Title

Parameter Learning in General Equilibrium: The Asset Pricing Implications

Journal

American Economic Review, 106(3)

ISSN

0002-8282

Authors

Collin-Dufresne, Pierre
Johannes, Michael
Lochstoer, Lars A

Publication Date

2016-03-01

DOI

10.1257/aer.20130392

Peer reviewed

American Economic Review 2016, 106(3): 664–698
http://dx.doi.org/10.1257/aer.20130392

Parameter Learning in General Equilibrium: The Asset Pricing Implications[†]

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

Parameter learning strongly amplifies the impact of macroeconomic shocks on marginal utility when the representative agent has a preference for early resolution of uncertainty. This occurs as rational belief updating generates subjective long-run consumption risks. We consider general equilibrium models with unknown parameters governing either long-run economic growth, rare events, or model selection. Overall, parameter learning generates long-lasting, quantitatively significant additional macroeconomic risks that help explain standard asset pricing puzzles. (*JEL* C52, D83, E13, E32, G12)

One of the classic tensions in macro-finance models is the apparent disconnect between macroeconomic risks and the risks reflected in asset prices. For example, Mehra and Prescott (1985) show in a standard rational expectations representative agent model that consumption is too smooth (its risk is too small) to explain the large risk premium and volatility of asset prices (the "equity-premium puzzle").

When calibrating macro-finance models, however, we are confronted with an inconvenient truth: there is significant uncertainty over the structural parameters governing the dynamics of the model and also over the model specification itself (see Hansen 2007). Although commonly assumed away in standard rational expectations models (e.g., Lucas 1978), the reality is that parameters such as the unconditional growth rate of the economy, the likelihood and severity of disaster events, or the persistence of long-run growth shocks are unknown. A long-standing critique of rational expectations models is the assumption that agents know "fixed, but unknown" parameters (e.g., Modigliani 1977).

*Collin-Dufresne: École Polytechnique Fédérale de Lausanne, Quartier UNIL-Dorigny, Extranef 209, CH-1015 Lausanne, Switzerland, Swiss Finance Institute, and CEPR (e-mail: pierre.collin-dufresne@epfl.ch); Johannes: Columbia Business School, 410 Uris Hall, 3022 Broadway, New York, NY 10027 (e-mail: MJ335@gsb.columbia.edu); Lochstoer: Columbia Business School, 424 Uris Hall, 3022 Broadway, New York, NY 10027 (e-mail: LL2609@gsb.columbia.edu). We thank David Backus, Mikhail Chernov, Darrell Duffie, Lars Hansen, Espen Henriksen, Thomas Mertens, Stavros Panageas, Stanley Zin, and seminar participants at Brigham Young University, the Chicago Initiative in Theory and Empirics (CITE conference), Columbia, HKUST, London School of Economics, NBER SI Asset Pricing meeting, Ohio State, SED meetings 2015, Stanford, UC Davis, UCLA, the University of Texas at Houston, the University of Minnesota, the University of York, the Inquire UK meeting at Imperial College, and the WFA for helpful comments. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†]Go to http://dx.doi.org/10.1257/aer.20130392 to visit the article page for additional materials and author disclosure statement(s).

At the same time, parameter or model learning is not inconsistent with rational decision making.[1]

This paper assumes economic agents learn about parameters from data, revising beliefs as new data arrive.[2] We show that the process of updating of beliefs can serve as a major amplification mechanism for the pricing of macroeconomic shocks, with the potential to be the primary driver of the high and time-varying risk prices historically observed (see, e.g., Cochrane and Hansen 1992, Campbell 1999). The fundamental mechanism underlying the importance of parameter uncertainty is driven by two assumptions: (i) Bayesian learning and (ii) recursive preferences (Epstein and Zin 1989—henceforth, EZ). Bayesian learning implies that posterior estimates of parameters are martingales.[3] Thus, revisions in parameter beliefs constitute permanent shocks to investors' expectations and are therefore a source of subjective long-run risks. Agents with recursive utility who have a preference for early resolution of uncertainty are averse to such long-run risks. Consequentially, uncertainty about future changes in parameter beliefs leads to higher risk prices, risk premia, and price volatility than in the known parameter case and, indeed, than in the standard time-separable utility case.

This paper analyzes asset pricing implications of priced parameter uncertainty in three distinct consumption-based exchange economies, each highlighting different aspects of how parameter learning impacts macroeconomic risk prices. The first case considers learning about the mean growth rate with i.i.d. normal log consumption growth. This work-horse model yields analytical solutions to utility and key asset pricing moments for arbitrary risk aversion as long as the elasticity of intertemporal substitution (EIS) is one. This case is useful for intuition and for contrasting how priced parameter uncertainty relates to the current main approach to deal with parameter learning, anticipated utility (AU) (Kreps 1998, Cogley and Sargent 2008), as well as to the long-run risk model of Bansal and Yaron (2004)—henceforth, BY—and to the parameter uncertainty approach in Weitzman (2007).

The analytical expressions show how parameter learning amplifies macroeconomic risk pricing and, specifically, that the amplification is a function of the degree to which the agent is averse to long-run shocks, the conditional volatility of updates in beliefs (related to the prior uncertainty), and the effective duration of the permanent revision in beliefs in terms of its effect on current utility.

AU pricing, which at each point in time ignores parameter uncertainty by calculating utility and asset prices assuming current mean parameter beliefs are the true values, does not price long-run shocks arising from belief updates. In fact, we show that rationally accounting for parameter uncertainty leads to large welfare losses relative to the AU approach, when agents are averse to long-run consumption shocks.

[1] Indeed, Lucas and Sargent (1978, p. 13) note, "…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."

[2] Model uncertainty can be viewed as a form of parameter uncertainty, since any model can be embedded in a more general specification with parameter restrictions. Unless explicitly noted, we will generally refer to both as parameter uncertainty/learning.

[3] This follows from the law of iterated expectations. For a parameter θ, $\mu_t \equiv E_t[\theta] = E_t[\mu_{t+j}], j > 0$.

Further, the mechanism here is entirely distinct from that of Weitzman (2007), who shows that variance uncertainty leads to fat-tailed consumption growth distributions that can have a first-order impact on asset pricing. Our approach does not rely on fat tails or the possibility of extreme parameter values, but rather on the fact that updates in beliefs about the parameters governing consumption dynamics have a large impact on the long-run subjective consumption distribution.

The permanent nature of rational learning introduces long-run consumption risks that seem similar to BY. In both cases, long-run shocks to consumption are priced due to the EZ preferences. However, the long-run shocks arising from parameter learning are *subjective*. Beeler and Campbell (2012) argue that long-run risk models imply excess consumption growth predictability and a counterfactually high EIS in Hall (1988)-type regressions. In the parameter learning case, consumption growth is in truth i.i.d. and therefore neither valuation ratios nor the risk-free rate predict future consumption growth. In sum, parameter learning-generated, subjective long-run risks are distinct from the objective long-run risks introduced in BY.

Second, we analyze a rare events model, in the spirit of Rietz (1988), Barro (2006), and Nakamura et al. (2013), via a two-state Markov switching model for consumption growth. The rare bad state is calibrated to the US Great Depression experience. This model has six parameters governing consumption dynamics and each uncertain parameter introduces at least one new state variable (governing parameter beliefs). To reduce the dimensionality of the problem, we propose a simple welfare-based metric to identify a smaller set of parameters which, when unknown, are the most asset pricing and utility relevant. This metric indicates that uncertainty about the transition probabilities leads to the biggest welfare loss and therefore has the largest asset pricing impact, while learning about state means has a moderate impact and learning about the variances of shocks has a negligible impact. This is an important intermediate result which shows that the impact of priced parameter uncertainty varies substantially across parameters. We therefore solve the full, sequential learning model with unknown transition probabilities.

Learning is naturally slow given the rare nature of the bad state. Even with an effective prior training sample of 300 years, the risk premium increases five-fold relative to a known parameters case, while the price of risk increases three-fold. Given the lack of data on crisis states, there is substantial uncertainty regarding the persistence of the bad state, which, in turn, implies a commensurate increase in risk premia in bad states as agents experience relatively large belief revisions depending on its realized duration. The model features a conditional equity premium of 22 percent and a conditional return volatility of 81 percent in this state, consistent with the historical data and much higher than the same model with known parameters. Finally, slow learning means that there are not counterfactually strong drifts in risk premia and valuation ratios even over the typical 100-year sample.

Third, we consider learning about model specification, a particular form of parameter uncertainty. Specifically, agents are uncertain whether the data are generated from an i.i.d. consumption growth model or from a long-run risk model with a small persistent component in consumption growth as in BY. There is considerable debate about the existence and degree of such risks, since a small persistent consumption growth component is difficult to detect empirically.

Interestingly, we find that even a small probability of the long-run risk model being true leads to a five-fold increase in the risk premium relative to the case where consumption growth is known to be i.i.d. This is due to *model disaster risk* which arises because the long-run risk model is associated with much lower utility level than the i.i.d. model. This negative skewness in utility outcomes is priced with EZ preferences and has similar effects on risk prices as the disaster models of Rietz (1988) and Barro (2006) even though it is not due to actual consumption disasters.

Further, the amount of model disaster risk is endogenously countercyclical, which leads to substantial countercyclical fluctuations in the price of risk and the risk premium. In particular, in recessions, when the small persistent component of consumption in the long-run risk model is low, the utility difference between the i.i.d. model and the long-run risk model is large. This leads to a high conditional price of risk and risk premium: in fact, higher than in either of the two alternative models. Conversely, in expansions, the price of risk and the risk premium are low as the conditionally high expected consumption growth in the BY model leads to a lower utility difference between the two models. Thus, even though each of the alternative models are homoskedastic, the price of risk and risk premium are strongly countercyclical in the case of model uncertainty.

In sum, parameter learning amplifies the impact of macroeconomic shocks on marginal utility when the representative agent has EZ preferences. In empirically relevant cases, this amplification is higher in bad states than in good states, leading to countercyclical risk premiums and Sharpe ratios as observed in the data. Further, learning about consumption disaster risk or model risk is slow, so that the amplification does not disappear quickly. It also does not lead to easily detectable and counterfactual predictability relations between returns and consumption shocks. We conclude that accounting for parameter uncertainty can be a first-order issue when relating macroeconomic risks to asset prices.

Our paper is related to a long list of papers studying the general equilibrium implications of parameter learning. The earlier literature focuses on the case of time-separable utility, which does not price the subjective long-run risks discussed above.[4] Much of the literature on equilibrium learning implications with preferences of the EZ type studies the impact of rational or near-rational learning about stationary but unobservable state variables whose dynamics have known parameters.[5] Johannes, Lochstoer, and Mou (forthcoming) empirically document large drifts in parameter estimates, high levels of parameter uncertainty, and that shocks to beliefs correlate with equity returns, for a Bayesian agent who updates beliefs about US consumption dynamics in real-time. Our paper addresses the issue of rational pricing of the risks generated by such updating of beliefs. The main asset pricing results in Johannes, Lochstoer, and Mou (forthcoming) ignore the pricing of these risks as they utilize AU pricing following Cogley and Sargent (2008).

[4] Previous papers on learning with time-separable utility function include: Detemple (1986); Dothan and Feldman (1986); Gennotte (1986); Veronesi (2000, 2004); Cogley and Sargent (2008); Jobert, Platania, and Rogers (2006); Pástor and Veronesi (2003, 2006, 2009, 2012); David (1997); Moore and Shaller (1996); and David and Veronesi (2013). See the online Appendix for a concise review.

[5] These include Brandt, Zeng, and Zhang (2004); Croce, Lettau, and Ludvigson (2014); Lettau, Ludvigson, and Wachter (2008); Ai (2010); Boguth and Kuehn (2013); Bansal and Shaliastovich (2010, 2011); and Benzoni, Collin-Dufresne, and Goldstein (2011).

Also related is the literature that looks at alternative preferences, which though distinct from recursive utility might display similar quantitative effects to those we describe here, if combined with learning.[6]

I. Intuition and a Benchmark Example

A. *Intuition for Priced Parameter Uncertainty*

This section shows that parameter uncertainty and rational updating generates subjective long-run consumption risks with strong asset pricing implications. To start, consider rational uncertainty over a vector of parameters θ affecting consumption growth dynamics and denote the time-t posterior density as $p(\theta \mid y^t)$, where y^t is the information set of the agent.[7] The posterior distribution summarizes all relevant information known at time t about the parameters. In many cases, the posterior distribution is a known function of a set of sufficient statistics of the observables (e.g., the posterior mean and variance of θ).

Importantly, because of the law of iterated expectation, the conditional expectation of any function of the unknown parameters is a martingale (i.e., unit root process) in the filtration of the representative agent (see Doob 1949). Thus, updates in parameter beliefs lead to long-lived, indeed *permanent*, shocks to such conditional expectations.[8] This is fundamentally why parameter learning endogenously generates long-run consumption risks.

To understand the impact on asset prices, consider a representative agent with EZ utility, V, over consumption, C. To focus on growth dynamics, it is useful to normalize by consumption levels: $VC_t \equiv V_t / C_t$, $vc_t \equiv \ln(VC_t)$, and $\Delta c_t = \ln(C_t / C_{t-1})$. The normalized value function is

$$(1)\quad vc_t = \begin{cases} \frac{1}{1-1/\psi} \ln\Big(1-\beta+\beta E_t\big\{\exp\big[(1-\gamma)(vc_{t+1}+\Delta c_{t+1})\big]\big\}^{(1-1/\psi)/(1-\gamma)}\Big), & \text{if } \psi \neq 1, \gamma \neq 1 \\ \frac{\beta}{1-\gamma} \ln E_t\big\{\exp\big[(1-\gamma)(vc_{t+1}+\Delta c_{t+1})\big]\big\}, & \text{if } \psi = 1, \gamma \neq 1 \end{cases},$$

where γ, ψ, and β capture relative risk aversion, the EIS, and time discounting, respectively. The stochastic discount factor (SDF) is

$$(2)\qquad M_{t+1} = \beta \delta_t \exp\big[-\gamma \Delta c_{t+1} - (\gamma - 1/\psi) vc_{t+1}\big],$$

where $\delta_t = E_t\Big[\exp\big\{(1-\gamma)(vc_{t+1} + \Delta c_{t+1})\big\}\Big]^{(\gamma-1/\psi)/(1-\gamma)}$ when $\gamma \neq 1$.

