

Parameter learning, tail risks, and risk premia decomposition*

Mykola Babiak**

Lancaster University Management School

Roman Kozhan***

Warwick Business School

Abstract

We examine the decomposition of equity and variance premiums on the state space of market returns. Empirically, we show that low but non-disastrous returns are the primary drivers of both risk premia, while left-tail returns have a stronger influence on the variance premium. Theoretically, we propose a production economy with uncertain transition probabilities in two-state regime-switching productivity growth. Rationally pricing parameter uncertainty is crucial for explaining novel decompositions while simultaneously capturing salient features of index option prices, equity returns, variance, and macroeconomic quantities. Intuitively, rational pricing of beliefs amplifies the impact of shocks, generating physical tail risks in the economy.

Keywords: parameter learning, tail risk, production economy, equity premium, variance premium, SVIX, option-implied volatilities

JEL: D81, E32, E44, G12

*We would like to thank Anmol Bhandari, Daniele Bianchi, Nicola Fusari, Andrea Gamba, Liu Hening, Michal Kejak, Mete Kilic, Ctirad Slavik, and seminar participants at Warwick Business School, Università Ca' Foscari Venezia and CERGE-EI for their helpful comments. We thank the High End Computing facility at Lancaster University for support with the supercomputing clusters. Mykola Babiak received financial support from the Charles University Grant Agency - GAUK (grant number 744218).

**Lancaster University Management School, Lancaster LA1 4YX, UK, E-mail: m.babiak@lancaster.ac.uk.

***Warwick Business School, University of Warwick, Coventry CV4 7AL, UK, E-mail: roman.kozhan@wbs.ac.uk.

1 Introduction

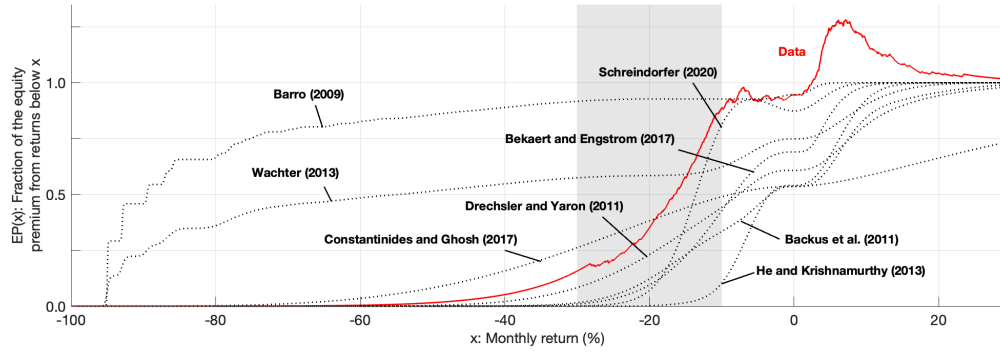
The macro-finance field has proposed several potential resolutions to major asset pricing puzzles in equity and derivatives markets. Yet, [Beason and Schreindorfer \(2022\)](#) demonstrate that leading theories are inconsistent with the shocks driving an equity premium in the data. By decomposing the equity premium on the market return state space, they find that most of it stems from negative but non-extreme returns. This novel decomposition challenges leading theories that attribute the premium to extreme tail or small negative returns. Our paper makes two key contributions. Empirically, we implement the decomposition of the variance premium and demonstrate that extreme and intermediate negative returns contribute more equally to the variance premium, though the latter returns remain dominant. Theoretically, we propose a production-based model with parameter uncertainty that explains these decompositions while capturing key macroeconomic moments, equity returns, and option prices.

We illustrate this challenge in Figure 1 by comparing empirical and model-based estimates of $EP(x)$ and $VP(x)$, which measure the fraction of the average equity and variance premiums attributable to returns below x .¹ The top panel shows that around three-quarters of the equity premium are associated with monthly returns between -30% and -10% , while returns below -30% account for less than one-sixth. The bottom panel demonstrates that intermediate and left-tail returns contribute slightly more than half and a third of the variance premium. Puzzling, however, is that well-known theories struggle to explain the substantial contribution of large negative but non-disastrous returns to the equity premium and the relatively balanced contributions of negative returns to the variance premium.

We capture the observed equity and variance premia decompositions by introducing parameter learning into a standard production economy. The model with known parameters

¹Following [Beason and Schreindorfer \(2022\)](#), we consider the models featuring habits ([Campbell and Cochrane, 1999](#)), long-run risks ([Bansal and Yaron, 2004](#)), disaster risk ([Rietz, 1988](#); [Barro, 2006](#)), undiversifiable idiosyncratic risk ([Constantinides and Ghosh, 2017](#)), constrained intermediaries ([He and Krishnamurthy, 2013](#)), and mixture shocks with generalized disappointment aversion preferences ([Schreindorfer, 2020](#)).

Panel A: Equity premium decomposition



Panel B: Variance premium decomposition

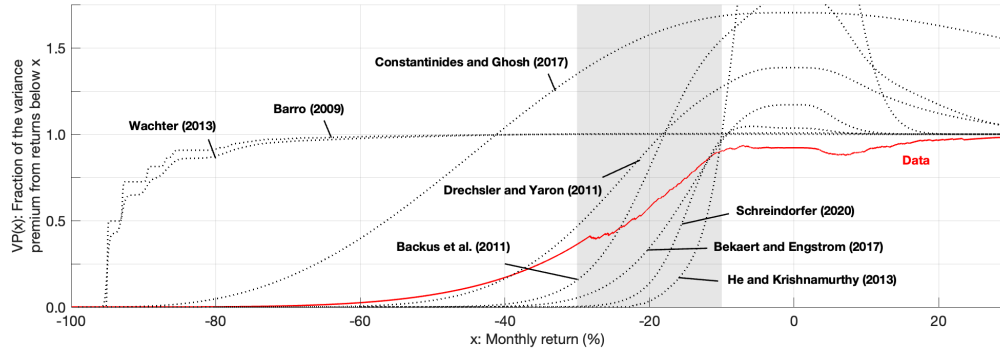


Figure 1. $EP(x)$ and $VP(x)$ in the data and models.

The figure shows the empirical and model-based $EP(x)$ (Panel A) and $VP(x)$ (Panel B) curves for a monthly horizon. The shaded area denotes monthly returns between -30% and -10% . The empirical curves are based on the data from January 1996 to December 2019.

generates abysmal risk premiums that are driven by small returns. Meanwhile, fully rational pricing of parameter uncertainty amplifies the impact of shocks on marginal utility and asset prices (Collin-Dufresne, Johannes, and Lochstoer, 2016; Babiak and Kozhan, 2024). Furthermore, the production economy allows for the endogenous feedback from productivity innovations to consumption through parameter beliefs and investment (Kozlowski, Veldkamp, and Venkateswaran, 2019, 2020), strengthening the amplification effects. Quantitatively, we demonstrate that this mechanism generates the realistic decompositions of equity and variance premiums while matching a wide array of stylized facts in equity and options markets.

Formally, we consider a real business cycle model with unknown parameters in a parsimonious two-state Markov switching process for aggregate productivity growth. A

Bayesian investor learns about the mean durations (e.g., transition probabilities) of regimes and rationally prices parameter uncertainty in equilibrium, as in [Collin-Dufresne et al. \(2016\)](#). Besides convex capital adjustment costs, we add no extra rigidities to the production model. In the interest of a convenient inspection of the mechanism, we also price exogenous dividends that enable us to generate procyclical cash-flow dynamics while isolating the impact of parameter learning from other forces.

The economic mechanism is as follows. Bayesian learning produces time-varying beliefs that create a channel through which macroeconomic shocks affect equilibrium utility and asset prices. The rational pricing of uncertainty amplifies the impact of belief updating. This creates additional risks in models with learning, leading to larger declines in consumption and higher agent’s marginal utility during bad times. Hence, the investor is willing to pay a large premium to hedge pessimistic beliefs. Variance swaps and out-of-the-money put options on the stock market index pay off in states of high realized variance and low valuations, which are associated with pessimistic beliefs and high marginal utility. Thus, variance swaps and puts earn high prices that transmit to the larger risk premia in a parameter learning setting compared to a full information case. We show that the distribution of returns associated with amplified risk premia is close to the empirical one.

It is noteworthy that the model is calibrated to match the moments of macroeconomic quantities and equity returns, while the risk premia decompositions and option-related results are not directly targeted. Nevertheless, a single shock model with priced parameter uncertainty can generate realistic unconditional moments of the variance premium and the level, skew, and smirk patterns in option-implied volatilities. It does so with empirically consistent dynamic properties of conditional equity variance, including its large volatility, moderate persistence, high skewness, and kurtosis. Finally, the model closely matches the salient features of the standard VIX index, the simple VIX index (SVIX), and the gap between the two. This contrasts leading representative-agent models that cannot match the behaviors of VIX and SVIX quantitatively or even qualitatively ([Martin, 2017](#)).

To better understand the mechanism, we look at conditional moments of consumption

and dividends. Endogenous consumption exhibits similar mean growth in expansions in models with known or unknown parameters but experiences markedly bigger declines in recessions in the parameter learning specification. In addition, consumption and dividends appear to be more correlated in bad states. The combination of the two effects leads to amplified physical tail risks in the economy with priced parameter uncertainty. We demonstrate that the model-implied declines in consumption are comparable to the U.S. experience during the Great Depression and certainly more conservative than consumption disasters ([Rietz, 1988](#); [Barro, 2006](#)).² The downside correlations between consumption and dividends also align well with the empirical evidence.

The model with parameter learning also delivers additional testable implications for conditional risk premiums. The most important is that the variance premium declines over time as the amount of parameter uncertainty investors face in our model decreases with a longer learning period. Empirically, we document 25 and 30 percent declines in the variance premium over the 10-year periods ending in 2012 and 2022. Theoretically, the variance premium in the parameter learning framework decreases, on average, by 25 percent over 50 years. Although the model-implied drift is smaller compared to the data, additional forces are likely to contribute to the observed decay. Overall, we believe the downward trend predicted by the model and observed in the data supports our mechanism. This insight relating priced parameter uncertainty to the declining variance premium is novel evidence, which relates to a recent study on declining alphas of synthetic options strategies ([Dew-Becker and Giglio, 2024](#)) and earlier literature on the declining equity premium.³

It is worth emphasizing that Bayesian learning alone is insufficient to generate a strong amplification of risk premia, and one needs to rationally price parameter uncertainty. To illustrate this point, we consider anticipated utility pricing commonly used by prior litera-

²The model-implied consumption declines are consistent with the empirical evidence in [Backus, Chernov, and Martin \(2011\)](#) showing that option prices imply relatively small and frequent jumps in consumption dynamics, counter to extreme declines predicted by disaster models.

³Notable studies documenting the declining equity premium include [Blanchard, Shiller, and Siegel \(1993\)](#); [Jagannathan, McGrattan, and Scherbina \(2001\)](#), and [Fama and French \(2002\)](#). [Lettau, Ludvigson, and Wachter \(2008\)](#) propose a fall in macroeconomic risk to explain the observed pattern.

ture ([Kreps, 1998](#); [Piazzesi and Schneider, 2009](#)). The anticipated utility agent learns about unknown parameters over time but treats current mean beliefs as true parameter values in decision making, thus, ignoring parameter uncertainty in equilibrium. We show that using the anticipated utility to deal with parameter uncertainty predicts virtually identical results to a full information case.

Several observations guide the construction of our theoretical model. To rationalize the puzzling equity premium decomposition, [Beason and Schreindorfer \(2022\)](#) suggest connecting the stock market tails events with parameter uncertainty and bad states of the real economy. [Collin-Dufresne et al. \(2016\)](#) show a strong impact of priced parameter uncertainty on marginal utility in endowment models, whereas [Babiak and Kozhan \(2024\)](#) demonstrate that this effect is further amplified by endogenous investment in production models. Therefore, incorporating parameter learning into a production-based economy is a natural step to reconcile the sources of equity and variance premiums in the data.

Our empirical analysis extends the novel equity premium decomposition of [Beason and Schreindorfer \(2022\)](#) to the second-moment risk premium. Our theoretical mechanism is related to the model of [Schreindorfer \(2020\)](#) with correlated endowments and a generalized disappointment risk-averse investor. Our model, however, endogenously produces physical tail risks and employs priced parameter uncertainty as an amplification of risk prices instead of asymmetric preferences.

[Liu and Zhang \(2022\)](#) are the first to explain the variance premium and equity index options in a production economy. In their model, an ambiguity-averse agent does not observe a hidden state of the economy. Their paper, however, does not analyze the equity and variance premia decompositions, a key focus of our work. In addition, priced parameter uncertainty is a key mechanism in our model as opposed to ambiguity aversion preferences in their setting. Although both papers assume incomplete investor information, our inference problem is opposite to their learning process. They assume hidden states and known parameters of productivity growth, whereas the agent in our model observes states but is uncertain about structural parameters.

Our mechanism connects to other production models with imperfect information. [Ai \(2010\)](#) is an early example of a real business cycle economy with learning. [Jahan-Parvar and Liu \(2014\)](#) combine a state uncertainty with ambiguity aversion preferences. Both articles focus on only equity returns. [Andrei, Mann, and Moyen \(2019\)](#) and [Davis and Segal \(2022\)](#) show how learning improves the behaviour of investments and valuation ratios. [Kozlowski, Veldkamp, and Venkateswaran \(2019\)](#) and [Kozlowski, Veldkamp, and Venkateswaran \(2020\)](#) show that learning about the whole distribution of productivity shocks can explain long-lasting effects of tail events on economic activity. [Winkler \(2020\)](#) examines an alternative approach of learning about endogenous prices and show that it provides substantial amplification of asset prices and real activity. In contrast to our paper, the extant literature focuses on standard equity moments or the real economy. They consider neither the risk premia decompositions nor option prices. In addition, their mechanisms differ from priced parameter uncertainty, a key driver of our paper’s results.

Our paper is also related to exchange economies focusing on the variance premium and option prices. Leading examples include the extensions of endowment models with habit ([Du, 2010](#); [Bekaert, Engstrom, and Ermolov, 2020](#)), rare disasters ([Liu, Pan, and Wang, 2005](#); [Seo and Wachter, 2019](#)), and long-run risks ([Eraker and Shaliastovich, 2008](#); [Bollerslev, Tauchen, and Zhou, 2009](#); [Drechsler and Yaron, 2011](#); [Zhou and Zhu, 2014](#); [Drechsler, 2013](#); [Shaliastovich, 2015](#)). Our paper is distinct from this literature because it emphasizes the key role of rational pricing of parameter uncertainty. Further, the model parameters are estimated from the post-war data and feature business cycle risks rather than rare disasters or long-run risks. Finally, none of these papers jointly target a broad set of features of macroeconomic quantities, equity returns, and index option prices.

Our framework also connects to Markov switching models with Bayesian learning. Most closely, [Cogley and Sargent \(2008\)](#) examine learning about transition probabilities. [Weitzman \(2007\)](#) study uncertainty about variance parameters. [Johannes, Lochstoer, and Mou \(2016\)](#) extend their approach to a multi-parameter inference problem. These papers, however, employ anticipated utility and focus only on equity moments. A number of papers

(Pakos, 2013; Gillman, Kejak, and Pakos, 2015; Andrei, Carlin, and Hasler, 2019; Andrei, Hasler, and Jeanneret, 2019) document quantitatively significant effects of learning about hidden persistence on equity premiums. None of these models, however, consider the risk premia decomposition, the variance premium, and option prices. Another notable example is Benzoni, Collin-Dufresne, and Goldstein (2011) who introduce state uncertainty and rare jumps in persistence into the long-run risks model to explain pre- and post-1987 crash option prices. One important difference between this paper and ours is that we study parameter uncertainty instead of state uncertainty. We also consider learning about business cycle risks instead of peso events in growth rates. Finally, we focus on additional asset pricing moments that are not targeted in their paper.

2 A decomposition of the equity and variance premia

This section presents the decomposition of the equity and variance premia on the market return state space. While Beason and Schreindorfer (2022) focus on the equity premium, we take a similar approach to examine the variance premium. The goal is to determine the contribution of market return states (e.g., the market return falling within an interval $[R, R + dR]$) to risk premia. Specifically, we estimate the functions $EP(x)$ and $VP(x)$, which measure the fraction of the average equity and variance premiums associated with returns below x . This section presents the empirical decompositions and compares them with those predicted by leading asset pricing models.

2.1 A state-space decomposition

We first decompose the equity premium following the procedure in Beason and Schreindorfer (2022).⁴ Let R_{t+1}^{cum} be the cum-dividend market return and $f_t(R)$ be its conditional probability density function (PDF) under a physical measure. No arbitrage assumption

⁴In our empirical analysis, we use the original codes of Beason and Schreindorfer (2022) and retrieve the options data from OptionMetrics. We thank the authors for making the replication code package publicly available. The replication files can be accessed via the following link:
https://www.journals.uchicago.edu/doi/suppl/10.1086/720396/suppl_file/20200045data.zip.

implies that there exists a risk-neutral density function $f_t^*(R)$ such that the risk-free rate can be written as $R_t^f = \int_{-1}^{\infty} R f_t^*(R) dR$, and hence, the equity premium equals to:

$$E_t [R_{t+1}^{cum}] - R_t^f = \int_{-1}^{\infty} R [f_t(R) - f_t^*(R)] dR. \quad (1)$$

Next, the absence of arbitrage implies that for any arbitrary (twice-differentiable) function $g(R)$, the forward price of the claim with payoff $g(R)$ is given by $E_t^*[g(R)]$ and the premium associated with the buy-and-hold strategy is $E_t[g(R)] - E_t^*[g(R)]$. We capture the variance premium by considering $g(R) = (\ln(1 + R))^2$, which corresponds to a quadratic swap strategy in [Kozhan, Neuberger, and Schneider \(2013\)](#).⁵ Thus, we can approximate the variance premium as:

$$E_t [g(R_{t+1}^{cum})] - E_t^* [g(R_{t+1}^{cum})] = \int_{-1}^{\infty} g(R) [f_t(R) - f_t^*(R)] dR. \quad (2)$$

In our analysis, we focus on unconditional premiums to avoid difficulties in estimating the conditional density functions. One can achieve this by taking the unconditional expectations in Equations (1) and (2) and defining the premia decompositions as:

$$EP(x) \equiv \frac{\int_{-1}^x R [f(R) - f^*(R)] dR}{\int_{-1}^{\infty} R [f(R) - f^*(R)] dR}, \quad (3)$$

$$VP(x) \equiv \frac{\int_{-1}^x (\ln(1 + R))^2 [f(R) - f^*(R)] dR}{\int_{-1}^{\infty} (\ln(1 + R))^2 [f(R) - f^*(R)] dR}, \quad (4)$$

where the unconditional density functions are given as $f(R) = E[f_t(R)]$ and $f^*(R) = E[f_t^*(R)]$. The normalization in the denominators in Equations (3) and (4) ensures that $EP(x)$ and $VP(x)$ approach zero when x goes to -1 and approach one for large values of x .

Moreover, the functions increase for states that contribute positively to the corresponding

⁵The premium captured by this choice of function g is the second non-central moment rather than a pure variance risk premium. Using the second non-central moment makes it possible to transition from the conditional premium in Equation (1) to the unconditional premium in Equation (4). Furthermore, as shown in [Kozhan, Neuberger, and Schneider \(2013\)](#), the variance risk premium defined by means of a function g is almost identical empirically to the traditional variance risk premium based on the VIX measure.

premium and decrease for states that contribute negatively.

