

ONLINE APPENDIX

Parameter Learning in General Equilibrium: The Asset Pricing Implications

Pierre Collin-Dufresne, Michael Johannes, and Lars A. Lochstoer*

September 21, 2015

Abstract

In this online appendix, we provide detailed descriptions of the numerical solution methods applied to solving the models in the main paper, as well as additional results and analysis as referenced in the main paper.

*Collin-Dufresne: Ecole Polytechnique Federale de Lausanne and Swiss Finance Institute and CEPR, Quartier UNIL-Dorigny, Extranef 209, CH-1015 Lausanne, Switzerland, pierre.collin-dufresne@epfl.ch. Johannes: Columbia Business School, 410 Uris Hall, 3022 Broadway, New York, NY 10027, MJ335@gsb.columbia.edu. Lochstoer: Columbia Business School, 424 Uris Hall, 3022 Broadway, New York, NY 10027, LL2609@gsb.columbia.edu.

Table of Contents

1. (*p. 2*) Additional literature review.
2. (*p. 4*) Analytical discrete- and continuous-time solutions to the model with learning about the mean growth rate and $EIS = 1$.
3. (*p. 18*) Truncated prior beliefs about μ , numerical solution methodology and analysis (including the case $EIS > 1$, as well as learning about the consumption volatility parameter).
4. (*p. 31*) A useful metric for assessing quantitative impact of uncertainty about a particular model parameter without solving the full sequential learning model.
5. (*p. 33*) Numerical solution methodology for the regime switching (Depression state) model.
6. (*p. 40*) Additional results for the regime switching (Depression state) model, including when mean and/or volatility parameters are unknown.
7. (*p. 46*) Numerical solution methodology for the model learning economy (Long-run risk vs. i.i.d. dynamics).
8. (*p. 48*) References.

1 Relation to existing literature

Our focus on parameter learning connects to a long-standing debate in macroeconomics. One common critique of full information, rational expectations models is precisely the assumption that agents know ‘fixed but unknown’ parameters, e.g., Modigliani (1977). Of course there is nothing about parameter or model learning inconsistent with rational expectations, as noted by Lucas and Sargent (1978, p. 68): “... *it has been only a matter of analytical convenience and not of necessity that equilibrium models have used the assumption of stochastically stationary "shocks" and the assumption that agents have already learned the probability distributions that they face. Both of these assumptions can be abandoned, albeit at a cost in terms of the simplicity of the model. . . . While models incorporating Bayesian learning and stochastic nonstationarity are both technically feasible and consistent with the equilibrium modeling strategy, almost no successful applied work along these lines has come to light. One reason is probably that nonstationary time series models are cumbersome and come in so many varieties.*” As discussed below, numerical solutions are generally required and can be quite complicated.

Hansen (2007) stresses the importance of studying how parameter and model uncertainty impacts asset valuation, forcing economic agents to face the inference problems as econometricians. Hansen (2007), and also Hansen and Sargent (2010), assume agents make decisions that are robust to model uncertainty and consider the case of an EIS of one. In contrast, we focus on Bayesian learning with EZ preferences, consider EIS values different from one and also consider the pricing of long-horizon risky claims—claims to infinite streams of consumption and dividends, as well as long-term bonds.

Other related papers considering general equilibrium implications of parameter learning include Veronesi (2000), Cogley and Sargent (2008), Jobert, et al. (2006), Benzoni, Collin-Dufresne and Goldstein (2011), Johannes, et al. (2014), and Kumar and Gvozdeva (2012). Relative to these, our paper focuses on the impact of priced parameter uncertainty on asset price moments in a general equilibrium model with EZ preferences.

Johannes, et al.’s (2010) main contribution is empirically documenting that shocks to beliefs about future consumption growth correlate with equity returns, consistent with the importance of updating. Their main asset pricing results rely on AU.

Pastor and Veronesi (2009, 2012) consider parameter learning with power utility preferences over terminal wealth. As shown in Timmermann (1993) and Lewellen and Shanken (2002), parameter learning about dividend dynamics induces excess return predictability in in-sample forecasting regressions. The hallmark of the learning channel, however, is poor out-of-sample performance of

such regressions, consistent with the data (see Goyal and Welch (2008)).

A number of papers consider state uncertainty, where the state evolves discretely via a Markov chain or smoothly via a Gaussian process.¹ Veronesi (2000) considers learning about mean-dividend growth rates with power utility, focussing on the role of information quality. Learning about a fixed parameter is a special case, and with common preference parameters, the equity premium falls and could even be negative with parameter uncertainty.

David (1997) considers learning about firm profitability, which switches between two states. Moore and Shaller (1996) consider consumption/dividend based Markov switching models with state learning and power utility. Veronesi (2004) studies the implications of learning about a peso state in a Markov switching model with power utility. David and Veronesi (2010) consider a Markov switching model with learning about states. Pastor and Veronesi (2003, 2006) study uncertainty about a fixed dividend-growth rate or profitability levels with an exogenously specified pricing kernel, in part motivated in order to derive cross-sectional implications.

In the case of EZ utility, Brandt et al. (2004) study alternative rules for learning about an unknown Markov state, assuming parameters are known. Chen and Pakos (2008), Lettau, et al. (2008) and Boguth and Kuehn (2012) consider economic agents who learn about states in a Markov switching model with known parameters. Ai (2010) studies Kalman state learning in a production-based long-run risks model. Bansal and Shaliastovich (2010, 2011) and Shaliastovich (2015) consider learning about the persistent component in a Bansal and Yaron (2004) style model with sub-optimal Kalman learning. Croce, Lettau and Ludvigson (2014) consider boundedly rational learning in an long-run risk model, where agents have limited information and show that such learning can help explain a downward-sloping term structure of equity risk premiums.

Alternative utility functions with a preference for early resolution of uncertainty will exhibit effects similar to those we document with parameter learning and EZ preferences, though the quantitative effects will depend on the exact specification. Examples include general Kreps-Porteus preferences and smooth ambiguity aversion preferences of Klibanoff, Marinacci, and Mukerji (2009) and Ju and Miao (2012), as well as the fragile beliefs setup of Hansen and Sargent (2010).²

Strzalecki (2011) discusses of the relation between ambiguity attitudes and the preference for the timing of the resolution of uncertainty. Learning under ambiguity (e.g., Epstein and Schneider (2007)) differs from Bayesian learning as the former depends on the sets of priors entertained by the agent, with higher weight being given to more pessimistic prior beliefs when forming predictive

¹Earlier contributions include Detemple (1986), Dothan and Feldman (1986), and Gennotte (1986) who show that you can separate the filtering problem from the pricing problem.

²Benzoni, Collin-Dufresne, Goldstein and Helwege (2015) investigate the implications for credit spreads of learning under fragile beliefs.

distributions.

2 Analytical solutions to "learning about the mean"-model when EIS = 1

2.1 Discrete-time case

In this section, log consumption growth is i.i.d.:

$$\Delta c_{t+1} = \mu + \sigma \varepsilon_{t+1}, \quad (1)$$

where $\varepsilon_{t+1} \stackrel{i.i.d.}{\sim} N(0, 1)$. Agents have conjugate, Normally distributed prior beliefs about the mean growth rate and update beliefs sequentially by observing consumption growth realizations. In particular, let $\mu \sim N(\mu_t, A_t \sigma^2)$. Bayes rule then implies that the subjective consumption dynamics are:

$$\Delta c_{t+1} = \mu_t + \sqrt{1 + A_t} \sigma \tilde{\varepsilon}_{t+1}, \quad (2)$$

$$\mu_{t+1} = \mu_t + \frac{A_t}{\sqrt{1 + A_t}} \sigma \tilde{\varepsilon}_{t+1}, \quad (3)$$

$$A_{t+1}^{-1} = A_t^{-1} + 1, \quad (4)$$

where $\tilde{\varepsilon}_{t+1} = (\Delta c_{t+1} - \mu_t) / (\sqrt{1 + A_t} \sigma)$ is a standard normal shock in the agent's filtration.

The representative agent has Epstein-Zin preferences with unit elasticity of intertemporal substitution (EIS, hereafter). Thus:

$$v c_t \equiv \ln V_t / C_t = \beta \ln E_t [e^{\alpha(v c_{t+1} + \Delta c_{t+1})}]^{1/\alpha}, \quad (5)$$

where β is the usual time discounting preference parameter and $\alpha = 1 - \gamma$, where γ is the agent's relative risk aversion to atemporal wealth gambles.

Next, consider the log v/c ratio for the case with sequential learning from consumption realiza-

tions and parameter uncertainty. First, guess that $vc_t = a_t + b\mu_t$.³ Next, verify:

$$vc_t = \beta \ln E_t \left[e^{\alpha(vc_{t+1} + \Delta c_{t+1})} \right]^{1/\alpha} \quad (6)$$

$$= \beta \ln E_t \left[e^{\alpha(a_{t+1} + b\mu_{t+1} + \mu_t + \sqrt{1+A_t}\sigma\varepsilon_{t+1})} \right]^{1/\alpha} \quad (7)$$

$$= \beta \left(a_{t+1} + b\mu_t + \mu_t + \alpha \left(\frac{1}{2} \right) \left(bA_t/\sqrt{1+A_t} + \sqrt{1+A_t} \right)^2 \sigma^2 \right) \quad (8)$$

\Downarrow

$$a_t + b\mu_t = \beta a_{t+1} + \beta b\mu_t + \beta\mu_t + \beta\alpha \left(\frac{1}{2} \right) \left(bA_t/\sqrt{1+A_t} + \sqrt{1+A_t} \right)^2 \sigma^2 \quad (9)$$

First, solve for b :

$$b = \beta b + \beta \quad (10)$$

\Updownarrow

$$b = \frac{\beta}{1-\beta}. \quad (11)$$

Next, solve for a_t :

$$a_t = \beta a_{t+1} + \beta\alpha \frac{1}{2} \left(b \frac{A_t}{\sqrt{1+A_t}} + \sqrt{1+A_t} \right)^2 \sigma^2 \quad (12)$$

$$= \beta a_{t+1} + \beta \frac{\alpha}{2} \left(\frac{1}{1-\beta} A_t + 1 \right)^2 \sigma^2 / (1+A_t). \quad (13)$$

Thus, we have that:

$$vc_t = a_t + b\mu_t, \quad (14)$$

$$\begin{aligned} a_t &= \beta a_{t+1} + \beta \frac{\alpha}{2} \left(\frac{1}{1-\beta} A_t + 1 \right)^2 \sigma^2 / (1+A_t) \\ &= \sum_{j=0}^{\infty} \beta^{j+1} \frac{\alpha}{2} \left(\frac{1}{1-\beta} A_{t+j} + 1 \right)^2 \sigma^2 / (1+A_{t+j}), \end{aligned} \quad (15)$$

$$b = \beta / (1-\beta). \quad (16)$$

When $A_t = 0$, we get the standard value function when the mean is known. In this case, $a_t = \frac{1}{2}\alpha\sigma^2 \frac{\beta}{1-\beta}$.

³We thank Mikhail Chernov for pointing out this discrete-time solution to us.

2.1.1 The stochastic discount factor and the price of risk

Shocks to the log stochastic discount factor (m_t) can be written (see, e.g., Hansen, Heaton, and Li (2008)):

$$m_{t+1} - E_t[m_{t+1}] = -\gamma(\Delta c_{t+1} - E_t[\Delta c_{t+1}]) - (\gamma - 1)(\Delta v c_{t+1} - E_t[\Delta v c_{t+1}]) \quad (17)$$

$$= -\gamma\sqrt{1 + A_t}\sigma\tilde{\varepsilon}_{t+1} - (\gamma - 1)\frac{\beta}{1 - \beta}(\mu_{t+1} - E_t[\mu_{t+1}]) \quad (18)$$

$$= -\left(\gamma + (\gamma - 1)\frac{\beta}{1 - \beta}\frac{A_t}{1 + A_t}\right)\sqrt{1 + A_t}\sigma\tilde{\varepsilon}_{t+1}. \quad (19)$$

Thus, defining the Price of Risk as the conditional volatility of the log stochastic discount factor, the Price of Risk is $\left(\gamma + (\gamma - 1)\frac{\beta}{1 - \beta}\frac{A_t}{1 + A_t}\right)\sqrt{1 + A_t}\sigma$. Note that this is greater than $\gamma\sigma$, which is the Price of Risk when there is no parameter uncertainty, when agents have a preference for early resolution of uncertainty (i.e., when $\gamma > 1$).

2.1.2 The consumption claim: risk premium and return volatility

First, given the unit EIS, the wealth-consumption ratio is constant and equal to $\beta/(1 - \beta)$. Thus, log returns to an asset that pays consumption as its dividend (the consumption claim) is then

$$\begin{aligned} r_{c,t+1} &= \Delta c_{t+1} + \ln\left(\frac{\beta}{1 - \beta} + 1\right) - \ln\frac{\beta}{1 - \beta} \\ &= \Delta c_{t+1} - \ln\beta. \end{aligned}$$

Since both log consumption growth and the log stochastic discount factor are conditionally normally distributed, we have from the Law of One Price that:

$$1 = E_t[M_{t+1}R_{c,t+1}] \quad (20)$$

\Downarrow

$$0 = E_t[m_{t+1}] + E_t[r_{c,t+1}] + \frac{1}{2}\sigma_t^2[m_{t+1}] + \frac{1}{2}\sigma_t^2[r_{c,t+1}] + cov_t(m_{t+1}, r_{c,t+1}). \quad (21)$$

Further, the log risk-free rate is $r_{f,t} = -E_t[m_{t+1}] - \frac{1}{2}\sigma_t^2[m_{t+1}]$. Thus, the log return risk premium, including a Jensen's inequality term, is

$$E_t[r_{c,t+1}] - r_{f,t} + \frac{1}{2}\sigma_t^2[r_{c,t+1}] = -cov_t(m_{t+1}, r_{c,t+1}) \quad (22)$$

$$= cov_t(\gamma\Delta c_{t+1} + (\gamma - 1)\Delta v c_{t+1}, \Delta c_{t+1}) \quad (23)$$

$$= \gamma\sigma^2(1 + A_t) + (\gamma - 1)\frac{\beta}{1 - \beta}A_t\sigma^2. \quad (24)$$

The conditional return volatility (under the agent's filtration) is simply $\sigma_t(r_{c,t+1}) = \sigma_t(\Delta c_{t+1}) = \sqrt{1 + A_t}\sigma$, which leaves the 'Sharpe ratio' of log returns to the consumption claim equal to the Price Of Risk, as given earlier. In fact, it will be useful to define this maximum 'Sharpe ratio' as

$$SR_t \equiv \left(\gamma + (\gamma - 1)\frac{\beta}{1 - \beta}\frac{A_t}{1 + A_t} \right) \sqrt{1 + A_t}\sigma. \quad (25)$$

2.1.3 The default-free, real term structure

The log, real risk-free rate is:

$$r_{f,t} = -\ln E_t[M_{t+1}] \quad (26)$$

$$= -E_t[m_{t+1}] - \frac{1}{2}\sigma_t^2[m_{t+1}] \quad (27)$$

$$= \mu_t - \ln \beta + \frac{\sigma^2(1 + A_t)}{2} - SR_t\sigma\sqrt{1 + A_t}. \quad (28)$$

It is useful to define

$$f_t \equiv -\ln \beta + \frac{\sigma^2(1 + A_t)}{2} - SR_t\sigma\sqrt{1 + A_t}. \quad (29)$$

The two-period, zero-coupon yield to maturity is then

$$e^{-2y_t^{(2)}} = E_t[M_{t+1}M_{t+2}] \quad (30)$$

$$= E_t[M_{t+1}E_{t+1}[M_{t+2}]] \quad (31)$$

$$= E_t[M_{t+1}e^{\mu_{t+1} + f_{t+1}}]. \quad (32)$$

Solving, we can write:

$$y_t^{(2)} = r_{f,t} + \frac{1}{2} \left(f_{t+1} - f_t - \frac{\sigma^2}{2} \frac{A_t^2}{1 + A_t} - \frac{A_t}{\sqrt{1 + A_t}} \sigma SR_t \right). \quad (33)$$

Continuing this type of recursion, we have the general expression for the yield on a τ -period zero-coupon bond:

$$y_t^{(\tau)} = r_{f,t} + \frac{1}{\tau} \sum_{k=1}^{\tau-1} \left(f_{t+k} - f_t - \frac{(\tau-k)^2 \sigma^2}{2} \frac{A_{t+k-1}^2}{1+A_{t+k-1}} - \frac{(\tau-k) \sigma A_{t+k-1}}{\sqrt{1+A_{t+k-1}}} SR_{t+k-1} \right). \quad (34)$$

2.2 Continuous-time case

It is useful to also show the solutions in the continuous-time case. First, some general expressions of interest also for the case when EIS is different from unity can be derived. Second, in the EIS = 1 case, the analytical formulas are more elegant and highlight how learning about the mean in the continuous-time limit has *no* impact on short-run consumption risks (the conditional volatility of consumption growth), and so all the extra risks are subjective long-run risks induced endogenously through the learning-channel.