[6] Examples of such alternatives include general Kreps and Porteus (1978) 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). Strzalecki (2013) discusses the relation between ambiguity attitudes and the preference for the timing of the resolution of uncertainty. Epstein and Schneider (2007) discuss the differences between learning under ambiguity and Bayesian learning.

[7] Here, "rational" uncertainty implies that the dynamics of the posterior distributions of beliefs follow the standard laws of probability, such as Bayes' rule.

[8] For an example, consider the case of a posterior mean, $\mu_t \equiv E[\theta \mid y^t]$. For any $j \geq 0$,

$$E\big[\mu_{t+j} \mid y^t\big] = E\big[E\big[\theta \mid y^{t+j}\big] \mid y^t\big] = E[\theta \mid y^t] = \mu_t,$$

which implies that μ_t can be expressed as $\mu_{t+1} = \mu_t + \eta_{t+1}$, where $E[\eta_{t+1} \mid y^t] = 0$.

Parameter uncertainty and the associated state variables generally affect both expectations of future consumption growth and normalized utility. When the relative risk aversion is equal to the reciprocal of the EIS, $\gamma = \psi^{-1}$, the SDF reduces to that of a constant relative risk aversion utility agent (i.e., $M_{t+1} = \beta \exp(-\gamma \Delta c_{t+1})$). That is, only the one-period ahead posterior conditional distribution of Δc_{t+1} matters for pricing. Instead, when relative risk aversion differs from the reciprocal of the EIS, $\gamma \neq \psi^{-1}$, the SDF has an additional term, $\delta_t \exp\{-(\gamma - 1/\psi) v c_{t+1}\}$, through which parameter belief updates impact the normalized value function/continuation utility to become priced risk factors. This impact can be large, even if the change in beliefs is small on a per period basis, as changes in beliefs permanently affect the subjective consumption dynamics.

B. *Learning and Asset Pricing with i.i.d. Consumption Growth*

To better understand the mechanism and analytically quantify the impact of parameter uncertainty, consider an i.i.d. normal consumption growth model:

$$y_{t+1} = \Delta c_{t+1} = \mu + \sigma \varepsilon_{t+1}, \tag{3}$$

where $\varepsilon_{t+1} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. For now, assume agents are uncertain about expected consumption growth μ, but that σ is known.

The agent starts with the conjugate prior, $\mu \sim \mathcal{N}(\mu_0, A_0 \sigma^2)$, and updates beliefs from realized consumption growth using Bayes' rule. Given this prior, the posterior is $p(\mu \mid y^t) = \mathcal{N}(\mu_t, A_t \sigma^2)$. From the agent's perspective, consumption dynamics evolve via (see the online Appendix for detailed derivations)

$$\Delta c_{t+1} = \mu_t + \sigma \sqrt{1 + A_t}\, \tilde{\varepsilon}_{t+1}, \tag{4}$$

where $\tilde{\varepsilon}_{t+1} = (\Delta c_{t+1} - \mu_t)/\sqrt{\sigma^2(1 + A_t)} \sim \mathcal{N}(0,1)$ under the agent's filtration, and

$$\mu_{t+1} = \mu_t + \frac{\sigma A_t}{\sqrt{1 + A_t}}\, \tilde{\varepsilon}_{t+1} \quad \text{and} \quad A_{t+1}^{-1} = A_t^{-1} + 1. \tag{5}$$

Note that μ_t and A_t are now state variables in the economy.[9]

The data are generated from a fixed parameters i.i.d. model, but the agent perceives consumption growth to have time-varying mean and variance. The conditional mean of Δc_{t+1} has a unit root, which implies shocks permanently shift expected consumption growth. Eventually, expected consumption growth settles down however, as its posterior variance, $A_t \sigma^2$, declines deterministically to zero (learning converges).

[9] Defining $\omega_t \equiv (A_t^{-1} + 1)^{-1}$, beliefs have the familiar shrinkage form, $\mu_{t+1} = \omega_t \Delta c_{t+1} + (1 - \omega_t)\mu_t$, combining observed data with prior beliefs according to their precisions.

When $\psi = 1$, many asset pricing relevant quantities can be computed analytically (see the online Appendix). The scaled log-value function is proportional to μ_t, $v c_t = a_t + b\,\mu_t$, where

$$a_t = \sum_{j=0}^{\infty} \beta^{j+1}\,\frac{1-\gamma}{2}\left(\frac{A_{t+j}}{1-\beta} + 1\right)^2 \frac{\sigma^2}{1+A_{t+j}} \tag{6}$$

and the proportionality factor is $b = \beta/(1-\beta)$. Since β is typically close to one, b is very large and utility is highly sensitive to changes in mean beliefs about μ.[10]

The conditional Sharpe ratio (SR) of the log return to the consumption claim is[11]

$$\begin{aligned} SR_t &= \frac{E_t[r_{c,t+1}] - r_{f,t} + \sigma_t^2[r_{c,t+1}]/2}{\sigma_t[r_{c,t+1}]} \\ &= \gamma\sqrt{1+A_t}\,\sigma + (\gamma-1)\frac{\beta}{1-\beta}\,\frac{A_t\sigma}{\sqrt{1+A_t}}. \end{aligned} \tag{7}$$

The first term is the familiar power utility term—risk aversion times the perceived conditional volatility of consumption growth. The second learning-induced term can be decomposed into three components: (i) the preference for the timing of the resolution of uncertainty, $\gamma - 1/\psi = \gamma - 1$; (ii) the utility-impact of the permanent belief shock, $\beta/(1-\beta)$; and (iii) the conditional standard deviation of shocks to beliefs (μ_t), $A_t\sigma/\sqrt{1+A_t}$. Clearly the price of consumption risk is increasing in the amount of parameter uncertainty (A_t) if $\gamma > 1$.[12] But, given that Bayesian learning is efficient, how long do the effects last?

Table 1 shows a dramatic difference between power utility, where premia decrease quickly due to learning, and EZ preferences, where the impact is long lived. The preference parameters are $\beta = 0.994$, $\gamma = 10$, as in BY, and $\psi = 1$. The power utility case has $\psi = \gamma^{-1} = 0.1$. Table 1 reports the annualized conditional Sharpe ratio of the consumption claim (price of risk) as a function of the prior standard deviation of μ, $\sigma_t(\mu)$. The first column reports $A_t^{-1} = t$, intuitively, the number of quarters of data used to update an initially flat prior. Throughout, priors are unbiased and centered at the truth, μ.[13] In all calibrations, we set the annual time-averaged mean and volatility of consumption growth to 1.8 percent and 2.2 percent, respectively, as observed using US real, per capita consumption from NIPA from 1929 to 2011. We solve all models at the quarterly frequency.

[10] A typical parameter value used in the literature for quarterly calibrations is $\beta = 0.994$ which yields $b = 167$.

[11] If W_t the wealth of the representative agent is defined a the present value of the future aggregate consumption stream $\{C_{t+1}, C_{t+2}, \ldots\}$ then the log return to the consumption claim is simply $r_c(t+1) = \log\frac{W_{t+1}}{W_t - C_t}$. Note that the Sharpe ratio in (7) differs slightly from the standard Sharpe ratio definition, since we use log returns. So the numerator does not correspond to the expected return of an actual trading strategy. However, this definition is more convenient analytically in this conditionally Gaussian setup. With this definition SR_t is also equal to the conditional volatility of the log pricing kernel $(\sigma_t(\log M_{t+1}))$ and to the market price of risk of consumption shocks.

[12] Indeed, since $\sqrt{1+A_t} \approx 1$ the price of consumption risk is $\approx \gamma + (\gamma-1)\beta A_t/(1-\beta)$ which is clearly $> \gamma$ when $\gamma > 1$.

[13] See Cechetti, Lam, and Mark (2000) or Cogley and Sargent (2008) for biased beliefs with power utility.

TABLE 1—LEARNING ABOUT THE MEAN CONDITIONAL MOMENTS

Degree of parameter uncertainty		Epstein-Zin ($\gamma = 10$, $\psi = 1$)			Power utility ($\gamma = 10$, $\psi = 1/\gamma$)		
		Annualized, conditional			Annualized, conditional		
$A_t = \frac{1}{t}$ t (quarters)	$\sigma_t(\mu) = \sqrt{A_t}\,\sigma$ percent	Sharpe ratio (SR_t)	Risk-free rate	Yield slope percent	Sharpe ratio (SR_t)	Risk-free rate	Yield slope percent
40	0.21	1.267	0.008	−0.28	0.273	0.167	−0.89
100	0.14	0.672	0.024	−0.13	0.271	0.167	−0.35
200	0.10	0.472	0.030	−0.07	0.271	0.167	−0.18
400	0.07	0.371	0.032	−0.03	0.270	0.168	−0.09
800	0.05	0.321	0.034	−0.02	0.270	0.168	−0.04
∞	0.00	0.270	0.035	0.00	0.270	0.168	0.00

Notes: This table shows various conditional, annualized moments for the "learning about the mean" economy. The first column gives the implicit training sample that corresponds to the current standard deviation of beliefs about μ if starting from a flat prior. The second column gives the actual conditional standard deviation of beliefs. The conditional, annualized Sharpe ratio of consumption claim, the real risk-free rate, and the spread between the five-year yield and the risk-free rate (the yield slope) are given for the case of Epstein-Zin preferences with $\gamma = 10$ and $\psi = 1$, as well as the power utility case where $\gamma = 10$ and $\psi = 1/10$.

After observing 10 years or 40 quarters of data, the price of risk is "only" 0.27 with power utility, equal to the known parameters case. Thus, it is true that, with power utility, the asset pricing implications of parameter uncertainty vanish rapidly and learning has minimal asset pricing implications. With EZ preferences, however, the price of risk is 1.27 after 10 years of learning and, even after 100 years of learning, the price of risk is 0.37, 37 percent higher than the known parameters case. Thus, even though the uncertainty about μ declines quickly, the pricing effects of parameter uncertainty are long-lasting when $\gamma > 1/\psi$.

Why does parameter uncertainty have such long-lived effects with EZ preferences? It is not due to short-run consumption growth volatility since $\sqrt{1 + A_t}$ quickly converges to one, which is also the reason why there is little effect with power utility. Further, belief shocks are relatively small since innovations to μ_t are approximately proportional to A_t and $A_t = t^{-1}$. However, parameter uncertainty declines at a decreasing rate, thus learning "slows down" over time. While A_t falls quickly, the combination of aversion to long-run risks, $\gamma - 1 > 0$, and permanent shocks that impact marginal utility with a multiplier of $\beta/(1-\beta)$ translate into high risk prices as the total multiplier is $(\gamma - 1)\,\beta(1-\beta)^{-1} = 1{,}491$.

Figure 1 graphs the impulse-response of subjective log consumption levels to a one standard deviation positive consumption shock for known and unknown μ, assuming $A_t = 0.01$ or 25 years of learning. With known parameters, consumption shocks only affect current consumption. With unknown parameters, a positive consumption shock increases current consumption *and* the expected growth rate μ_t, affecting consumption levels *indefinitely* into the future via the martingale property. The one-period consumption response is the same in both cases, but they have quite different long-run implications. The lower panel shows the response of the log SDF to this positive shock when learning about μ for both EZ and power utility. The response of the SDF is more than three times larger with EZ preferences.

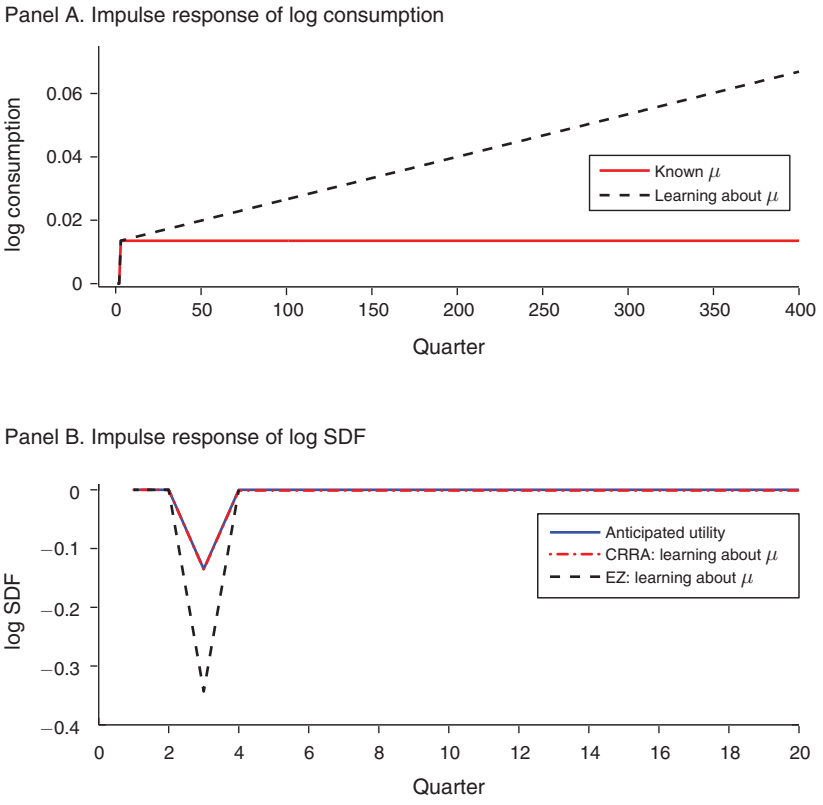

FIGURE 1. IMPULSE-RESPONSE OF CONSUMPTION AND THE SDF

Notes: The top plot shows the impulse-response to the subjective consumption dynamics, which are relevant for utility and pricing, in the case where the mean growth rate is known (solid line) and not known (dashed line). In particular, the shock occurs in quarter 3 and the lines correspond to the difference in expected log consumption in each future quarter based on information available at time 3 (after the shock) and information available at time 2 (before the shock). The low plot shows the corresponding impulse response for the log stochastic discount factor (SDF) for different cases where the agent is learning about μ. The solid line corresponds to the anticipated utility case, the dashed line corresponds to the Epstein-Zin case, while the dashed-dotted line corresponds to the constant relative risk aversion (CRRA) (power utility) case.

