Parameter Learning in General Equilibrium: The Asset Pricing Implications

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

February 7, 2012

*We thank David Backus, Mikhail Chernov, Lars Hansen, Stavros Panageas, and Stanley Zin for helpful comments. All errors are our own. Contact info: Lars A. Lochstoer, 405B Uris Hall, Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027. E-mail: LL2609@columbia.edu

Parameter Learning in General Equilibrium: The Asset Pricing Implications

Abstract

This paper studies the implications of parameter learning in general equilibrium, consumption-based asset pricing models. Learning about the structural parameters that govern aggregate consumption dynamics introduces long-run risks in the subjective consumption dynamics, as posterior mean beliefs are martingales and shocks to mean beliefs are permanent. These permanent shocks have particularly strong asset pricing implications for a representative agent with Epstein-Zin preferences and a preference for early resolution of uncertainty. We show that even a simple economy where aggregate consumption growth is truly i.i.d., but where the representative agent learns in a Bayesian fashion about the mean growth rate, yields a high equity premium, excess volatility, a low risk-free rate, and excess return predictability. Casual intuition might suggest that the asset pricing implications of parameter learning in this case are highly transient, as rational agents learn quickly. We show that this intuition is incorrect when agents have a preference for early resolution of uncertainty: 100 year sample moments from a model with a reasonably calibrated prior belief yields an equity risk premium of 4.4% – more than two and a half times the equity premium in the known-parameter, benchmark case. We also consider models with unknown parameters governing rare events, as well as learning in models with structural breaks. Unlike existing long-run risk models, the price-dividend ratios in these learning models do not predict future consumption growth. Further, the Hall (1978) risk-free rate regressions run on simulated model data yield estimates of the elasticity of intertemporal substitution close to zero, as in the data, despite the fact that the calibrated models have an elasticity of intertemporal substitution well above one.

1 Introduction

Conventional wisdom and existing research suggests that learning about fixed but unknown 'structural' parameters has a minor impact, and because of this most of the literature focuses on learning about stationary latent state variables.[1] To see this, assume the logarithm of consumption growth is normally distributed, $\Delta \ln(C_t) = y_t \sim \mathcal{N}(\mu, \sigma^2)$, and that the 'structural' parameter μ is unknown. Agents update normally distributed initial beliefs, $\mu \sim \mathcal{N}(\mu_0, \sigma_0^2)$, using Bayes rule which implies that the posterior of μ is $p(\mu|y^t) \sim \mathcal{N}(\mu_t, \sigma_t^2)$, where μ_t and σ_t^2 are given by standard recursions and y^t is data up to time t. If the representative agent has power utility preferences, the 'equity' premium on a single-period consumption claim is $\gamma(\sigma^2 + \sigma_t^2)$. Since σ_t^2 decreases rapidly when agents update using Bayes rule, the effect of structural parameter uncertainty on the equity premium is generally small to begin with and then rapidly dies out.[2] Thus, parameter uncertainty seems to have a negligible effect when allowing for a minimal amount of learning.

In this paper, we show that this conventional wisdom does not hold generally when the representative agent has Kreps-Porteus preferences, which allow for a preference for the timing of the resolution of uncertainty. A key feature of parameter uncertainty and rational learning is that mean parameter beliefs, or posteriors, are martingales. To see this, note that $\mu_t = E(\theta|y^t)$, where θ is a fixed parameter, is trivially a martingale by the law of iterated expectations. This implies that shocks to beliefs are permanent, affecting the conditional distribution of consumption growth indefinitely into the future. Parameter uncertainty thus generates a particularly strong form of long-run consumption risk (see Bansal and Yaron (2004)). For agents who care about the timing of the resolution of uncertainty, assets whose payoffs are affected by unknown parameters may therefore be particularly risky. The goal of this paper is to quantify the asset pricing implications of structural parameter uncertainty.

To quantify the impact of parameter uncertainty when the temporal resolution of uncertainty matters, we consider first the simplest setting where aggregate log consumption growth is i.i.d. normal, but where the representative agent is unsure about the true mean growth rate. We consider cases with unbiased beliefs, where priors are centered at the true values, to focus particularly on the impact on asset prices of *priced* parameter uncertainty (unlike, e.g., the analysis in Sargent and Cogley (2008)). We are particularly interested in

[1]See Veronesi (1999), Brennan and Xia (2002), among others.

[2]Weitzman (2007) argues that uncertainty about consumption volatility can be large and economically important. Bakshi and Skouliakis (2010) show that this is due to the prior distribution on the variance, and that the impact of volatility uncertainty with a prior with bounded support is negligible.

studying the dynamics of central asset pricing quantities like the ex-ante equity premium, as well ex-post realized quantities like average returns and volatility. We consider the pricing of equity as a levered consumption claim and assume the representative agent has Epstein-Zin preferences with a preference for early resolution of uncertainty. Although this model is too simple along many dimensions to be considered realistic, the learning dynamics reveal a number of interesting findings.

Even in this simple model, parameter uncertainty has a quantitatively large and long-lasting impact on the equity premium. As a benchmark, the average excess return on a levered consumption claim in the known-parameter benchmark case is roughly 1.7% per year, whereas over a 100 year sample in a reasonably calibrated parameter learning case the average excess equity returns are 4.4%. The equity premium does decline over the sample – in the first 10 years it is about 11%, while after 50 years is about 4.5%. Even after 100 years, the equity premium is 3%. This magnitude may at first seem almost implausibly large to the reader, as the agent after 100 years of learning is quite confident in her mean belief about the consumption growth rate. However, it is a direct effect of the combination of permanent shocks to long-run growth expectations and the preference for early resolution of uncertainty. The representative agent experiences a large amount of risk even after 100 years of learning as the shocks to the expected consumption growth rate, while small when viewed over a quarter, last forever and therefore has a large impact on the continuation utility.

These results show that the asset pricing implications of rational parameter learning can be quantitatively significant for a very long time, despite the fact that the posterior standard deviation of the beliefs of the mean growth rate decline rapidly. In fact, after 50 years, the standard deviation of shocks to mean beliefs is 5.8 times smaller than at the beginning of the sample, but the equity premium drop by a factor of 2, only. The standard deviation of the log pricing kernel – the price of risk – drop by a factor slightly less than 2 over the same period. Over the next 50 years, the standard deviation of shocks to mean beliefs drops by a factor of 1.9, while the price of risk drops by a factor of about 1.2. Two observations can be made here. First, the standard deviation of shocks to mean beliefs about the mean growth rate declines much faster in the beginning of the sample than after some time has relapsed. This is a standard result from Bayesian updating. Second, the price of risk in the economy declines at a much slower rate. The latter seems puzzling, but is in fact an endogenous outcome of the deterministically decreasing variance of beliefs and can be understood as follows.

While the agent's mean parameter beliefs are a Martingale, the effect on the continuation utility is nonlinear. In particular, in the beginning of the sample, when there is a lot of parameter uncertainty, discount rates are high. Thus, shocks to the beliefs about the mean growth rate are relatively quickly discounted in terms of their effect on wealth (utility). Thus, the *duration* of total wealth is relatively low. Towards the end of the sample, discount rates are lower and so the effect of an update in the mean belief about the growth rate has a larger effect on wealth. In other words, the duration of total wealth is relatively high. Since shocks to wealth appear in the pricing kernel when agents have a preference for the resolution of uncertainty, this affects the volatility of the pricing kernel. In sum, while the magnitude of the shocks to mean parameter beliefs decreases rapidly in rational learning, the impact on the continuation utility for a given shock to these beliefs is increasing. The net effect is a relatively slow decline in the price of risk.

Finally, parameter learning induces excess return predictability. This occurs both because the conditional risk premium declines over time and because of a small-sample correlation between future returns and the price-dividend ratio. The rationale for the latter is described in Timmermann (1996) and Lewellen and Shanken (2002). Importantly, while the model with parameter learning feature subjective long-run risks, there is no consumption growth predictability in the model. In particular, the price-dividend ratio does not predict future consumption growth in population or in small samples. Further, the long-run risk models have also been critiqued on the grounds that they assume a high value of agents' elasticity of intertemporal substitution, typically well above one. This in turn implies that consumption growth should be very responsive to changes in the real risk-free rate. However, Hall (1978) and several authors after him have estimated the elasticity of substitution to be close to zero. We run the same regressions as in Hall (1978) on simulated data from our models and show that we can replicate these low estimates even though the representative agent in fact has a high elasticity of intertemporal substitution. Again, the reason is that the asset prices, and in this case the risk-free rate, respond to agents' perceived consumption growth rate and not to the ex post true growth rate. Thus, the parameter learning model is not subject to Beeler and Campbell's (2011) main critiques of long-run risk models.

We consider two other cases of parameter uncertainty. The first considers learning about parameters governing the consumption dynamics in disasters. Learning about rare events is slow as there is, by definition, few historical observations to learn from. The second extension is an economy with structural breaks. In particular, we assume there is a small probability each quarter that the mean growth rate of the economy is redrawn from a given distribution. Such structural breaks restart the parameter learning problem and makes parameter uncertainty a perpetual learning problem. The paper proceeds as follows. In section 2 we describe in general how parameter learning is a natural source of long-run consumption risks. In Section 3, we describe the simple model with unknown mean growth rate. Section 4 considers the case of learning about disasters. Section 5 considers an economy with structural breaks.

2 Parameter learning as a source of long-run risk

Rational learning about fixed quantities such as parameters or model probabilities are natural sources of long-run risks. This is easy to see from the law of iterated expectations. In particular, given a vector of unknown parameters θ, any well-behaved function $h(\theta)$ of these parameters is a Martingale under the agents' filtration:

$$E\left[E\left[h(\theta)\,|\mathcal{F}_{t+1}\right]|\mathcal{F}_t\right] = E\left[h(\theta)\,|\mathcal{F}_t\right]. \tag{1}$$

Thus, we can write:

$$E\left[h(\theta)\,|\mathcal{F}_{t+1}\right] = E\left[h(\theta)\,|\mathcal{F}_t\right] + \eta_{t+1}, \tag{2}$$

where $E\left[\eta_{t+1}|\mathcal{F}_t\right] = 0$ and $E\left[\eta_{t+1}E\left[h(\theta)\,|\mathcal{F}_t\right]|\mathcal{F}_t\right] = 0$. Importantly, the shocks to beliefs, η_{t+1}, are permanent.

In this paper, we consider economies where there is a representative agent that derives utility from consumption. Thus, learning affects marginal intertemporal rates of substitution when agents are learning about the parameters governing aggregate consumption dynamics. In particular, updates in beliefs about such parameters are a source of permanent shocks to the conditional distribution of future aggregate consumption. These long-run consumption risks are priced risk factors when the agent has Kreps-Porteus preferences and a preference for the timing of the resolution of uncertainty.

In the following, we examine the asset pricing implications of such learning. We start with the simplest possible model, where consumption growth is truly i.i.d., but where the mean growth rate is unknown. This case gives most of the intuition needed in a transparent way. Also, this simple model works remarkably well in terms of matching a number of stylized facts, as the asset pricing implications of learning even in this simple environment remain quantitatively large even after 100 years of learning. We then move on to more complicated consumption dynamics, including learning about rare events and learning in an economy with structural breaks.

Throughout the paper, the representative agent is assumed to have Epstein-Zin utility, V, over consumption, C:

$$V_t = \left\{ (1-\beta) C_t^{1-1/\psi} + \beta \left(E_t \left[V_{t+1}^{1-\gamma} \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}}, \tag{3}$$

where γ governs relative risk aversion, ψ is the elasticity of intertemporal substitution, and β is the time-discounting parameter. As is well-known, the stochastic discount factor in this economy can be written:

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\beta \frac{PC_{t+1}+1}{PC_t} \right)^{\theta-1}, \tag{4}$$

where PC_t is the wealth-consumption ratio at time t and where $\theta = \frac{1-\gamma}{1-1/\psi}$. The first component of the pricing kernel, $\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}$, is of the usual power utility form, where the second component, $\left(\beta \frac{PC_{t+1}+1}{PC_t} \right)^{\theta-1}$, comes into play if the agent has a preference for the timing of the resolution of uncertainty (i.e., if $\theta \neq 1$; see Epstein and Zin (1989)).

