

NBER WORKING PAPER SERIES

ASSET PRICING WITH FADING MEMORY

Stefan Nagel  
Zhengyang Xu

Working Paper 26255  
<http://www.nber.org/papers/w26255>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 2019

We are grateful for comments to Kent Daniel, Stefano Giglio, Cosmin Ilut, Lars Lochstoer, Tarun Ramadorai, Pietro Veronesi, Rob Vishny, Jiang Wang, and Wei Xiong, seminar participants at Carnegie-Mellon, Colorado, Columbia, Dartmouth, European Central Bank, Imperial College, NYU, UCLA, University of Washington and conference participants at the Barcelona GSE Summer Forum, Bundesbank Household Finance Conference, CESifo Summer Institute, Cowles Conference on Expectations, Fordham-NYU-Imperial Rising Stars Conference, NBER Summer Institute, and the SED Meetings. We thank Robert Barro, Emmanuel Saez, Robert Shiller, and Stephen Wright for providing data on their websites and The Conference Board for providing data. Nagel gratefully acknowledges financial support from the Center for Research in Security Prices at the University of Chicago Booth School of Business. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Stefan Nagel and Zhengyang Xu. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Asset Pricing with Fading Memory  
Stefan Nagel and Zhengyang Xu  
NBER Working Paper No. 26255  
September 2019  
JEL No. E03,G12,G4

**ABSTRACT**

Building on evidence that lifetime experiences shape individuals' macroeconomic expectations, we study asset prices in an economy in which a representative agent learns with fading memory about unconditional mean endowment growth. With IID fundamentals, constant risk aversion, and memory decay calibrated to microdata, the model generates a high and strongly counter-cyclical objective equity premium, while the subjective equity premium is virtually constant. Consistent with this theory, experienced payout growth (a weighted average of past growth rates) is negatively related to future stock market excess returns and subjective expectations errors in surveys, and positively to analyst forecasts of long-run earnings growth.

Stefan Nagel  
University of Chicago  
Booth School of Business  
5807 South Woodlawn Avenue  
Chicago, IL 60637  
and NBER  
stefan.nagel@chicagobooth.edu

Zhengyang Xu  
Ross School of Business  
University of Michigan  
Ann Arbor, MI 48109  
zhengxu@umich.edu

## I. INTRODUCTION

The predictably counter-cyclical nature of the equity risk premium continues to be a major challenge in asset pricing. Researchers have proposed rational expectations models that generate time-variation in the equity premium by introducing modifications into the representative agent’s utility (Campbell and Cochrane 1999; Barberis, Huang, and Santos 2001) or by introducing persistence and stochastic volatility into the endowment growth process (Bansal and Yaron 2004). A key feature of these rational expectations models is that the representative agent knows the objective probability distribution she faces in equilibrium: subjective and objective expectations are the same. Therefore, the agent is fully aware of the counter-cyclical nature of the equity premium and knows the values of the parameters driving this process. This is a troubling feature of these models on two levels—conceptual and empirical. Conceptually, it is not clear how an agent could come to possess so much knowledge about parameters when econometricians struggle to estimate such parameters with much precision even from very long time-series samples (Hansen 2007). Empirically, surveys of investor return expectations from a number of sources fail to find evidence that investors’ return expectations are counter-cyclical. If anything, the survey data indicate pro-cyclicality (Vissing-Jorgensen 2003; Amromin and Sharpe 2013; Greenwood and Shleifer 2014).

Learning about the parameters of the asset payoff process offers a potential solution to both parts of this conundrum. Learning allows for subjective parameter uncertainty and a wedge between the subjective expectations of agents within the model and the objective expectations that an econometrician outside of the model could estimate from a large sample of data generated by the economy *ex post*. In such a learning model, the dynamics of subjective beliefs and asset prices depend crucially on the specification of agents’ memory. The modeling of memory determines how experienced historical data feeds into beliefs, decisions, and prices. A standard assumption would be that agents retain full memory of all observations experienced in the past. But full memory is, in an application, neither well defined—does it

start at the start of today’s electronic records of stock prices and dividends? when the first stock market opened? some other time?—nor necessarily empirically plausible.

In this paper, we study asset prices when the representative agent learns with gradually fading memory. Our approach is grounded in micro-evidence from household portfolio choice and survey expectations data showing that individuals learn from experience—that is, their expectations are shaped by data realized during their lifetimes, and most strongly by recently experienced data (Malmendier and Nagel 2011; Malmendier and Nagel 2016).<sup>1</sup> Malmendier and Nagel (2016) further show that when individuals learn from life-time experiences, the dynamics of their *average* expectation can be approximated very closely by a constant-gain learning scheme in which a data point’s influence on beliefs gradually fades over time as it recedes into the past. Abstracting from generational heterogeneity, we use this insight and consider a representative agent who learns with constant gain. This allows us to obtain a highly tractable model that nevertheless captures what may be, for the purposes of asset pricing, the essential aspect of learning from experience: the gradual loss of memory.

The decaying memory of observations in the past is the only modification to an otherwise standard Bayesian parameter learning model. As a consequence of the memory decay, learning is perpetual and there is a persistent time-varying wedge between the agent’s subjective beliefs and the objective beliefs implied by the true parameters of the process generating the asset payoffs. Importantly, however, by retaining everything else from the standard Bayesian set up, we are able to analyze the asset pricing effects of posterior subjective uncertainty under constant-gain learning. Based on this approach, the dynamics of asset prices and survey expectations can be reconciled within a simple setting with IID endowment growth, recursive utility with constant risk aversion, and a representative agent who learns with fading memory about the mean endowment growth rate.

We start the analysis by documenting several new facts about equity market returns

1. A growing body of evidence also suggests that extrapolation from experience helps understand the expectations and behavior of professionals. See, e.g., Greenwood and Nagel (2009) for mutual fund managers, Andonov and Rauh (2017) for pension fund return expectations, and Malmendier, Nagel, and Zhen (2018) for inflation expectations of members of the Federal Reserve’s Open Market Committee.

and subjective stock return expectations. We look at the data through the lens of a simple reduced-form framework that combines a present-value identity with constant-gain learning about the growth rate of dividends. Constant-gain learning implies that investors' subjective expectation of long-run dividend growth is equal to an exponentially weighted average of past dividend growth rates, which we label *experienced payout growth*. For the purpose of this initial analysis, we exogenously fix the subjective risk premium demanded by investors and the risk-free rate to be constant. Under objective expectations, then, the resulting equity premium is counter-cyclical. For example, after a string of positive growth innovations, investors are subjectively optimistic about the mean growth rate, the equity price is high, and subsequent returns are low because the investors' optimistic expectations are likely disappointed ex post.

How quickly memory of past realized growth rate observations decays is determined by a gain parameter in the belief-updating rule. This parameter plays a key role in our analysis as it determines the volatility and persistence of the price-dividend ratio and the strength of return predictability. We do not tweak this parameter to fit asset prices. Instead, we rely on the estimates in Malmendier and Nagel (2011, 2016) from survey data to pin down the value of this parameter at 0.018 for quarterly data. This means that the agent's posterior mean growth rate in the current quarter puts a weight of 0.018 on the most recent quarterly growth rate surprise and  $1 - 0.018$  on the posterior mean from the prior quarter. Experienced payout growth is therefore a slow-moving variable. We construct it empirically with dividend data going back to the 19th century.

We use the experienced payout growth series to uncover three novel empirical facts that are consistent with the predictions of this constant-gain learning model. First, experienced payout growth is strongly negatively related to subsequent stock returns in excess of the risk-free rate. Remarkably, unlike most existing equity return predictors in the literature, this one does not use information on price levels, just information from the past history of asset payouts. Second, using data on individuals' subjective expectations of stock market

excess returns, we find that they are basically unrelated to experienced payout growth. As a consequence, there is a wedge between subjective and objective expectations of excess returns that generates subjective expectations errors that are predictable by experienced payout growth. Third, subjective expectations of growth in fundamentals, proxied by long-term earnings forecasts of stock market analysts, are strongly positively related to experienced payout growth.

We then construct a structural asset-pricing model that matches these empirical regularities as well as other standard stylized asset pricing facts. A representative agent with recursive utility and constant risk aversion learns with fading memory about the mean of IID endowment growth. We keep the gain parameter fixed at the same value as in our reduced-form analysis. Equity is a levered claim to the endowment. Because of fading memory, learning is perpetual, the economy is stationary, and the agent's subjective uncertainty about long-run growth is high, which generates a high equity premium. The real risk-free rate varies slowly over time, but its volatility is low, consistent with empirical data.

The dynamics of risk premia in the model are in line with our empirical findings. Objectively, excess returns are predictable by experienced payout growth or the price-dividend ratio—without the subtle persistent components in the endowment process (and the agent's knowledge of these) in Bansal and Yaron (2004) (BY) or time-varying risk aversion built into the agent's preferences as in Campbell and Cochrane (1999) (CC). From the agent's subjective viewpoint, the world looks different. Subjective expected excess returns are virtually unrelated to experienced payout growth, consistent with the survey data. This wedge between subjective and objective expectations generates a strong negative relationship between subjective expectations errors and experienced payout growth. Thus, our model reconciles the evidence from returns data and surveys of investor return expectations.

Our model also addresses the Sharpe Ratio variability puzzle highlighted in Lettau and Ludvigson (2010). The leading rational expectations models by CC and BY imply that the conditional equity premium and conditional market return variance are almost perfectly

positively correlated. In CC, this happens because at times when risk aversion is high, it is also very volatile. In BY, the reason is that stochastic volatility in endowment growth is the driver of the time-varying equity premium. As a consequence, variation in conditional market return variance dampens the variability of the Sharpe Ratio relative to the variability of the conditional equity premium. Empirically, however, the first and second moments of market excess returns conditioned on the price-dividend ratio or experienced payout growth are basically uncorrelated. Our model is consistent with this empirical observation: subjective belief dynamics generate predictable variation in the objective equity premium, but without simultaneous variation in the volatility of equity returns, which generates a volatile Sharpe Ratio.

The model also matches the lack of out-of-sample predictability found empirically in Welch and Goyal (2008). In our model simulations, standard out-of-sample tests show no out-of-sample stock return predictability for empirically realistic sample sizes, even though returns are truly predictable under the objective distribution. While an econometrician can find predictable returns in sample by studying data ex post, it would be difficult, even with full memory, for the econometrician to construct a viable trading strategy in real time that takes advantage of the fading-memory agent.

To understand the economic mechanism generating investor belief and asset price dynamics in the model, it is useful to contrast it with a model in which constant-gain learning is an optimal Bayesian approach for tracking a random-walk component of growth rates. For appropriately chosen volatility of the random-walk increments, this alternative model would produce the same dynamics of investor beliefs about long-run growth. Thus, one can think of investors in our model as forming beliefs under the *perception* that endowment growth has a random-walk component. However, for the model's predictions about predictability of stock returns and subjective forecast errors it is crucial that the *actual* endowment growth rate is IID. The resulting wedge between subjective and objective expectations is key and makes the model predictions very different from models in which constant-gain learning is optimal

based on the actual law of motion.

This highlights an important difference relative to the large literature on constant-gain learning in macroeconomics (see Evans and Honkapohja (2001) for an overview). The typical motivation for constant-gain learning in this literature is structural change that renders long-distant historical data irrelevant, but it has been difficult to relate the gains that best explain survey forecasts of macroeconomic data to the gains that would be optimal given the degree of structural change in the forecasted time series (Branch and Evans 2006; Berardi and Galimberti 2017). In our case, fading memory reflects the fact that individuals rely on their life-time experiences when forming expectations, which is why we tie our gain parameter to earlier estimates from survey expectations microdata, not to the degree of structural change in the underlying time series. Our paper shares this focus on memory formation with a recent emerging literature on the topic in economics (e.g., Azeredo da Silveira and Woodford 2019; Bordalo, Gennaioli, and Shleifer 2019; and Wachter and Kahana 2019).

Our model builds on earlier work that has used different approaches to investigate the asset pricing implications of learning from experience. Collin-Dufresne, Johannes, and Lochstoer (2017), Ehling, Graniero, and Heyerdahl-Larsen (2018), Malmendier, Pouzo, and Vanasco (2017), Schraeder (2015), and Nakov and Nuño (2015) use an overlapping generations (OLG) approach to study learning-from-experience effects in asset pricing. The advantage of the OLG approach is that it maps very closely to the empirical work in Malmendier and Nagel (2011, 2016) that studied between-cohort heterogeneity in experiences, expectations, and choices. Moreover, these models produce interesting implications for cross-sectional heterogeneity in portfolio choices and wealth. The downside is that model solution requires simplifications that make a quantitative mapping to empirical data difficult.<sup>2</sup> By abstracting from cross-cohort heterogeneity, we also employ a simplified approach, but one that delivers

2. For example, Collin-Dufresne, Johannes, and Lochstoer (2017) use two overlapping cohorts and objective and subjective risk premia jump every 20 years when there is a generational shift; Ehling, Graniero, and Heyerdahl-Larsen (2018) assume log utility, Schraeder (2015) and Malmendier, Pouzo, and Vanasco (2017) work with CARA preferences in partial equilibrium with an exogenous risk-free rate, and the model in Nakov and Nuño (2015) has risk-neutral agents.



quantitatively realistic asset-pricing predictions. Our model is highly tractable and should therefore be well suited for further extensions such as, for example, to study production and real investment.

Our model shares many elements with full-memory Bayesian parameter learning models, especially the IID-normal model in Collin-Dufresne, Johannes, and Lochstoer (2016), but the fading memory feature avoids the arguably unrealistic implication of these models that learning effects disappear and risk premia decrease deterministically over time as the agent acquires more data (see, also, Timmermann (1993) and Lewellen and Shanken (2002) for partial equilibrium models with full memory). The fading memory model also avoids the problem of having to take a stand on what “year zero” is in an empirical implementation of a full-memory Bayesian learning model. And the gain parameter that determines memory can be pinned down based on microdata estimates.

Our model is also related to, but also in important ways different from recent models with extrapolative expectations. In Barberis, Greenwood, Jin, and Shleifer (2015) some investors extrapolate from stock price changes in recent years, which helps match the evidence in Greenwood and Shleifer (2014) that lagged stock market returns from the past few years are positively related to subjective expected returns. Jin and Sui (2018) build a representative agent model with return extrapolation. While these models can match the strong correlation of survey measures of subjective expected returns with lagged one-year stock market returns, they produce the counterfactual prediction that stock market excess returns are predictable by lagged one-year returns and that the price-dividend ratio quickly mean-reverts. The experienced growth measures in our setting put much greater weight on more distant observations in the past. As a consequence, the price-dividend ratio and objective expected returns in our model vary at a lower frequency, much close to the high persistence of the price-dividend ratio observed in the data. But our model cannot produce the correlation of one-year past returns with subjective expected returns. In Hirshleifer, Li, and Yu (2015) and Choi and Mertens (2013), extrapolation occurs at low frequency, like in our model. A key difference is

that in our model the agent perceives and prices subjective long-run growth rate uncertainty which allows us to generate a high equity premium in an IID economy. In Adam, Marcet, and Beutel (2017), agents know the expected growth rate of dividends, but they don't know the pricing function that maps expected fundamentals into prices, and they use an exponentially-weighted average of past price growth to forecast future prices. Matching the empirical equity risk premium in their model requires that subjective volatility of one-period ahead consumption growth far exceeds the actual volatility in the data. In our model, perceived short-run consumption volatility is very close to the objective volatility. The riskiness of equity in our model instead arises from subjective *long-run* growth rate uncertainty.

## II. FACTS ABOUT SUBJECTIVE AND OBJECTIVE EXPECTATIONS OF RETURNS AND PAYOFFS

Before looking at asset pricing with learning from experience within a structural asset-pricing framework, we start by laying out some empirical facts about stock market returns and investor return expectations from survey data that we want our asset-pricing model to match.

We consider a setting in which investors are learning about the mean growth rate  $\mu_d$  of log real stock market payouts,  $d$ ,

$$\Delta d_t = \mu_d + \epsilon_t, \tag{1}$$

where  $\epsilon$  is an IID shock. The microdata evidence in Malmendier and Nagel (2011, 2016) suggests that individuals form expectations from data they observe throughout their lifetimes and, within their life-time data set, with more weight on relatively recent data. In our analysis, we focus on the dynamics of the average individual's expectation in such a learning-from-experience setting. Malmendier and Nagel (2016) show that if individuals in different birth cohorts learn from their life-time experience, their average belief is described very well by a constant-gain learning rule. Applied in our setting here, this means that the perceived

growth rate  $\tilde{\mu}_d$  evolves as

$$\tilde{\mu}_{d,t+1} = \tilde{\mu}_{d,t} + \nu(\Delta d_{t+1} - \tilde{\mu}_{d,t}), \quad (2)$$

and where  $\nu$  is the (constant) gain parameter (see, e.g., Evans and Honkapohja (2001)).

As this expression shows,  $\tilde{\mu}_d$  is updated every period based on the observed surprise  $\Delta d_{t+1} - \tilde{\mu}_{d,t}$ . How much this surprise shifts the growth rate expectation depends on  $\nu$ . Malmendier and Nagel (2016) show that  $\nu = 0.018$  for quarterly data fits the dynamics of the average belief in inflation expectations microdata (and this value is also within the range of estimates obtained from microdata on household investment decisions in Malmendier and Nagel (2011)). Iterating on (2) one can see that  $\tilde{\mu}_{d,t}$  is an exponentially-weighted average of past  $\Delta d$  observations. In this way, the constant-gain updating scheme (2) captures the memory-loss implied by learning from experience and generational turnover. The more observations recede into the past, the lower the weight on these observations. In contrast, with full-memory Bayesian learning, the posterior mean would be formed by taking an equal-weighted average of all observed growth-rate realizations.