Parameter uncertainty also affects the real (log) risk-free rate, $r_{f,t} = \mu_t + f_t$, where

$$f_t = -\ln\beta + \frac{\sigma^2(1+A_t)}{2} - SR_t\sigma\sqrt{1+A_t}. \tag{8}$$

Since the EIS is one, the risk-free rate moves one-to-one with expected consumption growth (μ_t). f_t is a deterministic function of the posterior variance affecting precautionary savings. For $\gamma > 1$, parameter uncertainty increases precautionary savings (through the term $SR_t\sigma\sqrt{1+A_t}$), thus decreasing the risk-free rate. Over time, as

parameter uncertainty declines, the risk-free rate drifts upward, as shown in Table 1. For power utility, the risk-free rate appears constant.[14]

Parameter uncertainty also affects the shape of the default-free yield curve. Since μ_t is a martingale, the current mean belief does not affect the slope. The $\tau-$year (continuously compounded) slope is

$$(9)\quad y_{t,\tau} - r_{f,t} = \frac{1}{\tau}\sum_{k=1}^{\tau-1}\left(f_{t+k} - f_t - \frac{(\tau-k)^2\sigma^2}{2}\frac{A_{t+k-1}^2}{1+A_{t+k-1}} - \frac{(\tau-k)\,\sigma A_{t+k-1}}{\sqrt{1+A_{t+k-1}}}SR_{t+k-1}\right).$$

The first term, $f_{t+k} - f_t$, reflects expectations of future risk-free rates relative to today's rate and impacts the slope positively. The second term is a small Jensen's inequality term. The third term is the risk premium on long-term bonds, which is negative because negative shocks to consumption decrease the risk-free rate through a lower μ_t. As parameter uncertainty declines over time, this term shrinks. Overall, Table 1 shows that the slope is fairly flat (actually mildly negative for our parameters) in the EZ case. With power utility and a low EIS value, $\psi = 0.1$, the risk-free rate is quite volatile and the slope is more negative.

In summary, parameter learning and EZ utility with a preference for early resolution of uncertainty ($\gamma > 1/\psi$) substantially increases risk premia, decreases risk-free rates, and generates a roughly flat yield curve. The case of an EIS $\neq 1$, which requires a numerical solution, is discussed in the online Appendix, and can be easily summarized. When the substitution effect dominates the wealth effect (i.e., the EIS > 1), the price-consumption ratio increases upon a positive revision of the beliefs about the growth rate. Overall, the primary effect of increasing the EIS is an increase in excess return volatility, which, in turn, increases the risk premia, both of which are important for matching historical asset price data.

Relation to Bansal and Yaron (2004).—The subjective consumption dynamics with learning, equations (4) to (5), seem similar to the exogenous dynamics in BY. In particular, both feature small, highly persistent shocks to expected consumption growth.[15] Despite the similarities, there are important asset pricing differences.

First, unlike BY, who have persistent but transitory shocks to expected consumption growth, parameter learning generates permanent shocks (i.e., truly long-run risk) in subjective consumption growth. Second, consumption growth with parameter uncertainty (equation (3)) is unpredictable, unlike BY's model which implies significant predictability of consumption growth by price-dividend ratios and real risk-free rates for example. This "excess predictability," which has been seen as a weakness of long-run risk models (see Beeler and Campbell 2012), is not present with learning-induced long-run risks.

Related, Hall (1988) estimates the EIS to be close to zero in regressions of consumption growth on lagged risk-free rates, seemingly in contrast to the high EIS assumed here and to the predictions of long-run risk models such as BY. With parameter learning, however, simulated data regressions also generate EIS estimates close

[14] There is actually a slight (economically insignificant) upward drift of 0.1 bps over 10 to 100 years.

[15] In Bansal and Yaron, conditional expected consumption growth follows an AR(1) with monthly autoregression coefficient of 0.979: that is, with a half-life of about four years.

to zero, because consumption growth is actually i.i.d. and unpredictable in population. Thus, parameter uncertainty implies there is no contradiction between a high EIS and Hall's regression evidence. Finally, the real term structure in the model is on average only weakly downward-sloping, as opposed to strongly downward-sloping in BY. Backus, Chernov, and Zin (2014) and Beeler and Campbell (2012) argue that strongly downward-sloping real yield curves are counterfactual.

In sum, long-run risks due to parameter learning are substantively different from the long-run risks typically assumed in the literature along empirically relevant dimensions.

Relation to Anticipated Utility.—Anticipated utility is arguably the benchmark approach for dealing with parameter uncertainty, see, e.g., Cogley and Sargent (2008); Piazzesi and Schneider (2007); and Johannes, Lochstoer, and Mou (forthcoming). AU agents learn about parameters over time, but ignore parameter uncertainty when making decisions. That is, they calculate utility and asset prices treating the current posterior means as the "true" constant parameters. After observing additional data, agents update their parameter beliefs and recalculate utility and prices using the revised estimates, again assuming these estimates will never change. Thus, AU ignores the endogenous long-run risk channel generated by belief updating when pricing assets.

The AU case admits analytical solutions. To facilitate comparisons, for $\psi = 1$[16]

$$vc_t^{AU} = \frac{\beta}{1-\beta}\Big(\mu_t + \frac{1}{2}(1-\gamma)\,\sigma^2\Big), \tag{10}$$

the same as utility with i.i.d. normal consumption growth and known parameters (i.e., equation (6) with $A_t = 0$) and replacing μ with μ_t. This suggests there are key differences between the economics of priced parameter uncertainty and AU. To quantify these differences, we compute the fraction α_t of wealth a rational EZ agent facing parameter uncertainty would be willing to forgo (at current time t) to avoid parameter uncertainty and have parameters set to constants equal to current mean beliefs:

$$\alpha_t = 1 - \frac{\exp(vc_t)}{\exp\big(vc_t^{AU}\big)} = 1 - \exp\left(\frac{\sigma^2(1-\gamma)}{2}\sum_{j=0}^{\infty}\beta^{j+1}\left[\frac{\big(A_{t+j}/(1-\beta)+1\big)^2}{1+A_{t+j}} - 1\right]\right) \tag{11}$$

α_t is positive when the agent has a preference for early resolution of uncertainty ($\gamma > 1 = 1/\psi$) provided $A_t > 0$. Thus, under EZ preferences the AU assumption ignores a potentially major source of macroeconomic risk.

The third column reports the welfare loss of rationally accounting for parameter uncertainty relative to the AU case. The welfare loss can be quite large, 53 percent after 10 years of learning, and still 9 percent after 100 years of learning. In fact, *parameter uncertainty is the main component of macroeconomic risk*, as the agent

[16] The general solution for $\psi \neq 1$ is $vc_t^{AU} = \frac{1}{\rho}\log\frac{1-\beta}{1-\beta e^{\rho\mu_t+\rho(1-\gamma)\sigma^2/2}}$ with $\rho = 1 - 1/\psi$.

would pay more to eliminate it than he would be willing to pay in the known mean case to get rid of all risk (in the latter case the agent would pay 12 percent of total wealth to get rid of all consumption risk, i.e., to set $\sigma = 0$).

To price assets, AU uses the SDF for a payoff at time $t + j$ given by

$$M_{t,t+j}^{AU} = \tilde{\beta}_t^j \exp\left(-\gamma \Delta c_{t,t+j}\right), \tag{12}$$

where $\tilde{\beta}_t = \beta \exp\left(\mu_t + \frac{1}{2}(1-\gamma)\sigma^2\right)^{\gamma - 1/\psi}$. Thus, ex ante AU risk pricing is the same as if μ were known (assuming expected consumption growth is μ_t), generating an ex ante Sharpe ratio and risk premium of log consumption claim returns of $SR^{AU} = \gamma\sigma$ and $RP^{AU} = \gamma\sigma^2$, respectively. Thus, with AU, there is no ex ante pricing of parameter uncertainty.

The expected log return volatility with AU is σ, but, due to next period's belief update, the actual realized conditional log return volatility is $\sigma\sqrt{1 + A_t}$ the same as with rationally priced parameter uncertainty. For $\psi = 1$, the agents' perceived ex ante return volatility remains the same. However, if $\psi > 1$ and the substitution effect dominates, the *actual* return volatility becomes larger as the wealth-consumption ratio increases (decreases) upon a positive (negative) shock, while the opposite is the case if $\psi < 1$. Thus, the AU assumption can either over or understate actual return volatility depending on the EIS.

Further, with AU the risk-free rate is only a function of μ_t and σ^2:

$$r_{f,t}^{AU} = -\ln\beta + \frac{1}{\psi}\left(\mu_t + \frac{\sigma^2}{2}\right) - \frac{1}{2}\gamma\left(1 + \frac{1}{\psi}\right)\sigma^2. \tag{13}$$

This implies that the yield curve is flat, as the AU agent prices bonds as if consumption growth is i.i.d. The last three columns of Table 2 confirm that the price of risk, the risk-free rate, and the yield spread are always the same as in the corresponding known parameter E-Z case.

In sum, ignoring priced parameter uncertainty and using AU generates dramatically different economic implications when agents have a preference for the timing of the resolution of uncertainty.

Relation to Weitzman (2007).—The asset pricing implications of learning are due to the subjective long-run risk channel that arises endogenously from the *updating* of beliefs. This is in stark contrast to the mechanism identified in Geweke (2001) or Weitzman (2007), where learning about the volatility parameter σ in a discrete-time setting leads to extreme tail risk in subjective consumption growth (its predictive distribution has fat tails).[17]

[17] Geweke (2001) and Weitzman (2007) also consider learning about σ^2 with power utility and note that standard conjugate priors imply that the predictive distribution of consumption growth is t-distributed and thus expected utility may not exist. However, Bakshi and Skoulakis (2010) argue that their results are not robust as they disappear when the prior support of σ^2 is truncated even at very high values. Interestingly, we find similar results for EZ preferences, based on a numerical solution discussed in the online Appendix. Since learning about σ^2 is more rapid than learning about μ and volatility has only a second-order effect on continuation utility, we find that the asset pricing effects of learning about σ^2 are negligible with Epstein-Zin preferences when priors about variance are truncated at reasonable values.

Table 2—Learning about the Mean Anticipated Utility

Degree of parameter uncertainty		Anticipated utility ($\gamma = 10$, $\psi = 1$)			
			Annualized, conditional		
$A_t = \frac{1}{t}$ t (quarters)	$\sigma_t(\mu) = \sqrt{A_t}\sigma$ percent	Welfare loss of rational EZ versus AU (α_t) percent	Price of risk	Risk-free rate	Yield slope
40	0.21	53	0.27	0.035	0
100	0.13	29	0.27	0.035	0
200	0.10	17	0.27	0.035	0
400	0.07	9	0.27	0.035	0
800	0.05	5	0.27	0.035	0
∞	0.00	0	0.27	0.035	0

Notes: This table shows various conditional, annualized moments for the "learning about the mean" economy under anticipated utility (AU) pricing. The first column gives the implicit training sample that corresponds to the current standard deviation of beliefs about μ if starting from a flat prior. The second column gives the actual conditional standard deviation of beliefs. The third column gives the welfare loss of going from AU pricing to rationally accounting for parameter uncertainty ex ante (Epstein-Zin (EZ) pricing). The conditional, annualized Sharpe ratio of consumption claim (price of risk), the real risk-free rate, and the spread between the five-year yield and the risk-free rate (the yield slope) are given for the case of Epstein-Zin preferences with $\gamma = 10$ and $\psi = 1$.

Normal distributions also allow for extreme values of μ. To show that our results are not driven by extreme (and unreasonable) perceived tail events, consider a truncated normal prior $p(\mu \mid y^t) \sim \mathcal{N}_T(\mu_t, A_t\sigma^2; \underline{\mu}, \overline{\mu})$, where $\underline{\mu}$ ($\overline{\mu}$) is the lower (upper) truncation bound for μ. The prior is still conjugate and updating equations for the state variables μ_t and A_t are unchanged from the untruncated case, although μ_t and $A_t\sigma^2$ are no longer the conditional moments.

Analytical solutions are not available in this case and the model must be solved numerically using methods described in the online Appendix. We find that, when truncation bounds are set at 0 percent and 1 percent quarterly (i.e., 0 percent and 4 percent in annual terms), our previous results are essentially unchanged. This confirms that it is the permanent shocks arising from parameter *updating* that matter in our model, not the possibility of extreme perceived values of μ.

C. *Discussion*

The i.i.d. normal model conveys the main intuition for how parameter uncertainty, in combination with Epstein-Zin preferences, amplifies the impact of macroeconomic shocks on marginal utility and, therefore, asset prices. This endogenous long-run risk channel induced by parameter learning provides a novel mechanism that has distinct implications from seemingly similar models such as BY or learning approaches such as AU.

However, as equation (7) and Table 1 indicate, the model implies a *deterministic* decrease in the price of risk, and therefore risk premia, that is sufficiently fast to induce a counterfactually strong upward time trend in valuation ratios such as the price-dividend ratio when looking at samples of length like those typically considered in the literature.[18] Further, as argued by, e.g., Cochrane (2011), historical data

[18] While Fama and French (2002) argue that the market Sharpe ratio and risk premium indeed decreased over the available CRSP sample, the increase in valuation ratios over this sample, like price-dividend (including

indicate instead substantial *countercyclical* movement in the conditional market risk premium and Sharpe ratio.

In the following, we consider the impact of parameter uncertainty in two benchmark consumption-based asset pricing models—the long-run risk model of BY, and the disaster risk model following Rietz (1988) and Barro (2006). Parameter uncertainty in these models is highly realistic as they both feature more complicated consumption dynamics governing rare or difficult to measure components of consumption growth. In these models, the nonstationarities induced by parameter learning do not imply counterfactually strong drifts in valuation ratios or decreases in risk prices even over long samples. In both sets of models, parameter learning induces strong countercyclical variation in risk prices and premia.

