Asset Pricing with Persistence Risk

Daniel Andrei
McGill University

Michael Hasler
University of Toronto

Alexandre Jeanneret
HEC Montréal

Persistence risk is an endogenous source of risk that arises when a rational agent learns about the length of business cycles. Persistence risk is positive during recessions and negative during expansions. This asymmetry, which solely results from learning about persistence, causes expected returns, return volatility, and the price of risk to rise during recessions. Persistence risk predicts future excess returns, particularly at 3- to 7-year horizons. Its predictability is strongest around business-cycle peaks and troughs. We confirm the model's predictions in the data and provide evidence that persistence risk is priced in financial markets. (*JEL* D51, G11, G12)

Received October 13, 2017; editorial decision September 19, 2018 by Editor Stijn Van Nieuwerburgh.

Uncertainty and its effects on the economy have been subject to much research over the past decade. The driving force behind this surge in research interest is the finding that recessions are accompanied by

We are grateful to Harjoat Bhamra, Michael Brennan, John Campbell, Mike Chernov, John Cochrane, Lars Lochstoer, and Avanidhar Subrahmanyam for helpful comments that have substantially improved the paper. We thank Pat Akey, Jerome Detemple, Ian Dew-Becker (HEC-McGill Winter Finance Workshop discussant), Andrey Ermolov (EFA 2017 discussant), Jean-Sébastien Fontaine, Chanik Jo, Mariana Khapko, Gang Li, Erwan Morellec, Chayawat Ornthanalai, Marcel Rindisbacher, Alberto Rossi, Gustavo Schwenkler, and Alexandre Ziegler for their suggestions. We also thank seminar and conference participants at Boston University, Goethe University Frankfurt, HEC-McGill Winter Finance Workshop, University of Technology Sydney, University of New South Wales, University of Sydney, UCLA, the Einaudi Junior Conference, HEC Montréal, University of Zurich, HEC Lausanne, the EFA 2017, York University, the Asset Pricing Workshop in Zurich, the Bank of Canada, Queen's University, and the SITE Asset Pricing Theory and Computation Workshop for comments and suggestions. We thank McGill, UCLA, the University of Toronto, and HEC Montréal for financial support. M. H. and A. J. are particularly grateful to the Canadian Derivatives Institute for its generous financial support. Send correspondence to Daniel Andrei, Desautels Faculty of Management, McGill University, 1001 Sherbrooke Street West, Montréal, Québec H3A 1G5, Canada; telephone: +1 514-398-5365. E-mail: daniel.andrei@mcgill.ca.

Published by Oxford University Press on behalf of The Society for Financial Studies 2018.
doi:10.1093/rfs/Sample Advance Access publication October XXX, 2018

increases in uncertainty (Bloom, 2009; Jurado et al., 2015). A growing body of theoretical and empirical literature demonstrates the damaging effects of uncertainty on the real economy, such as depressed investment and hiring, sluggish economic performance, and declines in consumption expenditures, among many others.[1]

A basic question is how does uncertainty affect asset prices. Because uncertainty influences investors' decisions, it must affect equilibrium asset returns. In this paper, we show that investors' learning amplifies the impact of uncertainty on asset prices. This amplification seems counterintuitive: learning is expected to reduce uncertainty and to attenuate its impact. Yet we theoretically and empirically prove that learning amplifies asset price fluctuations and increases the price of risk, particularly during recessions.

We build a general equilibrium model in which a representative agent faces uncertainty about the path of future economic growth. We depart from the incomplete information literature, which assumes that the unknown dimension is the level of expected economic growth.[2] Instead, we assume that the level of expected economic growth is known, but its degree of persistence is not, that is, the agent is uncertain about the length of business cycles.

An environment with uncertainty about the length of business-cycles yields three novel implications. First, learning about persistence induces an asymmetry in the formation of beliefs. The agent revises her beliefs about persistence in bad times in the opposite direction as she does in good times.[3] In bad times, negative news about economic growth signals greater persistence. Under a preference for early resolution of uncertainty, greater persistence—that is, more long-run risk—amplifies the negative effect of news. In good times, negative news signals less persistence, which attenuates the negative effect of news. By the same logic, positive news signals less persistence in bad times (which amplifies its effect) but more persistence in good times (which attenuates its effect). Overall, learning creates an asymmetry that makes bad times riskier than good times and increases expected returns, return volatility, and the price of risk during recessions.

The second implication is that belief revisions are large around business-cycle peaks and troughs. Business-cycle peaks and troughs are *extreme times*, characterized by strongly positive or negative economic growth. During such times, changes in economic growth are highly

[1] For a comprehensive review of this literature, see Bloom (2014). See also Jurado et al. (2015) and Ludvigson et al. (2018) and the references therein.

[2] Extensive surveys of this literature can be found in Ziegler (2003) and Pastor and Veronesi (2009).

[3] Throughout the paper, we define good (bad) times as periods during which the level of expected economic growth is above (below) its long-term mean.

informative about the degree of persistence. Consequently, beliefs become more volatile. This leads to higher volatility and risk premiums not only during severe recessions but also during strong economic expansions.

The third implication is that the informativeness of news varies endogenously with the state of the economy. This causes the agent's subjective uncertainty about persistence to fluctuate: uncertainty decreases in states when news is more informative about persistence (and belief revisions are large), and it increases in states when news is less informative (and belief revisions are small). Therefore, learning about persistence affects aggregate risk and uncertainty in opposite ways: in states when belief revisions are large, aggregate risk is high but uncertainty about persistence is low.

The driver of these implications is a source of risk that arises solely when persistence is uncertain. This source of risk, which we call *persistence risk*, is the covariance between changes in beliefs about persistence and changes in expected economic growth. Because belief revisions are asymmetric, persistence risk is positive in bad times and negative in good times. In extreme times, persistence risk is large in absolute value. Persistence risk is priced in equilibrium and drives all asset pricing moments. Volatility, risk premiums, and Sharpe ratios increase during bad times, when persistence risk is positive, but also during very good times, when persistence risk is negative and large in magnitude.

We use the maximum likelihood estimation (Hamilton, 1994) to calibrate the model to real gross domestic product (GDP) growth and analyst forecast data over the period Q4:1968–Q4:2016. The estimation shows that the degree of persistence in output growth varies significantly over time. Using the estimated parameters, we generate model-implied time series for stock return volatility, the risk premium, and the Sharpe ratio. These model-implied asset pricing moments align well with their empirical counterparts.

The estimation provides time series of the uncertainty about persistence and persistence risk, which we use to test our model's predictions. First, the asymmetric formation of beliefs implies that the equity risk premium, Sharpe ratio and stock return volatility increase with persistence risk. Second, learning induces large belief revisions and increases asset pricing moments around business-cycle peaks and troughs, which implies U-shaped relationships between asset pricing moments and persistence risk. Third, large belief revisions contribute to faster resolution of uncertainty about persistence, which implies a negative relationship between asset pricing moments and the uncertainty about persistence. We find supporting evidence for these theoretical predictions.

Because high persistence risk yields a large risk premium, it must positively predict future excess returns. The data confirm this predictability, particularly at 3- to 7-year horizons. Furthermore, learning implies stronger return predictability in extreme times, when economic growth is far from its long-term mean. This implication, confirmed by the data, is unique to persistence risk. In contrast, the price-dividend ratio does predict future excess returns, but its predictability does not strengthen in extreme times. This provides evidence that learning is the channel through which persistence risk is priced in the equity market.

Our work belongs to the incomplete information literature. Most studies in this literature assume that the single unobservable dimension is the level of expected output growth.[4] As Johannes et al. (2016) argue, learning about the level of expected growth ignores confounding effects, which occur when uncertainty about one quantity makes learning about other quantities more difficult. Confounding effects magnify uncertainty, which is helpful for empirically relating changes in beliefs and asset prices. In contrast, we show that standard learning about a single parameter is sufficient to generate empirically relevant asset pricing implications. We obtain this result with a unidimensional learning framework in which we focus on the persistence of expected output growth rather than on its level. None of our asset pricing implications obtain with learning about the level of expected growth.

Another related strand of literature posits that investors' uncertainty about the conditional distribution of consumption growth explains a time-varying price of risk. Johannes et al. (2016) show that multidimensional learning induces significant variation in long-run beliefs. Collin-Dufresne et al. (2016b) show that learning generates long-run risk and implies a large equilibrium risk premium under a preference for early resolution of uncertainty. Kozlowski et al. (2015) show that belief updating generates a more persistent economic response from extreme shocks than from ordinary shocks: a negative tail event can trigger "secular stagnation."[5] Hansen and Sargent (2010, 2017) propose

[4] See Detemple (1986), Veronesi (1999, 2000), Brennan and Xia (2001), Dumas et al. (2009), and Ai (2010). Two exceptions are worth noting. Pakoš (2013) analyzes an economy in which a representative agent cannot distinguish between a mild recession and a "lost decade," which exogenously introduces asymmetry and a stronger response to news in bad times. In our case, the asymmetry arises endogenously. Andrei et al. (2018) build a model in which agents with isoelastic utility disagree about the length of business cycles. They do not study the implications of learning about persistence when agents have a preference for early resolution of uncertainty.

[5] In related work, Collin-Dufresne et al. (2016a) propose experiential learning (young people update beliefs more in response to aggregate shocks than do old people) as a mechanism for time variation in the price of risk. van Nieuwerburgh and Veldkamp (2006) show that learning about productivity implies economic growth asymmetries, with sharp downturns and gradual recoveries. Croce et al. (2015) provide a foundation for the

investors' fear of model misspecification: when the economic model is unknown, the price of risk fluctuates because "a pessimist thinks that good news is temporary and that bad news endures" (Hansen and Sargent, 2010, p. 129). In other words, investors choose a model with high persistence in bad times and a model with low persistence in good times.[6]

The common theme in all of the above papers is that agents do not know the true structure of the economy. This induces uncertainty about the model structure, which resolves slowly over time. In contrast, in our framework, the agent assigns probability one to a model with time-varying but unobservable persistence and updates beliefs with full knowledge of the structure of the economy and without fear of model misspecification. Although our model features standard learning about a single parameter, we show that the impact of business-cycle shocks depends on the state of the economy: in our framework, *ordinary* shocks trigger large belief revisions and significant fluctuations in asset prices around business-cycle peaks and troughs.

More broadly, several consumption-based models obtain a countercyclical price of risk, through various approaches. One approach assumes exogenous movements in the conditional volatility of shocks driving consumption growth. Prominent examples are the long-run risk model (Bansal and Yaron, 2004) and the time-varying disaster risk model (Wachter, 2013). A second approach assumes stochastic risk aversion, like in the habit-formation models. In Campbell and Cochrane (1999), the price of risk increases during recessions because the representative agent's risk aversion depends countercyclically on the history of consumption growth. Chan and Kogan (2002) show that such countercyclicality arises in a multiple-agent setting with habit formation and heterogeneous risk aversion. In contrast, our framework assumes away exogenous variation in the conditional volatility of consumption growth or stochastic risk aversion. We also show that learning about persistence implies a U-shaped relationship between asset-pricing moments and the state of the economy, whereas this relationship remains monotone in the long-run risk and habit formation models.

We propose a novel view of fluctuations in the price of risk. In our model, investors fear stocks during recessions because a decline in expected growth signals stronger persistence and thus increases the

---

downward-sloping term structure of equity risk based on bounded rationality and limited information.

[6]Bidder and Dew-Becker (2016) propose a similar argument, whereby ambiguity-averse investors fear a "worst-case" model in which shocks to expected consumption have a half-life of 70 years (in Bansal and Yaron (2004), the half-life of shocks is 3 years). See Hansen (2014) for a comprehensive review of this literature.

price of risk. This effect arises internally, solely from learning about persistence. We estimate the parameters of this learning model and focus on its quantitative implications in a general equilibrium with Epstein-Zin preferences. We develop, test, and find empirical support for a set of predictions specific to a model with learning about persistence. Overall, our study provides evidence of learning about persistence in financial markets and describes a mechanism through which the price of risk increases not only during recessions but also during strong economic expansions.

1. Theory

In this section, we introduce an economic model with uncertainty about persistence, formalize the agent's learning problem, and characterize the equilibrium asset prices.

The economy is defined over a continuous-time horizon $[0,\infty)$. A representative agent derives utility from consumption. The agent has Kreps-Porteus preferences (Epstein and Zin, 1989; Weil, 1990) with subjective discount rate β, relative-risk aversion γ, and elasticity of intertemporal substitution ψ. The indirect utility function is given by

$$J_t = \mathbb{E}_t\left[\int_t^\infty h(C_s,J_s)ds\right], \tag{1}$$

where the aggregator h is defined like in Duffie and Epstein (1992):

$$h(C,J) = \frac{\beta}{1-1/\psi}\left(\frac{C^{1-1/\psi}}{[(1-\gamma)J]^{1/\phi-1}} - (1-\gamma)J\right), \tag{2}$$

with $\phi \equiv \frac{1-\gamma}{1-1/\psi}$. The agent prefers early resolution of uncertainty when $1-\phi>0$.

The process of aggregate output in the economy evolves according to

$$d\delta_t/\delta_t = f_t dt + \sigma_\delta dW_t^\delta. \tag{3}$$

The agent observes the expected output growth f_t, which is measured by the average growth forecast obtained from a survey of professional forecasters. This average forecast mean reverts toward a long-term mean $\bar{f}$ according to

$$df_t = \theta_t(\bar{f}-f_t)dt + \sigma_f dW_t^f. \tag{4}$$

In addition to the average forecast f_t, the agent observes a measure of cross-sectional dispersion from the panel of forecasts. The forecast dispersion follows a square root process (Cox et al., 1985) with long-term mean $\bar{g}$, mean-reversion speed κ_g, and volatility parameter σ_g:

$$dg_t = \kappa_g(\bar{g}-g_t)dt + \sigma_g\sqrt{g_t}\left(\rho dW_t^f + \sqrt{1-\rho^2}dW_t^g\right). \tag{5}$$

Equation (5) is motivated by the fact that the forecast dispersion is typically measured by the difference between the 75th percentile and the 25th percentile of the forecasts. This difference is always positive, it has a positive long-term mean, and it mean reverts. These properties imply that the square root process is particularly well suited to model the forecast dispersion.[7]

The agent operates under incomplete information. Specifically, the agent does not observe the parameter θ_t in Equation (4), that is, the persistence of the average forecast. This parameter has two components: an observable long-term average and an unobservable, time-varying noise with zero mean. We thus define $\theta_t \equiv \bar{\theta} + \lambda_t$, where λ_t follows

$$d\lambda_t = -\kappa\lambda_t dt + \sigma_\lambda dW_t^\lambda. \tag{6}$$

The standard Brownian motions W_t^δ, W_t^f, W_t^g, and W_t^λ are mutually independent, and the parameters $\bar{f}$, $\bar{\theta}$, κ_g, κ, σ_δ, σ_f, σ_g, σ_λ, and ρ are known constants.

We make several simplifying assumptions. First, the conditional volatilities of shocks have no exogenous movements (σ_δ, σ_f, σ_g, and σ_λ are constant). Second, the agent knows the true structure of the economy (i.e., the model specification and its parameters) and perfectly trusts this structure, without fear of model misspecification or ambiguity aversion. Third, learning is unidimensional: the only unobservable variable is λ_t. These assumptions are made to isolate the effects of learning about persistence. Stochastic volatility, fear of model misspecification, and multidimensional learning have been the focus of separate studies.[8]

1.1 Learning

The agent continuously observes the aggregate output process δ_t, the average forecast f_t and the dispersion among forecasters g_t, but does not observe the persistence parameter λ_t. Denote the information set of the agent at time t by $\mathcal{F}_t$, the estimated unobservable component of the mean-reversion speed by $\widehat{\lambda}_t \equiv \mathbb{E}[\lambda_t|\mathcal{F}_t]$, and its posterior variance by $\nu_{\lambda,t} \equiv \mathbb{E}[(\lambda_t - \widehat{\lambda}_t)^2|\mathcal{F}_t]$. Standard Kalman filtering implies

$$\lambda_t \sim N(\widehat{\lambda}_t, \nu_{\lambda,t}), \tag{7}$$

[7] In Appendix A, we provide an information-processing interpretation for the forecast dispersion dynamics assumed in Equation (5). We show that, when two forecasters disagree about the long-term mean of the expected output growth rate (Berrada, 2006) and about the informativeness of a public news signal (Dumas et al., 2009), the resultant forecast dispersion dynamics can be written like in Equation (5).

[8] The long-run risk literature focuses on stochastic variation in the conditional volatility of shocks (Bansal et al., 2012; Bansal and Yaron, 2004). Hansen and Sargent (2010, 2017), among others, study fear of model misspecification. Johannes et al. (2016) analyze multidimensional learning.

where $N(m,v)$ is the Normal distribution with mean m and variance v.

In what follows, we refer to the estimate $\widehat{\lambda}_t$ as the *filter* and to the posterior variance $\nu_{\lambda,t}$ as the *uncertainty about persistence*. The following proposition obtains from filtering theory (Liptser and Shiryaev, 2001), the proof of which is provided in Appendix B.

Proposition 1 (**Learning about persistence**)**.** The filter evolves according to

$$d\widehat{\lambda}_t = -\kappa\widehat{\lambda}_t dt + \frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f} d\widehat{W}_t^f - \frac{\rho}{\sqrt{1-\rho^2}}\frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f} d\widehat{W}_t^g, \quad (8)$$

where $\widehat{W}_t^f$ and $\widehat{W}_t^g$ are independent standard Brownian motions under the filtration $\mathcal{F}_t$. Equations (B12) and (B13) provide their definitions.

The uncertainty about persistence is locally deterministic. It evolves according to

$$\frac{d\nu_{\lambda,t}}{dt} = \sigma_\lambda^2 - 2\kappa\nu_{\lambda,t} - \frac{(\bar{f}-f_t)^2\nu_{\lambda,t}^2}{\sigma_f^2(1-\rho^2)}. \quad (9)$$

The two independent Brownian motions driving the dynamics of the filter, $\widehat{W}_t^f$ and $\widehat{W}_t^g$, are defined such that under the filtration of the agent, f_t and g_t evolve according to

$$df_t = (\bar{\theta}+\widehat{\lambda}_t)(\bar{f}-f_t)dt + \sigma_f d\widehat{W}_t^f, \quad (10)$$

$$dg_t = \kappa_g(\bar{g}-g_t)dt + \sigma_g\sqrt{g_t}\left(\rho d\widehat{W}_t^f + \sqrt{1-\rho^2}d\widehat{W}_t^g\right). \quad (11)$$

At time t, the agent observes changes in the average forecast and in the forecast dispersion. These changes, which are generated by *forecast shocks* $d\widehat{W}_t^f$ and by *dispersion shocks* $d\widehat{W}_t^g$, are informative about the mean-reversion speed λ_t. Thus, both shocks drive the belief updating rule (Equation (8)).

To develop intuition and understand the implications of Proposition 1, it is useful to think of the unobservable degree of persistence as a regression coefficient between current values of a process and its previous values. When this regression coefficient is unobservable, the agent uses past observations to estimate it. Therefore, the agent's estimate of the degree of persistence depends on the history of the process. This result arises once the persistence of a mean-reverting process is unobservable and must be estimated. It implies that belief updating depends on the state of the economy.[9]

In the context of our model, the above intuition implies that the conditional covariance between changes in the filter $\widehat{\lambda}_t$ and changes in

[9] See also Xia (2001) for a similar intuition in a setting with learning about predictability.

the average forecast f_t is state-dependent. We label this conditional covariance *persistence risk*:

$$\text{Persistence risk} \equiv \frac{1}{dt}\text{Cov}_t[d\widehat{\lambda}_t, df_t] = (\bar{f} - f_t)\nu_{\lambda,t}. \tag{12}$$

Persistence risk differs from zero only in presence of uncertainty about persistence (when $\nu_{\lambda,t} > 0$). Persistence risk is positive in bad times (when $f_t < \bar{f}$) and negative in good times (when $f_t > \bar{f}$).[10] These properties yield three implications, which we state as corollaries.