3 Case 1: i.i.d. consumption growth, unknown mean

We start by assuming that aggregate log consumption growth is i.i.d. normal:

$$\Delta c_{t+1} = \mu + \sigma \varepsilon_{t+1}, \tag{5}$$

where $\varepsilon_{t+1} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. The representative agent does not know the mean growth rate, but starts the sample with a prior: $\mu \sim N(\mu_0, \sigma_0^2)$. We have to later truncate this prior to ensure finite utility, but for now we consider the untruncated case for ease of exposition. The volatility parameter, σ, is assumed known.[3] The agent updates his beliefs sequentially upon

[3] Weitzman (2007) argues that learning about the variance of consumption growth has first-order implications for quantities such as the equity risk premium and Sharpe ratio. In a recent critique, however, Bakshi and Skoulakis (2010) show that this result derives from assuming the support for the variance parameter is the positive real line. If one truncates the prior distribution at extremely high levels (but less than infinity), the effect on both the risk premium and the volatility of the pricing kernel is very small.

observing realized consumption growth using Bayes rule:

$$\mu_{t+1} = \frac{\sigma_t^2}{\sigma_t^2+\sigma^2}\Delta c_{t+1} + \left(1-\frac{\sigma_t^2}{\sigma_t^2+\sigma^2}\right)\mu_t, \tag{6}$$

$$\frac{1}{\sigma_{t+1}^2} = \frac{1}{\sigma_t^2}+\frac{1}{\sigma^2}. \tag{7}$$

In the agent's filtration, the aggregate consumption dynamics are thus:

$$\Delta c_{t+1} = \mu_t + \sqrt{\sigma^2+\sigma_t^2}\,\tilde{\varepsilon}_{t+1}, \tag{8}$$

where $\tilde{\varepsilon}_{t+1} \sim \mathcal{N}(0,1)$. Further, note that:

$$\begin{aligned}\mu_{t+1} &= \frac{\sigma_t^2}{\sigma_t^2+\sigma^2}\Delta c_{t+1} + \left(1-\frac{\sigma_t^2}{\sigma_t^2+\sigma^2}\right)\mu_t \\ &= \frac{\sigma_t^2}{\sigma_t^2+\sigma^2}\left(\mu_t+\sqrt{\sigma^2+\sigma_t^2}\,\tilde{\varepsilon}_{t+1}\right) + \left(1-\frac{\sigma_t^2}{\sigma_t^2+\sigma^2}\right)\mu_t \\ &= \mu_t + \frac{\sigma_t^2}{\sqrt{\sigma^2+\sigma_t^2}}\tilde{\varepsilon}_{t+1}. \end{aligned} \tag{9}$$

In words, in the agent's filtration the mean expected consumption growth rate is time-varying with a unit root. Comparing this to the consumption dynamics in Bansal and Yaron (2004), we note that learning induces truly long-run risk in that shocks to expected consumption growth (in the agent's filtration) are permanent versus Bansal and Yaron's persistent, but still transitory, shocks. The process does not explode, however, as the posterior variance is declining over time and will eventually (at $t=\infty$) go to zero. Note also that actual consumption growth is not predictable given its i.i.d. nature (Eq. (1)). Thus, the long-run consumption risks in this model do not imply excess consumption growth predictability – a critique often levied against long-run risk models (see, e.g., Beeler and Campbell (2012)).

Learning also increases consumption growth volatility: from the agent's perspective, the consumption growth variance is $\sigma^2+\sigma_t^2$. Setting $\sigma_0^2=\sigma^2$ as an upper bound, learning can maximally double the subjective conditional consumption growth variance.[4] Further, the posterior standard deviation decreases quickly, as shown in Figure 1. After ten years of quarterly consumption observations, the agent perceives the standard deviation of consumption growth to be only 1.012 times greater than the objective consumption growth standard

[4]If you start with a diffuse prior ($\sigma_{-1}^2=\infty$), you will after having observed *one* consumption growth outcome have $\sigma_0^2=\sigma^2$.

deviation. This fact have lead researchers working with power utility preferences to conclude that learning about the mean growth rate is not an important consideration for asset pricing. In particular, with power utility preferences the conditional volatility of the log pricing kernel is $\gamma\sqrt{Var_t\left(\Delta c_{t+1}\right)}$, and so, after ten years, learning will increase the maximum Sharpe ratio by only a tiny fraction.

[FIGURE 1 ABOUT HERE]

However, with a preference for early resolution of uncertainty ($\gamma > 1/\psi$), the agent strongly dislikes shocks to *expected* consumption growth as in Bansal and Yaron (2004). In particular, Bansal and Yaron show that even with a very small persistent component in consumption growth, the volatility of the pricing kernel can increase significantly relative to the power utility case. We make use of the same observation here. While the posterior variance decreases quickly, it takes a long time to converge to zero (see Figure 1). Even though it is small, the shocks to expected consumption growth in the agent's filtration are permanent – when learning about a fixed quantity, the conditional expectation is a Martingale. The second component of the pricing kernel (see Eq. 3), $\left(\beta\frac{PC_{t+1}+1}{PC_t}\right)^{\theta-1}$, then adds volatility in the following way. An increase in expected mean consumption growth, which occurs upon a higher than expected consumption growth realization, increases the wealth-consumption ratio when $\psi > 1$. In our main calibrations, $\gamma > 1$ and $\psi > 1$, which implies that $\theta < 0$, such movements in PC_{t+1} increase the total volatility of the pricing kernel. Since the shocks to mean consumption growth are permanent, they have a large impact on the wealth-consumption ratio.

In sum, while it is true that the extra variance that learning adds to the representative agents' perception of consumption growth quickly becomes small in magnitude, the type of risk that parameter learning represents is especially disliked by the agent with a preference for early resolution of uncertainty. This is because updates in mean beliefs about parameters constitute permanent shocks, which have a large impact on the continuation utility (or, as in Equation (3), the wealth-consumption ratio). This, coupled with the fact that the posterior variance of beliefs converges very slowly once the posterior variance gets small relative to the actual variance, implies that the asset pricing implications of parameter learning can be large and last for a long time. Next, we gauge the quantitative implications of parameter uncertainty in this general equilibrium model.

The dividend claim

In all the models, we assumed that the market return is a levered consumption claim:

$$R_{M,t} = \left(1 + \frac{D}{E}\right) R_{C,t}, \tag{10}$$

where $R_{C,t+1} = \frac{C_{t+1}}{C_t}\frac{1+PC_{t+1}}{PC_t}$ is the return to the consumption claim. The aggregate debt-to-equity ratio (D/E) in the U.S. postwar data is about 0.5, so the return we report is 1.5 times the return to the consumption claim. We don't explicitly write the dynamics for dividends and add idiosyncratic risk, as in e.g. Bansal and Yaron (2004). The rationale for looking directly at a levered consumption claim is two-fold. First, the dynamics of the consumption claim are more directly related to the learning problems we consider. Any dynamics in the idiosyncratic component of dividends obfuscates this relation. Second, while it is straightforward to price a claim to an exogenous dividend stream in our setup, different but common assumptions regarding the dividend dynamics can give vastly different asset pricing results. For instance, if one as in Abel (1999) models dividends as simply $\lambda \Delta c_t$, with $\lambda = 3$, the dividend claim would be much more sensitive to fluctuating expectations of the long-run mean of the economy. If, instead, one models dividends as cointegrated with consumption, as in most DSGE models, this long-run sensitivity is the same as for the consumption claim. While it is important to understand the joint (long-run) behavior of dividends and consumption, this is not the focus of this paper. We simply note that our definition of market returns is conservative in terms of its exposure to long-run risks relative to the long-run risk model of Bansal and Yaron (2004). Further, since we assume no idiosyncratic risk, the volatility of the market return will be low. However, the risk premium of this claim, which derives from the covariance of returns with the pricing kernel, is a quantity we can reasonably compare to the average excess equity returns in the data.

3.1 Results

We calibrate the true consumption dynamics to match the mean and volatility of time-averaged annual U.S. log, per capita consumption growth, as reported in Bansal and Yaron (2004): $E_T[\Delta c] = 1.8\%$ and $\sigma_T(\Delta c) = 2.72\%$. This implies true (not time-averaged) quarterly mean and standard deviation of 0.45% and 1.65%, respectively. The models are calibrated at a quarterly frequency. For the cases with parameter uncertainty, the prior beliefs about μ are assumed to be distributed as a truncated normal. The truncation ensures that utility is finite. The lower bound is set at a -1.2% annualized growth rate, while the upper bound is set at a 4.8% annualized growth rate. The true growth rate we will use is 0.45% per quarter and the prior beliefs are assumed to be unbiased.[5] Our baseline model has $\beta = 0.994$, $\gamma = 10$ and $\psi = 2$.

3.1.1 The effect of parameter uncertainty over time

First, we show how parameter uncertainty affects asset pricing moments over time. Note that the updating equation for the variance of beliefs (see Equation 7) is deterministic, and so this exercise captures the non-stationary aspect of parameter learning. At this point, we do not calibrate the prior dispersion, but simply start with a maximum standard deviation of prior beliefs, σ_0, set to 1.65% – i.e., equal to σ. This is the same as assuming investors at the beginning of the sample has observed only one consumption growth realization with a completely diffuse earlier prior. The prior mean belief is set to the true value of mean quarterly consumption growth, 0.45%.

Table 1 shows the ensuing decade by decade asset pricing moments averaged across 20,000 simulated 100-year economies which all start from the same initial prior. The prior standard deviation at the beginning of each decade is given in the second column of the table, as implied by the deterministic updating equation given in Equation (7). For instance, after 10 years of learning, the prior standard deviation over the mean drops from 1.65% to 0.26%, after 50 years the standard deviation of beliefs is 0.12% and after 100 years it is 0.09%. Thus, while the standard deviation of beliefs decreases very quickly the first 10 years, the decrease is quite slow thereafter.

[TABLE 1 ABOUT HERE]

Column 3 gives the annualized conditional volatility of the log pricing kernel, $\sigma_t(m_{t+1})$, which is a measure of the maximal Sharpe ratio attainable in the economy. The conditional volatility of the log pricing kernel is on average 1.05 in the first decade, 0.87 in the second decade, 0.61 in the fifth decade, and 0.48 in the tenth decade. This is compared to the conditional volatility of the log pricing kernel in the benchmark economy with known parameters,

[5] Note that the updating equations for the mean and variance parameters for the prior are the same regardless of whether the distribution is truncated or not – the truncation only affects the limits of integration and not the functional form of the priors. Thus, we retain the conjugacy of the standard normal prior. We solve the models numerically, working backwards from the known-parameters boundary values on a grid for μ and time t (or, equivalently, a grid for the posterior standard deviation, σ_t; see Johnson (2007)).

which is only 0.33. Thus, after 50 years of learning, the volatility of the pricing kernel is twice as high as in the fixed parameters case, while after 100 years of learning it is one and a half times as high as in the fixed parameter benchmark case. Clearly, parameter uncertainty in this economy has long-lasting effects.

The slow decrease in the volatility of the pricing kernel is striking compared to the very fast decline that occurs in a power utility model. The reason the decrease is so slow is that the sensitivity of the continuation utility to shocks to growth expectations is endogenously increasing over time, offsetting the decline in the posterior variance. The intuition is straightforward: when the prior variance is high, discount rates are endogenously high and so the wealth-consumption ratio is less sensitive to shocks to growth rates. As parameter uncertainty decreases, discount rates decrease and get closer to the expected growth rate, and thus the sensitivity of the wealth-consumption ratio to shocks to the expected consumption growth rate is higher. We explain these general equilibrium dynamics in detail in Section 3.1.4.

Columns $4-7$ in Table 1 show the mean risk-free rate, the difference between the 10-year zero-coupon default-free real yield and the short-term risk-free rate, the average market excess return and volatility. Though the mean belief about the growth rate averaged across the $20,000$ samples is at its true value, the risk-free rate is increasing through time. This is due to a decrease in the pre-cautionary savings component as the amount of risk decreases deterministically as the agent beliefs about the mean growth rate become more precise. This upward drift in the risk-free rate is reflected in yield spreads, which are positive the first 50 years or so of learning and effectively zero thereafter. This is notably different from the standard long-run risk models, which have strongly negatively sloped real yield curves (see Beeler and Campbell, 2012).

The annualized market risk premium is 11% in the first decade, 4.5% in the fifth decade, and 3% in the tenth decade, compared with 1.7% in the known parameters benchmark economy. A similar decreasing pattern holds for the standard deviation of market returns. Even after 100 years of learning, the excess volatility is still a sizable 24% of fundamental volatility; 6.2% versus the benchmark economy's 5%.

