FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

Ambiguity Aversion and Variance Premium

Jianjun Miao, Bin Wei, and Hao Zhou

Working Paper 2018-14
December 2018

Abstract: This paper offers an ambiguity-based interpretation of variance premium–the difference between risk-neutral and objective expectations of market return variance–as a compounding effect of both belief distortion and variance differential regarding the uncertain economic regimes. Our approach endogenously generates variance premium without imposing exogenous stochastic volatility or jumps in consumption process. Such a framework can reasonably match the mean variance premium as well as the mean equity premium, equity volatility, and the mean risk-free rate in the data. We find that about 96 percent of the mean variance premium can be attributed to ambiguity aversion. Applying the model to historical consumption data, we find that variance premium mostly captures depressions, deep recessions, and financial panics, with a postwar peak in 2009.

JEL classification: G12, G13, D81, E44

Key words: ambiguity aversion, learning, variance premium, regime shift, belief distortion

https://doi.org/10.29338/wp2018-14

Please address questions regarding content to Jianjun Miao, Department of Economics, Boston University, 270 Bay State Road, Boston MA 02215, USA, 617-353-6675, miaoj@bu.edu; Bin Wei, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, GA 30309-4470, 404-498-8913, bin.wei@atl.frb.org; or Hao Zhou, PBC School of Finance, Tsinghua University, 43 Chengfu Road, Haidian District, Beijing, 100083, China, zhouh@pbcsf.tsinghua.edu.cn.

1. Introduction

Much attention has been paid to the equity premium puzzle: the high equity premium in the data requires an implausibly high degree of risk aversion in a standard rational representative-agent model to match the magnitude (Mehra and Prescott (1985)). More recently, researchers have realized that such a standard model typically predicts a negligible premium for higher moments such as variance premium (defined as the difference between the expected stock market variances under the risk neutral measure and under the objective measure), even with a high risk aversion coefficient. This result, however, is at odds with the sizable variance premium observed in the data, generating the so called variance premium puzzle.[1]

The goal of this paper is to provide an ambiguity-based explanation for the variance premium puzzle. The Ellsberg (1961) paradox and related experimental evidence point out the importance of distinguishing between risk and ambiguity — roughly speaking, risk refers to the situation where there is a known probability measure to guide choices, while ambiguity refers to the situation where no known probabilities are available. In this paper, we show that ambiguity aversion helps generate a sizable variance premium to closely match the magnitude in the data. In particular, it captures about 96 percent of the average variance premium whereas risk can only explain about 4 percent of it.

To capture ambiguity-sensitive behavior, we adopt the recursive smooth ambiguity model developed by Hayashi and Miao (2011) and Ju and Miao (2012) who generalize the model of Klibanoff, Marinacci and Mukerji (2009). The Hayashi-Ju-Miao model also includes the Epstein-Zin model as a special case in which the agent is ambiguity neutral. 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. This feature generates a

[1] Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011), Londono (2010), Bollerslev, et al. (2011) show that variance premium predicts U.S. and global stock market returns. Further evidence of its predictive power to forecast Treasury bond and credit spreads can be found in Zhou (2010), Mueller, Vedolin, and Zhou (2011), as well as Buraschi, Trojani, and Vedolin (2009) and Wang, Zhou, and Zhou (2011).

large countercyclical variation of the pricing kernel.[2] Ju and Miao (2012) show that the large countercyclical variation of the pricing kernel is important for the model to resolve the equity premium and risk-free rate puzzles and to explain the time variation of equity premium and equity volatility observed in the data. The present paper shows that it is also important for understanding the variance premium puzzle.

The Hayashi-Ju-Miao model allows for a three-way separation among risk aversion, intertemporal substitution, and ambiguity aversion. This separation is important not only for a conceptual reason, but also for quantitative applications. In particular, the separation between risk aversion and intertemporal substitution is important for matching the low risk-free rate observed in the data as is well known in the Epstein-Zin model. In addition, it is important for long-run risks to be priced (Bansal and Yaron (2004)). The separation between risk aversion and ambiguity aversion allows us to decompose equity premium into a risk premium component and an ambiguity premium component (Chen and Epstein (2002) and Ju and Miao (2012)). We can then fix the risk aversion parameter at a conventionally low value and use the ambiguity aversion parameter to match the mean equity premium in the data. This parameter plays an important role in amplifying and propagating the impact of uncertainty on asset returns and variance premium.

Following Ju and Miao (2012), we assume that consumption growth follows a regime-switching process (Hamilton (1989)) and that the agent is ambiguity averse to the variation of the hidden regimes. The agent learns about the hidden state based on past data. Our adopted recursive smooth ambiguity model incorporates learning naturally. 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 in the utility function, unlike in the standard Bayesian analysis. It is this irreducibility of compound lotteries that captures sensitiv-

[2] Also see Hansen and Sargent (2011) for a similar result based on robust control.

ity to ambiguity or model uncertainty (Segal (1990), Klibanoff, Marinacci and Mukerji (2005), Hansen (2007), and Seo (2009)). We show that there are important quantitative implications for learning under ambiguity, while standard Bayesian learning has small quantitative effects on both equity premium and variance premium. This finding is consistent with that in Hansen (2007) and Ju and Miao (2012).

We decompose the variance premium under our framework into three components: (i) the difference between the Bayesian belief about the boom state and the corresponding uncertainty adjusted belief, (ii) the market variance differentials between recessions and booms, and (iii) two terms related to conditional covariance between the market variance and the pricing kernel. The variance premium is equal to the product of the first two components plus the last component.

The first component is positive because the uncertainty adjusted belief gives a lower probability to the boom state than the Bayesian belief whenever the agent is uncertainty averse. We show that ambiguity aversion leads the agent to put less weight on the boom state and more weight on the recession state, thereby lowering the uncertainty-adjusted belief relative to the Epstein-Zin model. The Epstein-Zin model in turn delivers a lower uncertainty-adjusted belief about the boom state than the standard time-additive constant relative risk aversion (CRRA) utility model, if the agent prefers early resolution of uncertainty.

The second component is positive as long as the conditional market variance is countercyclical, i.e., the conditional market variance is higher in a recession than in a boom. It is intuitive that agents are more uncertain about future economic growth in bad times, generating higher stock return volatility. Formally, our model implies that the price-dividend ratio is a convex function of the Bayesian belief about the boom state, as in Veronesi (1999) and Ju and Miao (2012). As a result, agents' willingness to hedge against changes in their perceived uncertainty makes them overreact to bad news in good times and underreact to good news in bad times. Because ambiguity aversion enhances the countercyclicality of the pricing kernel, it makes the price function more convex than the Epstein-Zin model. Consequently, ambiguity aversion amplifies the countercyclicality of the stock return variance, thereby raising the second component of the variance premium relative to the Epstein-Zin model.

The third component is also positive, because both the market variance and the pricing kernel are countercyclical and hence positively correlated. Ambiguity aversion enhances both countercyclicality and hence raises the third component as well.

Our model with ambiguity aversion generates a mean variance premium of 8.51 (in percentage squared, monthly basis), which is quite close to our empirical estimate of 10.80 and well within the typical range of 5-19 from existing empirical studies.[3] In contrast, models under full information with time-additive CRRA utility and with Epstein-Zin utility can only produce a mean variance premium of 0.07 and 0.31, respectively. Incorporating Bayesian learning in these models changes the mean variance premium to 0.10 and 0.32, respectively. Thus, risk aversion and intertemporal substitution contributes about 1 and 3 percent to the model implied variance premium, while ambiguity aversion contributes to about 96 percent.

Note that these results are achieved under ambiguity aversion without introducing stochastic volatility or volatility jumps in consumption growth. In addition, our calibration targets to match the mean equity premium and the mean risk-free rate, with variance premium only as an output.

We feed our calibrated models with the historical consumption growth data from 1890 to 2009. We find that in normal times variance premium is at a low range of 5-8, while it shot up to 41 during the Great Depression, 40 and 34 during the panics of 1893-1894 and 1907-1908, around 31 during the 1925 depression, of the range 18-32 in deep recessions of 1914-1915 and 1917-1918, and about 19 when U.S. joined Wold War II. After the war, the highest level of

[3] The range 5-19 of existing estimates for variance premium depends on whether to use index or futures as underlying, whether to include the recent crisis period or not, and whether to use expected or realized variance (see, e.g., Carr and Wu (2009), Bollerslev, Tauchen and Zhou (2009), Drechsler and Yaron (2011), Zhou and Zhu (2011), Bollerslev, Sizaova, Tauchen (2011), Drechsler (2011), among others).

variance premium ever reached is near 12 at the end of 2007-2009 global financial crisis. For comparison, the model with only risk aversion produces variance premium in a range of 0.06 to 0.53, while the model with only risk aversion and intertemporal substitution has a range of 0.19 to 2.1. Therefore, historically speaking, ambiguity aversion contributes almost 95 percent of the total variance premium, which is largely in line with the calibration evidence.

The existing approaches to generating realistic variance premium dynamics typically involve Epstein-Zin preferences combined with stochastic volatility-of-volatility in consumption (Bollerslev, Tauchen, and Zhou (2009)) or joint jumps in consumption volatility and growth (Todorov (2010) and Drechsler and Yaron (2011)). Alternatively, variance premium may arise from time-varying rare disasters (Gabaix (2011)) or extreme tail risk (Kelly (2011)). Within the time-separable expected utility framework, variance premium may also be generated by introducing a stochastic volatility process that is statistically correlated with the consumption process (Heston (1993) and Bates (1996), among others). Without having to rely on these exogenous factors, our model is able to generate large endogenous variance premium. Our approach highlights the important role of ambiguity aversion and learning in generating large variance premium dynamics.

Like our paper, Drechsler (2011) explores the implications of ambiguity aversion for variance premium. His paper is more ambitious than ours in that he also studies equity index option prices in addition to equity returns, conditional variance, and the risk-free rate. His model is more complicated than ours in that he incorporates stochastic volatility and jumps in the expected growth and growth volatility processes. As discussed earlier, our model is parsimonious and can generate empirically reasonable variance premium dynamics without relying on exogenously specified complex consumption or dividend dynamics.