As a preliminary step, we explore some basic asset pricing implications when investors form expectations as in (2). At time  $t$ , they price stocks based on their growth rate expectation  $\tilde{\mu}_{d,t}$ . For now, we further assume that they price in a constant risk premium  $\theta$  and a constant real risk-free rate  $r_f$  under their subjective beliefs. As we will show later, these assumptions are very close to the subjective belief dynamics that we obtain for a representative agent in a fully specified asset-pricing model with constant-gain learning and priced subjective uncertainty about long-run growth.

Now apply a Campbell and Shiller (1988) approximate present-value identity, used as in Campbell (1991) to decompose return innovations into changes in expectations about future growth rates and changes in return expectations. Under the investors' subjective

expectations, denoted  $\tilde{\mathbb{E}}[\cdot]$ , the innovation in stock returns is

$$r_{t+1} - \tilde{\mathbb{E}}_t r_{t+1} = (\tilde{\mathbb{E}}_{t+1} - \tilde{\mathbb{E}}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} \quad (3)$$

$$= \frac{\rho}{1-\rho} (\tilde{\mu}_{d,t+1} - \tilde{\mu}_{d,t}) + \Delta d_{t+1} - \tilde{\mu}_{d,t} \quad (4)$$

$$= \left(1 + \frac{\rho\nu}{1-\rho}\right) (\Delta d_{t+1} - \tilde{\mu}_{d,t}). \quad (5)$$

Under the investors' subjective beliefs there is no term for the revision of return expectations because subjective return expectations stay fixed at  $\theta + r_f$ . Under these subjective beliefs, all variance of unexpected returns is due to revisions in forecasts of future cash flows. Adding investors' subjectively expected return we obtain total realized returns

$$r_{t+1} = \left(1 + \frac{\rho\nu}{1-\rho}\right) (\Delta d_{t+1} - \tilde{\mu}_{d,t}) + \theta + r_f. \quad (6)$$

Now consider an econometrician who knows (from a large sample of data) the true growth rate  $\mu_d$ . Taking expectations of (6) under these objective beliefs yields

$$\mathbb{E}_t r_{t+1} - r_f = \theta + \left(1 + \frac{\rho\nu}{1-\rho}\right) (\mu_d - \tilde{\mu}_{d,t}), \quad (7)$$

where the term in parentheses times  $\mu_d - \tilde{\mu}_{d,t}$  represents the subjective growth-rate expectations revision that the econometrician anticipates, on average, in the next period, given her knowledge of  $\mu_d$ . This expression shows that the econometrician should find returns to be predictable. Specifically,  $\tilde{\mu}_{d,t}$  should predict future excess returns negatively.

Moreover, while subjective excess return expectations are constant, the expectations error  $\mathbb{E}_t r_{t+1} - \tilde{\mathbb{E}}_t r_{t+1}$  should be predictable by  $\tilde{\mu}_{d,t}$ . We can see this by subtracting the subjective equity premium  $\tilde{\mathbb{E}}_t r_{t+1} - r_f = \theta$  from (7). We obtain

$$\mathbb{E}_t r_{t+1} - \tilde{\mathbb{E}}_t r_{t+1} = \left(1 + \frac{\rho\nu}{1-\rho}\right) (\mu_d - \tilde{\mu}_{d,t}). \quad (8)$$

In summary, this reduced-form analysis suggests three empirical relationships that we now investigate before moving on to explaining these within a structural asset-pricing framework: (i) excess returns should be predictable by  $\tilde{\mu}_{d,t}$ ; (ii) subjective expectations errors should be predictable by  $\tilde{\mu}_{d,t}$ ; and, most basically, (iii) subjective cash-flow growth expectations should be positively related to  $\tilde{\mu}_{d,t}$ .

### *II.A. Measurement of experienced growth*

To estimate the relationship between a slow-moving predictor like  $\tilde{\mu}_{d,t}$  and future returns in (7), we want to use the full history of returns back to the start of the CRSP database in 1926. And to implement the constant-gain learning scheme in (2), we need a long history of past observations on stock market fundamentals. For example, to compute  $\tilde{\mu}_{d,t}$  in 1926, we then need data on stock market payout growth,  $\Delta d$ , stretching back at least an additional 50 years, up to the point where the weights become close to negligible.

From 1926 onwards, we obtain quarterly observations of aggregate  $\Delta d$  on the CRSP value-weighted index. We use a payout series, constructed as in Bansal, Dittmar, and Lundblad (2005), that includes repurchases in addition to dividends. Shifts from dividends to stock repurchases (e.g., motivated by tax changes) in the last few decades of the sample could otherwise distort the link between our payout measure and the stock market fundamentals that we want to proxy for. In the early decades in the sample, the role of repurchases is negligible. Prior to 1926, we use data on household dividend receipts from tax data in Piketty, Saez, and Zucman (2018) for the period 1913 to 1926, and data on aggregate corporate non-farm non-financial dividends from Wright (2004) for the period 1900 to 1913. We deflate payout growth with CPI inflation rates and calculate per-capita real growth rates. For the period from 1871 to 1900 we use per-capita real GDP growth rates from Barro and Ursua (2008) as proxy for  $\Delta d$ . Appendix A.1 provides more details on the construction of the  $\Delta d$  times series. We then use this series to calculate experienced payout growth based on the recursion in (2) and we label it  $\tilde{\mu}_{d,t}$ .

The experienced growth measure based on corporate payouts is likely imperfect. While the inclusion of repurchases may help alleviate distortions in the time-series properties caused by shifts in payout policies, some distortions likely remain. Moreover, the pre-1926 payout data is of lower quality than the CRSP data. For these reasons, we also construct an alternative measure, experienced real returns,  $\tilde{\mu}_{r,t}$ , that represents a weighted average of past log real stock market index returns. From the point it becomes available in 1926, we use quarterly returns on the CRSP value-weighted stock market index. Before that, we use data from Shiller (2005) back to 1871 to construct quarterly returns on the S&P Composite index up to 1926. We deflate returns with the CPI inflation series from Shiller (2005). For averages taken over long periods, real payout growth and real returns should be highly correlated and hence  $\tilde{\mu}_{r,t}$  should capture similar information as  $\tilde{\mu}_{d,t}$ .<sup>3</sup> While experienced returns have some advantages as a measure of experienced fundamentals growth, there are also some potential shortcomings because asset price movements unrelated to fundamentals could contaminate  $\tilde{\mu}_r$ . For example, sentiment shocks that are orthogonal to fundamentals could cause movements in asset prices (that we accumulate in  $\tilde{\mu}_r$ ) and future objective expected returns (a dependent variable in our regressions below). For this reason, we use both measures,  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$ , in our tests.

## *II.B. Return Predictability*

Table I presents predictive regressions along the lines suggested by (7). In Panel A, we use  $\tilde{\mu}_d$  as a predictor, in Panel B we use  $\tilde{\mu}_r$ . Both predictors are constructed with data up to the end of quarter  $t$ . The dependent variable is the quarterly return on the CRSP value-weighted index in quarter  $t + 1$  in excess of the three-month T-bill yield at the end of quarter  $t$ . Based on eq. (7), the present-value model in (7) would predict an OLS slope coefficient of -2.78

3. To check this, we simulated dividend growth and returns from eqs. (1) and (5) in 1,000 samples of 360 quarters plus 400 quarters as a burn-in period to compute  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$  at the start of the estimation sample. How well  $\tilde{\mu}_{r,t}$  tracks  $\tilde{\mu}_{d,t}$  depends on the gain parameter  $\nu$ . For the value  $\nu = 0.018$  that we work with here, the correlation is a very high 0.82. Other parameters like  $\mu_d$ , the variance of  $\epsilon$ ,  $\theta$ ,  $r_f$ , or  $\rho$  do not influence this correlation. Thus, the approach of using  $\tilde{\mu}_r$  to capture the time-series dynamics of  $\tilde{\mu}_d$  should work well. We confirm this again below when we study  $\tilde{\mu}_r$  in data simulated from our asset-pricing model.

(with  $\nu = 0.018$  and  $\rho = 0.99$ , which is the quarterly value implied by the value of  $\rho = 0.964$  for annual data reported in Campbell (2000)).

As Panel A shows, our estimates in the full 1927-2016 period are roughly in line with this predicted value. To account for small-sample biases in predictive regressions, we run bootstrap simulations as in Kothari and Shanken (1997) to compute a bias adjustment and a bootstrap  $p$ -value. Appendix B provides details on these bootstrap simulations. With  $\tilde{\mu}_d$  as the only predictor, we get an OLS coefficient estimate of  $-5.79$ . Bias-adjustment shrinks the coefficient only slightly to  $-5.71$ , which is bigger in magnitude than the predicted value of  $-2.78$ . Based on the bootstrapped  $p$ -value of  $< 0.01$ , we can reject the no-predictability null at high levels of confidence.

One potential issue with these regressions is that the experienced real growth variables could be distorted by recent unexpected inflation. If companies are sluggish to adjust nominal payout growth one-for-one with inflation, a burst of recently high inflation would temporarily depress the real experienced payout growth that we measure with our simple exponentially-weighted average, but not necessarily the real fundamentals growth that investors truly experience. For this reason, column (2) therefore adds the average log CPI inflation rate during quarters  $t - 3$  to  $t$  to the regression. The coefficient for experienced real payout growth gets somewhat more negative, but not by much.

Column (3) adds the log price-dividend ratio to the regression. As a test of the economic story that we propose here, adding the price-dividend ratio, or other fundamentals-price ratios, to the regression is not really meaningful. The price-dividend ratio should—following the usual present-value identity logic—pick up essentially any variation in objective expected returns, and so it should also absorb predictability associated with  $\tilde{\mu}_d$ . However, as a purely descriptive empirical matter, it is useful to know whether experienced dividend growth adds any forecasting power over and above the log price-dividend ratio,  $p - d$ . Column (3) suggests that it does. In fact, in the presence of experienced payout growth and inflation in the regression,  $p - d$  is not a significant predictor and does not raise the  $R^2$  compared with

TABLE I  
PREDICTING RETURNS WITH EXPERIENCED REAL GROWTH

Dependent variable is the log return of the CRSP value-weighted index in quarter  $t + 1$  in excess of the return on a 3-month T-bill. In Panel A, experienced payout growth denotes a long-run exponentially weighted average of overlapping quarterly observations of four-quarter per-capita repurchase-adjusted real dividend growth rates leading up to and including quarter  $t$ , constructed with weights implied by constant gain learning with quarterly gain  $\nu = 0.018$ . In Panel B, experienced returns are constructed analogously as an exponentially weighted average of quarterly log stock market index returns (S&P Composite before 1926; then CRSP value-weighted index). Inflation is measured as the average log CPI inflation rate during the four quarters  $t - 3$  to  $t$ ;  $p - d$  refers to the log price-dividend ratio of the CRSP value-weighted index at the end of quarter  $t$ . The table shows slope coefficient estimates, with bootstrap bias-adjusted coefficient estimates in brackets. Intercepts are not shown. Bootstrap  $p$ -values are shown in parentheses. The reported  $R^2$  are based on bias-adjusted estimates.

	(1)	(2)	(3)	(4)	(5)
	1927-2016	1927-2016	1927-2016	1946-2016	1946-2016
<i>Panel A: Predicting returns with experienced real payout growth</i>					
Experienced real payout growth	-5.79	-6.25	-5.82	-2.99	-1.29
[bias-adj. coeff.]	[-5.71]	[-6.14]	[-5.40]	[-2.82]	[-1.14]
( $p$ -value)	(0.00)	(0.00)	(0.00)	(0.04)	(0.29)
Inflation		-0.73	-0.71	-1.60	-2.13
[bias-adj. coeff.]		[-0.75]	[-0.73]	[-1.68]	[-2.14]
( $p$ -value)		(0.10)	(0.13)	(0.01)	(0.00)
$p - d$			-0.01		-0.04
[bias-adj. coeff.]			[0.01]		[-0.02]
( $p$ -value)			(0.56)		(0.04)
Observations	360	360	360	284	284
$R^2$	0.033	0.037	0.034	0.027	0.044
<i>Panel B: Predicting returns with experienced returns</i>					
Experienced real returns	-2.36	-2.58	-2.22	-3.17	-2.38
[bias-adj. coeff.]	[-1.74]	[-1.79]	[-2.06]	[-2.21]	[-2.26]
( $p$ -value)	(0.02)	(0.02)	(0.05)	(0.01)	(0.03)
Inflation		-0.60	-0.57	-2.47	-2.78
[bias-adj. coeff.]		[-0.57]	[-0.50]	[-2.42]	[-2.76]
( $p$ -value)		(0.19)	(0.21)	(0.00)	(0.00)
$p - d$			-0.01		-0.03
[bias-adj. coeff.]			[0.00]		[-0.01]
( $p$ -value)			(0.44)		(0.04)
Observations	360	360	360	284	284
$R^2$	0.015	0.017	0.016	0.045	0.061



column (2).

Column (4) re-runs the regressions of column (2) for the post-World War II sample to address a potential concern that the results could be driven by the Great Depression period. The estimated slope coefficient of  $-2.99$  is now very close to the predicted value of  $-2.78$ . Adding  $p - d$  in column (5) takes away a substantial part of the predictability in this shorter sample. Of course, as we have noted above,  $\tilde{\mu}_d$  or  $p - d$  should capture the same information and the learning-from-experience theory does not imply that  $\tilde{\mu}_d$  should necessarily have incremental predictive power over and above  $p - d$ .

Panel B uses experienced real returns,  $\tilde{\mu}_r$ , as a proxy for experienced stock market fundamentals growth. For the full sample in column (1), we obtain a bias-adjusted point estimate of  $-1.74$  that is highly statistically significant. The remaining columns show that adding inflation and  $p - d$  and focusing on the post-1945 sample slightly strengthens the predictive relationship between experienced real stock returns and future excess stock returns.

Figure I shows that experienced real payout growth and future excess returns are also strongly correlated at much longer prediction horizons. In this figure, we plot the predicted 5-year excess log return based on the bias-adjusted fitted values from column (1) in Panel A of Table I, and iterating on it using the AR(1) dynamics of the experienced growth updating rule (2) with AR coefficient  $1 - \nu = 0.982$ . For comparison, we then plot the actual future 5-year cumulative excess log returns in quarters  $t + 1$  to  $t + 20$ . As the figure shows, there is a strong positive correlation. Time periods in which predicted returns were low also tend to be periods when subsequent five-year excess returns were poor. For example, low experienced payout growth correctly forecasted high excess returns following the last three recessions in the early 90s, early 2000s, and in the financial crisis. That the cycles in expected excess returns line up so well with cycles in experienced growth is remarkable because we did not pick the  $\nu$  parameter value—and hence the degree of smoothing implied by the weights—to match asset prices, but we fixed it at a value obtained from earlier microdata evidence. Moreover, unlike most return predictors in the literature,  $\tilde{\mu}_d$  does not include price level

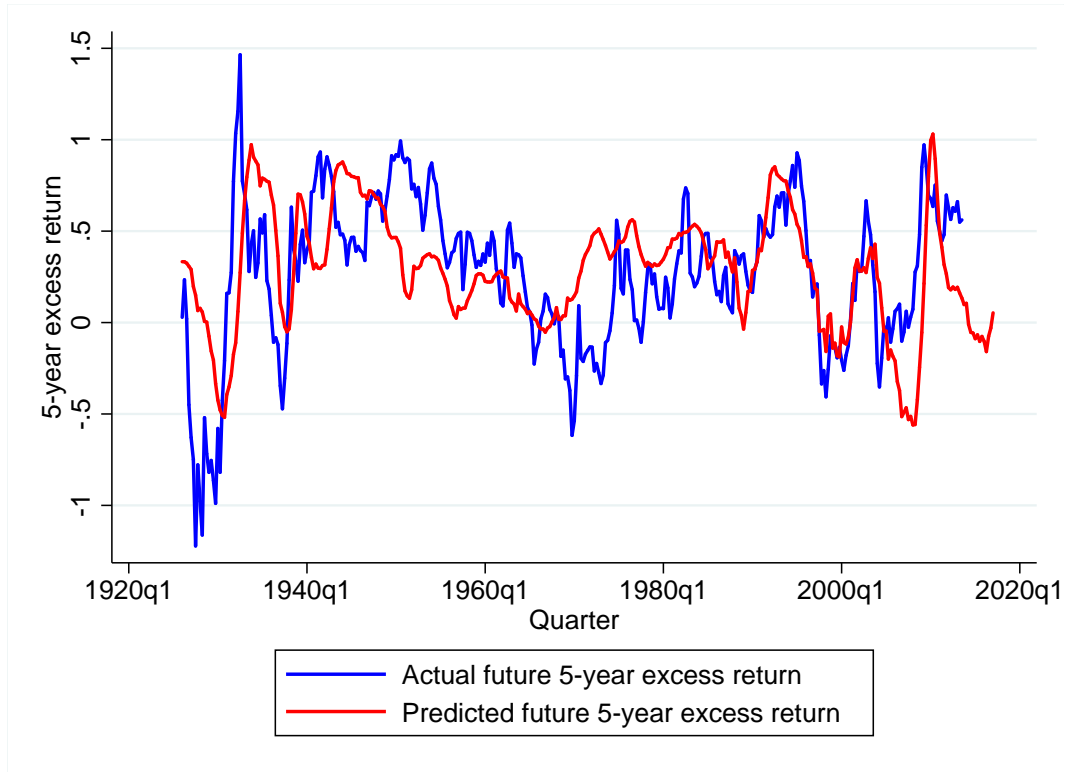


FIGURE I

Predicted five-year excess returns and subsequent actual cumulative five-year excess returns  
 Predicted returns are calculated based on bootstrap bias-adjusted coefficients from the predictive regression of log excess returns on experienced real payout growth shown in column (1) of Table I, Panel A, applied to experienced payout growth in quarter  $t$ , and iterating using an AR(1) with AR coefficient  $1 - \nu = 0.982$  to obtain 5-year forecasts. The actual cumulative five-year excess returns refers to the sum of excess log returns on the CRSP value-weighted index in quarters  $t + 1$  to  $t + 20$ .

information and hence it does not automatically pick up expected return shocks that affect prices.