3.1.2 Average moments over a long sample

Given that parameter uncertainty has a long-lasting impact on standard asset price moments, we next evaluate the asset pricing implications of the model, given a plausibly calibrated prior, for the standard long-sample asset price moments the literature typically considers. In particular, Table 2 shows 100-year standard sample moments averaged across the simulated economies, as well as the corresponding moments in the U.S. data taken from Bansal and Yaron (2004). We set the standard deviation of initial prior beliefs about the mean growth rate to 0.26%, which corresponds to a standard deviation of the annual growth rate of 1.04%. The Shiller data has real per capita consumption data available from 1889. The standard error of the estimated mean annual growth rate using this data up until a hundred years ago, in 1910, is in fact slightly higher at 1.12%. The prior mean beliefs are set equal to the true mean of consumption growth.

[TABLE 2 ABOUT HERE]

The third columns of Table 2 shows that the model with parameter uncertainty (*unknown* μ) yields a 100-year average excess annual market returns of 4.4%, compared to the 1.7% of the benchmark fixed parameter model (column 4; *known* μ). The risk premium in the data is higher still at 6.3% per year. The average annual volatility of the log pricing kernel in the learning model is 0.60, compared to 0.33 in the known parameters case. While the historical Sharpe ratio of equity returns is 0.33, the annual correlation between equity returns and consumption growth in the Shiller data is about 0.55 and so the pricing kernel need to have a volatility greater than or equal to 0.6 (=0.33/0.55) to match this value. Due mainly to no idiosyncratic component of dividends, the equity return volatility is too low in all the models relative to the data. The return volatility of the learning model is 7.35% versus the benchmark "fundamental" volatility of the known parameter case of 5%. Thus, the *excess volatility* (Shiller, 1980), measured as the ratio of standard deviation of returns in the learning case versus the standard deviation of returns in the no-learning case minus one, is 0.47 in the learning model. As reported by Bansal and Yaron (2004), the corresponding ratio of standard deviation of returns relative to the standard deviation of dividend growth minus one is 0.70. Thus, while the learning model does not generate quite as much excess volatility in relative terms, it goes a long way towards what is in the data. Due to the i.i.d. consumption growth assumption, the known parameter benchmark case features no excess volatility. Finally, the risk-free rate is low in the learning model and not too volatile, due to the high level of intertemporal elasticity of substitution, while the yield spread is on average slightly positive due to the on average upwards trend in real rates as agents become more sure of the mean growth rate. In sum, in terms of these unconditional sample moments, the simple learning model does quite well.

The two rightmost columns in Table 2 show the same moments for a model where the agent has power utility and thus is indifferent to the timing of the resolution of uncertainty. In this case, risk aversion is still 10, but the EIS is 0.1. The annual equity premium with no learning is 1.7%, but the equity premium with learning is -1.4%. This is due to the low EIS as an increase in investors perception of the expected growth rate in this case decreases the price-consumption ratio sufficiently to make stock returns negatively correlated with consumption growth (see Veronesi (2000)). Also, note that the learning does not increase the volatility of the log pricing kernel relative to the known parameters case, at least not to the second decimal, as expected. The indifference to the timing of the resolution of uncertainty means that the fact that shocks to growth expectations are permanent is immaterial for the conditional volatility of this investor's intertemporal marginal rate of substitution.

3.1.3 Predictability of returns, not consumption

The fixed parameter benchmark case features no predictability of excess returns or consumption growth by construction since consumption growth is assumed to be i.i.d. However, in the data excess equity market returns are predictable. A standard predictive variable is the price-dividend ratio. On the other hand, as emphasized by Beeler and Campbell (2012), aggregate consumption growth is not predicted by the price-dividend ratio in U.S. data. Further, Lettau and Ludvigson (2001) show that a measure of the wealth-consumption ratio also predicts excess returns but not long-horizon consumption growth. The latter point has been a bit of a sticking point for long-run risk models that rely on a small, but highly persistent component in consumption growth, as these models counterfactually imply that the price-dividend and price-consumption ratios should predict future, long-horizon consumption growth.

In the model presented here with parameter learning, there is no consumption growth predictability. The agent will ex post perceive the mean of consumption growth as changing, but in reality it is not (by assumption), and so the price-consumption ratio in the models with parameter uncertainty will not in population predict future consumption growth. Nevertheless, there is, in small samples, a correlation between the current price-consumption ratio and future consumption growth: if early consumption growth realizations happened to be high relative to the remainder of the sample, the price-consumption ratio will be negatively correlated with future consumption in-sample. Table 3 shows forecasting regression results for consumption growth and excess returns. The reported statistics are sample medians from the 20,000 simulated 100-year economies discussed previously.

[TABLE 3 ABOUT HERE]

Panel A of Table 3 shows that this small-sample correlation is not significant at the 1- or 5-year consumption growth forecasting horizons for the median economy. The average standard errors reported are Newey-West with lags accounting for autocorrelation on account of quarterly overlapping observations. Panel A also reports the risk-free rate regression of Hall (1978) on the simulated data. In particular, we regress quarterly consumption growth on the lagged risk-free rate. In a model with constant volatility of the pricing kernel, the coefficient on the real risk-free rate is a measure of the elasticity of intertemporal substitution, which in our model is 2. However, the reported median regression coefficient is -0.02 and insignificant, and the R^2 is low. This magnitude of the regression coefficient is consistent with what Beeler and Campbell (2012) show empirically. They also note that simulated data from the long-run risk model of Bansal and Yaron (2004) yields estimates of the EIS well in excess of 1. In the learning model, consumption growth is in fact unpredictable. The variation in the risk-free rate is due to time-variation in agents' perceived mean consumption growth rate, which is a function of their current beliefs. Thus, the long-run risk that arises through this learning channel does not result in counter-factual estimates of the EIS using the Hall-type regressions, even though the representative agent's elasticity of intertemporal substitution is in fact very high. In sum, the model with parameter uncertainty is a long-run risk model that addresses two of the main critiques Beeler and Campbell (2012) levy against long-run risk models.

Panel B of Table 3 addresses excess equity return predictability at the 1- and 5-year horizon. Here, the price-consumption ratio significantly predicts both 1- and 5-year equity returns with R^2's of 7% and 31%, respectively, over the median 100-year economy. These R^2 values are close to those reported in Beeler and Campbell (2012) who use the price-dividend ratio as the predictive variable. While not reported, the R^2 of the predictability regression is higher in the first 50 years than in the last 50 years as the effect of parameter uncertainty slowly wanes. This is broadly consistent with the evidence on excess return predictability using the price-dividend ratio as the predictive variable (see, e.g., Lettau and van Nieuwerburgh (2008)). Note that the return predictability arises both because excess returns are in fact predictable and because of an in-sample correlation between the price-dividend ratio and future returns. The in-sample relation is the same as that for consumption growth – if consumption has happened to be high, returns will also have been high, while the price-dividend ratio will have increased as investors' mean belief about the growth rate increases. Going forward, then, the returns are lower in an in-sample sense, and so there is a negative relation between the price-dividend ratio (or wealth-consumption ratio) and future excess returns (see also Timmermann (1996)). This evidence also implies that out-of-sample predictability is much lower than in-sample predictability, consistent with the empirical findings of Goyal and Welch (2006). Figure 2 show these dynamics by plotting a representative sample path of the ex ante annualized risk premium versus the ex post risk premium as predicted by the forecasting regression in Panel B of Table 3.

[FIGURE 2 ABOUT HERE]

3.1.4 Inspecting the mechanism

There are two particularly surprising results regarding the asset pricing implications of parameter learning when agents have a preference for early resolution of uncertainty. The first is that the volatility of the pricing kernel decreases at a much slower rate than the posterior variance of beliefs. The second is that after 100 years of learning, when the shocks to growth expectations are tiny – with a standard deviation of only 0.0041% per quarter – these long-run shocks increase the volatility of the pricing kernel by a factor of almost 1.5 relative to the known parameter benchmark economy. Here, we explain the economic rationale for both of these results in more detail.

[FIGURE 3 ABOUT HERE]

Long-lasting effects of learning. The analysis of the learning model points to a nonlinear relation between the level of parameter uncertainty, as measured by the level of the variance of beliefs over time (see Figure 1), and the impact of parameter learning as measured by standard asset pricing moments. Figure 1 shows that the posterior standard deviation initially decreases very rapidly – after 50 years it is 14 times smaller than the initial maximum prior dispersion of 1.65%.The top plot of Figure 3, however, shows that the standard deviation of the log pricing kernel – the price of risk – drop by a factor of about 2 over the same period. Over the next 50 years, the posterior standard deviation drops by a factor of 1.4, while the price of risk drops by a factor of about 1.3.

To understand these dynamics better, it is useful to consider the two components of the pricing kernel, as given in Equation (4), separately. In particular, the middle and bottom plots in Figure 3 show the annualized standard deviation of the two components of the log pricing kernel separately as a function of time. The "Power utility component" is $\ln \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}$, while the continuation utility component is $\ln \left(\beta \frac{PC_{t+1}+1}{PC_t} \right)^{\theta-1}$. The top plot shows that the volatility of the "power utility component" is always very close to the known-parameter benchmark price of risk, $\gamma \times \sigma = 0.33$. In the very beginning of the sample, the volatility is only slightly higher, which reflects the fact that subjective consumption growth volatility is slightly higher due to parameter uncertainty. This is the standard intuition we get from the power utility model: parameter learning has only a small, and highly transient, impact on the price of risk.

The bottom plot shows the conditional volatility of the "continuation utility component" as a function of time. With known parameters, this component has a conditional volatility equal to zero. In the parameter uncertainty case, however, the conditional volatility starts at about 0.8 and ends, after 100 years of learning, at about 0.15. Casual intuition would suggest a much quicker decline. In particular, from Equation (9), we have that the volatility of the shocks to the mean parameter belief, in the untruncated normal case, is $\frac{\sigma_t^2}{\sqrt{\sigma_t^2+\sigma^2}}$, which decreases by a factor of 10 over the 100 year sample calibrated with an initial quarterly prior dispersion of 0.26%. Thus, it would seem as though the amount of long-run risk decreases by a factor of 10 over the sample. Applying the intuition from the Bansal and Yaron model, this should relatively directly be reflected in a corresponding decrease in the volatility of the continuation utility component of the pricing kernel. However, in the case with parameter learning, the volatility dynamics are non-stationary which lead to an endogenous time-dependence in discount rates. In particular, endogenously high discount rates in the beginning of the sample make the consumption claim (total wealth) of relatively short *duration*. Thus, a given shock to mean parameter beliefs, μ_t, has a lower effect on total wealth early in the sample than later in the sample, when discount rates are lower and total wealth has relatively high duration. This intuition is confirmed in the top plot in Figure 4, which shows that the average path of the price-consumption ratio is increasing over time, indicating that discount rates decrease over time.[6] The middle plot of Figure 4 shows the numerical derivative of the log price-consumption ratio with respect to the mean parameter belief, μ_t, evaluated at the true mean, μ. This sensitivity is increasing over time. As mentioned, the volatility of the shocks to μ_t is rapidly decreasing over the sample. The net outcome of the two effects is showed in the bottom plot in Figure 4, which shows that $\frac{\sigma_t^2}{\sqrt{\sigma_t^2+\sigma^2}} \times \frac{\partial pc_t}{\partial \mu_t}|_{\mu_t=\mu}$ over time is decreasing, but at a slow rate corresponding more closely to the slow decline in the price of risk as shown in Figure 3.

[FIGURE 4 ABOUT HERE]

In sum, the asset pricing implications of parameter learning are large and long-lived due to the interaction of permanent shocks to beliefs about growth rates (subjective long-run consumption risk) and an endogenously increasing sensitivity of continuation utility with respect to these updates in beliefs.

Dynamics in the context of the Bansal and Yaron model. The approximate analytical solution to the Bansal and Yaron (2004) model provides a useful way to gain further intuition for the mechanics of the parameter learning case. Consider the homoskedastic case of the Bansal and Yaron model:

$$\Delta c_{t+1} = \mu + x_t + \sigma \varepsilon_{t+1}, \tag{11}$$

$$x_{t+1} = \rho x_t + \varphi \sigma \eta_{t+1}, \tag{12}$$

where both ε and η are i.i.d. normal shocks. We can, for intuition, think of these consumption dynamics as approximating the subjective consumption dynamics of the parameter learning case if we set ρ very high, say $\rho = 0.9999$, where x_t measures the time-variation in the long-run growth rate. The approximate solution to this model yields:

$$pc_t = A_0 + A_1 x_t. \tag{13}$$

[6]Since the subjective growth rate averaged across the 20,000 simulated economies is approximately constant, the increase in the P/C-ratio must come from a decrease in the discount rate.

