Ambiguity Aversion and the Variance Premium

Jianjun Miao
Department of Economics, Boston University
270 Bay State Road, Boston MA 02215, USA
ICFS, Southwestern University of Finance and Economics, China
CEMA, Central University of Finance and Economics, China
miaoj@bu.edu

Bin Wei
Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE
Atlanta, GA 30309, USA
bin.wei@atl.frb.org

Hao Zhou*
PBC School of Finance, Tsinghua University
43 Chengfu Road, Haidian District
Beijing 100083, China
zhouh@pbcfs.tsinghua.edu.cn

Published 8 October 2018

This paper offers an ambiguity-based interpretation of the 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 calibrated model can match the variance premium, the equity premium, and the risk-free rate in the data. We find that about 97% of the mean–variance premium can be attributed to ambiguity aversion. A three-way separation among ambiguity aversion, risk aversion, and intertemporal substitution, permitted by the smooth ambiguity preferences, plays a key role in our model’s quantitative performance.

Keywords: Ambiguity aversion; learning; variance premium; regime-shift; belief distortion.

JEL Classifications: D81, E44, G12, G13

*Corresponding author.
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 found that such a standard model typically predicts a negligible premium for higher moments such as the 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 97% of the average variance premium whereas risk can only explain about 3% 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 et al. (2009). The Hayashi–Ju–Miao model also includes the Epstein and Zin (1989) 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 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.

¹Bollerslev et al. (2009), Drechsler and Yaron (2011), Londono (2010), and 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 et al. (2011), Buraschi et al. (2009), and Wang et al. (2011). See Han and Zhou (2011) for the cross-sectional relationship between a stock’s expected return and its variance premium.

²Also see Hansen and Sargent (2010) for a similar result based on robust control.
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; 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 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 sensitivity to ambiguity or model uncertainty (Segal, 1990; Klibanoff et al., 2005; Hansen, 2007). 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 therefore raises the third component as well.

Our model with ambiguity aversion generates a mean–variance premium of 8.47 (in percentage squared, monthly basis), which is quite close to our empirical estimate of 9.46 in the data and well within the typical range of 5–19 from existing empirical studies. 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 0.25, respectively. Incorporating Bayesian learning in these models changes the mean–variance premium to 0.09 and 0.25, respectively. Thus risk aversion and intertemporal substitution contribute about 1% and 2% to the model-implied variance premium, respectively, while ambiguity aversion contributes to about 97%.

Note that these results are achieved under ambiguity aversion without introducing stochastic volatility or volatility jumps in consumption growth.

---

3The range of 5–19 of existing estimates for variance premium depends on whether the index or futures are used as underlying, whether the recent crisis period is included, and whether expected or realized variance is used. See, e.g., Carr and Wu (2009), Bollerslev et al. (2009), Drechsler and Yaron (2011), Zhou and Zhu (2011), Bollerslev et al. (2012), and Drechsler (2013).
In addition, our calibration targets to match the mean equity premium and the mean risk-free rate, with the variance premium only as an output. We emphasize that the estimation of the regime-switching consumption process using the century-long data from 1890 to 2015 is important for us to identify a recession state with a large negative consumption growth rate, which is like a disaster risk (Barro, 2006; Gabaix, 2011). We show that our model performs much worse when the calibration is based on the estimation using the postwar data. For the postwar period, the estimated expected consumption growth rate in the recession state is still positive. Thus there is not much severe downside uncertainty so that ambiguity aversion does not matter much.

The existing approaches to generating realistic variance premium dynamics typically involve Epstein–Zin preferences combined with stochastic volatility-of-volatility in consumption (Bollerslev et al., 2009) or joint jumps in consumption volatility and growth (Todorov, 2010; Drechsler and Yaron, 2011). Like our paper, Drechsler (2013) explores the implications of ambiguity aversion for the variance premium. His model is more complicated than ours in that he incorporates stochastic volatility and jumps into the expected growth and growth volatility processes. Moreover, he does not consider learning. All the preceding papers can capture the size and predictive power of the variance premium in the data.

Our model is parsimonious and can generate empirically plausible variance premium without relying on exogenously specified complex consumption or dividend dynamics. Our approach highlights the important role of ambiguity aversion and learning in generating large variance premium dynamics. Since our model has only one state variable (the posterior belief about the boom state), our model generates a weaker predictability result than the existing approaches.

Our model complements Drechsler’s (2013). One key difference 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 et al., 2005; Hansen and Sargent, 2010 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

2. Stylized Facts of Variance Premium

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

$$VP_t \equiv E^Q_t(\Sigma_{t+1}) - E_t(\Sigma_{t+1}),$$

where $Q$ represents the risk-neutral measure. We use the currently available data from January 1990 to December 2015 to measure the variance premium by closely following Drechsler (2013). The availability of the Chicago Board Options Exchange (CBOE) VIX index makes it straightforward to measure the risk-neutral expectation of stock market return variances. 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 30-day maturity.5 We compute $E^Q_t(\Sigma_{t+1})$ by squaring the CBOE VIX index and then dividing it by 12 to get a monthly quantity. We estimate $E_t(\Sigma_{t+1})$ as the conditional forecast of the realized variance in the following month. Following Drechsler (2013), we measure the realized variance of the returns on the S&P 500 index by summing up the squared five-minute log returns on the S&P 500 futures and on the S&P 500 index over the whole month. We obtain the high frequency data used to construct these realized variance measures from TICKDATA. Next we use a simple time-series model to obtain the one-step-ahead forecasts of the realized variance in the following month. Specifically, we regress the futures realized variance on the value of the squared VIX and on a lagged cash realized variance. We report the regression results in Panel A of Table 1.

Following Drechsler (2013), we truncate the estimates of the variance premium to zero whenever there are negative values since theoretically the physical measure of variance should be less than the risk-neutral one. For 305 out of 312 months in our sample, no truncation is needed as the estimates of the variance premium in these months are positive. All our results remain little changed if the estimates of the variance premium are used without

---


5See, e.g., Carr and Wu (2009) for the definition of model-free implied variance.
Panel B of Table 1 presents summary statistics for the squared VIX, the futures realized variance, the cash realized variance, estimates of the expected realized variance, and the variance premium. The results are similar to those reported in Drechsler (2013), though we have a larger sample.