Corollary 1. Belief updating is asymmetric: negative forecast shocks increase an agent's perceived persistence in bad times and decrease it in good times; conversely, positive forecast shocks decrease an agent's perceived persistence in bad times and increase it in good times.

This corollary follows from Equation (12), which shows that the covariance between changes in the filter and changes in the forecast depends on the distance between the long-run output growth and the current forecast, $\bar{f} - f_t$. We refer to this distance as the *output growth gap*. In bad times, when the output growth gap is positive, negative forecast shocks increase the perceived degree of persistence (i.e., $\widehat{\lambda}_t$ decreases). The opposite arises in good times, when negative forecast shocks decrease the perceived degree of persistence. Similarly, positive forecast shocks increase persistence in good times and decrease it in bad times.

Corollary 2. Forecast shocks trigger large belief revisions around business-cycle peaks and troughs, that is, when the expected growth is very high ($f_t \gg \bar{f}$) or very low ($f_t \ll \bar{f}$).

This corollary follows directly from the conditional variance of the filter,

$$\frac{1}{dt}\text{Var}_t[d\widehat{\lambda}_t] = \frac{1}{1-\rho^2}\frac{(\bar{f} - f_t)^2 \nu_{\lambda,t}^2}{\sigma_f^2}, \tag{13}$$

which increases when the output growth gap is large in absolute value. Intuitively, if f_t is far from its long-term mean, learning from forecast shocks and dispersion shocks is more effective for estimating λ_t because the signal-to-noise ratio in Equation (4) is high. Since shocks are highly informative about persistence in *extreme times* (when $f_t \gg \bar{f}$ or $f_t \ll \bar{f}$), beliefs become more volatile.

The dispersion in forecasts amplifies the conditional variance of the filter in Equation (13). As Proposition 1 shows, when the correlation ρ is nonzero, dispersion shocks are informative about the length of business

[10] Persistence risk can be negative in the same way that a stock can have a negative beta. This occurs in good times, when negative forecast shocks imply less persistence (for an agent who prefers early resolution of uncertainty, less persistence is good news).

cycles, and their informativeness increases with the magnitude of the output growth gap.

Overall, the conditional volatility of the filter is time-varying and increases when the forecast f_t moves away from its long-term mean $\bar{f}$, yielding larger belief revisions. This effect is distinct from other mechanisms proposed in the literature. In Kozlowski et al. (2015), tail events trigger large belief revisions; in Johannes et al. (2016), multidimensional learning magnifies uncertainty and generates stronger shocks to beliefs. Here, learning is unidimensional, and forecast shocks $d\widehat{W}_t^f$ are ordinary. These shocks, however, trigger large belief revisions when the growth forecast is either very high or very low.

Corollary 3. The uncertainty about persistence $\nu_{\lambda,t}$ is history-dependent. It decays faster in extreme times (when $f_t \gg \bar{f}$ or $f_t \ll \bar{f}$) and increases when the forecast f_t is close to its long-term mean $\bar{f}$ (i.e., in "normal times").

Corollaries 2 and 3 are closely related. This can be seen by rewriting the uncertainty-updating rule for $\nu_{\lambda,t}$ (Equation (9)) using the result of Corollary 2, Equation (13):

$$\frac{d\nu_{\lambda,t}}{dt} = \sigma_\lambda^2 - 2\kappa\nu_{\lambda,t} - \frac{1}{dt}\mathrm{Var}_t[d\widehat{\lambda}_t]. \tag{14}$$

According to the last term, uncertainty decays faster when the conditional variance of the filter is high. This arises in extreme times, when forecast shocks are highly informative about the degree of persistence. Once the forecast reverts to its long-term mean, the last term in (14) becomes negligible and uncertainty increases again. In short, uncertainty about persistence is time-varying. This differs from the standard uncertainty-updating rule encountered in previous work, where uncertainty is constant (e.g., Dumas et al., 2009; Scheinkman and Xiong, 2003).

Equation (14) indicates a negative relationship between the variance of the filter and uncertainty about persistence. A simple way to understand this (counterintuitive) relationship is to consider a sequence of negative forecast shocks in bad times. Each of these shocks helps resolve uncertainty but, in turn, magnifies the output growth gap. Beliefs become more volatile after each shock, but better learning decreases the uncertainty about persistence.

1.2 Equilibrium asset prices

Solving for the equilibrium involves writing the HJB equation:

$$\max_C \{h(C,J) + \mathcal{L}J\} = 0, \tag{15}$$

with the differential operator $\mathcal{L}J$ following from Itô's lemma. We guess the following value function (Benzoni et al., 2011):

$$J(C,f,\widehat{\lambda},\nu_\lambda)=\frac{C^{1-\gamma}}{1-\gamma}[\beta I(x)]^{\phi}, \tag{16}$$

where $I(x)$ is the wealth-consumption ratio, and $x\equiv[f\ \widehat{\lambda}\ \nu_\lambda]^\top$ denotes the state vector.[11]

Substituting the guess (16) into the HJB Equation (15) and imposing the market-clearing condition $C_t=\delta_t$, yields a partial differential equation (PDE) for the wealth-consumption ratio. We numerically solve this equation using Chebyshev polynomials (Judd, 1998). Appendix C describes the equilibrium and provides details about the numerical procedure.

To understand the impact of learning about persistence on the dynamics of the wealth-consumption ratio, we first characterize the signs of the partial derivatives of the wealth-consumption ratio with respect to two state variables:[12]

$$I_f>0,\ I_{\widehat{\lambda}}>0. \tag{17}$$

Let $\sigma_I(x_t)\equiv[\sigma_{If}(x_t)\ \sigma_{Ig}(x_t)]$ be the diffusion vector of the wealth-consumption ratio. As detailed in Appendix C, this vector has two elements:

$$\sigma_{If}(x_t)=\sigma_f\frac{I_f}{I}+\frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f}\frac{I_{\widehat{\lambda}}}{I}, \tag{18}$$

$$\sigma_{Ig}(x_t)=\frac{-\rho}{\sqrt{1-\rho^2}}\frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f}\frac{I_{\widehat{\lambda}}}{I}. \tag{19}$$

The first element, $\sigma_{If}(x_t)$, represents the sensitivity of the wealth-consumption ratio to forecast shocks $d\widehat{W}_t^f$; the second element, $\sigma_{Ig}(x_t)$, represents the sensitivity of the wealth-consumption ratio to dispersion shocks $d\widehat{W}_t^g$.

These sensitivities are directly driven by persistence risk. Equation (18) shows that learning about persistence increases the sensitivity of the wealth-consumption ratio to forecast shocks during bad times,

[11] The forecast dispersion g_t enters into the optimization problem only through the dispersion shocks $d\widehat{W}_t^g$. This is because the level of dispersion g_t does not affect the level of the forecast (which is observable); instead, *changes* in dispersion affect the agent's learning problem through the correlation ρ. Once learning takes place, the only relevant state variables for asset pricing are f_t, $\widehat{\lambda}_t$, and $\nu_{\lambda,t}$.

[12] The first inequality states that the wealth-consumption ratio increases with the forecast; the second inequality states that the wealth-consumption ratio decreases when the growth forecast is more persistent (i.e., when $\widehat{\lambda}_t$ is smaller). We verify these inequalities in Appendix C.2.

when persistence risk is positive (a consequence of Corollary 1). Furthermore, Equation (19) shows that the wealth-consumption ratio is also driven by dispersion shocks. This arises when dispersion shocks are correlated with forecast shocks. The sign and magnitude of both the correlation ρ and persistence risk jointly dictate the direction and magnitude of the sensitivity of the wealth-consumption ratio to dispersion shocks.

1.2.1 Risk-free rate and market price of risk. Following Duffie and Epstein (1992), the state price density is given by

$$\begin{aligned}\xi_t &= \exp\left[\int_0^t h_J(C_s,J_s)ds\right] h_C(C_t,J_t)\\ &= \exp\left[\int_0^t \left(\frac{\phi-1}{I(x_s)}-\beta\phi\right)ds\right]\beta^{\phi}C_t^{-\gamma}I(x_t)^{\phi-1}.\end{aligned} \tag{20}$$

The risk-free rate $r_{f,t}$ and the three-dimensional market price of risk m_t directly result from the dynamics of the state price density,

$$\frac{d\xi_t}{\xi_t} = -r_{f,t}dt - m_t^{\top}d\widehat{W}_t, \tag{21}$$

where $\widehat{W}_t \equiv [W_t^{\delta},\widehat{W}^f,\widehat{W}_t^g]^{\top}$ is the Brownian vector driving the output process, the growth forecast process, and the dispersion process. Itô's lemma yields

$$r_{f,t} = \beta + \frac{f_t}{\psi} - \frac{\gamma+\gamma\psi}{2\psi}\sigma_{\delta}^2 - \frac{1}{2}(1-\phi)\left[\sigma_{If}^2(x_t)+\sigma_{Ig}^2(x_t)\right], \tag{22}$$

$$m_t = \begin{bmatrix}\gamma\sigma_{\delta} & (1-\phi)\sigma_{If}(x_t) & (1-\phi)\sigma_{Ig}(x_t)\end{bmatrix}^{\top}. \tag{23}$$

Fluctuations in the growth forecast generate a procyclical risk-free rate, as observed from the second term in (22). The third term is the precautionary savings effect, which lowers the risk-free rate. When the agent prefers early resolution of uncertainty ($1-\phi>0$), fluctuations in the wealth-consumption ratio yield the last term in (22). The resultant effect is a lower risk-free rate.

The market price of risk in Equation (23) has three components. The first component is due to output growth shocks dW_t^{δ}. The second and third components arise when the agent prefers early resolution of uncertainty ($1-\phi>0$). Since persistence risk drives the wealth-consumption ratio diffusion terms $\sigma_{If}(x_t)$ and $\sigma_{Ig}(x_t)$, it also affects the price of risk required by the agent to bear forecast and dispersion shocks. The direct implication (from Corollary 1) is that the market price of risk increases during bad times, when persistence risk is positive. Furthermore, if $\rho<0$, the market price of dispersion risk is positive during bad times.

1.2.2 Asset prices. We assume that dividends are a levered claim on output (Abel, 1999):

$$D_t = e^{-\beta_d t}\delta_t^{\eta}, \tag{24}$$

where $\eta \geq 1$ is the leverage parameter, and $\beta_d > 0$ is a parameter that determines the expected growth rate of dividends. Leverage is motivated by the evidence that the volatility of dividend growth is larger than the volatility of output growth.[13] This specification does not change the learning problem of the agent.

The stock price is defined as a claim to the dividend process:

$$\frac{dD_t}{D_t} = \left(-\beta_d + \eta\mu_t + \frac{1}{2}\eta(\eta-1)\sigma_\delta^2\right)dt + \eta\sigma_\delta dW_t^\delta. \tag{25}$$

Define the price-dividend ratio as $\Pi(x_t)$. Its diffusion vector has two components:

$$\sigma_{\Pi f}(x_t) = \sigma_f \frac{\Pi_f}{\Pi} + \frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f}\frac{\Pi_{\widehat{\lambda}}}{\Pi}, \tag{26}$$

$$\sigma_{\Pi g}(x_t) = \frac{-\rho}{\sqrt{1-\rho^2}}\frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f}\frac{\Pi_{\widehat{\lambda}}}{\Pi}. \tag{27}$$

Without leverage ($\eta = 1$, $\beta_d = 0$), these two components coincide with $\sigma_{If}(x_t)$ and $\sigma_{Ig}(x_t)$ from Equations (18) and (19). With leverage, we expect (and verify in Appendix C.2) the partial derivatives of Π to satisfy the same inequalities used in Equation (17). The price-dividend ratio $\Pi(x_t)$ solves a partial differential equation that we provide in Appendix C.1.

1.2.3 Stock market volatility. The diffusion of stock returns, which we denote by σ_t, satisfies

$$\sigma_t = \begin{bmatrix}\eta\sigma_\delta & \sigma_{\Pi f}(x_t) & \sigma_{\Pi g}(x_t)\end{bmatrix}, \tag{28}$$

which, after replacement of (26) and (27), yields the stock return variance:

$$\|\sigma_t\|^2 = \eta^2\sigma_\delta^2 + \frac{\sigma_f^2\Pi_f^2}{\Pi^2} + \frac{2\Pi_f\Pi_{\widehat{\lambda}}}{\Pi^2}\underbrace{(\bar{f}-f_t)\nu_{\lambda,t}}_{=PR_t} + \frac{\Pi_{\widehat{\lambda}}^2}{\Pi^2}\underbrace{\frac{1}{1-\rho^2}\frac{(\bar{f}-f_t)^2\nu_{\lambda,t}^2}{\sigma_f^2}}_{=\frac{1}{dt}\mathrm{Var}_t[d\widehat{\lambda}_t]=\frac{1}{\sigma_f^2(1-\rho^2)}PR_t^2}, \tag{29}$$

[13] Using data from January 1969 to December 2016, the annualized CRSP dividend growth volatility is approximately 19%. Beeler and Campbell (2012) report values between 11% and 27%. See also Drechsler (2013) for numbers of similar magnitude. In comparison, the annualized output growth volatility is 1.4%.

where we denote by $PR_t \equiv (\bar{f} - f_t)\nu_{\lambda,t}$ our measure of persistence risk.

The stock return variance is directly driven by persistence risk (the third and fourth terms). Thus, Equation (29) provides a link between stock market volatility and Corollaries 1 and 2. Corollary 1 implies an asymmetric market response: the sensitivity of stock prices to news is stronger when the economy is in bad times and persistence risk is positive; in contrast, when the economy is in good times, persistence risk is negative and thus attenuates the sensitivity of stock prices to news. This asymmetric effect arises through the third term in Equation (29) and generates countercyclical stock market volatility.

In addition to this countercyclical effect, Corollary 2 implies an amplification of stock market volatility around business-cycle peaks and troughs. When the economic growth forecast is unusually low or unusually high, shocks are highly informative about persistence and trigger large belief revisions, thereby increasing the volatility of the filter (the last term in Equation (29)). In turn, this amplifies the volatility of stock returns.

In sum, a model with learning about persistence generates not only countercyclical volatility but also increased volatility around business-cycle peaks and troughs. The dispersion in forecasts amplifies the latter effect.

1.2.4 Equity risk premium and the aggregate price of risk. The equity risk premium is defined as the vector product between the diffusion of stock returns and the market price of risk, $RP_t \equiv \sigma_t m_t$. Using Equations (23) and (28), we obtain

$$RP_t = \gamma\eta\sigma_\delta^2 + (1-\phi)[\sigma_{If}(x_t)\sigma_{\Pi f}(x_t) + \sigma_{Ig}(x_t)\sigma_{\Pi g}(x_t)]. \tag{30}$$

The equity risk premium consists of two terms. The first term pertains to risk generated by output growth shocks and is constant in our model. The second term can be further developed using Equations (18), (19), (26), and (27), leading to

$$\begin{aligned} RP_t = \gamma\eta\sigma_\delta^2 + (1-\phi)\Bigg[& \frac{\sigma_f^2 I_f \Pi_f}{I\Pi} + \frac{I_f \Pi_{\widehat{\lambda}} + I_{\widehat{\lambda}}\Pi_f}{I\Pi} \underbrace{(\bar{f} - f_t)\nu_{\lambda,t}}_{=PR_t} \\ & + \frac{I_{\widehat{\lambda}}\Pi_{\widehat{\lambda}}}{I\Pi} \underbrace{\frac{1}{1-\rho^2}\frac{(\bar{f} - f_t)^2\nu_{\lambda,t}^2}{\sigma_f^2}}_{=\frac{1}{dt}\mathrm{Var}_t[d\widehat{\lambda}_t]=\frac{1}{\sigma_f^2(1-\rho^2)}PR_t^2} \Bigg]. \end{aligned} \tag{31}$$

The equity risk premium is countercyclical due to the presence of persistence risk in the second term in parentheses (Corollary 1). In addition, Corollary 2 implies a further increase in the risk premium in

very low and very high growth periods, because these are extreme times during which belief revisions are large (the last term in parentheses). This latter effect is further amplified by the dispersion in forecasts, through the correlation parameter ρ.

The aggregate price of risk in this economy (i.e., the Sharpe ratio) is defined as

$$SR_t \equiv \frac{RP_t}{\|\sigma_t\|}. \tag{32}$$

If the impact of learning about persistence is stronger in magnitude for the equity risk premium than for volatility, then we expect the Sharpe ratio to be countercyclical and U-shaped (quadratic) with respect to persistence risk.

In sum, a model with learning about persistence generates two main implications for the behavior of equilibrium asset prices. First, all asset pricing moments (the risk premium, volatility, and the Sharpe ratio) are countercyclical. This is due to the asymmetric nature of learning: negative shocks imply more persistence in bad times but less persistence in good times. It is a direct implication of Corollary 1.

The second implication, which is also unique to our model of learning about persistence, is that asset pricing moments are amplified around business-cycle peaks and troughs, that is, in unusually low and unusually high growth periods. These are periods when belief revisions are large. This implication follows from Corollary 2.

2. Predictions

An important question is whether the above implications are quantitatively large. In this section, we calibrate the model to U.S. output data and quantify the impact of learning about persistence on asset pricing moments.

2.1 Calibration

We use data from the Federal Reserve Bank of Philadelphia, at quarterly frequency from Q4:1968 to Q4:2016. The realized real GDP growth rate determines the growth rate of the output process δ_t.[14] The mean analysts' forecast for the 1-quarter-ahead real GDP growth rate constitutes our proxy for f_t. The difference between the 75th and 25th percentiles of analysts' forecasts for the 1-quarter-ahead real GDP growth rate (the analysts' forecasts dispersion) is our proxy for g_t.

[Insert Table 1 about here.]

[14] Using output rather than consumption data allows us to exploit a longer sample period (the time series of consumption forecasts starts only in Q3:1981).

Using the dynamics of the filter $\widehat{\lambda}_t$ from Equation (8), the dynamics of the uncertainty about λ_t from Equation (9), and the filtered Brownian shocks from Proposition 1, we generate model-implied paths for $\widehat{\lambda}_t$ and $\nu_{\lambda,t}$. We estimate the model using the maximum likelihood method (Hamilton, 1994) and determine the values of the parameters that provide the closest fit to realized observations. The uncertainty about λ_0 (i.e., the prior value of uncertainty) is set to the steady-state value,[15] whereas the prior value $\widehat{\lambda}_0$ is set to the long-term mean, which is zero. Appendix D provides details about the estimation.

[Insert Figure 1 about here.]