Unfortunately, it is not generally possible to study these models by allowing for parameter uncertainty in *all* parameters due the dimensionality of the problem, as parameter uncertainty typically adds more than one state variable for each uncertain parameter. Additionally, as shown above for the volatility parameters, parameter uncertainty does not appear to be important for all parameters. To understand the impact of parameter uncertainty in these classic models, we follow two steps. First, we propose a simple metric that allows us to identify which parameters are most relevant from the point of view of parameter uncertainty. The metric is essentially a measure of sensitivity of the value function to changes in parameters, taking into account prior uncertainty. We then analyze the fully dynamic model allowing for uncertainty only in these "most relevant" parameters. In rare-disaster model the most relevant parameters turn out to be related to the disaster transition probabilities. Second, we consider model uncertainty. In that approach agents are uncertain about a set of different models that they each estimate ignoring parameter uncertainty. That is, even though each model can have many parameters estimated with uncertainty, agents only consider the trade-off between the different model specifications.[19] We apply this approach to the long run risk model of BY and consider the uncertainty in its model specification relative to the i.i.d. model.

II. Parameter Uncertainty in Rare Disaster Models

Rare disaster models provide an excellent laboratory to understand priced parameter uncertainty as there are many parameters and the parameters describing disasters are highly uncertain given their "rare" occurrence. Relative to the previous case, we find that parameter learning in this model generates larger and longer lasting effects, negligible nonstationarities over samples of the lengths we typically analyze, and large countercyclicality in the price of macroeconomic risks.

We follow Rietz (1988), Barro (2006, 2009), and Nakamura et al. (2013), who propose dynamics with large and rare consumption disasters. Empirical papers document the difficulty in estimating disaster parameters, even with large data samples. For example, using more than a century of data and a broad panel of countries, Nakamura et al. (2013) estimate a disaster frequency of 2.8 percent per year and

repurchases) or price-earnings, is still relatively modest.

[19] This effectively makes uncertainty one-dimensional (in the case of two models) or n-dimensional in the case of $n + 1$ models.

a probability of exiting a disaster of 13.5 percent per year. The standard errors are high: the two standard error bounds for the average duration of the bad state are respectively 4.5 and 9 years. There is also a large amount of uncertainty over the size (mean and variance) of consumption disasters.

We consider a parsimonious two-state Markov switching model,

$$\Delta c_t = \mu_{s_t} + \sigma_{s_t}\varepsilon_t,$$

where $\varepsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, s_t is a Markov chain with transition matrix

$$\Pi = \begin{bmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{bmatrix}.$$

Without loss of generality, we label $s_t = 1$ the "good" or normal state and $s_t = 2$ the "bad" or rare event state. To focus on parameter uncertainty, we assume s_t is observed.

There are six parameters in this model. If all parameters were uncertain and using conjugate priors, there would be nine state variables, including time. To focus the results and avoid the curse of dimensionality, we would like to allow for parameter uncertainty in only a small number of parameters whose uncertainty matters most from an ex ante utility sense, and treat all the other parameters as known. Of course, the issue is then how to identify these parameters without solving a model where all parameters are uncertain.[20]

To do this, we propose using the fraction of wealth an agent would be willing to pay to avoid parameter uncertainty altogether by setting parameters equal to their prior mean.[21] This metric was used earlier to compare rational parameter uncertainty with AU. To operationalize the metric with multiple unknown parameters, we compare the continuation utility of an EZ agent who will learn the true parameter in one period, i.e., will face one period of parameter uncertainty, to that of an AU agent.

Formally, we define it 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. Both are defined recursively as a function of their initial priors and their certainty equivalent next period (see the online Appendix). The *parameter uncertainty premium*, α_t, is defined as in equation (11) as

$$\alpha_t = 1 - \frac{V_t^{One}}{V_t^{AU}}. \tag{14}$$

The crucial point, which makes this metric simple and easy to compute, is that the certainty equivalent next period only depends on the known-parameter value function (which is much simpler to calculate), because of our assumption that agents only face one period of parameter uncertainty.

[20] This question 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.

[21] This approach is related to Lucas (1987), who considers how much an agent would pay to eliminate business cycle risk and to Epstein and Farhi (2014), who consider what fraction of wealth an agent would pay to have all uncertainty resolved in one period.

In the online Appendix, we provide further intuition about this metric that confirms 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. While intuitive, the metric's value is from *quantifying* the impact of parameter uncertainty, without solving a full dynamic model.

A. *Calibration of Priors and Preference Parameters*

To analyze parameter uncertainty, we need to specify priors and preference parameters. We solve the model quarterly so that the time-averaged, annual log consumption growth mean and standard deviation are 1.8 percent and 2.2 percent, respectively, matching the observed values from 1929 to 2011. The "true" parameters are calibrated as follows. The bad state mimics the Great Depression, when real, per capita log consumption declined -4.6 percent per year from 1929 to 1933 with 2.94 percent volatility per year ($\mu_2 = -1.15$ percent and $\sigma_2 = 1.47$ percent quarterly). We set $\pi_{11} = 383/384$ and $\pi_{22} = 15/16$, corresponding to one four-year depression per century. The mean and variance in the normal growth state are $\mu_1 = 0.54$ percent and $\sigma_1 = 0.98$ percent.[22]

We use standard, conjugate priors distributions: beta, normal, and inverse gamma distributions for the transition probabilities, mean parameters, and variance parameters, respectively. Priors are centered at the true values to insure unbiased beliefs. We choose the remaining prior parameters to encode various training sample lengths (e.g., 100, 200, or 300 years), that form "fictitious" samples which the agent would have used to form beliefs starting from a flat prior. For instance, a 200-year prior means the agent has current mean beliefs equal to truth and a prior variance equal to those from a dataset with two depressions (each lasting 16 quarters) and two normal time periods (each lasting 384 quarters). Given unbiased priors, the length of the history the agent can observe (e.g., 200 years) is sufficient to set the scale parameters of all priors. To ensure existence of equilibrium for an EIS different from one, it is necessary to truncate the prior distributions.[23] The truncation bounds are wide, so as to not rule out economically sensible values of the parameters and are described in detail in the online Appendix.

The preference parameters are consistent with recent work using disaster risk models. Following Nakamura et al. (2013); Gourio (2012); Bansal, Kiku, and Yaron (2013); and Bansal et al. (2014), we use EIS values of 1.5 and 2. The literature provides a wide range of estimates for this parameter. Hall (1988) and Campbell (1999) argue the EIS is less than 1, but Attanasio and Weber (1989); Beaudry and van Wincoop (1996); Vissing-Jørgensen (2002); and Vissing-Jørgensen and Attanasio (2003) argue that the EIS is large and in fact greater than 1, with Gruber (2013) estimating a value of about 2. We set the risk aversion parameter γ to be 2 or 4 in our

[22] A two-state model with the "bad" state calibrated to the Great Depression clearly misses normal business cycle fluctuations. These could be introduced by adding an additional more frequent state. However, to retain parsimony we abstract from business cycle fluctuations in the conditional moments of consumption growth. The next section presents a model generating time variation in asset prices and risk premia at business cycle frequency.

[23] For instance, if the mean growth rate in the good state is allowed to be arbitrarily high, as is the case with a normal prior, utility will be infinite. Note, however, that conjugate priors remain conjugate under truncation, so the truncation does not significantly complicate the numerical model solution.

benchmark calibrations, similar to Barro (2006), and well within the upper bound of 10 argued as reasonable by Mehra and Prescott (1985). As in BY, and in earlier sections, we set β to 0.994.

B. *Which Parameters Are Important and Why?*

We first apply our metric to the rare events model to explore how uncertainty about different parameters affects utility and to identify the parameters for which uncertainty is particularly important. Table 3 quantifies the parameter uncertainty premium for each parameter. The benchmark model has known parameters equal to the posterior mean. Then, we introduce uncertainty about a single parameter, allowing for differing degrees of uncertainty by varying the degree of prior information based on training samples of 100 and 300 years.

The parameter uncertainty premium reflects the different utility levels arising from consumption dynamics with and without parameter uncertainty. To further understand the source of the premium, we also report in Table 3 the difference in the conditional mean and standard deviation of next period's log continuation utility normalized by consumption, vc_{t+1}, for cases with and without parameter uncertainty.[24] At time $t+1$ all parameters are known for both models, so the difference in these moments is entirely due to the parameter learning between time t and time $t+1$.

Overall, the parameter uncertainty premium is decreasing in the length of the prior training sample (prior precision), increasing in risk aversion, and not highly sensitive to the EIS levels considered.[25] Interestingly, Table 3 shows there is strong heterogeneity in the parameter uncertainty premium. Variance parameter uncertainty has a near zero premium, mean parameter uncertainty has a modest premium, and transition probability uncertainty has large premia. The premium is higher for uncertainty about parameters governing the dynamics of the bad (high marginal utility) state.

The parameter governing the persistence of the bad state, π_{22}, is associated with a particularly high and long-lasting parameter uncertainty premium: 28 percent if $\gamma = 2$ and 54 percent if $\gamma = 4$ for the 100-year prior when $\psi = 2$, and 11 percent if $\gamma = 2$ and 35 percent if $\gamma = 4$ for the 300-year prior when $\psi = 2$. This is partially due to a lower expected level of next period's utility as utility is concave in the transition probability, but also due to high conditional standard deviation in next period's utility induced by learning about this parameter. The high volatility of the continuation utility arises for two reasons. First, transition probabilities concerning rare states are hard to learn, which means prior uncertainty is relatively high. Second, the agent is particularly averse to persistent bad states and so utility with known parameters is very sensitive to the value of π_{22}.

The results for uncertainty about the mean growth rate illustrates that the effects of parameter uncertainty can be highly nonlinear in risk aversion. In this case, there

[24] If utility is nonlinear in a particular parameter, a mean-preserving spread in that parameter affects expected utility, which would also be reflected in the parameter uncertainty premium.

[25] Of course, we only consider levels of EIS such that $\gamma > 1/\psi$. Setting the EIS to the inverse of the risk aversion, as in power utility, would make the parameter uncertainty premium close to zero, as discussed earlier.

TABLE 3—WHICH PARAMETERS MATTER THE MOST? THE RARE EVENTS MODEL

	100-year prior			300-year prior		
	Par. unc. premium	Difference in (vc_{t+1})		Par. unc. premium	Difference in (vc_{t+1})	
Parameter	percent	Mean	SD	percent	Mean	SD
Panel A. Elasticity of intertemporal substitution (ψ) *of 2, risk aversion* (γ) *of 4*						
μ_1	1.1	0.00	0.05	0.4	0.00	0.02
μ_2	8.1	−0.02	0.13	2.6	−0.01	0.06
σ_1	0.0	0.00	0.00	0.0	0.00	0.00
σ_2	0.0	−0.00	0.00	0.0	−0.00	0.00
π_{11}	5.6	0.08	0.23	1.8	0.00	0.06
π_{22}	54	−0.40	0.55	35	−0.18	0.34
Panel B. Elasticity of intertemporal substitution (ψ) *of 1.5, risk aversion* (γ) of 4						
μ_1	0.9	0.00	0.04	0.3	0.00	0.01
μ_2	7.2	−0.03	0.11	2.3	−0.01	0.05
σ_1	0.0	0.00	0.00	0.0	0.00	0.00
σ_2	0.0	−0.00	0.00	0.0	−0.00	0.00
π_{11}	4.7	0.06	0.20	1.4	0.00	0.05
π_{22}	56	−0.42	0.55	35	−0.18	0.33
Panel C. Elasticity of intertemporal substitution (ψ) *of 2, risk aversion* (γ) *of 2*						
μ_1	0.4	0.00	0.08	0.1	0.00	0.04
μ_2	0.4	−0.00	0.03	0.1	−0.00	0.01
σ_1	0.0	0.00	0.00	0.0	0.00	0.00
σ_2	0.0	−0.00	0.00	0.0	−0.00	0.00
π_{11}	−0.1	0.02	0.15	0.2	0.00	0.05
π_{22}	28	−0.29	0.43	11	−0.09	0.19

Notes: This table shows the parameter uncertainty premium, as defined in the main text, with differing degrees of uncertainty over different parameters in the two-state switching regime model. In all cases $\beta = 0.994$. Par. unc. premium denotes parameter uncertainty premium, Difference in (vc_{t+1}) Mean and SD denote the difference in the conditional mean and standard deviation, respectively, of next period's log utility (normalized by consumption) between the anticipated utility case and the case where the agent ex ante accounts for parameter uncertainty and knows that the parameter is revealed next period.

is an order of magnitude increase in the parameter uncertainty premium when going from $\gamma = 2$ to $\gamma = 4$, though the premium even for the $\gamma = 4$ case is still only about one-sixth or less of the premium for the transition probability of the same state.

It is also apparent from the different priors that learning about the transition probabilities is slow and generates long-lasting implications. For instance, for the $\gamma = 4$, $\psi = 2$ case, the welfare loss of uncertainty in μ_2 drops by about two-thirds, from 8.1 percent (100-year prior) to 2.6 percent (300-year prior). For π_{22}, the welfare loss only drops by about one-third, from 54 percent to 35 percent. Intuitively, there is on average only one observation per century about the transition probabilities. For the mean growth rates, however, there are 384 and 16 observations on average per century of data, and so learning is quicker for the mean parameters. Further, the persistence of the bad state has particularly adverse effects on utility in the known parameter case.