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.

Figure 1’s top panel plots the monthly time series of our estimates of the 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–Long-Term Capital Management (LTCM) crisis. There are also huge run-ups of the variance premium around
May 2010 and August 2011, when heightened Greece sovereign default risk was threatening the Euro area financial stability. The sample mean of the variance premium is \(9.46\) (in percentage squared, monthly basis), with a standard deviation of \(8.81\). Nevertheless, the variance premium is not a very persistent process with an AR(1) coefficient of 0.53 at monthly frequency. Our estimates are broadly consistent with the existing estimates in the literature. For example, using data of five-minute log returns on the S&P 500 futures from January 1990 to December 2009, Drechsler (2013) 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 an almost zero variance premium (Bollerslev et al., 2009).

Recent empirical evidence has also suggested that the stock market return is predictable by the variance premium over a few quarters’ horizon.

Fig. 1. Variance premium and GDP growth.

Notes: The top panel plots our estimates of the variance premium based on the data from January 1990 to December 2015, expressed in percentage squared. The bottom panel plots the GDP growth rates (dashed line) and the variance premium (solid line) from 1990:Q1 to 2015:Q4, expressed in percentage. Both of the series are standardized to have a mean of zero and a variance of one. The shaded areas represent NBER recessions.
(Bollerslev et al., 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). 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 the 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 Fig. 1 plots variance premium together with the quarterly growth rate in GDP. As seen in the figure, there is a tendency for the variance premium to rise one to two quarters before a decline in GDP, while it typically narrows ahead of an increase in GDP. Indeed, the sample correlation equals $-0.05$ between current variance premium and two-quarter-ahead GDP (as first reported in Bollerslev and Zhou (2007)). In other words, the 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 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:

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

(2)

The regime-switching consumption process has been used widely in the asset pricing literature (see, e.g., Veronesi, 1999; Cecchetti et al., 2000). One may view this process as a nonlinear version of the long-run risk process studied by Bansal and Yaron (2004).
where $\varepsilon_t$ is an independently and identically distributed (iid) standard normal random variable, and $z_{t+1}$ follows a Markov chain that takes values 1 or 2 with transition matrix $(\lambda_{ij})$ 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

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

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 consumption growth as in Abel (1999). This parameter and the parameter $\sigma_d$ 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 $t$:

$$s^t = \{C_0, D_0, C_1, D_1, \ldots, C_t, D_t\}.$$  

In addition, he knows the parameters of the model (e.g., $\zeta$, $g_d$, $\sigma$, and $\sigma_d$). 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 $\mu_t = \Pr(z_{t+1} = 1 | s^t)$. The prior belief $\mu_0$ is given. By Bayes’ Rule, we can derive

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

where $f(y, i) = \frac{1}{\sqrt{2\pi}\sigma}\exp[-(y - \kappa_i)^2/(2\sigma^2)]$ is the density function of the normal distribution with mean $\kappa_i$ and variance $\sigma^2$. 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:

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

$$\mathcal{R}_t(V_{t+1}(C)) = \{\mathbb{E}_{\mu_t} (\mathbb{E}_{\pi_{zt+1}} [V_{t+1}^{1-\gamma}(C)]^{1-\gamma})^{1-\gamma}\},$$

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

\[
V_t(C) = [(1 - \beta)C_t^{1-\rho} + \beta V_{t+1}(C)^{1-\rho}]^{1/\gamma}.
\]

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

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

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 \( \gamma \) 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, \( \{\mathbb{E}_{\pi_{z,t}}[V_{t+1}(C)]\}^{1/\gamma} \). If the agent is ambiguity averse to the variation of economic regimes, he is averse to the variation of \( \{\mathbb{E}_{\pi_{z,t}}[V_{t+1}(C)]\}^{1/\gamma} \) across different regimes \( z_{t+1} \). Thus he evaluates the ex ante continuation value using a concave function with a curvature \( \eta \) so that he enjoys the certainty equivalent value \( R_t(V_{t+1}(C)) \). Only when \( \eta > \gamma \), the certainty equivalent to an ambiguity-sensitive agent is less than that to an agent with expected utility, i.e.,

\[
R_t(V_{t+1}(C)) < \{\mathbb{E}_t [V_{t+1}^{1-\gamma}(C)]\}^{1/\gamma},
\]

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

To understand the asset pricing implications of the aforementioned model, one needs only 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

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

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

$$\mathbb{E}_t [M_{t+1} R_{k,t+1}] = 1. \quad (8)$$

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 Eq. (8).

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

$$P_{c,t} = \varphi(\mu_t) D_t, \quad (9)$$

$$V_t = G(\mu_t) C_t. \quad (10)$$
By definition, we can write the equity return as:

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

This equation implies that the state beliefs drive changes in the price–dividend ratio, and hence dynamics of equity returns. In Sec. 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 returns by \( \varSigma_t \equiv \text{Var}_t[R_{e,t+1}] \). Variance premium is defined as

\[ \text{VP}_t = \mathbb{E}_t^Q(\varSigma_{t+1}) - \mathbb{E}_t(\varSigma_{t+1}) = \frac{\mathbb{E}_t[\varSigma_{t+1}M_{t+1}]}{\mathbb{E}_t[M_{t+1}]} - \mathbb{E}_t[\varSigma_{t+1}], \] (12)

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

\[ M_{t+1} = M_{t+1}^{EZ} M_{z,t}^A, \quad \text{for } z_{t+1} = z \in \{1, 2\}, \] (13)

where \( M_{t+1}^{EZ} \) is the Epstein–Zin pricing kernel defined by

\[ M_{t+1}^{EZ} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\rho} \left( \frac{V_{t+1}}{R_t(V_{t+1})} \right)^{\rho-\gamma}, \] (14)

