the equity premium, respectively. $\sigma(M)/E[M]$ is the ratio of the standard deviation to the mean of the pricing kernel. We set β = 0.975 in all cases.

Table 4. Decomposition of r_f, r_e and σ_{eq}

η	r_f^*	¢r_f^L	¢r_f	r_e^*	¢r_e^L	¢r_e	σ_{eq}^*	¢σ_{eq}^L	¢σ_{eq}
Panel A (Baseline): γ = 2.0, ρ = 1/1.5									
8.864	3.63	-0.01	-0.96	4.60	0.01	3.80	14.48	0.02	3.76
Panel B: ρ = 1/1.5, γ = 2.0									
2.0	3.62	-0.01	0.00	4.59	0.01	0.00	14.48	0.02	0.00
3.0	3.62	-0.01	-0.06	4.59	0.01	0.21	14.48	0.02	0.25
8.0	3.62	-0.01	-0.73	4.59	0.01	2.85	14.48	0.02	3.01
15.0	3.62	-0.01	-3.80	4.59	0.01	11.84	14.48	0.02	4.51
Panel C: ρ = 1/1.5, γ = 5.0									
5.0	3.05	-0.03	0.00	6.45	0.03	0.00	15.78	0.05	0.000
8.0	3.05	-0.03	-0.61	6.45	0.03	2.54	15.78	0.05	2.14
15.0	3.05	-0.03	-3.89	6.45	0.03	11.40	15.78	0.05	2.85
Panel D: ρ = 1/2, γ = 2.0									
2.0	3.30	-0.01	0.00	4.31	0.01	0.00	14.58	0.01	0.00
3.0	3.30	-0.01	-0.07	4.31	0.01	0.22	14.58	0.01	0.27
8.0	3.30	-0.01	-0.78	4.31	0.01	2.99	14.58	0.01	3.24
15.0	3.30	-0.01	-4.11	4.31	0.01	12.10	14.58	0.01	4.76
Panel E: ρ = 1/2, γ = 5.0									
5.0	2.73	-0.02	0.00	6.26	0.02	0.00	15.99	0.03	0.00
8.0	2.73	-0.02	-0.67	6.26	0.02	2.64	15.99	0.03	2.28
15.0	2.73	-0.02	-4.22	6.26	0.02	11.59	15.99	0.03	2.99

Notes: Except for the numbers in Column 1, all numbers are in percentage. The variables r_f^*, r_e^*, and σ_{eq}^* are the mean riskfree rate, the mean stock return, and the equity premium volatility, respectively, for benchmark model I. The variables r_f^L, r_e^L, and σ_{eq}^L are the mean riskfree rate, the mean stock return, and the equity premium volatility, respectively, for benchmark model II. We denote by ¢$r_f^L = r_f^L - r_f^*$ the change of the mean riskfree rate due to Bayesian learning, and by ¢$r_f = r_f - r_f^L$ the change of the mean riskfree rate due to learning under ambiguity. The other variables ¢r_e^L, ¢r_e, ¢σ_{eq}^L, ¢σ_{eq}, are de¯ned similarly. We set β = 0.975 in all cases.

Table 5. Predictability and Persistence of Excess Returns

	Baseline parameter values			Benchmark model I			Benchmark model II		
Horizon	Slope	R^2	Variance ratio	Slope	R^2	Variance ratio	Slope	R^2	Variance ratio
1	0.810	0.134	1.000	0.368	0.034	1.000	0.537	0.029	1.000
2	1.089	0.160	0.825	0.503	0.037	0.967	0.712	0.032	0.962
3	1.209	0.158	0.719	0.575	0.035	0.944	0.802	0.030	0.936
5	1.294	0.140	0.598	0.643	0.030	0.909	0.902	0.027	0.900
8	1.323	0.115	0.509	0.688	0.026	0.871	0.985	0.025	0.862

Notes: The slope and R^2 are obtained from an OLS regression of the excess returns on the log dividend yield at di®erent 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.