Thus, the sensitivity of the log price-consumption ratio to x_t is $A_1 = \frac{1-1/\psi}{1-\kappa_1\rho}$, where $\kappa_1 = \frac{\exp(\overline{pc})}{1+\exp(\overline{pc})}$ is an equilibrium quantity. The question is how this sensitivity depends on changes in the amount of long-run risk, as given by the parameter φ in the Bansal and Yaron model. If the intertemporal elasticity of substitution is greater than one, which is our case, we get that $\frac{d\kappa_1}{d\varphi} < 0$ and so $\frac{dA_1}{d\varphi} < 0$ (see Appendix for proof). That is, the unconditional level of the price-consumption ratio increases when the amount of long-run risk, φ, decreases. This in turns means that the sensitivity of the price-consumption ratio to changes in x_t increases as φ decreases, analogously to what we find in the parameter learning case.

Next, we turn to the level effect of the very small volatility of the long-run shocks the learning model implies after 100 years. After this long of a history of learning, the decrease in the posterior variance is very slow. Therefore, we can reasonably look at the magnitude of the long-run risk effect using the Bansal and Yaron model, which has constant volatility of long-run shocks, as a laboratory, assuming that $\rho = 0.9999$. In particular, the shocks to the log stochastic discount factor in the Bansal and Yaron economy is given by:

$$m_{t+1} - E_t\left[m_{t+1}\right] = -\gamma\sigma\varepsilon_{t+1} - \left(\gamma - 1/\psi\right)\kappa_1 \frac{\varphi}{1-\kappa_1\rho}\sigma\eta_{t+1}. \tag{14}$$

In the learning case, the two shocks are perfectly positively correlated (see Equations (8) and (9)). Thus, we have that:

$$\sigma_t\left(m_{t+1}\right) = \left(\gamma + \left(\gamma - 1/\psi\right)\kappa_1 \frac{\varphi}{1-\kappa_1\rho}\right)\sigma. \tag{15}$$

To mimic our quarterly calibration after 100 years of learning, we set $\rho = 0.9999$, $\gamma = 10$, $\psi = 2$, $\beta = 0.994$, $\sigma = 0.0165$, $\mu = 0.0045$ and $\varphi = 0.00411\%/\sigma = 0.2491\%$. Given these parameters, we find the equilibrium $\kappa_1 = 0.9955$. This yields $\sigma_t\left(m_{t+1}\right) = 0.2495$ which means the *annualized* log volatility is 0.499 versus 0.33 in the benchmark, known parameters case. Thus, the very high persistence of the shocks and the fact that the long-run risk shocks are perfectly correlated with the shocks to realized consumption growth combine to generate approximately a 1.5 time increase in the volatility of the log pricing kernel, relative to the benchmark case where there is no long-run risk. This is very close to the magnitude we find in the numerical solution for the non-stationary learning problem after 100 years of learning.

4 Case 2: Learning about rare events

Uncertainty about parameters that govern rare events is likely to be large, as rare events by their very nature yield few historical observations available for agents to learn from. In recent work, Barro, Nakamura, Steinsson, and Ursua (2011), hereafter BNSU, estimate that consumption disasters occur with a probability of 2.8% per year using the longest consumption series as available from a wide cross-section of countries. This enables them to estimate disaster parameters with some degree of accuracy. For instance, the standard error of their estimate of the probability of a world disaster is 1.6%. Consumption volatility in disasters is estimated to be 12%, and a disaster is estimated to on average lead to a −14% permanent negative shock to consumption. The latter quantity is estimated with a standard error of 4.2%. So, even after using all historical data available in both the time-series and the cross-section of countries, there is quite a bit of uncertainty about the parameter estimates.

We will consider a simpler model for consumption disasters relative to BNSU, but the parameters and the associated parameter uncertainty is calibrated to their estimates as far as possible. In particular, let:

$$\Delta c_{t+1} = g_{t+1}\left(1 - D_{t+1}\right) + z_{t+1} D_{t+1}, \tag{16}$$

where

$$g_{t+1} = \mu_N + \sigma_N \varepsilon_{t+1}, \tag{17}$$

$$z_{t+1} = \mu_D + \sigma_D \varepsilon_{t+1}, \tag{18}$$

where ε is i.i.d. standard normal and where D_{t+1} is 1 with probability λ and 0 with probability $1-\lambda$. We assume that investors observe D_{t+1}. Given the very large average initial consumption decline in a disaster, as estimated by BNSU, learning whether you are in a disaster or not would not add much as the agent would be able to tell pretty much immediately anyway. There are two other simplifying assumptions here. First, true consumption growth is still i.i.d., whereas BNSU estimate the average disaster lasts for 6 years. Second, we only consider the permanent shocks to consumption and not the transitory effects BNSU also consider. Keeping the i.i.d. nature of consumption growth as in our initial case means that any asset price dynamics comes from the learning channel alone.

We set the disaster probability, mean and volatility to $\lambda = 2.8\%$, $\mu_D = -14\%$, $\sigma_D = 12\%$, respectively. We calibrate the mean and volatility in the good state such that we match the same unconditional consumption moments as before. In particular, the unconditional mean and volatility of non-time-averaged quarterly consumption are 0.45% and 1.65%, respectively. We consider two cases of parameter uncertainty; first, about the mean in the disaster state, μ_D, and then about the probability of a disaster, λ.

4.1 Uncertain mean of disaster state

We take the estimation uncertainty from BNSU of the disaster mean of 4.2% as the initial prior standard deviation here. Thus, this exercise is forward-looking in the sense that this is the best estimate available using all data up until now. The prior is assumed to be unbiased with a mean of -14%.[7] Clearly, the agent will only learn about the disaster mean when a disaster state occurs, which is what makes learning much slower in this case. With 2.8 disasters per 100 years, learning is loosely speaking 35 times slower than in the simple i.i.d. case considered previously. Since learning in this case only happens quite rarely, there is little dynamics induced by learning in terms of excess volatility, return predictability, and time-variation in the price-consumption ratio. In particular, the price-consumption ratio is constant between disasters and return volatility only reflects realized consumption growth during normal times. Given this, it is clear that outside of disasters, parameter uncertainty about the disaster mean will mainly give an increase in the level of Sharpe ratios and the risk premium. In a disaster, however, there is quite a bit of learning as the initial prior uncertainty is large.

The two middle columns of Table 4 shows average moments from 20,000 simulated economies with 100 year samples given the initial prior, as well as the known disaster mean benchmark model. The preference parameters are $\gamma = 7$ and $\beta = 0.993$ and $\psi = 2$, similar to the parameters used in BNSU. The time-preference parameter β is calibrated to roughly match the real risk-free rate. This parameter is important as it determines the effective duration of a permanent shock in terms of its effect on the continuation utility.

[TABLE 4 ABOUT HERE]

In the uncertain disaster mean case, the risk premium is 4.7% versus 3.0% in the known disaster mean case. The volatility of the log pricing kernel is 1.07 and 0.72, respectively.

[7] As before, the prior is a truncated normal. The upper and lower truncation bounds are set at $+/-$ 3 standard deviations of the initial prior (4.2%) around the mean of the initial prior (-14%).

Thus, parameter uncertainty increases these quantities by about 50%. Note that this increase is very large relative to the small increase in subjective consumption growth volatility in the disaster state. In particular, for the initial observation, subjective consumption growth volatility in the disaster state is $\sqrt{0.042^2 + 0.12^2} = 0.127$ – only slightly higher than the objective volatility of 0.12. Again, it is the long-run risk aspect of the shocks to beliefs that makes the price of risk increase by as much as it does. The risk-free rate is low as in the data by construction and lower than in the known parameter case as there is more risk with uncertain parameters. The excess volatility is very small for the parameter uncertainty case, as learning only occurs very rarely when disasters occurs. Related, there is no economically significant return predictability in this model (not reported). Thus, uncertainty about parameters that one can only learn about during the rare event itself can have a large effect on the level of the risk premium and maximum Sharpe ratio, but will not lead to interesting dynamics in the price of risk and/or the risk premium in normal times. On the other hand, for these type of learning problems, it will take a very long time for agents to learn and thus for the asset pricing implications to become economically insignificant. In particular, the conditional volatility of the log pricing kernel after 100 years averaged across the 20,000 simulated economies is 0.96 times the initial conditional volatility of the log pricing kernel.

4.2 Uncertain probability of disasters

The posterior standard deviation of the BNSU estimate of the probability of a world disaster is 1.6%. We calibrate the model at the quarterly frequency, and so set $\lambda = 0.7\%$ with a prior standard deviation of 0.4%. The Beta-distribution yields a conjugate prior for a probability, and so we assume that the prior at time t is $\lambda \sim ß(a_t, b_t)$. Here a_t is the number of times a disaster has happened, while b_t is the number of times the normal state has occurred. The 0.4% standard error reported by BNSU is roughly consistent with having observed a total of 400 quarterly observations. Thus, we set $a_0 = 2.8$ and $b_0 = 397.2$. This means the mean belief is unbiased and equal to $0.7\% = \frac{a_0}{a_0+b_0}$.

The two rightmost columns of Table 4 report the average 100-year sample moments across 20,000 simulated economies for the model with uncertain disaster probability and the benchmark case of known parameters. The preference parameters are $\gamma = 7$ and $\psi = 2$, as in the case of unknown disaster mean, and $\beta = 0.9905$ is set to roughly match the level of the real risk-free rate. The average annualized volatility of the pricing kernel and risk premium are 1.3 and 6.6%, respectively. In the benchmark know disaster probability case, these quantities are 0.72 and 3.1%. Thus, learning about the disaster probability appears to have a stronger effect than learning about the disaster mean. Further, the volatility of returns is 5% versus 4.2% in the benchmark case, and so the excess volatility measure is 0.18 – still short of what is in the data (0.7), but more than 4 times that of the case of uncertain disaster mean.

There are two reasons why uncertainty about the disaster probability helps more with explaining standard asset pricing moments than learning about the disaster mean. First, the subjective distribution about the disaster probability is positively skewed with high kurtosis. Thus, there is a non-trivial probability assigned to the disaster probability being relatively high. This is very risky for the agent as disasters are very bad events. Second, the updating is continuous. In normal times, there are no disasters, which is reflected in b_t increasing while a_t stays constant. If a disaster occurs, b_t is constant but a_t increases. Thus, the nature of the uncertainty about λ changes every period and so asset prices change as well, leading to excess volatility. In fact, Table 5 shows that excess returns are predictable in this model over the 100-year samples, while consumption growth is not - much like what was the case for when agents learn about the mean in the initial simple i.i.d. model. Figure 5 shows the conditional moments of the model over time averaged across the simulated economies. The annualized Sharpe ratio (volatility of the pricing kernel) decreases from about 1.5 to 0.9 over the sample, while the annualized conditional risk-premium decreases from about 9% to 5%.

[FIGURE 5 ABOUT HERE]

The average moments do not reveal all of the dynamics of the disaster models, however. In particular, while it is clear there is a decrease in the price of risk and the risk premium on average due to decreased parameter uncertainty, the actual sample paths look more interesting. For each disaster, the subjective belief of the disaster probability increases markedly, which is reflected in the price-consumption ratio, the Sharpe ratio and the risk premium. As long as a disaster does not occur, the subjective mean of the disaster probability decreases. Thus, there is a "saw-tooth" pattern in asset prices and beliefs when learning about the disaster probability. Figure 6 shows a representative sample path for the conditional risk premium in this model, which shows that the "saw-tooth"-pattern in beliefs is reflected in the risk premium. Each vertical increase in the risk premium happens when a disaster occurs.

[FIGURE 6 ABOUT HERE]

5 Case 3: Structural breaks

With parameter learning, rational agents will eventually learn any fixed parameter. Of course, "eventually" may be in a really long time, but still such parameter learning does not embody the notion of a "new" paradigm which anecdotally may be an important component of agents' belief formation (see the discussion in Hong, Stein, and Yu (2007)). Structural breaks, studied earlier in the context of asset pricing by for instance Timmermann (2001) and Pastor and Veronesi (2001), is a way to make parameter learning a recurring problem. In this section, we consider a structural breaks version of the simple i.i.d. consumption growth economy.