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

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

Here, \( \mathbb{E}_i^i \) 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

\[ \frac{\mathbb{E}_t[\varSigma_{t+1}M_{t+1}]}{\mathbb{E}_t[M_{t+1}]} = \frac{\mu_i \mathbb{E}_t^i[\varSigma_{t+1}M_{t+1}^{EZ}] M_{1,t}^A + (1 - \mu_i) \mathbb{E}_t^i[\varSigma_{t+1}M_{t+1}^{EZ}] M_{2,t}^A}{\mu_i \mathbb{E}_t^i[M_{t+1}^{EZ}] M_{1,t}^A + (1 - \mu_i) \mathbb{E}_t^i[M_{t+1}^{EZ}] M_{2,t}^A}. \]

Using the fact that

\[ \mathbb{E}_t^i[\varSigma_{t+1}M_{t+1}^{EZ}] = \mathbb{E}_t^i[\varSigma_{t+1}] \mathbb{E}_t^i[M_{t+1}^{EZ}] + \text{Cov}_t^i(\varSigma_{t+1}, M_{t+1}^{EZ}), \]

\[ \text{Cov}_t^i(\varSigma_{t+1}, M_{t+1}^{EZ}) \]

Ambiguity Aversion and the Variance Premium
we can compute

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

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

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

where $\text{Cov}_t^i$ 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 boom state by

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

We can then rewrite the variance premium as

$$VP_t = (\mu_t - \hat{\mu}_t) (\mathbb{E}_t^2 [\Sigma_{t+1}] - \mathbb{E}_t^1 [\Sigma_{t+1}])$$

$$+ \frac{\hat{\mu}_t}{\mathbb{E}_t^1 [M_{t+1}^{EZ}]} \text{Cov}_t^1 (\Sigma_{t+1}, M_{t+1}^{EZ}) + \frac{1 - \hat{\mu}_t}{\mathbb{E}_t^2 [M_{t+1}^{EZ}]} \text{Cov}_t^2 (\Sigma_{t+1}, M_{t+1}^{EZ}).$$

Note that by Eq. (13), we can replace $M_{t+1}^{EZ}$ by $M_{t+1}$ in Eq. (16). This equation reveals that variance premium is determined by three components: (i) the expression $(\mu_t - \hat{\mu}_t)$, (ii) the expression $(\mathbb{E}_t^2 [\Sigma_{t+1}] - \mathbb{E}_t^1 [\Sigma_{t+1}])$, 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 Eq. (16).

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 countercyclicity 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 Sec. 5 shows that this component is quantitatively small for reasonable parameter values.

Now we examine the first component $(\mu_t - \hat{\mu}_t)$, 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

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

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

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

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

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

\[
\hat{\mu}_t^{\text{EZ}} \approx \frac{\mu_t e^{-\gamma \kappa_1} \mathbb{E}_t^1 \left[ G_{t+1}^{\rho - \gamma} \right]}{\mu_t e^{-\gamma \kappa_1} \mathbb{E}_t^1 \left[ G_{t+1}^{\rho - \gamma} \right] + (1 - \mu_t) e^{-\gamma \kappa_2} \mathbb{E}_t^2 \left[ G_{t+1}^{\rho - \gamma} \right]} < \hat{\mu}_t^{\text{CRRA}} \quad \text{when } \rho < \gamma.
\]

This equation shows that the component \( (\mu_t - \hat{\mu}_t^{\text{EZ}}) \) in the Epstein–Zin model is larger than \( (\mu_t - \hat{\mu}_t^{\text{CRRA}}) \) 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 Eq. (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 boom state in that

\[
\hat{\mu}_t < \hat{\mu}_t^{\text{EZ}} \quad \text{when } \eta > \gamma,
\]
where $\hat{\mu}_t$ is given by (15). Therefore, holding everything else constant, an ambiguity averse agent demands a higher variance premium than an ambiguity-neutral agent 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 and run Monte Carlo simulations to compute model moments as in Ju and Miao (2012). We first calibrate the model at an annual frequency in Sec. 5.1. We then study properties of unconditional and conditional moments of variance premium generated by our model in Secs. 5.2–5.4. 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, respectively. Models 1B and 2B are the corresponding models with partial information and Bayesian learning, respectively. The last two models are special cases of our full model with $\eta = \gamma = \rho$ and $\eta = \gamma \neq \rho$, respectively. Finally, in Sec. 5.5, we show that alternative calibrations by maximizing the power of CRRA and Epstein–Zin preferences fall short of simultaneously matching the equity premium and variance premium.

5.1. Baseline calibration

We calibrate the model to the sample period from 1890 to 2015 based on Robert Shiller’s data of the US consumption growth, dividends growth, stock returns, and risk-free rates. Most Robert Shiller’s data have been updated to the year of 2015; however, the consumption growth data are only available up to 2009 and the one-year interest rate data are updated until 2011. To update the consumption growth data to the year of 2015, we use the NIPA data of personal consumption expenditure in billions of chained 2009 dollars and convert it in terms of chained 2005 dollars to be consistent with the consumption growth data in earlier years. Robert Shiller’s one-year interest rate data since 1997 are based on six-month certificate of deposit rate from the Federal Reserve, which, however, is discontinued in June 2013. To update the one-year interest rate data to the year of 2015, we obtain “National Rate on Non-Jumbo Deposits (less than $100,000): 6 Month CD” and “National Rate

\[ \text{We run 10,000 simulations, each of which consisting of 126 data points. Increasing this number does not change our results significantly.} \]

\[ \text{The data is downloaded from Professor Robert Shiller’s website: http://www.econ.yale.edu/~shiller/data/ie_data.xls} \]
on Jumbo Deposits (greater or equal to $100,000): 6 Month CD” from FRED, the Federal Reserve Bank of St. Louis. Following Shiller’s method, we convert the January and July rates into an annual yield, and then take the average of annual one-year interest rates obtained from Non-Jumbo and Jumbo deposits. Table 2 presents summary statistics for the data and the variance premium we estimated in Sec. 2.