Beason and Schreindorfer (2022) point out the potential mismatch between the physical and risk-neutrals densities. The former defines the probability distribution of the cum-dividend returns, while option prices identify state prices for ex-dividend returns. Following their procedure, we overcome this issue by assuming that $(t + 1)$ -dividends are in investors' time- t information set. Thus, we can estimate the decompositions as functions of the ex-dividend return because the dividends will drop out in Equations (3) and (4).⁶

2.2 Empirical $EP(x)$ and $VP(x)$

Empirically, we use the S&P 500 index as a proxy for the aggregate stock market and focus on a one-month return horizon. The availability of options data from OptionMetrics limits the start date of our sample to January 1996. We perform the empirical analysis using the data until December 2019 to eliminate the impact of the COVID period. We sample daily to maximize the efficiency of our estimates, and hence, our sample comprises $T = 6,037$ overlapping 30 calendar day returns. We estimate the unconditional probability density function $f(R)$ using historical returns. Specifically, we calculate average payoffs $f(R) = E[1\{R < R_{t:t+30} < R + dR\}]$ where each one-month return observation $R_{t:t+30}$ has a probability of $1/T$. For robustness checks, we also estimate the smoothed PDF f^{smooth} . We then use market index options to obtain the risk-neutral probability density function $f_t^*(R)$. We estimate $f_t^*(R)$ daily using Breeden and Litzenberger (1978) method and calculate the unconditional density as $f^*(R) = \frac{1}{T} \sum_{t=1}^T f_t^*(R)$.

Panel A of Figure 1 closely replicates the $EP(x)$ curve in Beason and Schreindorfer (2022). The far left tail ($R < -30\%$) contributes about 0.15 to the equity premium, whereas the intermediate left tail ($-30\% < R < -10\%$) accounts for 0.65. Thus, most of the equity premium represents compensation for intermediate negative returns. Panel B of Figure 1 shows the resulting $VP(x)$ curve. Over a third of the variance premium represents

⁶We effectively present the decomposition on the state space of ex-dividend returns. Although the equity and variance premiums are defined in terms of cum-dividend returns, our assumption is inconsequential for comparing model-based and empirical $EP(x)$ and $VP(x)$ because both employ ex-dividend returns.

Table 1. Summary statistics of $EP(x)$ and $VP(x)$.

The table reports the contributions of different regions to the equity and variance premiums for a monthly horizon. It also shows the probability of returns in the intermediate region and the ratio of their risk-neutral to physical probabilities. The table reports the empirical moments when $f(R)$ is proxied by realized returns (the f_R column) or is a smoothed version (the f_R^{smooth} column) of the empirical PDF. The last five columns report selected percentiles of the estimates based on a block bootstrap with a block size of 21 trading days. The data is from January 1996 to December 2019.

	f_R	f_R^{smooth}	Percentiles				
			2.5%	5%	50%	95%	97.5%
$EP(-.3)$	0.157	0.155	0.085	0.092	0.154	0.373	0.497
$EP(-.1) - EP(-.3)$	0.728	0.709	0.465	0.499	0.738	1.458	1.845
$1 - EP(-.1)$	0.115	0.136	-1.320	-0.813	0.107	0.394	0.432
$VP(-.3)$	0.362	0.341	0.268	0.279	0.347	0.466	0.501
$VP(-.1) - VP(-.3)$	0.541	0.545	0.364	0.410	0.567	0.665	0.682
$1 - VP(-.1)$	0.097	0.113	-0.021	-0.001	0.089	0.171	0.189
$\int_{-0.3}^{-0.1} f(R) dR$	0.015	0.023	0.010	0.011	0.019	0.028	0.030
$\int_{-0.3}^{-0.1} f^*(R) dR / \int_{-0.3}^{-0.1} f(R) dR$	3.121	2.087	1.633	1.741	2.550	4.190	4.694

compensation for extremely negative returns. The contribution of the intermediate left tail is about 0.54, substantially smaller than in the case of the equity premium. In both decompositions, the contribution of returns above -10% is around 0.1. In sum, the variance premium decomposition demonstrates a dominant effect of intermediate market declines and a more significant role of more extreme returns.

Table 1 presents summary statistics of $EP(x)$ and $VP(x)$. Consistent with average values, the sampling distribution of $VP(-0.3)$ is shifted to the right side compared to $EP(-0.3)$. For instance, $VP(-0.3)$ exceeds 0.25 in more than 95% of samples, whereas $EP(-0.3)$ is above 0.25 in fewer than 50% of simulations. Consequently, $EP(-0.1) - EP(-0.3)$ exceeds 0.5 in almost 95% of bootstrap samples, $VP(-0.1) - VP(-0.3)$ is above 0.5 slightly more than 50% cases.

Figure 2 plots the joint sample distribution of $EP(-0.3)$ and $EP(-0.1) - EP(-0.3)$ as

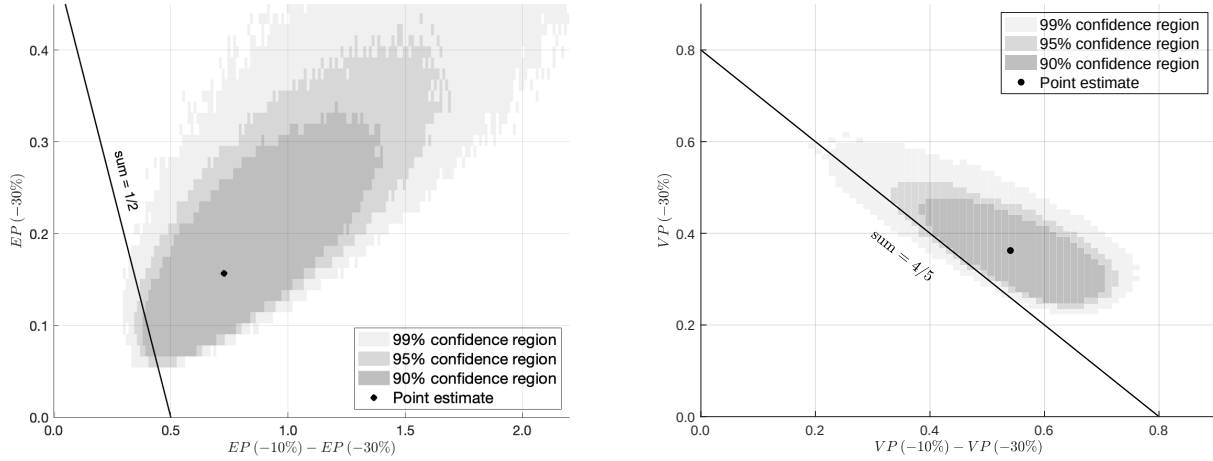


Figure 2. Sampling distribution of $EP(x)$ and $VP(x)$.

The figure shows the contributions of the two regions — monthly returns below -30% and between -30% and -10% — to the equity and variance premiums in the 1 million bootstrap samples. For the sampling distribution, we implement a block bootstrap with a block size of 21 trading days. The length of the artificially simulated data corresponds to the length of the empirical data from January 1996 to December 2019. The shaded areas mark the confidence regions, whereas the solid black line denotes the contributions to the equity (variance) premium with a sum of 0.5 (0.8).

well as $VP(-0.3)$ and $VP(-0.1) - VP(-0.3)$. Two comments are noteworthy. First, the negative returns — either intermediate or large declines — account for the bigger part of both premia in the data. For instance, the contribution of the returns below -10% is above one-half of the equity premium and four-fifths of the variance premium in more than 99% of all simulations. Second, returns below -30% account for a much smaller share of the equity premium than returns between -30% and -10% , whereas the two regions tend to have more equal shares. Thus, the key findings cannot be attributed to sampling error.

Finally, the last two rows in Table 1 report the probability and risk price of returns from -30% to -10% . The average statistics based on historical data are 0.015 and 3.121. It is noteworthy that the confidence bounds are quite wide for both estimates. Once the empirical PDF is smoothed, the probability of intermediate market returns increases to 0.023, leading to a substantial decline in the price of risk to 2.087.

2.3 Representative agent models

We now examine how well different representative agent models can explain the empirical $EP(x)$ and $VP(x)$. We consider the models from different streams of the literature: the rare disaster model of [Barro \(2006\)](#) and its extensions with the time-varying probability of disasters by [Wachter \(2013\)](#) and the “frequent disasters” by [Backus et al. \(2011\)](#), the external habit model of [Bekaert and Engstrom \(2017\)](#), the long-run risks model of [Drechsler and Yaron \(2011\)](#), the model with undiversifiable idiosyncratic risk of [Constantinides and Ghosh \(2017\)](#), the intermediary based model of [He and Krishnamurthy \(2013\)](#) and the GDA model with mixture shocks of [Schreindorfer \(2020\)](#).

Figure 1 compares $EP(x)$ and $VP(x)$ in the data and leading asset pricing models. Similarly to the evidence presented in [Beason and Schreindorfer \(2022\)](#), only the model with disappointment aversion preferences can partially capture the key patterns of the equity premium decomposition. However, the $VP(x)$ curves in all models differ from the empirical counterpart more substantially. The frameworks of [Backus et al. \(2011\)](#), [He and Krishnamurthy \(2013\)](#), [Bekaert and Engstrom \(2017\)](#), and [Schreindorfer \(2020\)](#) underestimate the contribution of larger (in magnitude) negative returns, whereas the rare disaster models associate the variance premium with extreme market declines. Moreover, all mechanisms except those using rare disasters and disappointment aversion preferences attribute a significant negative variance premium to positive returns. In unreported results, we verify that none of these models can capture a non-zero contribution of returns below -30% to the variance premium jointly with a significant role of returns between -30% and -10% and a negligible contribution of returns above -10% .

In their Section V, [Beason and Schreindorfer \(2022\)](#) conjecture two channels that could rationalize the sources of equity and variance premia in the data. First, they suggest parameter uncertainty as an amplification mechanism of shocks on asset prices. Second, they suggest that stock market tail events might coincide with bad states of the real economy and might not be a feature of investors’ preferences. Motivated by these conjectures, we

incorporate parameter learning into a real business cycle model to reconcile the sources of equity and variance premiums in the data.

3 The production economy

We consider a standard production economy with a representative household and firm. This section presents the model, the equilibrium prices, and a sketch of the numerical solution. The appendix provides a detailed description of the solution methodology.

3.1 Household, firm, and technology

The household has recursive utility of [Epstein and Zin \(1991\)](#) defined as:

$$U_t = \left\{ (1 - \beta)C_t^{1-1/\psi} + \beta \left(E_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}}, \quad (5)$$

where U_t is the continuation utility, C_t is consumption, $\beta \in (0, 1)$ is the discount factor, $\psi > 0$ is the elasticity of inter-temporal substitution, and $\gamma > 0$ is the risk aversion parameter. The firm produces the output using a constant return to scale Cobb-Douglas production function:

$$Y_t = K_t^\alpha (A_t N_t)^{1-\alpha},$$

where Y_t is the output, K_t is the capital, N_t is labor hours, and A_t is an exogenous, labor-enhancing technology level. We assume that the representative household supplies the fixed number of labor hours, $N_t = 1$. This simplification considerably reduces the computational costs and is inconsequential for the economic intuition of the results.

The firm's capital accumulation equation incorporates capital adjustment costs:

$$K_{t+1} = (1 - \delta)K_t + \varphi(I_t/K_t)K_t,$$

where $\delta \in (0, 1)$ is the capital depreciation rate and $I_t = Y_t - C_t$ is a gross investment. $\varphi(\cdot)$

is a concave function capturing adjustment costs $\varphi(x) = a_1 + \frac{a_2}{1-1/\xi} x^{1-1/\xi}$, where ξ is the elasticity of the investment rate. The lower value of ξ implies higher capital adjustment costs, while the extreme case of $\xi = \infty$ means that capital adjustment costs are zero.⁷

The log technology growth follows a two-state Markov switching model:

$$\Delta a_t = \mu_{s_t} + \sigma_{s_t} \cdot \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} N(0, 1), \quad (6)$$

where the mean μ_{s_t} and volatility σ_{s_t} of productivity growth depend on a two-state Markov chain $s_t \in \{1, 2\}$ with transition probabilities $\pi_{11} = 1 - \pi_{12}$ and $\pi_{22} = 1 - \pi_{21}$. Our choice of a parsimonious two-state model of productivity growth is motivated by a convenient inspection of the mechanism and the costly numerical model solution.

3.2 Equilibrium and asset prices

In equilibrium, the social planner chooses consumption and investment to maximize the household's utility subject to the resource constraint and the capital accumulation.⁸ The first-order condition of the household's maximization program implies the stochastic discount factor (SDF):

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-1/\psi} \left(\frac{U_{t+1}}{\left(E_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{1-\gamma}}} \right)^{1/\psi-\gamma}. \quad (7)$$

The recursive preferences separate the relative risk aversion and the elasticity of intertemporal substitution. We consider the household with a preference for early resolution

⁷Following [Boldrin, Christiano, and Fisher \(2001\)](#), we choose a_1 and a_2 such that there are no adjustment costs in the non-stochastic steady state. Specifically, $a_1 = \frac{1}{\xi-1} (1 - \delta - \exp(\bar{\mu}))$, $a_2 = (\exp(\bar{\mu}) - 1 + \delta)$, where $\bar{\mu}$ is the unconditional mean of the productivity growth rate. We find state values of remaining quantities from the conditions $\varphi\left(\frac{I}{K}\right) = 1$, $\varphi'\left(\frac{I}{K}\right) = 1$. In particular, the steady-state investment-capital ratio is $\frac{I}{K} = \exp(\bar{\mu}) - 1 + \delta$.

⁸Following a standard approach, the economy can also be decentralized: the household works for the firm and trades its shares and risk-free bonds to maximize the lifetime utility over a consumption stream, while the firm chooses investment to maximize its value, the present value of future cash flows.

of uncertainty by setting $\gamma > \frac{1}{\psi}$. This calibration is crucial for our results because belief updates will be priced in the equilibrium. The real risk-free rate is $1/R_t^f = E_t [M_{t+1}]$.

The first-order condition of the firm's maximization problem implies that

$$E_t [M_{t+1} R_{j,t+1}] = 1 \quad (8)$$

for any gross return $R_{j,t+1}$. In particular, this equation holds for the investment return, which can be interpreted as the return of a claim to the unlevered firm's payouts. Before advancing further, it is important to emphasize that a frictionless production economy fails to capture the firm's procyclical dividends, resulting in a low equity premium and variance. This countercyclical is driven by investment, which is more volatile than output and, hence, capital's share of output. A shock that increases capital will also cause higher investment. Since dividends net out investment expenditure from capital's share, they are countercyclical unless the model is modified. Although extending the model to generate procyclical dividends is possible, we price the calibrated dividends to provide a convenient interpretation of the economic mechanism and to isolation the effect of parameter learning from other mechanisms in the extended model.⁹

Following [Ai \(2010\)](#), [Kaltenbrunner and Lochstoer \(2010\)](#), and [Liu and Zhang \(2022\)](#), we define the process of aggregate dividends in the production economy as:

$$\Delta d_t = g_d + \lambda \Delta c_t + \sigma_d \cdot \eta_t, \quad \eta_t \stackrel{\text{iid}}{\sim} N(0, 1), \quad (9)$$

where λ is a leverage factor, g_d and σ_d are the dividend growth rate and volatility. We choose g_d and σ_d to match the first and second moments of annual dividend growth. We further set λ consistent with the extant literature to compare our results with those reported in other studies. Then, the equity return $R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_{t+1}} = \frac{1 + P_{t+1}/D_{t+1}}{P_t/D_t} e^{\Delta d_{t+1}}$ should satisfy Equation (8). We next define the main asset prices of interest.

⁹[Uhlig \(2007\)](#), [Belo, Lin, and Bazdresch \(2014\)](#), and [Favilukis and Lin \(2016\)](#) introduce wage rigidity to generate more volatile and procyclical dividends. [Babiak and Kozhan \(2024\)](#) demonstrate that financial leverage and asymmetric adjustment costs can also change the cyclicity of dividends.

Variance premium and implied volatilities. The variance premium at time t is $vp_t = VIX_t^2 - VOL_t^2$, where VIX_t^2 and VOL_t^2 are expectations of return variance under the risk-neutral $\mathbb{Q}$ and physical $\mathbb{P}$ probability measures. Following [Bollerslev, Tauchen, and Zhou \(2009\)](#) and [Carr and Wu \(2009\)](#), we define VIX_t^2 and VOL_t^2 as:

$$VIX_t^2 = \mathbb{E}_t^{\mathbb{Q}} [var_{t+1}(r_{t+2})] = \mathbb{E}_t \left[\frac{d\mathbb{Q}}{d\mathbb{P}} \cdot var_{t+1}(r_{t+2}) \right] \quad \& \quad VOL_t^2 = \mathbb{E}_t [var_{t+1}(r_{t+2})],$$

where $\mathbb{E}_t$ is the conditional expectation under the physical measure, r_{t+2} is the log equity return, $var_{t+1}(r_{t+2}) = \mathbb{E}_{t+1} [r_{t+2}^2] - [\mathbb{E}_{t+1} [r_{t+2}]]^2$, and $\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{M_{t+1}}{\mathbb{E}_t(M_{t+1})}$ is the Radon-Nykodim derivative. Two comments are noteworthy. First, the literature has adopted several definitions of the variance premium in discrete-time models. Section 5.3 verifies that our findings are robust to alternative specifications. Second, the definition of the variance premium is inconsequential for the variance premium decomposition in the data and the model because we only require realized returns and options price for the latter.

We also price European put options on the index. Using the equilibrium condition (8), the relative price $\mathcal{O}_t(K) = \frac{P_t^o(K)}{P_t(\pi_t)}$ of the one-quarter European put option with the strike price K , expressed as a ratio to the price of the equity P_t , should satisfy

$$\mathcal{O}_t(K) = \mathbb{E}_t \left[M_{t+1} \cdot \max \left(K - \frac{P_{t+1}}{P_t}, 0 \right) \right] = \mathbb{E}_t \left[M_{t+1} \cdot \max \left(K - \frac{P_{t+1}/D_{t+1}}{P_t/D_t} e^{\Delta d_{t+1}}, 0 \right) \right]. \quad (10)$$

We numerically evaluate and convert put prices into Black-Scholes implied volatilities.¹⁰

SVIX and entropy of the risk-neutral return distribution. Many recent models with long-run risks and rare disasters have been proposed to explain the large variance premium and the implied volatility skew. These option-related premiums are associated with a high degree of non-normality of the risk-neutral distribution of the market return. [Martin \(2017\)](#)

¹⁰Let r_t^f denote the annualized continuous interest rate, q_t is the dividend yield, $\tau = 1/4$ is the maturity time. The one-quarter implied volatility $\sigma_t^{\text{BS}} = \sigma_t^{\text{BS}}(K)$ solves the equation:

$$\mathcal{O}_t = e^{-r_t^f \tau} \cdot K \cdot N(-d_2) - e^{-q_t \tau} \cdot N(-d_1), \quad d_{1,2} = \left[-\ln(K) + \tau \left(r_t^f - q_t \pm (\sigma_t^{\text{BS}})^2 \right) \right] / [\sigma_t^{\text{BS}} \sqrt{\tau}].$$

shows that the difference between the entropy of the risk-neutral return distribution and the simple VIX index can alternatively measure the non-log-normality of the market return. Furthermore, the author shows that equilibrium models, which can successfully explain option prices, still have difficulty matching the difference between the risk-neutral and simple VIX series.¹¹ Adopting Equations (13) and (25) in [Martin \(2017\)](#), we define

$$SVIX_t = \frac{1}{R_t^f} \sqrt{\text{var}_t^Q(R_{t+1}^{ex})} \quad \& \quad S^Q = \sqrt{2\mathcal{L}_t^Q(R_{t+1}^{ex})} \quad (11)$$

where $R_{t+1}^{ex} = \frac{P_{t+1}}{P_t}$ is the simple ex-dividend return, $\text{var}_t^Q(x) = \mathbb{E}_t^Q[x^2] - [\mathbb{E}_t^Q[x]]^2$ and $\mathcal{L}_t^Q(x) = \ln E_t^Q[x] - E_t^Q[\ln x]$ denote the risk-neutral variance and entropy of x .