Uncertainty about the variance has negligible utility effects for two reasons. First, utility with known parameters is not very sensitive to variance parameters, which are second moments of the consumption dynamics. Second, variance parameters are relatively easy to learn so uncertainty about variance parameters is relatively low given that our priors have implicit in them long samples of prior learning. This

differs from the conclusion of Weitzman (2007) and is due to the truncation of the prior distribution for the variance parameters.[26]

C. *Long Sample Moments*

We now turn to the full-fledged parameter learning model where parameter learning resolves dynamically over time. Given the results in the previous section, we analyze the case where the transition probabilities (π_{11}, π_{22}) are unknown, while the other parameters are known and set equal to their true values.[27]

As in Campbell and Cochrane (1999) and Bansal and Yaron (2004), equity is a claim to an exogenous dividend stream:

$$\Delta d_{t+1} = \overline{\mu} + \lambda(\Delta c_{t+1} - \overline{\mu}) + \sigma_d \eta_{t+1}, \tag{15}$$

where d_t is the log of dividends, $\overline{\mu}$ 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). This specification implies that the long-run growth rate of dividends and consumption is equal, while the short-run response of dividends to consumption shocks is higher than that of consumption.[28] Using the same sample as earlier, we estimate the leverage parameter to be 2.5 by regressing annual real dividend growth on annual consumption growth. We set the idiosyncratic volatility σ_d such that annual dividend volatility is 12.9 percent, as in the data.[29] We compute standard asset pricing moments for the same-length sample (82 years) by averaging these moments across 20,000 simulated samples.

Table 4 shows the average log excess equity returns, equity return volatility, and real risk-free rates, as well as the average sample Sharpe ratio of gross equity returns and price of risk (Hansen and Jagannathan 1991 bound). All moments are annualized. In panel A, $\gamma = 4$ and $\psi = 2$, while in panel B, $\gamma = 4$ and $\psi = 1.5$. The data column gives the empirical sample averages. The real quarterly risk-free rate is constructed using nominal 90-day T-bill rates minus expected inflation, where the latter is obtained from an AR(4) for inflation estimated on the full sample. Market returns are obtained from the Center for Research in Security Prices (CRSP) files.

Overall, parameter learning leads to a substantial increase in risk premia, Sharpe ratios, and the price of risk, as well as a decrease in the risk-free rate. Importantly, the effects are quantitatively significant for these long-sample averages even with a prior that embodies 300 years of prior learning. Relative to the case of known

[26] See Bakshi and Skoulakis (2010) for a thorough discussion of this point and Table 3 above.

[27] The online Appendix gives extensive details for the numerical methods used to solve the model and presents results when the mean and/or variance parameters are unknown. We find that, consistent with the implications of Table 3 discussed previously, the impact of uncertainty about the mean and variance parameters is small relative to the impact from unknown transition probabilities also in the full learning model.

[28] In terms of exposure to parameter uncertainty, our dividend assumption is conservative relative to the more standard specification, $\Delta d_{t+1} = \lambda \Delta c_{t+1} + \sigma_d \eta_{t+1}$. In particular, the uncertainty about the long-run growth rate is the same for consumption and dividend, and $\overline{\mu} = E(s_1)\mu_1 + \big(1 - E(s_1)\big)\mu_2$. Alternatively, one could assume consumption and dividends are cointegrated, which introduces another state variable and is computationally costly.

[29] Aggregate dividends are constructed using monthly ex- and cum-dividend CRSP market returns to get monthly market dividends. These monthly dividends are summed within each year and the annual growth rates are inflation adjusted. We replicate this approach in the model and match the volatility of this measure of log annual dividend growth to the data with $\sigma_d = 6.65$ percent.

parameters, the risk-premium is about five to six times higher with parameter learning, the price of risk is about three to four times higher, and the equity Sharpe ratio is about two to three times higher.

As seen earlier in the i.i.d. case, parameter learning provides a powerful amplification mechanism. Unlike the i.i.d. normal case, the risk premium decreases quite slowly due to the infrequent nature of the rare events. In fact, there is on average only about a 10 percent increase in the price-dividend ratios over the 100-year samples due to a very slowly decreasing discount rate. In the data, the price-to-cash dividend ratio has increased by about 90 percent over the same period. Although some of this is likely due to other factors (such as stock repurchases) than learning, this shows that the amount of nonstationarity in valuation ratios due to learning is not unreasonable compared to historical data. Further, the amount of parameter uncertainty investors face in these calibrations is, in our view, conservative. For instance, using 300-year prior in 1929 assumes agents effectively had data back to the early 1600s, around the time when the first stock exchange, the Amsterdam Stock Exchange, was opened and before the industrial revolution.

The results in panel A versus panel B of Table 4 show that similar results arise whether $\psi = 2$ or $\psi = 1.5$. The risk pricing of shocks to beliefs, which is driven by $\gamma - 1/\psi$, is not strongly affected by the EIS at these parameter values. The main difference with a lower EIS is an increase in the real risk-free rate to a level higher than the data.[30]

The last column in Table 4 shows the sample moments that arise from an AU implementation of the model. In this case, we only show results with the 200-year prior. As expected, the risk premium, the price of risk, and the average real risk-free rates are roughly the same as in the known parameters case. Given that beliefs still fluctuate and since prices are recomputed based on revised beliefs each period, the return volatility is slightly higher under AU, which makes the average sample equity Sharpe ratio slightly lower. Thus, unlike the case of power utility analyzed in Cogley and Sargent (2008), AU does not yield a good approximation to the true equilibrium outcomes when agents have EZ preferences.

D. *Conditional Dynamics*

This section documents that parameter learning leads to strong countercyclicality in conditional risk premia, return volatility, and Sharpe ratios. For brevity, we consider only the economy with $\gamma = 4$, $\psi = 2$, and a 200-year prior.

Belief Dynamics and Valuations.—We illustrate the conditional dynamics though a 40-year sample where the economy enters a depression event in year 11 (quarter 41). The realized depression length is random, given the two-state Markov switching model, and we consider average durations (4 years) as well as extreme durations of 1 quarter and 12 years, which correspond roughly to 5 percent and 95 percent outcomes of the true distribution of depression durations.

[30] Note that this could be countered by allowing for a higher value of β: see Kocherlakota (1996).

Table 4—Sample Moments Learning about the Probability and Persistence of a Great Depression

	Data	Priors (training sample) 100 yrs	200 yrs	300 yrs	∞ yrs	Anticipated utility
Panel A. $\psi = 2$						
$E_T[r_m - r_f]$ (percent)	5.0	5.63	4.81	4.26	0.87	0.93
$\sigma_T[r_m - r_f]$ (percent)	20.2	18.1	18.4	18.5	17.3	18.8
$SR_T[R_M - R_f]$	0.35	0.37	0.32	0.30	0.14	0.13
$E_T[r_f]$ (percent)	0.7	0.81	1.38	1.69	2.92	2.93
$\sigma_T[M]/E_T[M]$	n/a	1.12	0.90	0.78	0.27	0.26
Panel B. $\psi = 1.5$						
$E_T[r_m - r_f]$ (percent)	5.0	4.92	4.27	3.78	0.67	0.72
$\sigma_T[r_m - r_f]$ (percent)	20.2	17.4	17.8	18.0	16.9	18.1
$SR_T[R_M - R_f]$	0.35	0.35	0.31	0.28	0.12	0.12
$E_T[r_f]$ (percent)	0.7	1.85	2.26	2.48	3.33	3.34
$\sigma_T[M]/E_T[M]$	n/a	1.06	0.85	0.74	0.25	0.24

Notes: This table gives average sample moments from 20,000 simulations of 400 quarters of data from the two-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. Lowercase letters denote log of uppercase variable. All statistics are annualized and, except for the Sharpe ratio, given in percent. The time-preference parameter β and risk aversion γ are set to 0.994 and 4, respectively, in all cases. The data column shows the historical excess market return moments for the United States from end of 1929 to end of 2011.

Figure 2 plots mean beliefs about the transition probabilities for these cases, as well as the paths of corresponding log wealth-consumption ratios. Panel A shows $E_t[\pi_{11}]$, which increases slightly until the onset of the depression event, at which point $E_t[\pi_{11}]$ falls sharply. There is, naturally, no learning about π_{11} during a depression, but instead, as panel B shows, the agent starts learning about the persistence of the bad state, $E_t[\pi_{22}]$. Belief revisions about π_{22} are much larger than for π_{11} as investors have less data and more dispersed beliefs. The longer the depression lasts, investors' mean beliefs about π_{22} increase. Note that an "average" depression implies that the beliefs about π_{22} are the same before and after the depression, while a longer (shorter) depression realization leads to, in expectation, a permanent increase (decrease) in the mean belief relative to the belief before the depression event commenced. In all cases, beliefs change with each additional realization of the depression event, causing large shocks to beliefs during the bad state that are absent when this parameter is known.

The time-series of beliefs are reflected via variation in the log wealth-consumption ratio. In particular, at the onset of the depression event, the wealth-consumption ratio falls by more than 30 percent, about twice as much as in the known parameters benchmark case: the amplification effect mentioned earlier. This larger response occurs for two reasons: (i) the sharp downward revision in beliefs about the probability of staying in the good state, π_{11}, and (ii) a strong increase in the discount rate. The latter is due to the higher volatility of belief updates in the bad state. Throughout the depression event the wealth-consumption ratio keeps declining as the persistence of this state is revised upward.

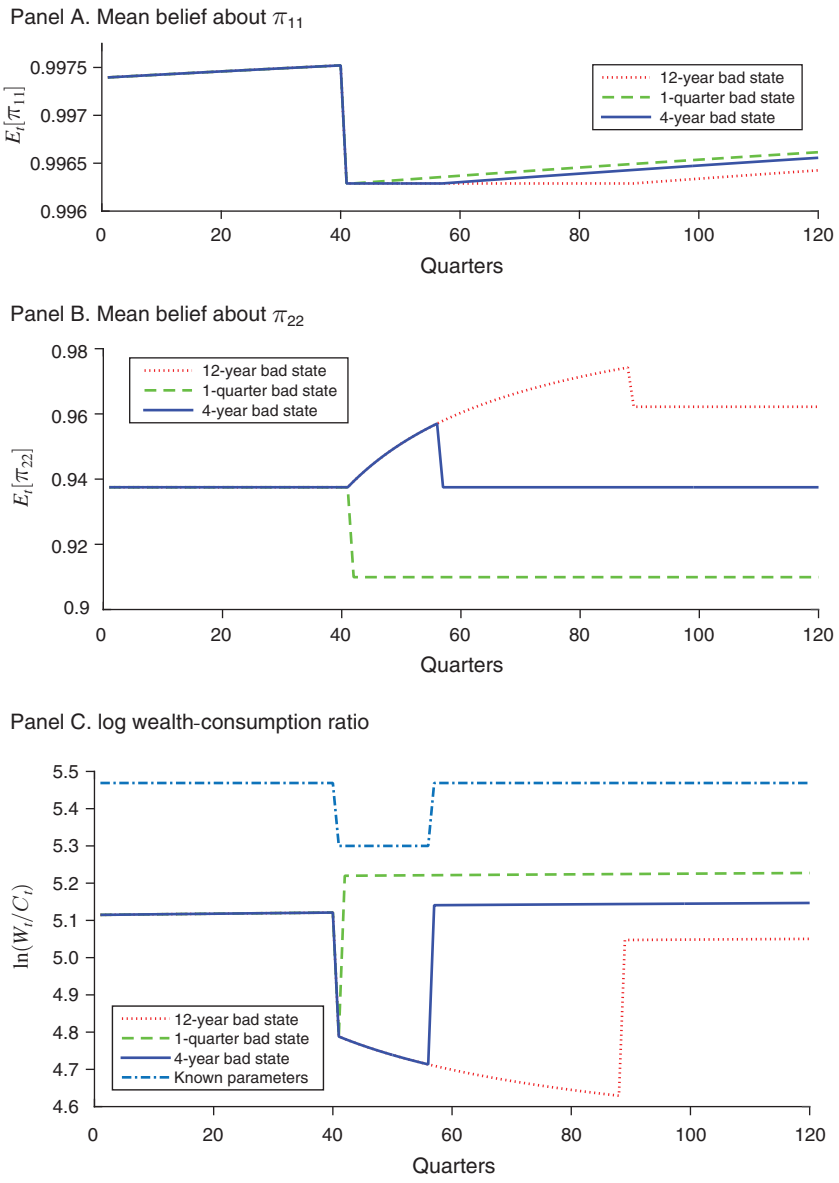

FIGURE 2. BELIEFS, THE W/C-RATIO AND DEPRESSION REALIZATIONS

Notes: Panel A shows simulated paths for the mean beliefs about the probability of staying in the good state, π_{11} for the case of a 200-year prior with unbiased initial mean beliefs. There are three simulated paths, with Depression realizations of 1 quarter, 4 years (the ex ante expected length), and 12 years. The tail outcomes correspond roughly to 5 percent and 95 percent outcomes. Panel B shows the same for the probability of staying in the Depression state, π_{22}. Panel C shows the corresponding log wealth-consumption (W/C) ratios, as well as the log wealth-consumption ratio for the benchmark case where parameters are known (dash-dotted line). The latter is only plotted for a four-year Depression realization.

The differences in postdepression wealth-consumption ratios as a function of depression duration is similar in magnitude to the drop in the wealth-consumption ratio that occurs in a depression when parameters are known. In other words, a significant additional risk with uncertain parameters is the permanent shock to wealth that is a function of the net belief update that results from observing the realized length of a depression event. This wealth shock is of the same order of magnitude as the depression event itself in the known parameters case and highlights how the assumption of known parameters potentially ignores a first-order source of macroeconomic risk.

The Conditional Risk Premium, Return Volatility, and Sharpe Ratio.—Table 5 shows the annualized conditional log excess returns and return volatility of the dividend claim right before and just after the onset of a depression event. These moments are given for the case of a 200-year prior with $\gamma = 4$ and $\psi = 2$, where the depression events starts in quarter 41 of the simulated sample.

With parameter uncertainty, the conditional log risk premium and return volatility increase to 22 percent and 81 percent at the onset of the bad state, up from 4.6 percent and 15 percent the quarter before. The annualized Sharpe ratio of simple returns increases from 0.32 to 0.66. For comparison, Martin (2015) argues that the risk premium at the onset of the financial crises briefly exceeded 55 percent when annualized realized market volatility and the VIX index were greater than 80 percent. Realized volatility in the fall of 1929 also exceeded 80 percent. The fixed parameters case also features strong increases in these moments, but the conditional risk premium and return volatility at the onset of the depression event are still less than one-half of that in the case of uncertain transition probabilities due to the amplification effect.