There are several formal treatments of stochastic differential utility and its implications for asset pricing (see, e.g., Duffie and Epstein (1992a,b), Duffie and Skiadas (1994), and Skiadas (2003)). In this Appendix we offer a simple derivation of the pricing kernel that obtains in an exchange economy where the representative agent has a KPEZ recursive utility with unit EIS and where he is learning about the constant growth rate of aggregate consumption. We first recall a few well-known results about stochastic differential utility.

2.2.1 Representation of Preferences and Pricing Kernel

We assume the existence of a standard filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ on which there exists a vector $z(t)$ of d independent Brownian motions.

Aggregate consumption in the economy is assumed to follow a continuous process, with stochastic growth rate and volatility:

$$d \log C_t = \mu_C(X_t) dt + \sigma_C(X_t) dz(t), \quad (35)$$

$$dX_t = \mu_X(X_t) dt + \sigma_X(X_t) dz(t), \quad (36)$$

where X_t is a n -dimensional Markov process (we assume sufficient regularity on the coefficient of the stochastic differential equation (SDE) for it to be well-defined, e.g., Duffie (2001) Appendix B). In particular μ_X is an $(n, 1)$ vector, σ_X is an (n, d) matrix.

Following Epstein and Zin (1989), we assume that the representative agent's preferences over a consumption process $\{C_t\}$ are represented by a utility index $U(t)$ that satisfies the following

recursive equation:

$$U(t) = \left\{ (1 - e^{-\beta dt}) C_t^{1-\rho} + e^{-\beta dt} E_t (U(t+dt)^{1-\gamma})^{\frac{1-\rho}{1-\gamma}} \right\}^{\frac{1}{1-\rho}}. \quad (37)$$

With $dt = 1$, this is the discrete time formulation of KPEZ, in which $\Psi \equiv 1/\rho$ is the EIS and γ is the risk-aversion coefficient.

To simplify the derivation let us define the function

$$u_\alpha(x) = \begin{cases} \frac{x^{1-\alpha}}{(1-\alpha)} & 0 < \alpha \neq 1 \\ \log(x) & \alpha = 1. \end{cases}$$

Further, let us define

$$g(x) = u_\rho(u_\gamma^{-1}(x)) \equiv \begin{cases} \frac{((1-\gamma)x)^{1/\theta}}{(1-\rho)} & \gamma, \rho \neq 1 \\ u_\rho(e^x) & \gamma = 1, \rho \neq 1 \\ \frac{\log((1-\gamma)x)}{(1-\rho)} & \rho = 1, \gamma \neq 1, \end{cases}$$

where

$$\theta = \frac{1-\gamma}{1-\rho}.$$

Then, defining the ‘normalized’ utility index J as the increasing transformation of the initial utility index $J(t) = u_\gamma(U(t))$, Equation 37 becomes:

$$g(J(t)) = (1 - e^{-\beta dt}) u_\rho(C_t) + e^{-\beta dt} g(E_t [J(t+dt)]). \quad (38)$$

Using the identity $J(t+dt) = J(t) + dJ(t)$ and performing a simple Taylor expansion we obtain:

$$0 = \beta u_\rho(C_t) dt - \beta g(J(t)) + g'(J(t)) E_t [dJ(t)]. \quad (39)$$

Slightly rearranging the above equation, we obtain a backward recursive stochastic differential equation which could be the basis for a formal definition of stochastic differential utility (see Duffie Epstein (1992), Skiadas (2003)):

$$E_t [dJ(t)] = -\frac{\beta u_\rho(C_t) - \beta g(J(t))}{g'(J(t))} dt. \quad (40)$$

Indeed, let us define the so-called ‘normalized’ aggregator function:

$$f(C, J) = \frac{\beta u_\rho(C) - \beta g(J)}{g(J)} \equiv \begin{cases} \frac{\beta u_\rho(C)}{((1-\gamma)J)^{1/\theta-1}} - \beta\theta J & \gamma, \rho \neq 1 \\ (1-\gamma)\beta J \log(C) - \beta J \log((1-\gamma)J) & \gamma \neq 1, \rho = 1 \\ \frac{\beta u_\rho(C)}{e^{(1-\rho)J}} - \frac{\beta}{1-\rho} & \gamma = 1, \rho \neq 1. \end{cases} \quad (41)$$

We obtain the following representation for the normalized utility index:

$$J(t) = E_t \left(\int_t^T f(C_s, J(s)) + J(T) \right). \quad (42)$$

Note the well-known fact that when $\rho = \gamma$ (i.e., $\theta = 1$) then $f(C, J) = \beta u_\rho(C) - \beta J$ and a simple application of Itô’s lemma shows that

$$J(t) = E_t \left(\int_t^T e^{-\beta(s-t)} \beta u_\rho(C_s) ds + e^{-\beta(T-t)} J(T) \right). \quad (43)$$

Further, Duffie-Epstein (1992b) show that the pricing kernel $(\Pi(t))$ for this economy has the following form (if there exists an ‘interior’ solution to the optimal consumption portfolio choice problem of the representative agent):

$$\Pi(t) = e^{\int_0^t f_J(C_s, J_s) ds} f_C(C_t, J_t). \quad (44)$$

It is the Riesz representation of the gradient of the normalized utility index at the optimal consumption (See Chapter 10 of Duffie (2001) for further discussion.) We now consider the case of unitary EIS ($\rho = 1$) and give an expression for the utility index and for the pricing kernel in this economy.

2.2.2 Equilibrium Prices when $\rho = 1$

Assuming the equilibrium consumption process given in equations (35)-(36) above, we obtain an explicit characterization of the utility index J and the corresponding pricing kernel Π .

For this we define, respectively, the operator

$$\mathcal{D}h(x) = h_x(x) \mu_X(x) + \frac{1}{2} \text{trace} \left(h_{xx} \sigma_X(x) \sigma_X(x)^T \right)$$

where h_x is the $(n, 1)$ Jacobian vector of first derivatives and h_{xx} denotes the (n, n) Hessian

matrix of second derivatives. With these notations, we find:

Proposition 1 *Suppose $I(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ solves the following equation:*

$$0 = I(x) \left((1 - \gamma)\mu_C(x) + (1 - \gamma)^2 \frac{\|\sigma_C(x)\|^2}{2} \right) + \mathcal{D}I(x) + (1 - \gamma)\sigma_C(x)\sigma_X(x)^\top I_x(x) - \beta I(x) \log I(x) \quad (45)$$

and satisfies the transversality condition $\lim_{T \rightarrow \infty} E [e^{-\beta T} \log C_T + e^{-\beta T} \log I(X_T)] = 0$ then the value function is given by:

$$J(t) = u_\gamma(C_t)I(x_t) \quad (46)$$

The corresponding pricing kernel is:

$$\Pi(t) = e^{-\int_0^t \beta(1 + \log I(x_s)) ds} (C_t)^{-\gamma} I(x_t) \quad (47)$$

Proof. From its definition in equations 42 and 41 we obtain:

$$\frac{dJ(t)}{J(t)} = (-(1 - \gamma)\beta \log C(t) + \beta \log((1 - \gamma)J(t))) dt + \sigma_J(t) dz(t) \quad (48)$$

for some \mathcal{F}_t -measurable diffusion process σ_J . An application of Itô's lemma to $e^{-\beta t} \log((1 - \gamma)J(t))$ shows that its solution satisfies the following integral equation for any $T > t$:

$$\log((1 - \gamma)J(t)) = E \left[\int_t^T e^{-\beta(s-t)} \left(\beta(1 - \gamma) \log C(s) + \frac{1}{2} |\sigma_J(s)|^2 \right) ds + e^{-\beta(T-t)} \log((1 - \gamma)J(T)) \right] \quad (49)$$

Further, if it satisfies the transversality condition $\lim_{T \rightarrow \infty} E[e^{-\beta T} \log((1 - \gamma)J(T))] = 0$, then $J(t)$ solves

$$\log((1 - \gamma)J(t)) = E \left[\int_t^\infty e^{-\beta(s-t)} \left(\beta(1 - \gamma) \log C(s) + \frac{1}{2} |\sigma_J(s)|^2 \right) ds \right] \quad (50)$$

Now suppose $I(x)$ satisfies the ODE given in equation 45. Applying Itô's lemma to $e^{-\beta t} \log((1 -$

$\gamma)J(t))$ using our guess $J(t) = u_\gamma(C_t)I(x_t)$ we find that

$$\begin{aligned} e^{-\beta T} \log((1 - \gamma)J(T)) - e^{-\beta t} \log((1 - \gamma)J(t)) &= - \int_t^T e^{-\beta s} \left(\beta(1 - \gamma) \log C(s) + \frac{1}{2} |\sigma_J(s)|^2 \right) ds \dots \\ &+ \int_t^T e^{-\beta s} ((1 - \gamma)\sigma_C(X_s) + \sigma_I(X_s)) dz(s) \end{aligned}$$

where

$$\sigma_I(x)^\top = \frac{\sigma_X(s)^\top I_X(x)}{I(x)}$$

and

$$\sigma_J = (1 - \gamma)\sigma_C + \sigma_I.$$

Suppose that (a) the stochastic integral is a martingale (sufficient conditions are

$E \left[\int_t^T e^{-2\beta s} (|\sigma_C(X_s)|^2 + |\sigma_I(X_s)|^2) ds \right] < \infty$) and (b) the transversality condition listed in the proposition is satisfied, then taking expectations and the limit when $T \rightarrow \infty$ in equation 50 above we obtain:

$$\log((1 - \gamma)J(t)) = E \left[\int_t^\infty e^{-\beta(s-t)} \left(\beta(1 - \gamma) \log C(s) + \frac{1}{2} |\sigma_J(s)|^2 \right) ds \right] \quad (51)$$

This shows that our candidate solution satisfies the recursive backward stochastic differential equation we are trying to solve. Uniqueness follows from the appendix in Duffie, Epstein, Skiadas (1992) (under some additional technical conditions listed therein).

The expression for the pricing kernel follows from its definition in equation 44, the expression for the aggregator in equation 41 and the expression for the value function just derived. ■

The next result investigates the property of equilibrium prices.

Proposition 2 *The risk-free interest rate is given by:*

$$r(x_t) = \beta + \mu_C(x_t) + \frac{\|\sigma_C(x_t)\|^2}{2} - \gamma \|\sigma_C(x_t)\|^2 + \sigma_C(x) \sigma_I(x) \quad (52)$$

Further, if the following transversality condition is satisfied $\lim_{T \rightarrow \infty} E [\Pi_T C_T] = 0$, then the value of the claim to aggregate consumption is given by:

$$V(t) = \frac{C(t)}{\beta} \quad (53)$$

It follows that

$$\frac{dV_t}{V_t} = (\mu_C(x_t) + \frac{1}{2}\|\sigma_C(x_t)\|^2)dt + \sigma_C(x_t)dz(t) \quad (54)$$

The risk premium on the claim to aggregate consumption is given by

$$\mu_V(x) + \beta - r(x) = (\gamma\sigma_C(x) - \sigma_I(x))^\top \sigma_C(x) \quad (55)$$

Proof. To prove the result for the interest rate, apply the Itô-Doeblin formula to the pricing kernel. It follows from $r(t) = -E[\frac{d\Pi(t)}{\Pi(t)}]/dt$ that:

$$r(x_t) = \beta + \beta \log I(x_t) + \gamma\mu_c(x_t) - \frac{1}{2}\gamma^2\|\sigma_c(x_t)\|^2 - \frac{\mathcal{D}I(x_t)}{I(x_t)} + \gamma\sigma_I^\top \sigma_C. \quad (56)$$

Now substitute the expression for $\log I(x)$ from the ODE for $I(x)$ in Proposition 1 to obtain the result.

To prove the result for the consumption claim, define $V(t) = \frac{C_t}{\beta}$. It follows from Itô's lemma that:

$$\frac{dV_t + C_t dt}{V_t} = (\mu_V(X_t) + \beta)dt + \sigma_C(X_t)dz_t$$

with $\mu_V(x) = (\mu_C(x_t) + \frac{1}{2}\|\sigma_C(x_t)\|^2)$. Then, using the definition of the risk-free rate we obtain:

$$\frac{dV_t + C_t dt}{V_t} = \left(r(X_t) + (\gamma\sigma_C(X_t) - \sigma_I(X_t))^\top \sigma_C(X_t) \right) dt + \sigma_C(X_t)dz_t$$

In turn since the state price density has dynamics:

$$\frac{d\Pi_t}{\Pi_t} = -r(X_t)dt - (\gamma\sigma_C(X_t) - \sigma_I(X_t)) dz_t$$

an application of the Itô's formula shows that

$$\Pi_T V_T - \Pi_t V_t + \int_t^T \Pi_s C_s ds = \int_t^T \Pi_s V_s ((1 - \gamma)\sigma_C(X_s) + \sigma_I(X_s)) dz_s,$$

which is a martingale (under appropriate regularity conditions for the stochastic integral to be a Martingale). Taking an expectation we thus obtain

$$\Pi_t V_t = E_t\left[\int_t^T \Pi_s C_s ds + \Pi_T V_T\right]$$

If, furthermore, the solution satisfies the transversality condition listed in the proposition (i.e., $\lim_{T \rightarrow \infty} E[\Pi_T C_T] = 0$), then we can let $T \rightarrow \infty$ and have indeed proved that $V_t = \frac{C_t}{\beta}$ is the value of the claim to aggregate consumption. ■

2.2.3 Application to learning

Suppose now that log consumption follows the following process:

$$d \log C_t = \mu dt + \sigma dz(t) \tag{57}$$

but we assume further that μ has to be estimated by the representative agent based on observing past consumption. Suppose that he starts with some Gaussian prior $\mu \sim N(m_0, \Sigma_0)$. Then it is well-known that his posterior is also Gaussian with mean and variance given by (m_t, Σ_t) with dynamics:

$$dm_t = \lambda_t (d \log C_t - m_t dt) \tag{58}$$

$$= \lambda_t \sigma d\tilde{z}(t) \tag{59}$$

where the second equation defines the innovation process \tilde{z}_t , a Brownian motion in the observation filtration of the agent, in terms of which the consumption process can be rewritten as

$$d \log C_t = m_t dt + \sigma d\tilde{z}(t). \tag{60}$$

Further, the posterior variance is:

$$d\Sigma_t = -\lambda_t^2 \sigma^2 dt \tag{61}$$

and the regression coefficient is given by:

$$\lambda_t = \frac{\Sigma_t}{\sigma^2}. \tag{62}$$

Therefore, note the dynamics of λ :

$$d\lambda_t = -\lambda_t^2 dt. \tag{63}$$

The solution of which is simply

$$\frac{1}{\lambda(t)} = \frac{1}{\lambda_0} + t.$$

Now we see that the state-vector in the information filtration of the agent is $X(t) = [m_t, t]$ (or equivalently, $[m_t, \Sigma_t]$).

The pricing kernel with EIS = 1 We now derive an expression for the $I(\cdot)$ function from proposition 1 above (for the case unitary EIS $\rho = 1$ and arbitrary risk-aversion γ). The ode given in equation 45 simplifies to (we drop arguments for simplicity):

$$0 = I((1 - \gamma)m + (1 - \gamma)^2 \frac{\sigma^2}{2}) + \frac{1}{2} I_{mm} \lambda(t)^2 \sigma^2 + (1 - \gamma) \sigma^2 \lambda(t) I_m - \beta I \log I + I_t. \quad (64)$$

We guess a solution of the form

$$\log I(m, t) = a(t) + b(t)m$$

Plugging into the pde and setting coefficients in m to zero we obtain two odes which are:

$$b'(t) - \beta b(t) + 1 - \gamma = 0 \quad (65)$$

$$a'(t) - \beta a(t) + \frac{(1 - \gamma)^2 \sigma^2}{2} + \frac{1}{2} b(t)^2 \lambda(t)^2 \sigma^2 + (1 - \gamma) \sigma^2 \lambda(t) b(t) = 0 \quad (66)$$

Now for the boundary conditions we note that (since $\lim_{t \rightarrow \infty} m_t = \mu$ and $\lim_{t \rightarrow \infty} \lambda(t) = 0$):

$$\lim_{t \rightarrow \infty} \log I(t) = \frac{(1 - \gamma)\mu + \frac{1}{2}(1 - \gamma)^2 \sigma^2}{\beta}$$

thus we find the boundary conditions:

$$\lim_{t \rightarrow \infty} b(t) = \frac{1 - \gamma}{\beta} \quad (67)$$

$$\lim_{t \rightarrow \infty} a(t) = \frac{(1 - \gamma)^2 \sigma^2}{2\beta} \quad (68)$$

Now a solution satisfying this is (uniqueness follows from the result on BSDE):

$$b(t) = \frac{1 - \gamma}{\beta} \quad (69)$$

$$a(t) = (1 - \gamma)^2 \sigma^2 \int_t^\infty e^{-\beta(s-t)} \left(\frac{1}{2} + \frac{\lambda(s)}{\beta} + \frac{\lambda(s)^2}{2\beta^2} \right) ds \quad (70)$$

$$= (1 - \gamma)^2 \sigma^2 \frac{1 + \frac{\lambda_t}{\beta} - e^{\beta/\lambda_t} \text{Ei}\left(-\frac{\beta}{\lambda_t}\right)}{2\beta} \quad (71)$$

where we have used the definition of the exponential integral function (the principal value of the integral $\text{Ei}(z) = \int_{-z}^\infty \frac{e^{-t}}{t} dt$). It is straightforward to verify that the transversality condition of proposition 1 is satisfied.