Table 1 reports the estimated parameters, and Figure 1 displays the time series of the state variables.[16] The estimated average degree of persistence is much lower than what is typically considered in the long-run risk literature. The long-term mean of the mean-reversion speed is $\bar{\theta}=1.29$ (it is approximately 0.25 in long-run risk models, e.g., Bansal and Yaron 2004). Overall, the estimation indicates a low degree of persistence in the growth forecast on average, but this persistence is significantly time varying (Figure 1, panel a), as confirmed by the high and statistically significant level of volatility σ_λ. The uncertainty about persistence $\nu_{\lambda,t}$ (Figure 1, panel b) also exhibits strong time variation. It explains 42.1% of the variation in persistence risk, while the output growth gap $\bar{f}-f_t$ explains the rest (Table D2 in Appendix D.1). Panels c and d of Figure 1 plot the mean analysts' forecast f_t and our measure of persistence risk, PR_t. Persistence risk reaches a maximum during the Great Recession of 2008, although the growth forecast is only mildly negative during the same period. Finally, the correlation ρ between the average forecast and the forecast dispersion is negative. Thus, the forecast dispersion is countercyclical. This finding is consistent with previous studies: van Nieuwerburgh and Veldkamp (2006) find that the dispersion of GDP forecasts across a panel of forecasters is higher during recessions; Bloom (2014) confirms this result for forecasts of U.S. industrial production growth, and the same holds for European countries (Bachmann et al., 2013). Thus, growth forecast dispersion increases during downturns.

We generate theoretical predictions using the calibration provided in Table 1. We set the risk aversion to $\gamma=10$ and the elasticity of

[15] We assume that the agent considers a (local) steady state when computing the prior on uncertainty about the mean-reversion speed. That is, the uncertainty about λ_t at time $t=0$ is the positive root of the polynomial obtained when $\frac{d\nu_{\lambda,t}}{dt}=0$. Uncertainty about persistence initially starts at this level, $\nu_{\lambda,0}$, and then dynamically evolves according to Equation (9).

[16] See also Table D1 in Appendix D.1 for the descriptive statistics and further discussion of these variables.

intertemporal substitution (EIS) to $\psi = 1.5$. The leverage parameter is set to $\eta = 7$, the subjective discount rate to $\beta = 0.02$, and $\beta_d = 0.15$. This implies a dividend growth volatility of 10%, which is a lower bound of what is typically considered.[17] The parameters β_d and β are chosen to obtain reasonable values for the average wealth-consumption ratio (Lustig et al., 2013), price-dividend ratio, and dividend growth rate, which are approximately 80, 40, and 4.5% in our model.

2.2 Asset pricing implications

We now use the calibrated economy to quantify the stock return volatility (Equation (29)), the equity risk premium (Equation (31)), and the equilibrium Sharpe ratio (Equation (32)).[18] The three panels in Figure 2 depict these theoretical asset pricing moments as a function of the forecast f_t. For these plots, we set the filter $\widehat{\lambda}_t$ at zero, which implies that the mean-reversion speed equals its long-term mean, $\widehat{\theta}_t = \bar{\theta}$. The three lines in each panel correspond to various levels of uncertainty about persistence, $\nu_{\lambda,t}$.

[Insert Figure 2 about here.]

Asset pricing moments are almost constant when $\nu_{\lambda,t} = 0$. In this case, there is no uncertainty about persistence and none of the asset pricing moments depends on the state of the economy. In contrast, in presence of uncertainty about persistence ($\nu_{\lambda,t} > 0$), asset pricing moments become, on average, countercyclical: they are higher when the growth forecast is lower. This relationship arises from the asymmetric formation of beliefs described in Corollary 1: negative forecast shocks imply more persistence when $f_t < \bar{f}$ but less persistence when $f_t > \bar{f}$. When the Epstein-Zin agent dislikes persistence, negative shocks become particularly bad in bad times, but not as bad in good times. This asymmetry generates countercyclical asset pricing moments. Moreover, higher uncertainty about persistence strengthens the negative relationship between asset pricing moments and the forecast. This occurs because higher uncertainty magnifies the asymmetric effect of learning.

Figure 2 further shows that the relationship between asset pricing moments and economic conditions is U shaped. This pattern arises because the informativeness of forecast shocks is stronger in extreme times, when the forecast is well above or well below its long-term mean. These are periods when forecast shocks trigger large belief revisions (Corollary 2); in turn, large belief revisions increase asset pricing

[17] Beeler and Campbell (2012) report values from 11% to 27%; the CRSP dividend growth volatility is 19%.

[18] Appendix C.3 provides additional results for the log price-dividend ratio and the risk-free rate.

moments. The U-shaped pattern is particularly pronounced when uncertainty about persistence is high, because higher uncertainty implies greater changes in beliefs. Overall, learning about persistence implies that asset pricing moments depend on the current state of the economy in a nonmonotonic way. In equilibrium, the asymmetric formation of beliefs increases risk in bad times. Risk further increases around business-cycle peaks and troughs, yielding a U-shaped relationship.

We reformulate these theoretical predictions in terms of persistence risk. As Equations (29) and (31) show, persistence risk is a direct driver of asset pricing moments, which allows us to develop a set of testable predictions. First, the negative relationship between asset pricing moments and the forecast implies a positive relationship between asset pricing moments and persistence risk. Second, the model implies a U-shaped (quadratic) relationship between asset pricing moments and persistence risk. Third, the relationship between asset pricing moments and persistence risk strengthens in extreme times, when shocks are more informative about persistence. Finally, uncertainty about persistence resolves more rapidly when belief revisions are large (Corollary 3), which induces a negative relation between uncertainty about persistence and asset pricing moments.

3. Evidence

We now turn to the empirical evaluation of our model. We start by examining how the model-implied asset prices compare to their empirical counterparts. We then test how the risk premium, the stock return volatility, and the Sharpe ratio vary with persistence risk and with uncertainty about persistence. Finally, we analyze the return predictability of persistence risk.

3.1 Data

The empirical analysis is based on quarterly U.S. data over the period Q1:1969–Q4:2016. The estimation performed in Section 2.1 provides time series of the filtered mean-reversion speed $\widehat{\lambda}_t$ and persistence risk, $PR_t \equiv (\bar{f} - f_t)\nu_{\lambda,t}$. Using our state variables, we construct model-implied time series for the risk premium, Sharpe ratio, stock return volatility, price-dividend ratio, and risk-free rate.

For the empirical counterparts of these asset pricing quantities, we compute quarterly stock returns and dividend growth from the value-weighted CRSP index, which covers NYSE, Amex, and Nasdaq data, and convert them into real terms using the consumer price index (CPI). We create a proxy for the ex ante risk-free rate by forecasting the ex post quarterly real return on 3-month Treasury bills with the previous

year's inflation and the most recent available 3-month nominal bill yield (Beeler and Campbell, 2012).

The price-dividend ratio is the price in the last month of the quarter divided by the sum of dividends paid in the last 12 months. Our proxy for the risk premium is the fitted value obtained by regressing excess stock returns on the lagged dividend yield (inverse of the price-dividend ratio), the lagged default premium (Baa yield minus the 10-year government bond yield), and stock return volatility.[19] Stock return volatility is the conditional volatility of real stock returns estimated with an exponential GARCH(1,1) to account for asymmetry in the sensitivity of variance to news (Nelson, 1991). Appendix E describes the data construction in detail. All variables are measured in real terms.

3.2 Descriptive analysis of asset pricing moments

In a first analysis, we compare the model-implied asset pricing moments with their empirical counterparts.

Table 2 presents unconditional descriptive statistics. The model generates an average risk premium of 6.1% and a stock return volatility of 19.8%. With the exception of the risk-free rate (for which the model delivers the volatility, but not the level) and the price-dividend ratio (for which the model delivers the level, but not the volatility), the asset pricing moments implied by our model are reasonably close to the data. Note that the standard deviation of the log price-dividend ratio is low but similar to that obtained in Bansal and Yaron (2004) and Bansal et al. (2012), although our model does not assume time variation in output growth volatility. In Table 3, we regress observed moments on model-implied moments. Except for the risk-free rate, all slopes are positive and statistically significant.

[Insert Table 2 about here.]

[Insert Table 3 about here.]

Overall, a model with learning about persistence generates reasonable asset pricing moments, both in terms of levels and dynamics, despite that we did not use financial market data in the estimation of our model: the model-implied asset pricing moments are obtained only after we calibrate the economy using data on the real GDP growth rate, the mean forecast, and the forecast dispersion across analysts.

[19] This choice is based on empirical evidence that the dividend yield (Fama and French, 1988), the default premium (Fama and French, 1989), and the level of stock market volatility (French et al., 1987) have predictive power for stock market returns.

3.3 Asset pricing with persistence risk

We now test the theoretical implications of learning about persistence. Specifically, the model predicts that the risk premium, the stock return volatility, and the Sharpe ratio vary with persistence risk, $PR_t \equiv (\bar{f} - f_t)\nu_{\lambda,t}$.

As a preliminary exercise, we split the sample into two parts according to the level of persistence risk, above and below the median. Table 4 compares the conditional asset pricing moments, in the model and the data. Asset pricing moments increase in times of higher persistence risk. The differences are statistically and economically significant in all cases. In times of high persistence risk, the risk premium, the stock return volatility, and the Sharpe ratio are higher by 2.01%, 2.12%, and 7.18%, respectively. In addition, the last column of Table 4 shows that the variance of asset pricing moments also increases when persistence risk is higher. This implies that asset pricing moments tend to fluctuate more when persistence risk increases.

[Insert Table 4 about here.]

Panel A of Table 5 presents regression results that confirm the positive relationship between asset pricing moments and persistence risk. Stock return volatility, the risk premium, and the Sharpe ratio increase in persistence risk, with statistically and economically significant coefficients. For instance, a 1-standard-deviation increase in persistence risk (0.0093) widens the empirical equity risk premium by 1.64%. These results confirm our model's prediction that aggregate risk increases when persistence risk is high.

3.3.1 U-shaped relationships. We now test the model prediction that the relationship between asset pricing moments and persistence risk is nonlinear. As suggested by our theory, panel B of Table 5 introduces a quadratic term in the regressions. The results offer strong empirical evidence in support of nonlinearity: the quadratic term is statistically significant for all moments, in the model and in the data. Furthermore, accounting for a quadratic term helps explain a larger fraction of the time variation in asset pricing moments, as measured by the increase in the R-squared coefficients from panels A to B. In the data, the explanatory power almost doubles for the risk premium and triples for the Sharpe ratio.

[Insert Table 5 about here.]

[Insert Table 6 about here.]

Our model predicts that periods of strongly negative persistence risk are also associated with a high risk premium and volatility. These are periods of intense economic growth. The reason is that news is particularly informative during both very bad and very good times, thus generating strong variability in beliefs about persistence around business-cycle peaks and troughs. We thus expect a U-shaped relationship between asset pricing moments and persistence risk, whereby large absolute values of persistence risk generate high asset pricing moments. Table 6 validates this prediction, both in the model and in the data. It reports the results from regressing asset pricing moments on persistence risk for two subsamples: the bottom quartile of persistence risk (i.e., very good times) in panel A and the remaining observations in panel B. The sign change across the two subsamples provides evidence of a U-shaped relationship, which is strongest for the risk premium and the Sharpe ratio. This nonmonotonicity confirms a prediction that is unique to our model with learning about persistence. As we discuss in the Internet Appendix, alternative specifications in a complete information environment (e.g., time-varying preferences, stochastic volatility, disaster risk) typically generate a monotone relationship between asset pricing moments and economic conditions.

3.3.2 Role of the uncertainty about persistence. Another specific implication of learning about persistence is that asset pricing moments decrease with the uncertainty about persistence. From (29)–(31), asset pricing moments increase in the conditional variance of the filter $\mathrm{Var}_t[d\widehat{\lambda}_t]$. Moreover, from Corollary 3, the conditional variance of the filter reduces the uncertainty about persistence $\nu_{\lambda,t}$ (uncertainty resolves faster with larger belief revisions). The combination of the two effects yields a negative relationship between asset pricing moments and uncertainty about persistence.

[Insert Table 7 about here.]

We test this prediction in panel A of Table 7. Uncertainty about persistence is indeed negatively related to the equity risk premium, the volatility, and the Sharpe ratio. All slope coefficients are significant, in the model and in the data. A 1-standard-deviation increase in uncertainty about persistence (0.2) reduces the equity risk premium by 2.38%. Thus, the data confirm the theoretical prediction that learning creates a negative relationship between asset pricing moments and uncertainty about persistence.

According to our model, this negative relationship strengthens in bad times, when negative forecast shocks magnify the output growth

gap $\bar{f}-f_t$. The increase in the output growth gap generated by a negative forecast shock induces both a steeper increase in asset pricing moments and a steeper decline in the uncertainty about persistence (because of the last term in Equation (14)). In contrast, in good times negative forecast shocks reduce the output growth gap and thus have a dampening effect on the relationship between uncertainty about persistence and asset pricing moments. We test this asymmetry in panel B of Table 7. To this aim, we interact $\nu_{\lambda,t}$ with a dummy that equals 1 in bad times ($f_t<\bar{f}$) and 0 otherwise. This interaction term is negative and statistically significant within the model, which indicates that the negative relationship indeed strengthens in bad times. In the data, the results go into the right direction, that is, the interaction terms are also negative, albeit not statistically significant.

The impact of $\nu_{\lambda,t}$ on asset pricing moments remains robust when including f_t to control for changes in economic conditions (panel C). This analysis confirms that the uncertainty about persistence drives asset pricing moments beyond the business-cycle fluctuations generated by f_t.[20] This implies that fluctuations in persistence risk capture not only variations in economic conditions but also changes in the uncertainty about persistence and provides further support for our channel of learning about persistence.

3.4 Return predictability

The previous section provides evidence that changes in persistence risk generate fluctuations in expected excess returns. Persistence risk should then predict future excess returns. We now test the model by evaluating the return predictability of persistence risk and contrast the results with the return predictability of the price-dividend ratio.

3.4.1 Return predictability with persistence risk. We consider the following regression specification, at a quarterly frequency:

$$\sum_{k=1}^{K}(r_{t+k}-r_{f,t+k})=a_K+b_K PR_t+\epsilon_{t+K}, \tag{33}$$

where r_{t+k} and $r_{f,t+k}$ are the log real return and real risk-free rate for quarter $t+k$. We consider different horizons: 1 year, 3 years, 5 years, and 7 years.

[Insert Table 8 about here.]

Table 8 reports the results, which are depicted in panel A for the model and in panel B for the data. Both panels show that persistence

[20] The results are also robust to controlling for NBER recessions.

risk generates strong return predictability. Persistence risk is highly statistically significant, particularly beyond the 1-year horizon. In the data, at the 5-year horizon, a 1-standard-deviation increase in persistence risk (0.0093) increases the cumulative excess return by 7.42%; the associated R^2 is 5.3%. The results are similar when we include controls,[21] which indicates that persistence risk contains additional information for explaining future excess returns, beyond what is already embedded in the price-dividend ratio, the current level of stock return volatility, or the level of macroeconomic uncertainty.

Importantly, our measure of persistence risk is only driven by shocks to the real economy (based on realizations and forecasts of output growth). Moreover, this predictability arises only when agents face uncertainty about persistence: if persistence is observable, $\nu_{\lambda,t}=0$, and the role of persistence risk disappears. Therefore, these results show that investors are compensated for the risk premium that they demand for bearing persistence risk.

3.4.2 Asymmetric predictability: High versus low informative times. When learning about persistence, not all forecast shocks are equally informative. When $f_t=\bar{f}$, changes in the forecast f_t are uninformative for the agent because she is unable to learn about the mean-reversion speed. As Corollary 2 shows, however, news become particularly informative for large absolute values of $\bar{f}-f_t$. During extreme times, learning about persistence has the greatest impact on asset prices, and the relationship between the risk premium and persistence risk strengthens. Thus, the predictability of future excess returns is expected to be concentrated during times when the expected economic growth forecast is far from its long-term mean.

[Insert Table 9 about here.]

We test this prediction by decomposing the data into two subsamples. We first consider the observations when f_t is far from the long-term average, as determined by the bottom and the top quintiles of f_t. We refer to such periods as "high informative times." Then we consider the remaining observations when f_t is relatively close to the long-term average, that is, "low informative times." Based on this sample split, we

[21] For robustness, we also consider a set of control variables that are expected to have predictive power in the data. Consistent with our construction of the risk premium in Section 3.1, the controls include the price-dividend ratio (Fama and French, 1988), stock market volatility (French et al., 1987), and the default premium (Fama and French, 1989). Because persistence risk is a measure of economic uncertainty, we also control for the macro uncertainty index of Jurado et al. (2015). Table II in the Internet Appendix reports the results in the data when using such controls.

estimate the following regression:

$$\sum_{k=1}^{K}(r_{t+k}-r_{f,t+k})=a_K+b_K\mathbb{1}_{HIT}PR_t+c_K\mathbb{1}_{LIT}PR_t+\epsilon_{t+K}, \qquad (34)$$

where $\mathbb{1}_{HIT}$ and $\mathbb{1}_{LIT}$ are dummies equal to one for observations during high (HIT) and low (LIT) informative times, respectively, and zero otherwise.

Table 9 shows that persistence risk is statistically significant for predicting future stock returns, but only during high informative times. This is when f_t is far from $\bar{f}$, either above or below.[22] Panel A shows that the finding is robust to forecast horizons ranging from 3 to 7 years. This prediction is well supported by the data (panel B), with or without controls.[23]

We also separately estimate the regressions during either high or low informative times and compare the R^2 statistics, following the predictability literature.[24] Table 10 confirms that the predictive power is concentrated when economic news is informative for updating the degree of persistence. At the 5-year horizon, whereas the R^2 in the data is 5.3% unconditionally, the conditional R^2 is 15.6% when f_t is far from $\bar{f}$ and 0% when f_t is close to $\bar{f}$.

[Insert Table 10 about here.]

Overall, these results endorse our theoretical message that persistence risk is an important determinant of future excess returns. Further, persistence risk better predicts future excess returns during high informative times—around business-cycle peaks and troughs—than during times when f_t is close to its long-term mean.

3.4.3 Return predictability with the price-dividend ratio. A large body of empirical evidence has consistently found that the price-dividend ratio predicts future stock returns.[25] Thus, we verify whether the price dividend ratio predicts future excess returns within our model. However, unlike for persistence risk, our model offers no theoretical justification for a stronger predictability of the price-dividend ratio in

[22] Separating observations by low and high values of $(\bar{f}-f_t)^2$, using the median or quartiles, yields similar results.

[23] In Table III in the Internet Appendix, we control for the log price-dividend ratio, the stock market volatility, the default premium, and the macro uncertainty index (Jurado et al., 2015). Including these controls strengthens our results.

[24] See Rapach et al. (2010), Henkel et al. (2011), and Dangl and Halling (2012).

[25] See Fama and French (1988), Hodrick (1992), Cochrane (2008), van Binsbergen and Koijen (2010), and Beeler and Campbell (2012), among others.

extreme times.[26] Thus, when separating low and high informative times, we expect the price-dividend ratio to remain equally informative about future returns. To test this, in specification (34), we replace persistence risk with the log price-dividend ratio. Panels A and B of Table 11 report results in the model and in the data.

[Insert Table 11 about here.]

The model implies strong predictability of excess returns with the price-dividend ratio. The relationships are negative and statistically significant at all horizons (1 year to 7 years), with R^2 coefficients ranging from 6% to 18%. These results are consistent with the data, whether or not we control for the level of stock market volatility, the default premium, and the macro uncertainty index of Jurado et al. (2015) (see Table IV in the Internet Appendix).

However, when we use the price-dividend ratio, we do not find asymmetry in the predictive relationship. The estimates are remarkably close across the two subsamples, both in the model and in the data. The finding that the return predictability of the price-dividend ratio remains similar across economic conditions indicates that the predictability of persistence risk is a unique implication of our model with uncertainty about persistence. Learning about persistence matters most during unusually bad and good economic conditions, and our measure of persistence risk captures this. Thus, persistence risk is a novel source of risk that is priced in the equity market, particularly around business-cycle peaks and troughs.