3.3 Numerical model solutions

The equilibrium utility and pricing ratios do not admit a closed-form solution in models with learning or full information. In this section, we discuss the numerical solution methods for models with known parameters and priced parameter uncertainty. A detailed description of methodologies and delegated to the Internet Appendix.

We reformulate the household's maximization problem in terms of stationary variables:

$$\tilde{U}_t = \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1 - \beta) \tilde{C}_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1}^{1-\gamma} \cdot e^{(1-\gamma)\Delta a_{t+1}} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\} \quad (12)$$

$$\tilde{C}_t + \tilde{I}_t = \tilde{K}_t^\alpha \quad (13)$$

$$e^{\Delta a_{t+1}} \tilde{K}_{t+1} = (1 - \delta) \tilde{K}_t + \varphi \left(\frac{\tilde{I}_t}{\tilde{K}_t} \right) \tilde{K}_t \quad (14)$$

$$\Delta a_t = \mu_{s_t} + \sigma_{s_t} \varepsilon_t, \quad \varepsilon_t \sim N(0, 1), \quad (15)$$

where A_t is a common trend and $\{\tilde{C}_t, \tilde{I}_t, \tilde{Y}_t, \tilde{K}_t, \tilde{U}_t\} = \left\{ \frac{C_t}{A_t}, \frac{I_t}{A_t}, \frac{Y_t}{A_t}, \frac{K_t}{A_t}, \frac{U_t}{A_t} \right\}$. The solution

¹¹[Martin \(2017\)](#) demonstrates that the models of [Campbell and Cochrane \(1999\)](#), [Bansal and Yaron \(2004\)](#), [Bansal, Kiku, and Yaron \(2010\)](#), [Bollerslev, Tauchen, and Zhou \(2009\)](#), and [Drechsler and Yaron \(2011\)](#) generate too small spread, whereas a rare disaster model of [Wachter \(2013\)](#) produces too large gap.

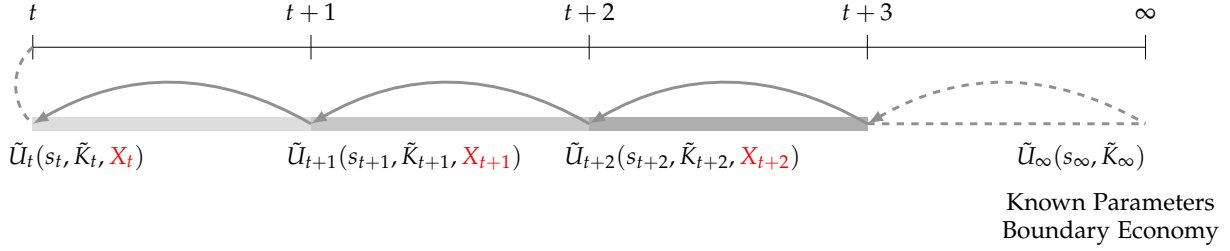


Figure 3. Priced parameter uncertainty.

The figure illustrates the agent's continuation utility in the production economy with priced parameter uncertainty. The equilibrium utility is the function of the state s_t , capital $\tilde{K}_t$, and sufficient statistics for unknown parameters X_t . To find $\tilde{U}_t = \tilde{U}_t(s_t, \tilde{K}_t, X_t)$ at time t , the agent uses the backward recursion starting from the known parameters boundary as shown by arrows in the diagram. The boundary economy, in turn, is solved assuming the agent knows the true parameters in the productivity growth process.

method for the model with known parameters is standard: we approximate the equilibrium utility as a function of the productivity growth regime and capital through value function iteration. The solution procedure for the model with priced parameter uncertainty consists of two steps illustrated in Figure 3. First, the household will have learned the true parameters when a long data history becomes available. Thus, the household begins solving the boundary economy with all known parameters. Second, the household applies the recursive equilibrium conditions to go backward and compute the utility starting from boundary conditions.

Here, we sketch the backward recursion in the second step. Let s_t denote the regime, $\tilde{K}_t$ is capital, and X_t is the vector of sufficient statistics for beliefs about unknown probabilities. We employ standard, conjugate prior distributions for unknown parameters in productivity growth. By assumption, the state s_t is observable. Thus, we can update $X_{t+1} = g(s_{t+1}, s_t, X_t)$ via standard Bayes' rule. Further, the law of motion of capital implies that $\tilde{K}_{t+1} = f(\Delta a_{t+1}, s_t, \tilde{K}_t, X_t)$. Given these, we can now write the equilibrium utility as:

$$\tilde{U}_{t+1}(s_{t+1}, \tilde{K}_{t+1}, X_{t+1}) = \tilde{U}_{t+1}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t)$$

to indicate that the utility evolution is the function of the two observable variables, s_{t+1}

and Δa_{t+1} . Using these notations, we can rewrite the recursive maximization as:

$$\begin{aligned} \tilde{U}_t(s_t, \tilde{K}_t, X_t) = \max_{\tilde{C}_t, \tilde{I}_t, N_t} & \left\{ (1 - \beta) \tilde{V}_t^{1 - \frac{1}{\psi}} \right. \\ & \left. + \beta \left(E_t \left[\tilde{U}_{t+1}^{1-\gamma}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, \tilde{K}_t, X_t \right] \right)^{\frac{1 - \frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} \end{aligned} \quad (16)$$

where the expectation on the right-hand side is equivalent to

$$\begin{aligned} & E_t \left[\tilde{U}_{t+1}^{1-\gamma}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, \tilde{K}_t, X_t \right] \\ = & \sum_{s_{t+1}=1}^2 E_t(\pi_{s_{t+1}, s_t} | s_t, X_t) E_t \left[\tilde{U}_{t+1}^{1-\gamma}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, \tilde{K}_t, X_t \right]. \end{aligned} \quad (17)$$

We have an analytical expression for the conditional expectation of transition probabilities and can use quadrature-type integration to evaluate the second expectation. The backward recursion is completely defined by Equations (12)-(17) and $X_t = g(s_{t+1}, s_t, \Delta a_{t+1}, X_t)$.

4 Quantitative analysis

This section calibrates the production economy, discusses an inference problem about unknown parameters, and presents the quantitative implications of the model.

4.1 Calibration

We choose the parameters in the model using the following guidelines. First, we estimate a two-state Markov switching process of productivity growth using the historical data from 1947:Q2 to 2019:Q4. Total factor productivity has experienced an extreme decline during the first two quarters of 2020. If we include those data points in the estimation, the recession state would feature disaster-like events. We refrain from doing so because we aim to study the impact of learning about business cycle dynamics on asset prices. Second, the subjective discount factor, the elasticity of intertemporal substitution, and relative risk

Table 2. Parameter values.

The table reports the parameter values in the production economy. It shows the maximum likelihood estimates of a two-state Markov-switching model for productivity growth, the preference parameters, the calibrated parameters in the production function, capital adjustment costs, and the dividend growth process. We obtain the estimates of the productivity process by applying the expectation-maximization algorithm (Hamilton, 1990) to quarterly total factor productivity growth rates from 1947:Q2 to 2019:Q4.

Productivity	π_{11}^a 0.968	π_{22}^a 0.723	μ_1^a 0.48%	μ_2^a −1.26%	σ_1^a 1.38%	σ_2^a 2.19%
Preferences	β^4 0.985	γ 8	ψ 1.5			
Capital	α 0.36	a_1 −0.001	a_2 0.8546	δ 0.02	ξ 24	
Dividends	λ 2.5	g_d −0.54%	σ_d 4.00%			

aversion are set to values that enable us to generate a low interest rate and a large equity premium. Third, all parameters in the production function and capital equation are set to the common values in the literature. Fourth, the parameters in calibrated dividends are chosen to match the first and second moments of dividends. Overall, all parameters are pinned down by empirical moments of macroeconomic quantities, dividends, and equity returns and are not selected to match option-related puzzles.

Table 2 reports the parameter values. Productivity is estimated to grow at the quarterly rate of $\mu_1^a = 0.48\%$ in the expansion and $\mu_2^a = -1.26\%$ in the recession. The productivity volatility comes out $\sigma_1^a = 1.38\%$ in the expansion, whereas it is moderately higher and about $\sigma_2^a = 2.19\%$ in the recession. The mean duration of the expansion is $(1 - \pi_{11}^a)^{-1} = 31.25$ quarters. The mean duration of the recession is $(1 - \pi_{22}^a)^{-1} = 3.61$ quarters, which is consistent with the sample mean duration of NBER recessions of around 3.44 quarters.¹²

The subjective discount factor is set to $\beta^4 = 0.985$ and the elasticity of intertemporal

¹²For comparison, Cagetti, Hansen, Sargent, and Williams (2002) estimate a two-state model using a different sample period and report mean productivity growth in the expansion and recession of 1.14% and −2.9% and productivity volatility of 1.92% per quarter. Using the data from 1947:Q2 to 2016:Q4, Liu and Zhang (2022) estimate a two-state model with regimes in the mean and volatility of productivity growth and report the estimates similar to those in Table 2. The only noticeable difference is that Liu and Zhang (2022) estimate productivity volatility in the recession much higher at 3.726% per quarter.

substitution to $\psi = 1.5$. The relative risk aversion parameter is set to 8. Consistent with the real business cycle literature, we set the capital share in a Cobb-Douglas production function at $\alpha = 0.36$ and the quarterly capital depreciation rate at $\delta = 0.02$. The constants a_1 and a_2 are chosen such that there are no adjustment costs in the non-stochastic steady state. The costs for adjusting capital are $\zeta = 24$, a value required to produce smooth consumption and volatile investment in the model.

We set $\lambda = 2.5$, a conservative number among parameter values chosen in the literature. We fix the remaining two parameters at $g_d = -0.54\%$ and $\sigma_d = 4.00\%$ to approximately capture the observed mean and volatility of aggregate stock market dividends. The choice of λ and σ_d implies a positive sample correlation between consumption and dividends, which corresponds well to the empirical point estimate of 0.24. This also leads to the annual dividend growth volatility of less than 8%, which is close to the estimate in our sample and substantially below the typically reported values.

4.2 *Learning and priors*

We consider the economies with known parameters and unknown transition probabilities. For the latter case, we calibrate the hyperparameters by embedding realistic and rather conservative prior information of the investor in the model. First, we consider various lengths of a prior learning period incorporating the information based on 100, 150, and 200 years of prior learning. Since we begin our asset pricing exercise shortly after World War II, training samples of 100 and 150 years effectively mean that the representative investor started learning about the unknown persistence of productivity growth in the middle and at the beginning of the nineteenth century. These dates approximately correspond to the beginning of the historical U.S. consumption and GDP growth series in the Barro-Ursua Macroeconomic Database. Second, we center the mean prior beliefs at the true maximum likelihood estimates obtained from the post-war data and calibrate the standard deviation of prior beliefs to reflect the length of the corresponding prior sample. Thus, our results are not driven by the pessimistic experience of the Great Depression and the two World

Wars but are the manifestation of rational pricing of hidden persistence in the post-war data. We discuss the impact of alternative priors in Section 5.2.

4.3 Macroeconomic quantities and equity returns

This section evaluates the impact of priced parameter uncertainty on macroeconomic quantities and equity returns. Panels A and B in Table 3 show that the models with parameter learning reasonably capture the first and second moments of consumption, dividends, and macroeconomic variables. Parameter learning amplifies the volatilities of consumption/investment growth and lowers correlations between quantities, helping the model to better capture empirical moments. In contrast, in the full information model, variables exhibit substantially higher correlations and lower volatilities than the data. Model-implied correlations between quantities and their persistence gradually increase with the length of a training period. This aligns with the observed pattern of empirical macroeconomic series, which appear more correlated in the post-war period than in the prior sample.

Panel C in Table 3 shows that the model with full information fails to produce the low risk-free rate, the large equity premium and excess return volatility, and the low mean of the log price-dividend ratio. Parameter uncertainty increases the equity premium and volatility more than six and two times. The table further shows that rational pricing of uncertainty about unknown parameters lowers the interest rates and equity valuations, allowing a better match of the average risk-free rate and price-dividend ratio. Even though the volatility of the log price-dividend ratio remains below a sample estimate, it is more than four times higher than in the full information setting. Overall, the model with parameter learning generates realistic moments of macroeconomic variables and equity returns.

4.4 Variance premium and option prices

Table 4 provides summary statistics for the variance premium and return variances. The model with known parameters produces a tiny variance premium. It also performs poorly

Table 3. Macroeconomic quantities and equity returns.

The table reports summary statistics of macroeconomic quantities and equity returns from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The model-based moments are the averages of small sample statistics based on 1000 simulations corresponding to the empirical data. Consumption and dividends are time-averaged to an annual frequency. The mean and standard deviation of investment, output, and returns are quarterly moments annualized by multiplying by an appropriate number. The correlation and autocorrelation are expressed in quarterly terms.

	Data	Priors (training sample)				Known parameters
		100 yrs	150 yrs	200 yrs	10000 yrs	
Panel A: Consumption and dividends						
$E_T(\Delta c_t)$	1.74	1.33	1.30	1.28	1.21	1.21
$\sigma_T(\Delta c_t)$	1.41	1.60	1.51	1.44	1.23	1.23
$ar1_T(\Delta c_t)$	0.42	0.16	0.21	0.26	0.46	0.47
$E_T(\Delta d_t)$	1.21	1.18	1.13	1.07	0.84	0.90
$\sigma_T(\Delta d_t)$	7.07	7.64	7.59	7.50	7.21	7.27
$ar1_T(\Delta d_t)$	0.21	0.19	0.21	0.22	0.25	0.26
$\rho_T(\Delta c_t, \Delta d_t)$	0.24	0.51	0.49	0.48	0.41	0.42
Panel B: Consumption, investment, and output						
$\sigma_T(\Delta i_t)$	4.69	4.80	4.64	4.52	4.17	4.17
$\sigma_T(\Delta y_t)$	2.52	2.02	2.02	2.02	2.02	2.02
$\rho_T(\Delta i_t, \Delta y_t)$	0.62	0.72	0.76	0.79	0.94	0.94
$\rho_T(\Delta c_t, \Delta y_t)$	0.35	0.53	0.57	0.62	0.84	0.84
$\rho_T(\Delta c_t, \Delta i_t)$	0.28	0.05	0.07	0.17	0.60	0.61
Panel C: Returns						
$E_T(r_t^f)$	0.49	1.16	1.36	1.48	1.88	1.89
$\sigma_T(r_t^f)$	0.91	0.26	0.25	0.25	0.24	0.24
$ar1_T(R_f)$	0.91	0.91	0.92	0.92	0.91	0.91
$E_T(r_t - r_t^f)$	6.64	6.72	5.47	4.45	1.04	1.05
$\sigma_T(r_t - r_t^f)$	16.24	21.22	19.44	17.55	10.66	10.62
$E_T(pd_t)$	3.50	2.77	2.94	3.09	3.88	3.90
$\sigma_T(pd_t)$	0.43	0.13	0.12	0.10	0.03	0.03
$ar1_T(pd_t)$	0.94	0.84	0.84	0.84	0.75	0.75

by generating too low volatility of the conditional return variance under both probability measures. Introducing priced parameter uncertainty amplifies the mean and volatility of the variance premium by a factor of more than ten and six, respectively, and yields values comparable to observed statistics. The results also demonstrate that parameter learning generates a large and volatile variance premium with realistic properties of conditional

Table 4. Variance premium.

The table reports summary statistics of the variance premium and return variances from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The empirical moments are based on the data from 1990:Q1 to 2019:Q4. The model-based moments are the averages of small sample statistics based on 1000 simulations corresponding to the empirical data. The mean, median, standard deviation, and maximum statistics are quarterly moments converted to monthly units by multiplying by an appropriate number. The skewness, kurtosis, and autocorrelation are expressed in quarterly terms.

	Data	Priors (training sample)				Known parameters
		100 yrs	150 yrs	200 yrs	10000 yrs	
$E_T(vp_t)$	10.05	7.16	5.53	3.94	0.30	0.52
$\sigma_T(vp_t)$	12.00	5.24	4.30	3.10	0.22	0.85
$E_T(VIX_t^2)$	35.05	40.41	35.22	28.66	9.68	9.75
$Md_T(VIX_t^2)$	24.13	30.47	26.29	21.52	8.74	8.67
$\sigma_T(VIX_t^2)$	30.99	30.50	27.13	21.24	2.59	3.10
$skew_T(VIX_t^2)$	2.34	3.17	3.29	3.28	3.11	3.22
$kurt_T(VIX_t^2)$	8.97	14.42	15.83	15.77	14.35	15.37
$ar1_T(VIX_t^2)$	0.50	0.62	0.61	0.60	0.61	0.60
$ar2_T(VIX_t^2)$	0.38	0.39	0.38	0.37	0.38	0.37
$E_T(VOL_t^2)$	25.01	33.25	29.69	24.72	9.38	9.24
$Md_T(VOL_t^2)$	18.07	24.91	22.11	18.59	8.53	8.45
$\sigma_T(VOL_t^2)$	23.86	25.29	22.84	18.14	2.37	2.26
$skew_T(VOL_t^2)$	3.42	3.16	3.29	3.28	3.11	3.20
$kurt_T(VOL_t^2)$	17.43	16.37	15.81	15.77	14.34	15.22
$ar1_T(VOL_t^2)$	0.55	0.62	0.61	0.60	0.61	0.60
$ar2_T(VOL_t^2)$	0.31	0.39	0.38	0.37	0.37	0.37

variances. With parameter uncertainty, the model-implied VIX^2 and VOL^2 exhibit a high volatility, a large positive skewness and kurtosis, and a moderate persistence. Thus, rationally accounting for parameter uncertainty also helps to reconcile salient moments of the variance premium and conditional variances.

Figure 4 shows the empirical and model-based one-quarter implied volatilities. With known parameters, implied volatilities are flat and approximately equal to the annualized equity return volatility. Parameter learning again improves the fit with the data. The higher degree of parameter uncertainty, as measured by a shorter prior sample, leads to an upward shift in the implied volatility line. Similarly to a level shift, the implied volatility curve becomes steeper at both ends of the moneyness range in response to more

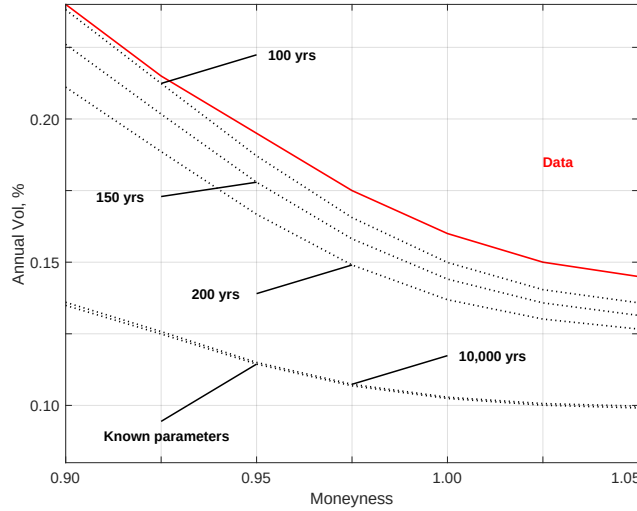


Figure 4. One-quarter implied volatilities.