Another difference between our model and the Drechsler (2011) model is that we adopt the recursive smooth ambiguity model with learning, while he adopts the continuous-time recursive multiple-priors model of Chen and Epstein (2002) without learning. The latter utility model is a dynamic generalization of Gilboa and Schmeidler (1989). There are many applications of the multiple-priors model in finance (e.g., Epstein and Wang (1994), Epstein and Miao (2003)). An alternative approach to modeling ambiguity is based on robust control proposed by Hansen and Sargent (2008) (see, e.g., Liu, Pan and Wang (2005) and Hansen and Sargent (2011) for applications in finance). An important advantage of the smooth ambiguity model over other models of ambiguity such as the multiple-priors model 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.[4]

The rest of the paper is organized as follows. Section 2 describes the stylized facts of variance premium. Section 3 introduces the model, followed by Section 4, a decomposition of variance premium. Section 5 conducts a calibration exercise to study variance premium and its predictability pattern, and to provide a historical reconstruction of variance premium from 1890 to 2009. Section 6 concludes.

2. Stylized Facts of Variance Premium

Variance premium is formally defined as the difference between the risk-neutral expectation $\mathbb{E}_t^{\mathbb{Q}}(\cdot)$ and the objective expectation $\mathbb{E}_t(\cdot)$ of the return variance Σ_{t+1}, that is,

$$VP_t \equiv \mathbb{E}_t^{\mathbb{Q}}(\Sigma_{t+1}) - \mathbb{E}_t(\Sigma_{t+1}). \tag{1}$$

The availability of Chicago Board Options Exchange (CBOE) VIX index makes it straightforward to measure the risk-neutral expectation of stock market returns.[5] The objective expecta-

[4] See Chen, Ju and Miao (2011) and Collard et al. (2009) for other applications of the smooth ambiguity model in finance.

[5] We use $\text{VIX}^2/12$ as a measure of risk-neutral expectation of return variance. The CBOE VIX index is based on the highly liquid S&P 500 index options along with the "model-free" approach explicitly tailored to replicate the risk-neutral variance of a fixed one-month maturity. See, e.g., Carr and Wu (2009) for the definition of model-free implied variance.

tion of variance can be measured as the forecasted realized variance using daily returns, where the forecast is based on the lagged realized variance and lagged implied variances. Both VIX and daily return are for the S&P 500 index as a proxy for the stock market. We use all currently available sample from January 1990 to December 2011 to measure the variance premium.

We find that there are three challenging facts for standard asset pricing models: (i) a large and volatile variance premium, (ii) short-run predictability of variance premium for stock returns which is complementary to the long-run predictability of dividend yield, and (iii) countercyclicality of variance premium—high in bad times and low in good times.

[Insert Figure 1 Here.]

Figure 1 top panel plots the monthly time series of variance premium, which tends to rise around the 1990 and 2001 economic recessions but reaches a much higher level during the 2008 financial crisis and around the 1997-1998 Asia-Russia-LTCM crisis. There are also huge run-ups of variance premium around May 2010 and August 2011, when heightened Greece sovereign default risk threatens the Euro area financial stability. The sample mean of variance premium is 10.80 (in percentages squared, monthly basis), with a standard deviation of 23.36. Nevertheless, variance premium is not a very persistent process with an AR(1) coefficient of 0.17 at monthly frequency. Our estimates are broadly consistent with the existing estimates in the literature. For example, using data of 5-minute log returns on the S&P 500 futures from January 1990 to December 2009, Drechsler (2011) reports that the mean, the standard deviation, and the AR(1) coefficient are equal to 10.55, 8.47, and 0.61, respectively. The sizable variance premium and its temporal variation are puzzling in that standard consumption-based asset pricing models would predict a zero variance premium (Bollerslev, Tauchen, and Zhou (2009)).

Recent empirical evidence has also suggested that stock market return is predictable by the variance premium over a few quarters horizon (Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011)). Such a finding contrasts the long-run multi-year return predictability that is typically associated with the traditional valuation ratios like dividend yield and price-earnings (P/E) ratio (see, e.g., Fama and French (1988), Campbell and Shiller (1988), among others). In fact, as first reported in Bollerslev and Zhou (2007), the short-run and long-run predictability seem to be complementary in the sense that when dividend yield or P/E ratio is included in the regressions, the predictability of variance premium is not crowded out but often is enhanced. It is not clear that any existing consumption-based asset pricing model can replicate such a puzzling phenomenon.

To further appreciate the economics behind the apparent connection between the variance premium and the underlying macroeconomy, the bottom panel of Figure 1 plots variance premium together with the quarterly growth rate in GDP. As seen from the figure, there is a tendency for variance premium to rise in one-to-two quarters before a decline in GDP, while it typically narrows ahead of an increase in GDP. Indeed, the sample correlation equals -0.08 between current variance premium and two-quarter-ahead GDP (as first reported in Bollerslev and Zhou (2007)). In other words, variance premium is countercyclical, which is the third puzzle that a standard consumption-based asset pricing model can hardly replicate (Drechsler and Yaron (2011)).

3. The Model

Consider a representative agent consumption-based asset pricing model studied by Ju and Miao (2012). There are three key elements of this model. First, consumption growth follows a Markov regime-switching process and dividends are leveraged claims on consumption. Second, the representative agent does not observe economic regimes and learns about them by observing past data. Third, and the most important, the representative agent has ambiguous beliefs about the economic regimes. His preferences are represented by the generalized smooth ambiguity utility model proposed by Hayashi and Miao (2011) and Ju and Miao (2012).

We now describe this model formally. Aggregate consumption follows a regime-switching process:[6]

$$\ln\left(\frac{C_{t+1}}{C_t}\right)=\kappa_{z_{t+1}}+\sigma\varepsilon_{t+1}, \tag{2}$$

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 $\kappa_1>\kappa_2$. Aggregate dividends are leveraged claims on consumption and satisfy

$$\ln\left(\frac{D_{t+1}}{D_t}\right)=\zeta\ln\left(\frac{C_{t+1}}{C_t}\right)+g_d+\sigma_d e_{t+1}, \tag{3}$$

where e_{t+1} is an iid standard normal random variable, and is independent of all other random variables. The parameter $\zeta>0$ can be interpreted as the leverage ratio on expected consumption growth as in Abel (1999). This parameter and the parameter σ_d allow 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 modeling of the dividend process is convenient because it does not introduce any new state variable in our model.

Assume that the representative agent does not observe economic regimes. He observes the history of consumption and dividends up to the current period $t: s^t=\{C_0,D_0,C_1,D_1,...,C_t,D_t\}$. In addition, he knows the parameters of the model (e.g., ζ, g_d, σ, and σ_d). But he has ambiguous beliefs about the hidden states. His preferences are represented by the generalized recursive smooth ambiguity utility model. To define this utility, we first derive the evolution of the posterior state beliefs. Let $\mu_t=\Pr\left(z_{t+1}=1|s^t\right)$.The prior belief μ_0 is given. By Bayes' Rule, we can derive:

$$\mu_{t+1}=\frac{\lambda_{11}f\left(\ln\left(C_{t+1}/C_t\right),1\right)\mu_t+\lambda_{21}f\left(\ln\left(C_{t+1}/C_t\right),2\right)(1-\mu_t)}{f\left(\ln\left(C_{t+1}/C_t\right),1\right)\mu_t+f\left(\ln\left(C_{t+1}/C_t\right),2\right)(1-\mu_t)}, \tag{4}$$

[6] The regime-switching consumption process has been used widely in the asset pricing literature (see, e.g., Cecchetti et al (2000) and Veronesi (1999)). One may view this process as a nonlinear version of the long-run risk process studied by Bansal and Yaron (2004).

where $f(y,i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-(y-\kappa_i)^2 / \left(2\sigma^2\right)\right]$ is the density function of the normal distribution with mean κ_i and variance σ^2. By our modeling of dividends in (3), dividends do not provide any new information for belief updating and for the estimation of the hidden states.

Let $V_t(C)$ denote the continuation utility at date t. Following Ju and Miao (2012), assume that $V_t(C)$ satisfies the following recursive equation:

$$\begin{aligned} V_t(C) &= \left[(1-\beta) C_t^{1-\rho} + \beta \left\{\mathcal{R}_t \left(V_{t+1}(C)\right)\right\}^{1-\rho}\right]^{\frac{1}{1-\rho}}, && (5) \\ \mathcal{R}_t\left(V_{t+1}(C)\right) &= \left\{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{\frac{1-\eta}{1-\gamma}}\right\}^{\frac{1}{1-\eta}}, && (6) \end{aligned}$$

where $\mathcal{R}_t(V_{t+1}(C))$ is an uncertainty aggregator that maps an s^{t+1}-measurable random variable $V_{t+1}(C)$ to an s^t-measurable random variable. Furthermore, $\pi_{z,t}$ denotes the likelihood distribution conditioned on the history s^t and on a given economic regime $z_{t+1} = z$ in period $t+1$, $\beta \in (0,1)$ represents the subjective discount factor, $1/\rho > 0$ represents the elasticity of intertemporal substitution (EIS), $\gamma > 0$ represents the degree of risk aversion, and $\eta \geq \gamma$ represents the degree of ambiguity aversion. We use $\mathbb{E}_{\mu_t}$ and $\mathbb{E}_{\pi_{z,t}}$ to denote conditional expectation operators with respect to the distributions $(\mu_t, 1-\mu_t)$ and $\pi_{z,t}$, respectively.