Using the consumption data, we first estimate five parameters \((\lambda_{11}, \lambda_{22}, \kappa_1, \kappa_2, \sigma)\) in the regime-switching process of consumption growth by the maximum likelihood method. Panel A of Table 3 reports the estimation result. We find that the boom state is highly persistent, with consumption growth in this state being 2.35%. The economy spends most of the time in this state with the unconditional probability of \((1 - \lambda_{22})/(2 - \lambda_{11} - \lambda_{22}) = 0.96\). The recession state is moderately persistent, but very bad, with consumption growth at \(-6.1\%\). The boom and recession states have an expected duration of 45.5 and 1.9 years. The long-run average rate of consumption growth is 2.01%.

The feature of a small probability of large negative consumption growth \((-6.1\%)\) is similar to a disaster risk. This can be estimated because there are three years with consumption growth below \(-6\%\) between 1890 and 2015; that is, \(-9.93\%\) in 1932, \(-8.45\%\) in 1908, and \(-6.56\%\) in 1930. In contrast, the worst year in the postwar period 1946–2015 is 2009 with a decline in consumption of 1.9%. Estimating a regime-switching process of consumption growth for the postwar period leads to a very different result as reported in Panel B of Table 3. We find that consumption growth is 2.94% in the boom

<table>
<thead>
<tr>
<th>Table 2. Summary statistics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>cons</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Std.-Dev.</td>
</tr>
<tr>
<td>Skewness</td>
</tr>
<tr>
<td>Kurtosis</td>
</tr>
<tr>
<td>AC(1)</td>
</tr>
</tbody>
</table>

Notes: This table presents summary statistics for log consumption growth rate (column “cons”), log dividend growth rate (column “div”), log price-dividend ratio (column “pd”), risk-free rate (column “rf”), equity premium (column “EP”), and variance premium (column “VP”). The last column is taken from Table 1 based on monthly data between 1990 and 2015. The other columns are based on annual data between 1890 and 2015. The mean, median and standard deviation (Std.-Dev) for cons, div, rf, and EP are expressed in annualized percentages. AC(1) represents first-order autocorrelation.
state, and 0.25% in the recession state, with an expected duration of 4.5 and 2.2 years, respectively. In Sec. 6.1, we will show that our model performs worse when it is calibrated based on these estimates and thus disaster risk is important for our model success.

Next we calibrate the other parameters \((\gamma, \rho, \eta, \beta, \zeta, g_d, \sigma_d)\) in the model listed in Panel C of Table 3 based on the full sample 1890–2015. The leverage ratio \(\zeta\) is set to 2.74 following Abel (1999) and \(g_d\) is chosen as \(-0.0349\) 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 11.5%, we choose \(\sigma_d = 0.0672\) 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 \(\beta\) and the ambiguity aversion parameter \(\eta\) to match the mean risk-free rate of 1.76% and the mean equity premium of 6.02% from the data reported in Table 2. We obtain \(\beta = 0.9838\) and \(\eta = 10.948\).

In sum, the calibrated parameter values listed in Panel C of Table 3 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 $\eta$ only, as 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 et al. (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 $1.7\%$ of the expected prize value when one sets $\gamma = 2$ and the prize–wealth ratio of $1\%$. Similarly, we can compute that the ambiguity premium is equal to $2.22\%$ of the expected prize for $\eta = 10.948$. Camerer (1999) reports that the ambiguity premium is typically in the order of $10\%$ to $20\%$ 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 our targets of the mean risk-free rate and the mean equity premium. Table 4 shows that the model-implied equity premium volatility is equal to $16.41\%$, which is quite close to $18.52\%$ in the data. As Shiller (1981) and Campbell (1999) point out, it is challenging for the

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model 1A</th>
<th>Model 1B</th>
<th>Model 2A</th>
<th>Model 2B</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_f$</td>
<td>1.76</td>
<td>5.58</td>
<td>5.57</td>
<td>2.83</td>
<td>2.82</td>
<td>1.75</td>
</tr>
<tr>
<td>$\sigma(R_f)$</td>
<td>5.68</td>
<td>1.60</td>
<td>1.31</td>
<td>0.60</td>
<td>0.49</td>
<td>1.05</td>
</tr>
<tr>
<td>$\mu_{eq}$</td>
<td>6.02</td>
<td>0.64</td>
<td>0.65</td>
<td>0.79</td>
<td>0.80</td>
<td>6.04</td>
</tr>
<tr>
<td>$\sigma(\mu_{eq})$</td>
<td>18.52</td>
<td>12.21</td>
<td>12.27</td>
<td>12.60</td>
<td>12.62</td>
<td>16.41</td>
</tr>
<tr>
<td>VP</td>
<td>9.46</td>
<td>0.07</td>
<td>0.09</td>
<td>0.25</td>
<td>0.25</td>
<td>8.47</td>
</tr>
<tr>
<td>$\sigma(VP)$</td>
<td>8.81</td>
<td>0.09</td>
<td>0.09</td>
<td>0.31</td>
<td>0.27</td>
<td>6.25</td>
</tr>
<tr>
<td>AC(1)</td>
<td>0.53</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Notes: The variables $R_f$, $\sigma(R_f)$, $\mu_{eq}$, $\sigma(\mu_{eq})$, VP, $\sigma(VP)$, and AC(1) denote the mean risk-free rate, the volatility of the risk-free rate, the mean equity premium, the equity premium volatility, the mean–variance premium, the volatility of the variance premium, and the first-order autocorrelation of the variance premium, respectively. The numbers in the top panel are in annualized percentage. The mean and standard deviation of the variance premium are converted to monthly values by multiplying $10^4/12$. Models 1A and 2A are the models with standard time-additive CRRA utility and with Epstein–Zin preferences under full information, respectively. Models 1B and 2B are the corresponding models with Bayesian learning, respectively. “Full Model” is our general model with ambiguity aversion.
standard rational model to explain the high equity volatility observed in the data, creating 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 of 2.82\%, mean equity premium of 0.80\%, and equity volatility of 12.62\%, which are far away from the data. However, the performance of Model 2B with the Epstein–Zin preferences is better than that of Model 1B with the time-additive CRRA utility for $\eta = \gamma = \rho = 2$.