4. Concluding Remarks

We propose a novel view of the fluctuations in the price of risk. When the degree of persistence of expected economic growth is unknown, rational learning generates a countercyclical price of risk: investors fear stocks during recessions because a decline in expected growth signals stronger persistence. Thus, investors' learning creates *persistence risk*, which amplifies asset price fluctuations in bad times but dampens them in good times. This asymmetry implies an increasing relationship between persistence risk and asset pricing moments. The model predicts time variation in asset pricing moments despite constant conditional moments for consumption and dividend growth.

An important question is whether these results can be obtained in alternative settings. In the Internet Appendix, we thoroughly address this question by comparing our model of learning about persistence

[26] Persistence risk is not the sole driver of the price-dividend ratio, as the price-dividend ratio jointly depends on f_t, $\widehat{\lambda}_t$, and $\nu_{\lambda,t}$.

with several alternative specifications. We summarize here the main conclusions. First, in a setup without incomplete information, we show that asset pricing moments remain constant through the business cycle. The same result obtains in a setup in which the agent learns about the level of expected output growth instead of its persistence. These conclusions emphasize the importance of learning about persistence for our results. We then turn to a model in which the degree of persistence is time varying but observable. This alternative setting delivers an asymmetry, but does not generate a U-shaped relationship between asset pricing moments and persistence risk. Finally, we compare our setting with alternative models of countercyclical price of risk, that is, the model with habit formation and heterogeneous preferences proposed by Chan and Kogan (2002) and the long-run risk model with stochastic growth volatility proposed by Bansal and Yaron (2004). Both models successfully generate a countercyclical price of risk, but they neither generate the U-shaped relationships between asset pricing moments and economic conditions nor the strong long-horizon predictability of future excess returns with the price-dividend ratio.

Our framework offers several directions for future research. It can be used, for instance, to analyze the impact of macroeconomic announcements. These announcements reveal information about the future path of economic growth—including its persistence—prompting investors to demand a higher risk premium during announcement days (Savor and Wilson, 2013). Our framework also can be used to study information acquisition (e.g., investors' choice on whether to focus on level or persistence). Finally, the effects of learning about persistence on the cross-section of asset returns, as well as the effects of investors' disagreement about persistence, are promising questions to be addressed in future research.

Appendix A. An Information-Processing Interpretation for the Forecast Dispersion Dynamics

The forecast dispersion dynamics assumed in Equation (5) can be endogenously obtained in the following framework. Assume that only two professional forecasters, 1 and 2, exist. From each forecaster's point of view, the dynamics of the output growth rate and the expected output growth rate satisfy, respectively,

$$\frac{d\delta_t}{\delta_t} = \tilde{f}_t dt + \sigma_\delta dB^\delta_t \tag{A1}$$

$$d\tilde{f}_t = \kappa_t(\bar{f}_i - \tilde{f}_t)dt + \sigma_f dB^f_{it}, \qquad \text{for } i \in \{1,2\}. \tag{A2}$$

In addition, forecasters observe a news signal, which they interpret differently (e.g., they might use different forecasting models). Each forecaster believes that the dynamics of the signal is

$$ds_t = \rho_{it} dB^f_{it} + \sqrt{1-\rho^2_{it}} dB^s_{it}, \tag{A3}$$

where the three-dimensional vector $(B^\delta_t, B^f_{it}, B^s_{it})$ is a standard Brownian motion.

In this framework, forecasters have heterogeneous beliefs about the long-term mean of the expected output growth rate (Berrada, 2006) and about the informativeness of the public news signal (Dumas et al., 2009). More precisely, each forecaster has a different parameter $\bar{f}_i$ in Equation (A2) and a different correlation parameter ρ_{it} (which can be time-varying) in Equation (A3).

Forecasters do not observe the expected output growth rate and therefore apply the Kalman filter to obtain a forecast $f_{it} \equiv \mathbb{E}^i_t[\tilde{f}_t]$, where $\mathbb{E}^i_t[\cdot]$ stands for the expectation conditional on forecaster i's information set at time t. After an application of standard filtering theory (Liptser and Shiryaev, 2001), the dynamics of the two forecasts can be written under the following general form:

$$df_{1t} = \kappa_t(\bar{f}_1 - f_{1t})dt + \sigma_{f1,t} dW_{1t}, \tag{A4}$$

$$df_{2t} = \kappa_t(\bar{f}_2 - f_{2t})dt + \sigma_{f2,t}\left(\rho_{ft} dW_{1t} + \sqrt{1-\rho^2_{ft}} dW_{2t}\right), \tag{A5}$$

where the Brownian motion W_{1t} is independent W_{2t}, and ρ_{ft} measures the correlation between the two forecasts. The functional forms for $\sigma_{f1,t}$, $\sigma_{f2,t}$, and ρ_{ft} are not necessary for the arguments made here. The dynamics of the forecast dispersion, which we denote by $g_t \equiv f_{1t} - f_{2t}$, satisfy

$$dg_t = \kappa_t(\bar{f}_1 - \bar{f}_2 - g_t)dt + (\sigma_{f1,t} - \rho_{ft}\sigma_{f2,t})dW_{1t} - \sqrt{1-\rho^2_{ft}}\sigma_{f2,t} dW_{2t}. \tag{A6}$$

We are interested in the correlation between the average forecast, $f_t \equiv \frac{1}{2}(f_{1t} + f_{2t})$, and the forecast dispersion:

$$\mathrm{Corr}_t[df_t, dg_t] = \frac{\sigma^2_{f1,t} - \sigma^2_{f2,t}}{\sqrt{\sigma^2_{f1,t} + 2\rho_{ft}\sigma_{f1,t}\sigma_{f2,t} + \sigma^2_{f2,t}}\sqrt{\sigma^2_{f1,t} - 2\rho_{ft}\sigma_{f1,t}\sigma_{f2,t} + \sigma^2_{f2,t}}}. \tag{A7}$$

Assuming that the processes $\sigma_{f1,t}$, $\sigma_{f2,t}$, and ρ_{ft} are such that

$$\mathrm{Vol}_t[dg_t] = \sqrt{\sigma^2_{f1,t} - 2\rho_{ft}\sigma_{f1,t}\sigma_{f2,t} + \sigma^2_{f2,t}} \equiv \sigma_g\sqrt{g_t} \qquad \text{and} \tag{A8}$$

$$\mathrm{Corr}_t[df_t, dg_t] \equiv \rho, \tag{A9}$$

yields

$$dg_t = \kappa_g(\bar{g}-g_t)dt+\sigma_g\sqrt{g_t}\left(\rho dW_t^f+\sqrt{1-\rho^2}dW_t^g\right), \tag{A10}$$

where $\kappa_g \equiv \kappa_t$, $\bar{g}\equiv \bar{f}_1-\bar{f}_2$, and the Brownian motion W_t^f is independent of W_t^g.

Equations (A7)–(A10) show that (1) as long as $\bar{f}_1 > \bar{f}_2$, the long-term mean of the forecast dispersion is positive, like in the data, and (2) as long as $\sigma_{f1,t}^2 < \sigma_{f2,t}^2$, the correlation between the average forecast and the forecast dispersion is negative, like in the data. To summarize, the dynamics of the forecast dispersion assumed in (5) are motivated by the belief heterogeneity forecasters may have about the long-term mean of the expected output growth rate and about the informativeness of a public news signal.

Appendix B. Learning

This appendix presents a proof of Proposition 1. We use the following standard result:

Theorem 1. (Liptser and Shiryaev, 2001) Consider an unobservable process u_t and an observable process s_t with dynamics given by

$$du_t = [a_0(t,s_t)+a_1(t,s_t)u_t]dt+b_1(t,s_t)dZ_t^u+b_2(t,s_t)dZ_t^s, \tag{B1}$$

$$ds_t = [A_0(t,s_t)+A_1(t,s_t)u_t]dt+B_1(t,s_t)dZ_t^u+B_2(t,s_t)dZ_t^s. \tag{B2}$$

All the parameters can be functions of time and of the observable process. The filter and the Bayesian uncertainty evolve according to (we drop the dependence of coefficients on t and s_t for notational convenience):

$$d\widehat{u}_t = (a_0+a_1\widehat{u}_t)dt+[(b\circ B)+\nu_t A_1^\top](B\circ B)^{-1}[ds_t-(A_0+A_1\widehat{u}_t)dt], \tag{B3}$$

$$\frac{d\nu_t}{dt} = a_1\nu_t+\nu_t a_1^\top+(b\circ b)-[(b\circ B)+\nu_t A_1^\top](B\circ B)^{-1}[(b\circ B)+\nu_t A_1^\top]^\top, \tag{B4}$$

where

$$b\circ b = b_1b_1^\top+b_2b_2^\top, \tag{B5}$$

$$B\circ B = B_1B_1^\top+B_2B_2^\top, \tag{B6}$$

$$b\circ B = b_1B_1^\top+b_2B_2^\top. \tag{B7}$$

Write the dynamics of the observable variables:

$$\begin{bmatrix} d\ln\delta_t \\ df_t \\ dg_t \end{bmatrix} = \Bigg(\underbrace{\begin{bmatrix} f_t-\frac{1}{2}\sigma_\delta^2 \\ \bar{\theta}(\bar{f}-f_t) \\ \kappa_g(\bar{g}-g_t) \end{bmatrix}}_{A_0}+\underbrace{\begin{bmatrix} 0 \\ \bar{f}-f_t \\ 0 \end{bmatrix}}_{A_1}\lambda_t\Bigg)dt+\underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{B_1}dW_t^\lambda+\underbrace{\begin{bmatrix} \sigma_\delta & 0 & 0 \\ 0 & \sigma_f & 0 \\ 0 & \sigma_g\sqrt{g_t}\rho & \sigma_g\sqrt{g_t}\sqrt{1-\rho^2} \end{bmatrix}}_{B_2}\begin{bmatrix} dW_t^\delta \\ dW_t^f \\ dW_t^g \end{bmatrix}. \tag{B8}$$

The unobservable variable λ_t follows:

$$d\lambda_t = \Big(\underbrace{0}_{a_0}+\underbrace{(-\kappa)}_{a_1}\lambda_t\Big)dt+\underbrace{\sigma_\lambda}_{b_1}dW_t^\lambda+\underbrace{\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}}_{b_2}\begin{bmatrix} dW_t^\delta \\ dW_t^f \\ dW_t^g \end{bmatrix}. \tag{B9}$$

An application of Equation (B4) yields the dynamics of the posterior uncertainty:

$$\frac{d\nu_{\lambda,t}}{dt} = \sigma_\lambda^2-2\kappa\nu_{\lambda,t}-\frac{(\bar{f}-f_t)^2\nu_{\lambda,t}^2}{\sigma_f^2(1-\rho^2)}. \tag{B10}$$

The uncertainty $\nu_{\lambda,t}$ does not admit a constant steady-state solution because of the stochastic term $(\bar{f}-f_t)$ in its dynamics. To obtain the dynamics of the filter, compute first

$$\begin{bmatrix} d\ln\delta_t \\ df_t \\ dg_t \end{bmatrix} - (A_0 + A_1\widehat{\lambda}_t)dt = \begin{bmatrix} \sigma_\delta dW_t^\delta \\ (\lambda_t - \widehat{\lambda}_t)(\bar{f}-f_t)dt + \sigma_f dW_t^f \\ \sigma_g\sqrt{g_t}\rho dW_t^f + \sigma_g\sqrt{g_t}\sqrt{1-\rho^2}dW_t^g \end{bmatrix}, \tag{B11}$$

then define $d\widehat{W}_t^f$ and $d\widehat{W}_t^g$ as independent Brownian motions:

$$d\widehat{W}_t^f \equiv dW_t^f + \frac{(\lambda_t - \widehat{\lambda}_t)(\bar{f}-f_t)}{\sigma_f}dt, \tag{B12}$$

$$d\widehat{W}_t^g \equiv dW_t^g - \frac{\rho}{\sqrt{1-\rho^2}}\frac{(\lambda_t - \widehat{\lambda}_t)(\bar{f}-f_t)}{\sigma_f}dt. \tag{B13}$$

Replace this in (B11) to obtain

$$\begin{bmatrix} d\ln\delta_t \\ df_t \\ dg_t \end{bmatrix} - (A_0 + A_1\widehat{\lambda}_t)dt = \begin{bmatrix} \sigma_\delta dW_t^\delta \\ \sigma_f d\widehat{W}_t^f \\ \sigma_g\sqrt{g_t}\rho d\widehat{W}_t^f + \sigma_g\sqrt{g_t}\sqrt{1-\rho^2}d\widehat{W}_t^g \end{bmatrix}, \tag{B14}$$

which can be further replaced in Equation (B3) to obtain the dynamics of the filter:

$$d\widehat{\lambda}_t = -\kappa\widehat{\lambda}_t dt + \frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f}d\widehat{W}_t^f - \frac{\rho}{\sqrt{1-\rho^2}}\frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f}d\widehat{W}_t^g. \tag{B15}$$

Appendix C. Equilibrium

The dynamics of the vector of state variables are

$$\begin{aligned} \begin{bmatrix} d\delta_t \\ df_t \\ dg_t \\ d\widehat{\lambda}_t \\ d\nu_{\lambda,t} \end{bmatrix} &= \begin{bmatrix} \delta_t f_t \\ (\bar{\theta}+\widehat{\lambda}_t)(\bar{f}-f_t) \\ \kappa_g(\bar{g}-g_t) \\ -\kappa\widehat{\lambda}_t \\ \sigma_\lambda^2 - 2\kappa\nu_{\lambda,t} - \frac{(\bar{f}-f_t)^2\nu_{\lambda,t}^2}{\sigma_f^2(1-\rho^2)} \end{bmatrix} dt \\ &+ \begin{bmatrix} \delta_t\sigma_\delta & 0 & 0 \\ 0 & \sigma_f & 0 \\ 0 & \sigma_g\sqrt{g_t}\rho & \sigma_g\sqrt{g_t}\sqrt{1-\rho^2} \\ 0 & \frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f} & -\frac{\rho}{\sqrt{1-\rho^2}}\frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} dW_t^\delta \\ d\widehat{W}_t^f \\ d\widehat{W}_t^g \end{bmatrix}. \end{aligned} \tag{C1}$$

Proof that $I(x_t)$ is the wealth-consumption ratio The following relationship results directly from replacing the conjectured form of the value function J in the aggregator $h(C,J)$ defined in Equation (2):

$$\frac{h(C,J)}{J} = \frac{\phi}{I(x_t)} - \beta\phi. \tag{C2}$$

Define

$$W_t = C_t I(x_t), \tag{C3}$$

and replace (20) in the product $\xi_t W_t$ to obtain

$$\xi_t W_t = (1-\gamma)\exp\left[\int_0^t \left(\frac{\phi-1}{I(x_s)} - \beta\phi\right) ds\right] \frac{C_t^{1-\gamma}}{1-\gamma} [\beta I(x_t)]^{\phi} \tag{C4}$$

$$= (1-\gamma)\exp\left[\int_0^t \left(\frac{\phi-1}{I(x_s)} - \beta\phi\right) ds\right] J. \tag{C5}$$

This is a function of J and of time. Applying Itô's lemma yields:

$$d(\xi_t W_t) = (1-\gamma)\exp\left[\int_0^t \left(\frac{\phi-1}{I(x_s)} - \beta\phi\right) ds\right] \left[dJ - J\left(\beta\phi - \frac{\phi-1}{I(x_t)}\right) dt\right]. \tag{C6}$$

We also know that

$$dJ = -h(C,J)dt + dM_t, \tag{C7}$$

$$= J\left(\beta\phi - \frac{\phi}{I(x_t)}\right) dt + dM_t, \tag{C8}$$

where M_t is a martingale. The second equality follows from (C2). Replace this in (C6):

$$d(\xi_t W_t) = (1-\gamma)\exp\left[\int_0^t \left(\frac{\phi-1}{I(x_s)} - \beta\phi\right) ds\right] \left[J\left(\beta\phi - \frac{\phi}{I(x_t)}\right) dt + dM_t - J\left(\beta\phi - \frac{\phi-1}{I(x_t)}\right) dt\right], \tag{C9}$$

$$= -(1-\gamma)\exp\left[\int_0^t \left(\frac{\phi-1}{I(x_s)} - \beta\phi\right) ds\right] \frac{J}{I(x_t)} dt + d\widetilde{M}_t \tag{C10}$$

$$= -\xi_t C_t dt + d\widetilde{M}_t, \tag{C11}$$

where $d\widetilde{M}_t$ is a martingale. The third equality follows from replacing the conjectured form of the value function. The last equation can be integrated on $[t,\infty)$. Then taking the expectation and assuming that the transversality condition holds, yields the total wealth (claim to future output):

$$W_t = \mathbb{E}_t\left[\int_t^{\infty} \frac{\xi_s}{\xi_t} C_s ds\right], \tag{C12}$$

which proves that $I(x_t)$ is indeed the wealth-consumption ratio.