The table also shows that the drop in the price-dividend ratio of the dividend claim at the onset of the depression is 55 percent in the model with parameter uncertainty and 50 percent in the model with known transition probabilities. In the Great Depression, the drop in the price-dividend ratio from the beginning of the Depression in 1929 to its lowest point in 1932 was 79 percent. For the recent financial crisis, the corresponding drop was 50 percent.

With AU, the conditional values are very close to the known parameters case. The slight differences arise as beliefs about the transition probabilities are updated in the ten years before the depression event and upon the event, but this effect on the conditional moments is minor relative to the effect of rationally priced parameter uncertainty.

Overall, the rare events model calibrated to US consumption dynamics and the Great Depression with realistic levels of parameter uncertainty performs well along a number of dimensions when compared to observed data and to the same model with known parameters. In particular, it predicts a high unconditional risk premium and a low risk-free rate, along with a very high risk premium and return volatility in the crisis period, and only requires a relative risk aversion coefficient of 4. Relative to the benchmark case with known transition probabilities, parameter uncertainty increases the Sharpe ratio and the risk premium by factors of about 2.5 and 5, respectively, underscoring the importance of accounting for certain types of parameter uncertainty, namely the persistence of rare, bad states, when relating macroeconomic risks to asset prices.

TABLE 5—LEARNING ABOUT THE PROBABILITY AND PERSISTENCE OF A GREAT DEPRESSION: CONDITIONAL MOMENTS

Moment	Model		
	Unknown $\pi's$	Known $\pi's$	Anticipated utility
$\frac{P/D_{\text{Beginning of depression}}}{P/D_{\text{Before depression}}}$	0.45	0.50	0.47
$E_t\left[r_{m,t+1} - r_{f,t+1} \mid \text{Before depression}\right]$	4.6%	0.50%	0.46%
$\sigma_t\left[r_{m,t+1} - r_{f,t+1} \mid \text{Before depression}\right]$	15%	15%	15%
$SR_t\left[R_{m,t+1} - R_{f,t+1} \mid \text{Before depression}\right]$	0.32	0.11	0.11
$E_t\left[r_{m,t+1} - r_{f,t+1} \mid \text{Beginning of depression}\right]$	22%	10%	9.0%
$\sigma_t\left[r_{m,t+1} - r_{f,t+1} \mid \text{Beginning of depression}\right]$	81%	38%	36%
$SR_t\left[R_{m,t+1} - R_{f,t+1} \mid \text{Beginning of depression}\right]$	0.66	0.45	0.43

Notes: This table gives conditional, annualized expected log excess returns and return volatility, as well as the annualized, conditional Sharpe ratio of simple returns, on the dividend claim for the two-state switching regime model. The moments are related to the simulated four-year Depression path of Figure 2. That is, they are based on a 200-year prior, where the Depression starts in quarter 41 of the sample. The *Before depression* numbers correspond to conditional moments in quarter 40, while the *Beginning of depression* numbers correspond to conditional moments in quarter 41. P/D refers to the level of the price-dividend ratio. The next-to-last column gives the corresponding numbers for the benchmark known-parameters case, while the last column gives the corresponding numbers using anticipated utility pricing.

III. Parameter Uncertainty in Long-Run Risk Models

The previous section discussed the implications of parameter uncertainty, when we can identify ex ante the few parameters that are most relevant. Here, we assume agents are uncertain about the model specification, assuming the parameters within the model are known. This "model" risk is just another form of parameter uncertainty, as mentioned above and shown below. We find that realistic model learning can lead to substantial, endogenously time-varying model risk as the likelihood of the different models shifts over time. This in turn generates higher and more volatile risk premia.

We focus on the case where agents are uncertain about the presence of long-run risks and thus compare an i.i.d. model to a long-run risk (LRR) model as in Bansal and Yaron (2004).[31] These two model alternatives are statistically hard to distinguish, but entail very different levels of risk and therefore of utility for the agent. The large difference in utility outcomes across the two models gives rise to priced model risk as the agent updates beliefs about which model is true.

In this case, consumption growth can be expressed as

$$\Delta c_{t+1} = (1 - M)\{\mu + \sigma\varepsilon_{t+1}\} + M\{\mu + x_t + \sigma\varepsilon_{t+1}\}, \tag{16}$$

[31] We consider the homoskedastic Case I model in Bansal and Yaron (2004). A similar problem is considered by Hansen and Sargent (2010), though their alternative model is not the i.i.d. case, but a case where there is still positive, but less autocorrelation in consumption growth than in the long-run risk model. Also, our focus is on the quantitative implications of rational learning for long-horizon claims when the agent has Epstein-Zin preferences, whereas Hansen and Sargent use "fragile beliefs" preferences for robustness.

where $x_{t+1} = \rho x_t + \varphi\sigma\varepsilon_{t+1}$ and $\varepsilon_{t+1} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. M is a parameter that takes the value zero if the i.i.d. model is true and one if the LRR model is true. This "nesting" shows how model uncertainty can be viewed as a form of parameter uncertainty. Unlike the earlier cases where parameters were continuously distributed, here M takes two values. As before, we let $\mu = 0.45$ percent, $\sigma = 1.35$ percent. In addition, we follow BY and set $\rho = (0.979)^3$ and $\varphi = 0.044\sqrt{1+\rho^{2/3}+\rho^{4/3}}$, accounting for the fact that our calibration is quarterly while BY's is monthly.

The agent does not know M, but instead learns from realized consumption data using Bayes' rule. The agent's current expected value of M equals the subjective probability that the i.i.d. model is true: $p_t = E_t[M]$. The agent is assumed to observe x_t. If the i.i.d. model is true, x_t is unrelated to consumption growth. If, on the other hand, the LRR model is true, x_t represents a small, persistent component in consumption growth.[32]

A. *Model Risk Intuition*

Valuable intuition about how model risk affects asset prices and risk premia can be obtained by following an approach similar to the one used to identify the "most relevant" parameters in Section IIB. Specifically, consider the case where the true model will be revealed in one period:

$$V_t^{One} = \left\{(1-\beta)\, C_t^{1-1/\psi} + \beta\, \mu_{One,t}^{1-1/\psi}\right\}^{1/(1-1/\psi)}, \tag{17}$$

where

$$\mu_{One,t} = \left\{p_t E_t\left[\left(V_{t+1}^{LRR}\right)^{1-\gamma}\right] + (1-p_t)\, E_t\left[\left(V_{t+1}^{iid}\right)^{1-\gamma}\right]\right\}^{1/(1-\gamma)}, \tag{18}$$

and V_{t+1}^{LRR} is the time $t+1$ value function for the LRR model, while V_{t+1}^{iid} is the $t+1$ value function for the case of i.i.d. consumption growth. Two things are immediately clear.

First, the distance between the models' value functions will be important for how model uncertainty impacts utility (a bigger difference all else equal lowers current utility as the certainty equivalent penalizes volatility). With $x_t = 0$, the utility in the LRR model is lower than the utility in the i.i.d. model (because of aversion to LRR). Further, a low value of x_t makes this difference even larger, while a high x_t makes it smaller. Thus, model risk is time-varying as a function of the level of x_t. Second, for low values of p_t the distribution of next period's value function becomes more negatively skewed, even though the expected level of next period's utility increases. This increased skewness is penalized in the certainty equivalent.

To illustrate these effects, Figure 3 shows the parameter uncertainty premium for different values of p_t, x_t, and γ. The AU case is calculated assuming $p_t = p^{AU}$ will be constant forever, which implies the consumption dynamics

[32] The dynamics of x_t can be thought of as the subjective dynamics that arises from the filtering of the observed consumption dynamics for some latent $\hat{x}_t$ that also follows an AR(1) process.

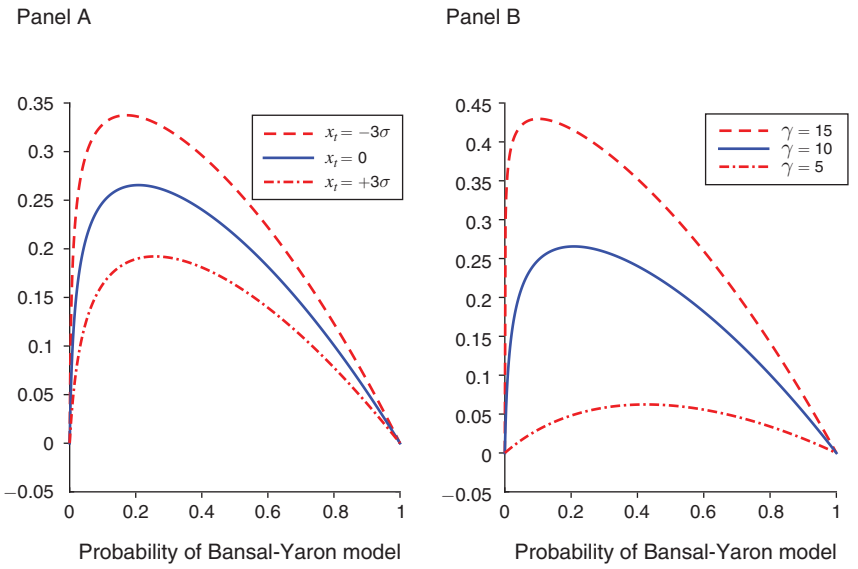

FIGURE 3. THE COST OF MODEL UNCERTAINTY: APPLYING THE METRIC

Notes: The figure shows the parameter uncertainty premium, as defined in the main text, when the agent is unsure whether consumption growth is i.i.d. or contains a small persistent component as in Bansal and Yaron (2004). The state variables are the current belief about the model p_t, where $p_t = 1$ means the agent is certain the BY model is true, and x_t, the current belief about expected consumption growth, conditional on the BY model being the correct model. In all cases, $\psi = 2$, $\beta = 0.994$. In panel A, $\gamma = 10$, while in panel B, $x_t = 0$.

$$\Delta c_{t+1}^{AU} = \mu + p^{AU} x_t + \sigma \varepsilon_{t+1}, \tag{19}$$

$$x_{t+1} = \rho x_t + \varphi \sigma \varepsilon_{t+1}.$$

Thus, for the AU case, a low value of p^{AU} implies there is little long-run risk in the consumption dynamics. Figure 3 shows that the parameter uncertainty premium can be substantial, more than 30 percent when $\gamma = 10$ and x_t is low. Interestingly, the premium peaks at $p^{AU} \approx 0.4$ when $\gamma = 5$, 0.2 when $\gamma = 10$, and 0.1 when $\gamma = 15$, even though the variance of outcomes is maximized at $p_t = 0.5$. This is due to the impact of negative skewness on the certainty equivalent, which is stronger with higher risk aversion.

Intuitively, with a low current p_t, the agent experiences *model disaster risk:* there is a small probability that a low utility model is true and so relative to the current utility level, which is high since p_t is low, there is a large potential utility drop. Figure 3 also shows that variation in the current level of x_t can lead to large variation in the parameter uncertainty premium. Since x_t will be low in recessions and high in expansions, this channel gives rise to endogenously countercyclical model uncertainty risk.

B. *Long Sample Moments*

We now analyze model learning over the full sample, with uncertainty resolving slowly. Learning is slow because it is hard to distinguish the two models empirically. Therefore, the asset pricing implications of model learning are long-lived. In particular, even if the i.i.d. model is true, asset prices reflect a substantial time-varying amount of long-run risks even after several centuries of learning. Thus, even if the probability of a bad model alternative is low, it has large, time-varying effects on the risk premium, return volatility, and Sharpe ratio.

Initial priors are given by $p_0 = \mathbb{P}(M = 0)$, and the agent updates beliefs via Bayes' rule upon observing consumption realizations:

$$p_{t+1} = \mathbb{P}\left(M = 0 \,|\, y^{t+1}\right) \propto p\left(y_{t+1} \,|\, y^t, M = 0\right) p_t. \tag{20}$$

The posterior probabilities are martingales. Letting $p(y_{t+1} \,|\, y^t, M = 0) = p_{BY}\left(y_{t+1} \,|\, y^t\right)$ and $p\left(y_{t+1} \,|\, y^t, M = 1\right) = p_{iid}\left(y_{t+1}\right)$, the belief recursion is

$$p_{t+1} = \frac{p_{BY}\left(y_{t+1} \,|\, y^t\right) p_t}{p_{BY}\left(y_{t+1} \,|\, y^t\right) p_t + p_{iid}\left(y_{t+1}\right)\left(1 - p_t\right)}, \tag{21}$$

where $p_{BY}\left(y_{t+1} \,|\, y^t\right) \sim \mathcal{N}\left(\mu + x_t, \sigma^2\right)$ and $p_{iid}\left(y_{t+1}\right) \sim \mathcal{N}\left(\mu, \sigma^2\right)$. The value function normalized by consumption is a function of p_t and x_t and is computed numerically using value function iteration, with boundary values given by the cases $p_t = 0$ and $p_t = 1$. See the online Appendix for a detailed description of the numerical solution methodology.

Table 6 shows average sample moments for different calibrations corresponding to combinations of the EIS and initial model probability: $\psi \in \{1.5, 2\}$ and $p_0 \in \{0.194,\ 0.0894,\ 0.0432\}$. We assume consumption growth is truly i.i.d., and the priors correspond to 100-, 200-, and 300-year training samples starting from an uninformative prior ($\mathbb{P}[M = 0] = 0.5$).[33] Since the true model used in these simulations is i.i.d., the ∞-year prior correspond to $p_0 = 0$. The data column contains the same empirical asset pricing moments described in the previous section.

