Do Subjective Expectations Explain Asset Pricing Puzzles?*

Gurdip Bakshi[a†] Georgios Skoulakis[b ‡]

[a]*Smith School of Business, University of Maryland, College Park, MD 20742, USA*

[b]*Smith School of Business, University of Maryland, College Park, MD 20742, USA*

First draft, May 2008

This draft, November 30, 2009

Abstract

The structural uncertainty model with Bayesian learning, advanced by Weitzman (AER 2007), provides a framework for gauging the effect of structural uncertainty on asset prices and risk premiums. This paper provides an operational version of this approach that incorporates realistic priors about consumption growth volatility, while guaranteeing finite asset pricing quantities. In contrast to the extant literature, the resulting asset pricing model with subjective expectations yields well-defined expected utility, finite moment generating function of the predictive distribution of consumption growth, and tractable expressions for equity premium and riskfree return. Our quantitative analysis reveals that explaining the historical equity premium and riskfree return, in the context of subjective expectations, requires implausible levels of structural uncertainty. Furthermore, these implausible prior beliefs result in consumption disaster probabilities that virtually coincide with those implied by more realistic priors. At the same time, the two sets of prior beliefs have diametrically opposite asset pricing implications.

KEY WORDS: subjective expectations; learning; structural uncertainty; priors; predictive density of consumption growth; equity premium; riskfree return.
JEL CLASSIFICATION CODES: D34, G12.

---

*The feedback of an anonymous referee and G. William Schwert (the editor) have dramatically improved the paper. The authors acknowledge helpful discussions with Doron Avramov, Isabelle Bajeux-Besnainou, Alexandre Baptista, Zhiwu Chen, Alex David, Dobrislav Dobrev, Lorenzo Garlappi, Steve Heston, Jerry Hoberg, Ravi Jagannathan, Nengjiu Ju, Pete Kyle, Dilip Madan, Anna Obizhaeva, Lubos Pastor, George Panayotov, Nagapurand Prabhala, Matt Pritsker, Mark Loewenstein, Jay Shanken, Steve Sharpe, Allan Timmermann, and Martin Weitzman. Earlier versions of the paper were presented at the Federal Reserve Board, University of Maryland, George Washington University, 2008 Latin American Meeting of the Econometric Society (Rio de Janeiro), the 2009 Society for Financial Econometrics Conference (Geneva), the 2009 Multinational Finance Society Conference (Crete), and the 2009 Society for Economic Dynamics Conference (Istanbul). We welcome comments, including references to related papers we have inadvertently overlooked.

†Tel.: +1-301-405-2261; fax: +1-301-405-0359. *E-mail address:* gbakshi@rhsmith.umd.edu
‡Tel.: +1-301-405-2934; fax: +1-301-405-0359. *E-mail address:* gskoulak@rhsmith.umd.edu

1. Introduction

The structural uncertainty model with Bayesian learning, developed by Weitzman (2007), provides an economic environment for understanding the effect of structural uncertainty on asset prices and risk premiums. Rooted in the theoretical underpinnings of Lucas (1978), the logical structure of the model formalizes the sense in which structural parameter uncertainty can dominate aggregate consumption risk in explaining aggregate stock market behavior.[1]

In the asset pricing setup of Weitzman (2007), the representative agent experiences uncertainty about the volatility of consumption growth, the support of which is assumed to be the entire positive real line, and updates his beliefs in a Bayesian fashion. Drawing on the work of Shephard (1994), Weitzman notes that the predictive distribution of consumption growth, which is Normal in the absence of structural uncertainty, gets transformed into a heavy-tailed Student-t distribution. This theoretical modification within the Lucas paradigm is enormous in its implications. First, the theoretical results imply that the agent can demand arbitrarily large compensation for bearing uncertainty about consumption growth volatility, the true structure of which remains unknown forever. Second, the observed magnitudes of equity risk premium and riskfree return are not to be regarded as puzzles requiring reconciliation, but rather as antipuzzles. The effect of structural uncertainty in a Bayesian framework is sufficiently powerful to reverse the direction of asset pricing puzzles, implying the futility of aligning theoretical models to fit observed asset pricing quantities.

How robust is the implication in Weitzman (2007) that structural uncertainty and learning can quantitatively exert a big influence on the tail behavior of consumption growth, and hence on asset pricing quantities? Are the arguments in Weitzman (2007) inextricably linked to the predictive Student-t distribution for consumption growth for which the moment generating function does not exist, or do they apply more broadly? For instance, how do the asset pricing implications change when the Student-t distribution is replaced by a comparable class of heavy-tailed Bayesian predictive distributions? Finally, what levels of *a priori* uncertainty are required to reconcile asset-return puzzles, and do such conditions appear reasonable?

[1] There is a long list of theoretical and empirical research based on the consumption based asset pricing model. See, for instance, Lucas (1978), LeRoy and Porter (1981), Mehra and Prescott (1985), Rietz (1988), Epstein and Zin (1991), Bakshi and Chen (1996), and Campbell and Cochrane (1999, 2000). Related treatments and refinements are exposited in Weil (1989), Campbell (1993), Cecchetti, Lam, and Mark (2000), Abel (2002), Aït-Sahalia, Parker, and Yogo (2004), Bansal and Yaron (2004), Barro (2006), Gabaix (2007), Lettau and Wachter (2007), Lettau, Ludvigson, and Wachter (2007), Cogley and Sargent (2008), and David (2008). Overview of these approaches are provided in Mehra and Prescott (2003, 2008), and Kocherlakota (1996). A strand of contributions in Bayesian asset pricing model building and learning include Timmermann (1993, 1996), Brennan and Xia (2001), Lewellen and Shanken (2002), Tsionas (2005), Jobert, Platania, and Rogers (2006), Adam, Marcet, and Nicolini (2008), Cogley (2009), and Pastor and Veronesi (2009).

At the heart of our contribution is an exact characterization of the predictive distribution of consumption growth that supports the finiteness of the moment generating function under subjective uncertainty and Bayesian learning. The operational version of the theory we offer is imperative for ensuring finite expected utility (e.g., Geweke (2001)) and for studying the quantitative asset pricing implications of the model, when the support of consumption growth volatility is a bounded interval.

We operationalize the theory by accommodating prior and posterior distributions of the precision of consumption growth that are both in the truncated Gamma class. The linchpin of our approach is a theorem that establishes the conjugacy of the prior and the posterior distributions for the precision of consumption growth, under a judicious choice of a transition mechanism. Our analysis shows that the derived predictive density of consumption growth is akin to the Student-t distribution in terms of its probabilistic law, except that it possesses finite moments of all orders. The model offers the advantage of a representation of asset pricing quantities under structural uncertainty and Bayesian learning that is amenable to convenient computation through the moment generating function. Our approach can be construed as a mathematical formalization of an asset pricing model that incorporates a compact support for consumption growth volatility, and addresses the intuition conveyed in Weitzman (2007) on the role of subjective expectations in determining asset prices. The model also maintains the non-ergodic aspect of learning as in Weitzman (2007), where uncertainty about consumption growth volatility does not vanish even with an arbitrarily large amount of past data.

The model with structural uncertainty is implemented with a view to investigate its potential to resolve asset pricing puzzles. First, we observe that plausible structural uncertainty, summarized by the uncertainty about consumption growth volatility, fails to produce a large equity premium. Second, the response of asset pricing quantities is flat over a broad range of configurations of structural uncertainty. Only when the maximum level of consumption growth volatility is unreasonably high can the model match the historical average equity premium and riskfree return. Third, we find that two sets of priors, that result in consumption disaster probabilities that are indistinguishable, nevertheless generate vastly different asset pricing implications. The gist is that rationalizing asset pricing phenomena through subjective expectations demands arguably excessive levels of structural uncertainty.

The paper proceeds as follows. Section 2 outlines the Weitzman (2007) framework for understanding the asset pricing implications of structural uncertainty. Section 3 departs from Weitzman (2007) and provides an alternative characterization of the predictive density of consumption growth that supports a finite moment generating function. Our generalization posits uncertainty about consumption growth volatility over a realistic compact support instead of the entire positive real line. The focus of Section 4 is to investigate whether structural uncertainty with realistic priors facilitates a resolution of well-documented asset pricing puzzles. Finally, Section 5 concludes the paper. Proofs are in the Appendix.

2. Review of the structural uncertainty framework in Weitzman (2007)

The following assumptions about preferences, subjective expectations, and learning about structural uncertainty are at the center of the theoretical analysis in Weitzman (2007).

Assumption 1 *The model is developed in terms of a representative agent who orders his preferences over random consumption paths, and maximizes utility subject to the usual budget constraint (Lucas (1978)),*

$$E_t\left(\sum_{j=0}^{\infty}\beta^j U(C_{t+j})\right), \qquad \text{where} \qquad U(C_t)=\frac{C_t^{1-\alpha}}{1-\alpha}. \tag{1}$$

C_t denotes consumption at time t, $0<\beta<1$ is the time-preference rate, and $E_t(.)$ is expectation operator with respect to the subjective distribution of future consumption growth rates. The agent has power utility function with coefficient of relative risk aversion $\alpha>0$.

This model differs from the previous literature in that it accommodates probability beliefs as reflected in subjective expectations about consumption growth. There is uncertainty in the economy with respect to the stochastic process of consumption growth,

$$X_{t+1}\equiv\ln\left(\frac{C_{t+1}}{C_t}\right)\in(-\infty,\infty). \tag{2}$$

In particular, the conditional volatility of consumption growth, denoted by σ_t, is stochastic and unobservable. Define the precision as

$$\theta_t \;\equiv\; \frac{1}{\sigma_t^2}\in(0,\infty), \tag{3}$$

which is the reciprocal of the variance of consumption growth.

Suppose observations on realized consumption growth are available starting in period τ, which should be thought of as representing the distant past. For each time period $t=\tau,\tau+1,\ldots$, denote by X^t the set of all observations on consumption growth up to and including period t: $X^t = \{x_\tau, \ldots, x_t\}$.

The intertemporal marginal rate of substitution between t and $t+1$ is given by $\beta \left(\frac{C_{t+1}}{C_t}\right)^{-\alpha} = \beta e^{-\alpha X_{t+1}}$. At each time t, the agent obtains the subjective distribution of X_{t+1} by conditioning on all past data X^t.

Assumption 2 *Conditional on precision θ_t, the consumption growth, X_t, is Normally distributed,*

$$X_t \sim N\left(\mu, \frac{1}{\theta_t}\right), \qquad \text{with} \qquad p\left(X_t|\theta_t\right) = \frac{1}{\sqrt{2\pi}} \theta_t^{\frac{1}{2}} e^{-\frac{\theta_t (X_t-\mu)^2}{2}}, \tag{4}$$

where $p\left(X_t|\theta_t\right)$ denotes the density of consumption growth conditional on θ_t. The mean consumption growth is μ, which is assumed to be a known constant.

It must be appreciated that the Weitzman (2007) framework can be refined to allow for learning about mean consumption growth without affecting model predictions about asset pricing puzzles. In particular, Theorem 2.25 in Bauwens, Lubrano, and Richard (1999) shows that when both the mean and the variance are unknown, the resulting predictive distribution is Student-t, but with inflated variance. Furthermore, Example 3 in Geweke (2001) emphasizes that in a model with *known* variance and *unknown* mean, the predictive and the primitive consumption growth distributions are both Gaussian. It is learning about volatility, as opposed to learning about the mean, that is fundamental to generating fat-tailed predictive consumption growth distributions that are important for addressing the asset pricing issues at hand.

The agent is uncertain about the true precision θ_{t+1} and maintains subjective beliefs, captured by the conditional density $p(\theta_{t+1}|X^t)$ to be described shortly. The modeling innovation of Weitzman (2007) is that the agent makes Bayesian inferences about θ_{t+1} given history X^t, combining elements from Harvey and Fernandes (1989), Shephard (1994), and Geweke (2001).

Assumption 3 *Assume that, in the initial period τ, the conditional density, $p\left(\theta_\tau|X^\tau\right)$, is Gamma$(a_\tau, b_\tau)$. Specifically,*

$$p\left(\theta_\tau|X^\tau\right) = \frac{b_\tau^{a_\tau}}{\Gamma[a_\tau]} \theta_\tau^{a_\tau - 1} e^{-b_\tau \theta_\tau}, \qquad \theta_\tau \in (0, \infty), \qquad a_\tau > 0, \qquad b_\tau > 0, \tag{5}$$

where $\Gamma[a] = \int_0^\infty z^{a-1} e^{-z} dz$ is the complete Gamma function with $a > 0$.

The evolution of precision θ_t is described in the following sequential fashion, so as to maintain the conjugacy of the system through time.

Assumption 4 *Given precision of consumption growth θ_t, the evolution of precision at date $t+1$, θ_{t+1}, is governed by the transition equation,*

$$\theta_{t+1} \;=\; \frac{1}{\omega}\,\eta_{t+1}\,\theta_t, \qquad\qquad 0<\omega<1, \tag{6}$$

where the multiplicative shock η_{t+1}, given the history X^t, follows a $\mathrm{Beta}\left(\omega a_t,(1-\omega)\,a_t\right)$ *distribution.*

Equation (6) specifies the learning scheme about structural uncertainty in the economy and is the driving force behind the fatter tails of the predictive distribution of consumption growth. The parameter ω is a constant that controls the speed of precision. An important aspect of the above specification, as adopted in Weitzman (2007), is that the support of the precision process is the entire positive real line. As a result, the consumption growth volatility can take arbitrarily large values with positive probability.