Partial differential equation for the wealth-consumption ratio Define the log wealth-consumption ratio:

$$i \equiv \log I. \tag{C13}$$

Substituting the guess (16) into the Hamilton-Jacobi-Bellman (HJB) Equation (15) and imposing the market-clearing condition, $C = \delta$, yields the following PDE for the log wealth consumption ratio:

$$\begin{aligned} 0 = {} & \frac{\gamma-1}{\phi}\left[-f + \frac{1}{2}\gamma\sigma_\delta^2\right] - \beta + e^{-i} \\ & + (\bar{\theta} + \widehat{\lambda})(\bar{f} - f) i_f - \kappa\widehat{\lambda} i_{\widehat{\lambda}} + \left[\sigma_\lambda^2 - 2\kappa\nu_\lambda - \frac{(\bar{f}-f)^2\nu_\lambda^2}{\sigma_f^2(1-\rho^2)}\right] i_{\nu_\lambda} \\ & + \frac{\sigma_f^2}{2} i_{ff} + \frac{(\bar{f}-f)^2\nu_\lambda^2}{2\sigma_f^2(1-\rho^2)} i_{\widehat{\lambda}\widehat{\lambda}} + \nu_\lambda(\bar{f}-f) i_{f\widehat{\lambda}} \\ & + \phi\frac{\sigma_f^2}{2} i_f^2 + \phi(\bar{f}-f)\nu_\lambda i_f i_{\widehat{\lambda}} + \phi\frac{(\bar{f}-f)^2\nu_\lambda^2}{2\sigma_f^2(1-\rho^2)} i_{\widehat{\lambda}}^2. \end{aligned} \tag{C14}$$

C.1 Levered equity

Define

$$P_t = D_t \Pi(x_t) = e^{-\beta_d t} C_t^{\eta} \Pi(x_t). \tag{C15}$$

Compute

$$\xi_t P_t = \underbrace{(1-\gamma)\exp\left(\int_0^t \left(\frac{\phi-1}{I(x_s)} - \beta\phi\right) ds - \beta_d t\right) J}_{\equiv \Delta(t)} C_t^{\eta-1} \frac{\Pi(x_t)}{I(x_t)}. \tag{C16}$$

One can clearly see that if $\eta = 1$ and $\beta_d = 0$, then $C_t^{\eta-1}$ drops out and the last fraction equals one, which brings us back to (C5). The case of interest is $\eta > 1$. Define

$$K(C_t, x_t) = C_t^{\eta-1} \frac{\Pi(x_t)}{I(x_t)} \tag{C17}$$

and thus,

$$d(\xi_t P_t) = \Delta(t)\left[-KJ\left(\frac{1}{I(x_t)} + \beta_d\right) dt + K dM_t + J dK + (dJ)(dK)\right], \tag{C18}$$

with dM_t being the same martingale used in (C8). We know that if P_t is the stock price, then we should also have

$$d(\xi_t P_t) = -\xi_t e^{-\beta_d t} C_t^{\eta} dt + d\widehat{M}_t, \tag{C19}$$

where $d\widehat{M}_t$ is a martingale. This means that the drifts in (C18) and (C19) have to be equal. This yields a partial differential equation (PDE) to be solved by $\Pi(x_t)$. Replacing $j(x_t) \equiv \ln \Pi(x_t)$ results in the following PDE:

$$\begin{aligned}
0 = {} & e^{-j} - \beta - \beta_d + \left(\eta - \frac{1}{\psi}\right) f + \frac{\sigma_\delta^2}{2}\left[\frac{\gamma(\gamma-1)}{\phi} + (\eta-1)(\eta-2\gamma)\right] \\
& + \frac{1-\phi}{2}\left(\sigma_f^2 i_f^2 + 2(\bar{f}-f)\nu_\lambda i_f i_{\widehat{\lambda}} + \frac{(\bar{f}-f)^2 \nu_\lambda^2}{\sigma_f^2(1-\rho^2)} i_{\widehat{\lambda}}^2\right) \\
& + \left[(\bar{\theta} + \widehat{\lambda})(\bar{f}-f) - (1-\phi)(\sigma_f^2 i_f + (\bar{f}-f)\nu_\lambda i_{\widehat{\lambda}})\right] j_f \\
& - \left[\kappa\widehat{\lambda} + (1-\phi)\left((\bar{f}-f)\nu_\lambda i_f + \frac{(\bar{f}-f)^2\nu_\lambda^2}{\sigma_f^2(1-\rho^2)} i_{\widehat{\lambda}}\right)\right] j_{\widehat{\lambda}} + \left[\sigma_\lambda^2 - 2\kappa\nu_\lambda - \frac{(\bar{f}-f)^2\nu_\lambda^2}{\sigma_f^2(1-\rho^2)}\right] j_{\nu_\lambda} \\
& + \frac{\sigma_f^2}{2} j_{ff} + \frac{(\bar{f}-f)^2\nu_\lambda^2}{2\sigma_f^2(1-\rho^2)} j_{\widehat{\lambda}\widehat{\lambda}} + \nu_\lambda(\bar{f}-f) j_{f\widehat{\lambda}} + \frac{\sigma_f^2}{2} j_f^2 + \frac{(\bar{f}-f)^2\nu_\lambda^2}{2\sigma_f^2(1-\rho^2)} j_{\widehat{\lambda}}^2 + \nu_\lambda(\bar{f}-f) j_f j_{\widehat{\lambda}}.
\end{aligned} \tag{C20}$$

This equation has a similar structure to (C14), except that it also involves the log wealth-consumption ratio i. It is a matter of algebra to verify that replacing $\eta = 1$ and $\beta_d = 0$ in (C20) gives exactly (C14).

In Section I of the Internet Appendix, we describe the Chebyshev collocation method (Judd, 1998) used to solve PDEs (C14) and (C20).

C.2 Evaluation of the partial derivatives I and Π

In this appendix, we discuss the two inequalities in Equation (17), which we restate here for convenience:

$$I_f > 0,\ I_{\widehat{\lambda}} > 0. \tag{C21}$$

These inequalities follow directly from the guess of the value function in (16). Taking the derivative of J with respect to any of the two state variables f and λ yields

$$J_{(\cdot)} = \phi J \frac{I_{(\cdot)}}{I}, \tag{C22}$$

with the product ϕJ being positive when $\gamma > 1 > 1/\psi$.

Because of nonsatiation, expected lifetime utility must rise as investment opportunities improve, and thus, $J_f > 0$. Using (C22), this reasoning yields $I_f > 0$.

Turning to the inequality $I_{\widehat{\lambda}} > 0$, assuming that the agent prefers early resolution of uncertainty, we expect that she prefers less persistence (i.e., a higher mean-reversion speed), which yields $J_{\widehat{\lambda}} > 0$ and $I_{\widehat{\lambda}} > 0$. In Section II of the Internet Appendix, we verify numerically that the sign of $I_{\widehat{\lambda}}/I$ is indeed positive with our calibration.

C.3 Price-dividend ratio and risk-free rate

This appendix provides additional results for Section 2.2. Figure C1 plots the log price-dividend ratio and the equilibrium risk-free rate. The log price-dividend ratio increases with the output growth forecast f_t (left panel). The relationship is almost linear, implying that Π_f/Π is positive and close to being a constant.

[Insert Figure C1 about here.]

After replacing (18) and (19) in (22), the risk-free rate becomes

$$r_{f,t} = \beta + \frac{f_t}{\psi} - \frac{\gamma + \gamma\psi}{2\psi}\sigma_\delta^2 - (1-\phi)\left[\frac{\sigma_f^2}{2}\left(\frac{I_f}{I}\right)^2 + \frac{I_f I_{\widehat{\lambda}}}{I^2}(\bar{f} - f_t)\nu_{\lambda,t} + \frac{1}{2\sigma_f^2(1-\rho^2)}\left(\frac{I_{\widehat{\lambda}}}{I}\right)^2 (\bar{f} - f_t)^2 \nu_{\lambda,t}^2\right] \tag{C23}$$

The right panel of Figure C1 depicts the equilibrium risk-free rate, which increases with the growth forecast f_t. Uncertainty about persistence decreases the risk-free rate, but its impact is weak.

Appendix D. Estimation

To fit our continuous-time model to the data, we first discretize the filtered dynamics in (C1) using the following approximations

$$\log(\delta_{t+\Delta}/\delta_t) = \left(f_t - \frac{1}{2}\sigma_\delta^2\right)\Delta + \sigma_\delta\sqrt{\Delta}v_{1,t+\Delta}, \tag{D1}$$

$$f_{t+\Delta} = e^{-\widehat{\theta}_t\Delta} f_t + \left(1 - e^{-\widehat{\theta}_t\Delta}\right)\bar{f} + \sigma_f\sqrt{\frac{1 - e^{-2\widehat{\theta}_t\Delta}}{2\widehat{\theta}_t}}v_{2,t+\Delta}, \tag{D2}$$

$$g_{t+\Delta} = e^{-\kappa_g\Delta} g_t + \left(1 - e^{-\kappa_g\Delta}\right)\bar{g} + \sigma_g\sqrt{g_t}\sqrt{\frac{1 - e^{-2\kappa_g\Delta}}{2\kappa_g}}\left(\rho v_{2,t+\Delta} + \sqrt{1-\rho^2}v_{3,t+\Delta}\right), \tag{D3}$$

$$\widehat{\lambda}_{t+\Delta}=e^{-\kappa\Delta}\widehat{\lambda}_t+\frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f}\sqrt{\frac{1-e^{-2\kappa\Delta}}{2\kappa}}\left(v_{2,t+\Delta}-\frac{\rho}{\sqrt{1-\rho^2}}v_{3,t+\Delta}\right) \tag{D4}$$

$$\nu_{\lambda,t+\Delta}=\nu_{\lambda,t}+\left[\sigma_\lambda^2-2\kappa\nu_{\lambda,t}-\left(\frac{(\bar{f}-f_t)\nu_{\lambda,t}}{\sigma_f\sqrt{1-\rho^2}}\right)^2\right]\Delta, \tag{D5}$$

where $\widehat{\theta}_t=\bar{\theta}+\widehat{\lambda}_t$ and $v_{1,t}$, $v_{2,t}$, $v_{3,t}$ are independent normally distributed random variables with mean 0 and variance 1. The time interval is $\Delta=1/4$. We use the realized real GDP growth as a proxy for the output growth $\log(\delta_{t+\Delta}/\delta_t)$, the mean analyst forecast for the 1-quarter-ahead real GDP growth as a proxy for the expected growth rate f_t, and the difference between the 75th and 25th percentiles of analysts' forecasts—the analysts' forecast dispersion—on the 1-quarter-ahead real GDP growth rate as a proxy for the dispersion g_t. The system above shows that, conditional on knowing the parameters of the model and the prior values $(\widehat{\lambda}_0,\nu_{\lambda,0})$, the time series of the realized GDP growth, GDP growth forecast, and GDP growth forecast dispersion allow us to sequentially back out the time series of the posterior values $(\widehat{\lambda}_t,\nu_{\lambda,t})$ and the noises $(v_{1,t},v_{2,t},v_{3,t})$ for $t=\Delta,2\Delta,3\Delta,\ldots$. For the initial values, we set $\widehat{\lambda}_0$ to zero, which corresponds to the long-term mean, while $\nu_{\lambda,0}$ is set to the positive root of the polynomial obtained when $\frac{d\nu_{\lambda,t}}{dt}=0$, which defines a local steady state.

The objective is to maximize the log-likelihood function $L(.;.)$

$$L(\Theta;u_\Delta,\ldots,u_{N\Delta})=\sum_{i=1}^{N}\log\left(\frac{1}{(2\pi)^{3/2}\sqrt{|\Sigma_{(i-1)\Delta}|}}\right)-\frac{1}{2}u_{i\Delta}^\top\Sigma_{(i-1)\Delta}^{-1}u_{i\Delta}, \tag{D6}$$

where $\Theta\equiv(\sigma_\delta,\bar{f},\sigma_f,\sigma_\lambda,\bar{\theta},\kappa,\rho,\sigma_g,\bar{g},\kappa_g)^\top$, N is the number of observations, $\top$ is the transpose operator, and $|.|$ is the determinant operator. The three-dimensional vector u satisfies

$$u_{t+\Delta}\equiv\begin{bmatrix}u_{1,t+\Delta}\\u_{2,t+\Delta}\\u_{3,t+\Delta}\end{bmatrix}=\begin{bmatrix}\log(\delta_{t+\Delta}/\delta_t)-\left(f_t-\frac{1}{2}\sigma_\delta^2\right)\Delta\\f_{t+\Delta}-e^{-\widehat{\theta}_t\Delta}f_t-\left(1-e^{-\widehat{\theta}_t\Delta}\right)\bar{f}\\g_{t+\Delta}-e^{-\kappa_g\Delta}g_t-(1-e^{-\kappa_g\Delta})\bar{g}\end{bmatrix}. \tag{D7}$$

Therefore, the conditional expectation and conditional variance-covariance matrix of $u_{t+\Delta}$ are

$$\mathbb{E}_t(u_{t+\Delta})=\begin{bmatrix}0\\0\\0\end{bmatrix} \tag{D8}$$

$$\Sigma_t\equiv\mathrm{Var}_t(u_{t+\Delta}) \tag{D9}$$

$$=\begin{bmatrix}\sigma_\delta^2\Delta & 0 & 0\\ 0 & \sigma_f^2\frac{1-e^{-2\widehat{\theta}_t\Delta}}{2\widehat{\theta}_t} & \sigma_g\sqrt{g_t}\sqrt{\frac{1-e^{-2\kappa_g\Delta}}{2\kappa_g}}\rho\sigma_f\sqrt{\frac{1-e^{-2\widehat{\theta}_t\Delta}}{2\widehat{\theta}_t}}\\ 0 & \sigma_g\sqrt{g_t}\sqrt{\frac{1-e^{-2\kappa_g\Delta}}{2\kappa_g}}\rho\sigma_f\sqrt{\frac{1-e^{-2\widehat{\theta}_t\Delta}}{2\widehat{\theta}_t}} & \sigma_g^2 g_t\frac{1-e^{-2\kappa_g\Delta}}{2\kappa_g}\end{bmatrix}. \tag{D10}$$

D.1 Descriptive statistics of the state variables

Table D1 reports statistics describing the level, time variation, and range of the state variables. The growth rate forecast f_t is 2.6% on average and fluctuates mostly between −1% to 6%. The demeaned mean-reversion speed $\widehat{\lambda}_t$ varies strongly over

time, fluctuating mostly between −0.7 and 0.7. The uncertainty about persistence $\nu_{\lambda,t}$ also varies substantially, fluctuating mostly between 0.4 and 1. Overall, these results suggest that the degree of persistence clearly fluctuates over time and suggests a relatively high level of uncertainty about it, on average. Similarly, persistence risk exhibits strong variations, although its mean is close to zero on average.

[Insert Table D1 about here.]

[Insert Table D2 about here.]

Table D2 documents the relative contribution of the uncertainty about persistence and the output growth gap in explaining the variation in persistence risk. To measure the contribution of uncertainty about persistence, we separate bad ($f_t < \bar{f}$) versus good ($f_t > \bar{f}$) times. The rationale is that the relation between persistence risk and $\nu_{\lambda,t}$ is expected to change sign depending on the output growth gap, which can be positive or negative. Specifically, the relation should be positive in bad times, as $\nu_{\lambda,t}$ multiplies $\bar{f} - f_t > 0$, and negative in good times, as $\nu_{\lambda,t}$ multiplies $\bar{f} - f_t < 0$.

Table D2 shows that uncertainty about persistence explains 42.1% of the variation in persistence risk. The conditional analysis indicates that the contribution is relatively stronger in bad (24.2%) than in good (17%) times. All coefficients are statistically significant and have the expected signs. When we include the output growth gap in the regression, the overall explanatory power becomes 91.6%. These results suggest that uncertainty about persistence and the output growth gap approximately explain an equal fraction of the variation in persistence risk.

Appendix E. Data Description

Real GDP growth rate and forecast data We proxy the output process with the realized Gross Domestic Product (GDP). We compute the log growth rate of the real quarterly GDP over the period 1968Q4–2016Q4. We consider the mean real GDP growth forecast for the next quarter from the Survey of Professional Forecasters as a measure of expected real GDP growth. We use the cross-sectional dispersion (75th percentile minus 25th percentile) of the real GDP growth forecast for the next quarter from the Survey of Professional Forecasters as a measure of dispersion. The first forecast observation consists of the expected real GDP growth rate for 1969Q1, as released in 1968Q4. The reported growth forecasts are annualized.

Realized and forecasted GDP data are seasonally adjusted. Real GDP growth and forecast data are obtained from the Federal Reserve Bank of Philadelphia. These series can be retrieved using the following link:

- Real GDP growth realizations and forecasts:
  https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/data-files/rgdp

Real risk-free rate We compute the real risk-free rate as the 3-month nominal yield adjusted by the expected inflation rate over the next 3 months. Like in Beeler and Campbell (2012), we first take the nominal yield on a 3-month Treasury bill $y_{3,t}$ in month t and subtract the 3-month inflation $\pi_{t,t+3}$ from period t to $t+3$ to form a measure of the ex post real 3-month interest rate. This is the dependent variable in the predictive regression below:

$$y_{3,t} - \pi_{t,t+3} = \beta_0 + \beta_1 y_{3,t} + \beta_2 \pi_{t-12,t} + \epsilon_{t+3} \tag{E1}$$

where the independent variables are the inflation over the previous year $\pi_{t-12,t}$ divided by four and the 3-month nominal yield $y_{3,t}$. The predicted value for the regression in month t gives the ex ante risk-free rate for month $t+1$. Our quarterly measure of the real risk-free rate is the annualized value at the beginning of the quarter, which we denote by $r_{f,t}$.

The nominal yield is the 3-month Treasury-bill secondary market rate, which we continuously compound as follows: $y_{3,t} = \ln(1+y_{3,t,obs}/100)/4$. Inflation is computed as the monthly log growth rate of the Consumer Price Index (CPI) from the Bureau of Labor Statistics, which is seasonally adjusted. Both series are at the monthly frequency. The data can be retrieved from the Federal Reserve Bank of St. Louis, using the following links:

- Three-month Treasury-bill rate : https://fred.stlouisfed.org/series/TB3MS
- Consumer Price Index : https://fred.stlouisfed.org/series/CPIAUCSL

Stock prices and dividends We compute the stock market price index and extract the dividends using CRSP data. The stock price index in month t is constructed as:

$$P_t = P_{t-1}(1+R_{noD,t}) \tag{E2}$$

where $R_{noD,t}$ denotes the return of a value-weighted index excluding distributions in month t.

The monthly dividend is given by

$$D_t = P_t\left(\frac{1+R_{D,t}}{1+R_{noD,t}} - 1\right) \tag{E3}$$

where $R_{D,t}$ denotes the return of a value-weighted index including distributions in month t.

The quarterly dividend is the sum of dividends within a quarter, which are not seasonally adjusted. We then calculate the log quarter-over-quarter growth rate of dividends. Dividend growth is converted from nominal to real terms using the CPI. We thus subtract log inflation to form real growth rates.

The price-dividend ratio is the price in the last month of the quarter divided by the sum of dividends paid in the last 12 months. We use the series of the value-weighted index including distributions (VWRETD) and the value-weighted index excluding distributions (VWRETX) from CRSP, which cover NYSE, Amex, and Nasdaq data.

Stock return volatility Stock return volatility is the volatility of real stock returns computed at the quarterly frequency. We first fit an AR(1) process on the quarterly log return of the stock price index and take the residuals. We then obtain the conditional volatility estimate, denoted by $Vol_{R,t}$, with the exponential GARCH(1,1) of (Nelson, 1991). We finally annualize the series. We use the quarterly value-weighted market price index including distributions from CRSP.

Realized and expected excess stock returns We first compute the quarterly real excess stock returns by subtracting the real risk-free rate from real returns. The real return is the log return of the market price index deflated by the CPI, whereas the real risk-free rate is constructed like in Section Appendix E. The realized real excess stock return $R_{X,t}$ in quarter t is thus given by:

$$R_{X,t} = \ln(1+R_{D,t}) - \pi_{t-1,t} - r_{f,t} \tag{E4}$$

where quarterly inflation $\pi_{t-1,t}$ is the log growth rate of the CPI in the final month of the current quarter over the final month in the previous quarter. We use the Consumer Price Index from the Bureau of Labor Statistics, which is seasonally adjusted.

To compute the expected excess returns, we regress the returns $R_{X,t}$ on the lagged dividend yield (measured at time $t-1$), the lagged default premium (Baa yield minus the 10-year government bond yield), and stock return volatility. The estimated expected real excess return in quarter t is the fitted value at time t, $\widehat{R}_{X,t}$. We then annualize the series. This approach follows Fama and French (1989)'s measurement procedure for estimating expected returns.

The default premium is defined as the Moody's seasoned Baa corporate bond yield relative to the yield on a 10-year constant-maturity Treasury, as available from the Federal Reserve Bank of St. Louis (https://fred.stlouisfed.org/series/BAA10YM).

The Sharpe ratio is given by the expected real excess stock returns $\widehat{R}_{X,t}$ divided by the volatility of real stock returns $Vol_{R,t}$. Both series are at the quarterly frequency.

References

Abel, A. B. (1999). Risk premia and term premia in general equilibrium. *Journal of Monetary Economics 43*(1), 3–33.

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 (2018). Asset pricing with disagreement and uncertainty about the length of business cycles. *Management Science Forthcoming*.

Bachmann, R., S. Elstner, and E. R. Sims (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics 5*(2), 217–49.

Bansal, R., D. Kiku, and A. Yaron (2012). An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review 1*(1), 183–221.

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.

Beeler, J. and J. Y. Campbell (2012). The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review 1*(1), 141–182.

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*(3), 552–573.

Berrada, T. (2006). Incomplete information, heterogeneity, and asset pricing. *Journal of Financial Econometrics 4*(1), 136–160.