Now, using the expression for the pricing kernel in Proposition 1, we can obtain the interest rate in closed form using Proposition 2.

Specifically we find.

$$r(t) = \beta + m_t + \frac{\sigma^2}{2} - \gamma\sigma^2 + \sigma^2 b\lambda(t) \quad (72)$$

and the dynamics of the pricing kernel:

$$\frac{d\Pi(t)}{\Pi(t)} = -r(t)dt - (\gamma\sigma - b\sigma\lambda(t))d\tilde{z}_t. \quad (73)$$

Note that interestingly the function $a(t)$ plays no role in the expression for the interest rate and the risk-premium. We can also solve for long-term zero-coupon bond prices (and hence yields in this model). Note that the risk-free zero coupon bond prices has price:

$$P(0, T) = E^Q[e^{-\int_0^T r_t dt}] \quad (74)$$

$$= E^Q[e^{-\int_0^T (\beta + m_t + \frac{\sigma^2}{2} - \gamma\sigma^2 + \sigma^2 b\lambda(t)) dt}] \quad (75)$$

where under the risk-neutral measure m has dynamics:

$$dm_t = \lambda_t \sigma (d\tilde{z}^Q(t) - (\gamma\sigma - b\sigma\lambda_t) dt) \quad (76)$$

$$= -\sigma^2 (\gamma\lambda_t - b\lambda_t^2) dt + \lambda_t \sigma d\tilde{z}^Q(t). \quad (77)$$

Since λ_t is deterministic, m_t is a Gaussian process and the solution to the risk-free zero coupon

bond is immediate.

$$P(0, T) = e^{-\int_0^T (\beta + m_0 + \frac{\sigma^2}{2} - \gamma\sigma^2 + \sigma^2 b\lambda(t)) dt} E^Q [e^{-\int_0^T (\int_0^t \sigma^2 (\gamma\lambda_u - b\lambda_u^2) du - \int_0^t \lambda_u \sigma dz_u^Q) dt}]. \quad (78)$$

Now, note that:

$$E^Q [e^{-\int_0^T \int_0^t \lambda_u \sigma dz_u^Q dt}] = E^Q [e^{-\int_0^T \int_u^T \lambda_t \sigma dt dz_u^Q}] \quad (79)$$

$$= e^{\frac{1}{2} \int_0^T (\int_u^T \lambda_t \sigma dt)^2 du}. \quad (80)$$

Thus, we get the final solution for the zero-coupon bond price:

$$P(0, T) = e^{-\int_0^T (\beta + m_0 + \frac{\sigma^2}{2} - \gamma\sigma^2 + \sigma^2 b\lambda(t)) dt + \int_0^T \int_0^t \sigma^2 (\gamma\lambda_u - b\lambda_u^2) du dt + \frac{1}{2} \int_0^T (\int_t^T \lambda_u \sigma du)^2 dt}. \quad (81)$$

Now, we can do a lot of the integrals explicitly. By plugging in the expression for λ

$$\frac{1}{\lambda(t)} = \frac{1}{\lambda_0} + t$$

we find

$$P(0, T) = (\lambda_0 T + 1)^{\frac{\sigma^2(2\gamma(\lambda_0 T + 1) - 1)}{2\lambda_0}} e^{-T(b\lambda_0\sigma^2 + \beta + m_0)} \quad (82)$$

and corresponding yield curve:

$$Y(0, T) = \frac{\sigma^2(\lambda_0 T(2b\lambda_0 + 2\gamma - 1) + (1 - 2\gamma(\lambda_0 T + 1)) \log(\lambda_0 T + 1))}{2\lambda_0 T} \dots \quad (83)$$

$$+ \beta - \gamma\sigma^2 + m_0 + \frac{\sigma^2}{2}. \quad (84)$$

3 Further discussion of the "learning about the mean"-economy

Here we discuss in detail (1) the effect of truncation of the prior beliefs about μ in the $EIS = 1$ case, and (2) the effect of increasing the EIS above 1 in the case when agents are learning about the mean growth rate in the economy discussed in Section 3 in the main paper.

The upshot on (1) is that fairly tight truncation bounds, such as $\mu \in (-0.3, 1.2\%)$ have a minimal effect on the impact of parameter uncertainty in long samples—the truncation only has noticeable effects for very dispersed prior beliefs. Thus, our results are not driven by extreme values of the parameter(s) agents are uncertain about.

Regarding point (2) the upshot is that increasing the EIS increases return volatility, which in turn increases the risk premium. In addition, there is an interesting interaction effect between the endogenous discount rate and the effect of shocks to the mean growth rate to the wealth-consumption ratio which makes the pricing kernel more sensitive over time to shocks to beliefs. This is different from the case where the $EIS = 1$, where the sensitivity of the pricing kernel to updates in beliefs is constant, and means that the asset pricing implications of learning decrease over time at a slower rate than the posterior variance of beliefs.

3.1 The effect of truncating the prior distribution

The normally distributed prior over μ employed in Section 3 of the main paper implies that the support of the beliefs about μ covers the whole real line. One may therefore wonder if the substantial increase in risk we document is due to the agent assigning positive probability on extreme (and unrealistic) values of μ . To investigate this, we in this section consider a truncated normal prior for μ . In particular, let $\bar{\mu}$ denote the upper truncation bound and $\underline{\mu}$ denote the lower truncation bound.

Conveniently, the updating equations for the hyperparameters, $\mu_{t+1} = \mu_t + \frac{A_t}{1+A_t} (\Delta c_{t+1} - \mu_t)$ and $A_{t+1}^{-1} = A_t^{-1} + 1$, remain the same. The reason for this is easiest to see by considering Bayes' rule when the truncation of a general, untruncated prior, $p(\theta)$, is achieved by multiplying by an indicator function which takes the value 1 if $\theta \in [\underline{\theta}, \bar{\theta}]$:

$$p(\theta|y^t) \mathbf{1}_{\underline{\theta} \leq \theta \leq \bar{\theta}} \propto p(y^t|\theta) p(\theta) \mathbf{1}_{\underline{\theta} \leq \theta \leq \bar{\theta}}. \quad (85)$$

The conjugacy of the prior comes from the functional forms of the original prior and the likelihood,

$p(\theta)$ and $p(y^t|\theta)$. Thus, if the likelihood function is normal and the prior is truncated normal, the posterior is truncated normal with the same truncation limits as the prior. The hyperparameters, μ_t and A_t , along with the truncation limits, completely describe the solution. Of course, the hyperparameter μ_t , for instance, no longer in general corresponds to the subjective conditional mean of consumption growth.

Unfortunately, we no longer can obtain analytical solutions with a truncated prior distribution, and therefore have to resort to a numerical solution of the model. So, before we give results for the economy with truncated prior beliefs and compare these to the case of untruncated prior beliefs, we describe the solution technique we employ.

3.1.1 Model solution

Note that since the posterior variance is deterministic ($A_{t+1}^{-1} = A_t^{-1} + 1$), we can replace the state variable A_t with time, t . Thus, the parameter uncertainty model is nonstationary. Building on Johnson (2007), who consider the simpler case of power utility, we develop a solution methodology for the case of parameter uncertainty and Epstein-Zin preferences where we solve the model using a backwards recursion from the known-parameters ($t = \infty$) boundary economies.⁴ In particular, the log value function to consumption ratio, $vc(\mu_t, t) = \ln(V_t/C_t)$, is found for each t on a grid for $\mu_t \in [\underline{\mu}, \bar{\mu}]$ using the recursion:

$$vc(\mu_t, t) = \beta \ln E_t \left[e^{\alpha(vc(\mu_{t+1}, t+1) + \Delta c_{t+1})} \right]^{1/\alpha}, \quad (86)$$

where the evolution equations of the state variables are $\mu_{t+1} = \mu_t + \frac{A_t}{1+A_t} (\Delta c_{t+1} - \mu_t)$ and $A_{t+1}^{-1} = A_t^{-1} + 1$. The boundary condition for a given value of μ is $vc(\mu) = \frac{\beta}{1-\beta} (\mu + \frac{1}{2}\alpha\sigma^2)$. This backwards recursion is fast and very accurate. The only additional requirement is continuity at the boundary when going from the boundary solutions where $t = \infty$ to a large t (we use $t = 5000$ as the point in time before the known-parameter boundary is reached). This general solution method is used throughout the paper.

3.1.2 Truncated vs. Untruncated Priors

We consider a truncated Normal prior over μ , where $\underline{\mu} = 0\%$ is the lower truncation bound for μ , and $\bar{\mu} = 1\%$ is the upper truncation bound. The true parameters in our quarterly calibration are $\mu = 0.45\%$ and $\sigma = 1.35\%$. Figure 1 shows the conditional, annualized price of risk and risk-free

⁴See Vazquez-Grande (2009) for a similar solution technique.

rate for the case of untruncated and truncated priors. The initial prior $A_0 = 1/40$ implies a prior dispersion consistent with 10 years of prior learning. Thereafter, the precision of the prior increases with time (in quarters given the quarterly calibration) according to $A_{t+1}^{-1} = 1 + A_t^{-1}$. The current mean belief is fixed at truth throughout, $\mu_t = \mu$.

**Figure 1 - The Effects of Prior Truncation:
The Price of Risk and Risk-free rate**

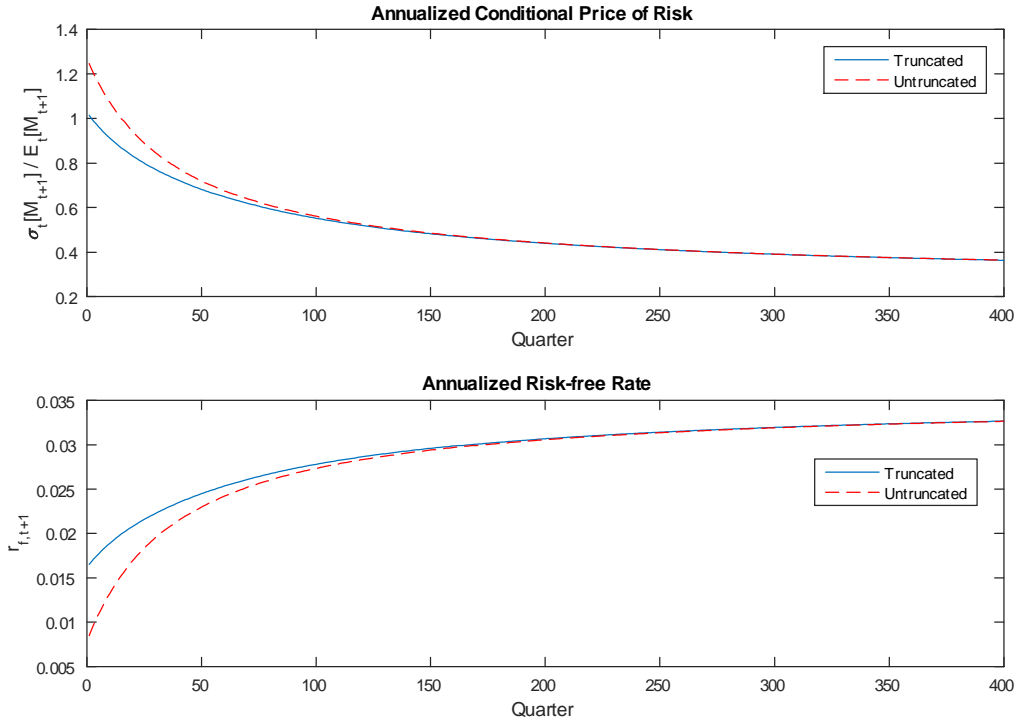


Figure 1: The graph shows the conditional price of risk and risk-free rate for the i.i.d. consumption growth economy when the mean growth rate is unknown and agent’s have unit elasticity of intertemporal substitution and $\gamma = 10$ and $\beta = 0.994$. The solid line shows the case when the prior distribution has upper and lower truncation bounds at 0% and 1%, respectively. The dashed line shows the case of no truncation bounds. The x-axis denotes sample length (in quarters) after an initial prior of $A_0 = 1/40$. The conditional mean belief about μ is fixed at the true value over time.

As the figure shows, truncation has some effect early in the sample—i.e., for very dispersed prior beliefs. This is intuitive, as this is when the truncation bounds are more influential in terms of the shape of the prior distribution. The price of risk is somewhat higher without truncation, which leads to a somewhat lower risk-free rate due to higher precautionary savings demand. After

25 years of learning (at 100 quarters), the difference between the price of risk and risk-free rate across the truncated prior versus untruncated prior economies is negligible. Thus, for prior beliefs that correspond to reasonable training samples, and thus also for asset pricing moments over long samples. the truncation is not important. This shows that the mechanism we present in this paper is not due to extreme, unrealistic parameter values.

It is also useful to show this in terms of the parameter uncertainty metric we propose in Section 3 of the main paper and that we discuss in more detail in Section 4 of this Online Appendix. The metric for evaluating the implications of parameter uncertainty allows for analytical solutions to the case of truncated prior beliefs. In particular, as an approximation to the full learning case we calculate utility assuming that the value of μ is revealed to investors next period. It is immediate that when agents have a preference for early resolution of uncertainty this earlier resolution of the value of μ leads to a lower bound on the utility loss arising from parameter uncertainty, relative to the case where learning evolves gradually over time. In particular, for the case of an *untruncated* prior over μ , the value function in this case is $vc_t^{one} = \frac{\beta}{1-\beta} \left(m_t + \frac{1}{2} (1-\gamma) \left(\frac{1}{1-\beta} A_t + 1 \right) \sigma^2 \right)$, which can be compared to the value function in the full learning case (Equations (14)-(16)). For the *truncated* case where the value of μ is revealed next period, the log value function is:

$$vc_t^{one_truncate} = \frac{\beta}{1-\beta} \left(m_t + \frac{1}{2} (1-\gamma) \left(\frac{1}{1-\beta} A_t + 1 \right) \sigma^2 \right) \dots$$

$$\times \left[\frac{\Phi \left(\frac{\bar{\mu}-m_t}{\sqrt{A_t}\sigma} - \frac{1-\gamma}{1-\beta} \sqrt{A_t}\sigma \right) - \Phi \left(\frac{\mu-m_t}{\sqrt{A_t}\sigma} - \frac{1-\gamma}{1-\beta} \sqrt{A_t}\sigma \right)}{\Phi \left(\frac{\bar{\mu}-m_t}{\sqrt{A_t}\sigma} \right) - \Phi \left(\frac{\mu-m_t}{\sqrt{A_t}\sigma} \right)} \right]^{1/(1-\gamma)}. \quad (87)$$

Table 1 shows that in the untruncated case the welfare loss relative to Anticipated Utility when μ is revealed next period is about 60-80% of that of the full learning case, depending on the amount of prior parameter uncertainty. Recall that the welfare loss (α) can be interpreted as as teh premium a rational agent would be willing to pay to have the unknown parameters set equal to his prior mean beliefs with certainty at time zero.

Thus, most of the risk associated with parameter uncertainty remains even if it is resolved early. For the truncated case where the truncation bounds are 0% and 1% (i.e., 0% and 4% in annual terms), the rightmost column of Table 1 shows that the welfare loss is identical up to two decimal places to that when the prior for μ is untruncated (The third decimal reveal that with truncation the welfare loss is slightly lower). Thus, it is the permanent shocks arising from parameter *updating* that matters, not the conditional volatility or higher order moments of the one-period ahead subjective

distribution of consumption growth as in Weitzman (2007).⁵

**Table 1 - Learning about the Mean
Truncated vs. Untruncated Prior**

Table 1: This table...