Table 4 shows that Model 2A and Model 2B yield very similar predictions. This finding 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 4 also shows that our model-generated volatility of the risk-free rate is lower than the data (1.05\% versus 5.68\%). 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 1919–1921. 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\% using the US postwar quarterly data. Thus, we view our model-generated low risk-free rate volatility as conceivable. By contrast, the widely used habit formation model (e.g., Jermann, 1998; Boldrin et al., 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.

Turning to the model predictions regarding the variance premium, Table 4 shows that our full model-implied mean–variance premium is 8.47 (percentage squared, monthly basis), which is about 90\% of the data, 9.46. By contrast, Models 1A and 1B with standard time-additive CRRA utility generate very small values of the mean–variance premium (0.07 and 0.09, respectively). Separating EIS from risk aversion, as in Models 2A and 2B, raises the mean–variance premium to 0.25. 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 previous numbers imply that risk aversion alone contributes about 1\% to the model-implied mean–variance premium, separating EIS from risk aversion contributes about 2\%, and the remaining 97\% is attributed to ambiguity aversion.
Model 1B by 144 times and in Model 2B by 52 times. Separating ambiguity aversion from risk aversion raises the mean value of the first component from 0.052, 0.147, and 7.588 in Model 1B, Model 2B, and the full model, respectively; it accounts for 0.5%, 1.4%, and 70.9% of the mean–variance premium in the data, respectively. This result shows that the full model with ambiguity aversion and learning raises the combined value of these two components in Model 1B by 144 times and in Model 2B by 52 times.

To understand why our model with ambiguity aversion and learning can generate a large mean–variance premium, we decompose variance premium in Table 5. As discussed in Sec. 4, we can decompose the variance premium into three components as in equation (16), repeated as follows

\[
\text{VP}_t = (\mu_t - \hat{\mu}_t)(\mathbb{E}_t^2[\Sigma_{t+1}] - \mathbb{E}_t^1[\Sigma_{t+1}]) + \frac{\hat{\mu}_t}{\mathbb{E}_t^1[M_{t+1}^{E2}]} \text{Cov}_t^1 (\Sigma_{t+1}, M_{t+1}^{E2}) + \frac{1}{\mathbb{E}_t^2[M_{t+1}^{E2}]} \text{Cov}_t^2 (\Sigma_{t+1}, M_{t+1}^{E2}).
\]

(17)

Separating risk aversion from EIS raises the mean value of the first component — belief distortions \((\mu_t - \hat{\mu}_t)\) — 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 component from 0.008 in Model 2B to 0.123 in the full model, a remarkable 14 times increase.

We now turn to the second component: the stock variance differential 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 component from 7.976 in Model 1B to 16.097 in Model 2B. Separating ambiguity aversion from risk aversion further raises it from 16.097 in Model 2B to 58.648 in the full model, generating a significant three times increase.

The mean value of the product of these two components is equal to 0.052, 0.147, and 7.588 in Model 1B, Model 2B, and the full model, respectively; it accounts for 0.5%, 1.4%, and 70.9% of the mean–variance premium in the data, respectively. This result shows that the full model with ambiguity aversion and learning raises the combined value of these two components in Model 1B by 144 times and in Model 2B by 52 times.

Table 5. Variance premium decomposition.

<table>
<thead>
<tr>
<th></th>
<th>Model 1B</th>
<th>Model 2B</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathbb{E}[F1])</td>
<td>0.006</td>
<td>0.008</td>
<td>0.123</td>
</tr>
<tr>
<td>(\mathbb{E}[F2])</td>
<td>7.976</td>
<td>16.097</td>
<td>58.648</td>
</tr>
<tr>
<td>(\mathbb{E}[F1 \times F2])</td>
<td>0.052</td>
<td>0.147</td>
<td>7.588</td>
</tr>
<tr>
<td>(\mathbb{E}[F3])</td>
<td>0.041</td>
<td>0.106</td>
<td>0.868</td>
</tr>
<tr>
<td>(\text{VP})</td>
<td>0.093</td>
<td>0.253</td>
<td>8.466</td>
</tr>
</tbody>
</table>

Notes: This table reports the decomposition of the mean–variance premium for Model 1B, Model 2B and our full model with ambiguity aversion as in Sec. 4. In the table, \(F1 \equiv \mu_t - \hat{\mu}_t, F2 \equiv \mathbb{E}_t^2[\Sigma_{t+1}] - \mathbb{E}_t^1[\Sigma_{t+1}],\) and \(F3 \equiv \hat{\mu}_t \frac{\text{Cov}_t^1(\Sigma_{t+1}, M_{t+1})}{\mathbb{E}_t^1[M_{t+1}]} + (1 - \hat{\mu}_t) \frac{\text{Cov}_t^2(\Sigma_{t+1}, M_{t+1})}{\mathbb{E}_t^2[M_{t+1}]}\). The mean–variance premium \(\text{VP} = \mathbb{E}[F1 \times F2] + \mathbb{E}[F3]\). The full model is our general model with ambiguity aversion. Models 1B and 2B are special cases of our full model with \(\eta = \gamma = \rho = 2\) and \(\eta = \gamma = 2 \neq \rho\), respectively.
Table 5 reveals that the third covariance component is quantitatively small. Its mean value is equal to 0.041, 0.106, and 0.868 in Model 1B, Model 2B, and the full model, respectively; it accounts for 0.4%, 1.0%, and 8.1% of the total mean–variance premium in the data, respectively. Ambiguity aversion raises this component in Model 1B by 21 times and in Model 2B by about eight times.

To further understand the intuition, Fig. 2 plots the conditional variance premium and the three components in the variance premium decomposition as functions of the Bayesian beliefs about the high-growth state. This figure shows that all three components are positive and hump-shaped. Ambiguity aversion amplifies each component significantly, generating large movements of the conditional variance premium. The intuition behind this result was

---

Fig. 2. Determinants of variance premium.