If agents knew the true (i.i.d.) consumption growth dynamics, the risk premium would be 0.72 percent, the return Sharpe ratio would be 0.12, and the price of risk (maximal attainable Sharpe ratio) would be 0.28 for $\psi = 2$. Instead, with model uncertainty, the risk premium is between 4.3 percent and 4.7 percent, depending on the prior, the Sharpe ratio is between 0.33 and 0.35, and the price of risk is between 0.55 and 0.56. Thus, even with low prior probabilities of the LRR model being true, there is a six-fold increase in the risk premium, three-fold increase in the Sharpe ratio, and a two-fold increase in the price of risk. These moments are quite close to values when the agents actually believe the LRR model to be true ($p_0 = 1$)! Indeed, in that case, the risk premium is 5 percent, Sharpe ratio is 0.38, and the price of risk is 0.64. The case with $\psi = 1.5$ shown in panel B is quite similar, though there is a very slight decrease in risk premium and Sharpe ratio.

[33] The ensuing priors, p_0, are calculated by simulating 20,000 economies for 300 years and taking the average model belief across samples after 100, 200, and 300 years.

TABLE 6—SAMPLE MOMENTS MODEL UNCERTAINTY: LONG-RUN RISK VERSUS I.I.D.

		Priors (training sample)					
	Data	100 years ($p_0 = 0.194$)	200 years ($p_0 = 0.0894$)	300 years ($p_0 = 0.0432$)	∞ years (i.i.d.) ($p_0 = 0$)	Objective LRR model ($p_0 = 1$)	Anticipated utility ($p_0 = 0.0894$)
Panel A. $\psi = 2$							
$E_T[r_m - r_f]$ (percent)	5.0	4.66	4.51	4.33	0.72	5.01	1.53
$\sigma_T[r_m - r_f]$ (percent)	20.2	17.8	17.8	17.9	14.9	17.9	15.2
$SR_T[R_M - R_f]$	0.35	0.35	0.34	0.33	0.12	0.37	0.18
$E_T[r_f]$ (percent)	0.7	1.82	1.86	1.91	2.78	1.81	2.17
$\sigma_T[M]/E_T[M]$	n/a	0.55	0.54	0.52	0.27	0.56	0.32
Panel B. $\psi = 1.5$							
$E_T[r_m - r_f]$ (percent)	5.0	4.44	4.28	4.09	0.72	4.75	1.38
$\sigma_T[r_m - r_f]$ (percent)	20.2	17.6	17.7	17.9	14.9	17.6	15.2
$SR_T[R_M - R_f]$	0.35	0.34	0.33	0.32	0.12	0.36	0.17
$E_T[r_f]$ (percent)	0.7	2.19	2.23	2.28	3.03	2.14	2.54
$\sigma_T[M]/E_T[M]$	n/a	0.54	0.52	0.51	0.27	0.56	0.31

Notes: This table gives average sample moments from 20,000 simulations of 82-year samples, corresponding to the sample length in the data. The data column is based on US data from 1929 to 2011. The remaining columns show variations of the model where agents are unsure whether true consumption growth is i.i.d. or contains a small persistent component as in Bansal and Yaron (2004). The priors are named by their implicit length of training sample and the prior probability of the long-run risk (LRR) model is given in parentheses. The rightmost column shows the case of anticipated utility for the 200-year prior. Panel A shows results when $\psi = 2, \gamma = 10$, and $\beta = 0.994$, while panel B shows results when $\psi = 1.5, \gamma = 10$, and $\beta = 0.994$. $E_T[x]$ 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. Lowercase letters denote log of uppercase counterparts. The subscript m refers to the dividend claim (the "market" portfolio), while the subscript f refers to the real risk-free rate. The parameters governing the dividend dynamics are $\lambda = 2.5$ and $\sigma_d = 0.0665$. All statistics are annualized.

Thus, the unconditional asset pricing moments reflect the worst-case model (even if unlikely), due to the priced model uncertainty. This is consistent with the "model disaster risk" intuition given in the previous section. When the probability of the worst-case model is low, the current utility level is relatively high and thus the worst-case model constitutes an unlikely, but large negative drop in utility. This negative skewness in model outcomes generates substantial extra risk for the EZ agent.

For comparison, the AU case is also shown for the case of the 200-year prior. In this case, where model uncertainty is not a priced risk, asset pricing moments are instead close to those of the true, i.i.d. consumption growth model, reflecting the low probability of the BY model. For instance, the Sharpe ratio of equities is 0.18 with AU (which prices in some long-run risk since the LRR model is given a small probability), while the corresponding Sharpe ratio under the i.i.d. model is 0.12.

C. *Conditional Dynamics*

This section shows that model uncertainty yields rich conditional dynamics in risk premia and Sharpe ratios, even though the consumption dynamics in the two reference models are homoskedastic. For brevity, we focus on the case where $\psi = 2$ and where the agent has a 200-year prior. Figure 4 plots the dynamics of the

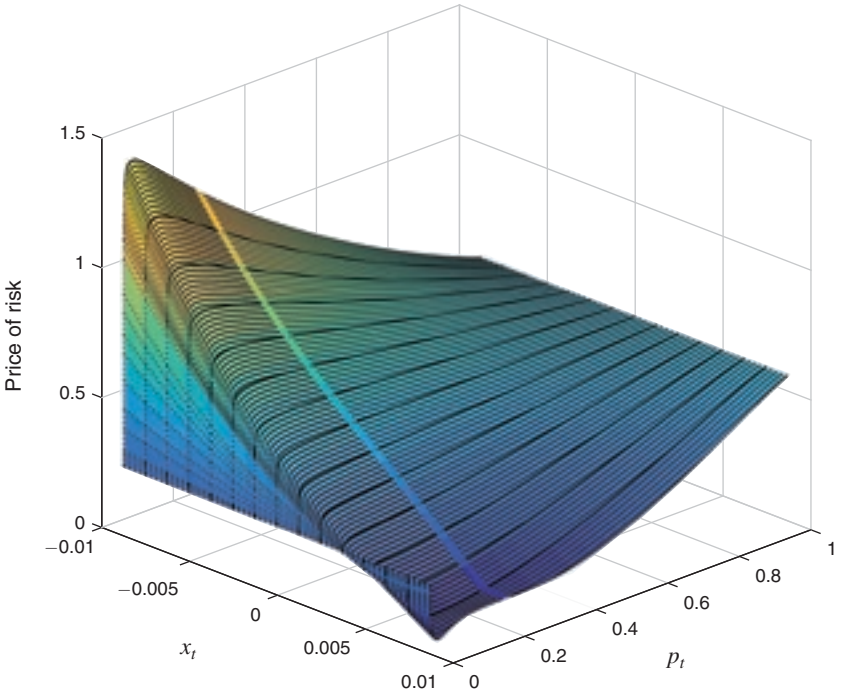

FIGURE 4. PRICE OF RISK FOR CASE OF MODEL UNCERTAINTY

Notes: The figure shows the annualized, conditional price of risk in the economy where the agent is unsure whether true consumption growth is i.i.d. or contains a small persistent component as in Bansal and Yaron (2004). The state variables are the current belief about the model p_t, where $p_t = 1$ means the agent is certain the BY model is true, and x_t, the current belief about expected consumption growth, conditional on the BY model being the correct model. Preference parameters are $\gamma = 10$, $\psi = 2$, and $\beta = 0.994$.

conditional price of risk $\left(\sigma\left(M_{t+1}\,|\,p_t, x_t\right)/E\left(M_{t+1}\,|\,p_t, x_t\right)\right)$ in this economy against the state variables p_t and x_t.

When $p_t = 1$, the agent is certain the BY economy is true, in which case, the annualized price of risk is constant and equal to 0.51.[34] In the i.i.d. case, $p_t = 0$, the price of risk is 0.27. For $p_t \in (0, 1)$ the price of risk is different than a simple weighted average of the two boundary case economies: a crucial feature of priced model uncertainty. In particular, at $x_t = 0$, the price of risk remains close to that in the worst-case LRR economy even for values of p_t close to zero, reflecting the model-disaster risk feature discussed above.

Unconditionally, the LRR model generates lower utility than the i.i.d. model, and the differences increase when $x_t < 0$, as future expected consumption growth rates are lower than in the i.i.d. model. Thus, model uncertainty is "worse" in these states of the world, generating a higher conditional price of risk. Furthermore, the two risks in the economy, the shock to consumption and the model probability update,

[34] In the exactly solved LRR model, the price of risk actually varies a tiny amount with x_t, but to the third decimal it is constant, as in the approximate solution for the homoskedastic case given in Bansal and Yaron (2004).

are reinforcing in these states. A low consumption growth value is bad in the i.i.d. model. However, when $x_t < 0$, a low consumption growth realization also increases the likelihood of the LRR model. In fact, Figure 4 shows that when $p_t = 0.05$ and x_t is three standard deviations below its mean, the price of risk is about 1.4, more than twice that of the riskiest alternative model of the world. On the other hand, when $x_t > 0$ the updates in beliefs hedge consumption shocks: a low consumption realization is bad (which is also the case in the i.i.d. model), but since $x_t > 0$, the low consumption growth increases the likelihood that the i.i.d. model is true, which is good. Therefore, the total price of risk in these states can fall below either of the limiting economies. In fact, when $p_t = 0.05$ and when x_t is three standard deviations above its mean, the annualized price of risk is only 0.07. Since x_t tends to be low/high in recessions/expansions, model uncertainty generates countercyclical risk prices.

Figure 4 also shows the tenuous nature of the full-information rational expectations assumption (see also Hansen 2007). It is not until p_t gets lower than 0.01 percent that the asset pricing implications of model uncertainty become negligible. It would take agents on average about 800 years to reach such a low model-belief starting from the initial prior $p_0 = 0.5$. In sum, one cannot outright dismiss a model as unimportant even though it is rejected by the data at conventional significance levels.[35]

D. *Feeding the Model Actual Consumption Data*

Finally, we consider the impact of model uncertainty on the postwar sample using the corresponding time series of US quarterly, real, per capita consumption growth. To be consistent with the model, we first remove autocorrelation of 0.25 induced by time-averaging of the data (see Working 1960) and then normalize the sample mean and variance of this modified consumption growth series to have mean and variance as assumed in the model calibration.[36]

The solid line in the top graph in Figure 5 shows the posterior probability of the LRR model, $\mathbb{P}(M = 0 | y^t)$, from 1947:III to 2010:IV, starting with the 200-year prior. The model probabilities vary substantially over the sample, from less than 0.05 to about 0.5, but there is not a clear time trend indicating that it is very hard to distinguish between the two models. Periods of either high consumption growth (late 1960s) or low consumption growth (the Great Recession) increase the probability of the LRR model relative to the i.i.d. model. At the end of the sample, the likelihood of the LRR model is 0.5 and at its maximum.

Panel B of Figure 5 shows the conditional price of risk, which varies substantially and is typically countercyclical in that it tends to increase in recessions (the bars in

[35] Of course, this conclusion depends on the agent having a preference for the timing of the resolution of uncertainty. With power utility preferences, the price of risk would be constant at $\gamma\sigma$. Basically, it is $\gamma - 1/\psi$ that matters for the pricing of shocks to the continuation utility. With $\gamma = 10$ this magnitude is 9.5 if $\psi = 2$, but only falls to 8 if $\psi = 0.5$. Thus, the main implications shown here are robust to the level of the EIS, as long as ψ is not very close to $1/\gamma$.

[36] We first construct $y_t = \Delta c_t - 0.25 \times \Delta c_{t-1}$, using actual real per capita quarterly consumption growth data from 1947:II to 2010:IV. The modified consumption growth series is then constructed as $\widetilde{\Delta c}_t \equiv \mu + \sigma_{iid} \times \frac{y_t - E_T[y_t]}{\sigma_T[y_t]}$, where $E_T[\cdot]$ and $\sigma_T[\cdot]$ denote the sample mean and variance, respectively.

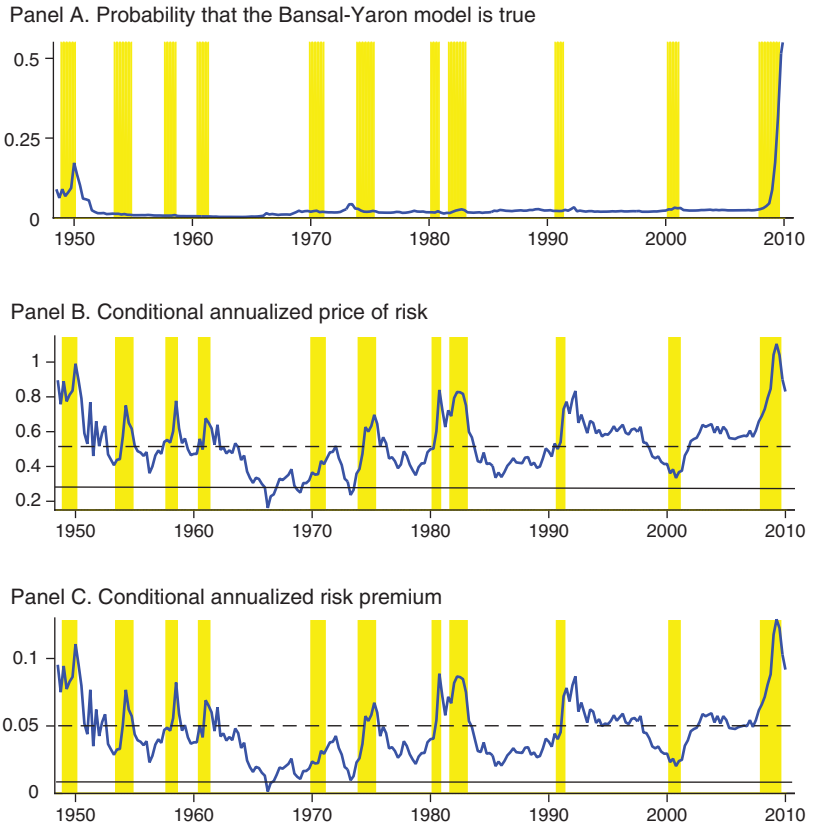

Figure 5. Model Uncertainty: Post-WW2 Sample Conditional Moments

Notes: The figure shows sample paths of the model probability (p_t), the annualized conditional price of risk, and the annualized conditional risk premium for the case of model uncertainty, where the agent is unsure whether true consumption growth is i.i.d. or contains a small persistent component as in Bansal and Yaron (2004). The shocks are taken from the post-WWII real per capita consumption growth as explained in the main text. The solid line corresponds to the case where the initial subjective probability of the BY model being true is set to 0.0894, corresponding to the 200-year prior. The dashed line in panels B and C corresponds to the conditional price of risk and risk premium, respectively, in the BY model. The bars correspond to NBER recessions.

the figure denote NBER recessions). This need not always be the case however. For instance, in the expansion of the late 1960s the price of risk *increases* as the LRR model becomes more likely. Through the recession of 2001, on the other hand, the price of risk decreases as the LRR model becomes more likely. This is due to high values of x_t perceived at the time based on the high growth in the 1990s. As can be seen from Figure 4, the high current x_t makes the prospect of facing BY model's consumption dynamics a conditionally less risky prospect as the agent then can enjoy higher expected consumption growth than in the i.i.d. case. The price of risk in this sample reaches its maximum of roughly 0.85 during the Great Recession, and its lowest point close to 0.3 in the mid 1960s. The dashed line denotes the constant price of risk of 0.64 in the benchmark LRR economy. Panel C of Figure 5 shows that the conditional risk premium largely inherits the dynamics of the price of risk. The

conditional, annualized risk premium varies substantially throughout the sample, from about 3.5 percent to 8 percent. Again, the dashed line shows the constant risk premium of 5 percent in the LRR economy.