The figure shows one-quarter implied volatilities from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The empirical curve corresponds to implied volatilities for S&P 500 index options from 1996:Q1 to 2019:Q4. The model-based curves are the averages of small sample implied volatilities based on 1000 simulations corresponding to the empirical data. Implied volatilities for the data and the models are annual. Strikes are expressed in moneyness (Strike Price/Spot Price).

uncertainty about unknown parameters. Hence, priced parameter uncertainty inflates the level and the slope of the implied volatility curve, capturing both skew and smirk patterns well. The impact of parameter learning remains significant after 200 years of prior learning. Indeed, the implied volatility curve in the model using a 200-year prior flattens and shifts downward but remains only slightly below the empirical line and well above the known parameter case. The effect of parameter uncertainty disappears after a very long learning period, as indicated by the convergence of the curves in the 10,000 years of learning and full information models.

Table 3 shows that the average equity premium generated by the model with priced parameter uncertainty drops by 20 percent when the prior length is increased by 50 years. The declining equity premium predicted by our mechanism is consistent with the empirical literature (Blanchard, Shiller, and Siegel, 1993; Jagannathan, McGrattan, and Scherbina, 2001; Fama and French, 2002). Lettau, Ludvigson, and Wachter (2008) present one partic-

ular explanation and advocate for a fall in macroeconomic risk. We present an alternative interpretation: the decrease in parameter uncertainty over time.

Furthermore, Table 4 reports a novel prediction of our model: the average variance premium drops around 25 percent over 50 years. Do we observe this decline in the data? If yes, is the degree of the decay reasonable? Although it is beyond the scope of this paper to examine the question of the declining variance premium, we provide evidence favoring a positive answer to both questions. Qualitatively, our results resonate well with Dew-Becker and Giglio (2024) demonstrating the decline in alphas of synthetic options strategies, which include the variance premium as a version of a delta-hedged put. Quantitatively, we estimate the variance premium from 1992:Q1 to 2022:Q1 using a rolling window procedure. The historical mean for decades ending in 2002:Q1, 2012:Q1, and 2022:Q1 are 13.97, 10.49, and 7.33, respectively. Consequently, the average variance premium has declined by around 25 and 30 percent over a decade-long period. In fact, in the data, it appears to decline, if anything, faster than in the production economy. Therefore, our results are not unreasonable and have support in the data in that the decline in parameter uncertainty could contribute to the decrease in the variance premium.

4.5 *SVIX and entropy of the risk-neutral return distribution*

We now examine the model's ability to account for salient features of the risk-neutral entropy of the return distribution and the simple volatility index. We retrieve the historical data for *SVIX* from Ian Martin's website. The data include the time series from 1996 to 2012, as in the original publication. Note that the empirical S^Q and *SVIX* are computed for the quarterly horizon and reported in annual volatility terms.

Table 5 shows that the empirical S^Q and *SVIX* have high mean and volatility, positive skewness, excess kurtosis, and moderate persistence. The difference between the series is positive on average and exhibits high volatility. The known parameter model has difficulty matching the properties of risk-neutral and simple VIX indexes. For instance, the mean and

median values of the model-based quantities are twice as small as in the data, while their volatility is more than five times smaller. In addition, the mean and volatility of $S^Q - SVIX$ predicted by the model are an order of magnitude below the empirical estimates.

The table also shows that priced parameter uncertainty leads to a substantial amplification of the first and second moments of variances. Focusing on the 100-year prior, the model with parameter learning closely matches the mean, median, and volatility of risk-neutral and simple variances. It generates a sufficiently large and volatile gap between S^Q and $SVIX$: 1.28 and 1.27 in the data versus 1.58 and 1.49 for the model with priced parameter uncertainty. The parameter learning model can also account for the spikiness and autocorrelation of variance measures as indicated by high and positive skewness, excess kurtosis, and relatively low autocorrelation. It is worth emphasizing that the productivity growth process primarily determines higher moments and autocorrelation, while parameter learning does not significantly alter these statistics.

4.6 $EP(x)$ and $VP(x)$

Here, we assess the ability of different models to capture the key patterns of the empirical $EP(x)$ and $VP(x)$. Similar to [Schreindorfer \(2020\)](#), we first compare the empirical moments with the corresponding population moments based on many observations. Then, we test the model performance in small samples where we generate many simulations of length corresponding to the empirical data and then report percentiles of small-sample statistics.

Figure 5 illustrates the population moment results. Note that we decompose the equity and variance premiums in the data for a quarterly horizon to align with the quarterly frequency in the production economy.¹³ In the model with known parameters, both curves are very low within the interval from $-30\% \cdot \sqrt{3}$ to $-10\% \cdot \sqrt{3}$, inconsistent with the primary role of such returns in the data. The model with priced parameter uncertainty does a good job of capturing the key patterns in $EP(x)$ and $VP(x)$. The model-implied curves begin to grow on the left side of an intermediate left tail region and steadily increase

¹³Appendix D provides the replication results of the empirical $EP(x)$ and $VP(x)$ for a quarterly horizon.

Table 5. SVIX and entropy of the risk-neutral return distribution.

The table reports summary statistics of $S^Q - SVIX$, S^Q , and $SVIX$ from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The empirical moments are based on the data from 1996:Q1 to 2012:Q1. The model-based moments are the averages of small sample statistics based on 1000 simulations corresponding to the empirical data. The mean, median, standard deviation, and maximum statistics are quarterly moments converted to annual units by multiplying by an appropriate number. The skewness, kurtosis, and autocorrelation are expressed in quarterly terms.

	Data	Priors (training sample)				Known parameters
		100 yrs	150 yrs	200 yrs	10000 yrs	
$E_T(S^Q - SVIX_t)$	1.28	1.58	1.22	0.89	0.17	0.16
$\sigma_T(S^Q - SVIX_t)$	1.27	1.49	1.28	0.98	0.12	0.11
$E_T(S^Q)$	22.41	21.63	19.89	17.88	10.98	10.89
$Md_T(S^Q)$	21.47	20.87	18.82	16.76	10.53	10.48
$\sigma_T(S^Q)$	7.89	3.01	3.58	3.50	1.22	1.16
$skew_T(S^Q)$	1.42	2.61	2.90	3.05	2.91	3.00
$kurt_T(S^Q)$	6.53	11.29	13.13	14.41	12.29	13.25
$ar1_T(S^Q)$	0.63	0.65	0.64	0.62	0.62	0.62
$ar2_T(S^Q)$	0.40	0.43	0.41	0.39	0.39	0.39
$E_T(SVIX_t)$	21.13	20.05	18.67	16.99	10.81	10.72
$Md_T(SVIX_t)$	20.51	18.74	17.13	15.51	10.32	10.28
$\sigma_T(SVIX_t)$	6.80	4.36	4.78	4.43	1.34	1.26
$skew_T(SVIX_t)$	1.33	3.08	3.12	3.21	2.93	3.03
$kurt_T(SVIX_t)$	6.14	14.09	14.55	15.53	12.49	13.46
$ar1_T(SVIX_t)$	0.64	0.61	0.62	0.61	0.62	0.62
$ar2_T(SVIX_t)$	0.42	0.38	0.39	0.38	0.39	0.39

within the gray-shaded area. Comparing the results for various priors, the effect of priced parameter uncertainty disappears slowly with $EP(x)$ and $VP(x)$ resembling the empirical counterparts even after 200 years of prior learning.

Table 6 provides summary statistics of return state contributions. It shows that the model with known parameters attributes 0 of the equity (variance) premium to returns below $-30\% \cdot \sqrt{3}$ and 0.901 (0.500) to returns above $-10\% \cdot \sqrt{3}$, the numbers outside of the empirical 95% confidence intervals. The contribution of the intermediate region to the equity premium is also significantly below the empirical counterpart. Introducing priced parameter uncertainty helps the production economy to match the moments well. The model with a 100-year prior produces the contributions of different intervals that are close

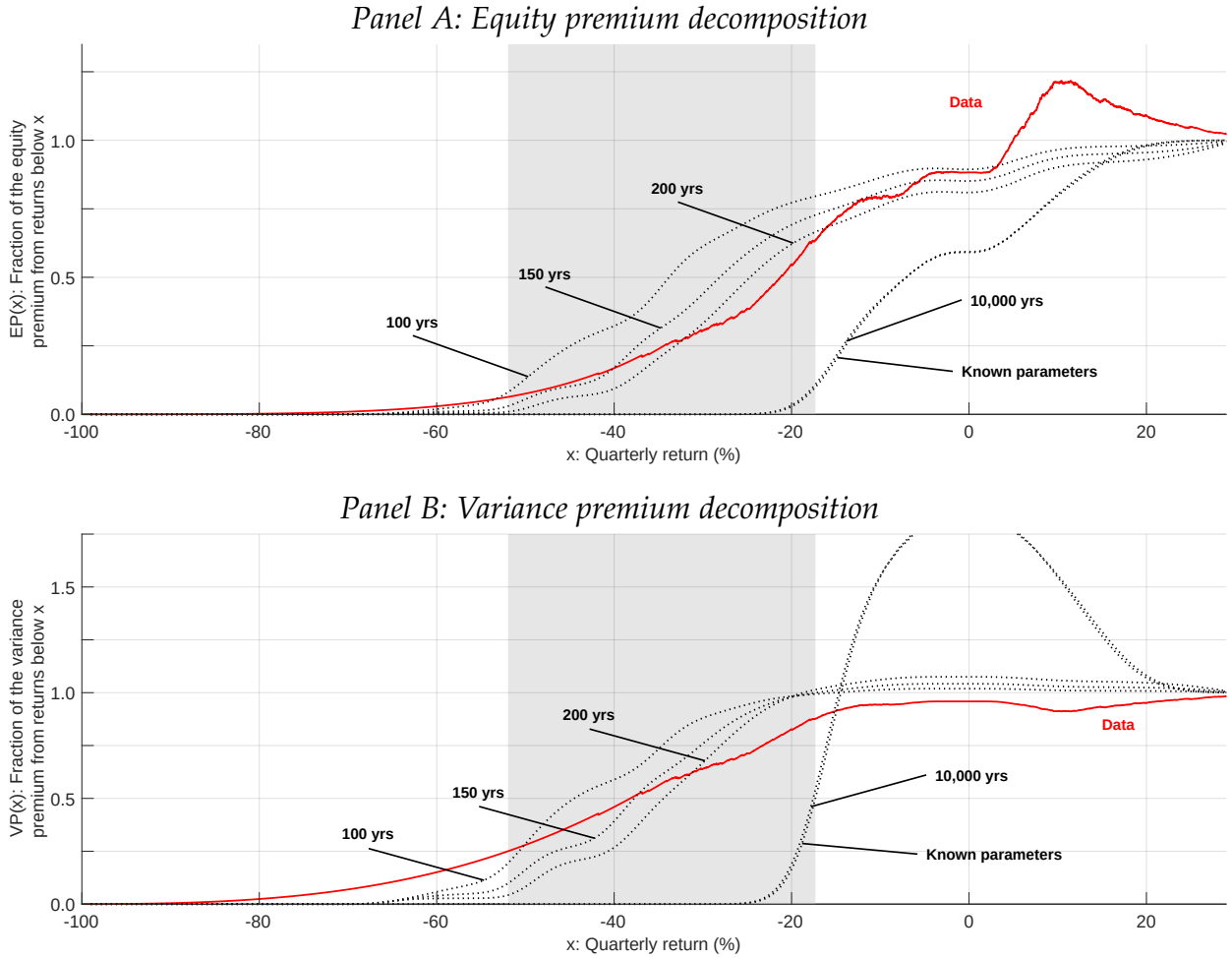


Figure 5. $EP(x)$ and $VP(x)$.

The figure shows the $EP(x)$ and $VP(x)$ curves from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The shaded area denotes the quarterly returns between $-30\% \cdot \sqrt{3}$ and $-10\% \cdot \sqrt{3}$. The empirical curve is based on the data from January 1996 to December 2019. The model-based curves are computed for the population of 1 million random draws.

to the mean estimates in the data and fall within the empirical 95% confidence intervals. Specifically, it predicts a non-negligible role of the far left tail – 0.087 and 0.206 versus 0.064 and 0.250 in the data – and a dominant impact of the intermediate left tail – 0.709 and 0.788 versus 0.571 and 0.626 in the data.

The bottom of Table 6 shows that more realistic risk prices drive the model improvement. The probability and price of the bad but non-disastrous events significantly undershoot the empirical estimates and are located on the left side of the 95% confidence inter-

Table 6. $EP(x)$ and $VP(x)$: population statistics.

The table reports the summary statistics of $EP(x)$, $VP(x)$, physical and risk-neutral PDFs from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The empirical moments are based on the data from 1996:Q1 to 2019:Q4. The table reports the empirical moments when $f(R)$ is proxied by realized returns (the f_R column) or is a smoothed version (the f_R^{smooth} column) of the empirical PDF. The entries in squared brackets are the empirical 95% confidence intervals. The model-based moments are computed for the population of 1 million random draws.

	Data		Priors (training sample)				Known parameters
	f_R	f_R^{smooth}	100 yrs	150 yrs	200 yrs	10000 yrs	
$EP(-.3\sqrt{3})$	0.064 [0.032, 0.226]	0.064	0.087	0.033	0.011	0.000	0.000
$EP(-.1\sqrt{3}) - EP(-.3\sqrt{3})$	0.571 [0.305, 1.532]	0.571	0.709	0.694	0.655	0.109	0.099
$1 - EP(-.1\sqrt{3})$	0.365 [-0.738, 0.644]	0.364	0.204	0.273	0.334	0.891	0.901
$VP(-.3\sqrt{3})$	0.250 [0.162, 0.356]	0.251	0.206	0.106	0.046	0.000	0.000
$VP(-.1\sqrt{3}) - VP(-.3\sqrt{3})$	0.626 [0.444, 0.756]	0.641	0.788	0.894	0.968	0.541	0.500
$1 - VP(-.1\sqrt{3})$	0.124 [0.012, 0.254]	0.108	0.017	0.005	0.014	0.459	0.500
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R) dR$	0.013 [0.003, 0.038]	0.019	0.026	0.026	0.025	0.004	0.003
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f^*(R) dR / \int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R) dR$	4.032 [1.453, 17.147]	2.765	2.520	2.217	2.073	1.773	1.755

vals. In the model with parameter learning, the probability and risk price of returns in $[-30\%, -10\%] \cdot \sqrt{3}$ increases substantially. Note that the simulated results are population moments based on a large number of observations, whereas the empirical counterparts are the historical averages based on a small sample. When we use the smoothed probability density function to compute the empirical estimates, the model-based values of 0.026 and 2.520 align better with the empirical averages of 0.019 and 2.765.

Table 7. $EP(x)$ and $VP(x)$: small sample statistics.

The table reports the summary statistics of $EP(x)$, $VP(x)$, physical and risk-neutral PDFs from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The empirical moments are based on the data from 1996:Q1 to 2019:Q4. The table reports the empirical moments when $f(R)$ is proxied by realized returns (the f_R column) or is a smoothed version (the f_R^{smooth} column) of the empirical PDF. The model-based moments are the medians with the 95% confidence intervals in squared brackets of small sample statistics based on 10000 simulations corresponding to the empirical data. The table reports the model-based moments when the physical PDF is proxied by realized returns in the simulation. The last row in the table shows p -value for the test $H_0 : EP(-.1) - EP(-.3) \geq 0.571 \quad \& \quad VP(-.1) - VP(-.3) \geq$

$$0.626 \quad \& \quad \int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f^*(R)dR \bigg/ \int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R)dR \geq 4.032.$$

	Data		Priors (training sample)				Known parameters
	f_R	f_R^{smooth}	100 yrs	150 yrs	200 yrs	10000 yrs	
$EP(-.3\sqrt{3})$	0.064	0.064	0.000 [-0.276, 0.731]	0.000 [0.000, 0.381]	0.000 [0.000, 0.120]	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
$EP(-.1\sqrt{3})$ $-EP(-.3\sqrt{3})$	0.571	0.571	0.896 [-0.427, 3.422]	0.843 [-1.886, 4.508]	0.781 [-3.166, 5.096]	0.147 [-2.696, 2.633]	0.142 [-2.625, 2.541]
$1 - EP(-.1\sqrt{3})$	0.365	0.364	0.044 [-2.482, 1.445]	0.126 [-3.380, 2.917]	0.209 [-4.146, 4.146]	0.853 [-1.633, 3.696]	0.858 [-1.541, 3.625]
$VP(-.3\sqrt{3})$	0.250	0.251	0.000 [-0.345, 0.892]	0.000 [0.000, 0.553]	0.000 [0, 0.176]	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
$VP(-.1\sqrt{3})$ $-VP(-.3\sqrt{3})$	0.626	0.641	0.877 [-0.288, 1.719]	0.863 [-0.583, 1.652]	0.860 [-1.046, 2.17]	0.402 [-5.576, 6.271]	0.389 [-5.665, 6.185]
$1 - VP(-.1\sqrt{3})$	0.124	0.108	0.069 [-0.316, 0.842]	0.103 [-0.580, 1.365]	0.122 [-1.144, 1.908]	0.598 [-5.271, 6.576]	0.611 [-5.185, 6.665]
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R)dR$	0.013	0.019	0.022 [0.000, 0.054]	0.022 [0.000, 0.054]	0.022 [0.000, 0.054]	0.000 [0.000, 0.022]	0.000 [0.000, 0.022]
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f^*(R)dR$ $\bigg/ \int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R)dR$	4.032	2.765	2.405 [1.006, 7.434]	2.144 [0.879, 6.619]	2.024 [0.792, 6.160]	0.592 [0.223, 0.989]	0.538 [0.222, 0.683]
p -value			0.161	0.167	0.159	0.000	0.000

The two comments are noteworthy before we present the small sample tests. First, the empirical statistics are based on the daily overlapping one-quarter returns. Due to the

Table 8. $EP(x)$ and $VP(x)$: small sample statistics (continuation).

The table reports the summary statistics of $EP(x)$, $VP(x)$, physical and risk-neutral PDFs from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The empirical moments are based on the data from 1996:Q1 to 2019:Q4. The table reports the empirical moments when $f(R)$ is proxied by realized returns (the f_R column) or is a smoothed version (the f_R^{smooth} column) of the empirical PDF. The model-based moments are the medians with the 95% confidence intervals in squared brackets of small sample statistics based on 10000 simulations corresponding to the empirical data. The table reports the model-based moments when the physical PDF is proxied by the model-implied density function of the return distribution. The last row in the table shows p -value for the test $H_0 : EP(-.1) - EP(-.3) \geq 0.571$ & $VP(-.1) - VP(-.3) \geq 0.641$ & $\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f^*(R)dR / \int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R)dR \geq 2.765$.