To interpret the above utility model, we first observe that in the deterministic case, (5) and (6) reduce to

$$V_t(C) = \left[(1-\beta) C_t^{1-\rho} + \beta V_{t+1}(C)^{1-\rho}\right]^{\frac{1}{1-\rho}}.$$

This justifies the interpretation of $1/\rho$ as EIS. When $\eta = \gamma$, (5) and (6) reduce to

$$V_t(C) = \left[(1-\beta) C_t^{1-\rho} + \beta \left\{\mathbb{E}_t\left[V_{t+1}^{\frac{1}{1-\gamma}}(C)\right]\right\}^{\frac{1-\rho}{1-\gamma}}\right]^{\frac{1}{1-\rho}},$$

where $\mathbb{E}_t$ is the expectation operator for the predictive distribution conditioned on history s^t. This is the Epstein and Zin (1989) model with partial information and justifies the interpretation of γ as a risk aversion parameter. In this case, the posterior and likelihood distributions $(\mu_t, 1-\mu_t)$ and $\pi_{z,t}$ can be reduced to a single predictive distribution in (6) by Bayes' Rule.

One can then analyze the model under the reduced information set $\{s^t\}$ and under the predictive distribution as in the standard expected utility model under full information.

When $\eta > \gamma$, the posterior and likelihood distributions cannot be reduced to a single distribution in (6). This irreducibility of compound distributions captures ambiguity aversion. Intuitively, given history s^t and the economic regime z_{t+1} in period $t+1$, the agent can compute the certainty equivalent of expected continuation value, $\left(\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{\frac{1}{1-\gamma}}$. If the agent is ambiguity averse to the variation of economic regimes, he is averse to the variation of $\left(\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{\frac{1}{1-\gamma}}$ across different regimes z_{t+1}. Thus, he evaluates the ex ante continuation value using a concave function with a curvature η so that he enjoys the certainty equivalent value $\mathcal{R}_t(V_{t+1}(C))$. Only when $\eta > \gamma$, the certainty equivalent to an ambiguity-sensitive agent is less than that to an agent with expected utility, i.e.,

$$\mathcal{R}_t(V_{t+1}(C)) < \left(\mathbb{E}_t\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{\frac{1}{1-\gamma}},$$

which implies that ambiguity is costly, compared to an agent with expected utility. If we identify expected utility as an ambiguity neutrality benchmark, then $\eta > \gamma$ fully characterizes ambiguity aversion (Klibanoff, Marinacci, and Mukerji (2005)). Alternatively, if we interpret uncertainty about economic regimes as second-order risk, then ambiguity aversion is equivalent to second-order risk aversion (see Ergin and Gul (2009), and Hayashi and Miao (2011)).

To understand the asset pricing implications of the above model, one only needs to understand the pricing kernel. Ju and Miao (2012) show that the pricing kernel for the generalized recursive smooth ambiguity utility model is given by

$$M_{t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\left(\frac{V_{t+1}}{\mathcal{R}_t(V_{t+1})}\right)^{\rho-\gamma}\left(\frac{\left(\mathbb{E}_{\pi_{z,t}}\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t(V_{t+1})}\right)^{-(\eta-\gamma)}. \tag{7}$$

Equation (7) reveals that there are two adjustments to the standard pricing kernel $\beta(C_{t+1}/C_t)^{-\rho}$. The first adjustment is present for the recursive expected utility model of Epstein and Zin

(1989). This adjustment is the middle term on the right-hand side of (7). The second adjustment is due to ambiguity aversion, which is given by the last term on the right-hand side of (7). This adjustment depends explicitly on the hidden state in period $t+1$, z_{t+1}, in that $\pi_{z,t}$ depends on the state $z_{t+1} = z$. It has the feature that an ambiguity averse agent with $\eta > \gamma$ puts a higher weight on the pricing kernel when his continuation value is low in a recession when $z = 2$. We will show later that this pessimistic behavior helps explain the equity premium, variance premium, and risk-free rate puzzles.

Given the above pricing kernel, the return $R_{k,t+1}$ on any traded asset k satisfies the Euler equation:

$$\mathbb{E}_t \left[M_{t+1} R_{k,t+1} \right] = 1. \tag{8}$$

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 (8).

Let $P_{e,t}$ denote the date t price of dividend claims. Using equations (7) and (8) and the homogeneity property of V_t, we can show that the price-dividend ratio $P_{e,t}/D_t$ is a function of the state beliefs, denoted by $\varphi(\mu_t)$, so is the ratio V_t/C_t. Specifically,

$$P_{e,t} = \varphi(\mu_t) D_t, \tag{9}$$

$$V_t = G(\mu_t) C_t. \tag{10}$$

By definition, we can write the equity return as:

$$R_{e,t+1} = \frac{P_{e,t+1} + D_{t+1}}{P_{e,t}} = \frac{D_{t+1}}{D_t} \frac{1 + \varphi(\mu_{t+1})}{\varphi(\mu_t)}. \tag{11}$$

This equation implies that the state beliefs drive changes in the price-dividend ratio, and hence dynamics of equity returns. In Section 5, we will show numerically that ambiguity aversion and learning under ambiguity help amplify consumption growth uncertainty, while Bayesian learning has a modest quantitative effect.

4. Variance Premium Decomposition

In this section, we explore model implications for variance premium. Denote the conditional variance of equity return by $\Sigma_t \equiv Var_t\left[R_{e,t+1}\right]$. Variance premium is defined as

$$VP_t = \mathbb{E}_t^{\mathbb{Q}}\left(\Sigma_{t+1}\right) - \mathbb{E}_t\left(\Sigma_{t+1}\right) = \frac{\mathbb{E}_t\left[\Sigma_{t+1}M_{t+1}\right]}{\mathbb{E}_t\left[M_{t+1}\right]} - \mathbb{E}_t\left[\Sigma_{t+1}\right], \tag{12}$$

where $\mathbb{Q}$ represents the risk-neutral measure. To understand the determinant of the variance premium, we rewrite (7) as

$$M_{t+1} = M_{t+1}^{EZ}M_{z,t}^{A}, \text{ for } z_{t+1} = z \in \{1,2\}, \tag{13}$$

where M_{t+1}^{EZ} is the Epstein-Zin pricing kernel defined by

$$M_{t+1}^{EZ} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\left(\frac{V_{t+1}}{\mathcal{R}_t\left(V_{t+1}\right)}\right)^{\rho-\gamma}, \tag{14}$$

and $M_{z,t}^{A}$ is the ambiguity adjustment of the pricing kernel defined by

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

Here $\mathbb{E}_t^{z}$ denotes the conditional expectation operator given that the state in period $t+1$ is $z_{t+1} = z \in \{1,2\}$.

Now, we use (13) to compute

$$\frac{\mathbb{E}_t\left[\Sigma_{t+1}M_{t+1}\right]}{\mathbb{E}_t\left[M_{t+1}\right]} = \frac{\mu_t\mathbb{E}_t^1\left[\Sigma_{t+1}M_{t+1}^{EZ}\right]M_{1,t}^{A} + (1-\mu_t)\,\mathbb{E}_t^2\left[\Sigma_{t+1}M_{t+1}^{EZ}\right]M_{2,t}^{A}}{\mu_t\mathbb{E}_t^1\left[M_{t+1}^{EZ}\right]M_{1,t}^{A} + (1-\mu_t)\,\mathbb{E}_t^2\left[M_{t+1}^{EZ}\right]M_{2,t}^{A}}.$$

Using the fact that

$$\mathbb{E}_t^{i}\left[\Sigma_{t+1}M_{t+1}^{EZ}\right] = \mathbb{E}_t^{i}\left[\Sigma_{t+1}\right]\mathbb{E}_t^{i}\left[M_{t+1}^{EZ}\right] + Cov_t^{i}\left(\Sigma_{t+1}, M_{t+1}^{EZ}\right),$$

we can compute

$$\begin{aligned} VP_t \;=\; & \frac{\mu_t\mathbb{E}_t^1\left[\Sigma_{t+1}\right]\mathbb{E}_t^1\left[M_{t+1}^{EZ}\right]M_{1,t}^{A} + (1-\mu_t)\,\mathbb{E}_t^2\left[\Sigma_{t+1}\right]\mathbb{E}_t^2\left[M_{t+1}^{EZ}\right]M_{2,t}^{A}}{\mu_t\mathbb{E}_t^1\left[M_{t+1}^{EZ}\right]M_{1,t}^{A} + (1-\mu_t)\,\mathbb{E}_t^2\left[M_{t+1}^{EZ}\right]M_{2,t}^{A}} \\ & + \frac{\mu_t Cov_t^1\left(\Sigma_{t+1}, M_{t+1}^{EZ}\right)M_{1,t}^{A} + (1-\mu_t)\,Cov_t^2\left(\Sigma_{t+1}, M_{t+1}^{EZ}\right)M_{2,t}^{A}}{\mu_t\mathbb{E}_t^1\left[M_{t+1}^{EZ}\right]M_{1,t}^{A} + (1-\mu_t)\,\mathbb{E}_t^2\left[M_{t+1}^{EZ}\right]M_{2,t}^{A}} \\ & - \mu_t\mathbb{E}_t^1\left[\Sigma_{t+1}\right] - (1-\mu_t)\,\mathbb{E}_t^2\left[\Sigma_{t+1}\right], \end{aligned}$$

where Cov_t^i denotes the conditional covariance operator given time t information and given the state at time $t+1$, $z_{t+1} = i \in \{1, 2\}$. Define the distorted belief about the high growth state by

$$\widehat{\mu}_t \equiv \frac{\mu_t \mathbb{E}_t^1 \left[M_{t+1}^{EZ} \right] M_{1,t}^A}{\mu_t \mathbb{E}_t^1 \left[M_{t+1}^{EZ} \right] M_{1,t}^A + (1 - \mu_t) \mathbb{E}_t^2 \left[M_{t+1}^{EZ} \right] M_{2,t}^A}. \tag{15}$$

We can then rewrite the variance premium as

$$\begin{aligned} VP_t &= (\mu_t - \widehat{\mu}_t) \left(\mathbb{E}_t^2 [\Sigma_{t+1}] - \mathbb{E}_t^1 [\Sigma_{t+1}] \right) \\ &\quad + \frac{\widehat{\mu}_t}{\mathbb{E}_t^1 \left[M_{t+1}^{EZ} \right]} Cov_t^1 \left(\Sigma_{t+1}, M_{t+1}^{EZ} \right) + \frac{1 - \widehat{\mu}_t}{\mathbb{E}_t^2 \left[M_{t+1}^{EZ} \right]} Cov_t^2 \left(\Sigma_{t+1}, M_{t+1}^{EZ} \right). \end{aligned} \tag{16}$$

Note that by equation (13), we can replace M_{t+1}^{EZ} by M_{t+1} in the above equation. This equation reveals that variance premium is determined by three components: (i) the expression $(\mu_t - \widehat{\mu}_t)$, (ii) the expression $\left(\mathbb{E}_t^2 [\Sigma_{t+1}] - \mathbb{E}_t^1 [\Sigma_{t+1}] \right)$, and (iii) the two terms related to the conditional covariance between the stock variance and the pricing kernel in the second line of the above equation.