Under Assumptions 3 and 4, it follows from the analysis in Section 3 in Shephard (1994) that the conditional density, $p\left(\theta_{t+1}|X^t\right)$, remains in the Gamma family and preserves conjugacy. Specifically,

$$\theta_t|X^t \sim \mathrm{Gamma}(a_t,b_t) \qquad \text{and} \qquad \theta_t|X^{t-1} \sim \mathrm{Gamma}(A_{t-1},B_{t-1}), \tag{7}$$

where a_t, b_t and A_{t-1}, B_{t-1} are related via the recursive equations,

$$a_t = A_{t-1}+\frac{1}{2}, \qquad A_{t-1}=\omega a_{t-1}, \qquad b_t = B_{t-1}+\frac{1}{2}(x_t-\mu)^2, \qquad B_{t-1}=\omega b_{t-1}. \tag{8}$$

The recursions involving A_t and B_t alone are $A_t = \omega\left(A_{t-1}+\frac{1}{2}\right)$ and $B_t=\omega\left(B_{t-1}+\frac{1}{2}(x_t-\mu)^2\right)$. Note that an extreme realization of consumption growth at time t causes a large squared deviation $(x_t-\mu)^2$ which, in turn, increases B_t. Thus, the overall impact is to decrease the mean A_t/B_t of the conditional distribution of the precision θ_{t+1} given the history X^t.

Moreover, recursive substitution in terms of past history τ of consumption growth yields:

$$A_t \;=\; \frac{1}{2}(\omega+\cdots+\omega^{t-\tau}) \;+\; \omega^{t-\tau}A_\tau = \frac{\omega}{2}\left(\frac{1-\omega^{t-\tau}}{1-\omega}\right) \;+\; \omega^{t-\tau}A_\tau, \tag{9}$$

$$B_t \;=\; \frac{1}{2}\left(\omega(x_t-\mu)^2 \;+\; \omega^2(x_{t-1}-\mu)^2+\cdots+\omega^{t-\tau}(x_{\tau+1}-\mu)^2\right) \;+\; \omega^{t-\tau}B_\tau. \tag{10}$$

In the spirit of Lemma 5 of Weitzman (2007), assume the availability of a large history of consumption growth starting in period τ, and let $\omega=(k-1)/k$, where k captures the effective sample size. Imposing a large (in absolute value) $\tau < 0$ in (9)-(10), the following relationships are obtained, since $0 < \omega < 1$:

$$A_t \approx \frac{\omega}{2(1-\omega)} = \frac{k-1}{2}, \qquad B_t \approx \frac{(k-1)V_t}{2}, \qquad \text{where } V_t = \sum_{j=0}^{t-\tau-1} \frac{1}{k}\left(1-\frac{1}{k}\right)^j (x_{t-j}-\mu)^2. \tag{11}$$

Thus, $2A_t \approx k-1$ and $B_t A_t^{-1} \approx V_t$. Hinging on a large history of past realizations of consumption growth, V_t is the state variable that estimates the volatility of consumption growth with declining weights $\frac{1}{k}\left(1-\frac{1}{k}\right)^j$.

Note that subjective uncertainty about consumption growth volatility does not vanish even with an arbitrarily large number of past observations, as emphasized in Weitzman (2007). Namely, the economy incorporates the aspect of non-ergodic learning.

With the aforementioned structural uncertainty and the posited Bayesian learning rule, the theoretical object of interest is the *predictive* distribution of X_{t+1} given *all* past observations:

$$g\left(X_{t+1}|X^t\right) \quad = \quad \int p\left(X_{t+1}|\theta_{t+1}, X^t\right) p\left(\theta_{t+1}|X^t\right) d\theta_{t+1} = \int p\left(X_{t+1}|\theta_{t+1}\right) p\left(\theta_{t+1}|X^t\right) d\theta_{t+1}. \tag{12}$$

Under the stated assumptions, the predictive distribution of consumption growth can be characterized as the fat-tailed Student-t distribution, with $2A_t$ degrees of freedom, represented by,

$$Y_{t+1} \sim t(2A_t), \qquad \text{where } Y_{t+1} \equiv \frac{X_{t+1}-\mu}{\sqrt{B_t A_t^{-1}}}, \tag{13}$$

with density

$$g(Y_{t+1}; 2A_t) = \frac{\Gamma[(2A_t+1)/2]}{\sqrt{\pi 2A_t}\,\Gamma[A_t]} \left(1+\frac{Y_{t+1}^2}{2A_t}\right)^{-(2A_t+1)/2}. \tag{14}$$

The Student-t distribution has the property that its moment generating function $\int_{-\infty}^{\infty} e^{\lambda X_{t+1}} g\left(X_{t+1}\right) dX_{t+1}$ does not exist. Unfortunately, this means that the expectation of the marginal rate of substitution, $E_t\left(\beta e^{-\alpha X_{t+1}}\right)$, taken under the subjective distribution of consumption growth, is not finite in the Weitzman (2007) economy.

Equations (13) and (14) for the predictive Student-t distribution reiterate equation (24) in Weitzman (2007) in the case of a long past history, except that the number of degrees of freedom is $k-1$, as also in Shephard (1994), instead of k. That is, conditional on the knowledge of a large history of consumption growth, $2A_t \approx k-1$ and $B_t A_t^{-1} \approx V_t$ from (11) and therefore $Y_{t+1} \sim t(k-1)$.

Given the marginal rate of substitution in the economy, the gross riskfree return and the equity premium, as derived in Weitzman (2007) in logarithmic form, are,

$$\ln(R^f_{t+1}) \;=\; -\ln(\beta) \;-\; \ln\left(E_t\left(e^{-\alpha X_{t+1}}\right)\right), \tag{15}$$

$$\ln(E_t(R^e_{t+1})) - \ln(R^f_{t+1}) \;=\; \ln\left(\frac{E_t\left(e^{X_{t+1}}\right)}{\beta E_t\left(e^{(1-\alpha)X_{t+1}}\right)}\right) \;-\; \ln\left(\frac{1}{\beta E_t\left(e^{-\alpha X_{t+1}}\right)}\right). \tag{16}$$

It follows that, as Weitzman (2007, p. 1112) acknowledges, an economy that gives rise to a predictive Student-t distribution for consumption growth X_{t+1} has undesirable features. As implied by equations (15)-(16), the riskfree return and the equity premium are not well-defined objects, as the moment generating function of X_{t+1} is not finite. Furthermore, such a framework does not guarantee finite expected utility. In addition, the risk-neutral density (e.g., Harrison and Kreps (1979)) given by $g(Y_{t+1})\, e^{-\alpha X_{t+1}}/E_t\left(e^{-\alpha X_{t+1}}\right)$, which is central to pricing contingent claims, is not well-defined.

To keep expected utility finite, Weitzman (2007, p.1112-1113) proposes, but does not formalize, restricting structural uncertainty by confining consumption growth precision between some minimum and maximum levels:

$$\theta_t \in [\underline{\theta}, \overline{\theta}], \qquad \text{for all } t, \qquad 0 < \underline{\theta} < \overline{\theta} < \infty, \tag{17}$$

which imposes a support both for the prior and the posterior for precision given X^t. However, our contention is that the proposed mechanism in Weitzman (2007) for the evolution of the process θ_t *does not* guarantee that the process takes values on the desired interval $\left[\underline{\theta}, \overline{\theta}\right]$ over time, while preserving the conjugacy of the system.[2]

Going beyond Weitzman (2007), our incremental contribution is two-fold. First, while imposing a compact support on precision away from zero for all t, we (i) establish the finiteness of, and (ii) provide an expression for the moment generating function of consumption growth under structural uncertainty and Bayesian learning. To be able to link asset pricing quantities to risk aversion and structural uncertainty is the impetus for ensuring the finiteness of the moment generating function. Second, exploiting the operational version of the theory developed here, we are able to compute the equity premium and the riskfree return

[2] To see this point, recall the precision transition equation (26) in Weitzman (2007) which reads $\theta_t = \zeta_t\,\theta_{t-1}$, where ζ_t is the multiplicative shock that corresponds to $\frac{1}{\omega}\eta_t$ in our notation. According to the density specification for ζ_t, given in equation (28) in Weitzman (2007), the shock ζ_t takes values over the entire interval $[0, k/(k-1)]$ with positive probability. Therefore, under the assumption that the support of θ_{t-1} is the interval $\left[\underline{\theta}, \overline{\theta}\right]$, it follows that the support of θ_t is the interval $[0, \overline{\theta}k/(k-1)]$. In other words, the specification in Weitzman (2007) does not guarantee that the precision process stays within $\left[\underline{\theta}, \overline{\theta}\right]$ over time.

through the moment generating function, a feature central to quantitative assessments. Such computations are not feasible within the setting of Weitzman (2007) due to the lack of finiteness of the moment generating function associated with the Student-t distribution.

3. A structural uncertainty asset pricing model under bounded volatility

This section presents the theoretical results used to address whether subjective expectations can help explain asset pricing puzzles. In Theorem 1, we operationalize the restriction that precision have compact support away from zero within a Bayesian learning framework and, in a key departure from Weitzman (2007), we develop the associated posterior density for the precision of consumption growth. Under this modification, Theorem 2 provides a closed form expression for the predictive density of consumption growth. Furthermore, the moment generating function of the predictive distribution of consumption growth is derived in Theorem 3.

3.1. Posterior density for precision in the truncated Gamma class: A new conjugacy result

To proceed, define ϑ_t to be the precision of consumption growth with finite support denoted by $\left[\underline{\vartheta},\overline{\vartheta}\right]$ with $\underline{\vartheta} > 0$. Bear in mind that ϑ_t is not to be confused with θ_t, as used in Shephard (1994) and Weitzman (2007) (see our equation (7)), since the latter follows a Gamma distribution and so its support is the entire positive real line. In other words, ϑ_t is the analogue to θ_t under the assumption of bounded support away from zero.

Under the following two assumptions, we establish the conjugacy of ϑ_{t+1} and ϑ_t in the class of doubly-truncated Gamma distributions with support $\left[\underline{\vartheta},\overline{\vartheta}\right]$ for all t.

Assumption 3' *For a given time t, the conditional distribution of precision ϑ_t given X^{t-1} is doubly-truncated Gamma with parameters A_{t-1} and B_{t-1} and truncation points $\underline{\vartheta}$ and $\overline{\vartheta}$ with density:*

$$p(\vartheta_t|X^{t-1}) \;\; \propto \;\; \vartheta_t^{A_{t-1}-1}\, e^{-B_{t-1}\vartheta_t}\, \mathbb{I}_{\left[\underline{\vartheta},\overline{\vartheta}\right]}(\vartheta_t), \tag{18}$$

where the support of the precision distribution is $\left[\underline{\vartheta},\overline{\vartheta}\right]$ and $\mathbb{I}_{[.,.]}(\cdot)$ denotes an indicator function. The distribution described in (18) is denoted by $TG(A_{t-1},B_{t-1};\underline{\vartheta},\overline{\vartheta})$.

Assumption 4' *For a given time t, the transition equation for precision* ϑ_{t+1} *is,*

$$\vartheta_{t+1} = \frac{1}{\omega}\,\delta_{t+1}\,\vartheta_t, \qquad \vartheta_t \in \left[\underline{\vartheta}, \overline{\vartheta}\right], \qquad 0 < \omega < 1. \tag{19}$$

The conditional distribution of the multiplicative shock δ_{t+1}*, given* ϑ_t *and the history* X^t*, is specified in equation (A1) of the Appendix.*

Theorem 1 *Suppose that, for a given time t, (i) Assumption 3' regarding the precision* ϑ_t*, and (ii) Assumption 4' regarding the multiplicative shock* δ_{t+1} *are satisfied. Then, the posterior distribution of precision* ϑ_{t+1} *given* X^t *is also* $TG(A_t, B_t; \underline{\vartheta}, \overline{\vartheta})$ *where,*

$$A_t = \omega\left(A_{t-1} + \frac{1}{2}\right), \qquad B_t = \omega\left(B_{t-1} + \frac{(x_t - \mu)^2}{2}\right). \tag{20}$$

Proof: See the Appendix. □

What we have derived in Theorem 1 is an exact result on the posterior distribution of precision ϑ_{t+1} given X^t, which is doubly-truncated Gamma, with truncation points $\underline{\vartheta}$ and $\overline{\vartheta}$ (see Coffey and Muller (2000)). An explicit characterization of the posterior distribution is imperative to developing a predictive distribution of consumption growth. In particular, our approach formalizes a predictive density framework that prevents extreme beliefs about consumption growth volatility to dominate the analysis. Corresponding to the truncation points of consumption growth volatility defined by

$$\underline{\sigma} \equiv \frac{1}{\sqrt{\overline{\vartheta}}} \qquad \text{and} \qquad \overline{\sigma} \equiv \frac{1}{\sqrt{\underline{\vartheta}}}, \tag{21}$$

we have $\sigma_t \in [\underline{\sigma}, \overline{\sigma}]$ for all t.

In our theoretical model, Assumptions 3' and 4' replace Assumptions 3 and 4 of Section 2. When the support of consumption growth precision is enlarged to accommodate $\underline{\vartheta} \to 0$ and $\overline{\vartheta} \to \infty$, we obtain, as a limiting case, the setting of Weitzman (2007).