	Data		Priors (training sample)				Known parameters
	f_R	f_R^{smooth}	100 yrs	150 yrs	200 yrs	10000 yrs	
$EP(-.3\sqrt{3})$	0.064	0.064	0.000 [0.000, 0.532]	0.000 [0.000, 0.381]	0.000 [0.000, 0.12]	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
$EP(-.1\sqrt{3})$ $-EP(-.3\sqrt{3})$	0.571	0.571	0.773 [0.256, 0.91]	0.740 [0.238, 0.889]	0.707 [0.188, 0.875]	0.109 [0.056, 0.189]	0.102 [0.056, 0.155]
$1 - EP(-.1\sqrt{3})$	0.365	0.364	0.180 [0.075, 0.610]	0.233 [0.097, 0.752]	0.286 [0.114, 0.811]	0.891 [0.811, 0.944]	0.898 [0.845, 0.944]
$VP(-.3\sqrt{3})$	0.250	0.251	0.001 [0.000, 0.690]	0.000 [0.000, 0.348]	0.000 [0.000, 0.106]	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]
$VP(-.1\sqrt{3})$ $-VP(-.3\sqrt{3})$	0.626	0.641	0.980 [0.277, 1.156]	0.996 [0.566, 1.526]	1.007 [0.697, 2.147]	0.509 [0.322, 1.721]	0.473 [0.356, 1.485]
$1 - VP(-.1\sqrt{3})$	0.124	0.108	0.004 [-0.156, 0.089]	0.000 [-0.526, 0.120]	-0.008 [-1.147, 0.187]	0.491 [-0.721, 0.678]	0.527 [-0.485, 0.644]
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R)dR$	0.013	0.019	0.029 [0.011, 0.039]	0.028 [0.014, 0.035]	0.027 [0.011, 0.034]	0.004 [0.002, 0.006]	0.004 [0.003, 0.005]
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f^*(R)dR$ $/ \int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R)dR$	4.032	2.765	2.267 [1.804, 3.412]	2.144 [1.775, 2.761]	2.024 [1.761, 2.500]	1.723 [1.707, 1.74]	1.720 [1.713, 1.729]
p -value			0.129	0.021	0.004	0.000	0.000

model's quarterly frequency, we cannot replicate this daily approach in simulations. One possible solution is to generate the same number of observations in each simulation. How-

ever, this would counterfactually imply that the agent learns about unknown parameters for $24 \times 12 \times 21 = 6048$ quarters in a single simulation. For this reason, we simulate models of length $24 \times 4 = 196$ quarters following the approach in [Schreindorfer \(2020\)](#). Note, however, that this leads to wide model-based confidence intervals because the simulated moments employ observations less than 3×21 times in the data (also much wider intervals compared to monthly model simulations considered in [Schreindorfer \(2020\)](#)). Second, the unconditional PDF under a physical measure in the data is based on historical frequencies of returns, while the unconditional PDF under a risk-neutral measure is the average of daily conditional PDFs obtained from option prices. In the model, the exact conditional PDFs are available for each period. Hence, we can use the true conditional physical PDF to obtain a more precise estimate of the unconditional distribution. Furthermore, this also alleviates the discrepancy in sampling daily and quarterly observations in the data and the model. Indeed, as we generate more returns, unconditional distribution based on sampled returns will converge to the true distribution.

Table 7 reports the results for the case when $f(R)$ is proxied by realized returns in simulations. The model with known parameters generates the median values of return contributions far from the empirical estimates. However, the confidence intervals of some statistics are very wide and might contain empirical estimates. The full information economy can still be rejected based on low risk prices of tail events. In specifications with priced parameter uncertainty, the median statistics better align with the empirical estimates. Consequently, we cannot reject the model with a 100-year prior based on any of the univariate confidence intervals.

The bottom line in the table further reports the p-values of a multivariate test considered in [Schreindorfer \(2020\)](#). For each model, it shows the fraction of samples across all simulations of the economy satisfying three conditions: the equity and variance premia contributions and risk prices of returns in $[-30\%, -10\%] \cdot \sqrt{3}$ are larger than empirical estimates. One can interpret these fractions as p-values for a one-sided test of the model generating a significant impact of tail returns as in the data. We see that parameter learning

models cannot be rejected even based on the joint test.

Table 8 shows the results for the case when $f(R)$ is estimated based on the exact probability distribution function in the model. Note that, in this case, the construction of model-based statistics aligns better with empirical estimates using the smoothed physical PDF. As shown in the table, median statistics change slightly while the confidence intervals narrow significantly compared to the previous construction. As a result, we can reject the known parameter specification based on almost any univariate confidence interval. The models with parameter learning also have a harder time explaining the data. Still, the specification with 100 years of prior learning cannot be rejected based on univariate or multivariate tests.

5 Inspecting the mechanism and additional results

We have demonstrated that a single-shock production economy with priced parameter uncertainty can jointly capture (1) basic macroeconomic moments, (2) basic asset pricing moments, (3) salient features of the variance premium and return variances including SVIX, (4) option-implied volatilities, and (5) the equity and variance premia decompositions. This section explains the economic mechanism of how the model achieves these results and discusses additional implications.

5.1 Macroeconomic tail risks

This section demonstrates that the key intuition stems from sizable physical tail risks driven by priced parameter uncertainty in the production economy. To illustrate this point, Table 9 reports conditional moments of consumption and dividend growth in expansion and recession, assuming unbiased initial beliefs. In expansions, the mean consumption and dividend growth rates are similar across different specifications with parameter learning and are slightly higher compared to the full information case. In recessions, however, consumption and dividends decline, on average, more than twice as much in models with

Table 9. Conditional moments.

The table reports conditional moments of log consumption and dividend growth from the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. The conditional moments are computed for the economy in expansion and recession where the initial capital is equal to the steady state value and initial mean beliefs are centered at the true parameter values. The mean, standard deviation, and correlation of consumption and dividends are expressed in quarterly terms.

	$s_t = \text{expansion}$					$s_t = \text{recession}$				
	Priors (training sample)				Known param.	Priors (training sample)				Known param.
	100 yrs	150 yrs	200 yrs	10000 yrs		100 yrs	150 yrs	200 yrs	10000 yrs	
$E_T(\Delta c_t)$	0.83	0.75	0.68	0.49	0.49	-3.65	-3.57	-3.28	-1.49	-1.49
$\sigma_T(\Delta c_t)$	0.47	0.40	0.39	0.47	0.46	0.66	0.62	0.62	0.65	0.65
$E_T(\Delta d_t)$	1.57	1.34	1.12	0.65	0.65	-9.66	-9.49	-8.70	-4.29	-4.29
$\sigma_T(\Delta d_t)$	4.19	4.17	4.13	4.18	4.18	4.36	4.32	4.26	4.33	4.33
$\rho_T(\Delta c_t, \Delta d_t)$	0.29	0.25	0.24	0.29	0.29	0.38	0.36	0.36	0.38	0.38

parameter learning. This occurs because priced parameter uncertainty amplifies the impact of macroeconomic shocks in equilibrium models. The production economy allows for endogenous feedback from productivity shocks to consumption through parameter beliefs and investment decisions, further amplifying the decline in consumption and dividends.

The bottom line of Table 9 also shows that the correlation between consumption and dividend growth strengthens in a recession. A higher positive correlation implies that low consumption growth (and hence high marginal utility values) in a low productivity growth state is more likely to occur with low dividend growth. This creates a higher covariance between the SDF and the payoff to the equity and variance claims. Hence, a larger part of the equity and variance premia is associated with negative stock returns. Quantitatively, we demonstrate that the magnitude of tail risk in the parameter learning model is consistent with the observed equity and variance premia decompositions, even though these moments are not directly targeted in the calibration.

It is noteworthy that Schreindorfer (2020) presents a similar mechanism to explain the option prices. There are two key differences between the two approaches. First, his model generates the state-dependent correlation between consumption and dividends because their innovations are exogenously modeled as a mixture of Gaussian distributions, con-

sistent with the empirical evidence. In contrast, the production economy in our setting generates this feature endogenously. Second, the author employs generalized disappointment aversion preferences as risk price amplification, whereas our model relies on priced parameter uncertainty.

We now demonstrate that the degree of tail risks in consumption and dividends in our model is consistent with the data. We examine the downside correlation between the log consumption and log dividend growth (Δc_t and Δd_t) defined as follows:

$$\rho(\Delta c) = \text{corr}[\Delta c_t, \Delta d_t | \Delta c_t \leq x].$$

Following the procedure in [Schreindorfer \(2020\)](#), we first aggregate the consumption and dividends in the data and simulations by time-averaging quarterly consumption and dividends to an annual frequency and then compute the downside correlation between the two series. Figure 6 demonstrates the downside correlations in the data and the model with a 100-year prior. The median of model-based downside correlations reasonably tracks its empirical counterpart. The data-based values tend to be within the 95% confidence interval. Furthermore, the figure confirms that the downside correlation is high in periods of low consumption growth, which is in line with the key mechanism of the model.

As an additional exercise, we compare the empirical distribution of aggregate consumption and dividends with those predicted by the calibration. The macro-finance literature follows a standard approach to calibrate the models for shorter time horizons (e.g., monthly) and target moments of annual (time-aggregated) quantities. Figure 7 compares the distribution of annual log consumption and dividend growth in the data and model simulations. For each simulation, we perform the Andersen-Darling test for equality of the empirical and the simulated distribution. Focusing on the parameter learning results in top panels, the proportions of simulations for which the test does not reject the null hypothesis of equality of the distributions at the 5% confidence level are 0.631 and 0.845 for consumption and dividend growth. The empirical distributions are generally within

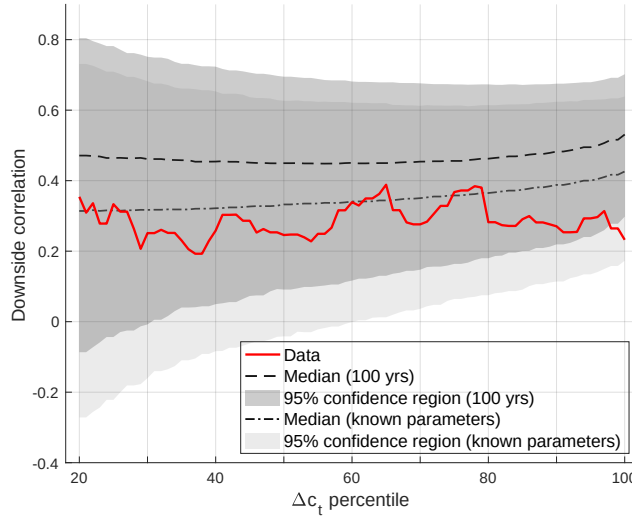


Figure 6. Downside correlations.

The figure shows downside correlations between annual log consumption and dividend growth from the data and production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100 years of prior learning) and known parameters. The empirical curve is based on the data from 1947:Q2 to 2019:Q4. The model-based lines are the medians of small sample statistics based on 1000 simulations corresponding to the empirical data. The shaded area denotes the 95% confidence region. We time-average quarterly consumption and dividends to an annual frequency and compute the correlation between the two series below the threshold for consumption (shown on the horizontal axis) consistent with [Schreindorfer \(2020\)](#).

95% confidence intervals predicted by model-based distributions. This demonstrates that the model with parameter learning produces distributions of cash flows close to those observed in the data.

We emphasize an additional aspect: the distribution of consumption growth exhibits a fatter tail in the parameter learning model compared to the data and the specification with known parameters. Although we have not observed such big declines in consumption and dividends in the post-war data from Q1:1947 to Q4:2019, the large declines in consumption and dividends at the onset of the COVID outbreak (as well as even more extreme observations before 1947) clearly indicate a non-zero likelihood of such events. Thus, we conclude that the observed data in the post-war period could be likely generated by our model.

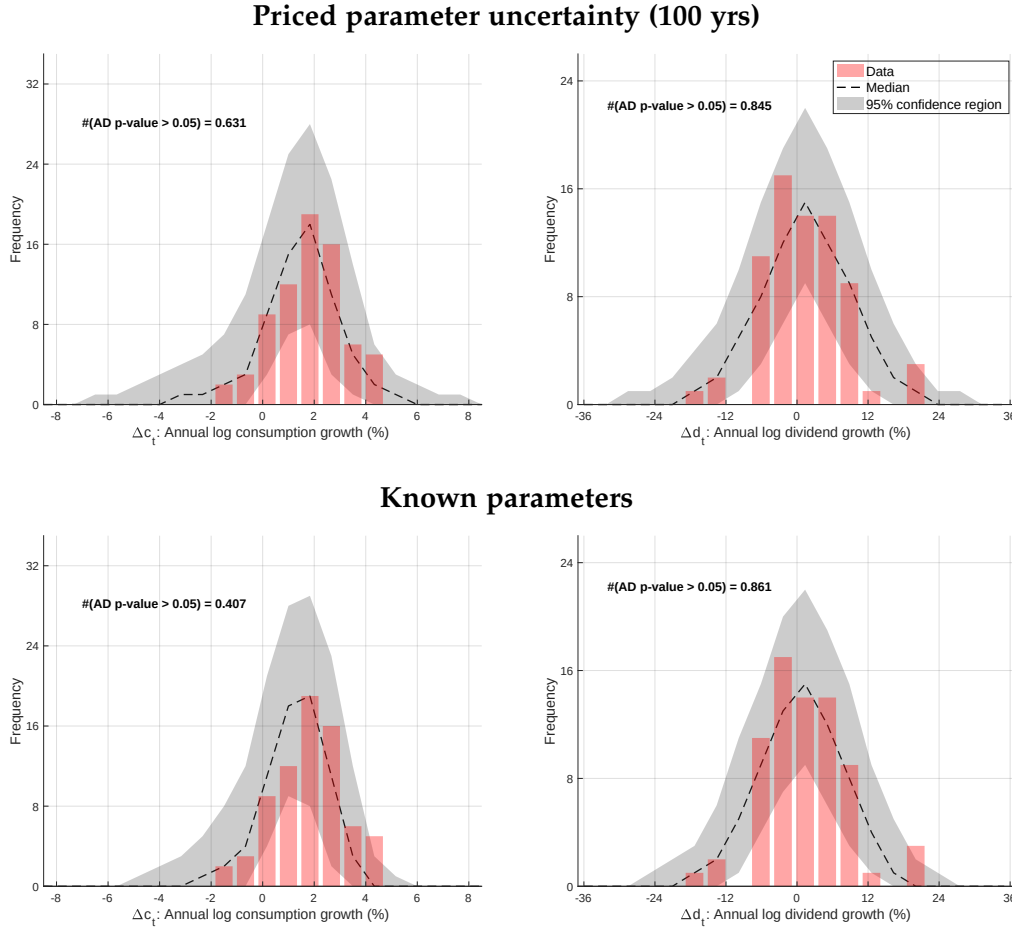


Figure 7. Distributions of annual log consumption and dividend growth.

The figure shows the distribution of annual log consumption and dividend growth from the data and production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100 years of prior learning) and known parameters. The empirical distribution is based on the data from 1947:Q2 to 2019:Q4. The model-based lines are the medians of small sample statistics based on 1000 simulations corresponding to the empirical data. The shaded area denotes the 95% confidence region. We time-average quarterly consumption and dividends to an annual frequency and plot the histogram of the two series.

5.2 Anticipated utility

This section emphasizes the key role of priced parameter uncertainty for our results. Anticipated utility (AU) pricing (Kreps, 1998; Piazzesi and Schneider, 2009) is a benchmark approach to dealing with parameter uncertainty in equilibrium models (Piazzesi and Schneider, 2009; Johannes et al., 2016). AU investors learn about unknown parameters over time and treat their current mean beliefs as true parameter values when making decisions. Thus, they ignore parameter uncertainty when computing utility and asset prices. We now solve

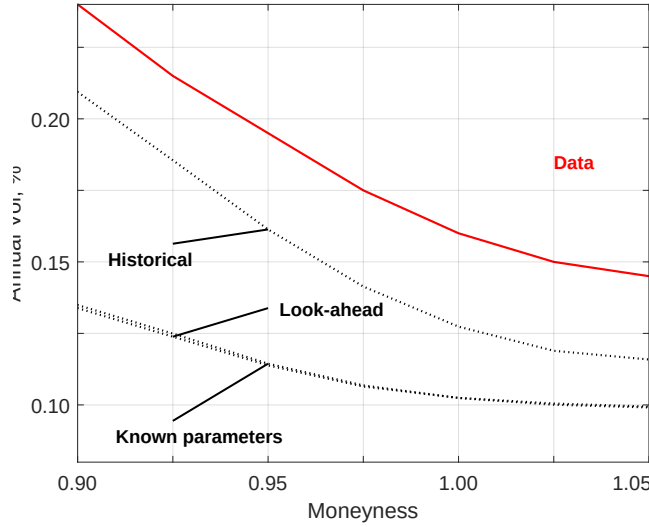


Figure 8. Implied volatilities: anticipated utility.

The figures show one-quarter implied volatilities from the data and the production economies with unknown transition probabilities under different priors, as well as known parameters. The models with parameter learning assume anticipated utility pricing and a 100-year prior period. We report the results for look-ahead priors equal to maximum likelihood estimates of transition probabilities and initial beliefs centered at the historical priors in 1947:Q1. In both cases with look-ahead and historical priors, we use the original calibration of the production economy. The empirical curve corresponds to implied volatilities for S&P 500 index options from 1996:Q1 to 2019:Q4. Implied volatilities for the data and the models are annualized. Strikes are expressed in moneyness (Strike Price/Spot Price).

the benchmark calibration with unknown transition probabilities by applying anticipated utility pricing. We report the results for the two types of initial beliefs: look-ahead unbiased priors employed in the main analysis and historical beliefs estimated from annual productivity growth from 1929 to 1946. Appendix C discusses the estimation procedure used to obtain these priors.

Table 10 reports the key moments. Panel A demonstrates that feeding forward-looking priors into the model produces almost identical moments to those implied by the known parameter specification. One notable difference between the two cases is the increased volatility of excess returns (12.23% versus 10.62%) and the price-dividend ratio (0.08% versus 0.03%). Panel A demonstrates further improvements in unconditional moments when using pessimistic priors based on historical data. Specifically, we observe an almost two-fold increase in the volatility of excess returns and a more than three-fold increase

Table 10. Sample moments: anticipated utility

This table reports asset pricing moments from the data and the production economies with unknown transition probabilities under different priors, as well as known parameters. The models with parameter learning assume anticipated utility pricing and a 100-year prior period. We report the results for look-ahead priors equal to the maximum likelihood estimates of transition probabilities and initial beliefs centered at the historical priors in 1947:Q1. In both cases, we use the original calibration of the production economy.