Degree of parameter uncertainty		Welfare loss (α) for ($\beta = 0.994, \gamma = 10, \psi = 1$)			
$A_t = \frac{1}{t}$ t (quarters)	$\sigma_t(\mu) = \sqrt{A_t}\sigma$	Untruncated case:		Truncated case:	
		Full, sequential learning case	Revelation of μ in next period	Revelation of μ in next period	Difference with untruncated case
40	0.21%	53%	43%	43%	-0.6%
100	0.13%	29%	20%	20%	-0.3%
200	0.10%	17%	11%	11%	0.0%
400	0.07%	9%	6%	6%	0.0%
800	0.05%	5%	3%	3%	0.0%
∞	0%	0%	0%	0%	0.0%

3.2 The Effect of the Intertemporal Elasticity of Substitution

Increasing the EIS from 1 to, say, 2 has in typical calibrations only a minor effect on the sensitivity of the pricing kernel to shocks to the continuation utility. For instance, if $\gamma = 10$, as in Bansal and Yaron (2004), increasing the EIS from 1 to 2, means that this sensitivity, $\gamma - 1/\psi$, only increases from 9 to 9.5. However, when the substitution effect dominates the wealth effect, the wealth-consumption ratio increases upon a high realization of consumption growth as the expected mean growth rate is revised upwards. These dynamics create excess return volatility, which in turn increases the risk premium. Further, interestingly, the endogenous sensitivity of the continuation utility to updates in beliefs in this case increases over time, which leads to a slower decline in asset return Sharpe ratios over time than in the cases with a lower EIS. Before we discuss these effects, we first explain how the model is solved.

⁵We show in numerical analysis in the Online Appendix that this conclusion also holds when considering the full learning problem and also when EIS is greater than one.

3.2.1 Model solution

In the case when $EIS \neq 1$, it is necessary to resort to numerical solution of the model. Further, since the wealth and substitution effects now no longer cancel, it is now also necessary to bound the support for the beliefs about the mean growth rate, μ , in order to ensure the existence of equilibrium.⁶ Therefore, we consider a truncated normal prior distribution for μ , where the truncation bounds $\underline{\mu}$ and $\bar{\mu}$ are set such that equilibrium exists for all possible 'boundary' ($t = \infty$) economies, where the agents have learned the true mean. This ensures the existence of equilibrium. As discussed earlier, the updating equations for the hyperparameters, $\mu_{t+1} = \mu_t + \frac{A_t}{1+A_t} (\Delta c_{t+1} - \mu_t)$ and $A_{t+1}^{-1} = A_t^{-1} + 1$, remain the same.

As for the truncated prior $EIS = 1$ case, we solve the model using a backwards recursion from the known-parameters ($t = \infty$) boundary economies. In these $EIS \neq 1$ cases, we solve for the log wealth-consumption ratio, pc , which is found for each t on a grid for $\mu_t \in [\underline{\mu}, \bar{\mu}]$ using the recursion:

$$e^{pc(\mu_t, t)} = E \left[\beta^\theta e^{-\gamma \Delta c_{t+1} + (\theta-1) \ln(\exp(pc(\mu_{t+1}, t+1)) + 1)} \mid \mu_t, t; A_0, \underline{\mu}, \bar{\mu} \right], \quad (88)$$

where $\theta = \frac{\alpha}{\rho} = \frac{1-\gamma}{1-1/\psi}$.

3.2.2 Asset pricing moments versus amount of parameter uncertainty

To assess the asset pricing implications of varying degrees of parameter uncertainty in cases where $EIS \neq 1$, it is necessary to calibrate the consumption dynamics as the model relies on a numerical solution. For this exercise, we choose the true consumption dynamics to match the mean and volatility of time-averaged annual U.S. log, per capita consumption growth, as reported in Bansal and Yaron (2004): $E_T[\Delta c] = 1.8\%$ and $\sigma_T(\Delta c) = 2.72\%$. This implies true (not time-averaged) quarterly mean and standard deviation of 0.45% and 1.65%, respectively, which are the numbers we use in the quarterly calibration. The lower bound for the quarterly growth rate μ is set at -0.3% , while the corresponding upper bound is set at 1.2%. The prior beliefs are assumed to be unbiased and the maximum level of prior uncertainty is $A_0 = 1$, as before. The market claim we consider is a levered consumption claim. In particular, we simply multiply the excess returns on the consumption claim with 1.5, which is consistent with the average aggregate leverage ratio in

⁶The reason is that a positive probability of an arbitrarily high (low) μ when the IES is greater than (smaller than) 1, leads to a violation of a transversality condition and the equilibrium does not exist (i.e., the wealth-consumption ratio is infinite). This is easiest to see by considering a deterministic economy with a constant growth rate. If the growth rate is higher than the risk-free rate, the wealth-consumption ratio, and thus utility, is infinite.

the U.S. stock market.⁷

Figure 2 - Price of Risk for case of unknown mean growth rate when $IES \neq 1$

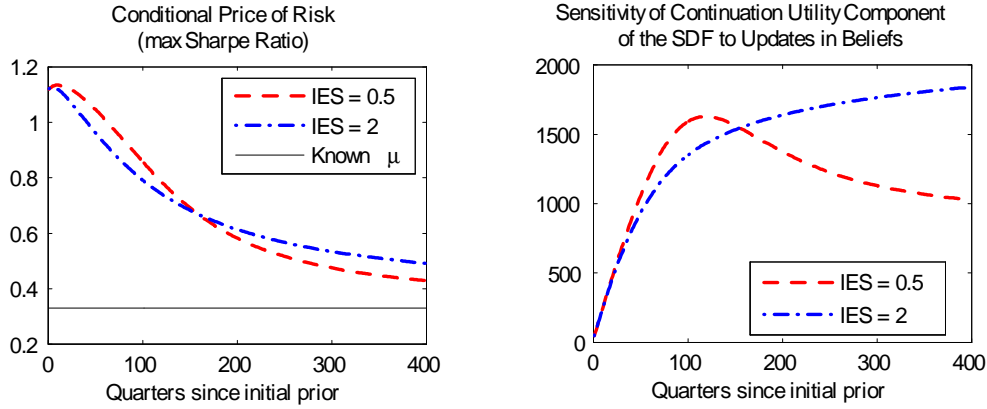


Figure 2: The graph shows the conditional price of risk as well as the sensitivity of the continuation utility component of the stochastic discount factor (SDF), as defined in the main text, to updates in the mean beliefs about the growth rate of the economy (μ). The horizontal axis gives the number of quarters passed since the initial prior. For both graphs the relevant statistics are evaluated for unbiased beliefs (i.e., at $E_t[\mu] = \mu$). The red dashed line gives the case when the intertemporal elasticity of substitution (IES) is 0.5, the blue dash-dotted line gives the case when the IES is 2, while the black solid line gives the case for which the mean is known (note that in the latter case the IES does not matter for the moments shown here).

The left plot in Figure 2 shows that the conditional, annualized price of risk (that is, $\sigma_t(M_{t+1})/E_t(M_{t+1})$; see Hansen and Jagannathan (1992)) is constant at 0.33 in the known parameters benchmark economy (solid line). This is true for any level of the EIS since consumption growth is i.i.d. In the two economies with parameter uncertainty, however, the price of risk decreases from about 1.15 to about 0.45. Thus, as in the $EIS = 1$ case discussed earlier, the price of risk decreases with the decreasing amount of parameter uncertainty (see Equation (4)). However, the component of the price of risk due to shocks to beliefs only decreases by a little less than a factor of 3 over the first 50 years in the cases shown in Figure 2, relative to a factor of 200 in the $EIS = 1$ case.⁸

⁷Note that we do not add idiosyncratic dividend growth shocks. Thus, the return volatility of the market claim ought not be compared directly to equity market return volatility. The risk premium, however, which is a function of the covariance of dividend and consumption growth, can reasonably be compared to the equity market risk premium.

⁸In the plot, the price of risk decreases from about 1.15 to about 0.6 over the first 200 quarters. Since the component due to learning is given by the amount in excess of 0.33 (the case of known parameters), the decrease is from about 0.8 to about 0.3, which is a little less than a factor of 3.

This much slower decline in the price of risk at the beginning of the sample is mainly due to the effect of truncation bounds, which alters the prior distribution. In particular, for high A_t the prior is close to a uniform distribution with upper and lower bounds $\underline{\mu}$ and $\bar{\mu}$. Thus, extreme values are ruled out, which decreases the price of risk, but at the same time, learning is slower as the signal to noise ratio in this case is decreased, relative to the untruncated case.

After about 100 quarters the effect of the truncation is negligible as, for $\mu_t = \mu$, the truncation bounds are 5 standard deviations away given the tighter prior at this time. At this point, there is an interesting relation between the price of risk and the level of the EIS. In particular, the price of risk decreases more slowly with time with a higher EIS (the red, dashed line has EIS = 0.5, while the blue, dashed-dotted line has EIS = 2). This is the result of an endogenous interaction between the amount of parameter uncertainty, the effect of such uncertainty on the volatility of the pricing kernel, and the level of the EIS.

The rightmost plot of Figure 2 shows the sensitivity of the continuation utility component of the log pricing kernel to shocks to beliefs as a function of time (prior variance) and for the two levels of the EIS (0.5 and 2). This sensitivity is calculated numerically from each model as $(\theta - 1) \frac{dpc_t}{d\mu_t} |_{\mu_t = \mu}$. From the analytical EIS = 1 case (see Equation (18)), the relevant sensitivity is given by $(\gamma - 1) \frac{\beta}{1 - \beta} = 1,495$ and is thus constant. From the rightmost plot of Figure 2, focusing on the dynamics after 100 quarters of learning has taken place and the effects of truncation are negligible, it is clear that this sensitivity is increasing for the high EIS case, but decreasing for the low EIS case. In fact, this time-varying sensitivity is due to an endogenous interaction between discount rates and the size of the shock to the growth rate, μ_t . As is well understood, the price-consumption ratio is more sensitive to shocks to the growth rate if discount rates are low (see, e.g., Pastor and Veronesi (2002)). When the EIS is greater than one and the substitution effect dominates, the price-consumption ratio increases when uncertainty decreases (see also Bansal and Yaron (2004)) and so, holding $\mu_t = \mu$, the price-consumption ratio in this case increases over time, meaning that discount rates decrease and, thus, the sensitivity to belief shocks increases. When the EIS is less than one, however, the price-consumption ratio is decreasing over time as discount rates now increase when uncertainty decreases, and this makes the pricing kernel less sensitive to shocks to beliefs over time. These endogenous dynamics are why after 100 years of learning the component of the price of risk that is due to belief shocks is almost twice as big for EIS = 2 versus EIS = 0.5, even though the direct risk price of shocks to the continuation utility $(\gamma - 1/\psi)$ when the EIS = 2 is only 1.19 times the same when the EIS = 0.5.

Figure 3 shows the conditional, annualized market risk premium, log return volatility, as well

APPENDIX: Figure 3 - Conditional Moments for case of unknown mean growth rate when $IES \neq 1$

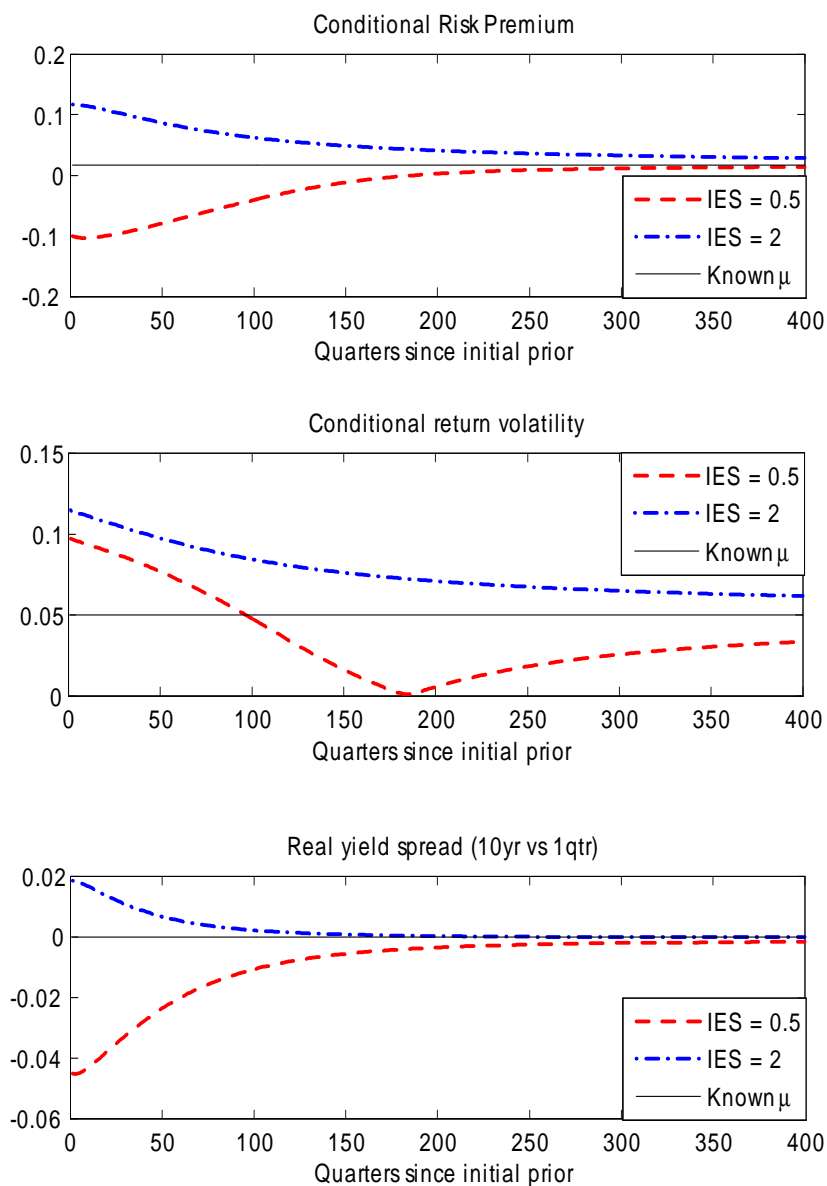


Figure 3: The graph shows the conditional risk premium, return volatility and real yield spread for the economy with unknown mean growth rate (μ). The horizontal axis gives the number of quarters passed since the initial prior. For all graphs the relevant statistics are evaluated for unbiased beliefs (i.e., $E_t[\mu] = \mu$). The red dashed line gives the case when the intertemporal elasticity of substitution (IES) is 0.5, the blue dash-dotted line gives the case when the IES is 2, while the black solid line gives the case for which the mean is known (note that in the latter case the IES does not matter for the moments shown here).

as the log real yield spread, for different economies versus the amount of parameter uncertainty, as measured by quarters passed since the initial prior, $A_0 = 1$. The conditional moments are always calculated assuming an unbiased prior, $\mu_t = \mu$. The yield spread is defined as the difference between the annualized 10-year yield on a real, default-free zero-coupon bond minus the current (quarterly), annualized real risk-free rate. The economies with parameter uncertainty have EIS $\psi = \{0.5, 2\}$, and as before we let $\gamma = 10$ and $\beta = 0.994$. As a benchmark, the black solid line depicts the case where μ is known (in which the EIS does not matter for any of the reported moments, given the assumed i.i.d. consumption growth process).

For the high EIS case ($\psi = 2$), the risk premium is initially about 12%, declining as posterior variance decreases, and after 100 years of learning about 3%. For comparison, the case where μ is known yields a risk premium of 1.7% (for any level of the EIS). Thus, there is again a significant effect of parameter uncertainty even after 100 years of learning, even though at this point the standard deviation of beliefs about μ is only 0.09%.

For the low EIS case ($\psi = 0.5$), however, the risk premium is initially negative and increasing as posterior variance decreases. When the EIS is less than one, the wealth effect dominates, and therefore a positive revision in beliefs about μ , resulting from high realized consumption growth, is accompanied by a *decrease* in the price-dividend ratio of this claim. For high levels of parameter uncertainty, the return on this claim is in fact negatively correlated with realized consumption growth, which leads to a negative risk premium. As the posterior variance tightens over time, the updates in the price-dividend ratio become smaller and the realized dividend dominates, which restores the positive correlation between the return to this claim and consumption growth.

The middle plot in Figure 3 shows how these dynamics play out for the volatility of returns. With a high EIS, when the substitution effect dominates, a positive shock leads to both high dividends and an increase in the price-dividend ratio of the claim, which in turn means return volatility is high (excess volatility). After 100 years of learning, the volatility is 6.2% versus 5% in the benchmark, known parameters economy. With a low EIS, the response of the price-dividend ratio and realized dividend growth offset to some extent and volatility is therefore always lower than in the case with an EIS > 1 . The lowest level of volatility occurs when the level of parameter uncertainty is such that the two offset almost exactly, as seen at around 180 quarters of learning.

The bottom plot in Figure 3 shows that the yield spread is positive initially and then basically zero for the case of high EIS, but negative initially for the case of a low EIS. With a low EIS, the volatility of the risk-free rate is higher, which leads to more volatile bond returns. The negative risk premium on real bonds, which are hedges in this economy as discussed earlier, is therefore higher in the case if a low EIS, which leads to a negative yield spread. With a high EIS, however, the effect

of expected future higher short-term rates, as precautionary savings decreases over time, dominates initially.