From a broader economic perspective, the structure of beliefs posited in (18) and (19) is essential to ensuring finite expected utility. Paramount for asset pricing formulations, the finiteness of the expected utility ensures a well-posed marginal utility function (e.g., Duffie (1992)). An alternative modeling approach that would guarantee finite expected utility is to truncate the support of the consumption growth distribution. However, this approach is hampered by the lack of a closed-form expression for the predictive density of consumption growth through Bayesian methods. Moreover, and of greater economic relevance, recall the argument in Weitzman (2007) that fat tails of the predictive distribution of consumption growth, as implied by structural uncertainty, can substantially bear on asset pricing quantities. Eliminating the tails, by truncating the support of the consumption growth distribution, would diminish the impact of structural uncertainty. Thus, our model overcomes the shortcomings of the alternative approach, while preserving finite expected utility and, at the same time, generating fat-tailed predictive consumption growth distributions.

Theorem 1 shows that when the conditional distribution ϑ_t given X^{t-1} is truncated Gamma, then the conditional distribution of ϑ_{t+1} given X^t is also truncated Gamma. The distribution of the multiplicative shock δ_{t+1} in (19) must consequently be outside of the beta distribution class which is defined over $(0,1)$, as in Shephard (1994) and Weitzman (2007). Our choice of the multiplicative shock distribution is designed to preserve the conjugacy of the prior and the posterior to be in the truncated Gamma class. Note that equation (A1) of the Appendix implies that δ_{t+1} is correlated with ϑ_t in our model.[3]

Equation (20) reveals that the evolution of A_t and B_t is characterized by the same recursion as in Weitzman (2007), where the precision process takes values over the entire positive real line. Nevertheless, the distinction is that, in order to make the theory operational, we propose a precision process that maintains bounded support away from zero which, as we illustrate shortly, is crucial for drawing asset pricing implications.

We wish to stress that, as in the framework of Weitzman (2007), the aspect of non-ergodic learning is maintained in our model, in the sense that, even for an arbitrarily long history of past data, uncertainty about consumption growth volatility remains influential. This is a consequence of the fact that, in both settings, the Bayesian agent is learning about the *time-varying* quantity of interest, namely the precision of consumption growth. Moving backward recursively and assuming that a long past history of length τ is available in (20), we arrive at:

$$A_t \approx \frac{k-1}{2}, \qquad B_t \approx \frac{(k-1)V_t}{2}, \qquad \text{where } V_t = \sum_{j=0}^{t-\tau-1} \frac{1}{k}\left(1-\frac{1}{k}\right)^j (x_{t-j}-\mu)^2. \tag{22}$$

[3] One aspect of the transition dynamics in (19) deserves further discussion. Shephard (1994) argues that when θ_t is Gamma distributed with support $(0,\infty)$, $\theta_T \to 0$ almost surely, as $T \to \infty$. To circumvent the undesirable feature that consumption growth volatility eventually becomes infinite, he modifies, using an appropriate dampening constant r_t, the transition equation to $\theta_t = e^{r_t}\eta_t\theta_{t-1}$ to prevent θ_t from reaching zero. Restricting precision ϑ_t on $[\underline{\vartheta},\overline{\vartheta}]$, we prevent consumption growth volatility from becoming arbitrarily large as the economy evolves. Although not reported, our simulations suggest that, for plausible values of ω, it takes several hundred years for the consumption growth volatility to become large even when precision is not truncated and takes values over $(0,\infty)$. Nevertheless, given the observed consumption growth history, precision paths associated with eventually infinite consumption growth volatility will be ignored by the Bayesian investor in his posterior calculations. For these reasons, we keep matters simple and do not introduce a dampening factor in the transition equation (19).

With such an understanding, we henceforth set $2A_t = k-1$ and $B_t A_t^{-1} = V_t$ in our implementations.

[Fig. 1 about here.]

Given that the support of the precision process is the bounded interval $\left[\underline{\vartheta}, \overline{\vartheta}\right]$, with $\underline{\vartheta} > 0$, the relevant question to ask is whether we are excluding priors that are economically relevant, and whether the proposed truncation points are overly restrictive. To address such a concern, we simulate 10 million draws from the posterior distribution θ_{t+1} given X^t, when no truncation is imposed, which is Gamma(A_t, B_t). For this exercise, we set the volatility of consumption growth to $\sqrt{V} = 2\%$ and the effective sample size to $k = 50$. Thus, $A = (k-1)/2 = 24.5$ and $B = (k-1)V/2 = 0.0098$. Accordingly, we present in Figure 1, the distribution of both the precision θ_{t+1} (see Panel A) and the volatility of consumption growth $\sigma_{t+1} = 1/\sqrt{\theta_{t+1}}$ (see Panel B). The message is that even with 10 million draws, it is not feasible to generate values of σ_{t+1} above 4.5% and θ_{t+1} below 735. This is an indication that the two precision distributions, without truncation and with truncation, are virtually indistinguishable. To substantiate the claim from a different angle, we calculate Prob$(\sigma_{t+1} \geq \overline{\sigma})$, for values of $\overline{\sigma}$ ranging from 5% to 500%, and report the probabilities:

$\overline{\sigma} = 1/\sqrt{\underline{\theta}}$	5%	10%	20%	50%	100%	200%	500%
$k = 50$	2.60E-12	7.71E-26	5.38E-45	1.07E-59	1.96E-74	3.50E-89	1.11E-108
$k = 100$	3.08E-23	1.62E-50	3.32E-89	6.62E-119	1.11E-148	1.77E-178	7.15E-218

With parameters corresponding to historical consumption growth, the calculations show that the probabilities are extremely small and decline rapidly in the tails. With $k = 100$, the probabilities Prob$(\sigma_{t+1} \geq \overline{\sigma})$ are even smaller. In conclusion, given the parameters specified above, imposing a truncation limit $\overline{\sigma}$ above 5% on the support of the distribution of consumption growth volatility does not materially alter the prior distribution.

3.2. Predictive density of consumption growth under bounded volatility

Before providing the analytical form of the predictive density of consumption growth in the next theorem, we invoke the following assumption, which is the analogue to Assumption 2 in Section 2.

Assumption 2' *Conditional on precision ϑ_{t+1}, the consumption growth, X_{t+1}, is Normally distributed with $X_{t+1} \sim N\left(\mu, \frac{1}{\vartheta_{t+1}}\right)$, where the mean μ is a known constant and the precision ϑ_{t+1} has support $[\underline{\vartheta}, \overline{\vartheta}]$.*

Theorem 2 *Under Assumptions 2', 3' and 4', the probability density function of the standardized consumption growth* $Y_{t+1} = \frac{X_{t+1}-\mu}{\sqrt{B_t A_t^{-1}}}$, *given the history* X^t, *is:*

$$g^{\mathrm{DT}}(Y_{t+1};\nu_t,\underline{\xi}_t,\overline{\xi}_t) = \frac{\gamma\left[\frac{\nu_t+1}{2},\frac{1}{2}\overline{\xi}_t\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)\right] - \gamma\left[\frac{\nu_t+1}{2},\frac{1}{2}\underline{\xi}_t\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)\right]}{\sqrt{\pi\nu_t}\left(\gamma\left[\nu_t/2,\overline{\xi}_t/2\right]-\gamma\left[\nu_t/2,\underline{\xi}_t/2\right]\right)\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)^{\frac{\nu_t+1}{2}}}, \tag{23}$$

where ν_t *represents the degrees of freedom and* $\underline{\xi}_t,\overline{\xi}_t$ *represent the truncation parameters, given by:*

$$\nu_t = 2A_t, \qquad \underline{\xi}_t = 2\underline{\vartheta}B_t, \qquad \overline{\xi}_t = 2\overline{\vartheta}B_t, \tag{24}$$

and $\gamma[a,j] = \int_0^j u^{a-1}e^{-u}du$ *is the lower incomplete Gamma function for* $a>0$ *and* $j>0$. *For reasons not yet articulated, we call the predictive density in (23) the dampened t distribution.*

Proof: See the Appendix. □

Suppressing the time subscripts for brevity, notice that when $\underline{\xi}\to 0$ and $\overline{\xi}\to\infty$, the density $g^{\mathrm{DT}}(Y;\nu,\underline{\xi},\overline{\xi})$ approaches the Student-*t* density with ν degrees of freedom. The Normal distribution with mean 0 and variance $\frac{\nu}{\xi}$ is obtained as a limiting case, when $\underline{\xi}\uparrow\xi$ and $\overline{\xi}\downarrow\xi$ with $\xi>0$.

An alternative parametrization of density $g^{\mathrm{DT}}(Y_{t+1};\nu,\underline{\xi},\overline{\xi})$ can be obtained by defining,

$$\ln\left(\underline{\xi}\right) = \ln(\nu) - \frac{1}{I}, \qquad \ln\left(\overline{\xi}\right) = \ln(\nu) + \frac{1}{I}, \tag{25}$$

where I is a positive constant.

Case 1 *Based on the mapping (25), when* $I\to 0$, *we have* $\underline{\xi}\to 0$ *and* $\overline{\xi}\to\infty$. *Accordingly, as* $I\to 0$ *the distribution in (23) approaches the Student-t distribution (Geweke (2001) and Weitzman (2007)).*

Case 2 *Correspondingly, from (25), as* $I\to\infty$ *it follows that* $\underline{\xi}\uparrow\nu$ *and* $\overline{\xi}\downarrow\nu$. *Thus, as* $I\to\infty$ *the distribution in (23) converges to the standard Normal distribution.*

Case 3 *Imposing* $2A_t = k-1$ *and* $B_tA_t^{-1} = V_t$ *in (23) of Theorem 2 yields:*

$$g^{\mathrm{DT}}(Y_{t+1};k-1,\underline{\vartheta}(k-1)V_t,\overline{\vartheta}(k-1)V_t), \tag{26}$$

which corresponds to the case of a large number of past observations on consumption growth from (22).

Our characterizations imply that the three-parameter density family $g^{\mathrm{DT}}(Y_{t+1};\nu_t,\underline{\xi}_t,\overline{\xi}_t)$ lies in between the Normal distribution and the Student-t distribution. But, is the predictive density (23) closer to the heavy-tailed Student-t distribution (12) or the thin-tailed Normal distribution? A parsimonious way to capture the difference between (14) and (23) is to compute the logarithm of the ratio of the two densities when a large past history of consumption growth is available (i.e., from Case 3, $2A_t = k-1$, $B_t A_t^{-1} = V_t$):

$$\Upsilon_{t+1} \equiv \ln\left(\frac{g(X_{t+1};k-1)}{g^{\mathrm{DT}}(X_{t+1};k-1,\underline{\vartheta}(k-1)V_t,\overline{\vartheta}(k-1)V_t)}\right), \qquad X_{t+1} = \mu + \sqrt{V_t}\,Y_{t+1}. \tag{27}$$

Figure 2 plots Υ_{t+1} versus consumption growth X_{t+1}, taking $\mu = 2\%$, $\sqrt{V} = 2\%$, $k = 50$, $\underline{\sigma} = 1/\sqrt{\overline{\vartheta}} = 0.01\%$ and $\overline{\sigma} = 1/\sqrt{\underline{\vartheta}} = 500\%$. The flat region in the bowl-shaped curve has the interpretation that the two predictive densities are equivalent over consumption growth of roughly $\pm$1,700%. Even when differences are observed in the tails, the distinction is of the order of 10^{-7}. Intuitively, this slight departure occurs as the Student-t distribution has raw moments up to order $\ell < k-1$, whereas the density in (23) has bounded algebraic moments of all orders (shown next in Theorem 3). Given the inherent closeness of the two predictive distributions, the density in (23), is labeled as the dampened t distribution. The source of the dampening is the difference between the incomplete Gamma function evaluated at the upper and the lower truncation points, as seen from (23). It is the incomplete Gamma trimming that ensures the finiteness of the moment generating function.[4]

[Fig. 2 about here.]

There is an alternative way to construct the predictive density (23). Suppose Z is a standard Normal variate and W follows a χ^2 distribution with ν degrees of freedom, truncated on the interval $[\underline{\xi},\overline{\xi}]$, where $0 < \underline{\xi} < \overline{\xi} < \infty$. Then, it is shown via Lemma 2 in the Appendix that $\frac{Z}{\sqrt{W/\nu}} \equiv Y$ obeys the dampened t distribution with density (23). It is this construction that we exploit to analytically derive the moment generating function of Y_{t+1} and show that it exists and is well-defined.

[4]Shephard (1994) has also considered the generalized error distribution as a building block for the conditional distribution of $X_t|\vartheta_t$. Such a treatment, along with a generalized Gamma distributed precision process, induces a predictive distribution that is generalized Student-t. However, the moment generating function still does not exist and therefore hinders the usefulness of the model for examining asset pricing puzzles. To our knowledge, when (i) $X_t|\vartheta_t$ follows a generalized error distribution, and (ii) the precision process follows (19) and lives on a compact support away from zero, the predictive distribution is unamenable to analytical characterization through Bayesian methods.

3.3. Finiteness of the moment generating function of consumption growth

Our innovation is that asset pricing quantities can be evaluated using the moment generating function of $g^{\mathrm{DT}}(Y_{t+1};\nu_t,\underline{\xi}_t,\overline{\xi}_t)$, a task that cannot be accomplished under the Student-t distribution. Suppressing time subscripts on ν_t, $\underline{\xi}_t$, $\overline{\xi}_t$, A_t and B_t we now state the next important result.