Notes: This figure plots the conditional variance premium and its decomposition in terms of three components. The solid lines correspond to the full model with ambiguity aversion. The dashed lines correspond to Model 2B with Epstein–Zin preference and Bayesian learning with $\gamma = \eta \neq \rho$. The dotted lines correspond to Model 1B with time-additive CRRA utility and Bayesian learning with $\gamma = \eta = \rho$. Except for the panel for $\hat{\mu}_t - \hat{\mu}_t$, the vertical axes are expressed in percentage squared and divided by 12.
discussed in Secs. 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 in more detail.

Figure 3 plots the price–dividend ratio and the conditional stock return variance as functions of the Bayesian beliefs about the boom 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 bottom panel of Fig. 3. As the economy spends most time in the boom 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 4. Our model-implied standard deviation of the variance premium 6.25 accounts for 71% of the data 8.81. In unreported simulation results, we find that the belief distortion component \((\mu_i - \hat{\mu}_i)\) explains most of the volatility of the variance premium. Though our model-implied volatility still falls short of the data, it is about 20 to 70 times as large as that in other models reported in Table 4. Regarding the autocorrelation coefficient, our model and the other models all deliver a lower level than the data 0.53. It appears that our model does not have a strong propagation mechanism to generate the persistence of the variance premium, but it has a powerful amplification mechanism to magnify the mean and volatility of the variance premium.

5.3. Return predictability

Recent empirical evidence has suggested that the variance premium is a highly significant predictor for stock market returns at short horizons, especially one quarter, with Newey–West \(t\)-statistics ranging from 2.86 to 3.53 and \(R^2\)'s from 6% to 8%, while the predictability over long horizons typically becomes quite weak (Bollerslev et al., 2009; Drechsler and Yaron, 2011; Drechsler, 2013). However, 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% over one-to-five year horizons (Drechsler and Yaron, 2011). More importantly, when the
Fig. 3. Price–dividend ratio and conditional variance.

Notes: This figure plots the price–dividend ratio (top panel) and conditional variance (bottom panel) as functions of the Bayesian posterior probabilities of the boom state. The solid lines correspond to the full model with ambiguity aversion. The dashed lines correspond to Model 2B with Epstein–Zin preferences and Bayesian learning with $\gamma = \eta \neq \rho$. The dotted lines correspond to Model 1B with time-additive CRRA utility and Bayesian learning with $\gamma = \eta = \rho$. 
variance premium and price-earning ratios are combined, there is complementarity in that the joint regression $R^2$ (16.76%) is higher than the sum of the two $R^2$’s of univariate regressions — 6.82% and 6.55% (see, e.g., (Bollerslev et al., 2009)).

Table 6 confirms some of the previous empirical findings on return predictability. Since our constructed empirical variance premium is at a monthly frequency in the sample 1990–2015, to run predictability regressions, we use monthly data of stock market returns on the S&P 500 index, the risk-free rate, and the dividend yield between 1990 and 2015, which are constructed as in Welch and Goyal (2008) and are available from either author’s website. The regressions in Panels A and B of Table 6 are forecasts at horizons of one to three years using the dividend yield and the variance premium as the univariate predictor, respectively. The dividend yield’s predictive power increases with the horizon, accounting for 29% of the monthly excess return variation. By contrast, the variance premium accounts for 3–5% of the monthly excess return variation and has no predictive power at the horizon of three years. Motivated by these empirical findings on the return predictability, we report our model-implied values of regression $R^2$’s, slope coefficients, and $t$-statistics, at horizons of one, two, and three years, based on the

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Dividend Yield</th>
<th>Variance Premium</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Model</td>
<td>Data</td>
</tr>
<tr>
<td>Panel A: Univariate predictive regression using dividend yield</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.20 (4.31)</td>
<td>0.82 (2.64)</td>
<td>0.12</td>
</tr>
<tr>
<td>2</td>
<td>0.43 (5.71)</td>
<td>1.12 (3.40)</td>
<td>0.22</td>
</tr>
<tr>
<td>3</td>
<td>0.68 (6.92)</td>
<td>1.24 (3.76)</td>
<td>0.29</td>
</tr>
<tr>
<td>Panel B: Univariate predictive regression using variance premium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3.26 (3.29)</td>
<td>6.35 (2.41)</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>4.09 (2.38)</td>
<td>8.67 (2.89)</td>
<td>0.03</td>
</tr>
<tr>
<td>3</td>
<td>1.85 (0.73)</td>
<td>9.73 (3.11)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Notes: Panels A and B present the results of the univariate predictive regressions of the excess returns on dividend yields and on variance premium, respectively. $t$-statistics are reported in the parentheses and are Newey–West (HAC) corrected with two lags. Column “Data” reports regression results using the actual monthly data from 1990.1 to 2015.12. The dependent variable is the excess return on the S&P 500 index over the following one-, two-, and three-year horizons. The regressions are conducted using the overlapping return series. Column “Model” reports the regression results obtained as the mean values of 10,000 Monte Carlo simulations of our full model at an annual frequency, each consisting of 126 data points.
baseline calibration at an annual frequency. As can be seen from Table 6, both the dividend yield and variance premium can predict return with highly significant t-statistics in univariate regressions. These results mimic the empirical regularity that both dividend yield (or P/E ratio) and variance premium are return predictors, but still fall short of replicating the data. It should be pointed out that as there is only one state variable $\mu_t$ in our model, when dividend yield and variance premium are combined, they crowd out each other as both become insignificant in regressions not reported here. More state variables need to be introduced to our model, to make both dividend yield and variance premium significant and complementary in joint regressions, as empirically reported by Bollerslev et al. (2009) and Drechsler and Yaron (2011).

We should mention that although our model cannot generate the 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.\footnote{These results are available upon request.}

### 5.4. Historical variance premium

To further elicit economic intuition on why and how variance premium changes with economic fundamentals, we feed our calibrated models with historical consumption growth data from 1890 to 2015 from Robert Shiller’s website. As shown in Fig. 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), the US initial engagement in World War II (1942), and the Great Recession (2008–2009). Also, the major belief switches or spikes, as shown in Fig. 4 bottom panel, do reflect those severe down times, while minor belief deviations from 1 seem to capture all other mild economic recessions.