In particular, we assume that log aggregate consumption growth within paradigm s is given by:

$$\Delta c_{t+1} = \mu_s + \sigma \varepsilon_{t+1}, \tag{19}$$

where ε is i.i.d. standard normal, σ is the constant volatility parameter, and where μ_s is a paradigm-specific mean growth rate, where s denotes the s'th paradigm. Each period there is a constant probability λ that there is a structural break. If a structural break occurs, a new mean growth rate μ_{s+1} is drawn from a normal distribution with mean μ and standard deviation σ_μ. The agent is assumed to know when a new paradigm has been drawn, but not the value of the mean of that paradigm, μ_{s+1}. Note that the fact that a redraw occurs does not imply that the new mean is far from the current mean. In fact, the assumption of a normal distribution for the μ_s's along with a relatively small σ_μ means that the most likely outcome is around the unconditional mean. Likely candidates for times of a redraw includes the beginning and end of world wars, technological revolutions (e.g., the dot-com era), as well as the recent financial crises. The assumption of a constant λ is a simplification that it is easy to extend. However, this assumption makes the analysis cleaner as any dynamics in the price of risk and the risk premium will in this case come from the learning channel alone.

This model in many ways looks much like the original long-run risk model of Bansal and Yaron (2004). In both cases, the true conditional mean of consumption growth is time-varying, but very persistent. There are two main differences. First, in the structural breaks model, the mean is constant within each regime, which means that the agents in the economy face a paradigm-specific parameter learning problem. The parameter learning induces quite different dynamics relative to what learning about the long-run risk component in the Gaussian state space model of Bansal and Yaron would. Second, we calibrate the regimes to have an expected duration of 50 years. Thus, these are in fact even longer-run risks than those assumed in Bansal and Yaron (2004).

5.1 Results from a calibrated model

We calibrate this model at the quarterly frequency. Thus, we set $\mu = 0.45\%$, $\sigma_\mu = 0.25\%$, $\lambda = 0.5\%$, $\sigma = 1.65\%$. We truncate the normal distribution for the redraw at $+/- \ 4 \times \sigma_\mu$ around the unconditional mean, μ. There is admittedly a lot of uncertainty about the distribution from which the paradigm-specific mean is drawn, but for now we assume the meta-parameters μ and σ_μ are known to isolate the effect of the structural breaks assumption. We set the preference parameters as in the initial i.i.d. case: $\beta = 0.994$, $\gamma = 10$, and $\psi = 2$. We also consider a case with $\psi = 1.5$ to show the effect of decreasing the preference for early resolution of uncertainty by lowering the elasticity of intertemporal substitution.

[TABLE 5 ABOUT HERE]

Table 5 shows the 100-year sample moments averaged across 20,000 simulated economies. For the model with $\psi = 2$ and unobserved paradigm means, the risk premium is 4.7% per year and the annualized volatility of the log pricing kernel is 0.72. The volatility of returns is 6.5% which implies an excess volatility of 0.22 relative to the 5% volatility of cash flow growth. The risk-free rate is 1.2% with a volatility of 0.2%. Thus, the structural breaks model does much better than the known parameters model with i.i.d. consumption growth, as given in Table 2. However, when compared to the case of structural breaks with a known paradigm mean, the learning model has a *lower* price of risk and risk premium. In particular, for the known means structural breaks case, the annualized volatility of the log pricing kernel is 1.25 and the annualized risk premium is 6.8%. The volatility of returns is lower, however, at 5.4%.

Given the results in the first simple i.i.d. consumption growth case, where parameter learning gives a risk premium of 4.4% versus only 1.7% in the known mean benchmark case, the fact that learning in the structural breaks case decreases the price of risk and the risk premium relative to if the paradigm mean is observed might seem surprising. However, the reason is straightforward. In the simple i.i.d. case, the agent was learning about a fixed quantity, and so shocks beliefs gave a permanent shock to consumption growth expectations. Thus, in this case learning yields slightly higher short-run risk and much higher long-run risk.

However, in the structural breaks case, while the individual paradigm means are constant, expected true consumption growth follows a stationary, mean-reverting process. In this case, the optimal learning smooths the beliefs about the conditional mean consumption growth relative to the unconditional, known mean μ. Thus, learning now yields *less* long-run risk and slightly more short-run risk. This is perhaps easiest to understand by referring to well-known Kalman filter results. Consider the Bansal and Yaron (2004) model, which is a Gaussian state-space model. If the conditional mean process (x_t) is unobserved, the filtered $\hat{x}_t$-process will have the same autocorrelation coefficient, but lower volatility than the true x_t process. In other words, *less* long-run risk.

The rightmost columns of Table 5 shows the case when $\psi = 1.5$, all else equal. Now, the price of risk and the risk premium decline for both the case of unknown and known paradigm means. In the learning case, the risk premium is now 4.0%, while it is 5.2% in the observed means case. Thus, as expected, a decrease in the preference for early resolution of uncertainty decreases the risk price for shocks to beliefs about future consumption dynamics.

5.1.1 Forecasting regressions and risk dynamics

While learning unconditionally decreases the price of risk in the structural breaks case relative to the case of known paradigm means, it does induce interesting dynamics. In particular, the known paradigm means case has constant price of risk and risk premiums. In the learning case, on the other hand, the risk premium and price of risk increases at the onset of a new paradigm as parameter uncertainty increases. Figure 9 shows a representative sample path for the annualized risk premium over a 100 year period. In this sample path, there are three breaks, around 5 years, 60 years, and 95 years. In each case, the risk premium shoots up from $2-3\%$ to more than 8%, as parameter uncertainty jumps up due to the redraw of the paradigm mean. The lower plot of Figure 9 shows that the price-consumption ratio also typically decreases at the onset of a new regime. However, note that the decrease depends on the consumption growth realizations early in the new paradigm. There are cases where a high initial consumption growth realization causes the price-consumption ratio to move up on account of the ensuing high subjective belief about the paradigm mean, even though the risk premium still increases.

[FIGURE 7 ABOUT HERE]

Table 6 shows the forecasting regressions from the structural breaks model. As before, Panel A considers consumption growth predictability. In the structural breaks case, consumption growth is in fact predictable over very long horizons. However, as the regressions show, using the price-consumption ratio or the real risk-free rate as the predictive variables lead to no significant predictability. Again, the estimate of the EIS from the risk-free rate regression is negative and comparable to that found in the same regression run on the historical data, even though the EIS is in fact 2. First, the actual predictability in consumption growth is quite small over conventional forecasting horizons and, second, the subjective estimates of the growth rate are quite volatile, especially at the start of a new paradigm, and so the relation between the risk-free rate, which reflects the subjective beliefs, and future consumption growth, which reflects the truth, is very weak.

[TABLE 6 ABOUT HERE]

Panel B of Table 6 shows that the price-consumption ratio is, in the median simulated economy, insignificantly related to future 1- and 5-year excess returns. However, as also shown in Panel B, if one conditions the start of a 50-year sample period as being the start of a new paradigm, the price-consumption ratio reemerges as a significant return predictor also in the structural breaks model. The reason for this is that the structural breaks with the different growth paradigms create time-variation in the price-consumption ratio that is unrelated to the risk premium. By conditioning on a redraw of the paradigm mean at the beginning of the samples across the simulated economies, the true mean of consumption growth is on average constant for the next 50 years, and so the relation between the price-consumption ratio and future returns reemerges as in the "Case 1"-model considered earlier with parameter uncertainty about the unconditional mean growth rate. This is consistent with Lettau and van Nieuwerburgh (2008) who show that if one estimates structural breaks in the aggregate price-dividend ratio and removes the paradigm means from this ratio, the resulting adjusted price-dividend ratio is a much stronger predictor of future excess returns than the actual price-dividend ratio.

In sum, the structural breaks model delivers high Sharpe ratios, a high risk premium, excess return volatility, and excess return predictability. It does this with a high elasticity of intertemporal substitution, but still replicating the Hall (1978) regressions in the data. Further, the price-consumption ratio does not significantly predict future consumption growth up to the 5-year forecasting horizon. In this sense, the structural breaks model addresses many of the critiques of the long-run risk models raised by Beeler and Campbell (2012).

The mean and volatility of the redraw distribution, μ and σ_μ, are both assumed known in this model, however. This is quite unrealistic as there effectively is only one observation every 50 years about this distribution. If one were to add parameter uncertainty about, say, μ as well, the asset pricing implications of the model are likely to take on aspects of the initial parameter uncertainty problem first considered in this paper. We conjecture that this would add risk, excess return volatility, and further return predictability.

References

[1] Ai, H. (2010), "Information about Long-Run Risk: Asset Pricing Implications," forthcoming, *Journal of Finance.*

[2] Backus, D., M. Chernov and I. Martin (2009), "Disasters Implied by Equity Index Options", *NYU Working Paper*

[3] Bakshi, G. and G. Skoulakis (2010), "Do Subjective Expectations Explain Asset Pricing Puzzles?", *Journal of Financial Economics,* December 2010, 117 - 140.

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

[5] Bansal, R. and I. Shaliestovich (2010), "Confidence Risk and Asset Prices," *American Economic Review P&P*, 100, 537 – 541.

[6] Barberis, N. (2000), "Investing for the Long Run When Returns Are Predictable", *Journal of Finance* 55(1), 225 - 264

[7] Barillas, F., Hansen, L. and T. Sargent, "Doubts or Variability?", *Journal of Economic Theory*, forthcoming.

[8] Barro, R. (2006), "Rare Disasters and Asset Markets in the Twentieth Century", *Quarterly Journal of Economics* 121, 823 - 866

[9] Barro, R., Nakamura, E., Steinsson, J., and J. Ursua (2011), "Crises and Recoveries in an Empirical Model of Consumption Disasters," Columbia Business School working paper

[10] Barro, R. and J. Ursua (2008), "Consumption Disasters since 1870", *Brookings Papers on Economic Activity*

[11] Beeler, J. and J. Campbell (2011), "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," Harvard Working paper.

[12] Boudoukh, J., M. Richardson, and R. Whitelaw (2008), "The Myth of Long Horizon Predictability," *Review of Financial Studies* 21, 1577 - 1605.

[13] Brandt, M. , Q. Zeng, and L. Zhang (2004), "Equilibrium stock return dynamics under alternative rules of learning about hidden states," *Journal of Economic Dynamics and Control*, 28, 1925 – 1954.

[14] Brennan, M. and Y. Xia (2002), "Dynamic Asset Allocation Under Inflation," *Journal of Finance* 57, 1201 – 1238.

[15] Bray, M. M. and N. E. Savin (1986), "Rational Expectations Equilibria, Learning, and Model Specification," *Econometrica* 54, 1129 - 1160.

[16] Cagetti, M., L. Hansen, T. Sargent and N. Williams (2002), "Robustness and Pricing with Uncertain Growth", *Review of Financial Studies* 15, 363 - 404

[17] Cecchetti, S., P. Lam and N. Mark (1990), "Mean Reversion in Equilibrium Asset Prices", *American Economic Review* 80, 398 - 418

[18] Cecchetti, S., P. Lam and N. Mark (1993), "The Equity Premium and the Risk Free Rate: Matching the Moments", *Journal of Monetary Economics* 31, 21 - 46

[19] Chen, H., S. Joslin, and N. Tran (2010), "Rare Disasters and Risk Sharing with Heterogeneous Beliefs," MIT Working paper

[20] Cogley, T. and T. Sargent (2008), "The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression?", *Journal of Monetary Economics* 55, 454 - 476

[21] Cogley, T. and T. Sargent (2009), "Anticipated Utility and Rational Expectations as Approximations of Bayesian Decision Making", *International Economic Review* 49, 185 - 221

[22] David, A. and P. Veronesi (2009), "What ties return volatilities to price valuations and fundamentals?" University of Calgary Working Paper.

[23] Detemple, J. (1986), "Asset pricing in a production economy with incomplete information." *Journal of Finance*, 41, 383–390.

[24] Dothan, M. U. and D. Feldman (1986), "Equilibrium interest rates and multiperiod bonds in a partially observable economy." *Journal of Finance*, 41, 369 – 382.

[25] Drechsler, I. and A. Yaron (2008), "What's Vol Got to Do with It?", Working Paper, Wharton School of Business, University of Pennsylvania

[26] Epstein, L. and S. Zin (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica* 57, 937 - 969

[27] Fama, E. and K. French (1988), "Dividend yields and expected stock returns," *Journal of Financial Economics* 22, 3 - 25.

[28] Fama, E. and K. French (2002), "The Equity Premium," *Journal of Finance* 57, 637 – 659.

[29] Froot, K. and S. Posner (2002), "The Pricing of Event Risks with Parameter Uncertainty", *GENEVA Papers on Risk and Insurance - Theory* 27, 153 - 165

[30] Hall, R. (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, December 1978, 86(6), pp. 971-987.