Theorem 3 *For the predictive density of* $Y_{t+1}=\frac{X_{t+1}-\mu}{\sqrt{BA^{-1}}}$ *specified in (23), the moment generating function is:*

$$\Psi_Y[\lambda] \;\equiv\; E_t\left(e^{\lambda Y_{t+1}}\right)=\int_{-\infty}^{\infty} e^{\lambda Y_{t+1}}\, g^{\mathrm{DT}}(Y_{t+1};\nu,\underline{\xi},\overline{\xi})\,dY_{t+1}, \tag{28}$$

$$= \frac{1}{c\left[\nu,\underline{\xi},\overline{\xi}\right]}\int_{\underline{\xi}}^{\overline{\xi}} e^{\frac{\lambda^2\nu}{2w}}\, w^{\frac{\nu}{2}-1}\, e^{-\frac{w}{2}}\,dw, \tag{29}$$

where $c\left[\nu,\underline{\xi},\overline{\xi}\right]=2^{\nu/2}\left(\gamma\left[\nu/2,\overline{\xi}/2\right]-\gamma\left[\nu/2,\underline{\xi}/2\right]\right)$, $\nu=2A$, $\underline{\xi}=2\underline{\vartheta}B$, *and* $\overline{\xi}=2\overline{\vartheta}B$. *The moment generating function of consumption growth* $X_{t+1}=\mu+\sqrt{BA^{-1}}\,Y_{t+1}$ *is then,*

$$\Psi_X[\lambda] \;=\; e^{\mu\lambda}\,\Psi_Y[\lambda\sqrt{BA^{-1}}]. \tag{30}$$

All odd-order moments of Y_{t+1} *are equal to 0. The even-order moments of* Y_{t+1} *are given by* $E_t\left(Y_{t+1}^{2\phi}\right)=\frac{(2\phi)!}{\phi!}\left(\frac{\nu}{2}\right)^{\phi}\frac{1}{c[\nu,\underline{\xi},\overline{\xi}]}\int_{\underline{\xi}}^{\overline{\xi}} w^{\frac{\nu}{2}-\phi-1}e^{-\frac{w}{2}}\,dw<\infty$, *for any positive integer* ϕ.

Proof: See the Appendix. □

Unlike the Student-t distribution, which does not possess a finite moment generating function, the dampened t distribution is shown to have a finite moment generating function for $0<\underline{\xi}<\overline{\xi}<\infty$.

At the crux of Theorem 3 is the statement that the moment generating function is the integral of a continuous function over a compact support, and consequently is finite (e.g., Lukacs (1960)). Well-defined preferences and marginal rate of substitution in the economy hinge on the finiteness of $\Psi_X[\lambda]$.

4. Can structural uncertainty resolve asset pricing puzzles?

Under structural parameter uncertainty, where the consumption growth volatility σ_t has support $[\underline{\sigma},\overline{\sigma}]$, and Bayesian learning, the operational version of the theory shows that the primitive distribution for consumption growth, which is Normal, gets transformed to a dampened t distribution. Tail thickening of the predictive consumption growth distribution induced by structural parameter uncertainty can, in principle, exert a substantial impact on asset pricing quantities through the marginal rate of substitution.

To see the restrictions imposed by the present theory, we note that the riskfree return and the equity risk premium can both be obtained from the moment generating function of consumption growth, $\Psi_X[\lambda]$. Guided by equations (15) and (16), we obtain:

$$\ln(R^f_{t+1}) \;=\; -\ln(\beta) \;-\; \ln\left(\Psi_X[-\alpha]\right), \tag{31}$$

$$\ln(E_t(R^e_{t+1})) - \ln(R^f_{t+1}) \;=\; \ln\left(\Psi_X[1]\right) \;-\; \ln\left(\Psi_X[1-\alpha]\right) \;+\; \ln\left(\Psi_X[-\alpha]\right), \tag{32}$$

where $\Psi_X[\lambda]$ is presented in equation (30) of Theorem 3. The special feature of this economy with subjective expectations is that it supports a finite moment generating function of X_{t+1}. Therefore, the contribution of our approach lies in that asset pricing quantities are computable, which is essential for conducting quantitative analysis.

Economic theory now suggests that the equity risk premium and the riskfree return are both linked to the parameters of the dampened t distribution. In particular, the asset pricing quantities are determined by (A_t, B_t, $\underline{\xi}_t$, $\overline{\xi}_t$), which jointly control the mean, volatility, and truncation points of the belief process. Accounting for subjective expectations expands the traditional consumption-based asset pricing model.

The tractability of our construction allows us to pose and answer outstanding questions of broad interest: (1) What is the quantitative impact of subjective expectations on asset pricing? (2) What level of a priori structural uncertainty is required for resolving asset pricing puzzles? (3) In what sense does structural uncertainty quantitatively map into consumption disaster fears? Operationalizing the theory, as established in Theorems 1 through 3, enables us to address these questions, a task not feasible in the environment of Weitzman (2007).

4.1. Plausible structural uncertainty does not profoundly affect asset pricing quantities

In our quantitative assessments, we rely on the assumption of a long history of past observations which yields $2A_t = k-1$ and $B_t A_t^{-1} = V_t$. We use benchmark values that are in line with the existing literature (e.g., Mehra and Prescott (1985), Barro (2006), and Weitzman (2007)):

$$\mu = 2\%, \qquad \sqrt{V} = 2\%, \qquad \beta = 0.98, \qquad k \in \{10, 30, 50\}, \qquad \underline{\sigma} = 1/\sqrt{\overline{\vartheta}} = 0.1\%.$$

The lower bound $\underline{\sigma} = 1/\sqrt{\overline{\vartheta}}$ is not essential, in the sense that asset pricing quantities are not affected by reasonable changes in $\underline{\sigma}$, and is, therefore, held fixed. The effective sample size k is selected to cover a range of reasonable values. In terms of subjective expectations, k measures how much confidence the investor places in the variance estimate, and at the same time it controls the degrees of freedom, and therefore the tail thickness, of the predictive distribution of consumption growth.

Table 1 first assesses the quantitative impact of subjective expectations on $\ln(R^f_{t+1})$ and $\ln(E_t(R^e_{t+1})) - \ln(R^f_{t+1})$ by adopting risk aversion $\alpha \in \{2, 3, 5, 10\}$. Considering that the average historical volatility of consumption growth is 2%, the maximum allowable level of consumption growth volatility is fixed at $\overline{\sigma} = 1/\sqrt{\underline{\vartheta}} = 50\%$.

According to the evidence in Table 1, the effect of subjective expectations on asset pricing quantities is barely noticeable over reasonable levels of $\underline{\sigma}$, $\overline{\sigma}$, and α. Moreover, the riskfree return and the equity premium are almost identical between the case $k = 50$ and the full information Normal distribution benchmark. Finally, lower levels of k tend to decrease both the riskfree return and the equity return, more so for the riskfree return.

We also find that the response of the riskfree return and the equity premium to $\overline{\sigma}$ is flat up to $\overline{\sigma} = 190\%$ for $\alpha \in (2, 10)$ with $k = 50$, which we consider to be a reasonable candidate for the effective sample size. To save on space, all results are henceforth reported for $k = 50$.

[Fig. 3 about here.]

On the other hand, when excessive levels of structural uncertainty are permitted to influence asset pricing, both $\ln(R^f_{t+1})$ and $\ln(E_t(R^e_{t+1})) - \ln(R^f_{t+1})$ change rapidly with small shifts in $\overline{\sigma}$ as seen from Figure 3. Possibly rectifying the shortcomings of the consumption-based asset pricing model, the presence of structural uncertainty now sharply lowers the riskfree return and raises the equity risk premium. Consider $\alpha = 3$ in Figure 3 to illustrate the main ideas. At a value of $\overline{\sigma} = 762.50\%$, $\ln(R^f_{t+1}) = 6.68\%$ and $\ln(E_t(R^e_{t+1})) - \ln(R^f_{t+1}) = 1.28\%$, while at $\overline{\sigma} = 766.50\%$ the counterpart values are -5.28% and 13.24%. The model behavior is knife-edge sensitive at large levels of structural uncertainty, regardless of the magnitude of risk aversion. Specifically, as the support of the precision distribution is enlarged, the posterior converges to the Student-t distribution, and, as a result, the equity premium (riskfree return) can admit arbitrarily large (small) values.

In summary, Table 1 and Figure 3 together give rise to two inferences about the implications of subjective expectations for asset pricing. First, the particularly parameterized structural uncertainty model operationalized here does not produce a significant change in the asset pricing quantities, unless excessive *a priori* uncertainty is entertained. Second, the model-implied magnitudes of the asset pricing quantities fall short of their historical counterparts, even though the dampened t distribution is virtually indistinguishable from the Student-t distribution. The conclusions remain robust even when consumption growth volatility is set to $\sqrt{V} = 3.5\%$ as in Barro (2006). For instance, when $k = 10$ and $\alpha = 10$, the model equity premium is 1.78% and the riskfree return is 13.56%.

4.2. It is still challenging to explain the equity premium of 6% and the riskfree return of 1%

Results presented in Table 2 are at the heart of asset pricing puzzles. Focus first on Panel A, where we numerically search over the maximum level of consumption growth volatility, $\overline{\sigma}$, required to match the 6% average equity risk premium. When α is restricted between 2 and 10, the range of $\overline{\sigma}$ required to match the equity risk premium varies between 199.3% and 1,193.5%. Most financial economists hold the view that α should lie somewhere between 2 and 10 (e.g., Mehra and Prescott (1985) and Kocherlakota (1996)).

The inability of the model to match economic theory with data is also implicit in the corresponding values of riskfree return when the equity risk premium is matched to 6%. In particular, the simultaneous justification of asset returns data requires levels of $\overline{\sigma}$ between 765.4% and 1,193.5% when α is confined between between 2 and 3. In sum, a mismatch still exists when asset pricing theory incorporating subjective expectations is applied to data from financial markets.

Panel B of Table 2 conducts the analog exercise where we search over the maximum level of consumption growth volatility, $\overline{\sigma}$, to match the riskfree return of 1% for a given α. The deficiencies of the model manifest themselves from a different perspective: declining values of $\overline{\sigma}$ that coincide with increasing α also imply increasing equity risk premium in order to maintain the target riskfree return of 1%.

To reiterate, a reasonably parameterized asset pricing model under subjective uncertainty about the volatility of consumption growth faces a formidable hurdle in explaining the equity risk premium and the real riskfree return simultaneously. The level of structural uncertainty required to resolve asset pricing puzzles in the model is, in our view, beyond any reasonable norm.

4.3. Priors yielding similar disaster probabilities have vastly different asset pricing implications

How does the maximum volatility level $\overline{\sigma} = 1/\sqrt{\underline{\vartheta}}$ relate to probabilities of consumption disasters, as mentioned in Weitzman (2007)? Consider the possibility of a rare disaster x^{rare}, which represents a deviation from mean consumption growth in multiples of volatility:

$$x^{\text{rare}} \;=\; \mu \;-\; h\sqrt{V}, \tag{33}$$

where the multiple $h \in \{3,4,5,6\}$, $\mu = 2\%$, and $\sqrt{V} = 2\%$. For instance, a 5-sigma downside event corresponds to $h = 5$ and implies a 8% drop in the level of consumption.

Our model setup facilitates computation of rare event probabilities. Specifically, it is shown in the Appendix that,

$$\text{Prob}\left(X_{t+1} \leq x^{\text{rare}}\right) = \text{Prob}\left(Y_{t+1} \leq -h\right) = \frac{1}{c\left[\nu, \underline{\xi}, \overline{\xi}\right]} \int_{\underline{\xi}}^{\overline{\xi}} N\left(-h\sqrt{w/\nu}\right) w^{\frac{\nu}{2}-1} e^{-\frac{w}{2}}\, dw, \tag{34}$$

where $\nu = k-1$ is the degrees of freedom, $\underline{\xi} = \frac{(k-1)V}{\overline{\sigma}^2}$, $\overline{\xi} = \frac{(k-1)V}{\underline{\sigma}^2}$, and $N(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-u^2/2}\, du$ is the standard Normal cumulative distribution function.

Table 3 presents the probabilities of rare disasters as a function of the maximum level of consumption growth volatility. Specifically, we present the probabilities corresponding to the Normal distribution, the dampened t distribution, and the Student-t distribution.

The crucial attribute to note is that the maximum allowable level of consumption growth volatility does not materially affect the probabilities of rare disasters in the model. Explaining the equity premium in the model with reasonable α requires a maximum level of consumption growth volatility in excess of 765%, and these implausible prior beliefs have in common disaster probabilities that are indistinguishable from those implied by more realistic priors.

Note that, at the same time, as the analysis of the previous two subsections reveals, the two different sets of prior beliefs, involving low and very high truncation levels $\overline{\sigma}$, have diametrically opposite asset pricing implications: one contradicting, and the other asserting, the antipuzzle view. Our explicit quantification of the equity premium, along with the rare disaster probabilities, possibly provides a reason to challenge the argument offered in Weitzman (2007) that large levels of equity premium are somehow compensation for structural uncertainty and consumption disaster fears, under the postulated asset pricing framework.