Overall, the evidence indicates that experienced growth of stock market fundamentals—whether proxied for by experienced real payout growth or experienced real returns—is strongly predictive of stock market excess returns, consistent with a model in which investors use experienced growth to forecast future growth in fundamentals.

### *II.C. Subjective Expectation Error Predictability*

Subjective belief dynamics are a key feature of the economic effects we explore in this paper. For this reason, we want to confront our model with data on individual investor return expectations from surveys. We need a relatively long time series of survey expectations because the experienced growth variables that we focus on to explain dynamics change only slowly over time. For this reason, we put together survey data from several sources that spans the period 1972 to 1977 and 1987 to 2016. We focus on surveys that target a representative sample of the U.S. population, supplemented with two surveys of brokerage and investment firm customers.

Several of the surveys in our data elicit respondents' expected stock market returns, in percent, over a one-year horizon (UBS/Gallup survey, 1998-2007, monthly; Vanguard Research Initiative survey of Vanguard customers in Ameriks, Kézdi, Lee, and Shapiro (2016), one survey in 2014; surveys of Lease, Lewellen, and Schlarbaum (1974) and Lewellen, Lease, and Schlarbaum (1977), annual, 1972 and 1973). To extend these series, we bring in data from three additional surveys that don't elicit the percentage expected return but ask respondents to provide the probability of a rise in the stock market over a one-year horizon (Michigan Survey of Consumers, monthly 2002-2016) or the categorial opinion whether they expect stock prices to rise, or stay about where they are, or decline over the next year (Conference Board Survey, monthly 1987-2016;<sup>4</sup> Roper Center Surveys, annual, 1974-1977). We impute a time-series of implied percentage expected return from these alternative series. Roughly, the approach involves projecting the average expected returns each period from the first set of surveys on the coarser expectations measures in the second set of surveys, using periods of overlapping coverage to estimate the projection. Appendix A.2 provides more detail.<sup>5</sup> We

4. The data was kindly provided by The Conference Board.

5. Greenwood and Shleifer (2014) use two different data sources to cover time periods prior to the 1990s. From the mid-1980s onwards, they use the American Association of Individual Investors Investor (AAII) Sentiment Survey. The AAI survey is conducted among members of the AAI and it records responses of members that self-select into participation. Respondents state whether they are "bullish" or "bearish" about the stock market. We prefer the Conference Board survey for this time period as it is based on a representative sample of the U.S. population. For the early part of their sample starting in the 1960s, Greenwood and Shleifer

average the expectations within calendar quarters to obtain a quarterly series.

We start the analysis by looking at the time-series relationship between experienced growth and subjective expected excess returns. In the present-value model we have sketched above, the level of asset prices is affected by the experience-driven optimism or pessimism of investors. But the subjective expected excess return on the stock market,  $\tilde{\mathbb{E}}_t r_{t+1} - r_f$ , is constant. At each point in time, assets are priced such that subjective expected returns equal the (constant) equity premium required by investors. This will also be approximately true in our full model below, though not exactly.

As Panel A of Table II shows, subjective expected excess returns also seem to be approximately constant empirically at the frequencies that are relevant for our theory. In this table, we show the results of regressions of one-year excess return expectations in quarter  $t$  on experienced real payout growth or experienced real returns up to the end of quarter  $t - 1$ . We calculate subjective expected excess returns by subtracting the average one-year Treasury yields measured at the beginning of the survey months within each quarter. As Panel A shows there is only a weak, and statistically not significant, positive relationship between  $\tilde{\mu}_d$  and subjectively expected excess returns. Column (4) repeats this analysis with experienced real returns as the key explanatory variable. The results are very similar to those in column (1).

Looking at past returns over a much shorter time window, Greenwood and Shleifer (2014) find that survey return expectations are positively related to returns. As column (2) shows, we also find this in our data (which partly overlaps with Greenwood and Shleifer's) when we introduce the past 12-month return on the CRSP value-weighted index as an explanatory variable. The estimated coefficient on this lagged return is about three standard errors bigger than zero and the  $R^2$  is substantially higher than in column (1). Column (3) and (5) show that when experienced growth variables and lagged returns are used jointly, the

use the Investors' Intelligence newsletter sentiment. For consistency over time, we prefer to stick to individual investor surveys in all time periods. The Roper and Lewellen et al. surveys give us at least partial coverage of the 1970s.

experienced growth effect remains very weak. The important take-away is that in terms of the lower frequency movements that are captured by the experienced growth variables and that we focus on in our analysis, the subjective equity premium in the survey data is close to acyclical. That there are short-run fluctuations in subjective return expectations with one-year lagged returns is also interesting, but this is not a fact that we try to explain in this paper.

We now turn to the prediction, based on equation (7), that  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$  should predict expectation errors. We calculate the expectation error  $r_{t+1} - \tilde{\mathbb{E}}_t r_{t+1}$  on the left-hand side of (7) by subtracting the one-year survey expected return from the realized one-year return from the beginning of quarter  $t + 1$  to the end of quarter  $t + 4$ . The fact that survey expectations in Panel A are unrelated to experienced growth combined with the fact in Table I that future returns are negatively related to experienced growth implies that the expectations error should be negatively related to experienced growth. However, since the survey data is restricted to the 1970s and 1987-2016, the samples in Table I and II cover very different periods. For this reason, it is still useful to check whether there is actually a negative relationship in the part of the sample in which survey data is available.

Panel B of Table II shows that this is the case. There is a strong negative relationship between  $\tilde{\mu}_d$  or  $\tilde{\mu}_r$  and the expectations error. When experienced growth is high, return expectations are predictably too optimistic. Since the prediction horizon is one year rather than the one-quarter horizon in the return prediction regressions in Table I, the coefficient that we would expect, if these relations are stable across samples, is about four times the coefficient in Table I. The results in Panel B show that this is approximately true.

#### *II.D. Subjective Long-Run Growth Expectations*

In the last part of our empirical analysis, we look at proxies for subjective payout growth expectations. The surveys that provide subjective return expectations unfortunately do not elicit respondents' views about future cash-flow growth. For this reason, we follow Sharpe

TABLE II  
SURVEY RETURN EXPECTATIONS AND EXPERIENCED REAL GROWTH

In Panel A, the dependent variable is the average subjective expected stock return of survey respondents in quarter  $t$  minus the one-year treasury yield at the end of quarter  $t - 1$ , which we regress on experienced real payout growth or experienced real returns leading up to and including quarter  $t - 1$ . Lagged one-year return refers to the return of the CRSP value-weighted index over the four quarters  $t - 4$  to  $t - 1$ . In Panel B, the dependent variable is the expectation error, i.e., the realized return on the CRSP value-weighted index during quarters  $t + 1$  to  $t + 4$  minus the subjective expected return of survey respondents in quarter  $t$ . We use expectations data up to and including quarter 2016:4. Newey-West standard errors are reported in parentheses (12 lags in Panel A; 6 lags in Panel B).

	(1)	(2)	(3)	(4)	(5)
<i>Panel A: Subjective expected excess returns</i>					
Experienced real payout growth	0.31 (0.97)		0.37 (0.91)		
Experienced real returns				0.86 (0.63)	0.38 (0.76)
Lagged one-year return		0.03 (0.01)	0.03 (0.01)		0.03 (0.01)
Constant	0.05 (0.01)	0.05 (0.00)	0.05 (0.01)	0.04 (0.01)	0.05 (0.01)
Observations	125	125	125	125	125
Adj. $R^2$	-0.004	0.075	0.073	0.031	0.073
<i>Panel B: Expectation error: Realized - subj. expected</i>					
Experienced real payout growth	-12.34 (6.75)		-12.59 (6.80)		
Experienced real returns				-14.27 (6.48)	-15.47 (7.08)
Lagged one-year return		-0.10 (0.14)	-0.12 (0.15)		0.07 (0.13)
Constant	0.12 (0.06)	0.03 (0.03)	0.14 (0.06)	0.25 (0.09)	0.27 (0.10)
Observations	125	125	125	125	125
Adj. $R^2$	0.055	0.002	0.060	0.106	0.102

(2002) and use financial analysts’ long-term earnings growth forecasts aggregated at the market level as a proxy for investors’ subjective payout growth expectations.<sup>6</sup> We start with analysts’ monthly median forecasts of long-run earnings per-share (EPS) growth. According to I/B/E/S, they represent forecasts over a horizon of between three to five years.

We aggregate across stocks by forming a value-weighted average each month, using current-year forecasted total earnings for each stock as weights. We then average the monthly aggregated growth rate forecasts within each quarter, which yields a time series from quarter 1984:3 to 2016:4. We construct expected long-term real EPS growth, expressed in terms of per-quarter growth rates, by subtracting CPI inflation expectations from the Survey of Professional Forecasters issued in the quarter prior to the issuance of the analyst forecast.<sup>7</sup> Appendix A.3 provides more detail on the procedure to construct these measures.

According to the constant-gain updating scheme in (2), subjective long-term growth expectations should be positively related to experienced payout growth. This is what we find. Figure II provides a visual impression of the positive co-movement of the two series. Column (1) in Table III shows the results of regressions of analysts’ long-term expectations in quarter  $t$  on experienced payout growth constructed from realized growth rates up to quarter  $t - 1$ . We find that there is a positive relationship with a coefficient of 0.30 that is statistically significant at conventional levels (s.e. 0.15).

While the results in column (1) are qualitatively consistent with constant-gain learning from payout growth, the coefficient estimate is lower than predicted. According to the constant-gain updating scheme in (2), subjective long-term growth rate expectations should be equal to experienced payout growth and hence the coefficient should be unity. As we will show below, in our full asset-pricing model, where dividends have some tendency to mean-revert, we actually expect a coefficient slightly lower than unity (about 0.89 with  $\tilde{\mu}_d$  as

6. La Porta (1996) and Bordalo, Gennaioli, La Porta, and Shleifer (2017) also use analyst long-term earnings growth expectations to measure investor expectations of future stock fundamentals, but they focus on a cross-sectional analysis while here we analyze time-series relationships at the aggregate level.

7. We use the SPF forecast from the prior quarter to make sure analysts have access to this information. Stock analysts are presumably not experts in inflation forecasting and rely on other forecasters for macro forecast inputs.

TABLE III  
LONG-RUN GROWTH EXPECTATIONS AND EXPERIENCED REAL GROWTH

The dependent variable is an aggregate of I/B/E/S analysts long-term earnings-per-share (EPS) growth forecast, deflated using CPI inflation expectations from the Survey of Professional forecasters from the quarter prior to the EPS forecast date. Each month, we collect the median forecasts for each individual stock, form a value-weighted average across stocks using each stock's monthly median forecasts of current fiscal year earnings as the weight, with the sample restricted to stocks with positive current year forecasted earnings. We then average the resulting monthly time-series observations within calendar quarters 1984:3 to 2016:4. We regress the dependent variable measured in quarter  $t$ , and expressed as a per-quarter growth rate, on experienced real payout growth or experienced real returns leading up to and including quarter  $t-1$ . The one-year EPS forecast is the expected growth rate from current fiscal year to next fiscal year earnings, averaged across stocks in the same way as the long-term forecast. Lagged one-year return refers to the return of the CRSP value-weighted index over the four quarters  $t-4$  to  $t-1$ . Newey-West standard errors are reported in parentheses (24 lags).

	(1)	(2)	(3)	(4)	(5)	(6)
Experienced real payout growth	0.30 (0.15)	0.61 (0.18)	0.62 (0.20)			
Experienced real returns				0.55 (0.14)	0.52 (0.13)	0.64 (0.16)
One-year EPS growth forecast		0.28 (0.09)	0.29 (0.11)		0.05 (0.05)	0.08 (0.05)
Lagged one-year return			-0.00 (0.00)			-0.01 (0.00)
Constant	0.02 (0.00)	0.02 (0.00)		0.01 (0.00)	0.01 (0.00)	0.01 (0.00)
Observations	130	130	130	130	130	130
Adjusted $R^2$	0.077	0.308	0.312	0.355	0.362	0.452



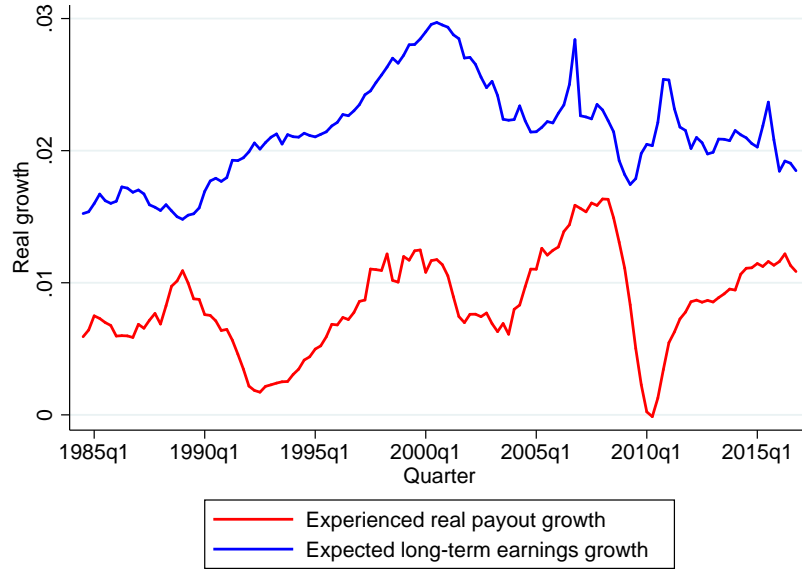


FIGURE II

Experienced real payout growth and analysts' long-run earnings growth expectations

The figure shows the value-weighted I/B/E/S long-term earnings-per-share growth forecast, value-weighted across stocks based on forecasted current year total earnings.

predictor and 0.68 with  $\tilde{\mu}_r$ ), but still substantially higher than the point estimate in column (1). There is one complication, however, in mapping the constant-gain learning scheme to the data. The setup in (1) and (2) is based on growth rates that are perceived as IID. In contrast, empirically, aggregate earnings changes have a sizable transitory component at short horizons of one to four quarters that is predictable by past earnings (Kothari, Lewellen, and Warner 2006) and by stock returns in recent quarters (Sadka and Sadka 2009). Analysts' aggregate short-term earnings forecasts vary over time to capture such short-horizon earnings growth predictability (Choi, Kalay, and Sadka 2016). If analysts' long-term forecasts,  $g_t^{LT}$ , are also influenced to some extent by such short-run predictability that is absent from our constant-gain learning framework, this could distort our estimates.<sup>8</sup> To be more precise,

8. Whether long-term forecasts are influenced by these short-term components is an open question. According to I/B/E/S, analysts should report long-term forecasts that represent expected growth "over the company's next full business cycle" (Thomson Reuters 2009), but it is not clear to what extent analysts are actually implementing this guidance.

suppose

$$g_t^{LT} = (1 - \beta)g_t^\infty + \beta g_t^{ST}, \quad (9)$$

where  $g_t^\infty$  is the unobserved long-run forecast undistorted by short-run expectations, which receives weight  $1 - \beta$ , and  $g_t^{ST}$  is the short-run forecast, which receives weight  $\beta$ . Our constant-gain learning framework suggests  $g_t^\infty = \tilde{\mu}_{d,t}$ , but if  $g_t^{ST}$  is correlated with  $\tilde{\mu}_{d,t}$ , a regression of  $g_t^{LT}$  on  $\tilde{\mu}_{d,t}$  will yield biased estimates.

For this reason, column (2) in Table III adds the aggregate one-year forecast as a control variable to the regression. The one-year forecast represents the expected real growth rate from current fiscal year to next fiscal year EPS, aggregated in the same way as we described for long-term growth expectations, and similarly expressed in terms of per-quarter growth rates. By controlling for the one-year forecast, we ask whether experienced payout growth explains long-term growth expectations holding fixed the potentially distorting short-term forecast components. In this specification, we obtain a much higher coefficient of 0.61 (s.e. 0.18) for the experienced payout growth variable and the one-year forecast obtains a coefficient of 0.28 (s.e. 0.09) which suggests that the long-term forecast puts a plausible weight of around one quarter on the short-term forecast. The implied coefficient in a regression of  $g_t^\infty$  on  $\tilde{\mu}_{d,t}$  is  $0.61/(1 - 0.28) = 0.85$ , which suggests that the results are very close to the predictions of the constant-gain learning model once one controls for the distorting effect of short-horizon predictability.

Column (3) adds the most recent one-year lagged return to the regression, which has little effect. Columns (4) to (6) repeat the analysis with experienced real returns in place of experienced payout growth and the results are very similar. In column (5), we obtain an implied coefficient in a regression of  $g_t^\infty$  on  $\tilde{\mu}_{r,t}$  of  $0.52/(1 - 0.05) = 0.64$ . Overall, the evidence on subjective expectations of growth in stock market fundamentals is broadly in line with the predictions of the constant-gain learning scheme.

### III. ASSET PRICING MODEL

We now develop these ideas more fully in a representative-agent endowment economy.

#### III.A. Learning with Fading Memory

Endowment growth follows an IID law of motion

$$\Delta c_{t+1} = \mu + \sigma \varepsilon_{t+1}, \quad (10)$$

where  $\{\varepsilon_t\}$  is a series of IID standard normal shocks. The agent knows that  $\Delta c_{t+1}$  is IID, and she also knows  $\sigma$ , but not  $\mu$ . The agent relies on the history of past endowment growth realizations,  $H_t \equiv \{\Delta c_t, \Delta c_{t-1}, \dots\}$ , to form an estimate of  $\mu$ .