We will show in the next section that the stock return variance is countercyclical, and hence the second component is positive. We will also show numerically that ambiguity aversion amplifies the countercyclicality significantly, making the second component large and volatile.

When the agent prefers early resolution of uncertainty, i.e., $\gamma > \rho$, the Epstein-Zin pricing kernel enhances the countercyclicality of the pricing kernel given in (14). Thus, the stock return variance and the Epstein-Zin pricing kernel are positively correlated, implying that the last covariance component in (16) is positive. Our numerical analysis in Section 5 shows that this component is numerically small for reasonable parameter values.

Now, we examine the first component $(\mu_t - \widehat{\mu}_t)$ which is due to belief distortions. In the special case with time-additive CRRA utility (i.e., $\eta = \rho = \gamma$), we can show that the uncertainty-adjusted belief in (15) is given by

$$\widehat{\mu}_t^{CRRA} = \frac{\mu_t \mathbb{E}_t^1\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]}{\mu_t \mathbb{E}_t^1\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right] + (1-\mu_t)\,\mathbb{E}_t^2\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right]} = \frac{\mu_t e^{-\gamma\kappa_1}}{\mu_t e^{-\gamma\kappa_1} + (1-\mu_t)\,e^{-\gamma\kappa_2}},$$

where the last equality follows from substitution of equation (2). Since we assume that state 1 is the boom state, i.e., $\kappa_1 > \kappa_2$, we deduce that $\widehat{\mu}_t^{CRRA} < \mu_t$.

For the Epstein-Zin utility ($\eta = \gamma \neq \rho$), the pricing kernel is given by M_{t+1}^{EZ} in equation (14). Plugging this equation and the utility function in (10) into (15), we can derive the uncertainty-adjusted belief as

$$\widehat{\mu}_t^{EZ} = \frac{\mu_t \mathbb{E}_t^1\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} G_{t+1}^{\rho-\gamma}\right]}{\mu_t \mathbb{E}_t^1\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} G_{t+1}^{\rho-\gamma}\right] + (1-\mu_t)\,\mathbb{E}_t^2\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} G_{t+1}^{\rho-\gamma}\right]}.$$

Suppose that the representative agent prefers early resolution of uncertainty so that $\rho < \gamma$. In this case, EIS $1/\rho$ is greater than $1/\gamma$. Suppose that $Cov_t^i\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}, G_{t+1}^{\rho-\gamma}\right] \approx 0$, which is verified in our numerical results below. We can then show that

$$\begin{aligned}
\widehat{\mu}_t^{EZ} &\approx \frac{\mu_t e^{-\gamma\kappa_1}\mathbb{E}_t^1\left[G_{t+1}^{\rho-\gamma}\right]}{\mu_t e^{-\gamma\kappa_1}\mathbb{E}_t^1\left[G_{t+1}^{\rho-\gamma}\right] + (1-\mu_t)\,e^{-\gamma\kappa_2}\mathbb{E}_t^2\left[G_{t+1}^{\rho-\gamma}\right]} \\
&< \widehat{\mu}_t^{CRRA} \text{ when } \rho < \gamma.
\end{aligned}$$

This equation says that the component $\left(\mu_t - \widehat{\mu}_t^{EZ}\right)$ in the Epstein-Zin model is larger than $\left(\mu_t - \widehat{\mu}_t^{CRRA}\right)$ in the standard time-additive CRRA utility model. Thus, holding everything else constant, the Epstein-Zin model can generate a larger variance premium than the standard time-additive CRRA utility model. The intuition is that, when $\rho < \gamma$, the agent puts more weight on the pricing kernel in the recession state as revealed by equation (14). Thus, the agent with Epstein-Zin preferences fears equity volatility more in recessions, generating a higher variance premium.

When the representative agent is also ambiguity averse (i.e., $\eta > \gamma$), there is an additional adjustment in the pricing kernel in (7) so that the agent puts an additional weight on the pricing kernel in recessions when he has low continuation values. As a result, ambiguity aversion further amplifies the distortion by further decreasing the belief about the high-growth state in that

$$\widehat{\mu}_t^A < \widehat{\mu}_t^{EZ} \text{ when } \eta > \gamma,$$

where $\widehat{\mu}_t^A$ is given by (15). Therefore, holding everything else constant, an ambiguity averse agent demands a higher variance premium than an ambiguity neutral one with Epstein-Zin preferences.

5. Results

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 as in Ju and Miao (2012).[7] We first calibrate the model at annual frequency in Section 5.1. We then study properties of unconditional and conditional moments of variance premium generated by our model in Sections 5.2-4. For comparison, we also solve four benchmark models. Models 1A and 2A are the models with standard time-additive CRRA utility and with Epstein-Zin preferences under full information. Models 1B and 2B are the corresponding models with partial information and Bayesian learning. The last two models are special cases of our full model with $\eta = \gamma = \rho$ and with $\eta = \gamma \neq \rho$, respectively.

5.1. Calibration

We calibrate the model to the sample period from 1890 to 2009 based on Robert Shiller's data of consumption growth, dividends growth, stock returns and risk-free rates.[8] In total, there are twelve parameters to be calibrated as listed in Table 1. Specifically, the five parameters $(\lambda_{11}, \lambda_{22}, \kappa_1, \kappa_2, \sigma)$ that govern the consumption process are estimated by the maximum likelihood method based on the annual per capita US consumption data from 1890

[7] We run 10,000 simulations. Increasing this number does not change our results significantly.
[8] The data is downloaded from Professor Robert Shiller's website:
http://www.econ.yale.edu/~shiller/data/ie_data.xls.

to 2009. Table 1 reveals that the high-growth state is highly persistent, with consumption growth in this state being 2.42 percent. The economy spends most of the time in this state with the unconditional probability of being in this state given by $(1 - \lambda_{22}) / (2 - \lambda_{11} - \lambda_{22}) = (1 - 0.5432) / (2 - 0.5432 - 0.9799) = 0.96$. The low-growth state is moderately persistent, but very bad, with consumption growth in this state being -5.6 percent. The long-run average rate of consumption growth is 2.08 percent.

[**Insert Table 1 Here**]

The leverage ratio ζ is set to 2.74 following Abel (1999), and then g_d is chosen as -0.0362 so that the average rate of dividend growth is equal to that of consumption growth. Furthermore, given that the volatility of dividend growth in the data is about 0.116, we choose $\sigma_d = 0.0629$ to match this volatility using (3).

The preference parameters are calibrated using the methodology in Ju and Miao (2012). First, we set $\gamma = 2$ which is widely used in macroeconomics and finance. We choose this small number in order to demonstrate that the main force of our model comes from ambiguity aversion, but not risk aversion. Following Bansal and Yaron (2004), we set EIS to 1.5 or $\rho = 1/1.5$. Finally, we select the subjective discount factor β and the ambiguity aversion parameter η to match the mean risk-free rate of 0.0191 and the mean equity premium of 0.0574 from the data reported in Table 2. We obtain $\beta = 0.9838$ and $\eta = 10.3793$.

In sum, the calibrated parameter values listed in Table 1 are broadly consistent with those in Ju and Miao (2012), with the difference reflecting different sample periods. Here we discuss the value of the ambiguity aversion parameter η only, since it is the most important parameter for our analysis. Because there is no consensus study of the magnitude of ambiguity aversion in the literature, it is hard to judge how reasonable it is. We may use the thought experiment related to the Ellsberg Paradox (Ellsberg (1961)) in a static setting designed in Chen, Ju and Miao (2011) and Ju and Miao (2012) to have a sense of our calibrated value. Ju and Miao

(2012) show that their calibrated value $\eta = 8.864$ implies that the ambiguity premium in the thought experiment is equal to 1.7 percent of the expected prize value when one sets $\gamma = 2$ and the prize-wealth ratio of 1 percent. Similarly, we can compute that the ambiguity premium is equal to 2.08 percent of the expected prize for $\eta = 10.379$. 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-style experiments. Given this evidence, our calibrated ambiguity aversion parameter seems small and reasonable. It is consistent with the experimental findings, though they are not the basis for our calibration.

5.2. Variance Premium and Equity Premium

To evaluate the performance of our model, we first examine model predictions of moments other than the mean risk-free rate and the mean equity premium. Table 2 shows that the model implied equity premium volatility is equal to 17.26 percent which is quite close to 18.80 percent in the data. As Shiller (1981) and Campbell (1999) point out, it is challenging for the standard rational model to explain the high equity volatility observed in the data, generating the so called equity volatility puzzle. By contrast, our model with ambiguity aversion can successfully match both the mean and volatility of equity premium. For comparison, when we shut down ambiguity aversion by setting $\eta = \gamma = 2$ as in Model 2B, we obtain the model implied mean risk-free rate 2.86 percent, mean equity premium 0.93 percent, and equity volatility 12.91 percent, which are far away from the data. By adjusting the risk aversion parameter γ or the EIS parameter ρ, these numbers may change, but still one cannot obtain a reasonably good fit for all three moments. However, the performance of the Epstein-Zin model is much better than that of the time-additive CRRA model with $\eta = \gamma = \rho = 2$.

Table 2 shows that Model 2A and Model 2B yield very similar predictions. This means that introducing Bayesian learning into the models with time-additive expected utility or Epstein-Zin utility has a small quantitative impact, confirming the findings reported in Hansen (2007) and Ju and Miao (2012).