5. Conclusions

In this paper, we contribute to the literature on asset pricing models with Bayesian learning by providing an operational version of the structural uncertainty approach that features a compact support for consumption growth volatility. Our setting incorporates realistic priors about consumption growth volatility and has substantial economic consequences. First, it guarantees finite asset pricing quantities. Second, it allows characterization of the Bayesian predictive density for consumption growth that is virtually indistinguishable from the heavy-tailed Student-t distribution, but importantly possesses a finite moment generating function. The resulting asset pricing model with subjective expectations yields finite expected utility and tractable expressions for equity premium and riskfree return. In this setting, embedding structural uncertainty induces a heavy-tailed predictive distribution for consumption growth without introducing jumps in the consumption growth process.

The availability of the moment generating function for the predictive distribution of consumption growth renders quantitative evaluation feasible. Applying economic theory in the context of subjective expectations reveals that explaining the historical equity premium and riskfree return requires extreme levels of a priori uncertainty about consumption growth volatility. Furthermore, these implausible prior beliefs give rise to disaster probabilities that almost coincide with those implied by more realistic priors. However, the two sets of prior beliefs have diametrically opposite asset pricing implications: one asserting, and the other contradicting, the antipuzzle view favored by Weitzman (2007).

The scope of the approach developed here can be expanded along several directions to enhance our understanding of how subjective expectations impact asset prices. First, the asset pricing framework can be extended to investigate higher-dimensional structural uncertainty, where the agent faces uncertainty not only about consumption growth volatility, but also about other aspects of the distribution. Second, the effect of more realistic consumption growth dynamics can be explored by accommodating serial correlation in the specification of the consumption growth process (e.g., Barsky and DeLong (1993), and Tsionas (2005)). Finally, further insights could be gained by considering richer pricing kernel specifications such as those resulting from alternative preference structures (e.g., Constantinides (1990), Epstein and Zin (1991), and Campbell and Cochrane (1999)). These extensions are left for future work.

Appendix: Proof of Results

Distribution of the multiplicative shock δ_{t+1} in equation (19): What we specify first is the distribution of the multiplicative shock δ_{t+1} that guarantees the conjugacy of the system while ensuring that the support of the precision process ϑ_t remains $[\underline{\vartheta}, \overline{\vartheta}]$ through time.

To appreciate our generalization, let us momentarily return to the Weitzman (2007) model where the precision process θ_t evolves according to the transition equation $\theta_{t+1} = \frac{1}{\omega}\eta_{t+1}\theta_t$ as in (6). Importantly, when θ_t satisfies $\theta_t|X^t \sim \text{Gamma}(a_t, b_t)$ and the multiplicative shock η_{t+1} follows a $\text{Beta}(\omega a_t, (1-\omega)\, a_t)$ distribution, then $\theta_{t+1}|X^t \sim \text{Gamma}(A_t, B_t)$ where $A_t = \omega a_t$ and $B_t = \omega b_t$. Such a structure preserves conjugacy of the system through time.

In order to provide the details of the specification of the multiplicative shock δ_{t+1} in (19), we introduce some additional notation. Conditionally on X^t, let $f_t(\cdot)$ and $F_t(\cdot)$ denote the probability density function (pdf) and cumulative distribution function (cdf) of θ_t, respectively. Similarly, given X^t, let $g_t(\cdot)$ and $G_t(\cdot)$ denote the pdf and cdf of θ_{t+1}, respectively. That is, $f_t(\cdot)$ is the $\text{Gamma}(a_t, b_t)$ density and $g_t(\cdot)$ is the $\text{Gamma}(\omega a_t, \omega b_t)$ density. Furthermore, let $h_t(\theta_t, \theta_{t+1})$ denote the joint pdf of (θ_t, θ_{t+1}), conditionally on X^t. This implies that the two marginal distributions associated with the joint density h_t are f_t and g_t, respectively.

Now denote by $\mathfrak{f}_t$ and $\mathfrak{F}_t$ the pdf and cdf of the truncated Gamma $\text{TG}(a_t, b_t; \underline{\vartheta}, \overline{\vartheta})$ distribution, respectively. Furthermore, denote by $\mathfrak{g}_t$ and $\mathfrak{G}_t$ the pdf and cdf of the truncated Gamma $\text{TG}(\omega a_t, \omega b_t; \underline{\vartheta}, \overline{\vartheta})$ distribution, respectively. When $\underline{\vartheta} \to 0$ and $\overline{\vartheta} \to \infty$, the TG density $\mathfrak{f}_t$ converges to the Gamma density f_t, and the TG density $\mathfrak{g}_t$ converges to the Gamma density g_t.

The distribution of the multiplicative shock δ_{t+1} is then defined as follows. Given ϑ_t and X^t, the conditional pdf of δ_{t+1} is:

$$p(\delta_{t+1}|\vartheta_t, X^t) = \frac{\vartheta_t}{\omega} \cdot \mathfrak{g}_t\left(\frac{1}{\omega}\vartheta_t\delta_{t+1}\right) \cdot \frac{h_t\left(F_t^{-1}(\mathfrak{F}_t(\vartheta_t)), G_t^{-1}(\mathfrak{G}_t(\frac{1}{\omega}\vartheta_t\delta_{t+1}))\right)}{f_t\left(F_t^{-1}(\mathfrak{F}_t(\vartheta_t))\right) \cdot g_t\left(G_t^{-1}(\mathfrak{G}_t(\frac{1}{\omega}\vartheta_t\delta_{t+1}))\right)}. \tag{A1}$$

While equation (A1) might seem complicated, it is a generalization of the Beta distribution used for the multiplicative shock in the case of the (untruncated) Gamma precision process θ_t. Indeed, in the limit $\underline{\vartheta} \to 0$ and $\overline{\vartheta} \to \infty$, the conditional density (A1) converges to the $\text{Beta}(\omega a_t, (1-\omega)a_t)$ distribution, namely the distribution of the multiplicative shock η_{t+1} in Weitzman (2007). To establish this property, we build on the following Lemma:

Lemma 1 *The joint pdf* $h_t(\cdot,\cdot)$ *of* (θ_t,θ_{t+1})*, conditionally on* X^t*, is given by*

$$h_t(\theta_t,\theta_{t+1}) = \frac{\omega}{\Gamma[\omega a_t]\,\Gamma[(1-\omega)\,a_t]} b_t^{a_t} e^{-b_t\theta_t} \cdot (\omega\,\theta_{t+1})^{\omega\,a_t-1}\,(\theta_t-\omega\,\theta_{t+1})^{(1-\omega)\,a_t-1}, \tag{A2}$$

where $0<\theta_{t+1}<\frac{\theta_t}{\omega}$.

Proof of Lemma 1. The pdf of θ_t, given X^t, is the Gamma(a_t,b_t) density, i.e., $p(\theta_t|X^t)=\frac{b_t^{a_t}}{\Gamma[a_t]}\theta_t^{a_t-1}e^{-b_t\theta_t}$. Since $\theta_{t+1}=\frac{1}{\omega}\eta_{t+1}\theta_t$ and $\eta_{t+1}\sim$ Beta$(\omega a_t,(1-\omega)a_t)$ with density $k_t(\eta)=\frac{1}{B[\omega a_t,(1-\omega)a_t]}\eta^{\omega a_t-1}(1-\eta)^{(1-\omega)a_t-1}$, it follows that:

$$p(\theta_{t+1}|\theta_t,X^t)=\frac{\omega}{\theta_t}k_t\left(\frac{\omega}{\theta_t}\theta_{t+1}\right)=\frac{\omega}{\theta_t}\frac{1}{B[\omega a_t,(1-\omega)a_t]}\left(\frac{\omega}{\theta_t}\theta_{t+1}\right)^{\omega a_t-1}\left(1-\frac{\omega}{\theta_t}\theta_{t+1}\right)^{(1-\omega)a_t-1}. \tag{A3}$$

By noting that $h_t\left(\theta_t,\theta_{t+1}\right)=p(\theta_t,\theta_{t+1}|X^t)=p(\theta_{t+1}|\theta_t,X^t)p(\theta_t|X^t)$ proves the result. ■

In the limiting case with $\underline{\vartheta}\to 0$ and $\overline{\vartheta}\to\infty$, we have $\mathfrak{f}_t\to f_t$, $\mathfrak{F}_t\to F_t$, and $\mathfrak{g}_t\to g_t$, $\mathfrak{G}_t\to G_t$. Hence, the density $p(\delta_{t+1}|\vartheta_t)$ converges to $\frac{\vartheta_t}{\omega}\frac{h_t\left(\vartheta_t,\frac{1}{\omega}\vartheta_t\delta_{t+1}\right)}{f_t(\vartheta_t)}$. The density $f_t\left(\vartheta_t\right)$ is given by $f_t\left(\vartheta_t\right)=\frac{b_t^{a_t}}{\Gamma(a_t)}\vartheta_t^{a_t-1}e^{-b_t\vartheta_t}$ and Lemma 1 implies that

$$h_t\left(\vartheta_t,\frac{1}{\omega}\vartheta_t\delta_{t+1}\right)=\frac{\omega}{\Gamma[\omega a_t]\,\Gamma[(1-\omega)a_t]}b_t^{a_t}\vartheta_t^{a_t-2}e^{-b_t\vartheta_t}\cdot\delta_{t+1}^{\omega a_t-1}\left(1-\delta_{t+1}\right)^{(1-\omega)a_t-1}. \tag{A4}$$

Thus, as $\underline{\vartheta}\to 0$ and $\overline{\vartheta}\to\infty$, the limit of $p(\delta_{t+1}|\vartheta_t)$ is:

$$\frac{\vartheta_t}{\omega}\frac{h_t\left(\vartheta_t,\frac{1}{\omega}\vartheta_t\delta_{t+1}\right)}{f_t\left(\vartheta_t\right)}=\frac{\Gamma[a_t]}{\Gamma[\omega a_t]\,\Gamma[(1-\omega)a_t]}\delta_{t+1}^{\omega a_t-1}\left(1-\delta_{t+1}\right)^{(1-\omega)a_t-1}, \tag{A5}$$

which is simply the Beta$(\omega a_t,(1-\omega)a_t)$ density of the multiplicative shock η_{t+1} in Weitzman (2007). Thus, the choice of δ_{t+1}, ϑ_t and ϑ_{t+1} ensures the internal consistency of the system and yet maintains the conjugacy of ϑ_t and ϑ_{t+1}, as verified next.

Proof of Theorem 1 (Posterior distribution of ϑ_{t+1} given X^t): It is shown that the conditional distribution of ϑ_t, given X^t, is TG$(A_t,B_t;\underline{\vartheta},\overline{\vartheta})$ where $A_t=\omega a_t=\omega\left(A_{t-1}+\frac{1}{2}\right)$ and $B_t=\omega b_t=\omega\left(B_{t-1}+\frac{(X_t-\mu)^2}{2}\right)$.

First, observe that the conditional distribution of ϑ_t given X^t is given by

$$
\begin{aligned}
p(\vartheta_t|X^t) &\propto p(X^t,\vartheta_t) = p(X_t|\vartheta_t)\,p(\vartheta_t|X^{t-1}) \\
&\propto \sqrt{\vartheta_t}e^{-\frac{(X_t-\mu)^2\vartheta_t}{2}}\cdot\vartheta_t^{A_{t-1}-1}e^{-B_{t-1}\vartheta_t}\mathbb{I}_{[\underline{\vartheta},\overline{\vartheta}]}(\vartheta_t) \\
&\propto \vartheta_t^{a_t-1}e^{-b_t\vartheta_t}\mathbb{I}_{[\underline{\vartheta},\overline{\vartheta}]}(\vartheta_t)
\end{aligned}
$$

where $a_t = A_{t-1}+\frac{1}{2}$ and $b_t = B_{t-1}+\frac{(X_t-\mu)^2}{2}$. Hence, the conditional distribution of ϑ_t given X^t is $\mathrm{TG}(a_t,b_t;\underline{\vartheta},\overline{\vartheta})$.