Bidder, R. and I. Dew-Becker (2016). Long-run risk is the worst-case scenario. *The American Economic Review 106*(9), 2494–2527.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica 77*(3), 623–685.

Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives 28*(2), 153–76.

Brennan, M. J. and Y. Xia (2001). Stock price volatility and equity premium. *Journal of Monetary Economics 47*(2), 249–283.

Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *The Journal of Political Economy 107*(2), 205–251.

Chan, Y. L. and L. Kogan (2002). Catching up with the joneses: Heterogeneous preferences and the dynamics of asset prices. *Journal of Political Economy 110*(6), 1255–1285.

Cochrane, J. H. (2008). The dog that did not bark: a defense of return predictability. *Review of Financial Studies 21*(4), 1533–1575.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016a). Asset pricing when 'this time is different'. *The Review of Financial Studies 30*(2), 505–535.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016b). Parameter learning in general equilibrium: The asset pricing implications. *The American Economic Review 106*(3), 664–698.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*(2), 385–407.

Croce, M. M., M. Lettau, and S. C. Ludvigson (2015). Investor information, long-run risk, and the term structure of equity. *Review of Financial Studies 28*(3), 706–742.

Dangl, T. and M. Halling (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics 106*(1), 157–181.

Detemple, J. (1986). Asset pricing in a production economy with incomplete information. *The Journal of Finance 41*(2), pp. 383–391.

Drechsler, I. (2013). Uncertainty, time-varying fear, and asset prices. *The Journal of Finance 68*(5), 1843–1889.

Duffie, D. and L. G. Epstein (1992). Asset pricing with stochastic differential utility. *Review of Financial Studies 5*(3), 411–436.

Dumas, B., A. Kurshev, and R. Uppal (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *The Journal of Finance 64*(2), 579–629.

Epstein, L. G. and S. E. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*(4), 937–969.

Fama, E. F. and K. R. French (1988). Dividend yields and expected stock returns. *Journal of Financial Economics 22*(1), 3–25.

Fama, E. F. and K. R. French (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics 25*(1), 23–49.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics 19*(1), 3–29.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

Hansen, L. P. (2014). Uncertainty outside and inside economic models. *Journal of Political Economy 122*(5), 945–87.

Hansen, L. P. and T. J. Sargent (2010). Fragile beliefs and the price of uncertainty. *Quantitative Economics 1*(1), 129–162.

Hansen, L. P. and T. J. Sargent (2017). Prices of macroeconomic uncertainties with tenuous beliefs.

Henkel, S. J., J. S. Martin, and F. Nardari (2011). Time-varying short-horizon predictability. *Journal of Financial Economics 99*(3), 560–580.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies 5*(3), 357–386.

Johannes, M., L. A. Lochstoer, and Y. Mou (2016). Learning about consumption dynamics. *The Journal of Finance 71*(2), 551–600.

Judd, K. L. (1998). *Numerical Methods in Economics*. MIT press.

Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *The American Economic Review 105*(3), 1177–1216.

Kozlowski, J., L. Veldkamp, and V. Venkateswaran (2015). The tail that wags the economy: Beliefs and persistent stagnation. Technical report, National Bureau of Economic Research.

Liptser, R. S. and A. N. Shiryaev (2001). *Statistics of Random Processes II*. Springer Verlag, New York.

Ludvigson, S. C., S. Ma, and S. Ng (2018). Uncertainty and business cycles: exogenous impulse or endogenous response? Technical report, National Bureau of Economic Research.

Lustig, H., S. Van Nieuwerburgh, and A. Verdelhan (2013). The wealth-consumption ratio. *Review of Asset Pricing Studies 3*(1), 38–94.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*(2), 347–70.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–08.

Pakoš, M. (2013). Long-run risk and hidden growth persistence. *Journal of Economic Dynamics and Control 37*(9), 1911–1928.

Pastor, L. and P. Veronesi (2009). Learning in financial markets. *Annual Review of Financial Economics 1*(1), 361–381.

Rapach, D. E., J. K. Strauss, and G. Zhou (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies 23*(2), 821–862.

Savor, P. and M. Wilson (2013). How much do investors care about macroeconomic risk? evidence from scheduled economic announcements. *Journal of Financial and Quantitative Analysis 48*(2), 343–375.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of Political Economy 111*(6), 1183–1219.

van Binsbergen, J. H. and R. S. J. Koijen (2010). Predictive regressions: a present-value approach. *Journal of Finance 65*(4), 1439–1471.

van Nieuwerburgh, S. and L. Veldkamp (2006). Learning asymmetries in real business cycles. *Journal of Monetary Economics 53*(4), 753–772.

Veronesi, P. (1999). Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *Review of Financial Studies 12*(5), 975–1007.

Veronesi, P. (2000). How does information quality affect stock returns? *The Journal of Finance 55*(2), 807–837.

Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance 68*(3), 987–1035.

Weil, P. (1990). Nonexpected utility in macroeconomics. *The Quarterly Journal of Economics 105*(1), 29–42.

Xia, Y. (2001). Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *The Journal of Finance 56*(1), 205–246.

Ziegler, A. C. (2003). *Incomplete Information and Heterogeneous Beliefs in Continuous-time Finance*. Springer-Verlag Berlin Heidelberg.

Tables

Table 1
Parameter estimates

Parameter	Symbol	Value
Volatility of output growth	σ_δ	0.0140^{***} (28.31)
Long-term growth rate	$\bar{f}$	0.0272^{***} (13.43)
Volatility of growth forecast	σ_f	0.0242^{***} (31.98)
Long-term reversion speed	$\bar{\theta}$	1.2947^{***} (10.98)
Volatility of reversion speed	σ_λ	0.7809^{***} (8.12)
Reversion speed of reversion speed	κ	0.2946^{***} (9.12)
Correlation between mean forecast and forecast dispersion	ρ	-0.1613^{***} (10.35)
Long-term dispersion	$\bar{g}$	0.0137^{***} (6.81)
Reversion speed of dispersion	κ_g	0.8348^{***} (6.80)
Volatility of dispersion	σ_g	0.0935^{***} (17.57)

This table reports the estimates of the model parameters, obtained using the maximum likelihood estimation, for the period Q4:1968 to Q4:2016. t-statistics are reported in parentheses. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

Table 2
Descriptive analysis of asset pricing moments

Variable	Mean	Std dev.	Median	5-percentile	95-percentile
Risk premium					
Data	0.0447	0.0574	0.0419	-0.0342	0.1515
Model	0.0607	0.0564	0.0467	0.0264	0.1826
Stock return volatility					
Data	0.1708	0.0443	0.1573	0.1263	0.2602
Model	0.1980	0.0730	0.1830	0.1347	0.3774
Sharpe ratio					
Data	0.2333	0.2993	0.2446	-0.2766	0.7366
Model	0.2751	0.0859	0.2555	0.1990	0.4797
log (P/D)					
Data	3.6390	0.4021	3.6304	3.0266	4.3087
Model	3.4219	0.1499	3.4432	3.1501	3.5995
Real risk-free rate					
Data	0.0071	0.0094	0.0073	-0.0231	0.0393
Model	0.0343	0.0119	0.0348	0.0088	0.0539

This table reports the unconditional asset pricing moments in the model and in the data. The statistics are based on quarterly data and are annualized. All values are in real terms. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 3
Empirical versus model-implied asset pricing quantities

	Risk premium	Stock return volatility	Sharpe ratio	log (P/D)	Risk-free rate
Constant	0.023***	0.138***	0.023	1.605***	0.002
t-stat	2.823	12.097	0.279	3.467	0.507
Slope	0.363***	0.168***	0.763***	0.594***	0.143
t-stat	2.680	2.877	2.640	4.336	1.217
R-squared	.127	.077	.048	.049	.008
N	192	192	192	192	192

This table reports the relationships between the model-implied asset pricing quantities and their empirical counterparts. Observed moments are regressed on model-implied moments. The t-statistics are computed with Newey and West (1987) standard errors. *$p<.1$; **$p<.05$; ***$p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 4
Conditional asset pricing moments

	High pers. risk		**Low pers. risk**		**High-minus-low**	
Variable	Mean	SD	Mean	SD	Mean diff.	SD diff.
Risk premium						
Data	0.0547	0.0652	0.0346	0.0466	0.0201**	0.0187***
Model	0.0797	0.0742	0.0416	0.0124	0.0381***	0.0618***
Stock return volatility						
Data	0.1814	0.0524	0.1602	0.0310	0.0212***	0.0214***
Model	0.2289	0.0907	0.1670	0.0237	0.0619***	0.0670***
Sharpe ratio						
Data	0.2692	0.3226	0.1973	0.2711	0.0718*	0.0515**
Model	0.3062	0.1074	0.2441	0.0367	0.0621***	0.0707***

This table reports the asset pricing moments conditional on persistence risk. The sample is split into two parts based on the median of $(\bar{f} - f_t)\nu_{\lambda,t}$. High values correspond to periods of high persistence risk, whereas low values reflect times of low persistence risk. The means and standard deviations (SD) are based on quarterly data and are annualized. *$p < .1$; **$p < .05$; ***$p < .01$. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 5
Asset pricing moments and persistence risk

	Risk premium		**Return volatility**		**Sharpe ratio**	
	Model	Data	Model	Data	Model	Data
A. Linear relationship						
PR_t	3.927^{***}	1.765^{***}	5.608^{***}	1.613^{***}	5.845^{***}	5.660^{**}
t-stat	5.529	3.029	7.444	4.739	6.032	2.471
R-squared	.422	.082	.512	.115	.403	.031
N	192	192	192	192	192	192
B. Quadratic relationship						
PR_t	2.333^{***}	0.992^{**}	3.696^{***}	1.282^{***}	3.372^{***}	2.297
t-stat	7.419	2.551	9.956	4.171	6.619	1.184
PR_t^2	232.8^{***}	112.8^{***}	279.2^{***}	48.35^{**}	361.0^{***}	491.0^{***}
t-stat	7.944	3.049	10.362	2.075	10.255	3.611
R-squared	.755	.158	.798	.139	.748	.084
N	192	192	192	192	192	192

This table reports the relations between the asset pricing moments and persistence risk, which is defined by $PR_t \equiv (\bar{f} - f_t)\nu_{\lambda,t}$. N is the number of observations. The t-statistics are computed with Newey and West (1987) standard errors. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 6
U-shaped relationship between asset pricing moments and persistence risk

	Risk premium		**Return volatility**		**Sharpe ratio**	
	Model	Data	Model	Data	Model	Data
A. Low persistence risk (bottom quartile)						
PR_t	-1.256**	-3.350***	-1.639	-0.426	-3.868**	-20.63***
t-stat	-2.214	-3.456	-1.601	-0.489	-2.453	-3.271
R-squared	.120	.142	.062	.005	.139	.143
N	48	48	48	48	48	48
B. Persistence risk (remaining observations)						
PR_t	7.108***	3.185***	9.685***	2.406***	11.17***	11.98***
t-stat	7.783	3.380	11.99	4.430	11.14	3.421
R-squared	.669	.139	.770	.132	.738	.075
N	144	144	144	144	144	144

This table reports the U-shaped relationship between the asset pricing moments and persistence risk, defined as $PR_t \equiv (\bar{f} - f_t)\nu_{\lambda,t}$. Panel A shows the results for low persistence risk (bottom quartile), and panel B reports results for the remaining observations (top-three quartiles of persistence risk). N is the number of observations. The t-statistics are computed with Newey and West (1987) standard errors. *$p<.1$; **$p<.05$; ***$p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 7
Asset pricing moments and uncertainty about persistence

	Risk premium		**Return volatility**		**Sharpe ratio**	
	Model	Data	Model	Data	Model	Data
A. Uncertainty about persistence						
$\nu_{\lambda,t}$	-0.099***	-0.119***	-0.125***	-0.056***	-0.172***	-0.648***
t-stat	-3.872	-6.239	-4.025	-3.693	-4.804	-6.828
R-squared	.122	.170	.117	.063	.159	.186
N	192	192	192	192	192	192
B. Asymmetric impact of uncertainty about persistence						
$\nu_{\lambda,t}$	-0.029***	-0.103***	-0.039**	-0.042**	-0.090***	-0.592***
t-stat	-3.330	-4.235	-2.466	-2.424	-3.668	-4.228
$\nu_{\lambda,t}\mathbf{1}_{f_t<\bar{f}}$	-0.118***	-0.028	-0.148***	-0.024	-0.140**	-0.101
t-stat	-3.121	-0.789	-3.316	-0.854	-2.538	-0.543
R-squared	.262	.203	.312	.111	.294	.201
N	192	192	192	192	192	192
C. Uncertainty about persistence and economic conditions						
$\nu_{\lambda,t}$	-0.080***	-0.111***	-0.100***	-0.050***	-0.146***	-0.622***
t-stat	-4.864	-6.248	-5.124	-3.190	-5.782	-6.826
f_t	-2.313***	-0.950***	-3.201***	-0.724***	-3.343***	-3.308***
t-stat	-6.238	-3.250	-8.390	-3.464	-6.907	-2.958
R-squared	.579	.244	.638	.136	.571	.219
N	192	192	192	192	192	192

This table reports the relationships between the asset pricing moments and uncertainty about persistence $\nu_{\lambda,t}$. Panel A reports the univariate regression results. Panel B interacts uncertainty about persistence with a dummy that equals 1 in bad times ($f_t<\bar{f}$) and 0 otherwise (the dummy is individually included in the regression although its coefficient is unreported). Panel C controls for the growth forecast f_t. N is the number of observations. The t-statistics are computed with Newey and West (1987) standard errors. *$p<.1$; **$p<.05$; ***$p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 8
Predictability of excess returns with persistence risk

	A. Predictability (model)				***B. Predictability (data)***			
	Excess return				Excess return			
	1Y	3Y	5Y	7Y	1Y	3Y	5Y	7Y
PR_t	0.806	13.89^{***}	11.62^{**}	14.86^{***}	2.194	5.944	7.979^{**}	10.06^{***}
t-stat	0.318	3.350	2.363	5.495	1.422	1.562	2.308	3.808
R-squared	.001	.157	.089	.102	0.013	0.042	0.053	0.094
N	188	180	172	164	188	180	172	164

This table shows the predictability of cumulative future excess stock returns with persistence risk, as measured by $PR_t = (\bar{f} - f_t)\nu_{\lambda,t}$. Panel A reports the regression estimates in the model, and panel B reports the results based on empirical data. Each column represents a different forecast horizon K, and N is the number of observations. t-statistics are computed with Newey and West (1987) standard errors with $2(K-1)$ lags. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 9
Predictability of excess returns with persistence risk: Asymmetry

	A. Predictability (model)				***B. Predictability (data)***			
	Excess return				Excess return			
	1Y	3Y	5Y	7Y	1Y	3Y	5Y	7Y
Low informative times (f_t close to $\bar{f}$)								
PR_t	-4.822	13.48	16.39	21.92	-4.419	7.261	-6.982	-0.955
t-stat	-0.619	1.016	1.326	1.322	-0.753	0.949	-0.706	-0.124
High informative times (f_t far from $\bar{f}$)								
PR_t	1.293	13.92***	11.24**	14.33***	2.765*	5.833	9.194**	10.89***
t-stat	0.505	3.323	2.342	5.255	1.827	1.446	2.544	4.092
R-squared	.005	.157	.090	.104	0.024	0.043	0.068	0.103
N	188	180	172	164	188	180	172	164

This table reports the return predictability of persistence risk during high versus low informative times. Persistence risk is given by $PR_t = (\bar{f} - f_t)\nu_{\lambda,t}$. High informative times correspond to observations when the growth forecast f_t falls into its bottom or top quintiles. Low informative times include the remaining observations. Panel A reports the results in the model, and panel B reports those in the data. Each column represents a different forecast horizon K, and N is the number of observations. t-statistics are computed with Newey and West (1987) standard errors with $2(K-1)$ lags. $^*p<.1$; $^{**}p<.05$; $^{***}p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 10
Predictive power (R^2) of persistence risk for excess returns

A. Predictive power (model)

	1Y	3Y	5Y	7Y
High informative times (f_t far from $\bar{f}$)	0.000	0.241	0.119	0.198
Low informative times (f_t close to $\bar{f}$)	0.006	0.085	0.064	0.040
Unconditional	0.001	0.157	0.089	0.102

B. Predictive power (data)

	1Y	3Y	5Y	7Y
High informative times (f_t far from $\bar{f}$)	0.047	0.105	0.156	0.267
Low informative times (f_t close to $\bar{f}$)	0.004	0.006	0.000	0.002
Unconditional	0.013	0.042	0.053	0.094

This table reports the conditional predictive power of persistence risk. Predictive power is measured with the R^2 of the regression of future excess returns on persistence risk, defined as $(\bar{f} - f_t)\nu_{\lambda,t}$. We run separate regressions for high informative times, which correspond to times when f_t is far from $\bar{f}$ (bottom and top quintiles), and low informative times, which correspond to times when f_t is close to $\bar{f}$. Panel A reports the results in the model, and panel B reports those in the data. Each column represents a different forecast horizon K. The construction of excess returns is discussed in Appendix E. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table 11
Return predictability with the price-dividend ratio

	A. Predictability (model)				***B. Predictability (data)***			
	Excess return				Excess return			
	1Y	3Y	5Y	7Y	1Y	3Y	5Y	7Y
Low informative times (f_t close to $\bar{f}$)								
Log (P/D)	-0.386**	-0.920***	-0.935***	-1.152***	-0.114**	-0.209***	-0.324***	-0.366***
t-stat	-2.374	-3.639	-4.416	-8.132	-2.027	-2.681	-4.429	-4.798
High informative times (f_t far from $\bar{f}$)								
Log (P/D)	-0.386**	-0.934***	-0.942***	-1.162***	-0.140**	-0.249***	-0.396***	-0.419***
t-stat	-2.351	-3.658	-4.279	-7.710	-2.226	-2.776	-4.728	-4.949
R-squared	0.061	0.177	0.148	0.159	0.078	0.100	0.182	0.207
N	188	180	172	164	188	180	172	164

This table reports the return predictability of the log price-dividend ratio during high versus low informative times. High informative times correspond to observations when the growth forecast f_t belongs to its bottom or top quintiles. Low informative times include the remaining observations. Panel A reports results in the model, and panel B reports those in the data. Each column represents a different forecast horizon K, and N is the number of observations. t-statistics are computed with Newey and West (1987) standard errors with $2(K-1)$ lags. $^*p<.1$; $^{**}p<.05$; $^{***}p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index. The sample spans the period Q1:1969–Q4:2016.

Table D1
Descriptive statistics of the state variables

	Mean	SD	5th percentile	95th percentile
$\log\left(\frac{\delta_{t+1/4}}{\delta_t}\right)$	0.0269	0.0162	−0.0330	0.0777
f_t	0.0256	0.0166	−0.0098	0.0553
$\widehat{\lambda}_t$	0.0456	0.4055	−0.6728	0.7076
$\nu_{\lambda,t}$	0.7612	0.1991	0.4258	0.9929
PR_t	0.0009	0.0093	−0.0146	0.0189

This table reports the descriptive statistics of the state variables in the economy: the log of real output growth $\left(\frac{\delta_{t+1/4}}{\delta_t}\right)$, the forecast f_t, the filtered demeaned mean-reversion speed $\widehat{\lambda}_t$, and the uncertainty about the mean-reversion speed $\nu_{\lambda,t}$.