Table 2 also shows that our model generated volatility of the risk-free rate is lower than the data (0.0110 versus 0.0578). Campbell (1999) argues that the high volatility of the real risk-free 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 0.018 using the US postwar quarterly data. Thus, we view our model generated low risk-free rate volatility as a success, rather than a shortcoming. By contrast, the widely used habit formation model (e.g., Jermann (1998) and Boldrin, Christiano and Fisher (2001)) typically predicts a too high volatility of the risk-free rate. To overcome this issue, Campbell and Cochrane (1999) calibrate their model by fixing a constant risk-free rate.

[Insert Table 2 Here.]

Turn to model predictions regarding variance premium. Table 2 shows that our full model implied mean variance premium is 8.51 (percentage squared, monthly basis), which is about 80 percent of the data, 10.80. By contrast, Models 1A and 1B with standard time-additive CRRA utility generate vary small values of the mean variance premium (0.0655 and 0.0997). Separating EIS from risk aversion as in Models 2A and 2B raises the mean variance premium to 0.3158 and 0.3150, respectively. These numbers are far below the data, explaining about 2.9 percent of the data.

Because the full model with ambiguity aversion nests Model 2B as a special case and Model 2B in turn nests Model 1B as a special case, the above numbers imply that risk aversion alone contributes about 1 percent to the model implied mean variance premium, separating EIS from risk aversion contributes about 3 percent, and the remaining 96 percent is attributed to ambiguity aversion.

[Insert Table 3 Here.]

To understand why our model with ambiguity aversion and learning can generate a large mean variance premium, we decompose variance premium in Table 3. As discussed in Section 4, the variance premium is determined by three factors and equal to the product of the first two factors plus the third factor. Separating risk aversion from EIS raises the first factor—belief distortions $(\mu_t - \widehat{\mu}_t)$—from 0.006 in Model 1B to 0.008 in Model 2B. Separating ambiguity aversion from risk aversion raises the mean value of the first factor from 0.008 in Model 2B to 0.1105 in the full model, a remarkable 12 times increase.

Turn to the second factor, the stock variance differentials between recessions and booms $(\mathbb{E}_t^2[\Sigma_{t+1}] - \mathbb{E}_t^1[\Sigma_{t+1}])$. Separating risk aversion from EIS raises the mean value of this factor from 7.274 in Model 1B to 16.53 in Model 2B. Separating ambiguity aversion from risk aversion further raises it from 16.53 in Model 2B to 61.486 in the full model, generating a significant 3 times increase.

The mean value of the product of these two factors is equal to 0.045, 0.156 and 7.306 in Model 1B, Model 2B and the full model, respectively, and accounts for 0.4, 1.4, 67.6 percent of the mean variance premium in the data, respectively. This result shows that the full model with ambiguity aversion and learning raises the product in Model 1B by 161 times and in Model 2B by 46 times.

Table 3 reveals that the third covariance factor is quantitatively small. Its mean value is equal to 0.054, 0.159, and 1.205 in Model 1B, Model 2B and the full model, respectively, and accounts for $0.5, 1.5, 11.1$ percent of the total mean variance premium in the data, respectively. Ambiguity aversion raises this factor in Model 1B by 21 times and in Model 2B by about 6 times.

[Insert Figures 2-3 Here.]

To further understand the intuition, Figure 2 plots the conditional variance premium and the three factors in the variance premium decomposition as functions of the Bayesian beliefs about the high-growth state. This figure shows that all three factors are positive and hump-shaped. Ambiguity aversion amplifies each factor significantly, generating large movements of conditional variance premium. The intuition behind this result has been discussed in Sections 1 and 4. One important property for this intuition to work is that the conditional stock return variance must be countercyclical. We now examine this issue below.

Figure 3 plots the price-dividend ratio and the conditional stock return variance as functions of the Bayesian beliefs about the high-growth state. This figure shows that the price-dividend ratio in Model 1B and Model 2B is almost linear. By contrast, it is strictly convex and shows a significant curvature in our full model with ambiguity aversion. In a continuous-time model with time-additive exponential utility similar to Model 1B, Veronesi (1999) proves theoretically that the price-dividend ratio is a convex function. This result implies that the agent overreacts to bad news in good times and underreacts to good news in bad times, generating large countercyclical movements of stock return volatility. In particular, the stock return variance is hump-shaped as illustrated in the top panel of Figure 3. Since the economy spends most time in the good state, the economy stays most of the time in the right arm of this panel. Our numerical results show that the countercyclical movements of stock variance are amplified remarkably by ambiguity aversion.

We now examine other statistics reported in Table 2. The standard deviation of variance premium in the data is 23.36, which is very large. Our model implied standard deviation is 7.53. Though it is still less than the data, it is about 20 times to 100 time as large as that in other models reported in Table 2. When using the data reported in Drechsler (2011), our model implied mean and standard deviation of variance premium are quite close to Drechsler's (2011) estimates, 10.55 and 8.47, respectively. Regarding autocorrelation coefficient, other models listed in Table 2 deliver too high coefficients compared to the data. Our model gives a lower level, 0.3778, though it is still higher than the data 0.17 but lower than 0.67 as reported in Drechsler (2011).

5.3. Return Predictability

Recent empirical evidence has suggested that variance premium is a highly significant predictor for stock market return at short horizons especially one quarter, with a Newey-West t-statistics ranging from 2.86 to 3.53 and R^2's from 6 to 8 percent, while the predictability completely disappears for longer horizons beyond one year (Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011), and Drechsler (2011)). On the other hand, traditional valuation ratios like price-dividend or price-earning ratios only have significant return predictability for horizons longer than one year, with R^2's increasing from 5 to 14 percent over one-to-five year horizons (Drechsler and Yaron (2011)). More importantly, when variance premium and price-earning ratio are combined together, there is complementarity in that the joint regression R^2 (16.76 percent) is higher than the sum of the two R^2's of univariate regressions—6.82 and 6.55 percent (see, e.g., Bollerslev, Tauchen, and Zhou (2009)).

[Insert Table 4 Here.]

Motivated by these important empirical findings on return predictability, we report here our model implied values of regression R^2's, slope coefficients, and t-statistics, at horizons of 1, 2, 3, and 5 years based on the benchmark calibration parameter settings. As can be seen from Table 4, both dividend yield and variance premium can predict return with highly significant t-statistics in univariate regressions. These results closely mimic the empirical regularity that both dividend yield (or P/E ratio) and variance premium are return predictors. It should be pointed out that since there is only one state variable μ_t in our model, when dividend yield and variance premium are combined together, they crowd out each other as both become insignificant. More state variable(s) need to be introduced into our model, to make both dividend yield and variance premium to be significant and complementary in joint regressions, as empirically reported by Bollerslev, Tauchen, and Zhou (2009) and Dreschler and Yaron (2011).

We should mention that although our model cannot generate predictability pattern for joint regressions, our model performs much better than models without ambiguity aversion, e.g., Models 1B and 2B. We find that even in univariate regressions, neither dividend yield nor variance premium is a significant predictor for Models 1B and 2B.[9]

5.4. Historical Variance Premium

To further elicit economic intuition on why and how variance premium changes with economic fundamental, we feed our calibrated models with historical consumption growth data from 1890 to 2009 from Robert Shiller's web site. As shown in Figure 4 top panel, the negative spikes of consumption growth over this long history capture the Great Depression (1929-1933), financial panics (1893-1894, 1907-1908), deep recessions (1914-1915, 1917-1918, 1925, 1937-1938), and U.S. initial engagement in World War II (1942). Also, the major belief switches or spikes, as shown in Figure 4 bottom panel, do reflect those severe down times, while minor belief deviations from 1 seem to capture all other mild economic recessions.

[Insert Figures 4 and 5 Here.]

We can see from Figure 5 that, in normal times of the past century, variance premium is at a low range of 5-8, while it shot up to 41 during the Great Depression, 40 and 34 during the financial panics of 1894 and 1908, around 31 during the 1925 depression, in the range 18-32 during deep recessions of 1914-1915 and 1917, and about 19 when U.S. joined the war in 1942. After the war, the highest level of variance premium ever reached is about 12, near the end of

[9] These results are available upon request.

2008-2009 global financial crisis, which is still dwarfed by the Great Depression and other pre-war negative spikes. In essence, the model-implied range of historical variance premium of 5-41, is driven mostly by ambiguity aversion to the uncertain economic recessions and depressions.

[Insert Figure 6 Here.]

For comparison, as seen from Figure 6, the variance premium in Model 1B only has a range of 0.06 to 0.53, while the variance premium in Model 2B has a range of 0.19 to 2.1. In other words, historically speaking, risk components are only about 1 and 4 percents of the total variance premium, while ambiguity component is almost 95 percent of the total variance premium. Such a historical decomposition is largely in line with the earlier result from the model calibration exercise.

[Insert Figure 7 Here.]

The market variance premium is only available since 1990, when the Chicago Board Options Exchange (CBOE) started the new VIX index. Figure 7 compares our consumption-based model-implied variance premium with the observed one of the last 20 years. Our ambiguity aversion model suggests that variance premium is largely flat around 6, with a bump up around 8 during the 1991 recession and a rise to 12 in 2009. The observed variance premium hovered around a high of 19 in early 1990s and shot up to 18 in 2008. However, the elevated variance premium level of 10-25 from 1997 to 2003 cannot be justified by our consumption-based ambiguity aversion model. Or, perhaps, there is nothing to be afraid of fundamentally during the period of 1997-2003.

Our reconstruction of a historical index of variance premium over the past one hundred years (1890-1990), when market variance premium is not observable, is also of interest for other empirical asset pricing exercises.