In our notation, the density of ϑ_t given X^t is $\mathfrak{f}_t(\cdot)$. Given the conditional density (A1) of the multiplicative shock δ_{t+1}, it follows that the joint density of $(\vartheta_t,\delta_{t+1})$, conditionally on X^t, is

$$
\psi_t(\vartheta_t,\delta_{t+1}) = \frac{\vartheta_t}{\omega}\cdot\mathfrak{g}_t\left(\frac{1}{\omega}\vartheta_t\delta_{t+1}\right)\cdot\frac{h_t\left(F_t^{-1}(\mathfrak{F}_t(\vartheta_t)),G_t^{-1}(\mathfrak{G}_t(\frac{1}{\omega}\vartheta_t\delta_{t+1}))\right)}{f_t\left(F_t^{-1}(\mathfrak{F}_t(\vartheta_t))\right)\cdot g_t\left(G_t^{-1}(\mathfrak{G}_t(\frac{1}{\omega}\vartheta_t\delta_{t+1}))\right)}\mathfrak{f}_t(\vartheta_t). \tag{A6}
$$

Based on the posited transition equation of the (truncated) precision process, $\vartheta_{t+1} = \frac{1}{\omega}\delta_{t+1}\vartheta_t$, we can now exploit the Jacobian of the appropriate two-dimensional transformation. Specifically, if (X,Y) has joint pdf $f_{X,Y}(x,y)$ and $(W,Z) = \left(X,\frac{1}{c}XY\right)$, then the joint pdf of (W,Z) is $f_{W,Z}(w,z) = c\frac{1}{w}f_{X,Y}\left(w,c\frac{z}{w}\right)$. Consequently, the joint pdf of $(\vartheta_t,\vartheta_{t+1})$, conditionally on X^t, is:

$$
\mathfrak{h}_t\left(\vartheta_t,\vartheta_{t+1}\right) = \mathfrak{f}_t(\vartheta_t)\cdot\mathfrak{g}_t(\vartheta_{t+1})\cdot\frac{h_t\left(F_t^{-1}(\mathfrak{F}_t(\vartheta_t)),G_t^{-1}(\mathfrak{G}_t(\vartheta_{t+1}))\right)}{f_t\left(F_t^{-1}(\mathfrak{F}_t(\vartheta_t))\right)\cdot g_t\left(G_t^{-1}(\mathfrak{G}_t(\vartheta_{t+1}))\right)}.
$$

The marginal density of ϑ_{t+1}, given X^t, is then obtained by integrating out ϑ_t: $p(\vartheta_{t+1}|X^t) = \int\mathfrak{h}_t\left(\vartheta_t,\vartheta_{t+1}\right)d\vartheta_t$. Using the transformation $\zeta_t = F_t^{-1}(\mathfrak{F}_t(\vartheta_t))$, we obtain $F_t(\zeta_t) = \mathfrak{F}_t(\vartheta_t)$ and then differentiation yields $f_t(\zeta_t)d\zeta_t = \mathfrak{f}_t(\vartheta_t)d\vartheta_t$. Hence, the above integral reduces to

$$
\int\mathfrak{h}_t\left(\vartheta_t,\vartheta_{t+1}\right)d\vartheta_t = \frac{\mathfrak{g}_t(\vartheta_{t+1})}{g_t\left(G_t^{-1}(\mathfrak{G}_t(\vartheta_{t+1}))\right)}\cdot\int h_t\left(\zeta_t,G_t^{-1}(\mathfrak{G}_t(\vartheta_{t+1}))\right)d\zeta_t.
$$

As previously discussed, the second marginal density associated with h_t is g_t, and so $\int h_t(\zeta_t,\zeta_{t+1})d\zeta_t = g_t(\zeta_{t+1})$. It follows that

$$
p(\vartheta_{t+1}|X^t) = \int\mathfrak{h}_t\left(\vartheta_t,\vartheta_{t+1}\right)d\vartheta_t = \mathfrak{g}_t(\vartheta_{t+1}), \tag{A7}
$$

and so the conditional distribution of ϑ_t, given X^t, is $\mathrm{TG}(A_t,B_t;\underline{\vartheta},\overline{\vartheta})$ where $A_t = \omega a_t = \omega\left(A_{t-1}+\frac{1}{2}\right)$ and $B_t = \omega b_t = \omega\left(B_{t-1}+\frac{(X_t-\mu)^2}{2}\right)$. ■

Proof of Theorem 2: The predictive distribution of X_{t+1}, given X^t, is

$$g\left(X_{t+1}|X^t\right) = \int p\left(X_{t+1}|\vartheta_{t+1}\right) p\left(\vartheta_{t+1}|X^t\right) d\vartheta_{t+1} \tag{A8}$$

$$\propto \int_{\underline{\vartheta}}^{\overline{\vartheta}} \vartheta_{t+1}^{\frac{1}{2}} \exp\left(-\frac{\vartheta_{t+1}(X_{t+1}-\mu)^2}{2}\right) \vartheta_{t+1}^{A_t-1} e^{-B_t\vartheta_{t+1}} d\vartheta_{t+1} \tag{A9}$$

$$= \int_{\underline{\vartheta}}^{\overline{\vartheta}} \vartheta_{t+1}^{\left(A_t+\frac{1}{2}\right)-1} \exp\left(-\left(B_t+\frac{(X_{t+1}-\mu)^2}{2}\right)\vartheta_{t+1}\right) d\vartheta_{t+1}. \tag{A10}$$

Since $\nu_t = 2A_t$, note that $B_t + \frac{(X_{t+1}-\mu)^2}{2} = B_t\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)$. Using the transformation $u = B_t\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)\vartheta_{t+1}$, we obtain that

$$g\left(X_{t+1}|X^t\right) \propto \frac{1}{\left(B_t\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)\right)^{\frac{\nu_t+1}{2}}} \int_{\underline{\vartheta}B_t\left(1+Y_{t+1}^2/\nu_t\right)}^{\overline{\vartheta}B_t\left(1+Y_{t+1}^2/\nu_t\right)} u^{\frac{\nu_t+1}{2}-1} e^{-u} du \tag{A11}$$

$$\propto \frac{\gamma\left[\frac{\nu_t+1}{2}, \frac{1}{2}\overline{\xi}_t\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)\right] - \gamma\left[\frac{\nu_t+1}{2}, \frac{1}{2}\underline{\xi}_t\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)\right]}{\left(1+\frac{Y_{t+1}^2}{\nu_t}\right)^{\frac{\nu_t+1}{2}}}, \tag{A12}$$

where $\underline{\xi}_t = 2\underline{\vartheta}B_t$, $\overline{\xi}_t = 2\overline{\vartheta}B_t$, and $\gamma[\rho,\kappa] = \int_0^\kappa z^{\rho-1}e^{-z}dz$ is the lower incomplete Gamma function. Hence, $Y_{t+1} = (X_{t+1}-\mu)/\sqrt{B_t A_t^{-1}}$ follows a dampened t distribution with ν_t degrees of freedom and truncation parameters $\underline{\xi}_t$ and $\overline{\xi}_t$. ■

Proof of an alternative representation of the predictive density of Y_{t+1} in equation (23):

Lemma 2 *Define,*

$$Y = \frac{Z}{\sqrt{W/\nu}}, \tag{A13}$$

where,

- *the random variable Z follows a standard Normal distribution;*
- *the random variable W follows a χ^2 distribution with ν degrees of freedom, truncated on the interval $(\underline{\xi},\overline{\xi})$, where $0 < \underline{\xi} < \overline{\xi} < \infty$;*
- *Z and W are independent random variables.*

Then Y has a density function that coincides with the predictive density (23) obtained from an asset pricing model subject to structural uncertainty.

Proof: We present a proof of this isomorphism as the construction allows us to establish that the moment generating function of Y_{t+1} is well-defined. The densities of Z and W are given by,

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad -\infty < z < \infty, \tag{A14}$$

$$f_W(w) = \frac{1}{c\left[\nu, \underline{\xi}, \overline{\xi}\right]} w^{\frac{\nu}{2}-1} e^{-\frac{w}{2}}, \qquad \underline{\xi} < w < \overline{\xi}, \tag{A15}$$

where $c\left[\nu, \underline{\xi}, \overline{\xi}\right] = 2^{\nu/2}\left(\gamma\left[\nu/2, \overline{\xi}/2\right] - \gamma\left[\nu/2, \underline{\xi}/2\right]\right)$ ensures that $f_W(w)$ is a proper density. Define the random variable $R = \sqrt{\frac{W}{\nu}}$. The density of R is seen to be,

$$f_R(r) = 2\nu r f_W(\nu r^2) = \frac{2\nu^{\frac{\nu}{2}}}{c\left[\nu, \underline{\xi}, \overline{\xi}\right]} r^{\nu-1} e^{-\frac{\nu r^2}{2}}, \qquad \sqrt{\underline{\xi}/\nu} < r < \sqrt{\overline{\xi}/\nu}.$$

Adopt the following two transformations,

$$Y = \frac{Z}{R}, \qquad Q = R, \tag{A16}$$

so that $(Y, Q) = g(Z, R)$ where $g(z, r) = (z/r, r)$. A transformation argument using the Jacobian of g^{-1} yields that the joint density of (Y, Q) is

$$\begin{aligned} f_{Y,Q}(y,q) &= f_{Z,R}(yq, q)q = f_Z(yq)\, f_R(q)\, q, \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(yq)^2}{2}} \frac{2\nu^{\frac{\nu}{2}}}{c\left[\nu, \underline{\xi}, \overline{\xi}\right]} q^{\nu-1} e^{-\frac{\nu q^2}{2}} q = \frac{1}{\sqrt{2\pi}} \frac{2\nu^{\frac{\nu}{2}}}{c\left[\nu, \underline{\xi}, \overline{\xi}\right]} q^{\nu} e^{-\frac{(\nu+y^2)q^2}{2}}, \end{aligned}$$

for $-\infty < y < \infty$ and $\sqrt{\underline{\xi}/\nu} < q < \sqrt{\overline{\xi}/\nu}$. To obtain the density of Y, integrate $f_{Y,Q}(y,q)$ over q as,

$$\begin{aligned} f_Y(y) &= \int_{\sqrt{\underline{\xi}/\nu}}^{\sqrt{\overline{\xi}/\nu}} f_{Y,Q}(y,q)\, dq, && \text{(A17)} \\ &= \frac{1}{\sqrt{2\pi}} \frac{2\nu^{\frac{\nu}{2}}}{c\left[\nu, \underline{\xi}, \overline{\xi}\right]} \int_{\sqrt{\underline{\xi}/\nu}}^{\sqrt{\overline{\xi}/\nu}} q^{\nu} e^{-\frac{(\nu+y^2)q^2}{2}}\, dq. && \text{(A18)} \end{aligned}$$

Letting $s=\frac{(\nu+y^2)q^2}{2}$ we obtain $q=\left(\frac{2}{\nu+y^2}\right)^{\frac{1}{2}}s^{\frac{1}{2}}$ and $dq=\frac{1}{2}\left(\frac{2}{\nu+y^2}\right)^{\frac{1}{2}}s^{-\frac{1}{2}}ds$. Therefore,

$$\begin{aligned}
\int_{\sqrt{\underline{\xi}/\nu}}^{\sqrt{\overline{\xi}/\nu}} q^{\nu} e^{-\frac{(\nu+y^2)q^2}{2}} dq &= \int_{\frac{\underline{\xi}}{2}\left(1+\frac{y^2}{\nu}\right)}^{\frac{\overline{\xi}}{2}\left(1+\frac{y^2}{\nu}\right)} \left(\frac{2}{\nu+y^2}\right)^{\frac{\nu}{2}} s^{\frac{\nu}{2}} e^{-s}\frac{1}{2}\left(\frac{2}{\nu+y^2}\right)^{\frac{1}{2}} s^{-\frac{1}{2}}\,ds,\\
&= \frac{2^{\frac{\nu-1}{2}}}{(\nu+y^2)^{\frac{\nu+1}{2}}}\int_{\frac{\underline{\xi}}{2}\left(1+\frac{y^2}{\nu}\right)}^{\frac{\overline{\xi}}{2}\left(1+\frac{y^2}{\nu}\right)} s^{\frac{\nu+1}{2}-1}e^{-s}\,ds. \qquad \text{(A19)}
\end{aligned}$$

Accordingly,

$$\int_{\sqrt{\underline{\xi}/\nu}}^{\sqrt{\overline{\xi}/\nu}} q^{\nu} e^{-\frac{(\nu+y^2)q^2}{2}} dq = \left(\frac{2^{\frac{\nu-1}{2}}}{\nu^{\frac{\nu+1}{2}}}\right)\frac{\left(\gamma\left[\frac{\nu+1}{2},\frac{1}{2}\overline{\xi}\left(1+\frac{y^2}{\nu}\right)\right]-\gamma\left[\frac{\nu+1}{2},\frac{1}{2}\underline{\xi}\left(1+\frac{y^2}{\nu}\right)\right]\right)}{\left(1+\frac{y^2}{\nu}\right)^{\frac{\nu+1}{2}}}. \qquad \text{(A20)}$$

Substitution of (A20) into (A18), along with the fact that $c\left[\nu,\underline{\xi},\overline{\xi}\right]=2^{\nu/2}\left(\gamma\left[\nu/2,\overline{\xi}/2\right]-\gamma\left[\nu/2,\underline{\xi}/2\right]\right)$, proves our assertion. That is,

$$f_Y(y) = \frac{\gamma\left[\frac{\nu+1}{2},\frac{1}{2}\overline{\xi}\left(1+\frac{Y_{t+1}^2}{\nu}\right)\right]-\gamma\left[\frac{\nu+1}{2},\frac{1}{2}\underline{\xi}\left(1+\frac{Y_{t+1}^2}{\nu}\right)\right]}{\sqrt{\pi\nu}\left(\gamma\left[\nu/2,\overline{\xi}/2\right]-\gamma\left[\nu/2,\underline{\xi}/2\right]\right)\left(1+\frac{Y_{t+1}^2}{\nu}\right)^{\frac{\nu+1}{2}}}. \qquad \text{(A21)}$$

Therefore, the random variable $Y=\frac{Z}{\sqrt{W/\nu}}$ has the same density as the predictive density of Y_{t+1} derived in (23). ∎