In sum, with an EIS greater than one, the model delivers excess return volatility and a higher risk premium than in the case when $EIS = 1$, as analyzed earlier. Further, the real term structure is, except for in cases with very high parameter uncertainty, essentially flat. Again, this is in sharp contrast to the objective long-run risks assumed in Bansal and Yaron (2004), which implies a strongly downward-sloping yield curve. A low EIS generates a counter-factual negative risk premium and also tends to deliver return volatility that's lower than dividend volatility, which is strongly counter-factual (see, e.g., Shiller (1981), Campbell and Shiller (1988)).

3.3 Unknown variance

In the preceding, the variance parameter σ^2 was assumed known to investors. It is straightforward to relax this assumption, though as pointed out in Weitzmann (2007) and Bakshi and Skouliakis (2010), it is necessary to truncate also the support for σ^2 in order to ensure finite utility. Weitzmann (2007) argues that learning about the variance parameter can lead to arbitrarily high risk premiums as the subjective distribution for consumption growth becomes fat-tailed. He further argues that learning about the mean, as in the preceding section, does not increase the fatness of the tails of the conditional consumption growth distribution and therefore cannot help in explaining asset pricing puzzles. Clearly, the latter intuition does not hold when considering a utility function that allows for a preference for early resolution of uncertainty.⁹

Bakshi and Skouliakis (2010) argue that Weitzmann's results, which are developed under power utility, are not robust to reasonable truncation limits for σ^2 . However, given that we focus primarily not on the fatness of the tails, but on permanent shocks to the conditional consumption growth distribution induced by the learning process itself, uncertain variance can potentially still have important asset pricing implications. In the following, we show that quantitatively large asset pricing implications of learning about the variance parameter indeed can arise, but that interesting asset pricing effects of learning about the variance parameter are substantially shorter-lived than those documented for the uncertain mean case.

We assume that the joint prior over the mean μ and the variance σ^2 is Normal-Inverse-Gamma:

$$p(\mu, \sigma^2 | y^t) = p(\mu | \sigma^2, y^t) p(\sigma^2 | y^t), \quad (89)$$

⁹In fact, with a truncated normal as the prior, the tails of the subjective distribution are actually less fat than for a normal distribution with the same dispersion, but due to the updating that generates long-run risks, the asset pricing implications were shown to be nontrivial.

where

$$p(\sigma^2|y^t) \sim IG\left(\frac{b_t}{2}, \frac{B_t}{2}\right), \quad (90)$$

$$p(\mu|\sigma^2, y^t) \sim N(\mu_t, A_t\sigma^2). \quad (91)$$

Given that log consumption growth is normally distributed, these prior beliefs lead to posterior beliefs that are of the same form (conjugate priors). The updating equations for investors' beliefs are:

$$A_{t+1}^{-1} = 1 + A_t^{-1}, \quad (92)$$

$$\frac{\mu_{t+1}}{A_{t+1}} = \frac{\mu_t}{A_t} + y_{t+1}, \quad (93)$$

$$b_{t+1} = b_t + 1, \quad (94)$$

$$B_{t+1} = B_t + \frac{(y_{t+1} - \mu_t)^2}{1 + A_t}. \quad (95)$$

In terms of pricing, note that this system can be reduced to three state-variables: μ_t , B_t , and t , given initial priors. We solve the model numerically and, as before, use the closed-form solution for the known parameters cases as the boundary values in a recursion that is solved backwards in time on a grid for μ_t and B_t . In order for the Inverse Gamma distribution to have a finite mean and variance, which is convenient, we set the maximum prior uncertainty as $b_0 = 5$. As mentioned, we need to truncate the distribution for σ^2 and we choose wide bounds: $\overline{\sigma^2} = 100 * \sigma^2$, $\underline{\sigma^2} = \sigma^2/100$. As before, the true quarterly variance is calibrated as $\sigma^2 = (1.65\%)^2$, and the model is solved at the quarterly frequency. The other parameters of the model are the same as in the case where the mean was the only unknown parameter: $\mu_0 = \mu = 0.45\%$, $A_0 = 1$, $\gamma = 10$, $\psi = 2$, and $\beta = 0.994$. We set $b_0 = 5$ and $\frac{B_0}{b_0-2} = \sigma^2$. The latter implies that the initial truncated prior for the variance is unbiased, with a standard deviation of $(1.85\%)^2$.

Figure 4 shows the conditional annualized volatility of the log pricing kernel as the average per quarter across 20,000 simulated economies over a 100 year sample. We plot three cases. Learning about the mean only, as discussed in the previous section, learning about the variance only, and learning about the mean and the variance parameters. First, consider the dashed line, which shows the case when learning about the variance only. The volatility of the pricing kernel is very high in the first decade, but then comes down quite quickly towards the benchmark, known parameter

**Figure 4 - Conditional Volatility of the Pricing Kernel:
Cases with unknown variance**

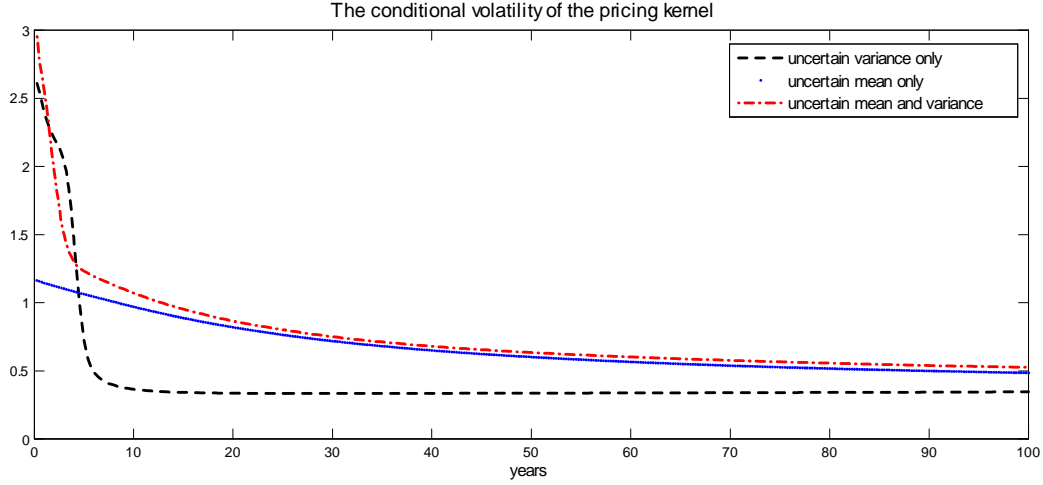


Figure 4: The graph shows the subjective conditional annualized volatility of the Epstein-Zin stochastic discount factor with preference parameters $\gamma = 10$, $\psi = 2$ and $\beta = 0.994$ over a 100 year sample period, averaged across 20,000 simulated economies at each time t . The dashed line corresponds to the case of unknown variance only, the dotted line corresponds to the case of unknown mean only, while the dash-dotted line corresponds to the case of unknown mean and variance.

value of 0.33.¹⁰ Pretty much all of this pattern comes from the continuation utility component of the pricing kernel and not from the power utility component. Thus, we confirm the results in Bakshi and Skouliakis (2010), the fatness of the tails given reasonable bounds on the variance parameter is not sufficient to strongly affect asset prices. However, the large updates in beliefs about the variance that occur in the first 10 years does have significant impact on the volatility of the pricing kernel through the effect on the continuation utility. After this, the impact of shocks to beliefs about the variance parameter have a very small impact. The dash-dotted line shows the case of unknown mean *and* variance. Here, we see that adding unknown variance yields a pricing kernel that is on average always more volatile than in the known variance, unknown mean case. However, there are only large differences in the first decade, relative to the case with only unknown mean (dotted line).

The risk premium for a 100 year long sample that start with priors corresponding to tossing out

¹⁰The somewhat uneven line for the variance cases in the 5 first years is due to the truncation bounds slightly affecting the form of the subjective distribution for the variance parameters when the level of uncertainty is very high.

the 10 first years plotted in Figure 4, is 1.8% for the case of unknown variance but known mean, relative to 1.7% for the benchmark known parameters economy. In the case of unknown mean and variance, the average risk premium over this sample is 4.9% compared to 4.4% for the case of unknown mean and known variance.

In sum, unknown variance has more of a second-order effect on asset pricing moments, unless uncertainty is very large, as would be the case in the decade after a structural break for instance. There are two reasons for this more short-lived effect. First, Bayesian learning implies that learning about variance is much faster than learning about the mean. Second, the variance is a second order moment, so generally less important for the continuation utility than changes in the mean.

4 A Metric for Quantifying the Impact of Parameter Uncertainty

Although the i.i.d. normal model is useful for intuition, realistic consumption-based asset pricing models feature more complicated consumption dynamics with many parameters and an EIS different from one. Prominent examples include the long-run risk models of Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012), as well as the rare disaster models of Barro (2009), Wachter (2013), and Barro, Nakamura, Steinsson and Ursua (2012). Since these models rely on more complicated and difficult-to-estimate features of consumption dynamics, parameter uncertainty is a particularly relevant concern.

This section presents a simple and intuitive metric for quantifying the impact of parameter uncertainty that does not require solving a full, long-horizon learning problem. This metric can guide the researcher in terms of deciding which parameters should be treated as unknown in ex-ante calculations of utility and asset prices, and transparently give the economic intuition for why uncertainty over certain parameters is economically important, while uncertainty over others is not. This problem is related to Chen, Dou, and Kogan (2013), who provide an approach for understanding the economic importance of statistically hard to measure parameters in asset pricing models.

4.1 The Metric

The metric quantifies how much an agent would pay to exchange consumption dynamics with parameter uncertainty for consumption dynamics with no parameter uncertainty, similar in spirit to Lucas (1985), who considers how much an agent would pay to eliminate risk.

The metric's usefulness lies in the alternative consumption stream specifications. First, the AU version of the model provides the benchmark known-parameters case, as this is the default option researchers would otherwise use. Second, we consider a simplified parameter learning economy in which parameters are fully revealed in the *next* period. This can be contrasted with the real learning process whereby agents gradually learn about parameters over time. This modification allows for fast and relatively easy calculation of the metric. Since the relevant calibrations in the literature imply that agents either have no preference for the timing of resolution of uncertainty or a preference for early resolution of uncertainty, the utility arising from learning true parameter values early (next period) is an upper bound for the utility in the gradual learning case. Therefore, the amount the agent would pay to receive the consumption implied by the AU approach relative to the learning problem with early resolution, is a lower bound relative to the amount the agent would pay in the case of gradual learning.

The metric is formally defined as follows. Let V_t^{AU} be the utility level corresponding to the AU case and V_t^{One} be the utility level for the case where parameters are uncertain but revealed in one period. The *parameter uncertainty premium*, π_t , is the fraction of wealth the agent would be willing to forego to avoid parameter uncertainty and have parameters set to current mean parameter beliefs:

$$\pi_t = 1 - \frac{V_t^{One}}{V_t^{AU}}. \quad (96)$$

The utility levels are

$$V_t^{One} = \{(1 - \beta) C_t^\rho + \beta (\mu_{One,t})^\rho\}^{1/\rho} \quad \text{and} \quad V_t^{AU}(\bar{\theta}) = \{(1 - \beta) C_t^\rho + \beta (\mu_{AU,t})^\rho\}^{1/\rho}, \quad (97)$$

and the crucial certainty equivalents in the EZ preferences are

$$\mu_{One,t} = \left\{ \int E_t [V_{t+1}(\theta)^\alpha | \theta] p(\theta | y^t) d\theta \right\}^{1/\alpha} \quad \text{and} \quad \mu_{AU,t} = \{E_t [V_{t+1}(\bar{\theta})^\alpha | \bar{\theta}] \times 1\}^{1/\alpha}. \quad (98)$$

Here, $\bar{\theta} = \int \theta p(\theta | y^t) d\theta$ where $p(\theta | y^t)$ denotes the prior distribution for the parameter vector θ , and $V_{t+1}(\theta)$ denotes the utility at time $t + 1$ when parameters are known and equal to θ . The ' $\times 1$ ' that appears in the expression for $\mu_{AU,t}$ is included to emphasize the fact that AU applies a point-mass prior centered at the mean parameter belief when calculating utility and pricing assets.

This metric is intuitive and easy to calculate for any preference parameters and for general forms of the prior $p(\theta | y^t)$ as it only requires integrating the utility in the known parameters case over the prior beliefs to calculate the certainty equivalent $\mu_{One,t}$.

4.2 The Intuition

The definition of π_t shows that the utility impact of parameter uncertainty is a function of (1) the sensitivity of the known-parameters utility $V_{t+1}(\theta)$ to changes in the parameter vector θ , (2) the belief or prior distribution, and (3) the form of the certainty equivalent. The first component is related to the measure proposed by Chen, Dou, and Kogan (2013), who argue that many asset-pricing models are highly sensitive to difficult-to-estimate parameters. The second component is also intuitive—greater prior dispersion, for example, leads to a larger impact of parameter uncertainty on utility. The third certainty equivalent component is important for understanding the pricing and utility effects of parameter uncertainty.

To gain intuition for the impact of the certainty equivalent component, consider a discrete parameter case where $\theta \in (\theta_L, \theta_H)$ with $p_t = \mathbb{P}[\theta = \theta^H | y^t]$. In this case, $\bar{\theta} = p_t \theta^H + (1 - p_t) \theta_L$. The certainty equivalent for resolution next period is

$$\mu_{One,t} = \{p_t E_t [V_{t+1}(\theta^H)^\alpha] + (1 - p_t) E_t [V_{t+1}(\theta_L)^\alpha]\}^{1/\alpha}, \quad (99)$$

where the conditional expectations within this expression are needed for other state-variables realized at time $t + 1$ that affect utility, such as consumption. Clearly, if $\theta^H = \theta^L$, $\mu_{One,t} = \mu_{AU,t}$. Holding constant the mean utility level the same across the two cases (i.e. holding $E_t [V_{t+1}(\bar{\theta})] = p_t E_t [V_{t+1}(\theta^H)] + (1 - p_t) E_t [V_{t+1}(\theta_L)]$ constant), it follows that a larger difference between expected utility levels for the θ^H -case versus the θ_L -case implies a lower value for $\mu_{One,t}$, given standard intuition for the effect of volatility on the certainty equivalent function. Higher risk aversion amplifies this effect. A similar intuition holds for more negatively skewed utility outcomes, which also are penalized in the certainty equivalent function.

This intuitively suggests that the utility and asset pricing impact of parameter uncertainty will be greatest for difficult-to-learn parameters (high prior dispersion) and for parameters that generate very low utility outcomes.¹¹ While this is intuitive, the key is that the metric allows researchers to easily *quantify* the impact, something that is quite costly to do in general. This quantification is the key to the metric. Researchers can thus, as we do in Section 4 of the main paper, use this metric to identify the parameters in their model for which accounting for parameter uncertainty is particularly important in utility and price calculations.

¹¹We also note that parameter uncertainty can affect the expected utility level, $E_t [V_{t+1}(\theta)]$, if utility is highly nonlinear in a parameter. This latter effect is then less about risk and more about intertemporal substitution effects.

5 Numerical Solution Methodology for the Switching Regime Model

5.1 The case of unknown transition probabilities

Aggregate consumption growth is given by:

$$\Delta c_{t+1} = \mu(s_{t+1}) + \sigma(s_{t+1}) \varepsilon_{t+1}, \quad (100)$$

where $\varepsilon_{t+1} \stackrel{i.i.d.}{\sim} N(0, 1)$ and where $s_t \in \{1, 2\}$ follows a 2-state observable Markov chain with constant transition probabilities:

$$\Pi = \begin{bmatrix} \pi_{11} & 1 - \pi_{11} \\ 1 - \pi_{22} & \pi_{22} \end{bmatrix}, \quad (101)$$

with $\pi_{ii} \in (0, 1)$. The regime changes are assumed to be independent of the Gaussian shocks.

The agent knows the parameters within each state $(\mu_1, \mu_2, \sigma_1, \sigma_2)$, but does not know the transition probabilities (π_{11} and π_{22}). At $t = 0$, the agent is given an initial, (potentially truncated) Beta-distributed prior over each of these parameters and thereafter updates beliefs sequentially upon observing the time-series of realized regimes, s_t . We denote the history of realized regimes up until time t as s^t . The prior Beta-distribution coupled with the realization of regimes, which are governed by constant probabilities, leads to a conjugate prior and so posterior beliefs are also Beta-distributed.