Figure 5 shows that, in normal times of the past century, the variance premium implied by our full model is at a low range of 5–8, while it shot up to 37 during the Great Depression, 35 and 33 during the financial panics of 1894 and 1908, around 29 during the 1925 depression, in the range of 16 to 26 during deep recessions of 1914–1915 and 1917, and about 18 when the US joined the war in 1942. After the war, the highest level of the variance premium ever reached was about 13, near the end of the 2007–2009 global
financial crisis, which is still dwarfed by the Great Depression and other prewar spikes. In essence, the model-implied range of historical variance premium of 6 to 37, is driven mostly by ambiguity aversion to the uncertain economic recessions and depressions. Figure 5 also presents the empirical variance premium during 1990–2015. It shows that although our model can capture the spikes in the variance premium during the 1990 recession and the 2007–2009 financial crisis, it still falls short of matching the data. Our model also fails to generate the spikes during the 1997–1998 Asia–Russia–LTCM crisis and the 2001 recession. This is because consumption growth does not drop much during these periods and hence belief distortions are not large.

For comparison, the top panel of Fig. 6 plots the variance premium in Model 1B, which has a range of 0.06–0.54. The middle panel plots the difference between the variance premium implied by Models 2B and 1B, which
shows the contribution due to Epstein–Zin utility. The bottom panel plots the difference between the variance premium implied by our full model and Model 2B, which shows the contribution due to ambiguity aversion. We can see that ambiguity aversion contributes to about 97% of the historical variance premium dynamics.

5.5. Alternative calibration

For our baseline calibration in Sec. 5.1, we fix the risk aversion coefficient at $\gamma = 2$, and calibrate the subjective discount factor $\beta$ and the ambiguity aversion parameter $\eta$ to match the mean risk-free rate and the mean equity premium. We obtain $\beta = 0.9838$ and $\eta = 10.948$. We then compare our full model with alternative models by shutting down ambiguity aversion with $\eta = \gamma = 2$. One may wonder whether the Epstein–Zin model with a larger ...
value of the risk aversion parameter can perform similarly to the model with ambiguity aversion. We now address this issue in this subsection. Note that Klibanoff et al. (2005), Hayashi and Miao (2011), and Ju and Miao (2012) have shown theoretically that the Epstein–Zin model is behaviorally distinct from the smooth ambiguity model adopted in this paper and that these two utility models are not observationally equivalent. The goal of this subsection is to explore the quantitative implications.

First, we re-calibrate both the risk aversion coefficient $\gamma$ and the subjective discount factor $\beta$ in Model 2B with Epstein–Zin utility to match the risk-free rate and the mean equity premium. Table 7 reports the new calibration results. Not surprisingly, the re-calibrated value of $\gamma$ is higher to match the equity premium. With the re-calibrated parameter values, the performance of

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{variancepremium.png}
\caption{Model-implied variance premium decomposition (1890–2015).}
\end{figure}

\textit{Notes:} The top panel with the dotted line plots the variance premium implied by Model 1B with time-additive CRRA utility and Bayesian learning ($\gamma = \eta = \rho$). The middle panel with the dashed line plots the difference between the variance premium implied by Model 2B with Epstein–Zin preferences and Bayesian learning ($\gamma = \eta \neq \rho$) and that implied by Model 1B. The bottom panel with the solid line plots the difference between the variance premium implied by the full model with ambiguity aversion and that implied by Model 2B. All vertical axes are expressed in percentage squared, monthly basis.
Model 2B improves in matching the risk-free rate, the equity premium, and their standard deviations. Moreover, Model 2B can better match the magnitude of the variance premium as well. However, even though it is a big improvement relative to the magnitude of 0.25 obtained in the previous calibration (see Table 4), Model 2B can only explain about one-half of the magnitude in the data, 9.46. Thus, the Epstein–Zin model still performs worse than the full model with ambiguity aversion.

Next we conduct an alternative calibration of the full model to see how well it can simultaneously match all three key moments: the mean risk-free rate, the mean equity premium, and the mean variance premium. Instead of having the risk aversion coefficient $\gamma$ fixed in the baseline calibration, we now:

### Table 7. Alternative calibration results.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Model 2B</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>6.9701</td>
<td>2.0400</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.9823</td>
<td>0.9826</td>
</tr>
<tr>
<td>$\eta$</td>
<td>NA</td>
<td>11.3585</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>Model 2B</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_f$</td>
<td>1.76</td>
<td>1.90</td>
<td>1.76</td>
</tr>
<tr>
<td>$\sigma(R_f)$</td>
<td>5.68</td>
<td>0.95</td>
<td>1.07</td>
</tr>
<tr>
<td>$\mu_{eq}$</td>
<td>6.02</td>
<td>5.66</td>
<td>6.63</td>
</tr>
<tr>
<td>$\sigma(\mu_{eq})$</td>
<td>18.52</td>
<td>15.95</td>
<td>16.67</td>
</tr>
</tbody>
</table>

Panel C: Variance premium

| VP         | 9.46     | 5.52     | 9.53       |
| $\sigma(VP)$ | 8.81   | 4.27     | 6.66       |
| AC(1)      | 0.53     | 0.34     | 0.29       |

Notes: In the alternative calibration of Model 2B, we calibrate the parameters $\gamma$ and $\beta$ to match the mean equity premium ($\mu_{eq}$) and the mean risk-free rate ($R_f$). In the alternative calibration of the Full Model, besides the parameters $\gamma$ and $\beta$, we further calibrate the parameter $\eta$ to match the mean–variance premium (VP). The calibrated parameter values are reported in Panel A. The model-implied moments are reported in Panels B and C. $\sigma(R_f)$ denotes the volatility of the risk-free rate; $\sigma(\mu_{eq})$ denotes the equity premium volatility; and $\sigma(\text{VP})$ denotes the volatility of the variance premium. The numbers in Panel B are in annualized percentage. The mean and standard deviation of variance premium are converted to monthly values by multiplying $10^4/12$. Model 2B is the Epstein–Zin model with Bayesian learning. The full model is our model with ambiguity aversion. For both models, $\rho = 2/3$. 