6. Conclusion

This paper provides an ambiguity-based interpretation of variance premium—the difference between risk-neutral and objective expectations of market return variance. To the first order, variance premium is a compounding effect of both belief distortion regarding unknown regimes and market variance differentials between regimes. The belief distortion represents the difference between Bayesian posterior belief and its uncertainty-adjusted belief. Based on our calibration, its mean value is very small under the standard time-additive CRRA utility. Separating intertemporal substitution from risk aversion as in the Epstein-Zin preference raises the mean belief distortion by about 30 percent. Further separating ambiguity aversion from risk aversion as in Hayashi and Miao (2011) and Ju and Miao (2012) raises the mean belief distortion by about 12 times. The mean market variance differentials between recession and boom regimes in the model with Epstein-Zin utility are more than two times as large as those in the model with time-additive CRRA utility. Introducing ambiguity aversion into Epstein-Zin utility further raises the mean market variance differentials by about 3 times. Overall, ambiguity aversion contributes about 96 percent of the model implied mean variance premium.

We show that our calibrated model with ambiguity aversion can generate a mean variance premium of 8.51 (percentage squared, monthly basis), which is about 80 percent of the empirical estimate of 10.80. In contrast, models only featuring risk aversion or risk aversion and intertemporal substitution can only produce variance premiums of 0.0997 and 0.3150. Our model can simultaneously explain the equity premium, equity volatility, and risk-free rate puzzles. We also show that in univariate regressions, either dividend yield or variance premium is a significant predictor for excess stock returns. However, in joint regressions, these two predictors crowd out each other, reflecting the fact that only one state variable drives return dynamics. This is the main limitation of our model, which is left for future research.

Model implied variance premium series using the consumption growth data from 1890 to

2009 reveals that the major spikes in variance premium mainly capture severe down times like the Great Depression, financial panics, deep recessions, and World War II engagement. After the war, the highest variance premium is reached in 2009, during the peak of global financial crisis.

In sharp contrast with the existing economic models that generate realistic variance premium by relying on either stochastic volatility-of-volatility in consumption or joint jumps in volatility and consumption, our model features a constant consumption growth variance and a regime-shift in consumption growth. Almost all the action comes from the agent's ambiguous belief about the unknown economic regimes and agent's aversion to such an ambiguity. Our model endogenously generates time-varying stock market variance and time-varying variance premium, which is predominantly driven by agent's fear of uncertain economic downside shifts.

References

Abel, Andrew B. (1999), "Risk Premia and Term Premia in General Equilibrium," *Journal of Monetary Economics,* vol. 43, 3–33.

Bansal, Ravi and Amir Yaron (2004), "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, vol. 59, 1481–1509.

Boldrin, Michele, Lawrence J. Christiano, and Jonas D. M. Fisher, 2001, Habit Persistence, Asset Returns and the Business Cycle, *American Economic Review* 91, 149-166.

Bollerslev, Tim, James Marrone, Lai Xu, and Hao Zhou (2011), "Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence," Working Paper.

Bollerslev, Tim, Natalia Sizova, and George Tauchen (2011), "Volatility in Equilibrium: Asymmetries and Dynamics Dependencies," *Review of Finance*, forthcoming.

Bollerslev, Tim, George Tauchen, and Hao Zhou (2009), "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, vol. 22, 4463–4492.

Bollerslev, Tim and Hao Zhou (2007), "Expected Stock Returns and Variance Risk Premia," Finance and Economics Discussion Series 2007-11 , Federal Reserve Board.

Buraschi, Andrea, Fabio Trojani, and Andrea Vedolin (2009), "The Joint Behavior of Credit Spreads, Stock Options and Equity Returns When Investors Disagree," Imperial College London Working Paper.

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 "*Handbook of Macroeconomics*," (edited by Taylor, John B. and Michael Woodford), 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 (1988), "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, vol. 43, 661–676.

Carr, Peter and Liuren Wu (2009), "Variance Risk Premiums," *Review of Financial Studies*, vol. 22, 1311–1341.

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*, vol. 90, 787–805.

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

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

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

Drechsler, Itamar (2011), "Uncertainty, Time-Varying Fear, and Asset Prices," *Journal of Finance,* forthcoming.

Drechsler, Itamar and Amir Yaron (2011), "What's Vol Got to Do With It," *Review of Financial Studies*, vol. 24, 1–45.

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 Tan Wang (1994), "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica* 62, 283-322.

Epstein, Larry G. and Stanley E. Zin (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, vol. 57, 937–969.

Ergin, Haluk and Faruk Gul (2009), "A Theory of Subjective Compound Lotteries," *Journal of Economic Theory*, vol. 144, 899–929.

Fama, Eugene F. and Kenneth R. French (1988), "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, vol. 22, 3–25.

Gilboa, Itzhak and David Schmeidler (1989), "Maxmin Expected Utility with Non-unique Priors," *Journal of Mathematical Economics* 18, 141-153.

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: Valuing Macroeconomic Risk," *American Economic Review*, vol. 97, 1–30.

Hansen, Lars P. and Thomas J. Sargent (2010), "Fragile Beliefs and the Price of Uncertainty," *Quantitative Economics* 1, 129-162.

Hansen, Lars P. and Thomas J. Sargent (2008), *Robustness*, Princeton University Press, Princeton and Oxford.

Hayashi, Takashi and Jianjun Miao (2011), "Intertemporal Substitution and Recursive Smooth Ambiguity Preferences," *Theoretical Economics* 6, 423–472.

Jermann, Urban (1998), "Asset Pricing in Production Economies" *Journal of Monetary Economics* 41, 257-275.

Ju, Nengjiu and Jianjun Miao (2012), "Ambiguity, Learning, and Asset Returns," *Econometrica* 80, 559-591.

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 (2009), “Recursive Smooth Ambiguity Preferences,” *Journal of Economic Theory* 144, 930-976.

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.

Londono, Juan-Miguel (2010), “The Variance Risk Premium around the World,” Working Paper, Tilburg University Department of Finance, The Netherlands.

Mehra, Rajnish and Edward C. Prescott (1985), “The Equity Premium: A Puzzle,” *Journal of Monetary Economics*, vol. 15, 145–161.

Mueller, Philippe, Andrea Vedolin, and Hao Zhou (2011), “Short-Run Bond Risk Premia,” Working Paper, Federal Reserve Board.

Segal, Uzi (1990), “Two-Stage Lotteries Without the Reduction Axiom,” *Econometrica*, vol. 58, 349–377.

Seo, Kyoungwon (2009), “Ambiguity and Second-Order Belief,” *Econometrica*, vol. 77, 1575–1605.

Shiller, Robert J. (1981), “Do Stock Prices Move too Much to be Justified by Subsequent Changes in Dividends?” *American Economic Review* 71, 421-436.

Todorov, Viktor (2010), “Variance Risk Premium Dynamics: The Role of Jumps,” *Review of Financial Studies* 23, 345-383.

Veronesi, Pietro (1999), “Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model,” *Review of Financial Studies* 12, 975-1007.

Wang, Hao, Hao Zhou, and Yi Zhou (2011), “Credit Default Swap Spreads and Variance Risk Premia,” Working Paper, Federal Reserve Board.

Zhou, Guofu and Yingzi Zhu (2011), “A Long-Run Risks Model with Long- and Short-Run Volatilities: Explaining Predictability and Volatility Risk Premium,” Working Paper, Olin School of Business, Washington University.

Zhou, Hao (2010), “Variance Risk Premia, Asset Predictability Puzzles, and Macroeconomic Uncertainty,” Working Paper, Federal Reserve Board.

Table 1 Parameter Values

The parameters $(\lambda_{11}, \lambda_{22}, \kappa_1, \kappa_2, \sigma)$ are estimated by the maximum likelihood method using the annual per capita US consumption data covering the period 1890-2009. The other parameter values are calibrated at the annual frequency.

Parameter	Value
Risk aversion	$\gamma = 2$
Elasticity of intertemporal substitution	$1/\rho = 1.5$
Ambiguity aversion	$\eta = 10.379$
Discount factor	$\beta = 0.9838$
Expected consumption growth in booms	$\kappa_1 = 0.0242$
Expected consumption growth in recessions	$\kappa_2 = -0.0560$
Volatility of consumption growth	$\sigma = 0.0317$
Transition probabilities	$\lambda_{11} = 0.9799$
	$\lambda_{22} = 0.5432$
Expected dividend growth	$g_d = -0.0362$
Leverage ratio	$\zeta = 2.74$
Volatility of dividend growth	$\sigma_d = 0.0629$

Table 2 Equity Premium and Variance Premium

This table reports various moments of the equity premium, the risk-free rate, and the variance premium in the data and predicted by various models. The numbers in the column labeled "Drechsler" are taken from Table I of Drechsler (2011). Note that the average and standard deviation of variance premium are converted to monthly values by multiplying $10^4/12$. R_f : mean risk-free rate; $\sigma(R_f)$: volatility of the risk-free rate; μ_{eq} : mean equity premium; $\sigma(\mu_{eq})$: equity premium volatility; VP : mean variance premium; $\sigma(VP)$: volatility of variance premium; $AC(1)$: first-order autocorrelation of variance premium.

	Data	Drechsler	Model 1A	Model 1B	Model 2A	Model 2B	Full Model
R_f	0.0191		0.0572	0.0572	0.0286	0.0286	0.0191
$\sigma(R_f)$	0.0578		0.0179	0.0139	0.0067	0.0053	0.0110
μ_{eq}	0.0574		0.0073	0.0074	0.0091	0.0093	0.0575
$\sigma(\mu_{eq})$	0.1880		0.1232	0.1240	0.1290	0.1291	0.1726
VP	10.80	10.55	0.0655	0.0997	0.3158	0.3150	8.5136
$\sigma(VP)$	23.36	8.47	0.0957	0.0913	0.4370	0.3581	7.5323
$AC(1)$	0.17	0.61	0.5163	0.4682	0.5163	0.4750	0.3778

Table 3 Variance Premium Decomposition

This table reports the decomposition of the mean variance premium for Model 1B, Model 2B and our full model as in Section 3. In the table, $F1 \equiv \mu_t - \widehat{\mu}_t$, $F2 \equiv \mathbb{E}_t^2[\Sigma_{t+1}] - \mathbb{E}_t^1[\Sigma_{t+1}]$, and $F3 \equiv \widehat{\mu}_t \frac{Cov_t^1[\Sigma_{t+1}, M_{t+1}]}{\mathbb{E}_t^1[M_{t+1}]} + (1 - \widehat{\mu}_t) \frac{Cov_t^2[\Sigma_{t+1}, M_{t+1}]}{\mathbb{E}_t^2[M_{t+1}]}$. The mean variance premium $VP = \mathbb{E}[F1 \times F2] + \mathbb{E}[F3]$. Note that the unconditional averages of these terms are reported.