The probability density function of the Beta-distribution is:

$$p(\pi|a, b) = \frac{\pi^{a-1} (1 - \pi)^{b-1}}{B(a, b)}, \quad (102)$$

where $B(a, b)$ is the Beta function (a normalization constant). The parameters a and b govern the shape of the distribution. Of particular interest is the expected value:

$$E[\pi|a, b] = \frac{a}{a + b}. \quad (103)$$

In our case, there is one uncertain probability corresponding to each regime and a standard application of Bayes rule shows that the updating equations basically count the number of times state i has been followed by state i versus the number of times state i has been followed by state

j . Given this sequential updating, we let the a and b parameters have a subscript for the relevant state (1 or 2), as well as a time subscript. In particular:

$$a_{i,t} = a_{i,0} + \#(\text{state } i \text{ has been followed by state } i), \quad (104)$$

$$b_{i,t} = b_{i,0} + \#(\text{state } i \text{ has been followed by state } j). \quad (105)$$

When solving this problem numerically, we use the known parameters boundary economies (at $T = \infty$ when the parameters have been learned) as terminal values in a backwards recursion, following Johnson (2007).¹² We use the following state variables in the numerical solution :

$$\tau_{1,t} = a_{1,t} - a_{1,0} + b_{1,t} - b_{1,0} \quad (106)$$

$$\lambda_{1,t} \equiv E_t[\pi_{11}] = \frac{a_{1,t}}{a_{1,t} + b_{1,t}} \quad (107)$$

$$\tau_{2,t} = a_{2,t} - a_{2,0} + b_{2,t} - b_{2,0} \quad (108)$$

$$\lambda_{2,t} \equiv E_t[\pi_{22}] = \frac{a_{2,t}}{a_{2,t} + b_{2,t}}, \quad (109)$$

where the initial prior beliefs $(a_{1,0}, b_{1,0}, a_{2,0}, b_{2,0})$ are given as parameter inputs to the economy.

The equilibrium, recursive expression for the wealth-consumption ratio (PC) is standard in the Epstein-Zin case and is (when $\psi \neq 1$) given by:

$$PC_t^\theta = E_t \left[\beta^\theta e^{(1-\gamma)\Delta c_{t+1}} (PC_{t+1} + 1)^\theta \right], \quad (110)$$

where the subscript t here denotes dependence on information known at time t and $E_t[\cdot]$ denotes the conditional expectation given all information available at time t . We note that the state variables are s_t and X_t , where $X_t \equiv [\tau_{11,t}, \lambda_{1,t}, \tau_{22,t}, \lambda_{2,t}]$ are sufficient statistics for the agent's priors. Further, note from Equations (104) through (109) that we can write $X_{t+1} = f(s_{t+1}, s_t, X_t)$. Given this, we write the recursion equation (Equation (110)) as:

$$PC(s_t, X_t)^\theta =$$

¹²Johnson uses this approach in a case with parameter learning and power utility. We extend this to the case of Epstein-Zin utility. A similar approach has also been used by Vasquez-Grande (2009).

$$\begin{aligned}
& \beta^\theta E \left[e^{(1-\gamma)(\mu(s_{t+1})+\sigma(s_{t+1})\varepsilon_{t+1})} (PC(s_{t+1}, s_t, X_t) + 1)^\theta |s_t, X_t \right] \\
= & \beta^\theta E \left[E \left[e^{(1-\gamma)(\mu(s_{t+1})+\sigma(s_{t+1})\varepsilon_{t+1})} (PC(s_{t+1}, s_t, X_t) + 1)^\theta |s_{t+1}, s_t, X_t \right] |s_t, X_t \right] \\
= & \beta^\theta E \left[e^{(1-\gamma)\mu(s_{t+1})+\frac{1}{2}(1-\gamma)^2\sigma(s_{t+1})^2} (PC(s_{t+1}, s_t, X_t) + 1)^\theta |s_t, X_t \right] \\
= & \beta^\theta \sum_{s_{t+1}=1}^2 \Pr(s_{t+1}|s_t, X_t) e^{(1-\gamma)\mu(s_{t+1})+\frac{1}{2}(1-\gamma)^2\sigma(s_{t+1})^2} (PC(s_{t+1}, s_t, X_t) + 1)^\theta. \quad (111)
\end{aligned}$$

where the second to last equality uses the fact that regime changes and the Gaussian shocks to consumption growth (s_{t+1} and ε_{t+1}) are independent. Next, we need to find $\Pr(s_{t+1}|s_t, X_t)$. Denote the conditional density of π_{s_{t+1}, s_t} as $g(\pi_{s_{t+1}, s_t}|s_t, X_t)$. Then:

$$\begin{aligned}
\Pr(s_{t+1}|s_t, X_t) &= \int_0^1 \pi_{s_{t+1}, s_t} g(\pi_{s_{t+1}, s_t}|s_t, X_t) d\pi_{s_{t+1}, s_t} \\
&= E[\pi_{s_{t+1}, s_t}|s_t, X_t]. \quad (112)
\end{aligned}$$

Given the state variables described above, this expectation equals $\lambda_{s_t, t}$ or $1 - \lambda_{s_t, t}$.

Thus, the numerical backward recursion will be as follows:

$$PC(s_t, X_t)^\theta = \beta^\theta \sum_{s_{t+1}=1}^2 E[\pi_{s_{t+1}, s_t}|s_t, X_t] e^{(1-\gamma)\mu(s_{t+1})+\frac{1}{2}(1-\gamma)^2\sigma(s_{t+1})^2} (PC(s_{t+1}, s_t, X_t) + 1)^\theta, \quad (113)$$

where the boundary values for the wealth-consumption ratio for this backwards recursion are given by the limiting economies at $\tau_{11, \infty}$ and/or $\tau_{22, \infty}$, where π_{11} and/or π_{22} are known. This backward recursion is solved on a grid for λ_1 and λ_2 . It is important to have very dense grid for λ of each state as it approaches 1 for the numerical solution to be accurate. We use 100 grid points for each λ and the exact grid used can be obtained upon request.

5.1.1 Solving for a dividend claim

Let exogenous dividend growth be given by:

$$\begin{aligned}
\Delta d_{t+1} &= \mu + \lambda(\Delta c_{t+1} - \mu) - \frac{1}{2}\sigma_d^2 + \sigma_d \varepsilon_{d, t+1} \\
&= \lambda \Delta c_{t+1} + (1 - \lambda)\mu - \frac{1}{2}\sigma_d^2 + \sigma_d \varepsilon_{d, t+1}, \quad (114)
\end{aligned}$$

where $\mu \equiv E(\Pr(s_\infty = 1|\pi_{11}, \pi_{22}))\mu_1 + E(\Pr(s_\infty = 2|\pi_{11}, \pi_{22}))\mu_2$ and $\Pr(s_\infty = i)$ is the ergodic probability of being in state i . Thus, the leverage factor λ affects (relatively) short-run movements,

whereas the uncertainty about true long-run (unconditional) growth, which is a function of the uncertainty about the transition probabilities, is the same for the dividend claim as for the consumption claim. Note that the long-run mean under the agent's filtration is in fact random and its' $t + 1$ value can be expressed as $\mu(s_{t+1}, s_t, X_t)$. Finally, dividends have an idiosyncratic component given by the standard normal shock ε_d , which is assumed uncorrelated with any other shocks in the economy.

Solving for the price-dividend ratio of this claim is analogous to solving for the consumption claim. Note how the uncertainty about the infinite horizon dividend growth rate is the same as that for infinite horizon consumption growth rate as the exposure of dividend growth to $\mu(s_{t+1}, s_t, X_t)$ is always one, unaffected by the leverage parameter, λ . In particular:

$$\begin{aligned}
PD(s_t, X_t) &= \\
&= \beta^\theta E \left[e^{(\lambda-\gamma)(\mu(s_{t+1})+\sigma(s_{t+1})\varepsilon_{t+1})+(1-\lambda)\mu(s_{t+1},s_t,X_t))} \left(\frac{PC(s_{t+1},X_{t+1})+1}{PC(s_t,X_t)} \right)^{\theta-1} \dots \right] \\
&\quad \times (PD(s_{t+1}, s_t, X_t) + 1) |s_t, X_t \\
&= \beta^\theta E \left[E \left[\begin{aligned} &e^{(\lambda-\gamma)(\mu(s_{t+1})+\sigma(s_{t+1})\varepsilon_{t+1})+(1-\lambda)\mu(s_{t+1},s_t,X_t))} \dots \\ &\left(\frac{PC(s_{t+1},X_{t+1})+1}{PC(s_t,X_t)} \right)^{\theta-1} (PD(s_{t+1}, s_t, X_t) + 1) \end{aligned} \middle| s_{t+1}, s_t, X_t \right] |s_t, X_t \right] \\
&= \beta^\theta E \left[e^{(\lambda-\gamma)\mu(s_{t+1})+\frac{1}{2}(\lambda-\gamma)^2\sigma(s_{t+1})^2+(1-\lambda)\mu(s_{t+1},s_t,X_t)} \left(\frac{PC(s_{t+1},X_{t+1})+1}{PC(s_t,X_t)} \right)^{\theta-1} \dots \right] \\
&\quad \times (PD(s_{t+1}, s_t, X_t) + 1) |s_t, X_t \\
&= \beta^\theta \sum_{s_{t+1}=1}^2 \Pr(s_{t+1}|s_t, X_t) \left\{ \begin{aligned} &e^{(1-\gamma)\mu(s_{t+1})+\frac{1}{2}(1-\gamma)^2\sigma(s_{t+1})^2+(1-\lambda)\mu(s_{t+1},s_t,X_t)} \dots \\ &\left(\frac{PC(s_{t+1},X_{t+1})+1}{PC(s_t,X_t)} \right)^{\theta-1} (PD(s_{t+1}, s_t, X_t) + 1) \end{aligned} \right\}. \quad (115)
\end{aligned}$$

5.1.2 Limiting economies – boundary values for general case

All parameters known The simplest limiting economy is given by the case where both π_{11} and π_{22} are known. Since the state is observed and all the parameters are known, s_t is the only state variable and thus the wealth-consumption ratio can only take on two values. Solving this limiting economy amounts to solving two nonlinear equations in two unknowns ($PC(s = 1)$ and

$PC(s = 2)$):

$$PC(s = 1)^\theta = \beta^\theta \left(\begin{array}{l} \pi_{11} e^{(1-\gamma)\mu_1 + \frac{1}{2}(1-\gamma)^2\sigma_1^2} (PC(s = 1) + 1)^\theta \dots \\ + (1 - \pi_{11}) e^{(1-\gamma)\mu_2 + \frac{1}{2}(1-\gamma)^2\sigma_2^2} (PC(s = 2) + 1)^\theta \end{array} \right), \quad (116)$$

$$PC(s = 2)^\theta = \beta^\theta \left(\begin{array}{l} (1 - \pi_{22}) e^{(1-\gamma)\mu_1 + \frac{1}{2}(1-\gamma)^2\sigma_1^2} (PC(s = 1) + 1)^\theta \dots \\ + \pi_{22} e^{(1-\gamma)\mu_2 + \frac{1}{2}(1-\gamma)^2\sigma_2^2} (PC(s = 2) + 1)^\theta \end{array} \right). \quad (117)$$

These equations are relatively straightforward to solve numerically, imposing the requirement ex ante that the wealth-consumption ratio is positive and real. In our case, it was possible to verify that there is only one economically reasonable solution by plotting the function space for the grid for the π 's, given the other parameters as assumed in the main paper.

We solve these limiting equations for a grid on π_{11} and π_{22} with lower and upper bounds set to the 0.01% and 99.9% percentile values of the initial prior distribution for the general case we ultimately want to solve for, as given by $a_{i,0}$ and $b_{i,0}$, $i \in \{1, 2\}$.

One transition probability known, one unknown Another set of boundary economies are given by the case where one of the transition probabilities is known and the other is unknown. This case corresponds to $\tau_{11,t} < \infty$, $\tau_{22,\infty}$ or $\tau_{11,\infty}$, $\tau_{22,t} < \infty$. We can find the wealth-consumption for these cases using the backward induction as given by Equation (113):

$$PC(s_t, X_t)^\theta = \beta^\theta \sum_{s_{t+1}=1}^2 E[\pi_{s_{t+1},s_t} | s_t, X_t] e^{(1-\gamma)\mu(s_{t+1}) + \frac{1}{2}(1-\gamma)^2\sigma(s_{t+1})^2} (PC(s_{t+1}, s_t, X_t) + 1)^\theta, \quad (118)$$

where for the transition probability whose value is known, trivially $E[\pi_{s_{t+1},s_t} | s_t, X_t] = \pi_{s_{t+1},s_t}$. For instance, if π_{11} is known, then $E[\pi_{11} | s_t, X_t] = \pi_{11}$ and $E[1 - \pi_{11} | s_t, X_t] = 1 - \pi_{11}$ for all t . Also, in this case we have that $X_t = [\tau_{22,t}, \lambda_{2,t}]$.

From these boundary values, we iterate backwards in time (which here is a 2-dimensional concept, as we are recording time spent in each regime) to find the solution for the finite τ 's we ultimately are interested in.

5.2 The case of unknown mean and variance parameters

Aggregate consumption growth is given by:

$$\Delta c_{t+1} = \mu_{s_t} + \sigma_{s_t} \varepsilon_t, \quad (119)$$

where $\varepsilon_t \stackrel{i.i.d.}{\sim} N(0, 1)$ and where $s_t \in \{1, 2\}$ follows a 2-state observable Markov chain with constant transition probabilities:

$$\Pi = \begin{bmatrix} \pi_{11} & 1 - \pi_{11} \\ 1 - \pi_{22} & \pi_{22} \end{bmatrix}, \quad (120)$$

with $\pi_{ii} \in (0, 1)$. The regime changes are assumed to be independent of the Gaussian shocks.

In this case, the transition probabilities are known, but the mean and scale parameters in each state (μ_1, σ_1, μ_2 and σ_2) are unknown, with independent truncated normal-inverse-gamma priors. The priors and updating equations are analogous with those in discussed for the mean and variance parameters in the i.i.d. consumption growth model of Section 3 of this Online Appendix. The only difference is that you now only learn about state 1 parameters when you are in state 1, etc.:

$$\mu_{i,t+1} = \mu_{i,t} + \mathbf{1}_{s_{t+1}=i} \frac{A_{i,t}}{1 + A_{i,t}} (\Delta c_{t+1} - \mu_{i,t}), \quad (121)$$

$$A_{i,t+1}^{-1} = A_{i,t}^{-1} + \mathbf{1}_{s_{t+1}=i}, \quad (122)$$

$$b_{i,t+1} = b_{i,t} + \mathbf{1}_{s_{t+1}=i} \quad (123)$$

$$B_{i,t+1} = B_{i,t} + \mathbf{1}_{s_{t+1}=i} \frac{(\Delta c_{t+1} - \mu_{i,t})^2}{1 + A_{i,t}} \quad (124)$$

where $i \in \{1, 2\}$ and $\mathbf{1}_{x=y}$ is an indicator function that equals 1 if the condition in subscript is true and 0 otherwise.

Note that the posterior dispersion parameters $A_{i,t}$ are a deterministic function of the time spent in regime i , as are $b_{i,t}$. Thus, the posterior variance parameters are not both needed as sufficient statistics for the priors, and the state vector related to parameter uncertainty is "only" 6-dimensional: $X_t \equiv [\mu_{1,t}, B_{1,t}, b_{1,t}, \mu_{2,t}, B_{2,t}, b_{2,t}]$, where the state of the Markov chain, s_t , is an additional state variable in the economy.

We can now write the wealth-consumption ratio as $PC(s_{t+1}, X_{t+1}) = PC(s_{t+1}, \Delta c_{t+1}, X_t)$ in order to clarify that the evolution equations are a function of the two observables, Δc_{t+1} and s_{t+1} . The equilibrium recursion used to solve the model is then:

$$PC(s_t, X_t)^\theta = \beta^\theta E \left[e^{(1-\gamma)\Delta c_{t+1}} (PC(s_{t+1}, \Delta c_{t+1}, X_t) + 1)^\theta | s_t, X_t \right] \quad (125)$$

$$= \beta^\theta E \left[E \left[e^{(1-\gamma)\Delta c_{t+1}} (PC(s_{t+1}, \Delta c_{t+1}, X_t) + 1)^\theta | s_{t+1}, s_t, X_t \right] | s_t, X_t \right] \quad (126)$$

$$= \beta^\theta \sum_{s_{t+1}=1}^2 \Pr(s_{t+1} | s_t, X_t^s) \times \dots E \left[e^{(1-\gamma)\Delta c_{t+1}} (PC(s_{t+1}, \Delta c_{t+1}, X_t) + 1)^\theta | s_{t+1}, s_t, X_t \right]. \quad (127)$$

In words, since the transition probabilities are not a function of the realized consumption growth, it is useful to solve numerically first the inner expectation, which conditions on next period's state realization, and then integrate over conditional distribution about the states afterwards. In particular, since we know the transition probabilities, this last operation is very simple and does not require quadrature-type numerical methods. For the inner expectation, however, we have no analytical, simplifying expressions. In particular, since the price consumption ratio is not conditionally independent of the realized consumption growth, we cannot separate the terms as we could in the previous case, where only the transition probabilities were unknown.