[31] Gabaix, X. (2009), "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance", *NYU Working Paper*

[32] Gennotte, G. (1986), "Optimal Portfolio Choice under Incomplete Information," *Journal of Finance*, 61, 733-749.

[33] Geweke, J. (2001), "A note on some limitations of CRRA utility," *Economics Letters* 71, 341 - 345.

[34] Goyal A. and I. Welch, "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," July 2008, *Review of Financial Studies* 21(4) 1455-1508.

[35] Hall, Robert E. (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy,* December 1978, 86(6), pp. 971-987.

[36] Hansen, L. (2007), "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk," Richard T. Ely Lecture, *The American Economic Review* 97, No. 2., 1 - 30

[37] Hansen, L. and T. Sargent (2009), "Robustness, Estimation and Detection," Working paper.

[38] Hansen, L. and T. Sargent (2010), "Fragile Beliefs and the Price of Uncertainty," *Quantitative Economics*, Vol. 1, Issue 1, pp. 129-162

[39] Hong, H., J. C. Stein,and J. Yu (2007), "Simple forecasts and paradigm shifts," *Journal of Finance*, 62, 1207 – 1242

[40] Johannes, M., L. A. Lochstoer, and Y. Mou (2011), "Learning about Consumption Dynamics," Working paper, Columbia University.

[41] Johnson, T. (2007), "Optimal Learning and New Technology Bubbles," *Journal of Monetary Economics*, 87, 2486 – 2511.

[42] Kandel, S. and R. Stambaugh (1990), "Expectations and Volatility of Consumption and Asset Returns", *Review of Financial Studies* 2, 207 - 232

[43] Kandel, S. and R. Stambaugh (1996), "On the Predictability of Stock Returns: An Asset-Allocation Perspective", *Journal of Finance* 51, 66 - 74

[44] Kreps, D. (1998), "Anticipated Utility and Dynamic Choice", *Frontiers of Research in Economic Theory*, (Cambridge: Cambridge University Press) 242 - 274

[45] Lettau, M. and S. Ludvigson (2001), "Consumption, Aggregate Wealth and Stock Returns," *Journal of Finance* 56, 815 - 849

[46] Lettau, M., S. Ludvigson and J. Wachter (2008), "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?", *Review of Financial Studies* 21(4), 1653 - 1687

[47] Lewellen, J. and J. Shanken (2002), "Learning, Asset-Pricing Tests, and Market Efficiency," *Journal of Finance* (57(3), 1113 - 1145.

[48] Lucas, R. (1978), “Asset Prices in an Exchange Economy,” *Econometrica* 46, 1429-1446

[49] Lucas, R. and T. Sargent (1979), "After Keynesian Macroeconomics," The Federal Reserve Bank of Minneapolis, *Quarterly Review* 321.

[50] Malmendier, U. and S. Nagel (2011), "Depression babies: Do Macroeconomic Experience Affect Risk-Taking?", *Quarterly Journal of Economics* 126, 373 – 416.

[51] Mehra, R. and E. Prescott (1985), "The Equity Premium: A Puzzle", *Journal of Monetary Economics* 15, 145 - 161

[52] Moore, B. and H. Schaller (1996), "Learning, Regime Switches, and Equilibrium Asset Pricing Dynamics", *Journal of Economic Dynamics and Control* 20, 979 - 1006

[53] Pastor, L. (2000), "Portfolio Selection and Asset Pricing Models," *Journal of Finance* 55, 179 - 223

[54] Pastor, L. and P. Veronesi (2003), "Stock valuation and learning about profitability," *Journal of Finance*, 58, 1749 – 1789.

[55] Pastor, L. and P. Veronesi (2006), "Was there a NASDAQ bubble in the late 1990s?," *Journal of Financial Economics*, 81, 61 – 100.

[56] Pastor, L. and P. Veronesi (2009), "Learning in Financial Markets," *Annual Review of Financial Economics.*

[57] Piazzesi, M. and M. Schneider (2010), "Trend and Cycle in Bond Risk Premia," Working Paper Stanford University

[58] Rietz, T. (1988), "The equity risk premium: A solution?", *Journal of Monetary Economics*, Volume 22, Issue 1, July 1988, Pages 117-131

[59] Romer, C. (1989), "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1908", *Journal of Political Economy* 97, 1 - 37

[60] Shaliastovich, I. (2008), "Learning, Confidence and Option Prices," Working Paper, Duke University

[61] Stambaugh, R. (1999), "Predictive Regressions." *Journal of Financial Economics,* 1999, 54, pp. 375–421.

[62] Timmermann, A. (1993), "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices." *Quarterly Journal of Economics*, 1993, 108, 1135-1145.

[63] Timmermann, A. (2001), "Structural breaks, incomplete information, and stock prices," *Journal of Business & Economic Statistics* 19, 233 – 315.

[64] Veronesi, P. (1999), "Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model," *Review of Financial Studies*, 12, 5, Winter 1999

[65] Veronesi, P. (2000), "How does information quality affect stock returns?" *Journal of Finance*, 55, .

[66] Weitzman, M. (2007), "Subjective Expectations and Asset-Return Puzzles," *American Economic Review*, 97(4), 1102-1130.

[67] Whitelaw, R. (2000), "Stock Market Risk and Return: An Equilibrium Approach", *Review of Financial Studies* 13, 521 - 547.

[68] Working, H. (1960), "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica* 28(4), 916 - 918.

[69] Xia, Y. (2001), "Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation", *Journal of Finance* 56(1), 205 - 246

Table 1 - Decade by decade

Table 1: This table gives average annualized sample moments from 20,000 simulations of 400 quarters of data from the model where the representative agent has a preference for early resolution of uncertainty ($\gamma = 10$ and $\psi = 2$). The sample moments are broken into decades, however, to illustrate the effect of parameter learning over time. In particular, the second column shows the prior dispersion parameter at the beginning of each decade ($\sigma_t(\mu)$). The remaining columns show the risk-free rate, the difference between the 10-year zero-coupon yield and the short-term risk-free rate, the equity premium, equity return volatility, and finally the volaility of the log pricing kernel.

	Prior uncertainy $\sigma_t(\mu)$	SDF volatility $\sigma_T[m_{t,t+1}]$	Asset price moments $[\gamma = 10, \psi = 2]$ $E_T[R_{f,t}]$	$E_T[y_t^{10}\text{-}r_{f,t}]$	$E_T[R_{M,t}\text{-}R_{f,t}]$	$\sigma_T[R_{M,t}\text{-}R_{f,t}]$
Decade 1	1.65%	1.05	−1.3%	1.4%	11.0%	10.7%
Decade 2	0.26%	0.87	−0.2%	0.6%	8.1%	9.4%
Decade 3	0.18%	0.75	0.6%	0.2%	6.3%	8.4%
Decade 4	0.15%	0.67	1.0%	0.1%	5.2%	7.8%
Decade 5	0.13%	0.61	1.3%	0.1%	4.5%	7.3%
Decade 6	0.12%	0.57	1.5%	0.0%	4.0%	7.0%
Decade 7	0.11%	0.54	1.7%	0.0%	3.6%	6.7%
Decade 8	0.10%	0.51	1.8%	0.0%	3.4%	6.5%
Decade 9	0.09%	0.49	1.8%	0.0%	3.2%	6.4%
Decade 10	0.09%	0.48	1.9%	0.0%	3.0%	6.2%
Known parameters	0.00%	0.33	2.5%	0.0%	1.7%	5.0%

Table 2 - 100 year sample moments

Table 2: This table gives average annualized sample moments from 20,000 simulations of 400 quarters of data from each model. The initial prior is centered around the true mean with a standard error of 0.26%, which corresponds to the dispersion that would obtain if one started with a flat prior and had learned for ten years. $E_T[x]$ denotes the sample mean of x, $\sigma_T[x]$ denotes the sample standard deviation of x, and m is the log stochastic discount factor, R_M denotes the simple "market" return, defined as 1.5 times the return to the consumption claim. R_f is the real simple risk-free rate, y_{10} is the continuously compounded annual yield on a zero-coupon default-free bond. "Excess volatility" is defined as the relative amount of return volatility in excess of the volatility of cash flow growth. The values in the "Data" column are taken from Bansal and Yaron (2004) and correspond to U.S. data from 1929 to 1998. In their data, dividend growth volatility is 11.5%, while return volatility is 19.4% which means "excess volatility" is $19.4/11.5 - 1 = 0.70$. All statistics are annualized.

		Preference for early resolution of uncertainty ($\gamma = 10$, $\psi = 2$)		**Indifferent to the timing of resolution of uncertainty** ($\gamma = 10$, $\psi = 1/\gamma$)	
	Data	Unknown μ	Known μ	Unknown μ	Known μ
$\sigma_T[m_t]$	≥ 0.6	0.60	0.33	0.33	0.33
$E_T[R_{M,t} - R_{f,t}]$	6.33	4.42	1.67	-1.39	1.67
$\sigma_T[R_{M,t} - R_{f,t}]$	19.42	7.35	5.00	5.31	5.00
Excess volatility	≈ 0.70	0.47	0.00	0.06	0.00
$E_T[R_{f,t}]$	0.86	1.34	2.51	15.1	15.1
$\sigma_T[R_{f,t}]$	0.97	0.67	0.00	2.12	0.00
$E_T[y^{10} - r_f]$	≈ 0	0.09	0.00	-1.26	0.00

Table 3 - Forecasting regressions

Table 3: This table shows the results from forecasting regressions of 1- and 5-year log consumption growth and excess market returns on the lagged log price-consumption ratio, as well as a regression of one quarter ahead consumption growth on the log risk-free rate. The β's reported are the median regression coefficient across 20,000 simulated paths from the model with $\gamma = 10$ and $\psi = 2$. Each sample path is 100 years long. The initial prior is centered around the true mean with a standard error of 0.26%, which corresponds to the dispersion that would obtain if one started with a flat prior and had learned for ten years. The median Newey-West t-statistic is also reported, where the number of lags equals the number of overlapping observations. The regressions use quarterly simulated data, so for the annual forecasting horizon there are 3 lags used. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. Finally, the median R^2 is also reported for each regression. The "data" columns are taken from Beeler and Campbell (2011), who use U.S. data from 1930 to 2008.

	Data		**Median model outcomes**		
Forecasting horizon	β^{data}	$\bar{R}^{2,\text{data}}$	β^{model}	$(s.e.)$	$\bar{R}^{2,\text{model}}$
Panel A: Consumption growth predictability					
Consumption growth vs. P/C-ratio: $\Delta c_{t,t+j} = \alpha + \beta pc_t + \varepsilon_{t,t+j}$					
1 year	0.01	6.0%	-0.01	(0.05)	0.0%
5 years	-0.00	0.0%	-0.12	(0.10)	3.0%
Consumption growth vs. risk-free rate: $\Delta c_{t,t+j} = \alpha + \beta r_{f,t} + \varepsilon_{t,t+j}$					
1 quarter	-0.12	not reported	-0.02	(0.49)	0.0%
Panel B: Excess return predictability					
Excess returns vs. P/C-ratio: $r_{t,t+j} - r_{f,t,t+j} = \alpha + \beta pc_t + \varepsilon_{t,t+j}$					
1 year	-0.09^{*}	4.4%	-0.25^{***}	(0.08)	7.4%
5 years	-0.41^{***}	26.9%	-1.16^{***}	(0.30)	30.9%

Table 4 - 100-year sample moments from disaster models

Table 4: This table gives average annualized sample moments from 20,000 simulations of 400 quarters of data from two version of parameter uncertainty in the consumption disaster model. $E_T[x]$ denotes the sample mean of x, $\sigma_T[x]$ denotes the sample standard deviation of x, and m is the log stochastic discount factor, R_M denotes the simple "market" return, defined as 1.5 times the return to the consumption claim. R_f is the real simple risk-free rate, y_{10} is the continuously compounded annual yield on a zero-coupon default-free bond. "Excess volatility" is defined as the relative amount of return volatility in excess of the volatility of cash flow growth. The values in the "Data" column are taken from Bansal and Yaron (2004) and correspond to U.S. data from 1929 to 1998. In their data, dividend growth volatility is 11.5%, while return volatility is 19.4% which means "excess volatility" is $19.4/11.5 - 1 = 0.70$. All statistics are annualized.