	Model 1B	Model 2B	Full Model
$\mathbb{E}[F1]$	0.006	0.008	0.1105
$\mathbb{E}[F2]$	7.274	16.53	61.486
$\mathbb{E}[F1 \times F2]$	0.045	0.156	7.306
$\mathbb{E}[F3]$	0.054	0.159	1.205
VP	0.099	0.315	8.513

Table 4 Predictability of Excess Returns

This table reports the results of three sets of predictive regressions using model simulated data. The slopes and R^2's are obtained from an OLS regression of the excess returns on the corresponding predictor variable(s) at various horizons (i.e., 1, 2, 3, and 5 years). The reported numbers are the mean values of 10000 Monte Carlo simulations, each consisting of 120 data points. Reported t-statistics are Newey-West (HAC) corrected.

	Dividend Yield		Variance Premium		
Horizon	Slope	t-statistics	Slope	t-statistics	R^2
1	0.82	2.64			0.12
2	1.12	3.40			0.14
3	1.24	3.76			0.14
5	1.33	3.72			0.12
1			6.35	2.41	0.11
2			8.67	2.89	0.13
3			9.73	3.11	0.12
5			10.53	3.11	0.11
1	0.78	0.49	0.81	0.07	0.14
2	0.89	0.47	1.37	0.11	0.15
3	0.99	0.45	1.37	0.09	0.15
5	1.13	0.41	1.18	0.07	0.13

Figure 1: Variance Premium and GDP Growth

The top panel plots variance premium or the implied-expected variance difference for the S&P500 market index from January 1990 to December 2011. The variance premium is based on the realized variance forecast from lagged implied and realized variances. The bottom panel plots the GDP growth rates (thin blue line) together with the variance premium (thick red line) from 1990Q1 to 2011Q4. Both of the series are standardized to have mean zero and variance one. The shaded areas represent NBER recessions.

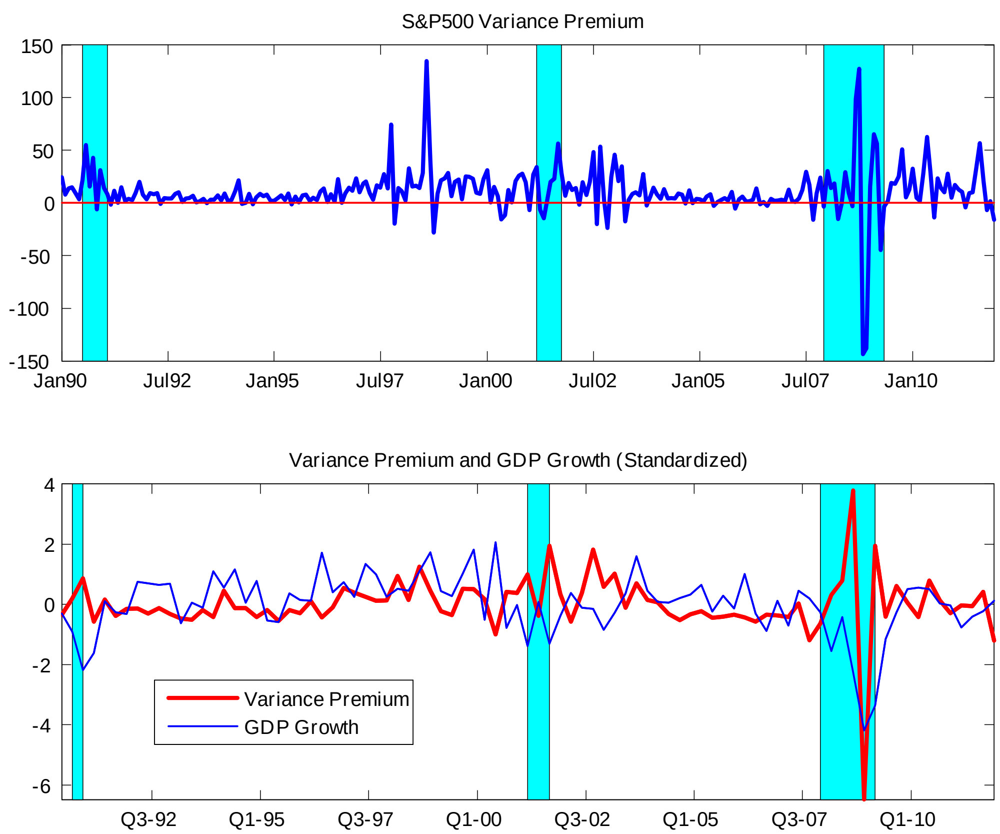

Figure 2: Determinants of Variance Premium

This figure plots the conditional variance premium and its decomposition in terms of three factors. The (blue) solid lines correspond to the full model. The (red) dashed lines correspond to Model 2B with Epstein-Zin preference and Bayesian learning (i.e., $\gamma = \eta \neq \rho$). The (black) dotted lines correspond to Model 1B with time-additive CRRA utility and Bayesian learning (i.e., $\gamma = \eta = \rho$). Except for the panel for $\mu_t - \hat{\mu}_t$, the vertical axes are in percentage squared divided by 12.

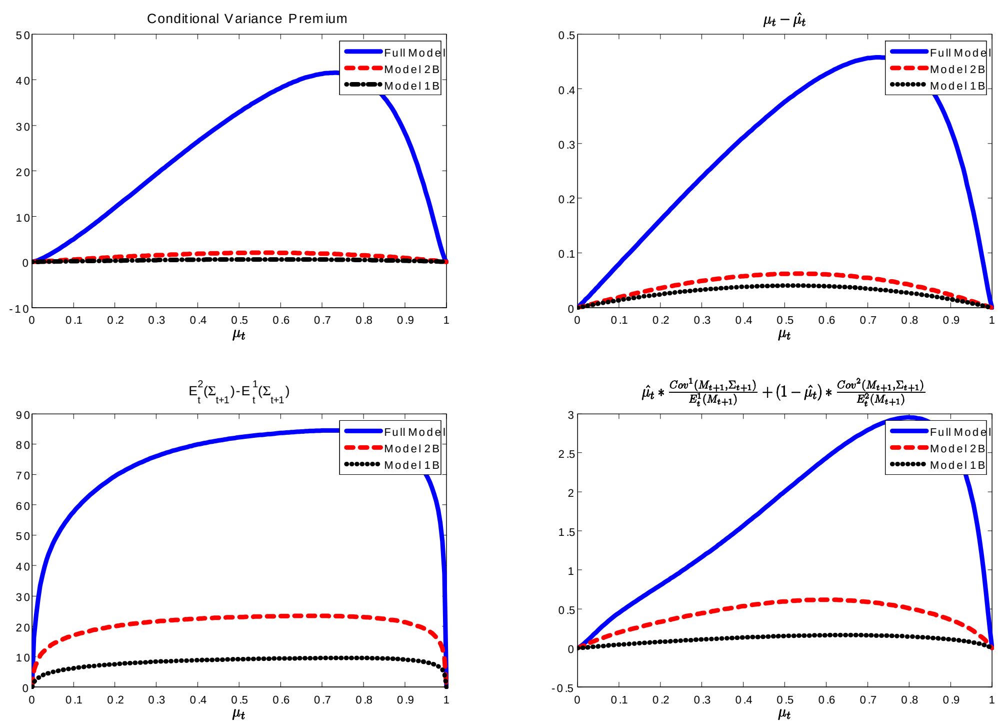

Figure 3: Price-Dividend Ratio and Conditional Variance

This figure plots the price dividend ratio (top panel) and conditional variance (bottom panel) as functions of the Bayesian posterior probabilities of the high-growth state. The (blue) solid lines correspond to the full model. The (red) dashed lines correspond to Model 2B with Epstein-Zin preference and Bayesian learning (i.e., $\gamma = \eta \neq \rho$). The (black) dotted lines correspond to Model 1B with time-additive CRRA utility and Bayesian learning (i.e., $\gamma = \eta = \rho$).

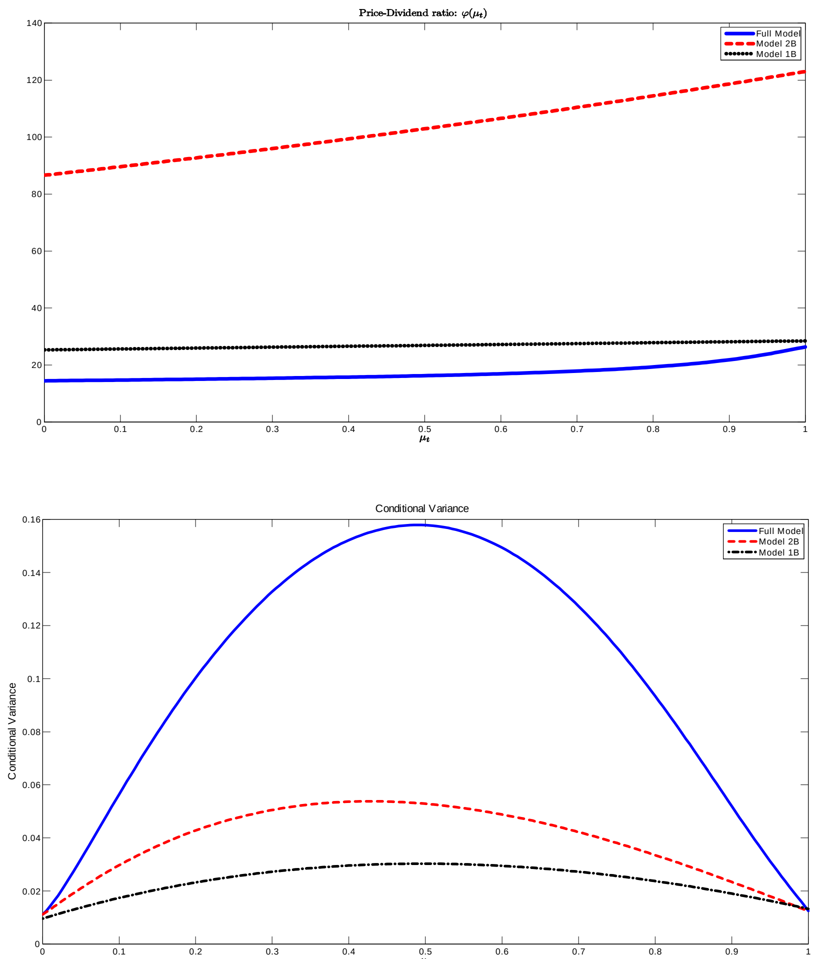

Figure 4: Consumption Growth and Posterior Beliefs (1890-2009)

This figure plots historical annual per-capita consumption growth rate (top panel) and model-implied posterior belief (bottom panel) during the sample period from 1890 to 2009. The shaded areas represent NBER recessions.