Proof of Theorem 3: Rather than use the predictive density (23) to derive the moment generating function, we appeal to the properties of $Y=\frac{Z}{\sqrt{W/\nu}}$, which shares the same density function as shown in Lemma 2. Recall that Z and W are independent random variables, where Z follows a standard Normal distribution and W follows a truncated chi-square distribution. Based on the law of iterated expectations,

$$\begin{aligned}
E\left(e^{\lambda Y}\right) &= E\left(\exp\left(\lambda Z/\sqrt{W/\nu}\right)\right), && \text{(A22)}\\
&= E\left(E\left(\exp\left(\left(\frac{\lambda\sqrt{\nu}}{\sqrt{W}}\right)Z\right)\middle| W\right)\right), && \text{(A23)}\\
&= E\left(\exp\left(\frac{\lambda^2\nu}{2W}\right)\right), \qquad \text{(from the mgf of standard normal)}, && \text{(A24)}
\end{aligned}$$

where the expectation in (A24) is to be taken with respect to the chi-squared distribution. Verifying (29) we have,

$$E\left(e^{\lambda Y}\right) = E\left(\exp\left(\frac{\lambda^2\nu}{2W}\right)\right) = \frac{1}{c\left[\nu,\underline{\xi},\overline{\xi}\right]}\int_{\underline{\xi}}^{\overline{\xi}} e^{\frac{\lambda^2\nu}{2w}}\, w^{\frac{\nu}{2}-1}\, e^{-\frac{w}{2}}\, dw, \tag{A25}$$

where $c\left[\nu,\underline{\xi},\overline{\xi}\right] = 2^{\nu/2}\left(\gamma\left[\nu/2,\overline{\xi}/2\right]-\gamma\left[\nu/2,\underline{\xi}/2\right]\right)$ is a constant of integration. Now,

$$E\left(Y^{2\phi}\right) = E\left(\left(\lambda Z/\sqrt{W/\nu}\right)^{2\phi}\right) = \nu^{\phi}E\left(Z^{2\phi}\right)E\left(\frac{1}{W^{\phi}}\right) = \nu^{\phi}\frac{(2\phi)!}{\phi!2^{\phi}}E\left[\frac{1}{W^{\phi}}\right], \tag{A26}$$

$$= \left(\frac{\nu}{2}\right)^{\phi}\frac{(2\phi)!}{\phi!}\frac{1}{c\left[\nu,\underline{\xi},\overline{\xi}\right]}\int_{\underline{\xi}}^{\overline{\xi}} w^{-\phi}\, w^{\frac{\nu}{2}-1}\, e^{-\frac{w}{2}}\, dw. \tag{A27}$$

Finally, the odd-order moments of Y are zero since the density is symmetric around zero. ■

Proof of the rare event probabilities in equation (34): Let $x^{\text{rare}} = \mu - h\sqrt{V}$, where the volatility multiple h represents the deviation from mean consumption growth. Thus, from Lemma 2,

$$\text{Prob}\left(X_{t+1}\le x^{\text{rare}}\right) = \text{Prob}\left(Y_{t+1}\le -h\right) = \text{Prob}\left(Z\sqrt{\nu/W}\le -h\right), \tag{A28}$$

$$= E\left(\text{Prob}\left(Z\le -h\sqrt{W/\nu}\middle|W\right)\right), \tag{A29}$$

$$= E\left(N\left(-h\sqrt{W/v}\right)\right), \tag{A30}$$

$$= \frac{1}{c\left[\nu,\underline{\xi},\overline{\xi}\right]}\int_{\underline{\xi}}^{\overline{\xi}} N\left(-h\sqrt{w/v}\right) w^{\frac{\nu}{2}-1}\, e^{-\frac{w}{2}}\, dw, \tag{A31}$$

where $N(d) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d} e^{-u^2/2}\, du$ is the standard Normal cdf. ■

References

Abel, A., 2002. An exploration of the effects of pessimism and doubt on asset returns. Journal of Economic Dynamics and Control 26, 1075–1092.

Adam, K., Marcet, A., Nicolini, J., 2008. Stock market volatility and learning. Working paper. European Central Bank.

Aït-Sahalia, Y., Parker, J., Yogo, M., 2004. Luxury goods and the equity premium. Journal of Finance 59, 2959–3004.

Bakshi, G., Chen, Z., 1996. The spirit of capitalism and stock market prices. American Economic Review 86, 133–157.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1509.

Barro, R., 2006. Rare disasters and asset markets in the twentieth century. Quarterly Journal of Economics 121, 823–866.

Barsky, R., DeLong, B., 1993. Why does the stock market fluctuate?. Quarterly Journal of Economics 108, 291–311.

Bauwens, L., Lubrano, M., Richard, J.-F., 1999. Bayesian Inference in Dynamic Econometric Models. Oxford University Press, New York, NY.

Brennan, M., Xia, Y., 2001. Stock price volatility and equity premium. Journal of Monetary Economics 47, 249–283.

Campbell, J., 1993. Intertemporal asset pricing without consumption data. American Economic Review 83, 487–512.

Campbell, J., Cochrane, J., 1999. Force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–251.

Campbell, J., Cochrane, J., 2000. Explaining the poor performance of consumption-based asset pricing models. The Journal of Finance 55, 2683–2878.

Cecchetti, S., Lam, P.-S., Mark, N., 2000. Asset pricing with distorted beliefs: Are equity returns too good to be true?. American Economic Review 90, 787–805.

Coffey, C., Muller, K., 2000. Properties of doubly-truncated Gamma variables. Communications in Statistics - Theory and Methods 29, 851–857.

Cogley, T., 2009. Is the market price of risk infinite?. Economic Letters 102, 13–16.

Cogley, T., Sargent, T., 2008. The market price of risk and the equity premium: A legacy of the great depression?. Journal of Monetary Economics 55, 454–476.

Constantinides, G., 1990. Habit formation: A resolution of the equity premium puzzle. Journal of Political Economy 98, 519–543.

David, A., 2008. Heterogeneous beliefs, speculation, and the equity premium. Journal of Finance 63, 41–83.

Duffie, D., 1992. Dynamic Asset Pricing Theory. Princeton University Press, Princeton, New Jersey.

Epstein, L., Zin, S., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. Journal of Political Economy 99, 263–286.

Gabaix, X., 2007. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. Working paper. Stern School of Business, New York University.

Geweke, J., 2001. A note on some limitations of CRRA utility. Economics Letters 71, 341–345.

Harrison, M., Kreps, D., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20, 381–408.

Harvey, A., Fernandes, C., 1989. Time series model for insurance claims. Journal of the Institute of Actuaries 116, 513–528.

Jobert, A., Platania, A., Rogers, L. C. G., 2006. A Bayesian solution to the equity premium puzzle. Working paper. Statistical Laboratory, University of Cambridge.

Kocherlakota, N., 1996. The equity premium: It's still a puzzle. Journal of Economic Literature 34, 42–71.

LeRoy, S., Porter, R., 1981. The present-value relation: Tests based on implied variance bounds. Econometrica 101, 555–574.

Lettau, M., Ludvigson, S., Wachter, J., 2007. The declining equity premium: What role does macroeconomic risk play?. Review of Financial Studies 21, 1653–1687.

Lettau, M., Wachter, J., 2007. Why is long-horizon equity less risky? A duration-based explanation of the value premium. Journal of Finance 62, 55–92.

Lewellen, J., Shanken, J., 2002. Learning, asset-pricing tests, and market efficiency. Journal of Finance 57, 1113–1145.

Lucas, R., 1978. Asset prices in an exchange economy. Econometrica 46, 1429–1445.

Lukacs, E., 1960. Characteristic Functions. Charles Griffin and Company, London.

Mehra, R., Prescott, E., 1985. The equity premium: A puzzle. Journal of Monetary Economics 15, 145–161.

Mehra, R., Prescott, E., 2003. The equity premium in retrospect. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. Elsevier Science, North Holland: Amsterdam.

Mehra, R., Prescott, E., 2008. Equity premium. In: Mehra, R. (Eds.), The Handbook of the Equity Risk Premium. Elsevier Science, North Holland: Amsterdam.

Pastor, L., Veronesi, P., 2009. Learning in financial markets. NBER Working Paper No. 14646, University of Chicago.

Rietz, T., 1988. The equity risk premium: A solution. Journal of Monetary Economics 22, 117–131.

Shephard, N., 1994. Local scale models: State space alternative to integrated GARCH processes. Journal of Econometrics 60, 181–202.

Timmermann, A., 1993. How learning in financial markets generate excess volatility and predictability in stock prices. Quarterly Journal of Economics 108, 1135–1145.

Timmermann, A., 1996. Excess volatility and predictability of stock prices in autoregressive dividend models with learning. Review of Economic Studies 63, 523–557.

Tsionas, E., 2005. Likelihood evidence on the asset return puzzle. Review of Economic Studies 72, 917–946.

Weil, P., 1989. The equity premium puzzle and the riskfree rate puzzle. Journal of Monetary Economics 24, 401–421.

Weitzman, M., 2007. Subjective expectations and asset-return puzzles. American Economic Review 97, 1102–1130.

Table 1

Riskfree return and equity premium when the maximum consumption growth volatility, $\overline{\sigma}$, is 50%

Reported are the riskfree return and the equity premium under structural uncertainty and Bayesian learning. The computation relies on the following parameter inputs:

$$\mu = 2\%, \quad \sqrt{V} = 2\%, \quad k \in \{10, 30, 50\}, \quad \beta = 0.98, \quad \underline{\sigma} = 1/\sqrt{\overline{\vartheta}} = 0.1\%, \quad \overline{\sigma} = 1/\sqrt{\underline{\vartheta}} = 50\%.$$

The Normal distribution corresponds to the case of no structural uncertainty and is reported in the row "Normal." We compute the riskfree return as

$$\ln(R^f_{t+1}) \quad = \quad -\ln(\beta) - \ln(\Psi_X[-\alpha]),$$

and the equity risk premium as

$$\ln(E_t(R^e_{t+1})) - \ln(R^f_{t+1}) \quad = \quad \ln(\Psi_X[1]) \; - \; \ln(\Psi_X[1-\alpha]) \; + \; \ln(\Psi_X[-\alpha]),$$

where $\Psi_X[\lambda]$ is presented in equation (30) of Theorem 3. α is the coefficient of relative risk aversion.

α	Effective sample size	Riskfree return	Expected return	Equity premium
2	$k = 10$	5.91%	6.02%	0.11%
	$k = 30$	5.93%	6.02%	0.09%
	$k = 50$	5.94%	6.02%	0.08%
	Normal	5.94%	6.02%	0.08%
3	$k = 10$	7.78%	7.94%	0.16%
	$k = 30$	7.83%	7.96%	0.13%
	$k = 50$	7.83%	7.96%	0.13%
	Normal	7.84%	7.96%	0.12%
5	$k = 10$	11.35%	11.62%	0.27%
	$k = 30$	11.48%	11.70%	0.22%
	$k = 50$	11.50%	11.71%	0.21%
	Normal	11.52%	11.72%	0.20%
10	$k = 10$	19.33%	19.87%	0.54%
	$k = 30$	19.86%	20.30%	0.43%
	$k = 50$	19.93%	20.35%	0.42%
	Normal	20.02%	20.42%	0.40%

Table 2

Matching the riskfree return and the equity premium using the asset pricing model with subjective expectations

Reported are the riskfree return and the equity premium under structural uncertainty and Bayesian learning by varying the maximum level of consumption growth volatility $\overline{\sigma} = 1/\sqrt{\underline{\vartheta}}$ and the risk aversion α along possible grid points. Fixing α, $\overline{\sigma} = 1/\sqrt{\underline{\vartheta}}$ is varied to match the equity premium of 6% in Panel A, and the riskfree return of 1% in Panel B. The computation relies on the following parameters:

$$\mu = 2\%, \qquad \sqrt{V} = 2\%, \qquad k = 50, \qquad \beta = 0.98, \qquad \underline{\sigma} = 1/\sqrt{\overline{\vartheta}} = 0.1\%.$$

The equity premium and the riskfree return are both expressed as annual percentages. Note that the maximum allowable consumption growth volatility $\overline{\sigma}$ is not expressed in percentage terms (i.e., $\overline{\sigma} = 11.935$ stands for 1193.5%).

Panel A: Matched equity premium of 6%			Panel B: Matched riskfree return of 1%		
α	$\overline{\sigma}$	Riskfree return	α	$\overline{\sigma}$	Equity premium
2	11.935	0.02%	2	11.931	5.02%
3	7.652	1.96%	3	7.654	6.96%
4	5.569	3.85%	4	5.574	8.85%
5	4.347	5.71%	5	4.353	10.71%
6	3.546	7.52%	6	3.553	12.52%
7	2.983	9.29%	7	2.990	14.29%
8	2.566	11.02%	8	2.574	16.02%
9	2.246	12.71%	9	2.253	17.71%
10	1.993	14.35%	10	2.000	19.35%
15	1.251	21.95%	15	1.258	26.95%
20	0.894	28.49%	20	0.900	33.49%

Table 3
Structural uncertainty and probabilities of rare disasters

The table presents the probabilities of rare disasters corresponding to the (i) Normal distribution, (ii) the dampened t distribution by varying the maximum level of consumption growth volatility $\overline{\sigma}$, and (iii) the Student-t distribution. Rare disaster x^{rare} represents the deviation from mean consumption growth in multiples of volatility:

$$x^{\text{rare}} \;=\; \mu \;-\; h\sqrt{V},$$

where $h \in \{3,4,5,6\}$, $\mu = 2\%$, and $\sqrt{V} = 2\%$. Under the assumption that $k = 50$ and $\underline{\sigma} = 1/\sqrt{\overline{\vartheta}} = 0.1\%$, we compute and report the rare disaster probabilities:

$$\text{Prob}\left(X_{t+1} \leq x^{\text{rare}}\right) \frac{1}{c\left[\nu, \underline{\xi}, \overline{\xi}\right]} \int_{\underline{\xi}}^{\overline{\xi}} N\left(-h\sqrt{w/v}\right) w^{\frac{\nu}{2}-1} e^{-\frac{w}{2}}\, dw,$$

where $\nu = k-1$ is the degrees of freedom, $\underline{\xi} = \frac{(k-1)V}{\overline{\sigma}^2}$, $\overline{\xi} = \frac{(k-1)V}{\underline{\sigma}^2}$, and $N(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-u^2/2}\, du$ is the standard Normal cumulative distribution function. The Normal distribution probabilities correspond to $k = \infty$, and the Student-t distribution probabilities correspond to $\underline{\sigma} = 1/\sqrt{\overline{\vartheta}} = 0$ and $\overline{\sigma} = 1/\sqrt{\underline{\vartheta}} = \infty$.