Thus, model uncertainty induces large time-varying risk premiums and Sharpe ratios, often increasing these quantities to levels above what would prevail in the worst-case model since model uncertainty itself is an additional priced risk. The endogenously time-varying distance between the models, due to the conditional mean consumption growth rate in the LRR economy, x_t, leads to time-varying Sharpe ratios and risk premia, despite the fact that fundamentals (consumption and dividend growth) are homoskedastic.

IV. Conclusion

Parameter uncertainty generates endogenous long-run risks, because updating beliefs about fixed parameters causes permanent shocks in the posterior means. Since agents with preferences for early resolution of uncertainty care about these long-run risks, parameter learning is a priced risk.

We first show this analytically in a simple i.i.d. model and then add parameter uncertainty to two popular asset pricing models (the Rietz-Barro disaster model and the Bansal-Yaron long-run risk model), which are natural candidates for parameter uncertainty. We find that parameter uncertainty can quantitatively dominate fundamental sources of risks in these models and generate realistic asset pricing moments both unconditionally and conditionally. In particular, countercyclical variation in equity Sharpe ratios and risk premia can arise even from homoskedastic fundamentals due to endogenously countercyclical fluctuations in the amount of risk arising from parameter uncertainty. We also show that our learning channel is different from that identified by previous papers on learning, such as Weitzman (2007) or Cogley and Sargent (2008), for example.

REFERENCES

Ai, Hengjie. 2010. "Information Quality and Long-Run Risk: Asset Pricing Implications." *Journal of Finance* 65 (4) 1333–67.

Attanasio, Orazio P., and Guglielmo Weber. 1989. "Intertemporal Substitution, Risk Aversion and the Euler Equation for Consumption." *Economic Journal* 99 (395): 59–73.

Backus, David, Mikhail Chernov, and Stanley Zin. 2014. "Sources of Entropy in Representative Agent Models." *Journal of Finance* 69 (1): 51–99.

Bakshi, Gurdip, and Georgios Skoulakis. 2010. "Do Subjective Expectations Explain Asset Pricing Puzzles?" *Journal of Financial Economics* 98 (3): 462–77.

Bansal, Ravi, Dana Kiku, Ivan Shaliastovich, and Amir Yaron. 2014. "Volatility, the Macroeconomy and Asset Prices." *Journal of Finance* 69 (6): 2471–2511

Bansal, Ravi, Dana Kiku, and Amir Yaron. 2013. "Risks for the Long Run: Estimation with Time Aggregation." Unpublished.

Bansal, Ravi, and Ivan Shaliastovich. 2010. "Confidence Risks and Asset Prices." *American Economic Review* 100 (2): 537–41.

Bansal, Ravi, and Ivan Shaliastovich. 2011. "Learning and Asset-Price Jumps." *Review of Financial Studies* 24 (8): 2738–80.

Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance* 59 (4): 1481–1509.

Barro, Robert J. 2006. "Rare Disasters and Asset Markets in the Twentieth Century." *Quarterly Journal of Economics* 121 (3): 823–66.

Barro, Robert J. 2009. "Rare Disasters, Asset Prices, and Welfare Costs." *American Economic Review* 99 (1): 243–64.

Beaudry, Paul, and Eric van Wincoop. 1996. "The Intertemporal Elasticity of Substitution: An Exploration Using a US Panel of State Data." *Economica* 63 (251): 495–512.

Beeler, Jason, and John Y. Campbell. 2012. "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment." *Critical Finance Review* 1 (1): 141–82.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein. 2011. "Explaining Asset Pricing Puzzles Associated with the 1987 Market Crash." *Journal of Financial Economics* 101 (3): 552–73.

Boguth, Oliver, and Lars-Alexander Kuehn. 2013. "Consumption Volatility Risk." *Journal of Finance* 68 (6): 2589–2615.

Brandt, Michael W., Qi Zeng, and Lu Zhang. 2004. "Equilibrium Stock Return Dynamics under Alternative Rules of Learning about Hidden States." *Journal of Economic Dynamics and Control* 28 (10): 1925–54.

Campbell, John. 1999. "Asset Prices, Consumption, and the Business Cycle." In *Handbook of Macroeconomics*, Vol. 1, edited by John B. Taylor and Michael Woodford, 1231–1303. Amsterdam: North-Holland.

Campbell, John Y., and John H. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy* 107 (2): 205–51.

Cecchetti, Stephen G., Pok-sang Lam, and Nelson C. Mark. 2000. "Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True?" *American Economic Review* 90 (4): 787–805.

Chen, Hui, Winston Wei Dou, and Leonid Kogan. 2013. "Measuring the 'Dark Matter' in Asset Pricing Models." Unpublished.

Cochrane, John H. 2011. "Discount Rates." *Journal of Finance* 66 (4): 1047–1108.

Cochrane, John H., and Lars Peter Hansen. 1992. "Asset Pricing Explorations for Macroeconomics." *NBER Macroeconomics Annual 1992, Vol.* 7, edited by Olivier Jean Blanchard and Stanley Fischer, 115–82. Cambridge, MA: MIT Press.

Cogley, Timothy, and Thomas J. Sargent. 2008. "Anticipated Utility and Rational Expectations as Approximations of Bayesian Decision Making." *International Economic Review* 49 (1): 185–221.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer. 2016. "Parameter Learning in General Equilibrium: The Asset Pricing Implications: Dataset." *American Economic Review.* http://dx.doi.org/10.1257/aer.20130392.

Croce, Mariano M., Martin Lettau, and Sydney C. Ludvigson. 2014. "Investor Information, Long-Run Risk, and the Term Structure of Equity." *Review of Financial Studies* 28 (3): 706–42.

David, Alexander. 1997. "Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility." *Journal of Financial and Quantitative Analysis* 32 (4): 427–62.

David, Alexander, and Pietro Veronesi. 2013. "What Ties Return Volatilities Price Valuations and Fundamentals?" *Journal of Political Economy* 121 (4): 682–746.

Detemple, Jérôme. 1986. "Asset Pricing in a Production Economy with Incomplete Information." *Journal of Finance* 41 (2): 383–91.

Doob, Joseph L. 1949. "Application of the Theory of Martingales." Colloque International du CNRS Paris 23: 23–27.

Dothan, Michael U., and David Feldman. 1986. "Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy." *Journal of Finance* 41 (2): 369–82.

Epstein, Larry G., Emmanuel Farhi, and Tomasz Strzalecki. 2014. "How Much Would You Pay to Resolve Long-Run Risk?" *American Economic Review* 104 (4): 2680–97.

Epstein, Larry G., and Martin Schneider. 2007. "Learning under Ambiguity." *Review of Economic Studies* 74 (4): 1275–1303.

Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57 (4): 937–69.

Fama, Eugene F., and Kenneth R. French. 2002. "The Equity Premium." *Journal of Finance* 57 (2): 637–59.

Gennotte, Gerard. 1986. "Optimal Portfolio Choice under Incomplete Information." *Journal of Finance* 41 (3): 733–49.

Geweke, John. 2001. "A Note on Some Limitations of CRRA Utility." *Economics Letters* 71 (3): 341–45.

Gourio, François. 2012. "Disaster Risk and Business Cycles." *American Economic Review* 102 (6): 2734–66.

Gruber, Jonathan. 2013. "A Tax-Based Estimate of the Elasticity of Intertemporal Substitution." *Quarterly Journal of Finance* 3 (1): 1–20.

Hall, Robert E. 1988. "Intertemporal Substitution in Consumption." *Journal of Political Economy* 96 (2): 339–57.

Hansen, Lars Peter. 2007. "Beliefs, Doubts, and Learning: Valuing Macroeconomic Risk." *American Economic Review* 97 (2): 1–30.

Hansen, Lars Peter, and Ravi Jagannathan. 1991. "Implications of Security Market Data for Models of Dynamic Economies." *Journal of Political Economy* 99 (2): 225–62.

Hansen, Lars Peter, and Thomas J. Sargent. 2010. "Fragile Beliefs and the Price of Uncertainty." *Quantitative Economics* 1 (1): 129–62.

Jobert, A., A. Platania, and L. C. G. Rogers. 2006. "A Bayesian Solution to the Equity Premium Puzzle." Unpublished.

Johannes, Michael, Lars A. Lochstoer, and Yiqun Mou. Forthcoming. "Learning about Consumption Dynamics." *Journal of Finance.*

Ju, Nengjiu, and Jianjun Miao. 2012. "Ambiguity, Learning, and Asset Returns." *Econometrica* 80 (2): 559–91.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. 2009. "Recursive Smooth Ambiguity Preferences." *Journal of Economic Theory* 144 (3): 930–76.

Kocherlakota, Narayana R. 1996. "The Equity Premium: It's Still a Puzzle." *Journal of Economic Literature* 34 (1): 42–71.

Kreps, David M. 1998. "Anticipated Utility and Dynamic Choice." In *Frontiers of Research in Economic Theory*, edited by Donald P. Jacobs, Ehud Kalai, and Morton I. Kamien, 242–74. New York: Cambridge University Press.

Kreps, David M., and Evan L. Porteus. 1978. "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica* 46 (1): 185–200.

Lettau, Martin, Sydney C. Ludvigson, and Jessica A. Wachter. 2008. "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?" *Review of Financial Studies* 21 (4): 1653–87.

Lucas, Robert E., Jr. 1978. "Asset Prices in an Exchange Economy." *Econometrica* 46 (6): 1429–45.

Lucas, Robert E., Jr. 1987. *Models of Business Cycles*. New York: Blackwell.

Lucas, Robert E., and Thomas J. Sargent. 1978. "After Keynesian Macroeconomics." In *After the Phillips Curve: Persistence of High Inflation and High Unemployment*, 49–72. Boston: Federal Reserve Bank of Boston.

Martin, Ian. 2015. "What is the Expected Return on the Market?" Unpublished.

Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics* 15 (2): 145–61.

Modigliani, Franco. 1977. "The Monetarist Controversy or, Should We Forsake Stabilization Policies?" *American Economic Review* 67 (2): 1–19.

Moore, Bartholomew, and Huntley Schaller. 1996. "Learning, Regime Switches, and Equilibrium Asset Pricing Dynamics." *Journal of Economic Dynamics and Control* 20 (6–7): 979–1006.

Nakamura, Emi, Jón Steinsson, Robert Barro, and José Ursúa. 2013. "Crises and Recoveries in an Empirical Model of Consumption Disasters." *American Economic Journal: Macroeconomics* 5 (): 35–74.

Pástor, Ľuboš, and Pietro Veronesi. 2003. "Stock Valuation and Learning about Profitability." *Journal of Finance* 58 (5): 1749–89.

Pástor, Ľuboš, and Pietro Veronesi. 2006. "Was There a NASDAQ Bubble in the Late 1990s?" *Journal of Financial Economics* 81 (1): 61–100.

Pástor, Ľuboš, and Pietro Veronesi. 2009. "Learning in Financial Markets." A*nnual Review of Financial Economics* 1: 361–81.

Pástor, Ľuboš, and Pietro Veronesi. 2012. "Uncertainty about Government Policy and Stock Prices." *Journal of Finance* 67 (4): 1219–64.

Piazzesi, Monika, and Martin Schneider. 2007. "Equilibrium Yield Curves." In *NBER Macroeconomics Annual 2006*, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, 389–472. Cambridge, MA: MIT Press.

Rietz, Thomas A. 1988. "The Equity Risk Premium: A Solution?" *Journal of Monetary Economics* 22 (1): 117–31.

Shaliastovich, Ivan. 2015. "Learning, Confidence and Option Prices." *Journal of Econometrics* 187 (1): 18–42.

Strzalecki, Tomasz. 2013. "Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion." *Econometrica* 81 (3): 1039–1074.

Veronesi, Pietro. 1999. "Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model." *Review of Financial Studies* 12 (5): 975–1007.

Veronesi, Pietro. 2000. "How Does Information Quality Affect Stock Returns?" *Journal of Finance* 55 (2): 807–37.

Veronesi, Pietro. 2004. "The Peso Problem Hypothesis and Stock Market Returns." *Journal of Economic Dynamics and Control* 28 (4): 707–25.

Vissing-Jørgensen, Annette. 2002. "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution." *Journal of Political Economy* 110 (4): 825–53.

Vissing-Jørgensen, Annette, and Orazio P. Attanasio. 2003. "Stock-Market Participation, Intertemporal Substitution, and Risk-Aversion." *American Economic Review* 93 (2): 383–91.

Weitzman, Martin L. 2007. "Subjective Expectations and Asset-Return Puzzles." *American Economic Review* 97 (4): 1102–30.

Working, Holbrook. 1960. "Note on the Correlation of First Differences of Averages in a Random Chain." *Econometrica* 28 (4): 916–18.