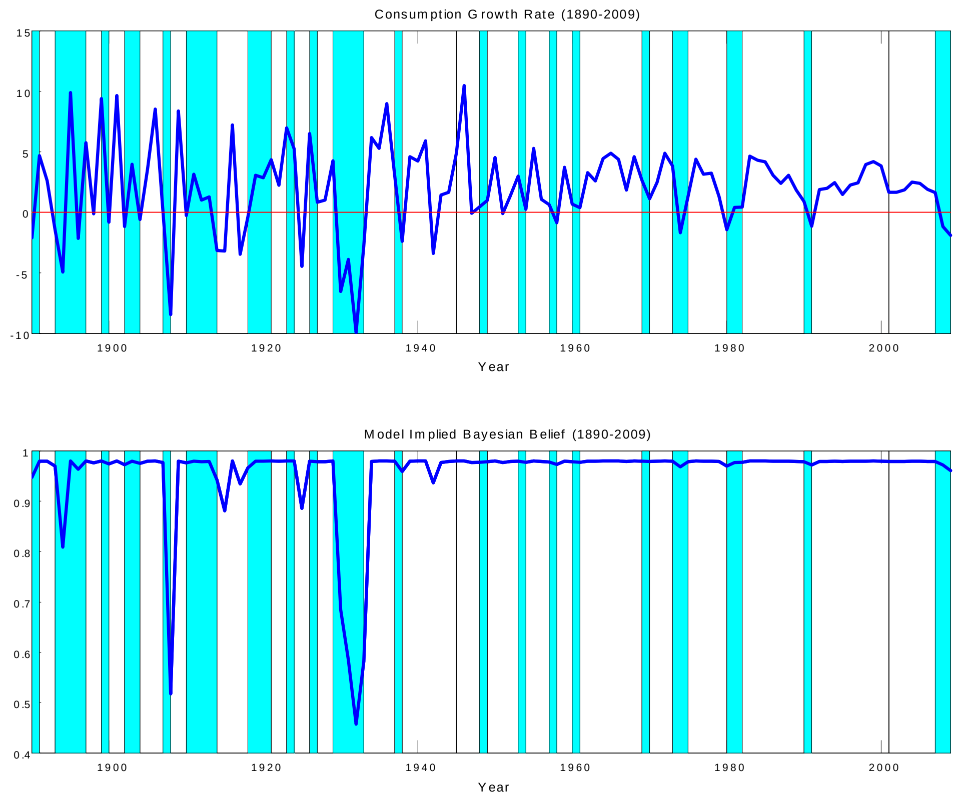

Figure 5: Ambiguity Aversion Driven Variance Premium (1890–2009)

This figure plots model-implied variance premium during the sample period from 1890 to 2009. The shaded areas represent NBER recessions.

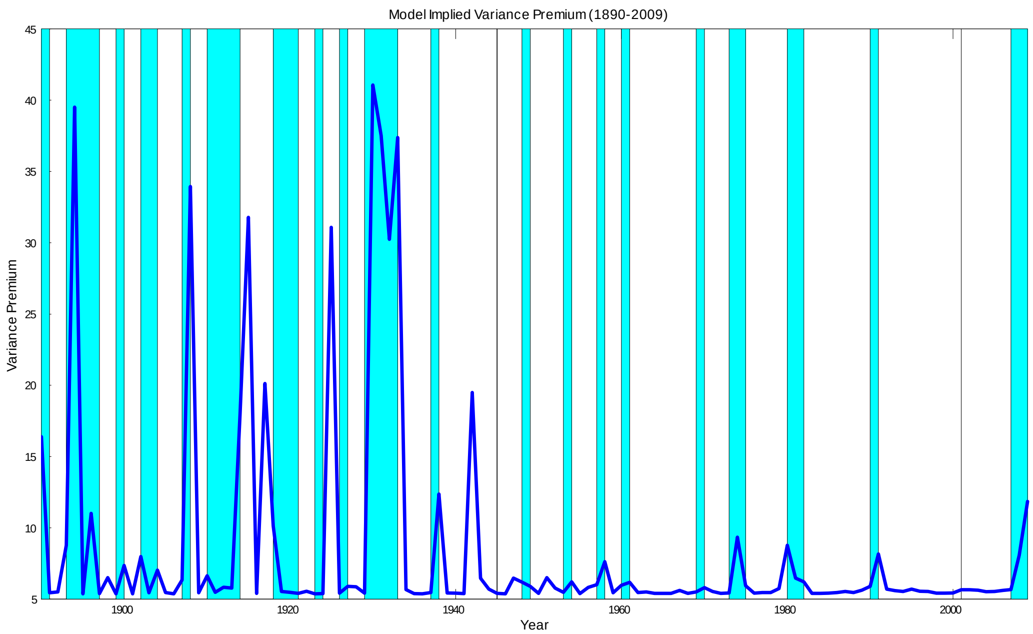

Figure 6: Model-Implied Variance Premium Decomposition (1890-2009)

This figure reports the decomposition of model-implied variance premium during the sample period from 1890 to 2009. The top panel with (black) dotted line plots the variance premium implied by Model 1B with time-additive CRRA utility and Bayesian learning (i.e., $\gamma = \eta = \rho$), labeled as "Standard VP." The middle panel with (red) dashed line plots the difference between the variance premium implied by Model 2B with Epstein-Zin preferences and Bayesian learning (i.e., $\gamma = \eta \neq \rho$) with that implied by Model 1B, labeled as "VP component due to Epstein-Zin utility." The bottom panel with (blue) solid line plots the difference between the variance premium implied by the full model with ambiguity aversion with that implied by Model 2B, labeled as "VP Component due to Ambiguity Aversion."

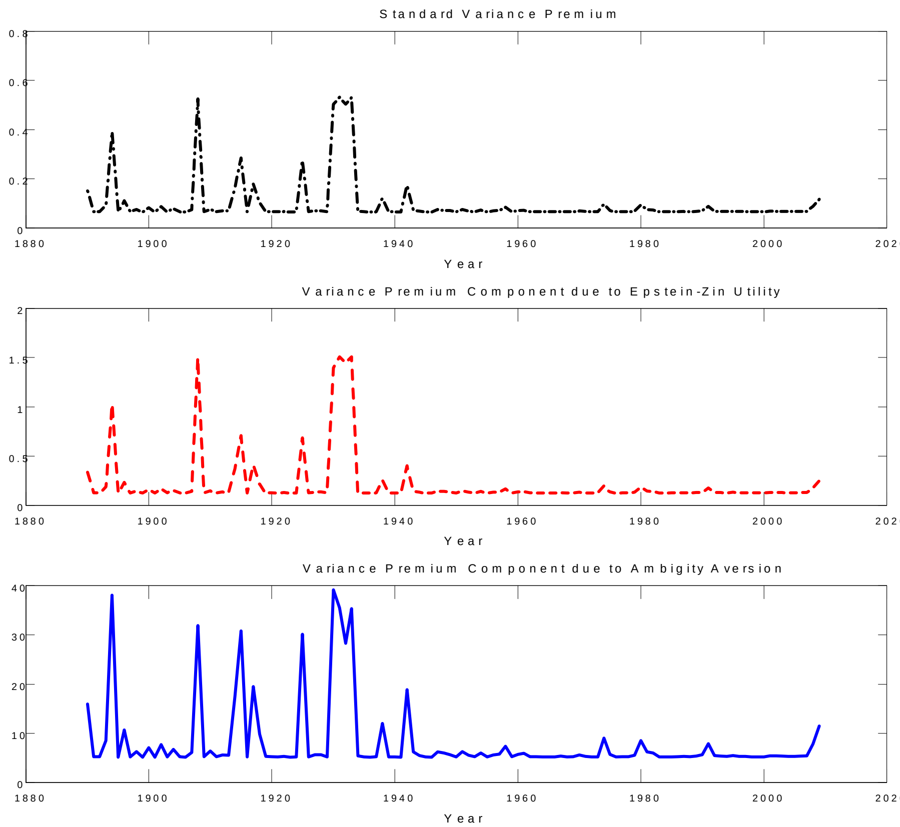

Figure 7: Model-Implied and Observed Variance Premium (1990-2009)

This figure plots the model-implied variance premium together with actual variance premium in the recent period between 1990 and 2009. The scatter plot corresponds to the actual variance premium as a 12 month average observation in the data. The (blue) solid lines correspond to the variance premium implied by the full model. The (red) dashed lines correspond to Model 2B with Epstein-Zin preference and Bayesian learning (i.e., $\gamma = \eta \neq \rho$). The (black) dotted lines correspond to Model 1B with time-additive CRRA utility and Bayesian learning (i.e., $\gamma = \eta = \rho$).