We assume that the agent learns with constant gain and hence fading memory. Unlike in standard constant-gain learning models, however, we retain the modeling of the full posterior distribution—and hence the agent’s subjective uncertainty—of the Bayesian approach. To do so, we use a weighted likelihood that has been used in the theoretical biology literature to model memory decay in organisms (Mangel 1990). An agent who has seen an infinite history of observations on  $\Delta c$ , but with fading memory, forms a posterior

$$p(\mu|H_t) \propto p(\mu) \prod_{j=0}^{\infty} \left[ \exp\left(-\frac{(\Delta c_{t-j} - \mu)^2}{2\sigma^2}\right) \right]^{(1-\nu)^j}, \quad (11)$$

where  $1 - \nu$  is a positive number close to one ( $\nu = 0$  is the standard full memory Bayesian case where observations are equally weighted). Thus,  $(1 - \nu)^j$  represents a (geometric) weight on each observation. This weighting scheme assigns smaller weights the more the observation recedes into the past. With a prior

$$\mu \sim \mathcal{N}(\mu_0, \sigma_0^2) \quad (12)$$

held before seeing any data we then obtain the posterior

$$\mu|H_t \sim \mathcal{N}\left(\frac{\sigma_0^2}{\nu\sigma^2 + \sigma_0^2}\tilde{\mu}_t + \frac{\nu\sigma^2}{\nu\sigma^2 + \sigma_0^2}\mu_0, \left(\frac{1}{\sigma_0^2} + \frac{1}{\nu\sigma^2}\right)^{-1}\right), \quad (13)$$

where

$$\tilde{\mu}_t = \nu \sum_{j=0}^{\infty} (1-\nu)^j \Delta c_{t-j}. \quad (14)$$

The variance of the posterior is the same as if the agent had observed, and retained fully in memory with equal weights,  $S \equiv 1/\nu$  realized growth rate observations. In our case, the *actual* number of observed realizations is infinite, but the loss of memory implies that the *effective* sample size is finite and equal to  $S$ .

Due to the limited effective sample size, the prior beliefs retain influence on the posterior. For now, however, we work with an uninformative prior ( $\sigma_0 \rightarrow \infty$ ) and hence the posterior

$$\mu|H_t \sim \mathcal{N}(\tilde{\mu}_t, \nu\sigma^2). \quad (15)$$

We will return to the informative prior case when we consider versions of the model that generalize our baseline assumption about the elasticity of intertemporal substitution. With this uninformative prior, the posterior mean is updated recursively as

$$\tilde{\mu}_t = \tilde{\mu}_{t-1} + \nu(\Delta c_t - \tilde{\mu}_{t-1}). \quad (16)$$

Thus, the  $\tilde{\mu}_t$  resulting from this weighted-likelihood approach with an uninformative prior is identical to the perceived  $\mu$  that one obtains from the constant-gain updating scheme (2) with gain  $\nu$ . However, in contrast to standard constant-gain learning specifications in macroeconomics that focus purely on the first moment, we obtain a full posterior distribution. For the purpose of asset pricing, the subjective uncertainty implied by the posterior distribution can be crucial.

We further get the predictive distribution

$$\Delta c_{t+j}|H_t \sim \mathcal{N}(\tilde{\mu}_t, (1+\nu)\sigma^2), \quad j = 1, 2, \dots, \quad (17)$$

where the variance of the predictive distribution reflects not only the uncertainty due to future  $\epsilon_{t+j}$  shocks, but also the uncertainty about  $\mu$ . We denote expectations under the predictive distribution with  $\tilde{\mathbb{E}}_t[\cdot]$ . To understand better how the stochastic nature of the endowment process looks like from the agent's subjective viewpoint, we can rewrite (16) as

$$\tilde{\mu}_{t+1} = \tilde{\mu}_t + \nu\sigma\sqrt{1+\nu}\tilde{\epsilon}_{t+1}, \quad \text{where} \quad \tilde{\epsilon}_{t+1} = \frac{\Delta c_{t+1} - \tilde{\mu}_t}{\sigma\sqrt{1+\nu}}, \quad (18)$$

and  $\tilde{\epsilon}_{t+1}$  is  $\mathcal{N}(0, 1)$  distributed and hence unpredictable under the agent's time- $t$  predictive distribution.

Under Bayesian learning with full memory, the agent's information would be represented by a filtration and posterior beliefs would follow a martingale under this filtration. With fading memory, however, the posterior in periods  $t+j > t$  will be formed based on information that is different, but not more informative about  $\mu$  than the information available to the agent at time  $t$ . Hence, the information structure is not a filtration. For this reason, the time- $t$  agent anticipates that  $\tilde{\mu}_{t+j}$  in the future may vary from period to period, but she knows that this variation will be stationary and there is no convergence to  $\mu$  in the long run. Consistent with stationarity, the time- $t$  agent perceives future increments  $\tilde{\epsilon}_{t+j}$ ,  $j = 1, 2, \dots$  in (18) as negatively serially correlated (see Appendix C for more details) and not as martingale differences.<sup>9</sup> When we calculate asset prices in this model, we therefore cannot directly rely on certain laws like the law of iterated expectations (LIE) that require a filtration or results that presume martingale posteriors. This requires special care in evaluating valuation equations.

9. At time  $t$ , the agent however cannot make use of this serial correlation by using  $\tilde{\epsilon}_t$  to forecast  $\tilde{\epsilon}_{t+1}$ , because  $\tilde{\epsilon}_t$  is not observable to the agent. To figure it out, the agent would need full memory to compare  $\tilde{\mu}_t$  with  $\tilde{\mu}_{t-1}$ , but under constant-gain learning this is not possible. As a consequence,  $\tilde{\mathbb{E}}_t[\tilde{\mu}_{t+1}] = \tilde{\mu}_t$  remains the agent's posterior mean.

### III.B. Valuation

The valuation approach we use throughout the paper is a “resale” valuation approach. To illustrate, consider the valuation at date  $t$  of a claim to consumption at date  $t + 2$ . Under resale valuation, the agent at  $t$  prices the asset under the time- $t$  predictive distribution of the stochastically discounted  $t + 1$  asset value,

$$P_{R,t} = \tilde{\mathbb{E}}_t \left[ M_{t+1|t} \tilde{\mathbb{E}}_{t+1} [M_{t+2|t+1} C_{t+2}] \right], \quad (19)$$

where we use  $M_{t+j|t}$  to denote the one-period stochastic discount factor (SDF) from  $t + j - 1$  to  $t + j$  that applies given the agent’s predictive distribution at  $t$ .

An alternative way of valuing this claim would be a “buy-and-hold” valuation, where the agent values the asset based on the stochastically discounted payoff under the time- $t$  predictive distribution

$$P_{H,t} = \tilde{\mathbb{E}}_t [M_{t+1|t} M_{t+2|t} C_{t+2}], \quad (20)$$

In a full-memory Bayesian setting, the LIE would apply in the valuation equation of  $P_{R,t}$  with the result that  $P_{R,t} = P_{H,t}$ , but with fading memory the information structure is not a filtration and the LIE typically fails to hold.<sup>10</sup> As a consequence  $P_{R,t} \neq P_{H,t}$ .

The valuation discrepancy between the two valuation approaches arises because the agent at  $t$  and at  $t + 1$  sees the statistical properties of the shock  $\tilde{\varepsilon}_{t+2}$  differently. The buy-and-hold valuation incorporates the negative serial correlation of  $\tilde{\varepsilon}_{t+1}$  and  $\tilde{\varepsilon}_{t+2}$ . In contrast, the resale valuation at  $t$  is based on the anticipation that the value of the asset at date  $t + 1$  will be determined by an agent—or a future self of the agent—who perceives  $\tilde{\varepsilon}_{t+2}$  as unpredictable. Thus, the resale valuation is based on a chain of valuations that each views the one-period ahead  $\tilde{\varepsilon}$  shock as unpredictable.

We work with the resale valuation approach below, for two reasons. First, the resale

10. For subjective expectations of linear functions of  $\Delta c$ , we still get an LIE, e.g.,  $\tilde{\mathbb{E}}_t \tilde{\mathbb{E}}_{t+1} \Delta c_{t+2} = \tilde{\mathbb{E}}_t \Delta c_{t+2}$ , but the LIE does not hold for nonlinear functions of  $\Delta c$ , e.g.,  $\tilde{\mathbb{E}}_t \tilde{\mathbb{E}}_{t+1} [\exp(a + b \Delta c_{t+2})] \neq \tilde{\mathbb{E}}_t [\exp(a + b \Delta c_{t+2})]$ .

valuation is time-consistent. In contrast, if the asset was priced at time  $t$  at the buy-and-hold valuation and the anticipation of a predictable  $\tilde{\varepsilon}_{t+2}$ , and time moves on to  $t + 1$ , the agent would, after memory loss, suddenly find  $\tilde{\varepsilon}_{t+2}$  unpredictable. Thus, the agent would then agree with a valuation based on an unpredictable  $\tilde{\varepsilon}_{t+2}$ , but this is not consistent with the buy-and-hold valuation at  $t$ . Second, the resale valuation also fits with the underlying motivation of our model as an approximation for experience-based learning in an overlapping generations model in which actual resale would occur when generations turn over.

### III.C. Kalman Filtering Interpretation

The updating scheme in (16) is reminiscent of optimal filtering in the case of a latent stochastic trend. Indeed, if the agent *perceived*  $\mu$  to follow a random walk—i.e., as  $\mu_t = \mu_{t-1} + \zeta_t$ , where  $\zeta$  is an IID-normal shock—rather than a constant as in (10), application of the steady-state Kalman filter yields exactly the same posterior distribution as in (15). With appropriate choice of the volatility of the  $\zeta$  shocks, and as long as the *actual* law of motion is still (10) with constant  $\mu$ , the dynamics of the posterior beliefs would be the same as in our fading memory model (see Appendix D for more details). The predictive distribution of one-period ahead endowment growth, and, as a consequence, asset prices under resale valuation would be the same as well.<sup>11</sup> This is useful for technical purposes because it allows us to map our framework into one in which the information structure is a filtration and it is Markovian. Through this mapping, we can use results from Hansen and Scheinkman (2012) to determine parameter restrictions that are sufficient for existence of equilibrium. While this reinterpretation is convenient for technical reasons, optimal Kalman filtering does not explain the micro-evidence on learning from experience that motivates our fading memory approach. For this reason, we stick to the fading memory interpretation in the discussion of our model.

11. However, for the asset-pricing predictions to remain the same, it is crucial that (10), with constant  $\mu$ , remains the *actual* law of motion. As we will show below, the time-varying wedge  $\tilde{\mu}_t - \mu$  between subjective and objective beliefs plays an important role in generating volatile asset prices and predictable excess returns. Without this wedge, there would be no return predictability.

### III.D. Stochastic Discount Factor

We assume that the representative agent evaluates payoffs under recursive utility as in Epstein and Zin (1989), with value function

$$V_t = \left[ (1 - \delta)C_t^{1-\frac{1}{\psi}} + \delta \tilde{\mathbb{E}}_t[V_{t+1}^{1-\gamma}]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}}, \quad (21)$$

where  $\delta$  denotes the time discount factor,  $\gamma$  relative risk aversion for static gambles and  $\psi$  the elasticity of intertemporal substitution (EIS). Note that the agent evaluates the continuation value under her subjective expectations  $\tilde{\mathbb{E}}_t[\cdot]$ . We apply the same resale valuation approach that we use for asset pricing to this continuation value as well.

In our baseline model, we set  $\psi = 1$ . Iterating on the value function as in Hansen, Heaton, and Li (2008), but here under the agent's predictive distribution, we then obtain the log SDF that prices assets under the agent's subjective beliefs,

$$m_{t+1|t} = \tilde{\mu}_m - \tilde{\mu}_t - \xi \sigma \tilde{\varepsilon}_{t+1}, \quad (22)$$

with

$$\tilde{\mu}_m = \log \delta - \frac{1}{2}(1 - \gamma)^2(\nu U_v + 1)^2(1 + \nu)\sigma^2, \quad (23)$$

$$\xi = [1 - (1 - \gamma)(\nu U_v + 1)]\sqrt{1 + \nu}, \quad (24)$$

$$U_v = \frac{\delta}{1 - \delta}. \quad (25)$$

Details are in Appendix E.1. This SDF implies the risk-free rate

$$r_{f,t} = -\tilde{\mu}_m + \tilde{\mu}_t - \frac{1}{2}\xi^2\sigma^2. \quad (26)$$



### III.E. Pricing the Consumption Claim

We can now solve for the consumption-wealth ratio,  $\zeta \equiv W_t/C_t$ , and the subjective risk premium for the consumption claim. Using (18), we can express the log of the return on the consumption claim as

$$r_{w,t+1} = \tilde{\mu}_t + \sqrt{1 + \nu\sigma}\tilde{\varepsilon}_{t+1} + \log\left(\frac{\zeta}{\zeta - 1}\right). \quad (27)$$

Applying the subjective pricing equation  $\tilde{\mathbb{E}}_t[M_{t+1|t}R_{W,t+1}] = 1$ , we can solve for the wealth-consumption ratio

$$\zeta = \frac{1}{1 - \delta}. \quad (28)$$

Thus, as in the rational expectations case,  $\psi = 1$  implies a constant and finite consumption-wealth ratio. In the posterior distribution in (15), extremely large values of  $\mu$  have greater than zero probability. The agent therefore also assigns some probability mass to extremely large future  $\tilde{\mu}_{t+j}$ . However, since  $\psi = 1$  implies that  $r_{f,t+j}$  moves one-for-one with  $\tilde{\mu}_{t+j}$ , the effect of high subjectively expected growth rates on the value of the consumption claim is exactly offset by a high future risk-free rate. As a consequence, the wealth-consumption ratio is constant and finite.

Evaluating the subjective pricing equation for  $R_{W,t+1}$  again, now using the fact that  $\zeta$  is constant, we can solve for the subjective risk premium of the consumption claim

$$\log \tilde{\mathbb{E}}_t[R_{w,t+1}] - r_{f,t} = \xi\sqrt{1 + \nu\sigma^2}, \quad (29)$$

which is constant over time. In contrast, the objective risk premium under the econometrician's measure, generated by data sampled from this economy, is time-varying: taking the objective and subjective expectations and variance of (27), we can calculate the wedge between subjective and objective expectations, and combining with (29), we obtain the objective

risk premium

$$\log \mathbb{E}_t[R_{w,t+1}] - r_{f,t} = \xi \sqrt{1 + \nu \sigma^2} - \frac{1}{2} \nu \sigma^2 + \mu - \tilde{\mu}_t, \quad (30)$$

where the time-varying wedge  $\mu - \tilde{\mu}_t$  reflects the disagreement between the econometrician and the agent about the conditional expectation of  $r_{w,t+1}$ . The wedge is observable to the econometrician who knows  $\mu$ , but since  $\tilde{\mathbb{E}}_t[\mu] = \tilde{\mu}_t$  the wedge is zero from the viewpoint of the agent at time  $t$ .

### III.F. Pricing the Dividend Claim

We now turn to pricing the dividend claim, which is the main focus of our analysis. Dividends in our model are a levered claim to the endowment. We assume that dividends and endowment are cointegrated. Specifically, we assume

$$\Delta d_{t+1} = \lambda \Delta c_{t+1} - \alpha (d_t - c_t - \mu_{dc}) + \sigma_d \eta_{t+1}, \quad \alpha > 0, \quad (31)$$

similar to Bansal, Gallant, and Tauchen (2007). We assume that  $\mu_{dc}$ ,  $\lambda$ , and  $\alpha$  are known to the agent. The agent's learning problem is focused on the unknown  $\mu$ .

Cointegration is economically realistic, and it is of particular importance in a model like ours with subjective growth rate uncertainty. Since the price of a dividend claim is convex in dividend growth rates, the subjective growth rate uncertainty in this model could cause the price to be infinite. For the consumption claim this issue was resolved by setting  $\psi = 1$ . However, leverage magnifies the convexity effect and without sufficiently strong cointegration, the price of the equity claim explodes even if the consumption claim has a finite price. In our quantitative implementation, we will assume that  $\alpha$  is very small and so dividends and consumption can drift away from each other quite far, but we keep  $\alpha$  sufficiently big to yield a finite price-dividend ratio with empirically reasonable moments. Appendix E.3 provides more detail.

The price of the  $n$ -period dividend strip is

$$P_t^n \equiv \tilde{\mathbb{E}}_t[M_{t+1}|_t \tilde{\mathbb{E}}_{t+1}[\cdots \tilde{\mathbb{E}}_{t+n-1}[M_{t+n}|_{t+n-1} D_{t+n}]]]. \quad (32)$$

As we discussed earlier, when we evaluate these expectations, we do so by iterating backwards from the payoff at  $t+n$ , evaluating one conditional expectation at a time without relying on the LIE. Taking logs and evaluating (32), we obtain

$$p_t^n - d_t = [1 - (1 - \alpha)^n] (c_t - d_t + \mu_{dc} + \frac{\lambda - 1}{\alpha} \tilde{\mu}_t) + n \tilde{\mu}_m + \frac{1}{2} (A_n \sigma^2 + B_n \sigma_d^2), \quad (33)$$

where, for very large  $n$ , approximately

$$A_n \approx n \left[ \sqrt{1 + \nu} \left( 1 + \nu \frac{\lambda - 1}{\alpha} \right) - \xi \right]^2, \quad (34)$$

and  $B_n$ , which does not grow with  $n$ , becomes very small relative to  $A_n$  (see Appendix E.3).

The analytical solution for dividend strip prices is useful for understanding the behavior of subjective and objective risk premia in this model. Consider the one-period return on the “infinite-horizon” dividend strip

$$R_{t+1}^\infty \equiv \lim_{n \rightarrow \infty} P_{t+1}^{n-1} / P_t^n. \quad (35)$$

As we show in Appendix E.3, we can use equation (33) to find the one-period subjective risk premium for this claim

$$\log \tilde{\mathbb{E}}_t[R_{t+1}^\infty] - r_{f,t} = \left[ 1 + \nu \frac{\lambda - 1}{\alpha} \right] \xi \sqrt{1 + \nu} \sigma^2. \quad (36)$$

For  $\gamma \geq 1$ ,  $\xi$  is a positive constant. We observe from the above that lowering  $\alpha$  raises the subjective risk premium because it enhances the persistence of the leverage effect by weakening the forces of cointegration. The subjective risk premium is positively related to  $\nu$

because higher  $\nu$  implies a smaller effective sample size used to estimate  $\mu$  and hence higher subjective uncertainty about  $\mu$ .