		Learning about disaster mean ($\beta = 0.993, \gamma = 7, \psi = 2$)		**Learning about disaster probability** ($\beta = 0.9905, \gamma = 7, \psi = 2$)	
	Data	Unknown μ_D	Known μ_D	Unknown λ	Known λ
$\sigma_T[m_t]$	> 0.5	1.07	0.72	1.33	0.72
$E_T[R_{M,t} - R_{f,t}]$	6.33	4.70	3.04	6.62	3.05
$\sigma_T[R_{M,t} - R_{f,t}]$	19.42	4.36	4.20	4.96	4.22
Excess volatility	≈ 0.70	0.04	0.00	0.18	0.00
$E_T[R_{f,t}]$	0.86	1.10	2.03	1.09	3.04
$\sigma_T[R_{f,t}]$	0.97	0.06	0.00	0.38	0.00
$E_T[y^{10} - r_f]$	≈ 0	0.00	0.00	0.05	0.00

Table 5 - Forecasting regression from disaster model

Table 5: This table shows the results from forecasting regressions of 1- and 5-year log consumption growth and excess market returns on the lagged log price-consumption ratio, as well as a regression of one quarter ahead consumption growth on the log risk-free rate. The β's reported are the median regression coefficient across 20,000 simulated paths from the model with learning about the disaster probability with $\gamma = 7$ and $\psi = 2$. Each sample path is 100 years long. The median Newey-West t-statistic is also reported, where the number of lags equals the number of overlapping observations. The regressions use quarterly simulated data, so for the annual forecasting horizon there are 3 lags used. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. Finally, the median R^2 is also reported for each regression. The "data" columns are taken from Beeler and Campbell (2011), who use U.S. data from 1930 to 2008.

	Data		**Median model outcomes**		
Forecasting horizon	β^{data}	$\bar{R}^{2,\text{data}}$	β^{model}	$(s.e.)$	$\bar{R}^{2,\text{model}}$
Panel A: Consumption growth predictability					
Consumption growth vs. P/C-ratio: $\Delta c_{t,t+j} = \alpha + \beta pc_t + \varepsilon_{t,t+j}$					
1 year	0.01	6.0%	-0.02	(0.06)	0.3%
5 years	-0.00	0.0%	-0.27	(0.29)	3.3%
Consumption growth vs. risk-free rate: $\Delta c_{t,t+j} = \alpha + \beta r_{f,t} + \varepsilon_{t,t+j}$					
1 quarter	-0.12	not reported	-0.29	(0.37)	0.0%
Panel B: Excess return predictability					
Excess returns vs. P/C-ratio: $r_{t,t+j} - r_{f,t,t+j} = \alpha + \beta pc_t + \varepsilon_{t,t+j}$					
1 year	-0.09^{*}	4.4%	-0.53^{***}	(0.14)	8.8%
5 years	-0.41^{***}	26.9%	-2.51^{***}	(0.58)	35.9%

Table 6 - 100 year sample moments for Structural Breaks case

Table 6: This table gives average annualized sample moments from 20,000 simulations of 400 quarters of data from models with structural breaks. $E_T[x]$ denotes the sample mean of x, $\sigma_T[x]$ denotes the sample standard deviation of x, and m is the log stochastic discount factor, R_M denotes the simple "market" return, defined as 1.5 times the return to the consumption claim. R_f is the real simple risk-free rate, y_{10} is the continuously compounded annual yield on a zero-coupon default-free bond. "Excess volatility" is defined as the relative amount of return volatility in excess of the volatility of cash flow growth. The values in the "Data" column are taken from Bansal and Yaron (2004) and correspond to U.S. data from 1929 to 1998. In their data, dividend growth volatility is 11.5%, while return volatility is 19.4% which means "excess volatility" is $19.4/11.5 - 1 = 0.70$. All statistics are annualized.

		$\gamma = 10, \quad \psi = 2$		$\gamma = 10, \quad \psi = 1.5$	
	Data	Unknown μ	Known μ	Unknown μ	Known μ
$\sigma_T[m_t]$	> 0.5	0.72	1.25	0.67	0.99
$E_T[R_{M,t} - R_{f,t}]$	6.33	4.65	6.81	3.95	5.17
$\sigma_T[R_{M,t} - R_{f,t}]$	19.42	6.45	5.44	5.94	5.21
Excess volatility	≈ 0.70	0.29	0.09	0.19	0.04
$E_T[R_{f,t}]$	0.86	1.24	-0.38	1.70	0.75
$\sigma_T[R_{f,t}]$	0.97	0.20	0.40	0.21	0.18

Table 7 - Forecasting regressions, structural breaks case

Table 7: This table shows the results from forecasting regressions of 1- and 5-year log consumption growth and excess market returns on the lagged log price-consumption ratio, as well as a regression of one quarter ahead consumption growth on the log risk-free rate. The β's reported are the median regression coefficient across 20,000 simulated paths from the structural breaks model with $\gamma = 10$ and $\psi = 2$. Each sample path is 100 years long unless otherwise specified. The median Newey-West t-statistic is also reported, where the number of lags equals the number of overlapping observations. The regressions use quarterly simulated data, so for the annual forecasting horizon there are 3 lags used. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. Finally, the median R^2 is also reported for each regression. The "data" columns are taken from Beeler and Campbell (2011), who use U.S. data from 1930 to 2008.

	Data		**Median model outcomes**		
Forecasting horizon	β^{data}	$\bar{R}^{2,\text{data}}$	β^{model}	$(s.e.)$	$\bar{R}^{2,\text{model}}$
Panel A: Consumption growth predictability					
Consumption growth vs. P/C-ratio: $\Delta c_{t,t+j} = \alpha + \beta pc_t + \varepsilon_{t,t+j}$					
1 year	0.01	6.0%	0.10	(0.08)	1.4%
5 years	-0.00	0.0%	0.21	(0.33)	4.2%
Consumption growth vs. risk-free rate: $\Delta c_{t,t+j} = \alpha + \beta r_{f,t} + \varepsilon_{t,t+j}$					
1 quarter	-0.12	not reported	-0.06	(0.30)	0.0%
Panel B: Excess return predictability					
Excess returns vs. P/C-ratio: $r_{t,t+j} - r_{f,t,t+j} = \alpha + \beta pc_t + \varepsilon_{t,t+j}$					
100 year sample medians:					
1 year	-0.09^{*}	4.4%	-0.26	(0.18)	1.9%
5 years	-0.41^{***}	26.9%	-0.99	(0.66)	7.3%
50 year sample medians, conditioning on structural break at beginning of sample					
1 year	-0.09^{*}	4.4%	-0.51^{**}	(0.24)	5.9%
5 years	-0.41^{***}	26.9%	-1.88^{***}	(0.69)	20.7%

Figure 1 - Posterior Standard Deviation

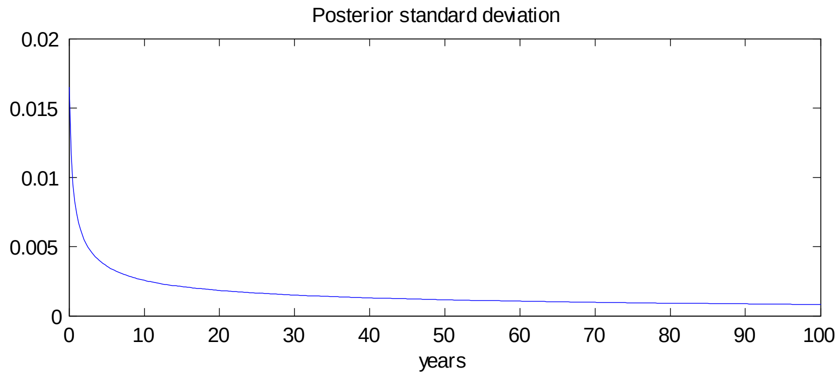

Figure 1: The graph shows the posterior standard deviation on the vertical axis and the time elapsed in years on the horizontal axis. The initial prior standard deviation is 0.0165.

Figure 2 - Ex ante vs ex post return predictability

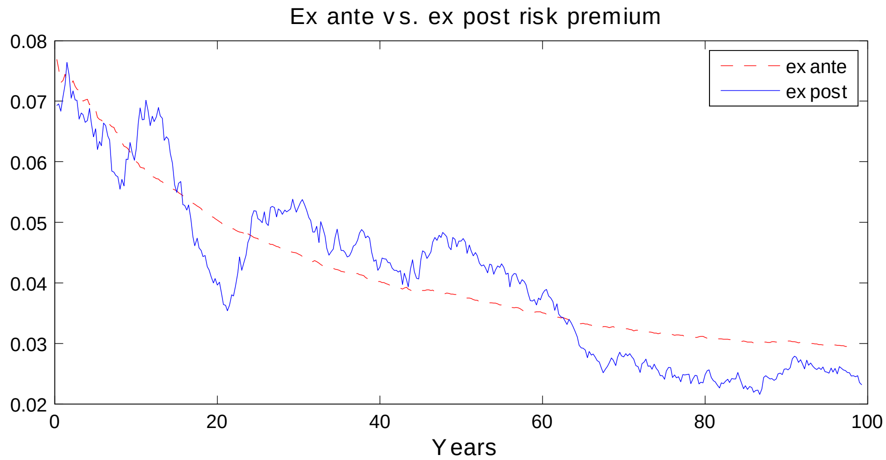

Figure 2: The graph shows a representative example of a sample path of the ex ante subjective risk premium, and the ex post estimated risk premium obtained from a regression of annual excess returns on the lagged price-consumption ratio. The sample length is 100 years and the initial prior is unbiased with dispersion equal to 0.26%. The solid line shows the ex post estimate, while the dashed line shows the ex ante value.

Figure 3 - Conditional Volatility of the Pricing Kernel: Preference for early resolution of uncertainty

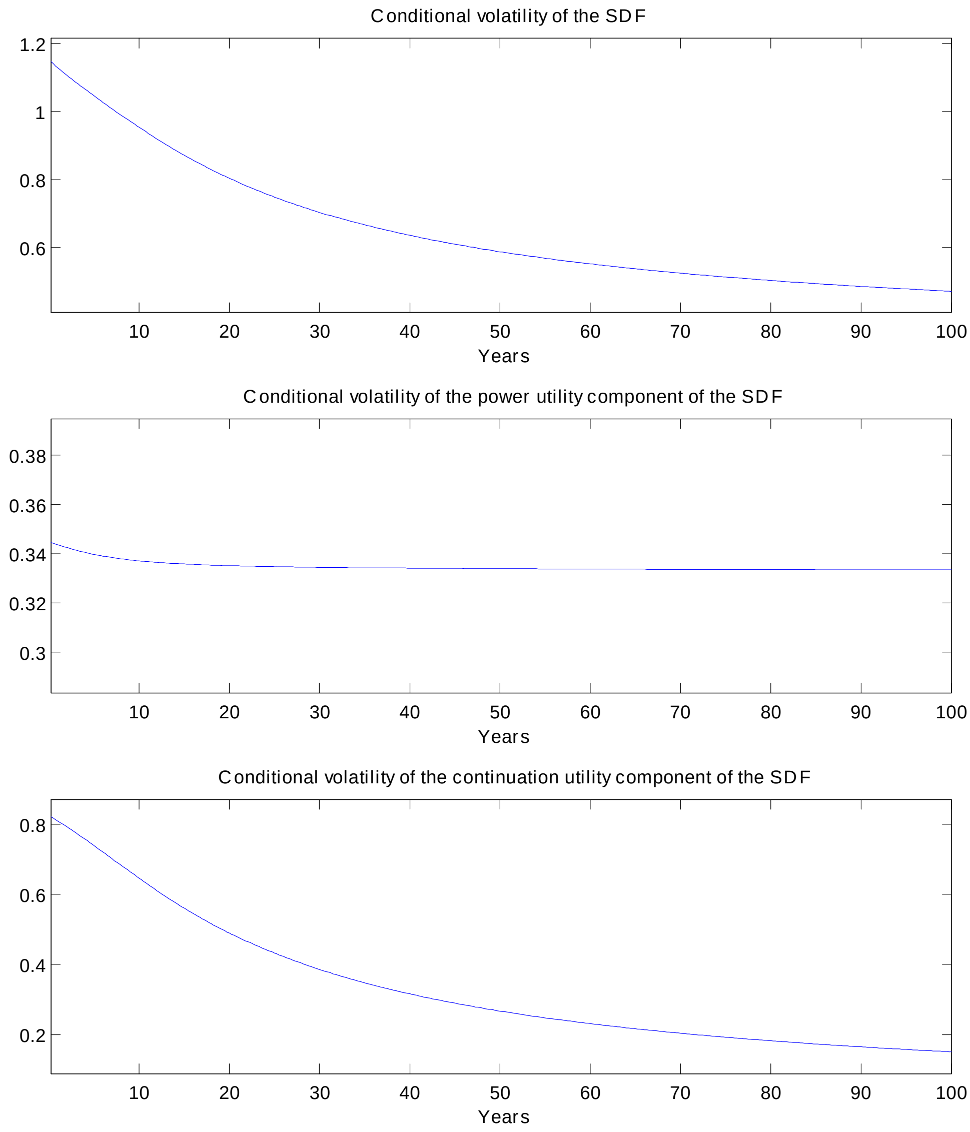

Figure 3: The top plot shows the subjective conditional annualized volatility of the Epstein-Zin stochastic discount factor with preference parameters $\gamma = 10, \psi = 2$ over a 100 year sample period. The plot shows the average conditional volatility across 20,000 simulated economies at each time t. The initial prior is unbiased with dispersion over mean consumption growth of $\sigma_{t=0} = 0.0165$. The middle plot shows the same for the "power utility component" of the stochastic discount factor $(\beta * exp(-\gamma \Delta c_{t+1}))$, while the bottom plot shows the conditional annualized volatility of the "continuation utility component" of the stochastic discount factor $((\beta \frac{PC_{t+1}+1}{PC_t})^{\theta-1})$.