	Data	Prior: 100 yrs		Known
		Look-ahead	Historical	parameters
Panel A: Returns and variances				
$E(R_f) - 1$	0.49	1.90	1.61	1.89
$\sigma(R_f)$	0.91	0.24	0.25	0.24
$E(R - R_f)$	6.64	1.08	3.86	1.05
$\sigma(R - R_f)$	16.24	12.23	18.67	10.62
$E(pd)$	3.50	3.90	3.33	3.90
$\sigma(pd)$	0.43	0.08	0.20	0.03
$E(VP)$	10.0	0.52	1.56	0.52
$\sigma(VP)$	12.00	0.86	1.11	0.85
$E_T(S^Q - SVIX_t)$	1.28	0.16	0.88	0.16
$\sigma_T(S^Q - SVIX_t)$	1.27	0.12	0.83	0.11
Panel B: $EP(x)$ and $VP(x)$				
$EP(-.3\sqrt{3})$	0.064 0.064 [0.032, 0.226]	0.001	0.014	0.000
$EP(-.1\sqrt{3}) - EP(-.3\sqrt{3})$	0.571 0.571 [0.305, 1.532]	0.222	0.528	0.099
$1 - EP(-.1\sqrt{3})$	0.365 0.364 [-0.738, 0.644]	0.777	0.459	0.901
$VP(-.3\sqrt{3})$	0.250 0.251 [0.162, 0.356]	0.017	0.074	0.000
$VP(-.1\sqrt{3}) - VP(-.3\sqrt{3})$	0.626 0.641 [0.444, 0.756]	0.918	0.967	0.500
$1 - VP(-.1\sqrt{3})$	0.124 0.108 [0.012, 0.254]	0.065	-0.042	0.500
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R) dR$	0.013 0.019 [0.003, 0.038]	0.009	0.019	0.003
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f^*(R) dR / \int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R) dR$	4.032 2.765 [1.453, 17.147]	1.306	1.815	1.755
p -value		0.000	0.000	0.000

in the average equity and variance premiums compared to the known parameter specification. The difference between VIX and SVIX becomes more than five times larger and more volatile. As equity returns and valuations become more volatile and risk-neutral properties of returns improve, a larger fraction of the average equity (variance) premium is attributable to larger market declines. Panel B shows that returns from -30% to -10% contribute only 0.099 (0.500) in the full information model, but 0.222 (0.918) and 0.528 (0.967) in the parameter learning models with look-ahead and historical priors. Although combining anticipated utility pricing with pessimistic beliefs improves the fit with the data, the moments appear far from empirical estimates. In addition, the bottom three lines in Panel B show that risk prices of left-tail returns still remain abysmal compared to the data, and hence, we can reject these specifications based on p-values close to zero. Figure 8 further illustrates the implied volatility curves in the data and models. Consistent with previous results, pessimistic priors inflate the level and slope of implied volatilities, but the effect is quantitatively small to match the empirical line.

Our results are not inconsistent with [Cogley and Sargent \(2008\)](#) who also examine learning about the mean duration of recessions and document some asset pricing advantages of anticipated utility. In their setting, a calibrated model with pessimistic priors based on the Great Depression can match the equity premium. As shown in Table 10 and Figure 8, pessimistic priors in our setting indeed lead to a several-fold increase in risk premiums and volatility of conditional return variances, and steeper and higher option-implied volatilities. Naturally, one can improve further the performance by introducing more complex inference problems with state, parameter, and model uncertainty. This would slow down learning due to confounding and, hence, amplify the impact of belief revisions on asset prices. In particular, it would be interesting to extend this approach to model uncertainty similarly to [Johannes, Lochstoer, and Mou \(2016\)](#) and study the implications for derivatives. Alternatively, anticipated utility with learning about the whole distribution of shocks may lead to long-lasting belief revisions that would translate into more pronounced asset pricing implications. We leave these interesting avenues for future research.

5.3 Alternative definitions of the variance premium

The empirical and theoretical literature on the variance premium has developed significantly in the past decade. Surprisingly, however, disagreement remains on a standardized estimation procedure for the variance premium in the data and a universally accepted definition in discrete-time models. The divergence in empirical methodologies arises from different approaches to estimating conditional expectations of return variance under the physical measure. Theoretical discrepancies primarily stem from the mismatch between data sampling (e.g., daily) and model frequency (e.g., monthly). Our definition in Section 3.2 and results in Section 4.4 are based on the original variance premium specification by [Bollerslev, Tauchen, and Zhou \(2009\)](#) and [Carr and Wu \(2009\)](#). This section demonstrates that our key findings are robust under alternative definitions proposed in other studies.

[Drechsler and Yaron \(2011\)](#) define the variance premium as $E_t^Q \left[\text{var}_{t+1}^Q(r_{t+2}) \right] - E_t \left[\text{var}_{t+1}(r_{t+2}) \right]$ to reflect not only the level difference in variance under different measures but also the difference in the drifts of conditional variance due to its time-varying nature. [Bekaert et al. \(2020\)](#) define the variance premium as $\text{var}_t^Q(r_{t+1}) - \text{var}_t(r_{t+1})$. This definition is more suitable for shorter horizons (such as a month or shorter). It differs from ours in that it does not skip the equity log return and assumes that the agents form their conditional expectations instantaneously and have perfect foresight of future variance. Next, we follow [Schreindorfer \(2020\)](#) and use the risk-neutral entropy S_t^Q to measure the risk-neutral variance expectation (to capture deviations from log-normality) and define the variance premium as $2S_t^Q - \text{var}_t(r_{t+1})$. Finally, [Lorenz, Schmedders, and Schumacher \(2020\)](#) define the variance premium as the squared deviation of a one-period ahead return from its mean under the physical measure $E_t^Q \left[(r_{t+1} - E_t(r_{t+1}))^2 \right] - E_t \left[(r_{t+1} - E_t(r_{t+1}))^2 \right]$ to reflect a discrete nature of returns in the model.

The results for alternative specifications are reported in Table 11. Note that we compute the empirical variance premium consistent with the theory-based definitions. For instance, we employ the forecast of future variance for the definition in [Drechsler and Yaron \(2011\)](#)

Table 11. Variance premium: alternative definitions.

The table reports summary statistics of the variance premium from the data and the production economies with unknown transition probabilities (assuming priced parameter uncertainty and 100, 150, 200, and 10000 years of prior learning) and known parameters. We report the empirical and model-based results for alternative definitions of the variance premium encountered in the literature. Following [Drechsler and Yaron \(2011\)](#), Panel A defines the variance premium as the difference between the risk-neutral and physical expectations of return variance, which is computed by the risk-neutral and physical expectations, respectively. Panel B follows the definition of [Bekaert et al. \(2020\)](#) where the variance premium is simply the difference between the risk-neutral and physical return variances. Panel C modifies this definition by computing the risk-neutral variance as the entropy of returns, similar to the VIX definition in [Schreindorfer \(2020\)](#). Panel D adopts a discrete-time analogous of realized return variance, which is measured as the squared deviation of a one-period ahead return from its mean, consistent with [Lorenz et al. \(2020\)](#). The empirical moments are based on the data from 1990:Q1 to 2019:Q4. The model-based moments are the averages of small sample statistics based on 1000 simulations corresponding to the empirical data. The mean and standard deviation are quarterly moments converted to monthly units by multiplying by an appropriate number.

	Data	Priors (training sample)				Known parameters
		100 yrs	150 yrs	200 yrs	10000 yrs	
Panel A: $vp_t = \mathbb{E}_t^Q[\text{var}_{t+1}^Q(r_{t+2})] - \mathbb{E}_t[\text{var}_{t+1}(r_{t+2})]$ (Drechsler and Yaron, 2011)						
$E_T(vp_t)$	10.05	9.56	7.20	5.38	1.04	0.87
$\sigma_T(vp_t)$	12.00	19.31	14.19	9.70	0.64	0.73
Panel B: $vp_t = \text{var}_t^Q(r_{t+1}) - \text{var}_t(r_{t+1})$ (Bekaert et al., 2020)						
$E_T(vp_t)$	9.26	7.91	5.42	3.83	0.84	0.83
$\sigma_T(vp_t)$	27.19	29.72	22.36	15.62	1.14	1.06
Panel C: $vp_t = 2S_t^Q - \text{var}_t(r_{t+1})$ (Schreindorfer, 2020)						
$E_T(vp_t)$	9.26	6.54	4.58	3.33	0.77	0.75
$\sigma_T(vp_t)$	27.19	25.97	19.06	13.22	0.99	0.92
Panel D: $vp_t = E_t^Q[(r_{t+1} - E_t(r_{t+1}))^2] - E_t^P[(r_{t+1} - E_t(r_{t+1}))^2]$ (Lorenz et al., 2020)						
$E_T(vp_t)$	9.26	11.27	7.76	5.40	0.96	0.94
$\sigma_T(vp_t)$	27.19	22.61	17.28	12.25	0.95	0.88

and the one-month realized variance measure for others. The summary statistics of the variance premium in the data vary across different definitions, an observation previously pointed out by [Bekaert et al. \(2020\)](#). Nevertheless, the key takeaways of our main analysis remain unchanged. Priced parameter uncertainty strongly amplifies the magnitude and variation of the variance premium. The model with a 100-year prior produces variance premium statistics close to empirical counterparts. Finally, the decay rate in model-based

statistics is similar across different definitions and equals 25 to 30 percent per 50 years of learning – all consistent with our findings in Section 4.4.

6 Conclusion

This paper studies the decomposition of the average equity and variance premiums on the market return state space. Empirically, we document that left-tail returns contribute significantly to both premia, with low but non-disastrous returns playing a dominant role. This dominance is more pronounced in the equity premium decomposition, whereas more extreme returns account for a larger fraction of the variance premium in relative terms. Theoretically, we propose a real business cycle model with parameter learning to explain the equity and variance premia decompositions while simultaneously matching a wide array of salient moments of macroeconomic variables, equity returns, and option prices. We show that the model's success is attributable to fully rational pricing of parameter uncertainty in equilibrium, which amplifies the impact of productivity shocks on marginal utility and asset prices. Specifically, priced parameter uncertainty gives rise to endogenous physical tail risks in consumption and dividends, which helps match the magnitude of market returns driving the risk premia in the data.

References

- Ai, H. (2010). Information quality and long-run risk: Asset pricing implications. *The Journal of Finance* 65(4), 1333–1367.
- Andrei, D., B. Carlin, and M. Hasler (2019). Asset pricing with disagreement and uncertainty about the length of business cycles. *Management Science* 65(6), 2900–2923.
- Andrei, D., M. Hasler, and A. Jeanneret (2019). Asset pricing with persistence risk. *The Review of Financial Studies* 32(7), 2809–2849.
- Andrei, D., W. Mann, and N. Moyen (2019). Why did the q theory of investment start working? *Journal of Financial Economics* 133(2), 251–272.
- Babiak, M. and R. Kozhan (2024). Parameter learning in production economies. *Journal of Monetary Economics* 144, 103555.
- Backus, D., M. Chernov, and I. Martin (2011). Disasters implied by equity index options. *The Journal of Finance* 66, 1969–2012.
- Bansal, R., D. Kiku, and A. Yaron (2010). Long-run risks, the macroeconomy, and asset prices. *American Economic Review, Papers and Proceedings* 100(2), 542–546.
- Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance* 59 (4), 1481–1509.
- Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics* 121 (3), 823–866.
- Beason, T. and D. Schreindorfer (2022). Dissecting the equity premium. *Journal of Political Economy* 130(8), 2203–2222.
- Beeler, J. and J. Y. Campbell (2012). The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review* 1, 183–221.
- Bekaert, G. and E. Engstrom (2017). Asset return dynamics under habits and bad environment–good environment fundamentals. *Journal of Political Economy* 60, 713–760.
- Bekaert, G., E. Engstrom, and A. Ermolov (2020). The variance risk premium in equilibrium models. Technical report, National Bureau of Economic Research.

- Belo, F., X. Lin, and S. Bazdresch (2014). Labor hiring, investment and stock return predictability in the cross section. *Journal of Political Economy* 122, 129–77.
- Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2011). Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics* 101, 552–573.
- Blanchard, O. J., R. Shiller, and J. J. Siegel (1993). Movements in the equity premium. *Brookings Papers on Economic Activity* 1993(2), 75–138.
- Boldrin, M., L. J. Christiano, and J. D. M. Fisher (2001). Habit persistence, asset returns, and the business cycle. *American Economic Review* 91, 149–66.
- Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *The Review of Financial Studies* 22 (11), 4463–4492.
- Breeden, D. and R. H. Litzenberger (1978). State contingent prices implicit in option prices. *Journal of Business* 51, 3–24.
- Cagetti, M., L. P. Hansen, T. Sargent, and N. Williams (2002). Robustness and pricing with uncertain growth. *The Review of Financial Studies* 15(2), 363–404.
- Campbell, J. and J. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107, 205–251.
- Carr, P. and L. Wu (2009). Variance risk premiums. *Review of Financial Studies* 22, 1311–1341.
- Carvalho, C. M., M. S. Johannes, H. F. Lopes, and N. G. Polson (2010). Particle learning and smoothing. *Statistical Science* 25(1), 88–106.
- Carvalho, C. M., M. S. Johannes, H. F. Lopes, and N. G. Polson (2011). Particle learning: Simulation-based bayesian inference. *Bayesian Statistics* 9, 1–21.
- Christoffersen, P. and K. Jacobs (2004). The importance of the loss function in option valuation. *Journal of Financial Economics* 72, 291–318.
- Cogley, T. and T. J. Sargent (2008). The market price of risk and the equity premium: A legacy of the great depression? *Journal of Monetary Economics* 55(3), 454–476.
- Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016). Parameter learning in general equilibrium: The asset pricing implications. *American Economic Review* 106 (3), 664–98.
- Constantinides, G. M. and A. Ghosh (2017). Asset pricing with countercyclical household

- consumption risk. *Journal of Finance* 72, 415–460.
- Davis, J. and G. Segal (2022, 10). Trendy Business Cycles and Asset Prices. *The Review of Financial Studies*. hhac084.
- Dew-Becker, I. and S. Giglio (2024). The decline of the variance risk premium: evidence from traded and synthetic options. *Working Paper*.
- Drechsler, I. (2013). Uncertainty, time-varying fear, and asset prices. *Journal of Finance* 68(5), 1837–1883.
- Drechsler, I. and A. Yaron (2011). What’s vol got to do with it. *Review of Financial Studies* 24, 1–45.
- Du, D. (2010). General equilibrium pricing of options with habit formation and event risks. *Journal of Financial Economics* 99, 400–426.
- Epstein, L. G. and S. E. Zin (1991). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy* 99(2), 263–286.
- Eraker, B. and I. Shaliastovich (2008). An equilibrium guide to designing affine pricing models. *Mathematical Finance* 18, 519–543.
- Fama, E. F. and K. R. French (2002). The equity premium. *The Journal of Finance* 57(2), 637–659.
- Favilukis, J. and X. Lin (2016). Wage rigidity: A quantitative solution to several asset pricing puzzles. *Review of Financial Studies* 29 (1), 148–192.
- Gillman, M., M. Kejak, and M. Pakos (2015). Learning about rare disasters: Implications for consumption and asset prices. *Review of Finance* 19(3), 1053–1104.
- Hall, R. E. (2001). The stock market and capital accumulation. *American Economic Review* 91(5), 1185–1202.
- Hamilton, J. (1990). Analysis of time series subject to changes in regimes. *Journal of Econometrics* 45, 39–70.
- He, Z. and A. Krishnamurthy (2013). Intermediary asset pricing. *American Economic Review* 103, 732–770.

- Jagannathan, R., E. McGrattan, and A. Scherbina (2001). The declining us equity premium.
- Jahan-Parvar, M. and H. Liu (2014). Ambiguity aversion and asset prices in production economies. *Review of Financial Studies* 27, 3060–3097.
- Johannes, M., L. A. Lochstoer, and Y. Mou (2016). Learning about consumption dynamics. *Journal of Finance* 71, 551–600.
- Johannes, M. and N. Polson (2009). Particle filtering. In *Handbook of Financial Time Series*, pp. 1015–1029. Springer.
- Johnson, T. (2007). Optimal learning and new technology bubbles. *Journal of Monetary Economics* 87, 2486–2511.
- Kaltenbrunner, G. and L. Lochstoer (2010). Long-run risk through consumption smoothing. *Review of Financial Studies* 23, 3190–224.
- Kozhan, R., A. Neuberger, and P. Schneider (2013). The skew risk premium in the equity index market. *The Review of Financial Studies* 26(9), 2174–2203.
- Kozlowski, J., L. Veldkamp, and V. Venkateswaran (2019). The tail that keeps the riskless rate low. *NBER Macroeconomics Annual* 33(1), 253–283.
- Kozlowski, J., L. Veldkamp, and V. Venkateswaran (2020). The tail that wags the economy: Beliefs and persistent stagnation. *Journal of Political Economy* 128(8), 2839–2879.
- Kreps, D. M. (1998). Anticipated utility and dynamic choice. *Econometric Society Monographs* 29, 242–274.
- Lettau, M., S. Ludvigson, and J. Wachter (2008). The declining equity premium: What role does macroeconomic risk play. *Review of Financial Studies* 21, 1653–1687.
- Liu, H. and Y. Zhang (2022). Financial uncertainty with ambiguity and learning. *Management Science* 68(3), 2120–2140.
- Liu, J., J. Pan, and T. Wang (2005). An equilibrium model of rare-event premia and its implication for option smirks. *Review of Financial Studies* 18, 131–164.
- Lorenz, F., K. Schmedders, and M. Schumacher (2020). Nonlinear dynamics in conditional volatility. In *Proceedings of Paris December 2020 Finance Meeting EUROFIDAI-ESSEC*.
- Martin, I. (2017). What is the expected return on the market? *The Quarterly Journal of*

- Economics* 132(1), 367–433.
- Pakos, M. (2013). Long-run risk and hidden growth persistence. *Journal of Economic Dynamics and Control* 37 (9), 1911–1928.
- Piazzesi, M. and M. Schneider (2009). *Trend and cycle in bond premia*, Volume 424. Citeseer.
- Rietz, T. A. (1988). The equity risk premium: A solution. *Journal of Monetary Economics* 22(1), 117–131.
- Schreindorfer, D. (2020). Macroeconomic tail risks and asset prices. *The Review of Financial Studies* 33(8), 3541–3582.
- Seo, S. B. and J. A. Wachter (2019). Option prices in a model with stochastic disaster risk. *Management Science* 65(8), 3449–3469.
- Shaliastovich, I. (2015). Learning, confidence, and option prices. *Journal of Econometrics* 187(1), 18–42.
- Stock, J. H. and M. W. Watson (1999). Business cycle fluctuations in us macroeconomic time series. *Handbook of macroeconomics* 1, 3–64.
- Uhlig, H. (2007). Explaining asset prices with external habits and wage rigidity in a dsge model. *American Economic Review* 97, 239–43.
- Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance* 68(3), 987–1035.
- Weitzman, M. (2007). Subjective expectations and asset-return puzzles. *American Economic Review* 97, 1102–1130.
- Winkler, F. (2020). The role of learning for asset prices and business cycles. *Journal of Monetary Economics* 114, 42–58.
- Zhou, G. and Y. Zhu (2014). Macroeconomic volatilities and long-run risks of asset prices. *Management Science* 61(2), 413–430.

Internet Appendix to

“Parameter learning, tail risks, and risk premia decomposition”

Abstract

This appendix provides a detailed description of the data, the numerical methods used to solve different models, and additional results not included in the main body of the paper.

Contents

A. Data description	A-1
B. Numerical solution methodology	A-4
B.1 Known parameters	A-4
B.2 Priced parameter uncertainty	A-6
B.3 Anticipated utility	A-12
C. Historical prior estimation	A-12
D. $EP(x)$ and $VP(x)$ for a quarterly horizon	A-14

A Data description

Macroeconomic quantities and productivity. We construct quarterly Solow residuals to obtain quarterly productivity growth from 1947:Q2 to 2019:Q4. We follow the procedure adopted from [Stock and Watson \(1999\)](#). We retrieve quarterly real non-farm gross domestic product (A358RX1Q020SBEA), quarterly nominal private non-residential fixed investment (PNFI) and corresponding price index (B008RG3Q086SBEA), quarterly labor proxied by hours of non-farm business sector (HOANBS), annual non-residential capital stock (K1NTOTL1ES000) and corresponding price index (B008RG3A086NBEA) from FRED. The data are U.S. series provided by NIPA. We deflate nominal capital and investment by the corresponding price indices. Following [Hall \(2001\)](#), we interpolate annual capital using quarterly investment to obtain quarterly values according to the recursion:

$$K_{t+1} = (1 - \delta_K)K_t + I_t,$$

where capital depreciation δ_K is constant within a year and is chosen to match annual capital values. Using a capital share of $\alpha = 0.36$, we construct the Solow residuals as:

$$\ln A_t = \ln Y_t - \alpha \ln K_t - (1 - \alpha) \ln N_t.$$

We rescale the Solow residuals by $1 - \alpha = 0.64$ to obtain labor-augmenting technology.