Model 2B improves in matching the risk-free rate, the equity premium, and their standard deviations. Moreover, Model 2B can better match the magnitude of the variance premium as well — it can now generate a mean–variance premium as large as 5.52. However, even though it is a big improvement relative to the magnitude of 0.25 obtained in the previous calibration (see Table 4), Model 2B can only explain about one-half of the magnitude in the data, 9.46. Thus, the Epstein–Zin model still performs worse than the full model with ambiguity aversion.

Next we conduct an alternative calibration of the full model to see how well it can simultaneously match all three key moments: the mean risk-free rate, the mean equity premium, and the mean–variance premium. Instead of having the risk aversion coefficient $\gamma$ fixed in the baseline calibration, we now...
re-calibrate $\gamma$ together with the other two parameters, $\beta$ and $\eta$, to match the aforementioned three moments. In other words, different from the baseline calibration exercise, the mean–variance premium is now a target moment used for calibration.

The last column in Panels B and C of Table 7 reports the results of this alternative calibration of the full model. From the table, we can see that the re-calibrated value of $\eta$ is larger than the previous value, 11.358 versus 10.948, while the re-calibrated values of $\gamma$ and $\beta$ are very close to the previous values. The higher ambiguity aversion estimate leads to a very similar risk-free rate (1.75% versus 1.76%), and a slightly higher equity premium (6.63% versus 6.02%). However, most importantly, it helps match closely the mean–variance premium in the data. In the baseline calibration, the full model can only generate a mean–variance premium of 8.51. By contrast, the new calibration enhances the performance by increasing the mean–variance premium to 9.53, very close to the value of 9.46 empirically observed in the data. Moreover, the model-generated volatility of the variance premium (6.66) is also closer to the data (8.81).

6. Further Analysis
6.1. Disaster risk

In Sec. 5.1, we have reported estimation results for consumption growth using the full sample 1890–2015 and the subsample 1946–2015. We find that one cannot identify large negative consumption growth in the recession state for the subsample because large downside shifts are very rare during the postwar period. Moreover, a small sample with only 69 observations cannot give a reliable econometric estimation. In this section, we will show the importance of the disaster risk or a large negative consumption growth by re-calibrating the model based on the estimates of the consumption process for the subsample 1946–2015.

Following the same calibration procedure as in Sec. 5.1, we set $\gamma = 2$, $\rho = 2/3$, $\beta = 0.986$, $\eta = 267.18$, $g_d = -0.0359$, $\zeta = 2.74$, and $\sigma_d = 0.0424$. These parameter values allow the model to roughly match the target of the risk-free rate and the mean equity premium. But the model-implied equity volatility 7.02% is much lower than the data 18.52%. More importantly, the model-implied mean and volatility of the variance premium (0.12 and 0.03) are much lower than the data (9.46 and 8.81). The model generated time series of the variance premium is almost flat. The key reason is that consumption growth in the recession rate is estimated to be 0.25% for the
postwar period, which is not an economic disaster. Also, the recession and boom states are not too much different. This means that the agent does not face too much uncertainty and thus a very large ambiguity aversion parameter is needed to match the mean equity premium. But the model-implied second moments such as the equity volatility and the variance premium are much smaller than the data.

6.2. Index option prices and implied volatilities

Our model has implications for index option prices and implied volatilities as in Drechsler (2013). Drechsler (2013) studies implied-volatility curves for index options with one-, three-, and 12-month maturities in a continuous time model. Since our model is in discrete time and calibrated at the annual frequency, we consider index option with one-year maturity only.

We use the following equation to compute index call option prices

\[ O_t = E_t \left[ M_{t+1} \max \{ P_{e,t+1} - K, 0 \} \right], \]

where \( K \) is the strike price. We then plug the aforementioned option prices into the Black–Scholes formula to compute the implied volatility. Figure 7 presents the model and empirical implied volatilities for strikes ranging in moneyness (strike/spot price) from 0.75 to 1.25. The empirical implied volatilities are the average daily implied volatilities for options on the S&P 500 index, obtained from the daily volatility surface files in the Option Metrics database between January 1996 and August 2015. When we plot the model-based implied-volatility curves, we use model-based option prices calculated by setting the model’s state variable \( \mu_t \) at its unconditional mean.

As a comparison, Fig. 7 also presents the implied-volatility curve for Model 2B (the Epstein–Zin model with Bayesian learning) by setting \( \eta = \gamma = 2 \).

Figure 7 shows that our full model with ambiguity aversion does a good job matching the downward-sloping shape of the implied-volatility curve, but over-estimates the level of the implied volatilities. By contrast, the implied-volatility curve for Model 2B is almost flat and lower than the empirical curve. Thus our model outperforms Model 2B in this dimension. But we acknowledge that our discrete-time model calibrated at the annual frequency is not quite suitable to study the issue of implied volatilities. In particular, we cannot study options with short maturities like one month or three months. Moreover, the Black–Scholes formula applies to continuous-time models. Thus, developing a continuous-time version of our model is critical for option pricing and will be left for future research.
7. Conclusion

This paper provides an ambiguity-based interpretation of the variance premium as a compounding effect of both belief distortions regarding unknown economic regimes and market variance differentials between regimes. Our calibrated model with ambiguity aversion can generate a mean–variance premium of 8.46 (percentage squared, monthly basis), quite close to the empirical estimate of 9.46. About 97% of the model-implied variance premium can be attributed to ambiguity aversion. Our model can simultaneously explain the equity premium, equity volatility, and risk-free rate puzzles, and can also generate a sizable volatility 6.25 of the variance premium, compared to the data 8.81. Model implied variance premium series
using the consumption growth data from 1890 to 2015 reveals that the major spikes in variance premium 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 the 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 beliefs about the unknown economic regimes and the 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 the agent’s fear of uncertain economic downside shifts.

References


Han, B., and Y. Zhou, 2011, Variance Risk Premium and Cross-Section of Stock Returns, Working Paper, University of Texas at Austin.