Figure 4 - Sensitivity of log P/C ratio to changes in mean beliefs: Preference for early resolution of uncertainty

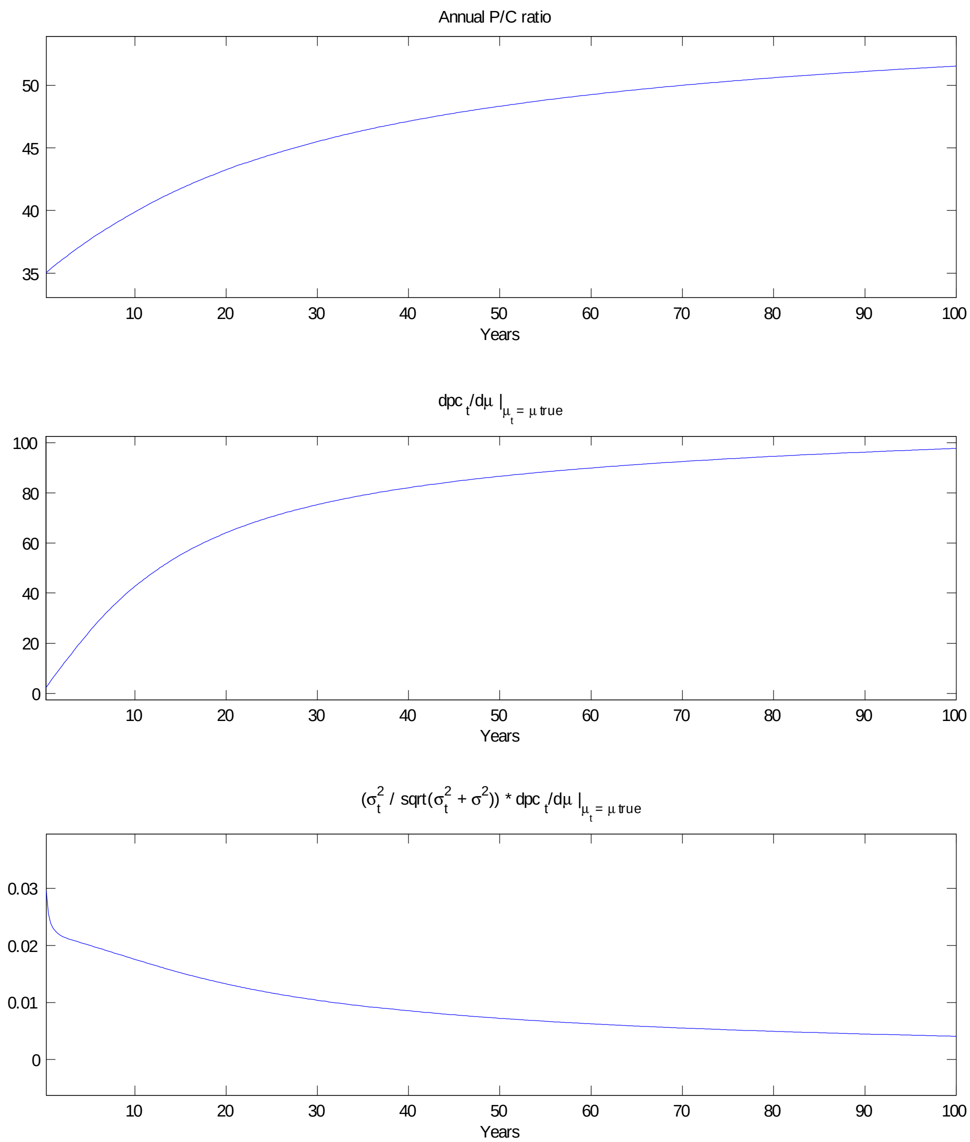

Figure 4: The top plot shows the annual wealth-consumption ratio (P/C) over a 100 year period, starting with an unbiased prior and initial dispersion $\sigma_{t=0} = 0.0165$. The plot gives the average outcome over 20,000 simulated economies with preference parameters $\gamma = 10, \psi = 2$. The middle plot shows the derivative of the log wealth-consumption ratio (pc) with respect to the mean beliefs about the consumption growth rate, evaluated at the true mean of consumption growth, versus years passed since the initial prior. The bottom plot shows the same derivative multiplied by the standard deviation of shocks to beliefs about the *mean* consumption growth rate, assuming a normal untruncated prior.

Figure 5 - Average conditional moments for case of learning about disaster probability

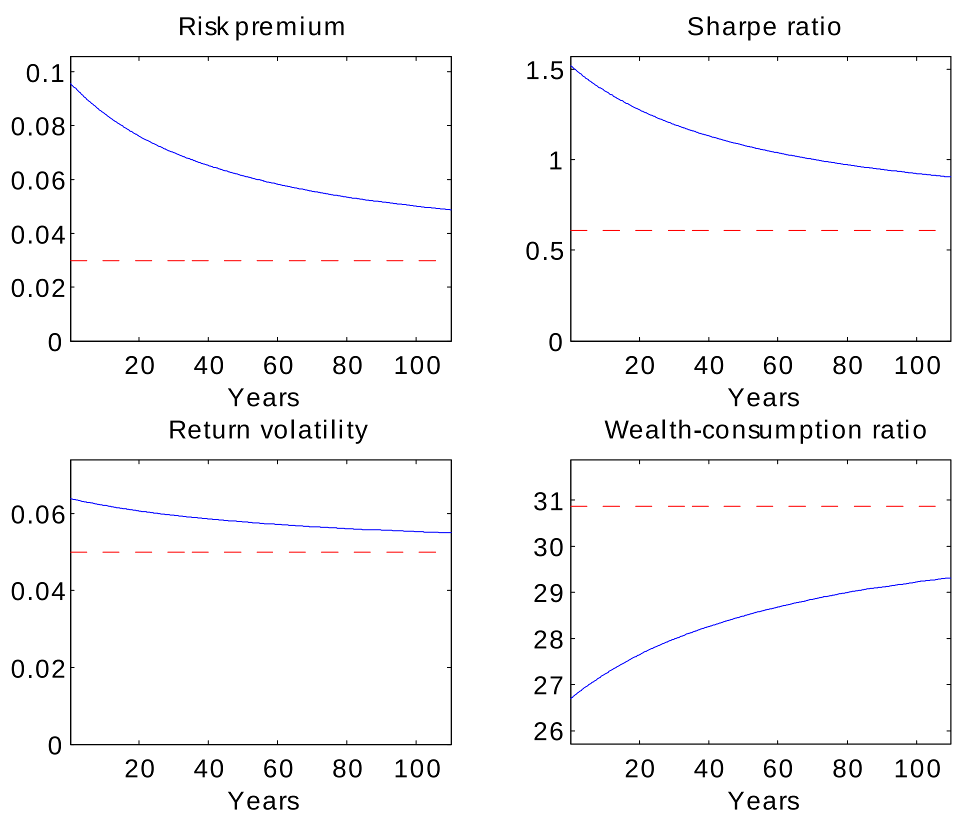

Figure 5: The figure shows annualized conditional asset pricing moments averaged across 20,000 simulated economies where the disaster probability is uncertain. Preference parameters $\gamma = 7, \psi = 2$.

Figure 6 - Sample path of the risk premium for case of learning about the disaster probability

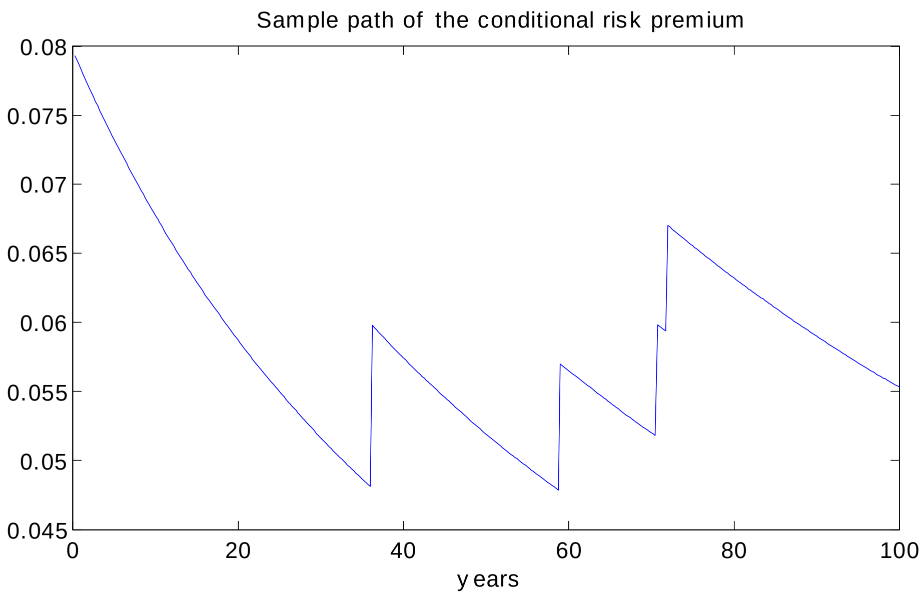

Figure 6: The figure shows a representative sample path of the annualized conditional risk premium from the model with learning about the disaster probability.

Figure 7 - Sample path of the risk premium and P/C ratio for structural break case

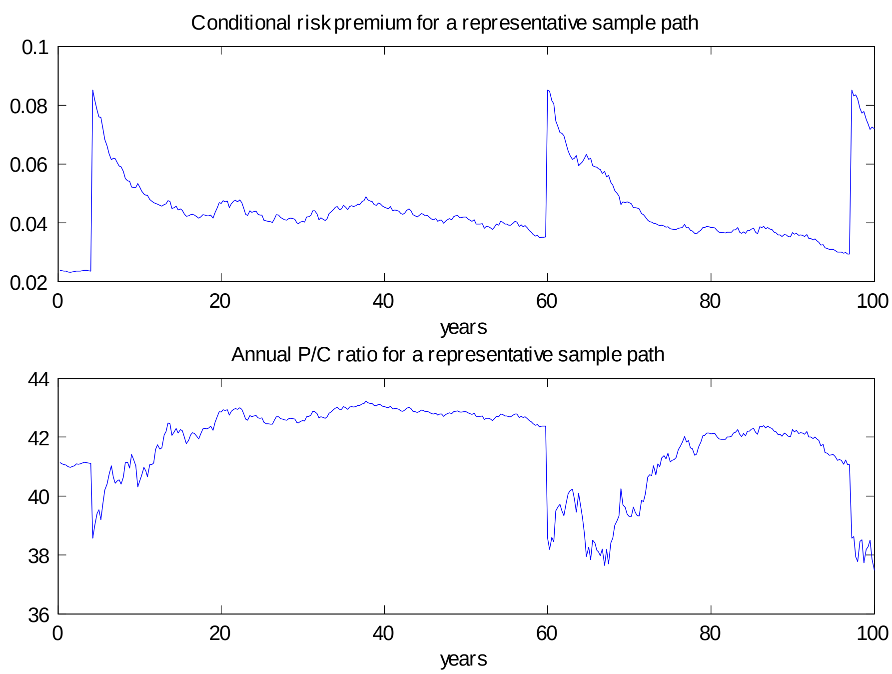

Figure 7: The figure shows a representative sample path of the annualized conditional risk premium in the top plot and the annual price-consumtion ratio in the bottom plot – both from the model with structural breaks and learning about the mean in each paradigm.