Aggregate consumption. We construct quarterly real per capita consumption growth series for the longest available sample from 1947:Q2 to 2019:Q4. Following [Bansal and Yaron \(2004\)](#), consumption is defined as a sum of nondurable goods and services. We retrieve nominal personal consumption expenditures on nondurable goods (PCND) and services (PCESV) from the [Federal Reserve Economic Data \(FRED\)](#) at the Federal Reserve Bank of St. Louis. The data are U.S. consumption series provided by National Income and Product Accounts (NIPA). We obtain real per capita time series by deflating by the end-of-quarter consumer price index (CPIAUCSL) and dividing by the population (POP).

Aggregate dividends, risk-free rate, excess market returns. We construct quarterly real per capita dividend growth series from 1947:Q2 to 2019:Q4. We follow the procedure adopted from [Andrei, Hasler, and Jeanneret \(2019\)](#). We retrieve the quarterly series of the value-weighted index including distributions (vwretd) and the value-weighted index excluding distributions (vwretx) from the Center for Research in Security Prices (CRSP). The data cover NYSE, Amex, and Nasdaq. We denote the returns of indices with or without distributions by $R_{D,t}$ and $R_{noD,t}$. We start by constructing quarterly price index as:

$$P_t = P_{t-1}(1 + R_{noD,t}).$$

We then compute the quarterly dividend series as:

$$D_t = P_t \left(\frac{1 + R_{D,t}}{1 + R_{noD,t}} - 1 \right).$$

We obtain real per capita time series by deflating by the end-of-quarter consumer price index (CPIAUCSL) and dividing by the population (POP). We remove the large seasonal component in dividend growth by using the X-12-ARIMA Seasonal Adjustment Program.

We construct the ex-ante real risk-free rate following [Beeler and Campbell \(2012\)](#). We retrieve monthly series of the three-month nominal treasury bill secondary market rate (TB3MS) and consumer price index (CPIAUCSL) from FRED. The nominal rate is continuously compounded as $y_{3,t} = \ln(1 + y_{3,t,obs}/100) / 4$. We estimate the following regression:

$$y_{3,t} - \pi_{t,t+3} = \beta_0 + \beta_1 y_{3,t} + \beta_1 \pi_{t-12,t} + \varepsilon_{t+3},$$

where $y_{3,t}$, $\pi_{t,t+3}$, and $\pi_{t-12,t}$ are monthly series of nominal three-month bill rate, three-month inflation from t to $t + 3$, and twelve-month inflation from $t - 12$ to t , respectively. We use the predicted values from this regression as a proxy for the ex-ante real risk-free rate. We obtain the real stock market return by deflating the return of the value-weighted index, including distributions with the consumer price index. The excess stock market

return is the difference between the real stock market return and the real risk-free rate.

Variance premium. For the empirical variance premium, we obtain the daily data of the VIX index, S&P 500 index futures, and the S&P 500 index from the Chicago Board of Options Exchange (CBOE) from 1990:Q1 to 2019:Q4. We follow [Drechsler \(2013\)](#) to calculate the variance risk premium. The variance risk premium is the difference between the risk-neutral and physical expectations of stock market return variance for a given horizon. We construct our own quarterly VIX index (expressed in annual volatility units) using the S&P index options. We use the last available observation of a particular quarter, take it to a second power, and divide by 4 to obtain a series expressed in quarterly variance units. We compute the objective expectation of return variance as the one-quarter-ahead forecast from a predictive regression of the futures realized variance on the squared VIX and an index realized variance measure. The quarterly futures and index realized variance measures are calculated as the sum over a quarter of the squared daily log returns on the S&P 500 futures and S&P 500 index. We estimate the predictive regression on an expanding window basis with the first two years as a burn-in period. We adopt the same procedure to construct the monthly variance premium.

Implied volatilities. For the empirical analysis of implied volatility surface, we use European options written on the S&P 500 index and traded on the CBOE. The option data set covers the period from January 1996 to December 2019 and is from OptionMetrics. Option data elements include the type of options (call/put) along with the contract's variables (strike price, time to expiration, Greeks, Black-Scholes implied volatilities, closing spot prices of the underlying) and trading statistics (volume, open interest, closing bid and ask quotes), among other details. To construct the empirical implied volatility curves, we first compute the moneyness for each observed option using the daily S&P 500 index on a particular trading day. We filter out all data entries with non-standard settlements. We use the remaining observations to construct the implied volatility surface for a range of moneyness and maturities. In particular, we follow [Christoffersen and Jacobs \(2004\)](#) and perform polynomial extrapolation of volatilities in the maturity time and strike prices.

This strategy makes use of all available options and not only those with a specific maturity time. The fitted values are further used to construct the implied volatility curves.

Decomposition of the equity and variance premiums. We construct the empirical decomposition of the equity and variance risk premiums using the replication files of [Beason and Schreindorfer \(2022\)](#), which can be accessed via the following link: https://www.journals.uchicago.edu/doi/suppl/10.1086/720396/suppl_file/20200045data.zip. In our analysis, we use options data from OptionMetrics covering the period from January 1996 to December 2019.

B Numerical solution methodology

This section provides the solution methodology of the production economy with known parameters and unknown transition probabilities assuming fully rational or anticipated utility pricing of parameter uncertainty.

B.1 Known parameters

Productivity growth is given by:

$$\Delta a_t = \mu_{s_t} + \sigma_{s_t} \cdot \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} N(0, 1)$$

where s_t is a two-state Markov chain with transition probabilities $\pi_{11} = 1 - \pi_{12}$ and $\pi_{22} = 1 - \pi_{21}$. The regime switches in s_t are independent of ε_t . Denoting the stationary variables $\{\tilde{C}_t, \tilde{I}_t, \tilde{Y}_t, \tilde{K}_t, \tilde{U}_t\} = \left\{ \frac{C_t}{A_t}, \frac{I_t}{A_t}, \frac{Y_t}{A_t}, \frac{K_t}{A_t}, \frac{U_t}{A_t} \right\}$, the household's problem is:

$$\tilde{U}_t = \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1 - \beta) \tilde{C}_t^{1 - \frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1}^{1-\gamma} \cdot \left(\frac{A_{t+1}}{A_t} \right)^{1-\gamma} \right] \right)^{\frac{1 - \frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} \quad (\text{B.1})$$

$$\tilde{C}_t + \tilde{I}_t = \tilde{K}_t^\alpha \bar{N}^{1-\alpha}, \quad \bar{N} = 1, \quad (\text{B.2})$$

$$e^{\Delta a_{t+1}} \tilde{K}_{t+1} = (1 - \delta) \tilde{K}_t + \varphi \left(\frac{\tilde{I}_t}{\tilde{K}_t} \right) \tilde{K}_t, \quad (\text{B.3})$$

$$\Delta a_{t+1} = \mu_{s_{t+1}} + \sigma_{s_{t+1}} \cdot \varepsilon_{t+1}, \quad \varepsilon_{t+1} \stackrel{\text{iid}}{\sim} N(0, 1), \quad (\text{B.4})$$

$$\tilde{C}_t \geq 0, \quad \tilde{K}_{t+1} \geq 0. \quad (\text{B.5})$$

The true parameters are assumed to be known and, hence, s_t and $\tilde{K}_t$ are the only state variables in the economy. Equation (B.1) can be rewritten as:

$$\begin{aligned} & \tilde{U}_t(s_t, \tilde{K}_t) \\ &= \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1 - \beta) \tilde{C}_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1}(s_{t+1}, \tilde{K}_{t+1})^{1-\gamma} \cdot e^{(1-\gamma)\Delta a_{t+1}} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} \end{aligned} \quad (\text{B.6})$$

We perform the following steps to solve Equation (B.6):

1. We find the de-trended steady state capital $\tilde{K}_{ss}$, assuming the productivity growth equals the steady state growth predicted by a Markov switching model. The state space for capital normalized by technology is $[0.2\tilde{K}_{ss}, 3.5\tilde{K}_{ss}]$. We use $n_k = 200$ points on a grid for capital.
2. For any level of capital $\tilde{K}_t$ at time t , we construct a grid for $\tilde{I}_t$ with uniformly distributed points between 0 and $\tilde{K}_t^\alpha \bar{N}^{1-\alpha}$. Specifically, we use $n_i = 4000$ points.
3. For the expectation, we use Gauss-Hermite quadrature with $n_{gh} = 16$ points. Using quadrature weights and nodes, we evaluate the right hand side of Equation (B.6) and update a new value function $\tilde{U}_t = \tilde{U}_t(s_t, \tilde{K}_t)$ given the old one $\tilde{U}_{t+1} = \tilde{U}_{t+1}(s_{t+1}, \tilde{K}_{t+1})$.
4. We iterate Steps 1-3 by updating the continuation utility at each iteration until a suitable convergence is achieved. Specifically, the stopping rule is that the distance between the new and old value functions satisfies $|\tilde{U}_{t+1} - \tilde{U}_t|/|\tilde{U}_t| < 10^{-12}$.

Having found the household's utility, we price a claim on calibrated stock market dividends:

$$\Delta d_t = g_d + \lambda \Delta c_t + \sigma_d \cdot \eta_t, \quad \eta_t \stackrel{\text{iid}}{\sim} N(0, 1).$$

The equilibrium condition for the price-dividend ratio is given by:

$$\begin{aligned} & \tilde{P}\tilde{D}_t \\ = & E_t \left[\beta e^{\left(\lambda - \frac{1}{\psi}\right)(\Delta \tilde{c}_{t+1} + \Delta a_{t+1}) + g_d + 0.5\sigma_d^2} \left(\frac{\tilde{U}_{t+1} \cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t(\tilde{U}_{t+1} \cdot e^{\Delta a_{t+1}})} \right)^{\frac{1}{\psi} - \gamma} (\tilde{P}\tilde{D}_{t+1} + 1) \right]. \end{aligned} \tag{B.7}$$

We use the value function iteration algorithm to solve this equation. The stopping rule is that the distance between the new and old value functions satisfies $|\tilde{P}\tilde{D}_{t+1} - \tilde{P}\tilde{D}_t| / |\tilde{P}\tilde{D}_t| < 10^{-12}$.

B.2 Priced parameter uncertainty

The numerical solution for the case of priced parameter uncertainty consists of two main steps.¹⁴ First, we solve for the equilibrium pricing ratios when true parameters are known by the household (by assumption, these are learned at $T = \infty$). We find the solution for this simplest limiting economy with known parameters by applying the methods outlined in Appendix B.1. Second, we use the known parameters boundary economy as a terminal value in the backward recursion to obtain the equilibrium model solution at each time t . In our solution, we set the period when the economy reaches the known parameters boundary equal to $T = 2000$ periods or 500 years. In this section, we present details of the solution methodology employed in the second step for the model with unknown transition probabilities. It is straightforward to extend the numerical methodology to the cases with (1) unknown transition probabilities and mean growth rates, and (2) unknown transition probabilities, mean growth rates, and volatilities.

¹⁴Johnson (2007) uses the solution methodology with the power utility. Collin-Dufresne, Johannes, and Lochstoer (2016) extend this approach to the case of Epstein-Zin utility in the endowment economy. In this paper, we further examine priced parameter uncertainty in the production economy with Epstein-Zin utility.

Productivity growth is given by:

$$\Delta a_t = \mu_{s_t} + \sigma_{s_t} \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} N(0, 1),$$

where s_t is a two-state Markov chain with a transition matrix:

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

with $\pi_{ii} \in (0, 1)$. The regime switches in s_t are independent of the Gaussian shocks ε_t . In the case of unknown transition probabilities, the representative household knows the true values of the parameters within each state $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ and observes states (s_t) but does not know the transition probabilities (π_{11}, π_{22}) . At time $t = 0$, the household holds priors about unknown probabilities in the transition matrix and updates beliefs each period upon realization of new series and regimes.

We assume a Beta distributed prior for unknown transition probabilities and, hence, posterior beliefs are also Beta distributed. We use two pairs of hyperparameters (a_1, b_1) and (a_2, b_2) for unknown transition probabilities π_{11} and π_{22} . At time t , the household uses Bayes' rule and the fact that states are observable to update hyperparameters for each state

$$a_{i,t} = a_{i,0} + \#(\text{state } i \text{ has been followed by state } i), \quad (\text{B.8})$$

$$b_{i,t} = b_{i,0} + \#(\text{state } i \text{ has been followed by state } j), \quad (\text{B.9})$$

given the initial prior beliefs $a_{i,0}$ and $b_{i,0}$.

Having found the limiting boundary economies from the first step, we then perform a backward recursion using the following state variables: the regime of the economy s_t ,

capital $\tilde{K}_t$, and the vector $X_t = \{\tau_{1,t}, \lambda_{1,t}, \tau_{2,t}, \lambda_{2,t}\}$ of sufficient statistics for the priors:

$$\tau_{1,t} = a_{1,t} + b_{1,t}, \quad \lambda_{1,t} = E_t[\pi_{11}] = \frac{a_{1,t}}{a_{1,t} + b_{1,t}}, \quad (\text{B.10})$$

$$\tau_{2,t} = a_{2,t} + b_{2,t}, \quad \lambda_{2,t} = E_t[\pi_{22}] = \frac{a_{2,t}}{a_{2,t} + b_{2,t}}. \quad (\text{B.11})$$

Note from Equations (B.3) and (B.8)-(B.11) that we can write $\tilde{K}_{t+1} = f(\Delta a_{t+1}, s_t, \tilde{K}_t, X_t)$ and $X_{t+1} = g(s_{t+1}, s_t, X_t)$. Given this, we can now write the equilibrium utility as

$$\tilde{U}_{t+1}(s_{t+1}, \tilde{K}_{t+1}, X_{t+1}) = \tilde{U}_{t+1}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t)$$

to indicate that the utility evolution is the function of the two observable variables, s_{t+1} and Δa_{t+1} . Ultimately, the recursive equation (B.1) can be rewritten as:

$$\begin{aligned} \tilde{U}_t(s_t, \tilde{K}_t, X_t) = \max_{\tilde{C}_t, \tilde{I}_t, N_t} & \left\{ (1 - \beta) \tilde{C}_t^{1 - \frac{1}{\psi}} \right. \\ & \left. + \beta \left(E_t \left[\tilde{U}_{t+1}^{1-\gamma}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, \tilde{K}_t, X_t \right] \right)^{\frac{1 - \frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}, \end{aligned} \quad (\text{B.12})$$

where the expectation on the right-hand side is equivalent to:

$$\begin{aligned} & E_t \left[\tilde{U}_{t+1}^{1-\gamma}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, \tilde{K}_t, X_t \right] \\ = & E_t \left[E_t \left[\tilde{U}_{t+1}^{1-\gamma}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, \tilde{K}_t, X_t \right] \middle| s_t, \tilde{K}_t, X_t \right] \\ = & \sum_{s_{t+1}=1}^2 \mathbb{P}(s_{t+1}|s_t, X_t) \times E_t \left[\tilde{U}_{t+1}^{1-\gamma}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, \tilde{K}_t, X_t \right] \\ = & \sum_{s_{t+1}=1}^2 E_t(\pi_{s_{t+1}, s_t} | s_t, X_t) \times E_t \left[\tilde{U}_{t+1}^{1-\gamma}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, \tilde{K}_t, X_t \right]. \end{aligned} \quad (\text{B.13})$$

The first and second equalities follow from the independency of regime changes and Gaussian shocks to productivity growth (s_{t+1} and ε_{t+1}). The third equality follows from

$$\mathbb{P}(s_{t+1}|s_t, X_t) = \int_0^1 \pi_{s_{t+1}, s_t} g(\pi_{s_{t+1}, s_t} | s_t, X_t) d\pi_{s_{t+1}, s_t} = E_t(\pi_{s_{t+1}, s_t} | s_t, X_t),$$

where $g(\pi_{s_{t+1},s_t}|s_t, X_t)$ denotes the conditional density of π_{s_{t+1},s_t} .

Using the definition of state variables, we have an analytical expression for the conditional expectation of transition probabilities in Equation (B.13), which is either $\lambda_{s_t,t}$ or $1 - \lambda_{s_t,t}$. For the conditional expectation in Equation (B.13), we do not have a closed form because the continuation utility depends on the realized productivity growth through $\tilde{K}_{t+1}$. Therefore, we use quadrature-type numerical methods to evaluate this expectation as follows:

$$\begin{aligned} & E_t \left[\tilde{U}_{t+1}^{1-\gamma} (s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, \tilde{K}_t, X_t \right] \\ & \approx \sum_{j=1}^J \omega_\varepsilon(j) \left[\tilde{U}_{t+1}^{1-\gamma} (s_{t+1}, \Delta a(j), s_t, \tilde{K}_t, X_t) \cdot e^{(1-\gamma)\Delta a(j)} \middle| s_{t+1}, s_t, \tilde{K}_t, X_t \right], \end{aligned} \quad (\text{B.14})$$

where $\omega_\varepsilon(j)$ is the quadrature weight corresponding to the quadrature node $n_\varepsilon(j)$ used for the integration of a standard normal shock ε_{t+1} in productivity growth. The observed realized productivity growth, $\Delta a(j)$, and a state variable, $\tilde{K}_{t+1}(j)$, are updated as follows:

$$\Delta a(j) = \mu_{s_{t+1}} + \sigma_{s_{t+1}} \cdot n_\varepsilon(j) \quad (\text{B.15})$$

$$e^{\Delta a(j)} \tilde{K}_{t+1}(j) = (1 - \delta) \tilde{K}_t + \varphi \left(\frac{\tilde{I}_t}{\tilde{K}_t} \right) \tilde{K}_t, \quad (\text{B.16})$$

$$\tilde{I}_t = \tilde{K}_t^\alpha N_t^{1-\alpha} - \tilde{C}_t. \quad (\text{B.17})$$

The backward recursion can be performed by using Equations (B.12)-(B.17).

We also solve for the price-dividend ratio of the equity claim on aggregate dividends, which are defined as leverage to aggregate consumption. Let exogenous aggregate dividends be given by:

$$\Delta d_{t+1} = g_d + \lambda \Delta c_{t+1} + \sigma_d \varepsilon_{d,t+1},$$

where $g_d = (1 - \lambda) \left(E(\mathbb{P}(s_\infty = 1 | \pi_{11}, \pi_{22})) \mu_1 + E(\mathbb{P}(s_\infty = 2 | \pi_{11}, \pi_{22})) \mu_2 \right)$ and $\mathbb{P}(s_\infty = i | \pi_{11}, \pi_{22})$ is the ergodic probability of being in the state i conditional on the transition probabilities π_{11} and π_{22} . Note that the long-run mean of dividends growth, g_d , is chang-

ing under the household's filtration, though the true long-run growth is constant. The subjective beliefs about the true parameter values induce fluctuations in g_d , which can be expressed as $g_d = g_d(s_{t+1}, s_t, X_t)$.