While the subjective risk premium is constant, the objective risk premium is

$$\log \mathbb{E}_t[R_{t+1}^\infty] - r_{f,t} = \left[1 + \nu \frac{\lambda - 1}{\alpha}\right] \xi \sqrt{1 + \nu \sigma^2} - \frac{1}{2} \nu \left(1 + \nu \frac{\lambda - 1}{\alpha}\right)^2 \sigma^2 + \left(1 + \nu \frac{\lambda - 1}{\alpha}\right) (\mu - \tilde{\mu}_t), \quad (37)$$

and hence time-varying with the wedge  $\tilde{\mu}_t - \mu$ : the more optimistic the agent relative to the econometrician, the lower the objective expected excess return. Thus, learning induces return predictability. And unlike Bayesian learning with full memory as in Collin-Dufresne, Johannes, and Lochstoer (2016) where return predictability dies out in the long-run, learning with constant gain means that the learning effects (and hence return predictability) are perpetual.

As equation (37) shows, leverage  $\lambda > 1$  magnifies the time-variation in the objective risk premium. With  $\lambda = 1$ , or for a consumption strip, all variation would come purely from excess variation in the risk-free rate: objective expected returns on the consumption strip are constant because the ratio of its price to consumption is constant and the objective expected growth rate of  $\Delta c$  is constant, while the risk-free rate rises with  $\tilde{\mu}_t$ . With leverage, however, the price of a dividend claim rises with  $\tilde{\mu}_t$ , which produces additional variation in the objective risk premium.

Like the long-horizon claim, shorter-horizon claims also have a constant subjective risk premium. For example, a one-period claim with return  $R_{t+1}^1 \equiv D_{t+1}/P_t^1$ , has the constant subjective risk premium

$$\log \tilde{\mathbb{E}}_t[R_{t+1}^1] - r_{f,t} = \lambda \xi \sqrt{1 + \nu \sigma^2}. \quad (38)$$

The ex-dividend price of the equity claim, i.e., the claim to the entire stream of dividends, is simply the sum of prices of dividend strips

$$P_t = \sum_{n=1}^{\infty} P_t^n. \quad (39)$$

We compute the sum in (39) numerically using the analytical solution for dividend strips. Details are in Appendix E.4.

Our result that the subjective risk premium is constant for short and long-maturity dividend strips does not imply that the subjective risk premium for the whole stream of dividends is constant. We solve for the subjective risk premium of the equity claim numerically using methods from Pohl, Schmedders, and Wilms (2018) (see Appendix E.4). As we report below, we find a slightly positive relationship between  $\tilde{\mu}_t$  and the subjective equity risk premium. This arises from the fact that the contribution of long-horizon claims to the overall value of the portfolio of strips gets bigger when  $\tilde{\mu}_t$  is higher: due to exponentiating, the effect of a rise of  $\tilde{\mu}_t$  on long-horizon equity is bigger than on short-horizon equity. As a consequence, the claim on the whole stream behaves more like long-horizon equity when  $\tilde{\mu}_t$  is high and is subjectively priced more like long-horizon equity, i.e., with a higher risk premium if  $\alpha < \nu$ .

### *III.G. Solving the Model with $\psi > 1$ and an Informative Prior*

When  $\psi \neq 1$ , and the prior is diffuse, the consumption-wealth ratio is no longer finite. For example, if  $\psi > 1$ , the effect of high subjectively expected growth rates on the value of the consumption claim is no longer fully offset by a high future risk-free rate, which causes the consumption-wealth to explode. Earlier work has resolved this through truncation of the state space (Collin-Dufresne, Johannes, and Lochstoer 2017) or limiting the time horizon over which growth is uncertain (Pástor and Veronesi 2003; Pástor and Veronesi 2006). We take a different approach by endowing the agent with an informative prior. In our fading memory model, this approach is effective in preventing the explosion of the consumption-wealth ratio because future agents never gain more precise information about  $\mu$  than the current agent has. As a consequence, the weight on the prior does not decay and the current agent anticipates that the posterior means of the agents pricing the asset at times in the future will always have a similarly strong tilt towards the prior mean.

Our approach is similar in spirit to the state-space truncation approach—both methods

effectively pull the perceived distribution of future posterior means towards economically plausible growth rates—but it is far more tractable in our setting. We center the prior distribution at the true mean  $\mu$ , but this is not essential. Since we work with high prior variance, the prior will remain almost uninformative. Therefore, setting the prior mean to any value in an economically plausible neighborhood around  $\mu$  would deliver similar results.

Given the prior in (12) with  $\mu_0 = \mu$ , we obtain the posterior

$$\mu|H_t \sim \mathcal{N}\left(\phi\tilde{\mu}_t + (1-\phi)\mu, \left(\frac{1}{\sigma_0^2} + \frac{1}{\nu\sigma^2}\right)^{-1}\right), \quad \text{where } \phi \equiv \frac{\sigma_0^2}{\nu\sigma^2 + \sigma_0^2}. \quad (40)$$

The perceived consumption growth can be represented as

$$\Delta c_{t+1} = \phi\tilde{\mu}_t + (1-\phi)\mu + \sqrt{1 + \phi\nu}\sigma\tilde{\varepsilon}_{t+1}, \quad (41)$$

where  $\tilde{\varepsilon}_{t+1}$  is  $\mathcal{N}(0,1)$  distributed under the agent’s time- $t$  predictive distribution. With an informative prior we have  $\phi < 1$  and so the volatility of the subjectively unexpected endowment growth is lower than in the diffuse prior case where  $\phi = 1$ . We solve this version of the model with log-linearization. Details, including parameter restrictions sufficient for existence of equilibrium, are provided in Appendix F.

## IV. CALIBRATION AND EVALUATION

Table IV summarizes the parameter values we use in our baseline quantitative analysis. We fix the gain parameter  $\nu$  at the value that Malmendier and Nagel (2016) estimated from survey data on inflation expectations. For the endowment process and preferences, we set most parameters to the same values as in Bansal, Kiku, and Yaron (2012) and Collin-Dufresne, Johannes, and Lochstoer (2017).

We set  $\sigma_d$  to a relatively low value of 1% quarterly. At this value, dividend volatility in the model will be smaller than in the data. However, it will allow us to roughly match the

volatility of  $\tilde{\mu}_d$  in the data, which is more important for our purposes. We cannot match both at the same time because in our model  $\Delta d$  is IID whereas in the data  $\Delta d$  has substantial negative autocorrelation, which implies that a lot of this dividend growth volatility cancels out when we form the long-run weighted average  $\tilde{\mu}_d$ . For this reason, it makes more sense to calibrate our IID dividend process to the volatility of  $\tilde{\mu}_d$ , which reflects permanent shocks, rather than the volatility of  $\Delta d$  in the data, which is influenced by a substantial transitory component.

In our baseline specification, we set  $\psi = 1$ , but we also report some results for the  $\psi = 1.5$  case. We choose the remaining parameters  $\gamma$  and  $\alpha$  to get a realistic equity premium and equity volatility. We work with a relatively low risk aversion of  $\gamma = 4$ . The value  $\alpha = 0.001$  satisfies the condition required for a finite price of the dividend claim in Appendix E.3 and it implies dividends can wander quite far away from consumption (but not as far as in models without cointegration such as Bansal, Kiku, and Yaron (2012)).

#### *IV.A. Unconditional moments*

We simulate the model at a quarterly frequency. Table V reports the annualized population moments estimated from an extremely long sample simulated from the model. We also show empirical moments for the 1927 to 2016 period for comparison.

As we anticipated, the volatility of  $\Delta d$  in the model is lower than in the data, but the volatility of  $\tilde{\mu}_d$  is close to the data, and actually even a bit higher. This reinforces our earlier point that much of the volatility of  $\Delta d$  in the data is due to transitory components.

In terms of unconditional asset pricing moments, the model produces a high equity premium (7.16%) and Sharpe Ratio (0.54) that are quite close to the empirical estimates in the first column. Return volatility and the volatility of the log price-dividend ratio are lower than in the data. The version of the model with  $\psi = 1.5$  in the third column gets somewhat closer to the empirical values.

TABLE IV  
 BASELINE MODEL PARAMETERS

This table reports the parameters values we use in the baseline calibration of our model at a quarterly frequency. The gain parameter  $\nu$  is fixed at the value that Malmendier and Nagel (2016) estimated from survey data on inflation expectations. For endowment process parameters and preferences, we set most at the same values as in Bansal, Kiku, and Yaron (2012) and Collin-Dufresne, Johannes, and Lochstoer (2017).

Parameter	Symbol	Value	Source
<i>Belief updating</i>			
Gain	$\nu$	0.018	MN (2016) (survey data)
<i>Endowment process</i>			
Leverage ratio	$\lambda$	3	CJL (2017)
Dividend cointegration parameter	$\alpha$	0.001	
Mean consumption growth	$\mu$	0.45%	CJL (2017)
Consumption growth volatility	$\sigma$	1.35%	CJL (2017)
Dividend growth volatility	$\sigma_d$	1%	
<i>Preferences</i>			
Risk aversion	$\gamma$	4	
EIS	$\psi$	1	
Time discount factor	$\delta$	0.9967	BKY (2012)



TABLE V  
UNCONDITIONAL MOMENTS

The second column in this table presents the model population moments obtained as average across 1,000 simulations of the model for 50,000 periods plus a 2,000-period burn-in period to compute  $\tilde{\mu}$ ,  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$  at the start of each sample, using a diffuse prior ( $\phi = 1$ ) and  $\psi = 1$ . The third column shows results for the model with  $\psi = 1.5$  and an informative prior with  $\phi = 0.99$ . The first column shows the corresponding empirical moments from the data for the 1927 to 2016 period. For the empirical versions of  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$ , we use data from 1871 to 1926 as pre-sample information to calculate their values at the start of the sample in 1927. Consumption growth is calculated from quarterly real per-capita consumption expenditure on nondurables and services in chained 2012 dollars from NIPA for 1947-2016, annual NIPA data on nondurables and services expenditure from 1929 to 1947, and annual real per-capita consumption expenditure from Barro and Ursua (2008) for 1926 to 1929. In both columns, returns are annualized as follows: The means of risky returns are multiplied by four and standard deviation multiplied by two. For the risk-free rate,  $\tilde{\mu}$ , and  $\tilde{\mu}_d$  we multiply quarterly means and standard deviations by four. We estimate the empirical moments of  $\Delta c$  from four-quarter changes of quarterly log nondurables and services consumption and those of  $\Delta d$  from four-quarter changes in the log of repurchase-adjusted dividends on the CRSP value-weighted index. The simulated statistics for  $p - d$  use a four-quarter trailing sum of dividends in the calculation of  $p - d$ , just like in the empirical version of  $p - d$ .

	Data 1927-2016	Model $\psi = 1, \phi = 1$	Model $\psi = 1.5, \phi = 0.99$
$\mathbb{E}(\Delta c)$	1.84	1.80	1.80
$\sigma(\Delta c)$	2.72	2.70	2.70
$\mathbb{E}(\Delta d)$	2.38	1.80	1.80
$\sigma(\Delta d)$	13.31	8.35	8.35
$\sigma(\tilde{\mu})$	-	0.51	0.51
$\rho(\tilde{\mu})$	-	0.98	0.98
$\sigma(\tilde{\mu}_d)$	1.32	1.55	1.55
$\rho(\tilde{\mu}_d)$	0.97	0.98	0.98
$\text{corr}(\tilde{\mu}, \tilde{\mu}_d)$	-	0.96	0.96
$\text{corr}(\tilde{\mu}, \tilde{\mu}_r)$	-	0.85	0.77
$\mathbb{E}(R_m - R_f)$	8.11	7.16	7.75
$\sigma(R_m - R_f)$	22.41	13.31	16.35
$SR(R_m - R_f)$	0.36	0.54	0.47
$\mathbb{E}(p - d)$	3.40	2.81	2.98
$\sigma(p - d)$	0.44	0.14	0.22
$\rho(p - d)$	0.97	0.91	0.94
$\mathbb{E}(r_f)$	0.67	1.64	0.61
$\sigma(r_f)$	2.47	0.51	0.34

The model also produces a low subjective real risk-free rate with low volatility.<sup>12</sup> The volatility of  $r_f$  in the data (2.47%) is higher than in the model (0.51%), but one should keep in mind that the inflation expectations are estimated with error and this measurement error contributes at least some of the empirically observed volatility in  $r_f$ . The low volatility of  $r_f$  is a virtue of the model (which is why Campbell and Cochrane (1999), for example, specifically reverse-engineer their model to produce a constant risk-free rate).

Overall, the model provides a reasonably good fit to standard unconditional asset pricing moments. However, the most interesting predictions of the model concern time-variation in objective and subjective conditional moments, which we turn to next.

#### *IV.B. Predictability of Excess Returns*

We now evaluate time-variation in the objective equity premium. In our model, this time-variation is induced by subjective belief dynamics rather than time-varying risk aversion or time-varying objective risk that generate time-varying risk premia in rational expectations models and this mechanism leads us to construct new return predictors  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$ .

Table VI presents the results from predictive regressions of log excess returns on the equity claim in data simulated from the model. The estimates in this table are the model-implied counterpart to the empirical predictive regression results in Table I. The first block of rows presents mean coefficients and adj.  $R^2$  from regressions with  $\tilde{\mu}_d$  as predictor variable. In column (1), the prediction horizon is one quarter, as in Table I. We find coefficients on  $\tilde{\mu}_d$  that are about half as big as those we found in the empirical data including the Great Depression, but quite close to the estimate from the post-WWII sample. Since the volatility of  $\tilde{\mu}_d$  is somewhat higher in the calibrated model than in the full empirical sample, this means that the variation in the objective risk premium in our model is at least roughly similar to

12. For the purpose of this moments comparison, we calculate the subjective real risk-free rate in the data using expected inflation expectation an AR(1) constant-gain learning inflation forecast with gain  $\nu = 0.018$ . Malmendier and Nagel (2016) show that this AR(1) constant-gain learning forecast fits household inflation expectations from the Michigan Survey of Consumers well (and we can construct it in the decades before survey data becomes available).

TABLE VI  
PREDICTIVE REGRESSIONS IN SIMULATED DATA

This table reports the mean return predictability regression coefficients and adj.  $R^2$  across 10,000 simulations of the model for 360 quarters plus a 400-quarter burn-in period to compute  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$  at the start of each simulated sample. The dependent variable is the log excess return on the equity claim. The predictors  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$  are constructed as the exponentially-weighted average of experienced payout growth and experienced log returns, respectively, with gain parameter  $\nu = 0.018$ . Each block of rows represents regressions with a different (single) predictor variable. Columns (1) to (3) show results using a diffuse prior ( $\phi = 1$ ) and  $\psi = 1$ . Columns (4) to (6) show the corresponding results for  $\psi = 1.5$  and an informative prior with  $\phi = 0.99$ .

	$\psi = 1, \phi = 1$			$\psi = 1.5, \phi = 0.99$		
	1Q	1Y	5Y	1Q	1Y	5Y
	(1)	(2)	(3)	(4)	(5)	(6)
$\tilde{\mu}_d$	-2.40	-9.24	-38.37	-3.07	-11.78	-48.84
$R_{adj}^2$	0.01	0.05	0.21	0.01	0.05	0.22
$\tilde{\mu}_r$	-1.51	-5.79	-23.77	-1.48	-5.64	-23.12
$R_{adj}^2$	0.01	0.03	0.15	0.01	0.03	0.14
$p - d$	-0.06	-0.22	-0.93	-0.05	-0.20	-0.82
$R_{adj}^2$	0.01	0.04	0.18	0.01	0.05	0.21

the variation in the data. Similar comments apply to the estimates with  $\tilde{\mu}_r$  as predictor presented in the second block of rows.

Columns (2) and (3) show the regression coefficients when returns are measured over longer horizons of one and five years, respectively. These estimate the persistence in the expected returns in the model. For example, the coefficients in column (3) where the prediction horizon is 20 times longer than in column (1) are only slightly smaller than 20 times the coefficients in column (1). On this dimension, the model also fits well with the empirical data. We didn't explicitly report the empirical predictive regressions for horizons beyond one quarter, but Figure I earlier showed how predicted 5-year returns line up well with realized 5-year returns.

The bottom block of rows shows regressions with  $p - d$  as predictor. The regression coefficients of around  $-0.06$  are somewhat larger than in the empirical data, but certainly of the right order of magnitude.

Columns (4) to (6) repeat this analysis with  $\psi = 1.5$  and an informative prior with

TABLE VII  
DYNAMICS OF SUBJECTIVE EXPECTATIONS OF EXCESS RETURNS IN SIMULATED DATA

This table reports the mean estimates from regressing subjective expected excess returns and expectation errors on  $\tilde{\mu}_d$ ,  $\tilde{\mu}_r$ , and 1-year lagged log returns from 10,000 simulations of 360 quarters with a 400-quarter burn-in period, using the model with a diffuse prior ( $\phi = 1$ ) and  $\psi = 1$ . In Panel A, the dependent variable is  $(\tilde{\mathbb{E}}_t[R_{m,t+1}]^4 - (R_{f,t})^4)$ , which we regress on experienced real returns leading up to and including quarter  $t$ , and/or lagged one-year log returns over the four quarters  $t - 3$  to  $t$ . In Panel B, the dependent variable is the expectation error, defined as  $\prod_{i=1}^4 R_{m,t+i} - (\tilde{\mathbb{E}}_t[R_{m,t+1}])^4$ .