The numerical solution for the inner conditional expectation is as follows:

$$E \left[e^{(1-\gamma)\Delta c_{t+1}} (PC(s_{t+1}, \Delta c_{t+1}, X_t) + 1)^\theta | s_{t+1}, s_t, X_t \right] \approx \sum_{j=1}^J w_\varepsilon(j) \left[\sum_{k=1}^K w_{\sigma_s^2}(k) \sum_{l=1}^L w_{\mu_{s_{t+1}}}(l) \times e^{(1-\gamma)\Delta c(j,k,l)} \times \dots (PC(s_{t+1}, \Delta c(j,k,l), X_t) + 1)^\theta | s_{t+1}, s_t, X_t \right] \quad (128)$$

given the current priors and state, as well as conditional on next period's state. Here, $w_\varepsilon(j)$ is the quadrature weight for a standard normal variable corresponding to the quadrature point $n_\varepsilon(j)$, $w_{\sigma_s^2}(l)$ and $w_{\mu_s}(k)$ are the quadrature weights for the truncated inverse gamma distributed variance and the truncated normally distributed mean corresponding to the quadrature points $n_{\sigma_s^2}(k)$ and $n_{\mu_s}(l)$. Thus, realized consumption growth, which is what the agent observes when updating beliefs about $\mu_{s_{t+1}}$ and $\sigma_{s_{t+1}}$ is:

$$\Delta c(j, k, l) = n_{\mu_{s_{t+1}}}(l) + \sqrt{n_{\sigma_s^2}(k)n_\varepsilon(j)}. \quad (129)$$

5.3 Existence of equilibrium

The existence proof relies on the concavity of the value function and the fact that the value function is finite in all the boundary, known parameter cases. Thus, a sufficient existence condition

amounts to checking existence for all values in the parameter grid in the known parameters boundary economies. A similar approach can be used to prove existence for all the cases considered in this paper.

6 Additional results for the switching regime model

Here, we first show the effects of uncertain mean and/or variance parameters in the switching regime model. Thereafter, we perform some sensitivity analysis in terms of the effect of the rate of time preference parameter (β) on the asset pricing moments.

6.1 Uncertain mean and variance of the Depression state

Here, we consider the case where the transition probabilities are known, but where instead the mean and variance parameters of the Depression state (μ_2 and σ_2) are unknown.¹³

We assume the same setup as given in the numerical solution section above. In particular, the joint prior over the mean and variance in the state is normal-inverse-gamma:

$$p(\mu_2, \sigma_2^2 | \Delta c^\tau) = p(\mu_2 | \sigma_2^2, \Delta c^\tau) p(\sigma_2^2 | \Delta c^\tau), \quad (130)$$

where

$$p(\sigma_2^2 | \Delta c^\tau) \sim IG\left(\frac{b_\tau}{2}, \frac{B_\tau}{2}\right), \quad (131)$$

$$p(\mu_2 | \sigma_2^2, \Delta c^\tau) \sim N(\mu_\tau, A_\tau \sigma_2^2), \quad (132)$$

and where τ counts time spent in state 2 and where Δc^τ denotes the history of consumption growth realizations in the Depression state. Obviously, there is no learning about μ_2 and σ_2^2 from observing consumption growth realizations in the good state. The updating equations are then the natural special case of the updating equations given in the previous section (Equations (121)-(124)).

In terms of pricing, note that this system can be reduced to three state-variables: μ_τ , B_τ , and τ , given initial priors. We solve the model numerically and, as before, use the closed-form solution for the known parameters cases as the boundary values in a recursion that is solved backwards in

¹³In a related paper, Lu and Siemer (2011) consider an economy where agents use an adaptive learning rule to learn about whether there is a disaster or not, as well as the mean growth rate in the disaster state. This mean growth rate is drawn at the beginning of each disaster and so it is not a fixed parameter as in the case we consider here.

time, where time again is counted in terms of time spent in the Depression state, τ , on a grid for μ_τ and B_τ .

To ensure existence of equilibrium, it is necessary to truncate the distribution for the unknown parameters. For example, the normal distribution for the mean in the disaster state implies that there is a positive probability that the disaster state has, in fact, an arbitrarily high mean growth rate. As is well known, the growth rate of the economy has, in conjunction with preference parameters, to satisfy a transversality condition, so an unbounded support for μ_2 is inadmissible. Further, as pointed out by Geweke (2002), Weitzman (2007), and Bakshi and Skoulakis (2010), it is also necessary to truncate the support for σ^2 . The updating equations for the state variables are not affected by the truncation, although of course the numerical integration will take the truncation into account.

As discussed earlier, our focus in this paper is on unbiased priors and we choose wide truncation limits relative to the initial prior to limit the effects of truncation on our results.¹⁴ In particular, we let the truncation limits on the mean growth rate be $+/- 4$ standard deviations away from the true mean, where the standard deviation in question is that which arises after an initial prior learning period of 100 years. Thus, $\underline{\mu}_2 = \mu_2 - 4 \times \sigma_2 \times \sqrt{A_0}$, $\bar{\mu}_2 = \mu_2 + 4 \times \sigma_2 \times \sqrt{A_0}$, where $A_0 = 16$ (given the assumption of one previously observed Depression in the 100 year training period) and $a_{t=0} = \mu_2$. This implies that the prior standard deviation of beliefs about the Depression mean is 0.3675%. Similarly, we set $\bar{\sigma}^2 = 9 * \sigma^2$, $\underline{\sigma}^2 = 1e - 6$. We set $b_0 = 16$, reflecting the 100 year prior learning prior, and $\frac{B_0}{b_0-2} = \sigma^2$ so the prior is unbiased.¹⁵ Simulating data from the truncated Inverse-Gamma distribution, we find that this implies that the prior standard deviation of beliefs about the variance σ_2^2 is $(0.91\%)^2$, while the standard deviation of beliefs about the standard deviation σ_2 is 0.27%.

Table 2 shows average 100-year sample asset pricing moments from this economy for $\gamma = \{3.9, 5\}$, $\beta = 0.994$, and $\psi = 2$, as well as for the 100-, 200-, and 300-year prior training sample periods. The remaining parameter values $\mu_2 = -1.15\%$, $\sigma_2 = 1.47\%$, $\pi_{11} = 0.9975$, $\pi_{22} = 0.9375$, $\mu_1 = 0.54\%$ and $\sigma_1 = 0.98\%$. We price a dividend claim to compare to market returns and assume $\Delta d_{t+1} = \bar{\mu} + \lambda(\Delta c_{t+1} - \bar{\mu}) + \sigma_d \eta_{t+1}$, where d_t is the log of dividends, $\bar{\mu}$ is the unconditional mean consumption growth rate, λ is a leverage parameter, and η_t is an i.i.d. standard normal shock (independent of ε_t). We set the leverage parameter to 2.5. We set the idiosyncratic volatility σ_d such that annual dividend volatility is 11.5%, as in Bansal and Yaron (2004).

¹⁴At the same time, we do not choose truncation limits that are very close to violating transversality conditions. Bakshi and Skoulakis (2010) show that the results in Weitzman (2007) are due exactly to this, i.e. that the quantitatively significant results come from the tiny probability of extreme levels of consumption volatility.

¹⁵Through simulation we verify that the truncation bounds are such that $E[\sigma_2^2 | b_0, B_0] = \frac{B_0}{b_0-1}$ is very close to a correct expression also for the *truncated* Inverse-Gamma distribution.

The calibration is slightly different from that in the main paper, but this does not affect the conclusions we draw in this section. The table also gives the asset pricing implications of assuming that only the mean or only the variance are unknown. This is achieved by setting $A_0 = 0$ and $b_0 = \infty$, respectively.

Panel *A* of Table 2 shows the case where only the disaster mean μ_2 is unknown. When $\gamma = 3.9$ the average sample equity premium is 1.80% versus 1.05% in the known parameters benchmark case. This compares to 5.67% for the case of unknown transition probabilities. Thus, while an uncertain disaster mean adds risk to this calibration, the risk amplification is much less than in the case of unknown persistence parameters. The Sharpe ratio is 0.19, while the risk-free rate is somewhat high at 2.75%, again giving a much poorer fit to the data than the case of unknown transition probabilities.

Figure 5 shows as an example a simulated path of the annual wealth-consumption from this economy, which shows how updates in beliefs only happen in the Depression state. In this case, the updating leads to time-varying beliefs about μ_2 (see Equation (132)), which is reflected in time-variation in the wealth-consumption ratio.¹⁶ Thus, the learning about the mean economy generates somewhat more interesting dynamics in valuation ratios during a disaster than the case when learning about the transition probabilities.

In Barro, Nakamura, Steinsson, and Ursua (2011), the standard error about the annual disaster mean is reported to be 0.7%, which corresponds to about 0.18% for a quarterly mean. This level of uncertainty corresponds roughly to a 200-year prior learning period before the last 100-years of data these authors base their estimates on. In the 200-year prior case, the equity premium is 1.45% and the Sharpe ratio is 0.17, while for the 300-year prior the risk premium is 1.33% and the Sharpe ratio is 0.16. Thus, while the effects of parameter uncertainty are decreasing over time, the decrease is very slow simply because one can only learn about a parameter that governs dynamics in a rare event when the event occurs.

The rightmost set of columns shows the case when the risk aversion coefficient is increased from 3.9 to 5. Now, the 100-year prior yields a risk premium of 4.95% for the case of unknown mean with a 100-year prior, decreasing to 3.57% for the 300-year prior, which is still substantially higher than the 2.74% in the known mean benchmark case. Note that the modest increase in risk aversion in this experiment has a large impact on the asset pricing moments. This nonlinearity is due to

¹⁶We did not feed actual regimes and consumption shocks into this model as quarterly consumption data is not available for the pre-WW2 period (including, of course, the Great Depression).

Figure 5 - Mean beliefs about disaster mean and the wealth-consumption ratio

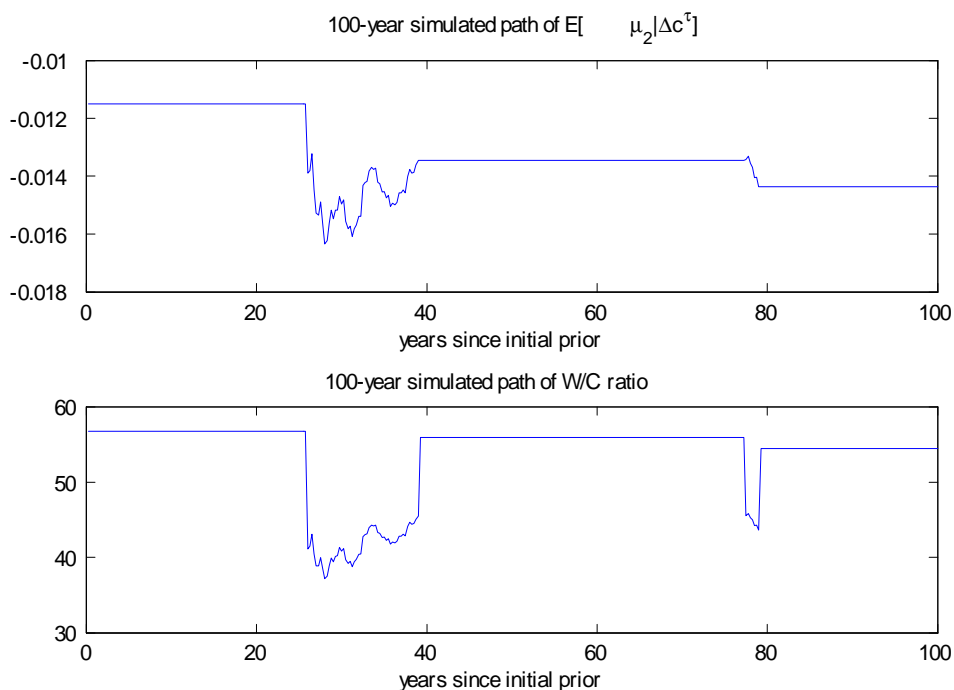


Figure 5: The figure shows a representative simulated 100-year sample path from the 2-state regime switching model, where the mean consumption growth rate in the bad state is unknown. Top plot shows the mean belief about this parameter, while the lower plot shows the annualized wealth-consumption ratio (P/C), given a 100-year training period for the prior.

the negative skewness and ensuing non-normality of the consumption dynamics the Depression calibration implies (see, e.g., Rietz (1988)). To summarize Panel A of Table 2, unknown Depression mean can almost double the risk premium and increase the Sharpe ratio by up to a factor of 1.5 relative to the known parameter case. This is in contrast to the previous case with unknown transition probabilities which increased the risk premium by up to a factor of 5 and the Sharpe ratio by almost a factor of 3.

Figure 5 shows a 100-year path of the mean belief about the Depression mean, a_t , as well as the wealth-consumption ratio, when the model is fed regimes from the last century of U.S. macro data, as discussed earlier for the case of unknown transition probabilities. The calibration has $\gamma = 3.9$ and a 100-year prior and uncertainty over only the mean parameter. The wealth-consumption ratio falls at the onset of the Great Depression, as before, but now the updating about the mean growth rate in the Depression state leads to more dynamics in the wealth-consumption ratio while in the Depression state. Once the normal state re-emerges, the wealth-consumption ratio is constant as

there is nothing to learn outside of the Depression event for the uncertain parameters considered here.

Panel *B* of Table 2 shows the case where only the variance in the bad state is unknown. This case is quickly summarized. For either risk aversion assumptions, the asset pricing implications arising from the 100-, 200-, and 300-year priors are in all cases the same as those for the known parameters benchmark model. As is well-known, second moments are much more precisely estimated than first moments, and this is reflected in the speed of learning about the variance parameter. Note that this result is robust to small changes in the truncation bounds for σ_2^2 .

Panel *C* of Table 2 shows the case where both the mean variance in the bad state are unknown. In this case, the risk premium and Sharpe ratio are slightly higher than for the case where only the mean was unknown. For instance, for $\gamma = 3.9$, the risk premium increases from 1.80% to 1.83%, while for the $\gamma = 5$ case, the risk premium increases from 4.95% to 5.13%. Thus, the unknown variance does interact with the unknown mean, in particular to create a little higher return volatility, but the effect is not very large. In sum, in the case considered here where the priors are unbiased, uncertainty about a truncated volatility parameter does not have large asset pricing impact given a reasonably calibrated prior. Of course, with biased priors, a drifting variance parameter can affect both ex ante and ex post risk premiums considerably (see Johannes, Lochstoer, and Mou (2010)). The exercise in this paper, however, is to establish the properties of the risk pricing of parameter uncertainty.

6.2 Sensitivity Analysis

In this section, we perform some sensitivity analysis in terms of the effect of the rate of time preference parameter (β) on the asset pricing moments. In particular, we consider the case of unknown transition probabilities, as in the main text, and solve this model for various values of β . As is clear from Equation (7) in the main text, the effective duration of revisions in beliefs in terms of their utility impact is strongly positively related to β .

Table 3 shows the sample asset price moments for various calibrations of the 2-state Depression model with a 100-year prior (see main paper for further explanation of the model). The columns 2 and 3 show the results with the original calibration for known and unknown parameters, i.e., when $\beta = 0.994$ (0.98 annually), $\gamma = 4$, and $\psi = 2$. Columns 4 and 5 show how the moments change when we reduce β to 0.99 (0.96 annually). The risk premium and price of risk are reduced from 5.6% and 1.12 to 3.5% and 0.71, respectively. Thus, since the effective duration of the belief shock in terms of its impact on utility is now much lower, the amount of risk is lower. Note that this

also affects the economy in the known parameters case, as the Depression state is persistent and therefore embodies some objective long-run risk. Here the risk premium decreases from 0.88% to 0.56%. Importantly, the lower value of β leads to a risk-free rate puzzle. With unknown transition probabilities, the average level of the risk-free rate increases from 0.8% (as in the data) to 3.5%. Columns 6 and 7 show the case where β again is low (0.99), but where the risk aversion is increased to 5. This allows the learning model to once again match (slightly overshoot) the risk premium, which is now 5.5% versus 1.8% in the known parameters case.

In sum, the value of β is important for the pricing of parameter belief updates, as clearly shown in Equation (7) in the main paper. The value we use in our main calibrations is the same as that used in Bansal and Yaron (2004), which in fact is conservative relative to the more recent long-run risk literature (e.g., Bansal, Kiku, and Yaron (2012)). The savings literature, e.g., Hubbard, Skinner and Zeldes (1995) use an annual β of 0.97 (0.9924 quarterly). Even with the more conservative quarterly β of 0.99, the main results of our paper carry through. Parameter learning provides a substantial amplification mechanism for the pricing of macro shocks—leading to a 3 times increase in the risk premium relative to the known parameters case even with a very conservative value of β .

7 Numerical solution method for model uncertainty case

In this case, consumption growth can be written:

$$\Delta c_{t+1} = (1 - M) \{\mu + \sigma \varepsilon_{t+1}\} + M \{\mu + x_t + \sigma \varepsilon_{t+1}\}, \quad (133)$$

where $x_{t+1} = \rho x_t + \varphi \sigma \varepsilon_{t+1}$ and $\varepsilon_{t+1} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$. M is a parameter that takes the value 0 if the i.i.d. model is the true model and 1 if the LRR model is true.