Table D2
Persistence risk decomposition

	PR_t			PR_t
	Good times	Bad times	All times	All times
$\nu_{\lambda,t}\mathbf{1}_{f_t>\bar{f}}$	-0.008***		-0.008***	-0.002***
t-stat	-10.08		-10.08	-5.857
$\nu_{\lambda,t}\mathbf{1}_{f_t<\bar{f}}$		0.008***	0.008***	0.003***
t-stat		9.457	9.457	8.329
$(\bar{f}-f_t)$				0.465***
t-stat				14.92
R-squared	.170	.242	.421	.916
N	192	192	192	192

This table reports the relations between persistence risk, which is defined by $PR_t \equiv (\bar{f}-f_t)\nu_{\lambda,t}$, and its components: $(\bar{f}-f_t)$ and $\nu_{\lambda,t}$. N is the number of observations. The t-statistics are computed with Newey and West (1987) standard errors. *$p<.1$; **$p<.05$; ***$p<.01$. The sample spans the period Q1:1969–Q4:2016.

Figures

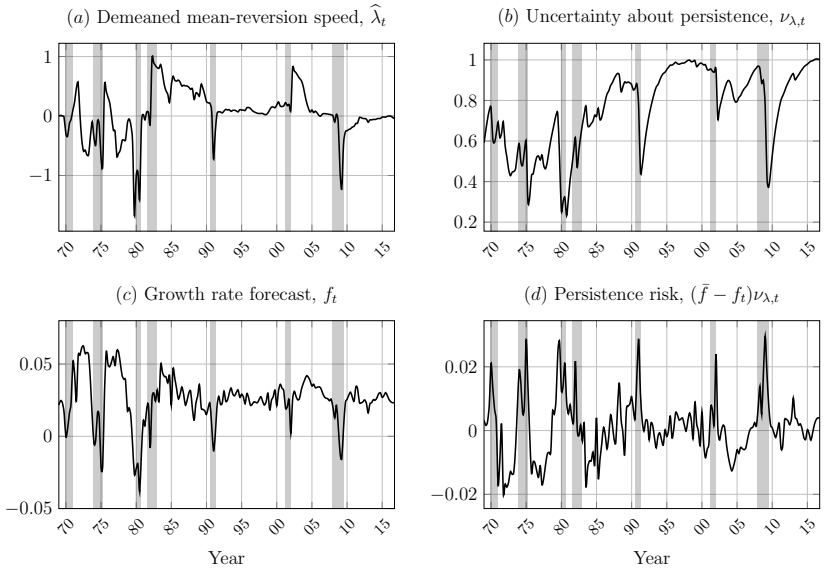

Figure 1
Historical paths of the state variables
This figure plots the time series of the state variables, from Q4:1968 to Q4:2016. Panel (a) reports the filtered demeaned mean-reversion speed of the output growth forecast. Panel (b) reports the uncertainty about persistence. Panel (c) reports the 1-quarter-ahead forecast of output growth. Panel (d) reports persistence risk.

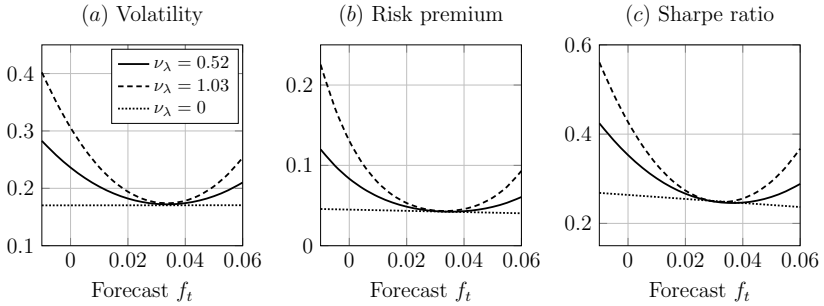

Figure 2
Asset pricing with learning about persistence
This figure shows how the stock return volatility, the risk premium, and the Sharpe ratio vary with the economic growth forecast. For the three panels, we fix $\widehat{\lambda}_t = 0$ and plot asset pricing moments against the growth forecast f_t for three different values of $\nu_{\lambda,t}$ (the upper bound of $\nu_{\lambda,t}$ is $\sigma_\lambda^2/(2\kappa) = 1.03$). We use the calibration provided in Table 1.

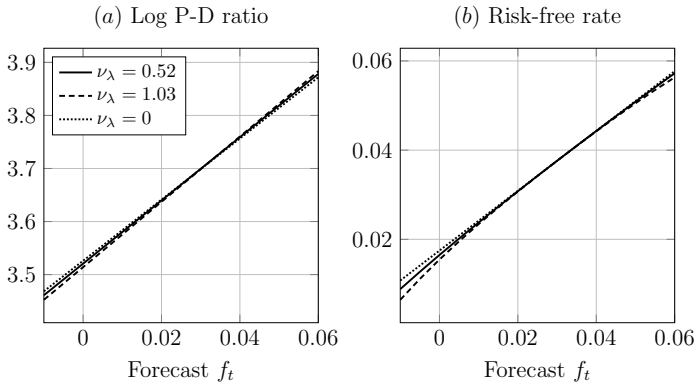

Figure C1
Behavior of the price-dividend ratio and risk-free rate with learning about persistence
This figure shows how the price-dividend ratio and the equilibrium risk-free rate vary with the forecast f_t. For the two plots, we fix $\widehat{\lambda}_t = 0$. Unless otherwise specified, we consider the calibration provided in Table 1.

Figure Legends

Figure 1
Historical paths of the state variables
This figure plots the time series of the state variables, from Q4:1968 to Q4:2016. Panel (*a*) reports the filtered demeaned mean-reversion speed of the output growth forecast. Panel (*b*) reports the uncertainty about persistence. Panel (*c*) reports the 1-quarter-ahead forecast of output growth. Panel (*d*) reports persistence risk.

Figure 2
Asset pricing with learning about persistence
This figure shows how the stock return volatility, the risk premium, and the Sharpe ratio vary with the economic growth forecast. For the three panels, we fix $\widehat{\lambda}_t = 0$ and plot asset pricing moments against the growth forecast f_t for three different values of $\nu_{\lambda,t}$ (the upper bound of $\nu_{\lambda,t}$ is $\sigma_\lambda^2/(2\kappa) = 1.03$). We use the calibration provided in Table 1.

Figure C1
Behavior of the price-dividend ratio and risk-free rate with learning about persistence
This figure shows how the price-dividend ratio and the equilibrium risk-free rate vary with the forecast f_t. For the two plots, we fix $\widehat{\lambda}_t = 0$. Unless otherwise specified, we consider the calibration provided in Table 1.

Internet Appendix for "Asset Pricing with Persistence Risk"

Daniel Andrei, Michael Hasler, and Alexandre Jeanneret

This Appendix is divided in four sections. In Section I, we describe the Chebyshev collocation method that we employ to solve the partial differential equations (C14) and (C20). In Section II, we verify numerically the inequalities $I_{\widehat{\lambda}}/I>0$ and $\Pi_{\widehat{\lambda}}/\Pi>0$. In Section III we compare our model to several alternative specifications. Finally, in Section IV we present additional empirical results to complement Section 3 of the paper.

I. Numerical solution for the PDEs (C14) and (C20)

The PDE for $i(f,\widehat{\lambda},\nu_\lambda)$ is solved numerically using the Chebyshev collocation method (Judd, 1998). That is, we approximate the function $i(f,\widehat{\lambda},\nu_\lambda)$ as follows:

$$i(f,\widehat{\lambda},\nu_\lambda)\approx P(f,\widehat{\lambda},\nu_\lambda)=\sum_{i=0}^{I}\sum_{k=0}^{K}\sum_{l=0}^{L}a_{i,k,l}T_i[f]\times T_k[\widehat{\lambda}]\times T_l[\nu_\lambda], \quad (5)$$

where $T_m[\cdot]$ is the Chebyshev polynomial of order m. The interpolation nodes are obtained by meshing the scaled roots of the Chebyshev polynomials of order $I+1$, $K+1$, and $L+1$. We scale the roots of the Chebyshev polynomials of order $I+1$, $K+1$, and $L+1$ such that they cover approximately 95% of the unconditional distributions of the 3 state variables.

The polynomial $P(f,\widehat{\lambda},\nu_\lambda)$ and its partial derivatives are then substituted into the PDE, and the resultant expression is evaluated at the interpolation nodes. This yields a system of $(I+1)\times(K+1)\times(L+1)$ equations with $(I+1)\times(K+1)\times(L+1)$ unknowns (the coefficients $a_{i,k,l}$). This system of equations is solved numerically.

Once we solve for the wealth-consumption ratio i, we replace it in Equation (C20); then we solve for the price-dividend ratio using the same procedure.

We generate a grid of 11^3 points. The mean squared PDE residuals over the set of 1,331 interpolation nodes is of order 10^{-8}. That is, the Chebyshev collocation method yields an accurate solution to the PDE.

II. Numerical verification of Equation (17)

We verify numerically that the sign of $I_{\widehat{\lambda}}/I$ is indeed positive with our calibration. We tabulate in the upper part of Table I values of $I_{\widehat{\lambda}}/I$

Table I
Values of the coefficients $I_{\widehat{\lambda}}/I$ (above) and $\Pi_{\widehat{\lambda}}/\Pi$ (below)

f \ $\widehat{\lambda}$	-0.6	-0.3	0	0.3	0.6
-0.02000	0.00132	0.00474	0.00663	0.00733	0.00718
-0.01000	0.00347	0.00560	0.00658	0.00668	0.00620
0.00000	0.00395	0.00530	0.00578	0.00562	0.00508
0.01000	0.00342	0.00428	0.00452	0.00433	0.00388
0.02000	0.00236	0.00288	0.00304	0.00292	0.00265
0.03000	0.00119	0.00138	0.00151	0.00152	0.00143
0.04000	0.00026	0.00005	0.00010	0.00020	0.00028

f \ $\widehat{\lambda}$	-0.6	-0.3	0	0.3	0.6
-0.02000	0.14000	0.18000	0.18900	0.18200	0.16700
-0.01000	0.14000	0.17000	0.17400	0.16400	0.14600
0.00000	0.12700	0.14800	0.14900	0.13900	0.12200
0.01000	0.10500	0.11900	0.11800	0.11000	0.09660
0.02000	0.07630	0.08500	0.08500	0.07950	0.07100
0.03000	0.04850	0.05190	0.05250	0.05050	0.04640
0.04000	0.02860	0.02440	0.02410	0.02430	0.02360

The tables report numerical evaluations of $I_{\widehat{\lambda}}/I$ and $\Pi_{\widehat{\lambda}}/\Pi$. For both panels, we fix $\nu_{\lambda,t}=0.52$ and build a grid for f_t and $\widehat{\lambda}_t$. Unless otherwise specified, we consider the calibration provided in Table 1.

for a two-dimensional grid on the state variables f_t and $\widehat{\lambda}_t$. The results indicate that $I_{\widehat{\lambda}}/I$ tends to decrease in good times (i.e., a larger f_t). The fact that $I_{\widehat{\lambda}}/I$ becomes smaller in good times is related to the following effect: positive shocks in good times signal not only higher persistence (which is bad news for the agent) but also a longer economic boom (which is good news). However, because the term $I_{\widehat{\lambda}}/I$ remains positive in good times, the second effect appears to be small.

The lower part of Table I tabulates values for the price-dividend ratio Π. We find coefficients $\Pi_{\widehat{\lambda}}/\Pi$ that are an order of magnitude higher than $I_{\widehat{\lambda}}/I$. The coefficients are positive at all times and also tend to decrease in good times.

III. Alternative setups

In this section, we compare our model with several alternative specifications. We analyze a model without learning in Section III.1. In Section III.2, we consider a setup in which the agent learns about the level of expected output growth. In Section III.3, we build a setting with time-varying, but observable persistence. In Section III.4, we

compare our model against two benchmarks that successfully deliver a countercyclical price of risk: the habit model (Chan and Kogan, 2002) and the long-run risk model (Bansal and Yaron, 2004).

III.1 A model without learning

We first analyze the most basic benchmark: an economy with complete information. In this standard case, the agent observes both the aggregate output and its expected growth:[1]

$$d\delta_t/\delta_t = f_t dt + \sigma_\delta dW_t^\delta \tag{6}$$

$$df_t = \bar{\theta}(\bar{f} - f_t)dt + \sigma_f dW_t^f. \tag{7}$$

Like in our main model, the two Brownian motions W_t^δ and W_t^f are mutually independent.

The volatility of stock returns and the risk premium are given by

$$\|\sigma_t\|^2 = \eta^2\sigma_\delta^2 + \left(\sigma_f \frac{\Pi_f}{\Pi}\right)^2 \tag{8}$$

$$RP_t = \gamma\eta\sigma_\delta^2 + (1-\phi)\sigma_f^2 \frac{I_f}{I}\frac{\Pi_f}{\Pi}. \tag{9}$$

The volatility, the risk premium and the Sharpe ratio in this economy are depicted by dashed lines in the three panels of column (a) in Figure I, as functions of the expected growth rate f_t. We use the calibration from Table 1. For comparison, we add solid lines that represent the asset pricing moments of our model with learning about persistence.

It is clear from the plots and directly from Eqs. (8)–(9), that a model with perfect information does not produce any variation in asset pricing moments.

III.2 A model with learning about the level of expected growth

We now turn to the case with learning about the level of expected output growth. Uncertainty about the level of expected growth is the premise of a large incomplete information literature and therefore constitutes an important benchmark.[2]

The aggregate output and the expected growth rate evolve like in (6)–(7), but the expected growth rate f_t is unobservable. The agent learns

[1] The closest model in the literature is Case I in Bansal and Yaron (2004).

[2] See Detemple (1986), Veronesi (1999), Veronesi (2000), Scheinkman and Xiong (2003), Dumas et al. (2009), and Ai (2010), among many others. Ziegler (2003) and Pastor and Veronesi (2009) provide comprehensive surveys of this literature.

from output realizations.[3] Standard filtering implies

$$d\widehat{f}_t = \bar{\theta}(\bar{f} - \widehat{f}_t)dt + \frac{\bar{\nu}_f}{\sigma_\delta} d\widehat{W}_t^\delta, \tag{10}$$

where $d\widehat{W}_t^\delta$ is the output growth innovation under the filtration of the agent, and $\bar{\nu}_f$ is the steady-state level of uncertainty about f_t:

$$\bar{\nu}_f \equiv \sigma_\delta \sqrt{(\bar{\theta}\sigma_\delta)^2 + \sigma_f^2} - \bar{\theta}\sigma_\delta^2. \tag{11}$$

Note that the uncertainty about the level of expected output growth converges to a stationary solution. This is a common result in most settings with incomplete information.

The equilibrium solution follows standard steps. We write directly the volatility of stock returns and the risk premium in this economy:

$$\sigma_t^2 = \left(\eta\sigma_\delta + \frac{\bar{\nu}_f}{\sigma_\delta}\frac{\Pi_{\widehat{f}}}{\Pi}\right)^2 \tag{12}$$

$$RP_t = \left(\eta\sigma_\delta + \frac{\bar{\nu}_f}{\sigma_\delta}\frac{\Pi_{\widehat{f}}}{\Pi}\right)\left(\gamma\sigma_\delta + (1-\phi)\frac{\bar{\nu}_f}{\sigma_\delta}\frac{I_{\widehat{f}}}{I}\right). \tag{13}$$

Eqs. (12)–(13) show that a model with learning about the level of expected growth does not generate fluctuations in volatility, the risk premium, or the Sharpe ratio, beyond the fluctuations that arise from the partial derivatives of the price-dividend ratio and of the wealth-consumption ratio. We expect these fluctuations to be relatively weak (they are zero with a log-linear approximation). Hence, standard models with learning about the level of expected growth do not generate significant time variation in asset pricing moments.[4]

The three plots of column (*a*) in Figure I confirm this result. For these plots, we use the same calibration used in Table 1. The dotted lines show that learning about the level of expected growth does not generate variations in asset pricing moments.

[3] An alternative is to assume that the agent observes a forecast of the expected growth rate and that this forecast is measured with error—i.e., the agent learns from two signals. This alternative setup yields similar results. We choose the current setup to remain closer to the existing literature.

[4] An exception is the class of models in which the expected growth rate follows a process with unobservable regime shifts (Veronesi, 1999, 2000). These models successfully generate time-varying uncertainty, volatility clustering, and time-varying expected returns. However, the implications of these alternative models are the opposite of ours: in Veronesi (1999), the reaction to news is high in good times and low in bad times. Furthermore, models with regime shifts tend to generate an inverse U-shaped relationship between fundamentals and asset pricing moments (see also Cagetti et al., 2002).

III.3 A model with time-varying but observable persistence

We consider an alternative model in which the persistence is time varying but observable. Starting with the two processes (3)–(4) of our main model, this alternative specification features an additional, observable state variable λ_t:

$$d\lambda_t = -\kappa\lambda_t dt + \sigma_\lambda \delta_\lambda dW_t^f + \sigma_\lambda \sqrt{1-\delta_\lambda^2} dW_t^\lambda. \tag{14}$$

We allow for an exogenous correlation δ_λ between the expected growth and its persistence. We consider positive values for this correlation (i.e., persistence increases after negative shocks). This model therefore has a built-in asymmetry: it generates stronger persistence during bad times and thus induces countercyclicality in asset pricing moments. Our aim is to evaluate whether it is able to generate results that are comparable with those obtained by a model with learning about persistence.

The price-dividend ratio in this economy $\Pi(f_t, \lambda_t)$ depends on two state variables, f_t and λ_t. Because these variables are driven by the Brownian motions W_t^f and W_t^λ, the price-dividend ratio has two diffusion components:

$$\sigma_{\Pi f} = \sigma_f \frac{\Pi_f}{\Pi} + \delta_\lambda \sigma_\lambda \frac{\Pi_\lambda}{\Pi} \tag{15}$$

$$\sigma_{\Pi \lambda} = \sqrt{1-\delta_\lambda^2}\, \sigma_\lambda \frac{\Pi_\lambda}{\Pi}. \tag{16}$$

Assuming that $\delta_\lambda > 0$, the first diffusion term, $\sigma_{\Pi f}$, is larger because of fluctuations in λ_t: negative expected growth shocks increase persistence, whereas positive expected growth shocks decrease persistence; the impact of these shocks is therefore amplified by the positive correlation δ_λ. Yet a direct comparison of this diffusion term with its counterpart in the presence of learning about persistence (Equation 26) reveals that our key mechanism arises only when λ_t is unobservable. It is only in this case that persistence risk directly enters into the diffusion of the price-dividend ratio in Equation (26) and magnifies it during bad times. In contrast, when λ_t is observable, the amplification occurs at all times. Consequently, this model cannot generate significant asymmetry in asset pricing moments.

Nevertheless, it might be the case that an (indirect) asymmetry arises through the partial derivatives of the price-dividend ratio in (15)-(16). To investigate this possibility, we solve the model for two different values of the correlation parameter, $\delta_\lambda \in \{0.5, 1\}$. We use the parameters from Table 1. The solution method follows the same steps taken in our main model.