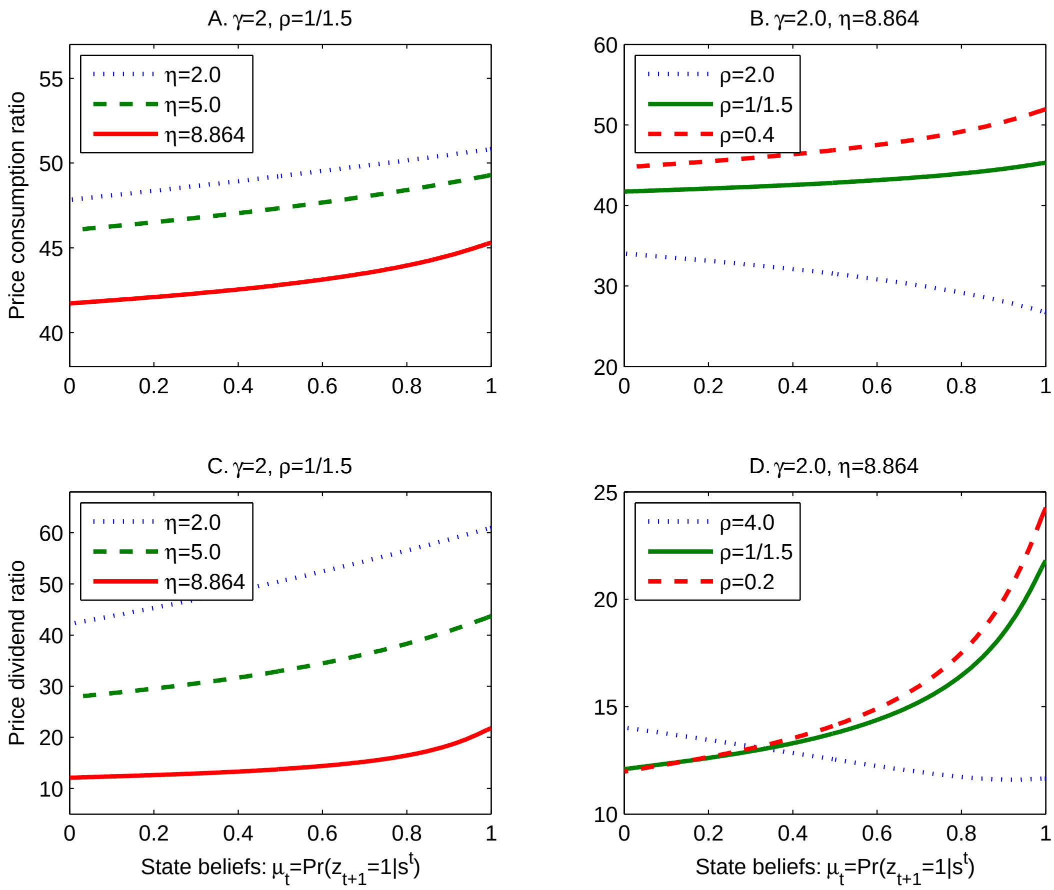

Figure 1: **The price-dividend ratio and the price-consumption ratio.** Panels A and B plot the price-consumption ratio as a function of the posterior probabilities of the high-growth state. Panels C and D plot the price-dividend ratio as a function of the posterior probabilities of the high-growth state.

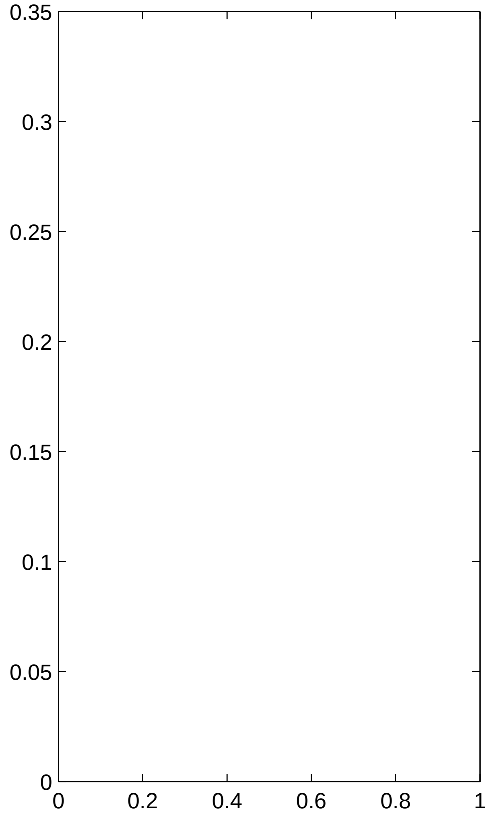

Figure 2: **Conditional expected equity premium and conditional volatility of equity premium.** Panels A and B plot the conditional expected equity premium and conditional volatility of equity premium as functions of the posterior probabilities of the high-growth state.