The state variables are x_t , the conditional level of long-run risk given the Bansal-Yaron model, and the probability that the Bansal-Yaron model is the true model, relative to the iid consumption growth model, p_t .

The boundary cases, $p_t = 0$ and $p_t = 1$, are solved as follows. For iid consumption growth, the price-consumption ratio, PC_t , is found using the fact that this ratio is constant in this case. Thus:

$$\left(\frac{PC(p_t = 0)}{1 + PC(p_t = 0)} \right)^\theta = E_t [\beta^\theta e^{(1-\gamma)\Delta c_{t+1}}] = \beta^\theta e^{(1-\gamma)\mu + \frac{1}{2}(1-\gamma)^2\sigma^2}. \quad (134)$$

For the case where $p_t = 1$, we have that:

$$PC(p_t = 1, x_t)^\theta = E_t \left[\beta^\theta e^{(1-\gamma)\Delta c_{t+1}} (PC(p_t = 1, x_{t+1}) + 1)^\theta \right], \quad (135)$$

where

$$\Delta c_{t+1} = \mu + x_t + \sigma \varepsilon_{t+1}, \quad (136)$$

$$x_{t+1} = \rho x_t + \varphi \sigma \varepsilon_{t+1}, \quad (137)$$

where the parameters are defined in the main text. The price-consumption ratio is found on a grid for x by iterating on Equation (135) given an initial guess of $PC(p_t = 1, x_t)$.

Given these boundary solutions, the price-consumption ratio for the general cases where $0 < p_t < 1$ are found by iterating on the equation:

$$PC(p_t, x_t)^\theta = E_t \left[\beta^\theta e^{(1-\gamma)\Delta c_{t+1}} (PC(p_{t+1}, x_{t+1}) + 1)^\theta \right], \quad (138)$$

on a grid for x and p , given an initial guess of $PC(p_t, x_t)$, with the state variable evolution given by

$$\Delta c_{t+1} = \mu + p_t x_t + \sigma \varepsilon_{t+1}, \quad (139)$$

$$x_{t+1} = \rho x_t + \varphi x \varepsilon_{t+1}, \quad (140)$$

$$p_{t+1} = \frac{p_{BY}(y_{t+1}|y^t) p_t}{p_{BY}(y_{t+1}|y^t) p_t + p_{iid}(y_{t+1}) (1 - p_t)}, \quad (141)$$

We note that it is important to have an incredibly fine grid as p approaches zero in order to ensure a nicely behaved PC ratio at this boundary. We use 480 grid points for p with a strongly nonlinear grid for the reason just mentioned. The exact grid used can be obtained upon request.

Once the solution for the price-consumption ratio is found, we solve for the price-dividend ratio of the claim to the exogenous dividend stream:

$$\Delta d_t = \mu + \lambda (\Delta c_t - \mu) - \frac{1}{2} \sigma_d^2 + \sigma_d \varepsilon_{d,t}, \quad (142)$$

where $\varepsilon_{d,t}$ is iid standard normal and uncorrelated with the shock to consumption. In particular,

we solve in a similar manner as that just described for the consumption claim, the recursion:

$$PD(p_t, x_t) = E_t \left[\beta^\theta e^{-\gamma \Delta c_{t+1} + \Delta d_{t+1}} \left((PC(p_{t+1}, x_{t+1}) + 1) / PC(p_t, x_t) \right)^{\theta-1} (1 - PD(p_{t+1}, x_{t+1})) \right]. \quad (143)$$

We set, as the initial guess, the price-dividend ratio equal to the price-consumption ratio.

References

- [1] H. Ai. Information about long-run risk: Asset pricing implications. *Journal of Finance*, 65:1333–1367, 2010.
- [2] G. Bakshi and G. Skoulakis. Do subjective expectations explain asset pricing puzzles? *Journal of Financial Economics*, december:117–140, 2010.
- [3] R. Bansal, D. Kiku, and A. Yaron. Risks for the long run: Estimation with time aggregation. *Wharton working paper*, 2012.
- [4] R. Bansal and I. Shaliastovich. Confidence risks and asset prices. *American Economic Review, papers & proceedings*, 100(2):537–541, 2010.
- [5] R. Bansal and A. Yaron. Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance*, 59:1481–1509, 2004.
- [6] R. Barro. Rare disasters, asset prices and welfare costs. *American Economic Review*, 99:243–264, 2009.
- [7] R. Barro, E. Nakamura, J. Steinsson, and J. Ursua. Crises and recoveries in an empirical model of consumption disasters. *American Economic Journal: Macroeconomics*, 5:35–74, 2013.
- [8] L. Benzoni, P. Collin-Dufresne, and R. S. Goldstein. Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics*, 101:552–573, 2011.
- [9] L. Benzoni, P. Collin-Dufresne, R. S. Goldstein, and J. Helwege. Modeling credit contagion via the updating of fragile beliefs. *forthcoming The Review of Financial Studies*, 2015.
- [10] O. Boguth and L. Kuehn. Consumption volatility risk. *Journal of Finance*, 2012.

- [11] M. Brandt, Q. Zeng, and L. Zhang. Equilibrium stock return dynamics under alternative rules of learning about hidden states. *Journal of Economic Dynamics and Control*, 28:1925–1954, 2004.
- [12] J. Campbell and R. Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1:195–228, 1988.
- [13] H. Chen, W. Dou, and L. Kogan. Measuring the 'dark matter' in asset pricing models. MIT working paper, September 2013.
- [14] T. Cogley and T. J. Sargent. Anticipated utility and rational expectations as approximations of bayesian decision making. *International Economic Review*, 49(1):185–221, February 2008.
- [15] T. Cogley and T. J. Sargent. The market price of risk and the equity premium: A legacy of the great depression? *Journal of Monetary Economics*, 55(3):454–476, April 2008.
- [16] M. Croce, M. Lettau, and S. Ludvigson. Investor information, long-run risk and the term structure of equity. *Review of Financial Studies*, 2014.
- [17] A. David. Fluctuating confidence in stock markets:implicatins for returns and volatility. *Journal of Financial and Quantitative Analysis*, VOL32:427–462, 1997.
- [18] A. David and P. Veronesi. What ties return volatilities price valuations and fundamentals. *Journal of Political Economy*, 121:682 – 746, 2013.
- [19] J. Detemple. Asset pricing in a production economy with incomplete information. *Journal of Finance*, VOL 41:383–393, 1986.
- [20] M. U. Dothan and D. Feldman. Equilibrium interest rates and multiperiod bonds in a partially observable economy. *Journal of Finance*, 41:369–382, 1986.
- [21] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton: University Press, 2001.
- [22] D. Duffie and L. G. Epstein. Stochastic differential utility. *Econometrica*, 60:353–394, 1992a.
- [23] D. Duffie and L. G. Epstein. Asset pricing with stochastic differential utility. *The Review of Financial Studies*, 5:411–436, 1992b.
- [24] D. Duffie and C. Skiadas. Continuous-time security pricing: A utility gradient approach. *Journal of Mathematical Economics*, 23:107–132, 1994.

- [25] L. G. Epstein and M. Schneider. Learning under ambiguity. *The Review of Economic Studies*, 74(4):pp. 1275–1303, 2007.
- [26] L. G. Epstein and S. E. Zin. Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57:937–969, 1989.
- [27] G. Genotte. Optimal portfolio choice under incomplete information. *Journal of Finance*, 61:733–749, 1986.
- [28] J. Geweke. A note on some limitations of crra utility. *Economics Letters*, 71:341–345, 2001.
- [29] A. Goyal and I. Welch. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508, 2008.
- [30] E. Gvozdeva and P. Kumar. Asset pricing with unknown consumption and financial risk. *working paper, University of Houston*, 2012.
- [31] L. Hansen. Beliefs, doubts and learning: Valuing macroeconomic risk, richard t. ely lecture. *The American Economic Review*, 97(2):1–30, 2007.
- [32] L. Hansen and R. Jaganathan. Implications of security market data for models of dynamic economies. *Journal of Political Economy*, 99:225–262, 1991.
- [33] L. Hansen and T. Sargent. Fragile beliefs and the price of uncertainty. *Quantitative Economics*, 1:129–162, 2010.
- [34] R. G. Hubbard, J. Skinner, and S. P. Zeldes. Precautionary saving and social insurance. *Journal of Political Economy*, 103(2):pp. 360–399, 1995.
- [35] A. Jobert, A. Platania, and L. Rogers. A bayesian solution to the equity premium puzzle. University of Cambridge working paper, March 2006.
- [36] M. Johannes, L. A. Lochstoer, and Y. Mou. Learning about consumption dynamics. *forthcoming at the Journal of Finance*, 2014.
- [37] T. Johnson. Optimal learning and new technology bubbles. *Journal of Monetary Economics*, 87:2486–2511, 2007.
- [38] N. Ju and J. Miao. Ambiguity, learning and asset returns. *Econometrica*, 80:559–591, 2012.

- [39] P. Klibanoff, M. Marinacci, and S. Mukerji. Recursive smooth ambiguity preferences. *Journal of Economic Theory*, 144:930–976, 2009.
- [40] S. L. Lettau, M. and J. Wachter. The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies*, 21(4):1653–1687, 2008.
- [41] J. Lewellen and J. Shanken. Learning, asset-pricing tests, and market efficiency. *Journal of Finance*, 57:1113–1145, 2002.
- [42] Y. Lu and M. Siemer. Asset pricing with learning about disaster risk. *Boston University Working paper*, 2011.
- [43] R. Lucas. Models of business cycles. *New York: Blackwell*, 1987.
- [44] R. Lucas and T. Sargent. After keynesian macroeconomics. in *After the Phillips curve: Persistence of high inflation and high unemployment*, Federal Reserve Bank of Boston, Boston, MA, 1978.
- [45] B. Moore and H. Schaller. Learning, regime switches and equilibrium asset pricing dynamics. *Journal of Economic Dynamics and Control*, 20:979–1006, 1996.
- [46] L. Pastor and P. Veronesi. Stock valuation and learning about profitability. *Journal of Finance*, 58:1749–1789, 2003.
- [47] L. Pastor and P. Veronesi. Was there a nasdaq bubble in the late 1990s? *Journal of Financial Economics*, 81:61–100, 2006.
- [48] L. Pastor and P. Veronesi. Learning in financial markets. *Annual Review of Financial Economics*, 2009.
- [49] L. Pastor and P. Veronesi. Uncertainty about government policy and stock prices. *Journal of Finance*, 64:1219–1264, 2012.
- [50] T. Rietz. The equity risk premium: A solution? *Journal of Monetary Economics*, 22:117–131, 1988.
- [51] I. Shaliastovich. Learning, confidence and option prices. *forthcoming, Journal of Econometrics*, 2015.

- [52] R. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436, 1981.
- [53] C. Skiadas. Robust control and recursive utility. *Finance and Stochastics*, 7:475–489, 2003.
- [54] T. Strzalecki. Temporal resolution of uncertainty and recursive models of ambiguity aversion. *Econometrica*, 81(3):1039–1074, 2013.
- [55] A. Timmermann. How learning in financial markets generates excess volatility and predictability in stock prices. *Quarterly Journal of Economics*, 108:1135–1145, 1993.
- [56] F. Vasquez-Grande. Effects of learning the long-run asset pricing model. *Chicago Booth School of Business working paper*, 2009.
- [57] P. Veronesi. How does information quality affect stock returns. *Journal of Finance*, 55:807–837, 2000.
- [58] M. Weitzman. Subjective expectations and asset-return puzzles. *American Economic Review*, 97(4):1102–1130, 2007.

**Table 2 - 100 year sample moments
Learning about the mean and variance in a Great Depression**

Table 2: This table gives average sample moments from 20,000 simulations of 400 quarters of data from the 2-state switching regime model of consumption growth, where the mean and/or variance parameters μ_2 and σ_2^2 are unknown. $E_T[x]$ denotes the average sample mean of x , $SR_T[x]$ denotes the average sample Sharpe ratio of x , and $\sigma_T[x]$ denotes the average sample standard deviation of x . R_m and R_f denote the gross market return and real risk-free rate. Lower case letters denote log of upper case variable. All statistics are annualized and, except for the Sharpe ratio, given in percent. In all cases, the time-preference parameter β is set to 0.994. The relative risk aversion and intertemporal elasticity of substitution are given for each panel.

Panel A: Learning about the disaster mean									
$\psi = 2, \gamma = 3.9$	$T_0 = 100\text{yrs}$	$T_0 = 200\text{yrs}$	$T_0 = 300\text{yrs}$	$T_0 = \infty$	$\psi = 2, \gamma = 5$	$T_0 = 100\text{yrs}$	$T_0 = 200\text{yrs}$	$T_0 = 300\text{yrs}$	$T_0 = \infty$
$E_T[r_m - r_f]$	1.80	1.45	1.33	1.05	$E_T[r_m - r_f]$	4.95	3.92	3.57	2.74
$\sigma_T[r_m - r_f]$	15.84	15.50	15.36	14.94	$\sigma_T[r_m - r_f]$	15.98	15.78	15.69	15.41
$SR_T[R_M - R_f]$	0.19	0.17	0.16	0.14	$SR_T[R_M - R_f]$	0.37	0.31	0.29	0.24
$E_T[r_f]$	2.75	2.89	2.92	2.98	$E_T[r_f]$	1.61	2.15	2.31	2.60
Panel B: Learning about the disaster variance									
$\psi = 2, \gamma = 3.9$	$T_0 = 100\text{yrs}$	$T_0 = 200\text{yrs}$	$T_0 = 300\text{yrs}$	$T_0 = \infty$	$\psi = 2, \gamma = 5$	$T_0 = 100\text{yrs}$	$T_0 = 200\text{yrs}$	$T_0 = 300\text{yrs}$	$T_0 = \infty$
$E_T[r_m - r_f]$	1.05	1.05	1.05	1.05	$E_T[r_m - r_f]$	2.74	2.74	2.74	2.74
$\sigma_T[r_m - r_f]$	14.94	14.94	14.94	14.94	$\sigma_T[r_m - r_f]$	15.41	15.41	15.41	15.41
$SR_T[R_M - R_f]$	0.14	0.14	0.14	0.14	$SR_T[R_M - R_f]$	0.24	0.24	0.24	0.24
$E_T[r_f]$	2.98	2.98	2.98	2.98	$E_T[r_f]$	2.60	2.60	2.60	2.60
Panel C: Learning about the disaster mean and variance									
$\psi = 2, \gamma = 3.9$	$T_0 = 100\text{yrs}$	$T_0 = 200\text{yrs}$	$T_0 = 300\text{yrs}$	$T_0 = \infty$	$\psi = 2, \gamma = 5$	$T_0 = 100\text{yrs}$	$T_0 = 200\text{yrs}$	$T_0 = 300\text{yrs}$	$T_0 = \infty$
$E_T[r_m - r_f]$	1.83	1.45	1.33	1.05	$E_T[r_m - r_f]$	5.13	3.94	3.57	2.74
$\sigma_T[r_m - r_f]$	15.83	15.48	15.35	14.94	$\sigma_T[r_m - r_f]$	15.97	15.77	15.68	15.41
$SR_T[R_M - R_f]$	0.19	0.17	0.16	0.14	$SR_T[R_M - R_f]$	0.38	0.31	0.29	0.24
$E_T[r_f]$	2.68	2.87	2.92	2.98	$E_T[r_f]$	1.35	2.11	2.30	2.60

Table 3 - Sensitivity to value of β
Learning about the probability and persistence of a Great Depression

Table 3: This table gives average sample moments from 20,000 simulations of 400 quarters of data from the 2-state switching regime model of consumption growth, where the transition probabilities are unknown. $E_T[x]$ denotes the average sample mean of x , $SR_T[x]$ denotes the average sample Sharpe ratio of x , and $\sigma_T[x]$ denotes the average sample standard deviation of x . R_m and R_f denote the gross market return and real risk-free rate. Lower case letters denote log of upper case variable. All statistics are annualized and, except for the Sharpe ratio, given in percent. The time-preference parameter are given in the table.

Panel A:	Unknown π 's	Known π 's	Unknown π 's	Known π 's	Unknown π 's	Known π 's
	$\gamma = 4$	$\gamma = 4$	$\gamma = 4$	$\gamma = 4$	$\gamma = 5$	$\gamma = 5$
$\psi = 2, \gamma = 4$	$\beta = 0.994$	$\beta = 0.994$	$\beta = 0.99$	$\beta = 0.99$	$\beta = 0.99$	$\beta = 0.99$
$E_T[r_m - r_f]$ (%)	5.64	0.88	3.46	0.56	5.51	1.75
$\sigma_T[r_m - r_f]$ (%)	18.0	17.3	18.1	16.7	17.2	17.1
$SR_T[R_M - R_f]$	0.37	0.14	0.26	0.12	0.39	0.18
$E_T[r_f]$ (%)	0.80	2.92	3.54	4.61	2.68	4.34
$\sigma_T[M] / E_T[M]$	1.12	0.27	0.71	0.23	1.19	0.41