The equilibrium condition for the price-dividend ratio is standard in the Epstein-Zin:

$$PD_t = E_t \left[\begin{aligned} & \beta \left(\frac{\tilde{C}_{t+1}}{\tilde{C}_t} \right)^{-\frac{1}{\psi}} \left(\frac{1-N_{t+1}}{1-N_t} \right)^{\left(1-\frac{1}{\psi}\right)\nu} \left(\frac{A_{t+1}}{A_t} \right)^{-\frac{1}{\psi}} \left(\frac{\tilde{U}_{t+1} \cdot \left(\frac{A_{t+1}}{A_t} \right)}{\mathcal{R}_t(\tilde{U}_{t+1} \cdot \left(\frac{A_{t+1}}{A_t} \right))} \right)^{\frac{1}{\psi}-\gamma} \cdots \\ & \times \left(\frac{D_{t+1}}{D_t} \right) (PD_{t+1} + 1) \end{aligned} \right] \quad (\text{B.18})$$

First, we solve for the price-dividend ratio defined by Equation (B.18) for the simplest limiting economy with all known parameters. In this case, s_t and $\tilde{K}_t$ are the only state variables in the economy. We find the price-dividend ratio on a grid of π_{11} and π_{22} . Second, we use the known parameters boundary price-dividend ratios as terminal values in the backward recursion to obtain the equilibrium pricing functions at the time t . We rewrite all variables as a function of state variables and use quadrature-type numerical methods to evaluate expectations. We update the long-run dividend growth, $g_d(s_{t+1}, s_t, X_t)$. The equilibrium recursion used to solve the model is then:

$$\begin{aligned} & PD_t(s_t, \tilde{K}_t, X_t) \\ &= E_t \left[\begin{aligned} & \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{C}_{t+1}+\Delta a_{t+1})} \left(\frac{1-N_{t+1}}{1-N_t} \right)^{\left(1-\frac{1}{\psi}\right)\nu} \left(\frac{\tilde{U}_{t+1}e^{\Delta a_{t+1}}}{\mathcal{R}_t(\tilde{U}_{t+1}e^{\Delta a_{t+1}})} \right)^{\frac{1}{\psi}-\gamma} \cdots \left| s_t, \tilde{K}_t, X_t \right. \\ & \times e^{g_d(s_{t+1}, s_t, X_t) + 0.5\sigma_d^2} (PD_{t+1}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) + 1) \end{aligned} \right] \\ &= E_t \left[E_t \left[\begin{aligned} & \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{C}_{t+1}+\Delta a_{t+1})} \left(\frac{1-N_{t+1}}{1-N_t} \right)^{\left(1-\frac{1}{\psi}\right)\nu} \left(\frac{\tilde{U}_{t+1}e^{\Delta a_{t+1}}}{\mathcal{R}_t(\tilde{U}_{t+1}e^{\Delta a_{t+1}})} \right)^{\frac{1}{\psi}-\gamma} \left| s_{t+1}, s_t, \right. \\ & \times e^{g_d(s_{t+1}, s_t, X_t) + 0.5\sigma_d^2} (PD_{t+1}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) + 1) \left| \tilde{K}_t, X_t \right. \end{aligned} \right] \left| s_t, \right. \\ & \left. \left| \tilde{K}_t, X_t \right. \right] \end{aligned}$$

$$\begin{aligned}
&= \sum_{s_{t+1}=1}^2 \mathbb{P}(s_{t+1}|s_t, X_t) \dots \\
&\times E_t \left[\beta e^{\left(\lambda - \frac{1}{\psi}\right)(\Delta \tilde{c}_{t+1} + \Delta a_{t+1})} \left(\frac{1-N_{t+1}}{1-N_t}\right)^{\left(1-\frac{1}{\psi}\right)\nu} \left(\frac{\tilde{U}_{t+1} e^{\Delta a_{t+1}}}{\mathcal{R}_t(\tilde{U}_{t+1} e^{\Delta a_{t+1}})}\right)^{\frac{1}{\psi}-\gamma} \right. \\
&\quad \left. \times e^{g_d(s_{t+1}, s_t, X_t) + 0.5\sigma_d^2} (PD_{t+1}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) + 1) \right]_{s_{t+1}, s_t, \tilde{K}_t, X_t} \\
&= \sum_{s_{t+1}=1}^2 E_t(\pi_{s_{t+1}, s_t} | s_t, X_t) \\
&\times E_t \left[\beta e^{\left(\lambda - \frac{1}{\psi}\right)(\Delta \tilde{c}_{t+1} + \Delta a_{t+1})} \left(\frac{1-N_{t+1}}{1-N_t}\right)^{\left(1-\frac{1}{\psi}\right)\nu} \left(\frac{\tilde{U}_{t+1} e^{\Delta a_{t+1}}}{\mathcal{R}_t(\tilde{U}_{t+1} e^{\Delta a_{t+1}})}\right)^{\frac{1}{\psi}-\gamma} \right. \\
&\quad \left. \times e^{g_d(s_{t+1}, s_t, X_t) + 0.5\sigma_d^2} (PD_{t+1}(s_{t+1}, \Delta a_{t+1}, s_t, \tilde{K}_t, X_t) + 1) \right]_{s_{t+1}, s_t, \tilde{K}_t, X_t}
\end{aligned}$$

Again, the conditional expectation of transition probabilities under the household's filtration permits an analytical formula, while the inner expectation in the expression above can be evaluated using the quadrature-type integration methods.

The numerical solutions are computationally intensive with a high-dimensional vector of state variables, which include one variable for a regime of the economy in a two-state Markov process, one variable for the endogenous capital, and hyperparameters governing parameter beliefs in the production economies with unknown transition probabilities. Furthermore, each iteration in the backward recursion requires solving the optimization problem to find optimal investment. For this reason, we perform all computations on a high-performance computing cluster at Lancaster University. We solve the models using parallel computing techniques, efficiently programmed in Matlab. Specifically, we parallelize our code at each iteration of a backward recursion such that for each iterative step a large number of equilibrium conditions can be solved on a high-dimensional space of state variables in parallel before moving to the next iteration. Furthermore, we build the MEX files for the model solutions to accelerate the computations and to provide performance comparable to codes in C++. In terms of the grids, we use the grid of normalized capital $[0.2\tilde{K}_{ss}, 3.5\tilde{K}_{ss}]$ with $n_k = 200$ grid points and employ $n_i = 4000$ uniformly distributed points to find the optimal investment. For normally distributed productivity shocks, we use Gauss-Hermite quadrature with $n_{gh} = 16$ points. We use 7 and 27 points for the grids

of π_{11} and π_{22} in the models with unknown transition probabilities.

B.3 Anticipated utility

In the anticipated utility case, the representative household learns about unknown parameters but ignores parameter uncertainty when making decisions. The numerical solution proceeds as follows. At each time t , the household holds current beliefs and solves for the continuation utility and the levered equity claim prices in the rational expectations model in which the true parameter values in the productivity growth process are centered at the time t posterior means. In the next period $t + 1$, the household updates beliefs upon observing new data and resolves the rational expectations economy in which the true parameters are centered at the time $t + 1$ posterior means. In sum, the numerical algorithm reduces to applying the methodology for the full information case with a set of model parameters, which are equal to the mean beliefs at each point in time.

C Historical prior estimation

This section provides details of the data and the empirical strategy used to estimate the historical priors for the post-war data. We begin by constructing the annual productivity growth series from 1929 to 1946 using the methodology outlined in Appendix A. The start date is restricted by the historical time series provided by NIPA. The end date corresponds to the first observation of quarterly data.

We consider the most general inference problem with hidden regimes and uncertainty about all parameters in the two-state Markov model. We apply Monte Carlo methods called particle filters to sequentially generate particles approximating the posterior distribution. Assuming beta priors for transition probabilities and normal and inverse gamma priors for mean and variance parameters, particle filtering provides state and parameter learning in the two-state economy. Internet Appendix of [Johannes, Lochstoer, and Mou \(2016\)](#) provides a general discussion of the methods, whereas [Johannes and Polson \(2009\)](#),

Table A1. Priors and posteriors

This table reports the initial priors in 1929 and posteriors in 1946 from a particle filtering algorithm applied to annual productivity growth from 1929 to 1946. We employ beta priors for transition probabilities and normal and inverse gamma priors for mean and variance parameters in the two-state Markov switching processes. It also reports the initial priors in 1947:Q1.

Panel A: Priors for annual parameters in 1929		
Parameter	Mean	St. dev
μ_1	8.00%	4.00%
μ_2	-10.00%	4.00%
σ_1	$(6.00\%)^2$	$(2.00\%)^2$
σ_2	$(8.00\%)^2$	$(4.00\%)^2$
π_{11}	0.80	0.16
π_{22}	0.50	0.24
Panel B: Posteriors for annual parameters in 1946		
μ_1	6.32%	2.40%
μ_2	-9.12%	3.92%
σ_1	$(6.00\%)^2$	$(2.00\%)^2$
σ_2	$(7.92\%)^2$	$(3.92\%)^2$
π_{11}	0.87	0.11
π_{22}	0.57	0.19
Panel C: Priors for quarterly parameters in 1947:Q1		
μ_1	0.79%	0.30%
μ_2	-1.14%	0.49%
σ_1	$(1.50\%)^2$	$(0.50\%)^2$
σ_2	$(1.98\%)^2$	$(0.98\%)^2$
π_{11}	0.9658	0.0240
π_{22}	0.8689	0.0656

Carvalho, Johannes, Lopes, and Polson (2010), and Carvalho, Johannes, Lopes, and Polson (2011) present a detailed overview of the algorithms.

Furthermore, we need to specify the priors to begin the estimation in 1929. Following a standard approach in the literature, we make the priors as uninformative as possible while trying to avoid the state identification problem. Panel A in Table A1 reports prior parameters for the two-state and productivity growth process. We set wide mean beliefs and standard deviations due to a wide range of realized productivity growth rates. The priors for transition probabilities are the same in the two cases.

Panel B in Table A1 summarizes the 1946 posteriors of annual parameters. We now convert these annual estimates into quarterly using the same reasoning as Johannes, Lochstoer, and Mou (2016). We take the mean beliefs about the transition probabilities to the power of $1/4$ to obtain the probabilities in the quarterly transition matrix. Similarly, we compute 95% confidence intervals of the annual estimates of each probability. We then take the lower and upper bounds to the power of $1/4$ to obtain approximately $\pm$ two standard deviations around the mean beliefs of quarterly transition probabilities. We transform the mean and standard deviation of annual mean beliefs into the quarterly estimates by dividing them by $4 \cdot 2$. The average productivity growth was more than twice stronger over the 1929-1946 sample. Thus, we choose to divide the 1946 posteriors of mean beliefs by $4 \cdot 2$ to better reflect a likely structural break in productivity growth data. Turning to variance beliefs, we follow Johannes, Lochstoer, and Mou (2016) and divide the mean and standard deviation by $2 \cdot 2$ to reflect the change in the quality of data.

Panel C in Table A1 summarizes the results of converting annual to quarterly priors. We treat these estimates as historical priors in 1947:Q1 for the anticipated utility model. We further compare the results based on historical priors with look-ahead priors equal to the maximum likelihood estimates over the post-war period.

D $EP(x)$ and $VP(x)$ for a quarterly horizon

This section replicates the empirical equity and variance premia decompositions for a quarterly horizon. Figure A1 shows $EP(x)$ and $VP(x)$ in the data. Table A2 reports summary statistics of the return state contributions and risk prices of returns between $-30\% \cdot \sqrt{3}$ and $-10\% \cdot \sqrt{3}$. Figure A2 illustrates the sampling distribution of contributions of extreme and intermediate negative returns.

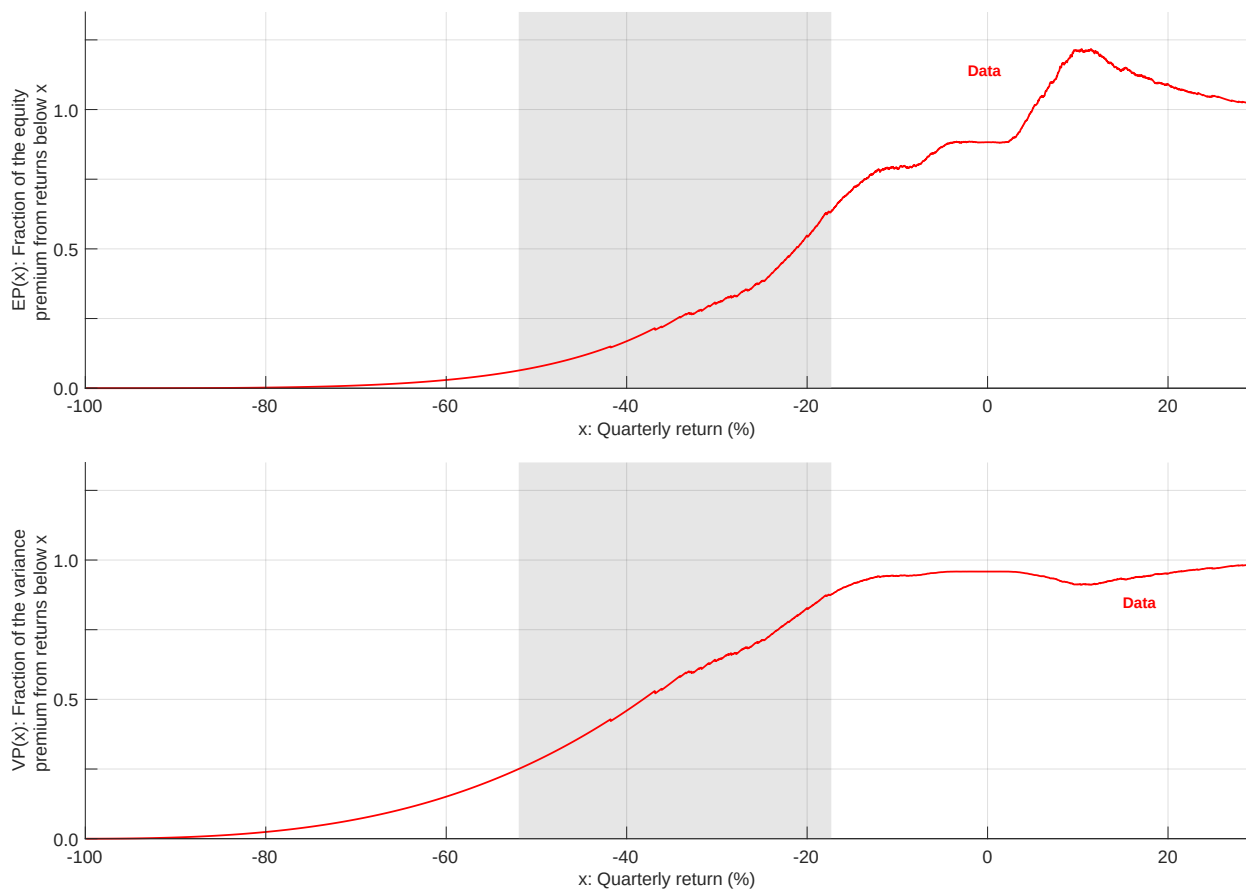


Figure A1. $EP(x)$ and $VP(x)$ in the data: quarterly horizon.

The figure shows the empirical $EP(x)$ and $VP(x)$ curves for a quarterly horizon. The shaded area denotes quarterly returns between $-30\% \cdot \sqrt{3}$ and $-10\% \cdot \sqrt{3}$. The empirical curves are based on the data from January 1996 to December 2019.

Table A2. Summary statistics of $EP(x)$ and $VP(x)$: quarterly horizon.

The table reports the contributions of different regions to the equity and variance premiums for a quarterly horizon. It also shows the probability of returns in the intermediate region and the ratio of their risk-neutral to physical probabilities. The table reports the empirical moments when $f(R)$ is proxied by realized returns (the f_R column) or a smoothed version (the f_R^{smooth} column) of the empirical PDF. The last five columns report selected percentiles of the estimates based on a block bootstrap with a block size of $21 \cdot 3$ trading days. The data is from January 1996 to December 2019.

	f_R	f_R^{smooth}	Percentiles				
			2.5%	5%	50%	95%	97.5%
$EP(-.3\sqrt{3})$	0.064	0.064	0.032	0.035	0.061	0.163	0.226
$EP(-.1\sqrt{3}) - EP(-.3\sqrt{3})$	0.571	0.571	0.305	0.350	0.568	1.186	1.532
$1 - EP(-.1\sqrt{3})$	0.365	0.364	-0.738	-0.335	0.370	0.603	0.644
$VP(-.3\sqrt{3})$	0.250	0.251	0.162	0.171	0.230	0.327	0.356
$VP(-.1\sqrt{3}) - VP(-.3\sqrt{3})$	0.626	0.641	0.444	0.493	0.646	0.739	0.756
$1 - VP(-.1\sqrt{3})$	0.124	0.108	0.012	0.032	0.127	0.228	0.254
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R)dR$	0.013	0.019	0.003	0.004	0.016	0.034	0.038
$\int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f^*(R)dR / \int_{-.3\sqrt{3}}^{-.1\sqrt{3}} f(R)dR$	4.032	2.765	1.453	1.624	3.415	12.281	17.147

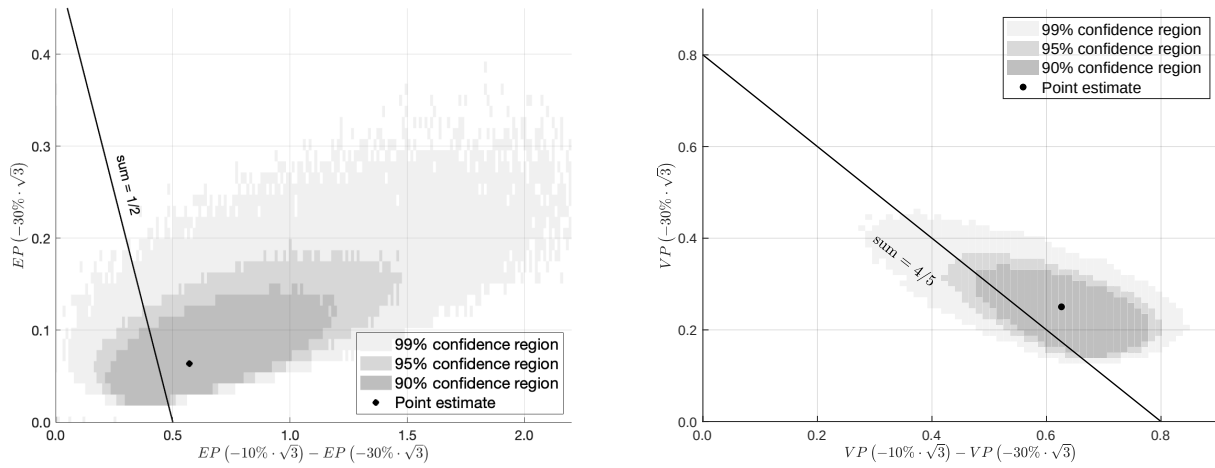


Figure A2. Sampling distribution of $EP(x)$ and $VP(x)$: quarterly horizon.

The figure shows the contributions of the two regions — quarterly returns below $-30\% \cdot \sqrt{3}$ and between $-30\% \cdot \sqrt{3}$ and $-10\% \cdot \sqrt{3}$ — to the equity and variance premiums in the 1 million bootstrap samples. For the sampling distribution, we implement a block bootstrap with a block size of $21 \cdot \sqrt{3}$ trading days. The length of the artificially simulated data corresponds to the length of the empirical data from January 1996 to December 2019. The shaded areas mark the confidence regions, whereas the solid black line denotes the contributions to the equity (variance) premium with a sum of 0.5 (0.8).