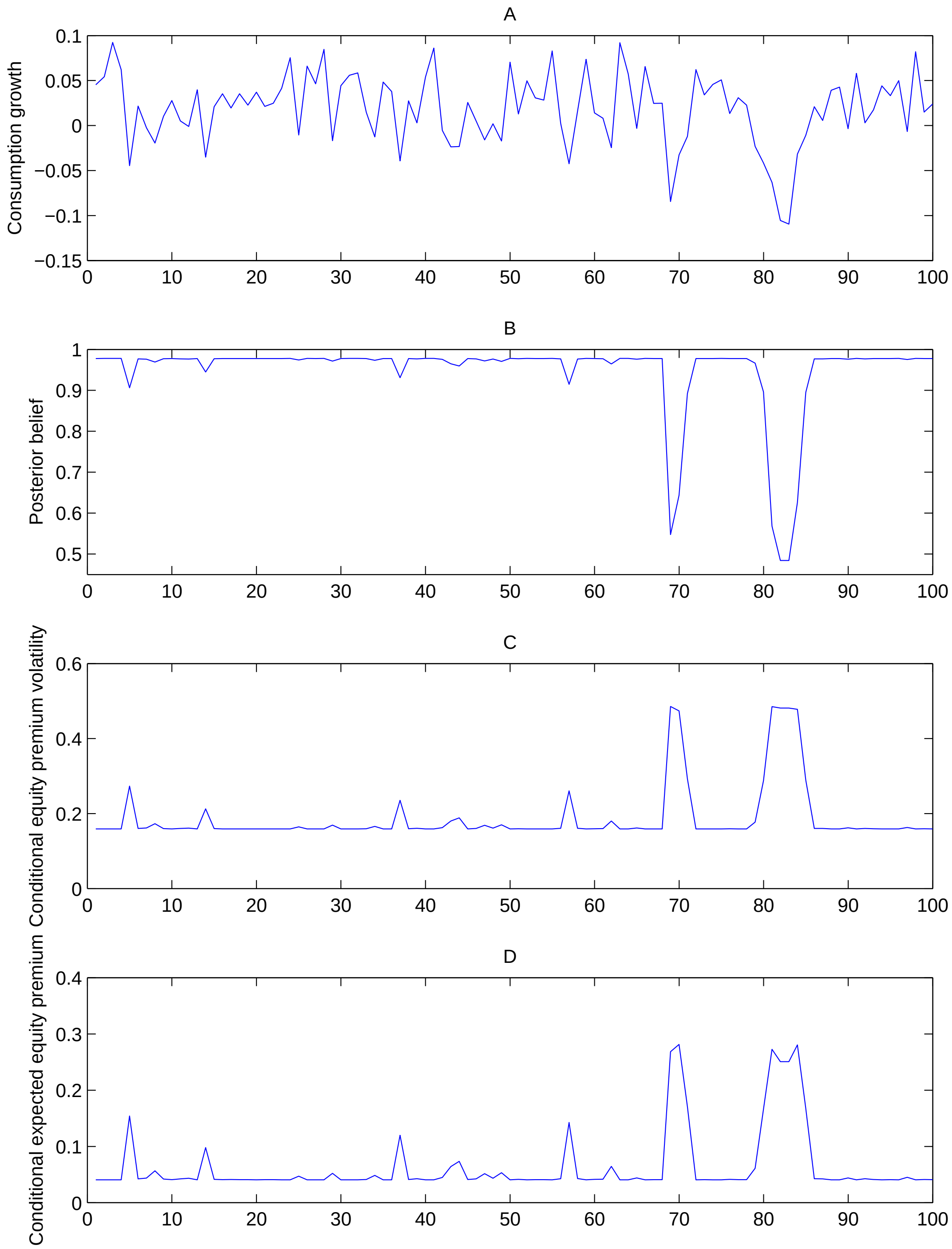

Figure 3: **Simulated time series of consumption growth, posterior probabilities of the high-growth state, conditional volatility of stock returns, and conditional expected equity premium.** Parameter values are set as the baseline values given in Table 3.