	(1)	(2)	(3)	(4)	(5)
<i>Panel A: Subjective expected excess returns</i>					
$\tilde{\mu}_d$	0.83		0.83		
$\tilde{\mu}_r$				0.60	0.62
$r_{t-3,t}$		0.01	0.00		0.00
$R_{adj}^2$	0.93	0.08	0.93	0.83	0.83
<i>Panel B: Expectation error: Realized - subj. expected</i>					
$\tilde{\mu}_d$	-10.75		-11.05		
$\tilde{\mu}_r$				-6.80	-7.24
$r_{t-3,t}$		-0.06	0.02		0.03
$R_{adj}^2$	0.05	0.01	0.06	0.04	0.04

$\phi = 0.99$ . Return predictability with  $\tilde{\mu}_d$  gets somewhat stronger and closer to the empirical magnitudes that we found in the full sample in Table I. But overall, the effects of changing the EIS are quite small.

#### IV.C. Predictability of Subjective Expectations Errors

To replicate the subjective expectations error regressions that we ran on the empirical data, we use the model-implied subjective equity premium as dependent variable, calculated as explained in Section III.F. Table VII is the simulated counterpart to the empirical results in Table II. Panel A shows that the model-generated data yields a relationship between subjective expected excess returns and experienced returns that is weakly positive and quantitatively similar to the empirical estimates. The empirical point estimate in column (1) of Table II, with  $\tilde{\mu}_d$  as predictor, is 0.31, while the regression on model-generated data in

column (1) of Table VII yields a mean coefficient of 0.83. With  $\tilde{\mu}_r$  as predictor in column (4), the simulated data yields a mean coefficient of 0.60, close to the empirical estimate of 0.86 in Table II, column (4). But it is also useful to keep in mind that the volatility of subjective expected returns in the simulated data is tiny relative to the volatility of objective expected returns. Therefore, the take-away from this analysis is that subjective expected excess returns both in the data and the model are nearly constant.<sup>13</sup>

In Panel B, with subjective expectations errors as dependent variable, the mean coefficient on  $\tilde{\mu}_d$  in column (1) is very similar to its empirical counterpart of  $-12.34$  in Table II. With  $\tilde{\mu}_r$  as a predictor in column (4), the model-implied coefficients are about half as big as the empirical ones, so the empirical experienced real returns variable may capture some additional expectations errors beyond those implied by our model.

Summing up, the dynamics of subjective and objective expected returns in the model are broadly consistent with the empirical data. Objectively, excess returns are strongly predictable by  $\tilde{\mu}_d$  and  $\tilde{\mu}_r$ , while subjective expected excess returns are largely acyclical.

#### *IV.D. Subjective Long-run Growth Expectations*

As the final piece of our comparison of the model with our earlier empirical estimates from the reduced-form framework, we look at the dynamics of subjective dividend growth expectations.

By iterating on eqs. (16) and (31) we can construct, at every point in time in our simulations, the representative agent's subjective expectation of average log dividend growth over the next 20 quarters (1,000 simulations of the model for 50,000 periods). The 20-quarter forecast horizon corresponds roughly to the analyst forecast horizon in our analysis in Table III. In each simulation run, we regress this subjective dividend growth expectation at time  $t$  on  $\tilde{\mu}_{d,t}$  or  $\tilde{\mu}_{r,t}$ . We obtain coefficients of 0.899 and 0.683. These values are extremely close to the implied coefficients of 0.85 and 0.64 that we obtained in Table III after adjusting

13. A comparison of the predictive regression coefficients in Tables VI and VII shows that objective expected returns at a one-year horizon have a volatility that is more than 10 times as high (absolute regression coefficient of 9.24 times volatility of quarterly  $\tilde{\mu}_d$  of  $0.39\% \approx 3.60\%$ ) as the volatility of subjective expected returns (regression coefficient of 0.83 times volatility of quarterly  $\tilde{\mu}_d \approx 0.32\%$ ).

for the distortionary effects of transitory predictable components in earnings on the analyst forecasts. Thus, our model quantitatively matches the dynamics of long-term forecasts of growth in stock market fundamentals by stock market analysts.

## V. RELATIONSHIP BETWEEN FIRST AND SECOND MOMENTS OF EQUITY RETURNS

Another interesting implication of our model that we have not discussed so far is that its predictions about the relationship between conditional first and second moments of equity market returns are sharply different from those of the leading rational expectations asset pricing models by CC and BY. In CC and BY, the conditional equity premium is approximately linearly and positively related to conditional equity return variance. In CC, this happens because at times when risk aversion is high, it is also very volatile. In BY, the reason is that stochastic volatility in endowment growth is the driver of the time-varying equity premium.<sup>14</sup> Therefore, variables like  $p - d$  that predict excess returns should also predict equity market return variance with the same sign—but empirical data does not support this prediction. This gives rise to what Lettau and Ludvigson (2010) have termed the *Sharpe Ratio variability puzzle*: These models make the empirically counterfactual prediction that positive co-movement of conditional first and second moments should dampen the variability of the Sharpe Ratio compared with the variability in expected excess returns.

Table VIII shows a similar empirical result when we use experienced payout growth,  $\tilde{\mu}_d$ , as a predictor variable. The dependent variable in these regressions is the sum of squared daily log returns of the CRSP value-weighted index in quarter  $t + 1$ , i.e., an estimate of quarterly realized variance. Recall that  $\tilde{\mu}_d$  predicts returns with a negative sign. However, as column (1) and (4) show, there is a weakly positive, but not statistically significant relationship between  $\tilde{\mu}_d$  and next-quarter variance both in the full sample and the post-WW II sample.

14. For the BY model, compare equations (A13) and (A14) in the appendix of BY. For the CC model one can infer the positive and close to linear relationship by comparing Figures 4 and 5 in CC.

The magnitude of the point estimate is small: a one standard deviation increase in  $\tilde{\mu}_d$  is associated with an increase of 0.15 percentage points in conditional quarterly return variance, which is about 10% of the sample standard deviation of the return variance. Using  $p - d$  as a predictor in columns (2) and (5), the picture is mixed. In the full sample,  $p - d$  predicts volatility negatively, but the coefficient estimate is not statistically significant. In the post-WW II period, the estimate is statistically significant, but it is positive, which is inconsistent with the CC and BY models. Columns (3) and (6) control for the risk-free rate because  $p - d$  and the real risk-free rate together should span the state variables in the BY model, but this has little effect on the coefficient of the  $p - d$  ratio. Overall, there is no support for the prediction of the CC and BY models that the conditional equity premium has strong positive correlation with conditional market return variance. At low frequencies captured by the slow-moving predictors  $\tilde{\mu}_d$  and  $p - d$ , there is not any co-movement between equity premium and volatility.

In contrast to CC and BY, our model is consistent with this empirical result. Learning with fading memory generates predictable variation in the objective equity premium, but without simultaneous variation in equity market return variance.<sup>15</sup> The decoupling objective and subjective beliefs allows the model to also decouple conditional equity return volatility and the objective conditional equity premium.

## VI. LACK OF OUT-OF-SAMPLE RETURN PREDICTABILITY

Welch and Goyal (2008) show that the simple trailing sample mean of past returns often beats an out-of-sample predictive regression forecast as a predictor of future returns. Since the representative agent in our model discards historical information at a relatively high rate (the half-life in terms of the observation's weight in the log likelihood is about 10 years), one

15. Our model actually produces a positive relationship between  $\tilde{\mu}_d$  and conditional variance of the equity claim return, but the effect is very weak, somewhat similar to the weak empirical relationship in Table VIII. Subjective and objective conditional variance of dividend strip returns are constant over time. However, higher subjectively expected growth implies higher weight of riskier longer-horizon dividend strips in the equity claim's value and hence higher variance of the return on the equity claim.

TABLE VIII  
PREDICTING MARKET RETURN VARIANCE

Dependent variable is the sum of squared daily log returns of the CRSP value-weighted index in quarter  $t + 1$ . Experienced real payout growth denotes a long-run exponentially weighted average of overlapping quarterly observations of four-quarter per-capita repurchase-adjusted real dividend growth rates leading up to and including quarter  $t$ , constructed with weights implied by constant gain learning with quarterly gain  $\nu = 0.018$ ;  $p - d$  refers to the log price-dividend ratio of the CRSP value-weighted index at the end of quarter  $t$ . For readability of the estimates, we use  $(p - d)/100$  as predictor variable. Newey-West standard errors with six lags are shown in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
	1927-2016	1927-2016	1927-2016	1946-2016	1946-2016	1946-2016
Experienced real payout growth	0.47 (0.41)			0.41 (0.35)		
$(p - d)/100$		-0.30 (0.42)	-0.30 (0.40)		0.41 (0.12)	0.40 (0.12)
Real risk-free rate			0.21 (0.26)			-0.17 (0.14)
Constant	0.00 (0.00)	0.02 (0.02)	0.02 (0.01)	0.00 (0.00)	-0.01 (0.00)	-0.01 (0.00)
Observations	360	360	360	284	284	284
Adjusted $R^2$	0.012	0.008	0.016	0.015	0.033	0.044

might suspect that a predictive regression run in real time, but with full memory of past data, should be able to identify the agent's errors and hence predict returns out-of-sample better than the sample mean. However, as we show now, this is not the case.

We apply the Goyal and Welch analysis to simulated data from our model. We run 10,000 simulations of a 360-quarter sample period with a 400-quarter burn-in period to compute  $\tilde{\mu}_d$  at the start of each sample. Within each 360-quarter sample, we then examine the in-sample explanatory power of the predictive regression by plotting the cumulative squared demeaned excess returns minus the cumulative squared full-sample regression residual from the beginning to the end of the sample. The predictive regression is run at quarterly frequency with the sum of four-quarter log excess returns from  $t + 1$  to  $t + 4$  as dependent variable and  $p - d$  or  $\tilde{\mu}_d$  as predictor. The blue line in the upper half of each plot in Figure III shows the average path across all simulations of this in-sample cumulative squared errors difference.



The upward slope of this line and the fact that it ends up, on the right-hand side, above zero, both with  $p - d$  as a predictor (top) and  $\tilde{\mu}_d$  (bottom), indicates that the in-sample  $R^2$  is greater than zero.

To assess the out-of-sample performance with each 360-quarter sample, we use a recursively expanding window, starting out at 80 quarters, to estimate the sample mean and the predictive regression. We calculate the next period out-of-sample squared prediction error of the trailing sample mean as a forecaster minus the squared prediction error of the fitted predictive regression. We cumulate these squared error differences forward and we average the resulting paths across all simulations. The red lines in the lower half of each plot in Figure III show that this path is, on average, in negative territory, which means that the predictive regression forecast underperforms the trailing sample mean as a forecaster. That the slope is still negative on average towards the end of the sample period shows that even after having observed almost 90 years of data, the trailing sample mean is typically still a better forecaster.

Thus, even though there is true return predictability in this model under the econometrician's objective probability measure, this predictability is not exploitable in real-time for typical sample sizes. The data generated by the model is therefore consistent with the lack of out-of-sample predictability found empirically by Welch and Goyal (2008).

The out-of-sample exercise also demonstrates that it would not be easy for the agent within the model to recognize that the loss of memory and the resulting reliance on relatively recent experiences in estimating endowment growth rates is detrimental to forecast performance. In this sense, one can interpret our model as a near-rational model.

## VII. CONCLUSION

We have shown that learning with fading memory can reconcile asset prices and survey expectations in a highly tractable framework. In our model, asset prices are volatile because subjective growth expectations are time-varying. Risk premia are high because subjective

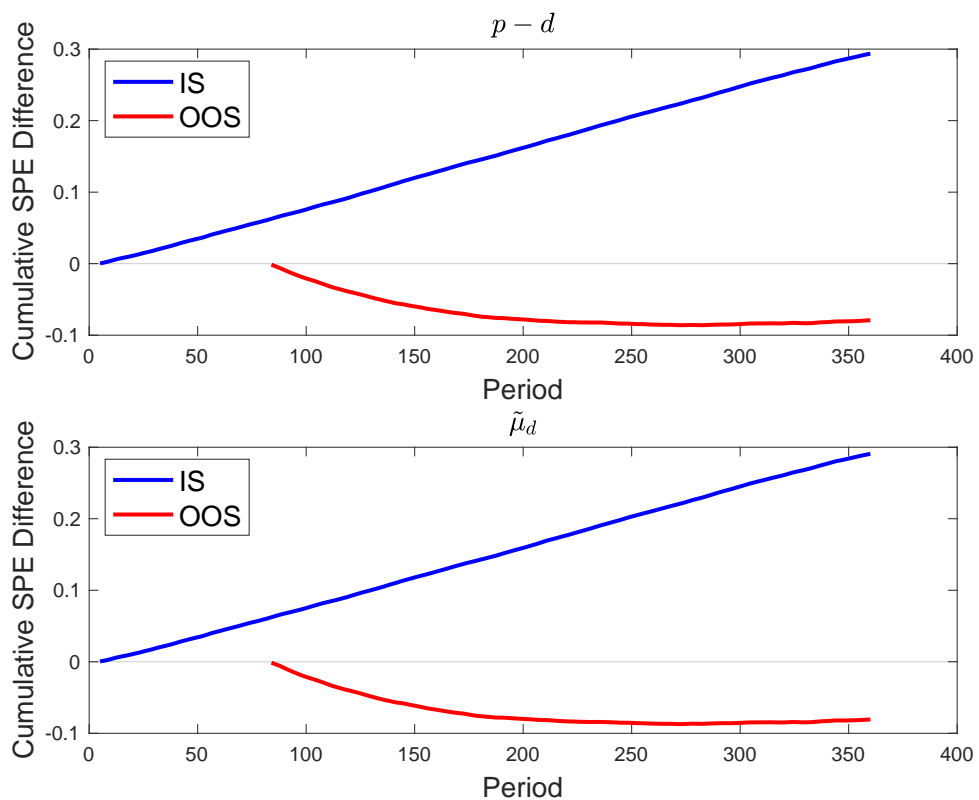


FIGURE III  
 Out-of-sample predictive performance

In-sample and out-of-sample performance of predictors  $p-d$  and  $\tilde{\mu}_d$  from 10,000 simulations of 360 quarters with a 400-quarter burn-in period to compute  $\tilde{\mu}_d$  at the start of each sample. The IS line plots the cumulative squared demeaned excess returns minus the cumulative squared full-sample regression residual. The OOS line plots the cumulative squared prediction errors of conditional mean minus the cumulative squared prediction errors of predictors. Both lines are the average path across simulations.

long-run growth rate uncertainty is high. The model produces realistic asset price behavior in a simple setting with IID endowment growth and constant risk aversion. While objective expected excess returns are strongly counter-cyclical, subjective beliefs about stock market excess returns are slightly pro-cyclical. As a consequence, subjective expectations errors are predictable, as they are in the survey data. As predicted by the model, long-run weighted averages of past real per-capita payout growth or past real stock index returns are a good empirical predictor of excess returns and subjective expectations errors. Unlike in leading rational expectations explanations of return predictability, and consistent with the data, movements in objective expected excess returns in our model are not associated with movements in conditional market return volatility.

Because memory of past data fades away, subjective beliefs about long-run growth fluctuate perpetually in our model. That these belief fluctuations persist is plausible because it would be difficult for an agent to detect that the loss of memory is detrimental to her investment decisions. While returns generated by our model economy are predictable to an econometrician examining a sample ex post, standard out-of-sample tests show that they are not predictable in real time in typical sample sizes, consistent with the empirical lack of out-of-sample predictability. Overall, these results suggest that subjective belief dynamics could be central to asset pricing and that learning with fading memory can provide a unifying account of many asset pricing phenomena and the evidence on subjective beliefs about stock returns in investor surveys.

## REFERENCES

- Adam, Klaus, Albert Marcet, and Johannes Beutel, 2017, “Stock price booms and expected capital gains,” *American Economic Review* 107, 2352–2408.
- Ameriks, John, Gábor Kézdi, Minjoon Lee, and Matthew D Shapiro, 2016, “Heterogeneity in Expectations, Risk Tolerance, and Household Stock Shares,” Working paper, University of Michigan.
- Amromin, Gene, and Steven A Sharpe, 2013, “From the Horse’s Mouth: Economic Conditions and Investor Expectations of Risk and Return,” *Management Science*.
- Andonov, Aleksandar, and Joshua Rauh, 2017, “The Return Expectations of Institutional Investors,” Working paper, Stanford University.
- Azeredo da Silveira, Rava, and Michael Woodford, 2019, “Noisy Memory and Over-Reaction to News,” *AEA Papers and Proceedings* 109, 557–61.
- Balke, Nathan S, and Robert J Gordon, 1989, “The Estimation of Prewar Gross National Product: Methodology and New evidence,” *Journal of Political Economy* 97, 38–92.
- Bansal, R., R. Dittmar, and C. Lundblad, 2005, “Consumption, Dividends, and the Cross Section of Equity Returns,” *Journal of Finance* 60, 1639–1672.
- Bansal, Ravi, A. Ronald Gallant, and George Tauchen, 2007, “Rational Pessimism, Rational Exuberance, and Markets for Macro Risks,” *Review of Economic Studies* 74, 1005–1033.
- Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, “An Empirical Evaluation of the Long-Run Risks Model for Asset Prices,” *Critical Finance Review* 1, 183–221.
- Bansal, Ravi, and Amir Yaron, 2004, “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles,” *Journal of Finance* 59, 1481–1509.
- Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015, “X-CAPM: An Extrapolative Capital Asset Pricing Model,” *Journal of Financial Economics* 115, 1–24.
- Barberis, Nicholas, Ming Huang, and Tano Santos, 2001, “Prospect Theory and Asset Prices,” *Quarterly Journal of Economics* 116, 1–53.
- Barro, Robert J, and Xavier Sala-i-Martin. *Economic Growth* (2nd edition). (MIT Press, Cambridge, MA, 2004).
- Barro, Robert J, and Jose F Ursua, 2008, “Consumption Disasters since 1870,” *Brookings Papers on Economic Activity* 1, 225–350.
- Beeler, Jason, and John Y. Campbell, 2012, “The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment,” *Critical Finance Review* 1, 141–182.
- Berardi, Michele, and Jaqueson K Galimberti, 2017, “Empirical Calibration of Adaptive Learning,” *Journal of Economic Behavior & Organization* 144, 219–237.
- Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2017, “Diagnostic Expectations and Stock Returns,” *Journal of Finance*, forthcoming.

- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2019, “Memory, Attention, and Choice,” Working paper, Harvard University.
- Branch, William A., and George W. Evans, 2006, “A Simple Recursive Forecasting Model,” *Economics Letters* 91, 158–166.
- Campbell, John Y., 1991, “A Variance Decomposition for Stock Returns,” *Economic Journal* 101, 157–179.
- Campbell, John Y., 2000, “Asset Pricing at the Millenium,” *Journal of Finance* 55, 1515–1567.
- Campbell, John Y., and John H. Cochrane, 1999, “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy* 107, 205–251.
- Campbell, John Y., and Robert J. Shiller, 1988, “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors,” *Review of Financial Studies* 1, 195–228.
- Choi, Jung Ho, Alon Kalay, and Gil Sadka, 2016, “Earnings News, Expected Earnings, and Aggregate Stock Returns,” *Journal of Financial Markets* 29, 110–143.
- Choi, James J., and Thomas Mertens, 2013, “Extrapolative Expectations and the Equity Premium,” Working paper, Yale University.
- Collin-Dufresne, Pierre, Michael Johannes, and Lars A Lochstoer, 2016, “Parameter Learning in General Equilibrium: The Asset Pricing Implications,” *American Economic Review* 106, 664–698.
- Collin-Dufresne, Pierre, Michael Johannes, and Lars A Lochstoer, 2017, “Asset Pricing When ‘This Time is Different’,” *Review of Financial Studies* 30, 505–535.
- Ehling, Paul, Alessandro Graniero, and Christian Heyerdahl-Larsen, 2018, “Asset Prices and Portfolio Choice with Learning from Experience,” *Review of Economic Studies*, forthcoming.
- Epstein, Larry G., and Stanley E. Zin, 1989, “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica* 57, 937–969.
- Evans, George W., and Seppo Honkapohja. *Learning and Expectations in Macroeconomics* (Princeton University Press, Princeton, NJ, 2001).
- Greenwood, Robin, and Stefan Nagel, 2009, “Inexperienced Investors and Bubbles,” *Journal of Financial Economics* 93, 239–258.
- Greenwood, Robin, and Andrei Shleifer, 2014, “Expectations of returns and expected returns,” *Review of Financial Studies* 27, 714–746.
- Hamilton, James D. *Time Series Analysis*, 1994).
- Hansen, Lars Peter, 2007, “Beliefs, Doubts and Learning: Valuing Macroeconomic Risk,” *American Economic Review* 97, 1–30.

- Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, “Consumption Strikes Back? Measuring Long-Run Risk,” *Journal of Political Economy* 116, 260–302.
- Hansen, Lars Peter, and José A Scheinkman, 2012, “Recursive Utility in a Markov Environment with Stochastic Growth,” *Proceedings of the National Academy of Sciences* 109, 11967–11972.
- Hirshleifer, David, Jun Li, and Jianfeng Yu, 2015, “Asset Pricing in Production Economies with Extrapolative Expectations,” *Journal of Monetary Economics* 76, 87–106.
- Jin, Lawrence J, and Pengfei Sui, 2018, “Asset Pricing with Return Extrapolation.”
- Kothari, SP, Jonathan Lewellen, and Jerold B Warner, 2006, “Stock Returns, Aggregate Earnings Surprises, and Behavioral Finance,” *Journal of Financial Economics* 79, 537–568.
- Kothari, S.P., and Jay Shanken, 1997, “Book-to-market, dividend yield, and expected market returns: A time-series analysis,” *Journal of Financial Economics* 44, 169–203.
- La Porta, Rafael, 1996, “Expectations and the Cross-Section of Stock Returns,” *The Journal of Finance* 51, 1715–1742.
- Lease, Ronald C, Wilbur G Lewellen, and Gary G Schlarbaum, 1974, “The individual investor: attributes and attitudes,” *Journal of Finance* 29, 413–433.
- Lettau, Martin, and Sydney C Ludvigson, “Measuring and Modeling Variation in the Risk-Return Trade-Off,” in *Handbook of Financial Econometrics: Tools and Techniques*, pp. 617–690, (2010).
- Lewellen, Jonathan, and Jay Shanken, 2002, “Learning, Asset-Pricing Tests and Market Efficiency,” *Journal of Finance* 57, 1113–1145.
- Lewellen, Wilbur G, Ronald C Lease, and Gary G Schlarbaum, 1977, “Patterns of investment strategy and behavior among individual investors,” *Journal of Business* 50, 296–333.
- Malmendier, Ulrike, and Stefan Nagel, 2011, “Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?,” *Quarterly Journal of Economics* 126, 373–416.
- Malmendier, Ulrike, and Stefan Nagel, 2016, “Learning from Inflation Experiences,” *Quarterly Journal of Economics* 131, 53–87.
- Malmendier, Ulrike, Stefan Nagel, and Yan Zhen, 2018, “The Making of Hawks and Doves,” Working paper, University of Chicago.
- Malmendier, Ulrike, Demian Pouzo, and Victoria Vanasco, 2017, “Investor Experiences and Financial Market Dynamics,” Working paper, UC Berkeley.
- Mangel, Marc, 1990, “Dynamic Information in Uncertain and Changing Worlds,” *Journal of Theoretical Biology* 146, 317–332.
- Nagel, Stefan, 2016, “The Liquidity Premium of Near-Money Assets,” *Quarterly Journal of Economics* 131, 1927–1971.
- Nakov, Anton, and Galo Nuño, 2015, “Learning from Experience in the Stock Market,” *Journal of Economic Dynamics and Control* 52, 224–239.

- Pástor, Ľuboš, and Pietro Veronesi, 2003, “Stock Valuation and Learning About Profitability,” *Journal of Finance* 58, 1749–1790.
- Pástor, Ľuboš, and Pietro Veronesi, 2006, “Was there a Nasdaq bubble in the late 1990s?,” *Journal of Financial Economics* 81, 61–100.
- Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, 2018, “Distributional National Accounts: Methods and Estimates for the United States,” *Quarterly Journal of Economics* 133, 553–609.
- Pohl, Walt, Karl Schmedders, and Ole Wilms, 2018, “Higher-Order Effects in Asset-Pricing Models with Long-Run Risks,” *Journal of Finance*, forthcoming.
- Sadka, Gil, and Ronnie Sadka, 2009, “Predictability and the Earnings–Returns Relation,” *Journal of Financial Economics* 94, 87–106.
- Schraeder, Stefanie, 2015, “Information Processing and Non-Bayesian Learning in Financial Markets,” *Review of Finance* 20, 823–853.
- Sharpe, Steven A, 2002, “Reexamining Stock Valuation and Inflation: The Implications of Analysts’ Earnings Forecasts,” *Review of Economics and Statistics* 84, 632–648.
- Shiller, Robert J. *Irrational Exuberance* (second edition). (Princeton University Press, Princeton, NJ, 2005).
- Thomson Reuters, 2009. Methodology for estimates.
- Timmermann, Allan G, 1993, “How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices,” *Quarterly Journal of Economics*, 1135–1145.
- Vissing-Jorgensen, Annette, “Perspectives on Behavioral Finance: Does ”Irrationality” Disappear with Wealth? Evidence from Expectations and Actions,” in *NBER Macroeconomics Annual*.
- Wachter, Jessica A, and Michael J Kahana, 2019, “A Retrieved-Context Theory of Financial Decisions,” Working paper, University of Pennsylvania.
- Welch, Ivo, and Amit Goyal, 2008, “A Comprehensive Look at The Empirical Performance of Equity Premium Prediction,” *Review of Financial Studies* 21, 1455–1508.
- Wright, Stephen, 2004, “Measures of Stock Market Value and Returns for the US Nonfinancial Corporate Sector, 1900–2002,” *Review of Income and Wealth* 50, 561–584.

## Appendix (for online publication)

### A. DATA

#### A.1. Financial market data

*Population.* We obtain quarterly data on the population of the United States from 1947 onwards from the FRED database at the Federal Reserve Bank of St. Louis. For 1871 to 1946, we obtain annual population numbers from the Historical Statistics of the United States, available at <https://hsus.cambridge.org/>.

*Stock index returns.* We calculate quarterly stock index returns from 1871 to 1925 using data from Shiller (2005) back to 1871. From 1926 onwards, we use quarterly returns on the value-weighted CRSP index.

*Inflation.* We use the Consumer Price Index (CPI) series (and the pre-cursors of the official CPI) in Shiller’s data set to deflate returns and the payout growth series that we describe next.

*Payouts.* From 1926 onwards, we calculate quarterly aggregate dividends using the lagged total market value of the CRSP value-weighted index and the difference between quarterly returns with and without dividends. Furthermore, working with the monthly CRSP individual stock files and following Bansal, Dittmar, and Lundblad (2005), we use reductions in the shares outstanding (after adjusting for stock splits, stock dividends, etc. using the CRSP share adjustment factors) as reported by CRSP to calculate stock repurchases. To eliminate the effect of data errors (there are instances where the shares outstanding drop by a huge amount and jump back up a few months later), we drop observations where the shares outstanding fall by more than 10 percent within one month. Repurchases account for an economically significant share of payouts only from the 1980s onwards. We aggregate the sum of dividends and repurchases across firm within each month and then at the aggregate level within each quarter. Dividing by the size population at each point in time, we obtain per-capita payouts. To avoid seasonality effects, we compute growth rates as the four-quarter change in log per-capita aggregate payouts, divided by four.

Prior to 1926, we use annual data on aggregate household dividend receipts from tax data in Piketty, Saez, and Zucman (2018) for the period 1913 to 1926. While this data source covers only the portion of dividends received by households, its advantage is that it is based on high-quality administrative data. As long as the share of total aggregate dividends received by households does not change much from year to year, the growth rates calculated from this data set should approximate well the growth rates in total aggregate dividends. Figure A.1 suggests that this is the case: in years when the Piketty-Saez data overlaps with CRSP, the per-capita growth rates obtained from the two data sets are closely aligned.

For the period from 1900 to 1913, we use a series of annual aggregate corporate non-farm non-financial dividends from Wright (2004). Figure A.1 shows that the per-capita growth rates calculated from the Wright data are, with a few exceptions, very close to those from the Piketty-Saez data. For the period from 1871 to 1900, we use real per-capita GDP growth rates from Barro and Ursua (2008) (which are in turn based on Balke and Gordon (1989)) as proxy for  $\Delta d$  from 1871 to 1900.



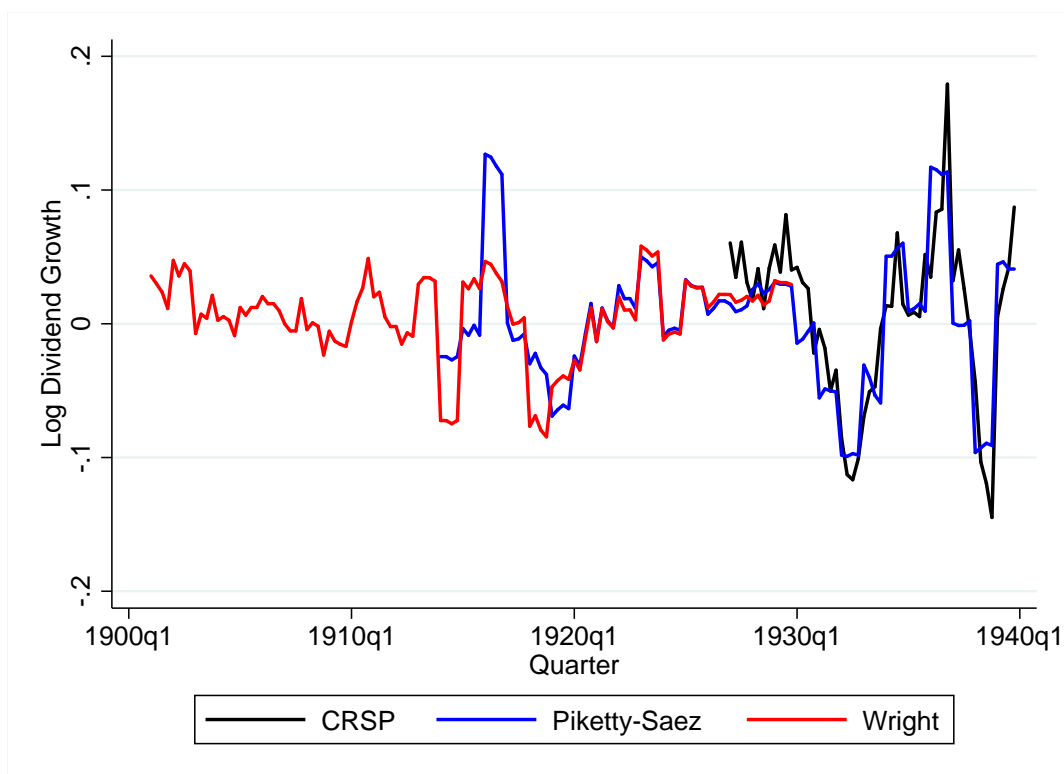


FIGURE A.1  
Dividend growth from three data sources

*Bond and bill yields.* To calculate subjectively expected excess return from survey expectations, we use the one-year constant maturity Treasury yield, obtained from FRED database at the Federal Reserve Bank of St. Louis. To calculate quarterly stock market index excess returns back to 1926, we use the three-month T-Bill yield series from Nagel (2016), extended until the end of 2016 with 3-month T-bill yields from the FRED database (where we convert the reported discount yields into effective annual yields).

#### A.2. Survey data on return expectations

Three surveys provide us with direct measures of percentage expected stock market returns over a one-year horizon: UBS/Gallup (1998-2007, monthly); Vanguard Research Initiative (VRI) survey of Ameriks, Kézdi, Lee, and Shapiro (2016) (2014, one survey); and Surveys of Lease, Lewellen, and Schlarbaum (1974) and Lewellen, Lease, and Schlarbaum (1977) (one survey per year in 1972 and 1973). In part of the sample, the UBS/Gallup survey respondents report only the return they expect on their own portfolio. We impute market return expectations by regressing expected market returns on own portfolio expectations using the part of the sample where both are available and using the fitted value from this regression when the market return expectation is not reported. The VRI survey asks about

expected growth in the Dow Jones Industrial Average (DJIA). Since the DJIA is a price index, we add to the price growth expectation the dividend yield of the CRSP value-weighted index at the time of the survey. Figure A.2 shows the time series of expectations from these surveys.

In the next step, we use the data from the Michigan Survey of Consumers (MSC). The MSC elicits the perceived probability that an investment in a diversified stock fund would increase in value over a one-year horizon. For comparability with the other surveys above, which are all based on surveys of people that hold stocks, we restrict the sample to respondents that report to hold stocks (as the MSC does for the aggregate stock market beliefs series that they publish on their website). To impute percentage expectations, we regress the percentage expectations from the UBS/Gallup and VRI surveys on the MSC probability. The red line in Figure A.2 shows the resulting fitted value. In the periods when the series overlaps with the UBS/Gallup and VRI samples, the fit is very good, indicating that the simple imputation procedure delivers reasonable results. In time periods when the UBS/Gallup and VRI surveys are not available, we use this fitted value.

Finally, we bring in data from the Conference Board (1986-2016, monthly) and Roper surveys (1974-1997, one survey per year). These surveys elicit respondents simple categorical beliefs about whether the “stock prices” will likely increase, decrease, or stay the same (or whether they are undecided, which we include in the “same” category). We construct the ratio of the proportion of those who respond with “increase” to the sum of the proportions of “decrease” and “same.” We then regress the expected return series that we obtained from the surveys above on this ratio. More precisely, since the Conference Board and Roper surveys ask about stock price increases, we subtract the current dividend yield of the CRSP value weighted index from the dependent variable in this regression and we add it back to the fitted value. The green line in Figure A.2 shows the fitted value from this regression for the Conference Board series and the four squares show the fitted value for the Roper surveys. Except for a relatively short period around the year 2000, the fitted series tracks the expected returns from UBS/Gallup, MSC, and VRI very well. In time periods when the UBS/Gallup, MSC and VRI surveys are not available, we use this fitted value.

### *A.3. Analysts long-term earnings expectations*

We collect the median of analysts’ stock-level long-term median earnings-per-share (EPS) growth forecasts from the I/B/E/S Unadjusted US Summary Statistics database, focusing on U.S. firms with earnings forecasts in U.S. dollars. The data is available at a monthly frequency starting in December 1981, but shares outstanding information that we need to aggregate across stocks becomes available only in September 1984 for all but 10% of stocks covered by I/B/E/S. We therefore start the sample in September 1984. We also collect the median forecasts of current fiscal year EPS. We use those to calculate forecasted current year total earnings by multiplying forecasted EPS with shares outstanding. We aggregate across stocks by forming a value-weighted average each quarter using total forecasted current fiscal year earnings as the weight. Stocks with negative values of this weight variable are excluded. We then aggregate the monthly long-term EPS growth forecasts across months within each calendar quarter.