		Dampened t distribution			
h	Normal distribution	$\overline{\sigma} = 4\%$	$\overline{\sigma} = 5\%$	$\overline{\sigma} = 6\%$	Student-t distribution
3	1.349898E-03	2.100851E-03	2.100852E-03	2.100852E-03	2.100852E-03
4	3.167124E-05	1.045948E-04	1.045951E-04	1.045951E-04	1.045951E-04
5	2.866516E-07	3.716510E-06	3.716606E-06	3.716606E-06	3.716606E-06
6	9.865877E-10	1.094241E-07	1.094470E-07	1.094470E-07	1.094470E-07

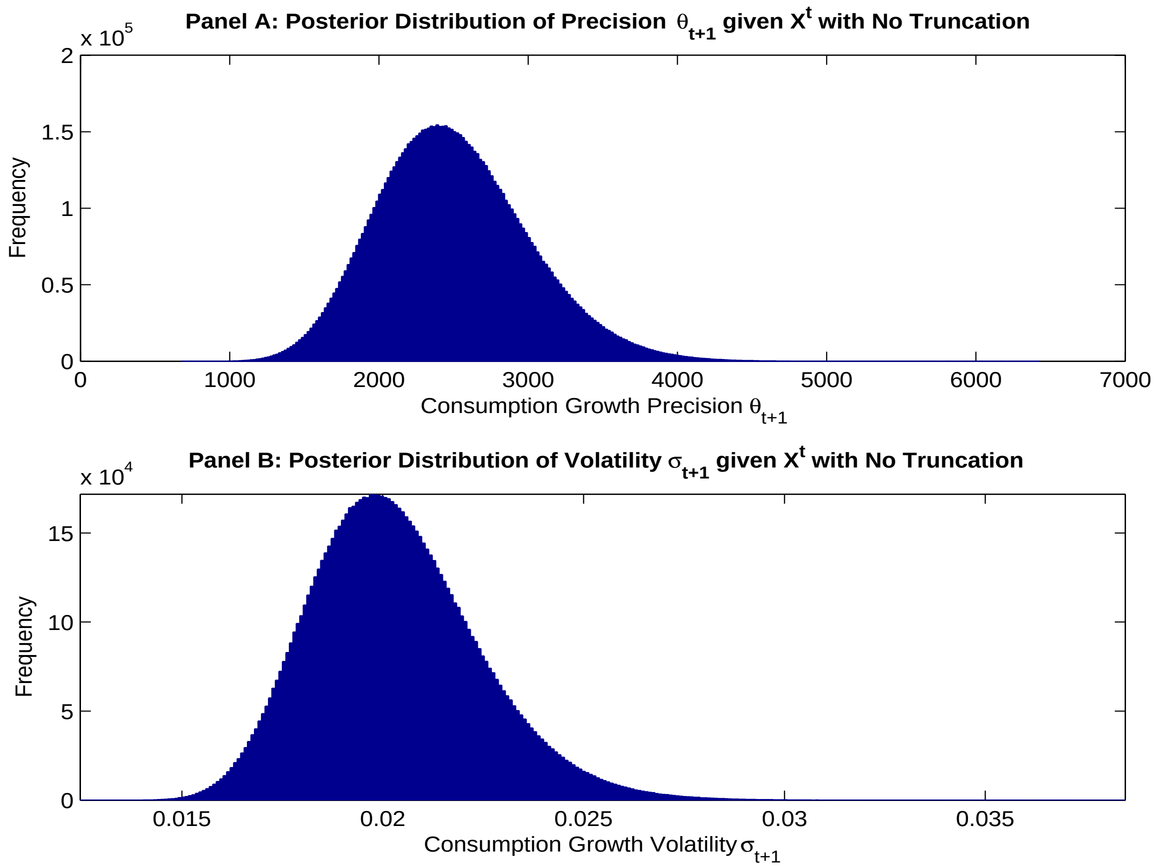

Fig. 1. Simulating the posterior distribution of θ_{t+1} and σ_{t+1} given X^t

Plotted is the posterior distribution of θ_{t+1} given X^t under no truncation, which is Gamma(A,B) given the availability of a large past history. In the computations, the volatility of consumption growth is $\sqrt{V} = 2\%$ and the effective sample size is $k = 50$. Therefore,

$$A = (k-1)/2 = 24.5, \qquad\qquad B = (k-1)V/2 = 0.0098.$$

Panel A presents the distribution of the precision of consumption growth, θ_{t+1}, while Panel B presents the distribution of the volatility of consumption growth, $\sigma_{t+1} = 1/\sqrt{\theta_{t+1}}$. Both plots are based on a simulation of 10 million draws.

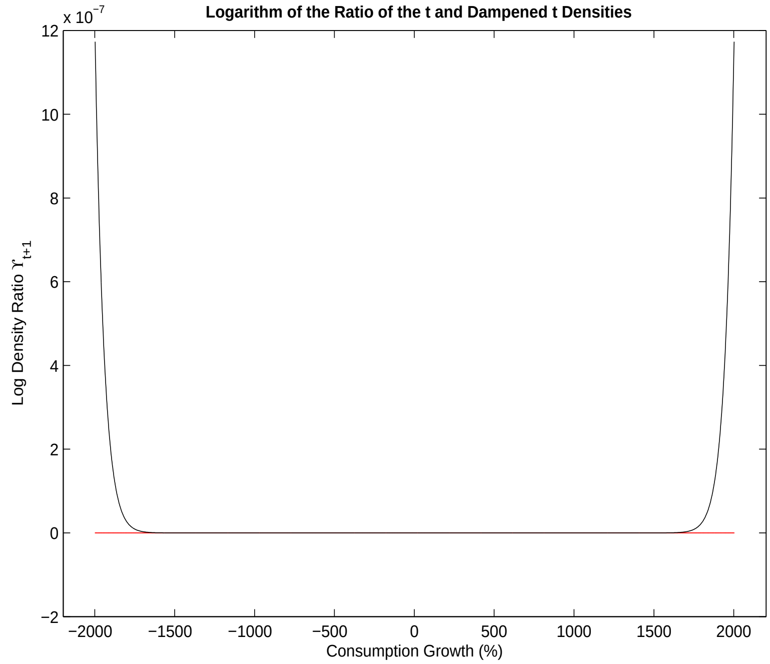

Fig. 2. Plot of $\ln\left(\frac{g(X_{t+1};k-1)}{g^{\mathrm{DT}}(X_{t+1};k-1,\underline{\vartheta}(k-1)V,\overline{\vartheta}(k-1)V)}\right)$ versus X_{t+1}

Plotted is the logarithm ratio of densities defined below (under the availability of a large past history):

$$\Upsilon_{t+1} \equiv \ln\left(\frac{g(X_{t+1};k-1)}{g^{\mathrm{DT}}(X_{t+1};k-1,\underline{\vartheta}(k-1)V_t,\overline{\vartheta}(k-1)V_t)}\right), \qquad X_{t+1} = \mu + \sqrt{V_t}\,Y_{t+1},$$

as a function of consumption growth X_{t+1}. Given that Y_{t+1} is Student-t distributed based on (12)-(13), the density of X_{t+1} is seen to be:

$$g(X_{t+1};k-1) = \frac{1}{\sqrt{V}}\,\frac{\Gamma[k/2]}{\sqrt{\pi(k-1)}\,\Gamma[(k-1)/2]}\left(1+\frac{\left((X_{t+1}-\mu)/\sqrt{V}\right)^2}{k-1}\right)^{-k/2}.$$

The form of the density $g^{\mathrm{DT}}(X_{t+1};k-1,\underline{\vartheta}(k-1)V,\overline{\vartheta}(k-1)V)$ is similarly obtained, as Y_{t+1} follows the dampened t distribution with density presented in (23). Here $\nu = 2A = k-1$, $\underline{\xi} = \underline{\vartheta}(k-1)V$ and $\overline{\xi} = \overline{\vartheta}(k-1)V$. For the purposes of this graph,

$$\mu = 2\%, \qquad \sqrt{V} = 2\%, \qquad k = 50, \qquad \underline{\sigma} = 1/\sqrt{\overline{\vartheta}} = 0.01\%, \qquad \overline{\sigma} = 1/\sqrt{\underline{\vartheta}} = 500\%.$$

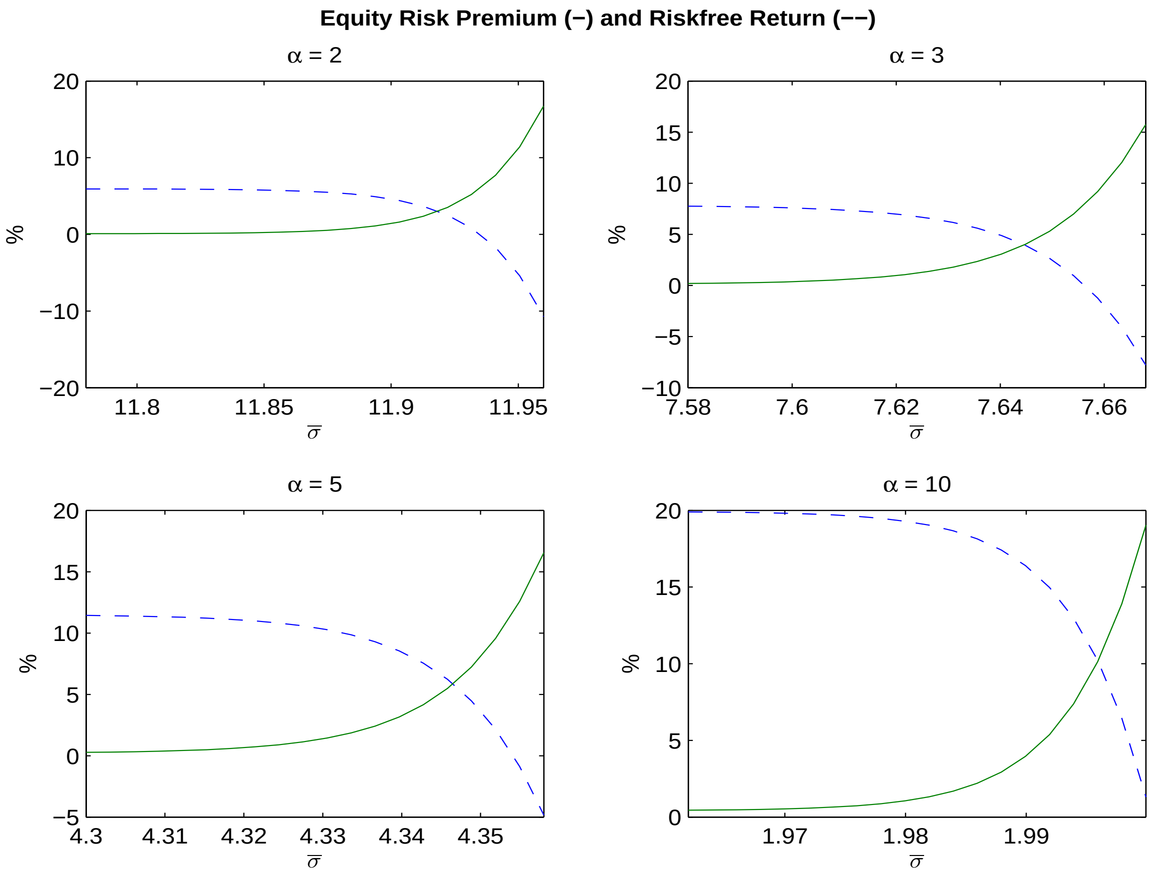

Fig. 3. Equity premium and riskfree return versus maximum consumption growth volatility $\overline{\sigma}$

The dashed-curve depicts the riskfree return, and the solid-curve depicts the equity risk premium on the y-axis, while changing $\overline{\sigma}$ on the x-axis. Note that the value, say, 11.9 on the x-axis corresponds to 1,190%. For these plots we set,

$$\mu = 2\%, \qquad \sqrt{V} = 2\%, \qquad k = 50, \qquad \beta = 0.98, \qquad \underline{\sigma} = 1/\sqrt{\vartheta} = 0.01\%.$$

Each plot corresponds to fixed risk aversion $\alpha \in \{2, 3, 5, 10\}$. We compute the riskfree return as

$$\ln(R^f_{t+1}) = -\ln(\beta) - \ln(\Psi_X[-\alpha]),$$

and the equity premium as

$$\ln(E_t(R^e_{t+1})) - \ln(R^f_{t+1}) \quad = \quad \ln(\Psi_X[1]) \; - \; \ln(\Psi_X[1-\alpha]) \; + \; \ln(\Psi_X[-\alpha]),$$

where $\Psi_X[\lambda]$ is presented in equation (30) of Theorem 3.