In this model, the stock return volatility is

$$\|\sigma_t\|^2=\eta^2\sigma_\delta^2+\left(\sigma_f\frac{\Pi_f}{\Pi}+\delta_\lambda\sigma_\lambda\frac{\Pi_\lambda}{\Pi}\right)^2+(1-\delta_\lambda^2)\sigma_\lambda^2\left(\frac{\Pi_\lambda}{\Pi}\right)^2, \tag{17}$$

whereas the risk premium is given by

$$\begin{aligned}RP_t=&\gamma\eta\sigma_\delta^2+(1-\phi)\times\\&\times\left[\left(\sigma_f\frac{I_f}{I}+\delta_\lambda\sigma_\lambda\frac{I_\lambda}{I}\right)\left(\sigma_f\frac{\Pi_f}{\Pi}+\delta_\lambda\sigma_\lambda\frac{\Pi_\lambda}{\Pi}\right)+(1-\delta_\lambda^2)\sigma_\lambda^2\frac{I_\lambda}{I}\frac{\Pi_\lambda}{\Pi}\right].\end{aligned} \tag{18}$$

The three panels in column (*b*) of Figure I depict the results. We plot the asset pricing moments as functions of f_t for $\delta_\lambda\in\{0.5,1\}$. This alternative model delivers an asymmetry, albeit less pronounced than the asymmetry from learning about persistence. More important, this model does not deliver the U-shaped relationships when $f_t\gg\bar{f}$ or $f_t\ll\bar{f}$. This is one of the main implications of our model with learning about persistence (Corollary 2), and is supported by the data (Tables 6 and 7). It is therefore learning about persistence that induces strong variations in equilibrium risk premiums and return volatility, not the fact that the degree of persistence is time varying but observable.

III.4 Alternative models of countercyclical price of risk

We finally compare our model with two alternative explanations for the countercyclical price of risk. The first is the model with habit formation and heterogeneous preferences proposed by Chan and Kogan (2002). The second is the long-run risk model with stochastic growth volatility proposed by Bansal and Yaron (2004).

The model of Chan and Kogan (2002) (henceforth CK) belongs to a large literature that studies habit formation.[5] CK show that a setting with heterogeneous preferences supports the assumption of Campbell and Cochrane (1999) that the risk aversion of the representative agent is countercyclical. When investors' risk aversion is heterogeneous, positive shocks transfer wealth from the more risk-averse investors to the less risk-averse investors. Consequently, the aggregate risk aversion becomes countercyclical. The key state variable that drives asset pricing moments in CK is the relative consumption ω_t, defined as the log-aggregate consumption relative to the external standard of living (a low ω_t represents a bad state of the world). The three panels in column (*c*) of Figure I plot asset pricing moments against the relative consumption ω_t. For these plots, we use the calibration from CK.

[5] See Constantinides (1990), Abel (1990), Campbell and Cochrane (1999), Menzly et al. (2004), and Santos and Veronesi (2010).

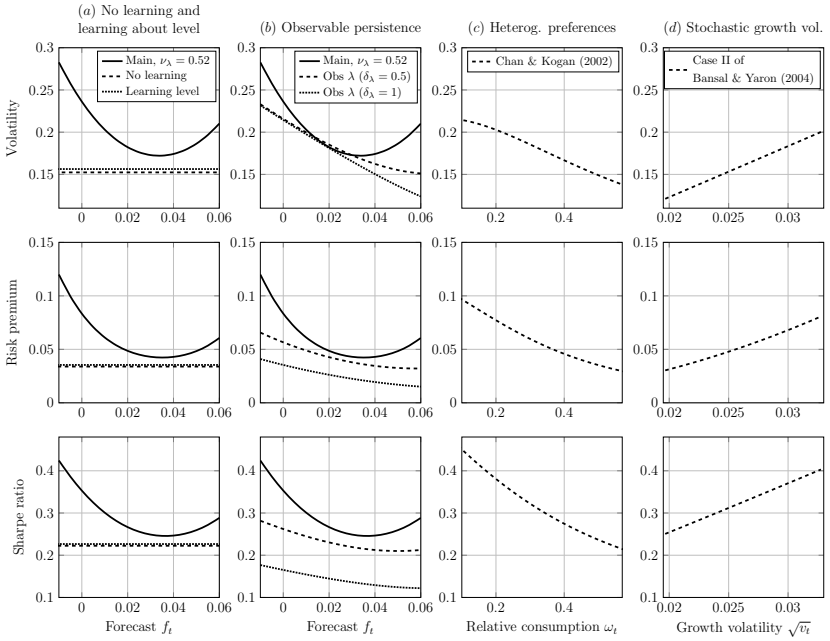

Figure I
Comparison with alternative models.
This figure compares the asset pricing implications of five alternative models with the implications of our model with learning about persistence. Column (a) plots the volatility, risk premium, and the Sharpe ratio in a model without incomplete information (Section III.1) and in a model with learning about the level of expected growth (Section III.2). Column (b) considers a model with time-varying but observable persistence (Section III.3). Column (c) plots results for the heterogeneous preferences model of Chan and Kogan (2002) (CK). Column (d) plots results for the stochastic growth volatility for Case II in Bansal and Yaron (2004) (BYII). Unless otherwise specified, we consider the calibration provided in Table 1. For columns (c) and (d), we use the calibrations from CK and BYII, respectively. The variable v_t is the continuous-time equivalent of the discrete-time variable σ_t^2 in BYII.

Another important benchmark is the long-run risk model (Bansal and Yaron, 2004)—an economy with a slow-moving component in aggregate consumption and a representative agent with Epstein and Zin (1989) preferences. Importantly, for a long-run risk model to produce countercyclical asset pricing moments, Bansal and Yaron (2004) rely on exogenous movements in the volatility of aggregate consumption growth, or "fluctuating economic uncertainty." This feature is introduced in Case II of their model (henceforth BYII). The three panels in column (d) of Figure I plot asset pricing moments against the volatility of consumption growth $\sqrt{v_t}$. For these plots, we use the calibration from BYII.

Both the heterogeneous preferences model of CK and the long-run risk model of BYII successfully generate a countercyclical price of risk.

Both models have strengths and weaknesses, and a discussion of those is beyond the scope of this paper.[6] Here, we focus on the unique theoretical predictions that our model with learning about persistence delivers. As Figure I shows, these two alternative models do not generate (with their original calibrations) the U-shaped relationships implied by learning about persistence.

Furthermore, learning about persistence delivers long-horizon return predictability with the price-dividend ratio. While the CK model does generate a negative relationship between the price-dividend ratio and future excess stock returns, the explanatory power of this relationship is smaller than in the data. The BYII model generates return predictability through the persistent variation in the volatility of consumption growth, but the strength of the long-horizon predictability is also weaker than in the data (Beeler and Campbell, 2012). In contrast, the strength of return predictability in our model is the same as that found in the data. Finally, our model features a novel source of risk that is priced in the equity market, persistence risk. We show that persistence risk predicts future returns, particularly around business-cycle peaks and troughs. This unique prediction is validated in the data.

III.4.1 Term structures of equity risk and return. An additional dimension for comparison between our model and CK and BYII is the term structure of equity risk and return. Figure II depicts the term structures of dividend strip return volatility, risk premium, and Sharpe ratio in our model (column *a*), in the CK model (column *b*), and in the BYII model (column *c*).

The model of learning about persistence yields term structures of volatility and the risk premium that are flatter than those obtained in CK and BYII. This is because the state variables in the model of learning about persistence are significantly less persistent than those in CK and BYII.[7] van Binsbergen et al. (2012) show that the strong persistence of the state variables in the long-run risk and habit formation models yields markedly upward-sloping term structures of dividend strip return volatility and the risk premium. For the Sharpe ratio, the CK model generates a flat term structure (with a single source of risk, the risk premium is a linear function of volatility), whereas the model of learning about persistence and the BYII model yield increasing term structures.

[6] Refer to Beeler and Campbell (2012), Bansal et al. (2012), and Cochrane (2017). See Xiouros and Zapatero (2010) for a discussion of the CK model.

[7] The relative consumption in CK has a half-life of approximately 12 years. The expected growth and variance of growth in BYII have half-lives of approximately 3 years and 4.5 years. In our model, the expected growth and the mean-reversion speed have half-lives of approximately 6 months and 2.5 years.

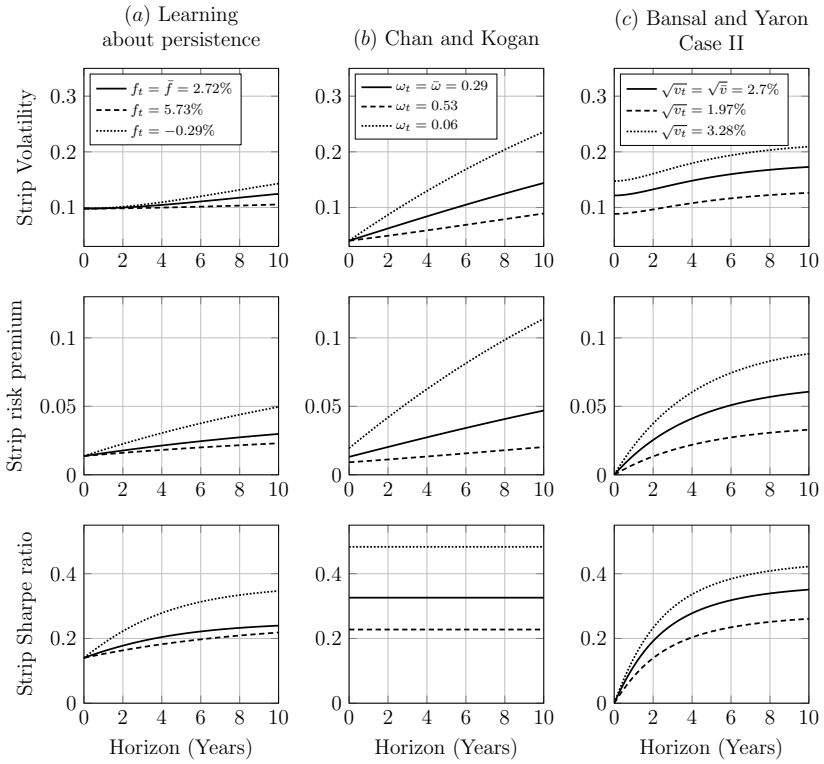

Figure II
Term structures of equity risk and return.
This figure depicts the term structures of dividend strip return volatility, the risk premium, and the Sharpe ratio in the model of learning about persistence (column *a*), in the model of Chan and Kogan (2002) (column *b*), and in Case II of Bansal and Yaron (2004) (column *c*). For the left plots, we use the calibration provided in Table 1, and we fix $\widehat{\lambda}_t = 0$ and $\nu_{\lambda,t} = 0.52$. For the middle and right plots, we use the calibrations from CK and BYII. The variable ω_t is the relative consumption. The variable v_t is the continuous-time equivalent of the discrete-time variable σ_t^2 in BYII. The upper and lower bounds for f_t, ω_t, and v_t correspond to the 97.5 and 2.5 percentiles of their distributions.

While the model of learning about persistence yields flatter term structures, we admit that it is unable to generate the downward-sloping term structures documented by van Binsbergen et al. (2012) and van Binsbergen et al. (2013). Theoretical foundations for downward-sloping term structures of risk include financial leverage, bounded rationality and limited information, post-disaster recoveries, and labor relations.[8]

[8] See Belo et al. (2015), Croce et al. (2015), Hasler and Marfe (2016), and Marfè (2017), and refer to van Binsbergen and Koijen (2017) for a recent survey on this topic.

IV. Additional empirical results

This section reproduces the predictability analysis of the paper (Section 3.4) with a set of control variables. The controls include the price-dividend ratio (Fama and French, 1988), stock market volatility (French et al., 1987), the default premium (Fama and French, 1989), and the macro uncertainty index (Jurado et al., 2015). Table II reproduces Table 8, Table III reproduces Table 9, and Table IV reproduces Table 11.

Table II
Predictability of excess returns with persistence risk

Panel A: Without controls					**Panel B: With controls**			
	Excess return				Excess return			
	1Y	3Y	5Y	7Y	1Y	3Y	5Y	7Y
PR_t	2.194	5.944	7.979^{**}	10.06^{***}	3.800^{*}	8.240^{*}	9.121^{***}	11.48^{***}
t-stat	1.422	1.562	2.308	3.808	1.865	1.889	2.922	3.045
R-squared	0.013	0.042	0.053	0.094	0.216	0.239	0.313	0.316
N	188	180	172	164	188	180	172	164

This table shows the predictability of cumulative future excess stock returns with persistence risk, as measured by $PR_t = (\bar{f} - f_t)\nu_{\lambda,t}$. Panel A reports the regression estimates in the data (like in panel B of Table 8), and panel B reports the results when controlling for the log price-dividend ratio (Fama and French, 1988), the level of stock market volatility (French et al., 1987), the default premium (Fama and French, 1989), and the macro uncertainty index of Jurado et al. (2015). Each column represents a different forecast horizon K, and N is the number of observations. t-statistics are computed with Newey and West (1987) standard errors with $2(K-1)$ lags. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table III
Asymmetric excess return predictability with persistence risk (with and without controls)

	Panel A: Without controls				**Panel B: With controls**			
	Excess return				Excess return			
	1Y	3Y	5Y	7Y	1Y	3Y	5Y	7Y
Low informative times (f_t close to $\bar{f}$)								
PR_t	-4.419	7.261	-6.982	-0.955	-5.789	5.256	-4.112	1.665
t-stat	-0.753	0.949	-0.706	-0.124	-0.891	0.982	-0.869	0.231
High informative times (f_t far from $\bar{f}$)								
PR_t	2.765*	5.833	9.194**	10.89***	4.853**	8.567*	10.44***	12.42***
t-stat	1.827	1.446	2.544	4.092	2.515	1.812	2.772	3.192
R-squared	0.024	0.043	0.068	0.103	0.239	0.240	0.325	0.323
N	188	180	172	164	188	180	172	164

This table reports the return predictability of persistence risk during high versus low informative times in the data. Persistence risk is given by $PR_t = (\bar{f} - f_t)\nu_{\lambda,t}$. High informative times correspond to observations when the growth forecast f_t falls into its bottom or top quintiles. Low informative times include the remaining observations. Panel A reports the results without controls (like in panel B of Table 9), and panel B reports the results when controlling for log price-dividend ratio, the level of stock market volatility, the default premium, and the macro uncertainty index of Jurado et al. (2015). Each column represents a different forecast horizon K, and N is the number of observations. t-statistics are computed with Newey and West (1987) standard errors with $2(K-1)$ lags. $^*p<.1$; $^{**}p<.05$; $^{***}p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

Table IV
Predictability of excess returns with the price-dividend ratio (with and without controls)

	Panel A: Without controls				**Panel B: With controls**			
	Excess return				Excess return			
	1Y	3Y	5Y	7Y	1Y	3Y	5Y	7Y
Low informative times (f_t close to $\bar{f}$)								
Log(P/D)	-0.114**	-0.209***	-0.324***	-0.366***	-0.176***	-0.291***	-0.413***	-0.412***
t-stat	-2.027	-2.681	-4.429	-4.798	-2.705	-3.650	-4.408	-3.877
High informative times (f_t far from $\bar{f}$)								
Log(P/D)	-0.140**	-0.249***	-0.396***	-0.419***	-0.194***	-0.322***	-0.476***	-0.461***
t-stat	-2.226	-2.776	-4.728	-4.949	-2.788	-3.812	-4.472	-4.302
R-squared	.078	.100	.182	.207	0.207	0.205	0.344	0.281
N	188	180	172	164	188	180	172	164

This table reports the return predictability of the log price-dividend ratio during high versus low informative times in the data. High informative times correspond to observations when the growth forecast f_t falls into its bottom or top quintiles. Low informative times include the remaining observations. Panel A reports the results without controls (like in panel B of Table 11), and panel B reports the results when controlling for the level of stock market volatility, the default premium, and the macro uncertainty index of Jurado et al. (2015). Each column represents a different forecast horizon K, and N is the number of observations. t-statistics are computed with Newey and West (1987) standard errors with $2(K-1)$ lags. *$p<.1$; **$p<005$; ***$p<.01$. Appendix E details the construction of the empirical moments. Stock prices represent the value-weighted CRSP index, and the sample spans the period Q1:1969–Q4:2016.

References

Abel, A. B. (1990, May). Asset prices under habit formation and catching up with the joneses. *American Economic Review 80*(2), 38–42.

Ai, H. (2010). Information quality and long-run risk: Asset pricing implications. *The Journal of Finance 65*(4), 1333–1367.

Bansal, R., D. Kiku, and A. Yaron (2012). An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review 1*(1), 183–221.

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.

Beeler, J. and J. Y. Campbell (2012). The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review 1*(1), 141–182.

Belo, F., P. Collin-Dufresne, and R. S. Goldstein (2015). Dividend dynamics and the term structure of dividend strips. *The Journal of Finance 70*(3), 1115–1160.

Cagetti, M., L. P. Hansen, T. Sargent, and N. Williams (2002). Robustness and pricing with uncertain growth. *Review of Financial Studies 15*(2), 363–404.

Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *The Journal of Political Economy 107*(2), 205–251.

Chan, Y. L. and L. Kogan (2002). Catching up with the joneses: Heterogeneous preferences and the dynamics of asset prices. *Journal of Political Economy 110*(6), 1255–1285.

Cochrane, J. H. (2017). Macro-finance. *Review of Finance 21*(3), 945–985.

Constantinides, G. M. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy 98*(3), 519–543.

Croce, M. M., M. Lettau, and S. C. Ludvigson (2015). Investor information, long-run risk, and the term structure of equity. *Review of Financial Studies 28*(3), 706–742.

Detemple, J. (1986). Asset pricing in a production economy with incomplete information. *The Journal of Finance 41*(2), pp. 383–391.

Dumas, B., A. Kurshev, and R. Uppal (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *The Journal of Finance 64*(2), 579–629.

Epstein, L. G. and S. E. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*(4), 937–969.

Fama, E. F. and K. R. French (1988). Dividend yields and expected stock returns. *Journal of Financial Economics 22*(1), 3–25.

Fama, E. F. and K. R. French (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics 25*(1), 23–49.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics 19*(1), 3–29.

Hasler, M. and R. Marfe (2016). Disaster recovery and the term structure of dividend strips. *Journal of Financial Economics 122*(1), 116–134.

Judd, K. L. (1998). *Numerical Methods in Economics*. MIT press.

Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *The American Economic Review 105*(3), 1177–1216.

Marfè, R. (2017). Income insurance and the equilibrium term structure of equity. *The Journal of Finance 72*(5), 2073–2130.

Menzly, L., T. Santos, and P. Veronesi (2004). Understanding predictability. *Journal of Political Economy 112*(1), 1–47.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–08.

Pastor, L. and P. Veronesi (2009). Learning in financial markets. *Annual Review of Financial Economics 1*(1), 361–381.

Santos, T. and P. Veronesi (2010). Habit formation, the cross section of stock returns and the cash-flow risk puzzle. *Journal of Financial Economics 98*(2), 385–413.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of Political Economy 111*(6), 1183–1219.

van Binsbergen, J., M. Brandt, and R. Koijen (2012). On the timing and pricing of dividends. *American Economic Review 102*(4), 1596–1618.

van Binsbergen, J., W. Hueskes, R. Koijen, and E. Vrugt (2013). Equity yields. *Journal of Financial Economics 110*(3), 503 – 519.

van Binsbergen, J. H. and R. S. Koijen (2017). The term structure of returns: Facts and theory. *Journal of Financial Economics 124*, 1–21.

Veronesi, P. (1999). Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *Review of Financial Studies 12*(5), 975–1007.

Veronesi, P. (2000). How does information quality affect stock returns? *The Journal of Finance 55*(2), 807–837.

Xiouros, C. and F. Zapatero (2010). The representative agent of an economy with external habit formation and heterogeneous risk aversion. *Review of Financial Studies 23*(8), 3017–3047.

Ziegler, A. C. (2003). *Incomplete Information and Heterogeneous Beliefs in Continuous-time Finance*. Springer-Verlag Berlin Heidelberg.