We also collect the median forecasts for next fiscal year EPS. We calculate the growth

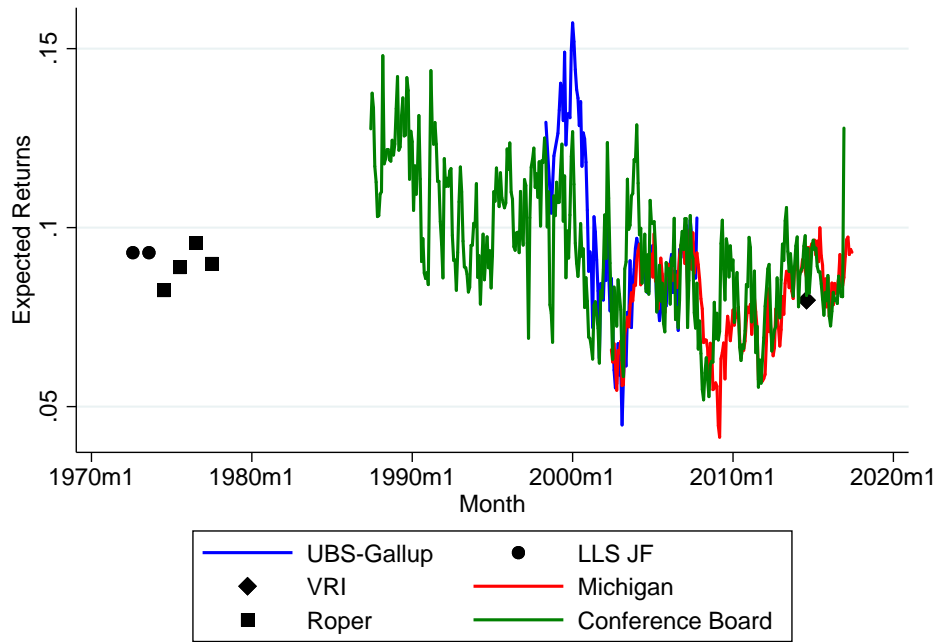


FIGURE A.2  
Expected return imputation

rate from current year to next-year earnings. We then aggregate the data, across stocks, then within quarter, in the same way as described above for the long-term forecasts.

We convert these nominal expected growth rates to real expected growth rates by deflating with inflation forecasts from the Survey of Professional Forecasters (SPF), which we obtain from the Federal Reserve Bank of Philadelphia.<sup>16</sup> We use the one-year and ten-year CPI median inflation forecasts in the quarter prior to the quarter in which the analyst forecast is made. The 10-year forecast is available in the SPF only starting in quarter 1991:4. Prior to this date we use the extended series constructed by the Philadelphia Fed from the Blue Chip Economic Indicators and the Livingston survey that is available on the same website. In quarters when the extended series has missing values, we substitute the value from the previous quarter.

Long-term earnings growth rates in the I/B/E/S data are meant to represent an expected growth rate over the next three to five years, or a “full business cycle” (see, e.g., Sharpe (2002)). To deflate with an expected inflation rate that approximates this forecast horizon, we subtract the average of the one-year and ten-year inflation forecast from the long-term earnings forecast. We subtract the one-year inflation forecast from the one-year earnings

16. Available at <https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/historical-data/inflation-forecasts>

growth forecast. As a last step, we convert the real expected earnings growth series to quarterly frequency by dividing the annualized numbers by four.

## B. BOOTSTRAP SIMULATIONS FOR PREDICTIVE REGRESSIONS

Our bootstrap simulations closely follow those in Kothari and Shanken (1997), but extended to multiple predictor variables. We start by estimating AR(1) processes for the predictor variables and we add  $1/T$  to the slope coefficient to perform first-order bias-adjustment (and we adjust the intercept accordingly). We also estimate the predictive regression for returns by OLS and record the residuals.

Using these bias-adjusted coefficients from the estimated AR(1) for the predictors, we then simulate a VAR(1) with a diagonal coefficient matrix, where the innovations are the bootstrapped residuals from the estimated AR(1). As in Kothari and Shanken (1997), we condition on the first observation of the predictor time series. We preserve contemporaneous correlations of the innovations by drawing vectors of residuals for the different predictors.

Based on the simulated predictor series, we then also simulate two return series by combining the predictor time-series with bootstrapped residuals from the predictive regression. For the first return series, we set the predictive regression slope coefficients equal to the OLS predictive regression estimate, i.e., we simulate under the alternative. For the second series, we set the predictive regression slope coefficients equal to zero, i.e., in this case we are simulating under the null hypothesis of no predictability.

We use the described approach to simulate 10,000 bootstrap samples of predictors and the two returns series. We then run the predictive regressions on the bootstrap samples and record the regression coefficients and  $t$ -statistics. We obtain the predictive regression bias-adjustment by comparing the mean slope coefficients from the bootstrap samples with the first return series (alternative) to the OLS estimate. We obtain the  $p$ -values by comparing the sample predictive regression  $t$ -statistic to the quantiles of the distribution of the  $t$ -statistic in the bootstrap regressions with the second return series (null).

## C. PROPERTIES OF THE PREDICTIVE DISTRIBUTION

We describe now the properties of  $\tilde{\varepsilon}_{t+j}$ ,  $j = 1, 2, \dots$  under the time- $t$  predictive distribution.

We first show that the subjective conditional variance of  $\tilde{\varepsilon}_{t+j}$  is decreasing in the forecast horizon. First note the perceived consumption growth process has the following autocovariance structure

$$\text{cov}_t(\Delta c_{t+i}, \Delta c_{t+j}) = \nu \sigma^2, \quad j > i \geq 1, \quad (\text{A.1})$$

which arises from the agent's uncertainty about  $\mu$ . From the definition of  $\tilde{\varepsilon}$ , we obtain

$$\begin{aligned}
\tilde{\varepsilon}_{t+i} &= \frac{\Delta c_{t+i} - \tilde{\mu}_{t+i-1}}{\sqrt{1+\nu\sigma}}. \\
&= \frac{\Delta c_{t+i} - \nu\Delta c_{t+i-1} - (1-\nu)\tilde{\mu}_{t+i-2}}{\sqrt{1+\nu\sigma}}. \\
&= \frac{\Delta c_{t+i} - \nu\sum_{j=1}^{i-1}(1-\nu)^{j-1}\Delta c_{t+i-j} - (1-\nu)^{i-1}\tilde{\mu}_t}{\sqrt{1+\nu\sigma}}. \tag{A.2}
\end{aligned}$$

Because of the constant autocovariance structure of perceived consumption growth,

$$\begin{aligned}
\tilde{\text{var}}_t(\tilde{\varepsilon}_{t+i-1}) &= \tilde{\text{var}}_t\left(\frac{\Delta c_{t+i-1} - \nu\sum_{j=1}^{i-2}(1-\nu)^{j-1}\Delta c_{t+i-1-j}}{\sqrt{1+\nu\sigma}}\right) \\
&= \tilde{\text{var}}_t\left(\frac{\Delta c_{t+i} - \nu\sum_{j=1}^{i-2}(1-\nu)^{j-1}\Delta c_{t+i-j}}{\sqrt{1+\nu\sigma}}\right), \tag{A.3}
\end{aligned}$$

and

$$\begin{aligned}
\tilde{\text{var}}_t(\tilde{\varepsilon}_{t+i}) &= \tilde{\text{var}}_t(\tilde{\varepsilon}_{t+i-1}) + \nu^2(1-\nu)^{2i-4} - \frac{2\nu^2}{1+\nu}(1-\nu)^{2i-4} \\
&= \tilde{\text{var}}_t(\tilde{\varepsilon}_{t+i-1}) - \frac{1-\nu}{1+\nu}\nu^2(1-\nu)^{2i-4}, \quad i \geq 2. \tag{A.4}
\end{aligned}$$

This leads to

$$\tilde{\text{var}}_t(\tilde{\varepsilon}_{t+i}) = 1 - \frac{1-\nu}{1+\nu}\nu^2\frac{1-(1-\nu)^{2i-2}}{1-(1-\nu)^2}, \tag{A.5}$$

which decreases over time and converges to  $1 - \frac{1-\nu}{1+\nu}\frac{\nu}{2-\nu}$ .

Using these results, we can calculate the time- $t$  perception of the autocovariance of future  $\tilde{\varepsilon}_{t+i}$ ,

$$\begin{aligned}
\tilde{\text{cov}}_t(\tilde{\varepsilon}_{t+i}, \tilde{\varepsilon}_{t+i+1}) &= \tilde{\text{cov}}_t\left(\frac{\Delta c_{t+i} - \nu\sum_{j=1}^{i-1}(1-\nu)^{j-1}\Delta c_{t+i-j}}{\sqrt{1+\nu\sigma}}, \frac{\Delta c_{t+i+1} - \nu\sum_{j=1}^{i-1}(1-\nu)^{j-1}\Delta c_{t+i+1-j}}{\sqrt{1+\nu\sigma}}\right) \\
&= \tilde{\text{cov}}_t\left(\frac{\Delta c_{t+i} - \nu\sum_{j=1}^{i-2}(1-\nu)^{j-1}\Delta c_{t+i-j}}{\sqrt{1+\nu\sigma}}, \frac{\Delta c_{t+i+1} - \Delta c_{t+i}}{\sqrt{1+\nu\sigma}} + (1-\nu)\frac{\Delta c_{t+i} - \nu\sum_{j=1}^{i-2}(1-\nu)^{j-1}\Delta c_{t+i-j}}{\sqrt{1+\nu\sigma}}\right) \\
&= -\frac{1}{1+\nu} + (1-\nu)\tilde{\text{var}}_t(\tilde{\varepsilon}_{t+i}) \\
&= -\frac{\nu^2}{1+\nu} - \frac{(1-\nu)^2}{1+\nu}\nu^2\frac{1-(1-\nu)^{2i-2}}{1-(1-\nu)^2} < 0. \tag{A.6}
\end{aligned}$$

i.e., it is negative, as we claimed in the main text.

## D. KALMAN FILTERING INTERPRETATION

Here we show that our model can be mapped into a full memory model that is equivalent in terms of the relevant subjective belief dynamics and asset prices. In this equivalent version of the model, the agent perceives a latent AR(1) trend growth rate and she uses the Kalman filter to optimally track this latent trend, while objectively the trend growth rate is constant. The agent uses full memory and the information structure is a filtration and it is Markovian. In the agent's subjective view, past data gradually loses relevance for forecasting not because of fading memory but because it is perceived as irrelevant given the perceived stochastic drift over time in the trend growth rate.

### *D.1. Diffuse prior*

Suppose the agent at time  $t$  perceives the law of motion

$$\Delta c_t = \mu_t + \xi_t, \quad \xi_t \sim \mathcal{N}(0, \sigma_\xi^2), \quad (\text{A.7})$$

$$\mu_{t+1} = \mu_t + \zeta_{t+1}, \quad \zeta_{t+1} \sim \mathcal{N}(0, \sigma_\zeta^2), \quad (\text{A.8})$$

where the agent knows  $\sigma_\xi^2$  and  $\sigma_\zeta^2$ , but not  $\mu_t$ . With diffuse prior and an infinite history,  $H_t$ , of observed data on  $\Delta c$ , the predictive distribution can be obtained from the steady-state Kalman filter (see, e.g., Hamilton (1994)) as

$$\mu_{t+1}|H_t \sim \mathcal{N}(\hat{\mu}_{t+1|t}, \omega^2 + \sigma_\zeta^2), \quad (\text{A.9})$$

where the optimal forecast  $\hat{\mu}_{t+1|t} \equiv \hat{E}(\mu_{t+1}|H_t)$  evolves as

$$\hat{\mu}_{t+1|t} = \hat{\mu}_{t|t-1} + K(\Delta c_t - \hat{\mu}_{t|t-1}), \quad (\text{A.10})$$

with

$$K = \frac{\omega^2 + \sigma_\zeta^2}{\omega^2 + \sigma_\zeta^2 + \sigma_\xi^2}, \quad (\text{A.11})$$

and

$$\omega^2 = K\sigma_\xi^2. \quad (\text{A.12})$$

Thus the predictive distribution of  $\Delta c_{t+1}$  at time  $t$  is

$$\Delta c_{t+1} \sim \mathcal{N}(\hat{\mu}_{t+1|t}, \omega^2 + \sigma_\zeta^2 + \sigma_\xi^2). \quad (\text{A.13})$$

To map into our fading memory setup, we choose

$$K = \nu, \quad \sigma_\xi^2 = (1 - \nu^2)\sigma^2, \quad \sigma_\zeta^2 = \nu^2(1 + \nu)\sigma^2. \quad (\text{A.14})$$

The time- $t$  predictive distribution of  $\Delta c_{t+1}$  then is exactly the same as in our fading memory setting. The time- $t$  predictive distribution of  $\Delta c_{t+j}$  for  $j > 1$  is different from the fading memory setting, though, because here the agent perceives  $\tilde{\mu}_t$  as a martingale and the predictive distribution inherits these martingale dynamics, while in our fading memory setting the

predictive distribution converges to a stationary one at long horizons. For pricing, however, this difference in the perceived distribution for  $j > 1$  does not matter, because under resale valuation, pricing is based on a chain of valuations of one-period ahead payoffs from selling the asset. Thus, pricing in this perceived stochastic trend setting here is the same as in our fading memory setting.

## D.2. Informative prior

Suppose the agent at time  $t$  perceives the law of motion

$$\Delta c_t = \mu_t + \xi_t, \quad \xi_t \sim \mathcal{N}(0, \sigma_\xi^2), \quad (\text{A.15})$$

$$\mu_{t+1} = (1-h)\mu + h\mu_t + \zeta_{t+1}, \quad \zeta_{t+1} \sim \mathcal{N}(0, \sigma_\zeta^2), \quad (\text{A.16})$$

where  $0 \leq h < 1$  and the value of  $h$  is known to the agent. Steady-state Kalman filter updating yields optimal forecasts of the state as

$$\hat{\mu}_{t+1|t} = (1-h)\mu + h\hat{\mu}_{t|t-1} + K(\Delta c_t - \hat{\mu}_{t|t-1}), \quad (\text{A.17})$$

with

$$K = h \frac{\sigma_\zeta^2 + h^2 \omega^2}{\sigma_\zeta^2 + \sigma_\xi^2 + h^2 \omega^2}, \quad (\text{A.18})$$

$$\omega^2 = K \sigma_\xi^2 / h. \quad (\text{A.19})$$

Iterating yields

$$\hat{\mu}_{t+1|t} = \frac{1-h}{1-h+K} \mu + K \sum_{j=0}^{\infty} (h-K)^j \Delta c_{t-j}. \quad (\text{A.20})$$

The predictive distributions are

$$\mu_{t+1}|H_t \sim \mathcal{N}(\hat{\mu}_{t+1|t}, h^2 \omega^2 + \sigma_\zeta^2), \quad (\text{A.21})$$

and

$$\Delta c_{t+1}|H_t \sim \mathcal{N}(\hat{\mu}_{t+1|t}, h^2 \omega^2 + \sigma_\zeta^2 + \sigma_\xi^2). \quad (\text{A.22})$$

We can map this into our fading memory setup with informative prior by choosing

$$K = \phi\nu, \quad h = 1 - \nu + \phi\nu, \quad \sigma_\xi^2 = \frac{1-\nu}{1-\nu+\phi\nu} (1+\phi\nu)\sigma^2 \quad (\text{A.23})$$

to obtain equivalence in terms of the relevant subjective belief dynamics and asset prices.

## E. MODEL SOLUTION FOR $\psi = 1$

### *E.1. SDF*

Following Hansen, Heaton, and Li (2008), we start with value function iteration

$$v_t = \frac{\delta}{1-\gamma} \log \tilde{\mathbb{E}}_t[e^{(1-\gamma)(v_{t+1} + \Delta c_{t+1})}], \quad (\text{A.24})$$

where  $v_t = \log(V_t/C_t)$  and  $V_t$  is the continuation value. We conjecture the solution to be linear in the state variable, i.e.

$$v_t = \mu_v + U_v \tilde{\mu}_t. \quad (\text{A.25})$$

Plugging in the conjectured solution we get

$$U_v = \frac{\delta}{1-\delta}, \quad (\text{A.26})$$

and

$$\mu_v = \frac{1}{2}(1-\gamma)U_v(\nu U_v + 1)^2(1+\nu)\sigma^2. \quad (\text{A.27})$$

We obtain the log SDF

$$\begin{aligned} m_{t+1|t} &= \log \left( \delta \frac{C_t}{C_{t+1}} \frac{(V_{t+1})^{1-\gamma}}{\tilde{\mathbb{E}}_t[(V_{t+1})^{1-\gamma}]} \right) \\ &= \log \delta - \Delta c_{t+1} + (1-\gamma) \log(V_{t+1}) - \log \tilde{\mathbb{E}}_t[(V_{t+1})^{1-\gamma}] \\ &= \log \delta - \Delta c_{t+1} + (1-\gamma)(v_{t+1} + c_{t+1}) - \log \tilde{\mathbb{E}}_t(e^{(1-\gamma)(v_{t+1} + c_{t+1})}) \\ &= \tilde{\mu}_m - \tilde{\mu}_t - \xi \sigma \tilde{\varepsilon}_{t+1}, \end{aligned} \quad (\text{A.28})$$

where

$$\tilde{\mu}_m = \log \delta - \frac{1}{2}(1-\gamma)^2(\nu U_v + 1)^2(1+\nu)\sigma^2, \quad (\text{A.29})$$

$$\xi = [1 - (1-\gamma)(\nu U_v + 1)]\sqrt{1+\nu}. \quad (\text{A.30})$$

### *E.2. Consumption claim valuation*

Let  $\zeta \equiv W_t/C_t$ . The return on the consumption claim is

$$R_{W,t+1} \equiv \frac{W_{t+1}}{W_t - C_t} = \frac{C_{t+1}}{C_t} \frac{\zeta}{\zeta - 1}, \quad (\text{A.31})$$

and in logs,

$$\begin{aligned} r_{w,t+1} &= \Delta c_{t+1} + \log(\zeta/(\zeta - 1)) \\ &= \tilde{\mu}_t + \sqrt{1+\nu}\sigma \tilde{\varepsilon}_{t+1} + \log(\zeta/(\zeta - 1)). \end{aligned} \quad (\text{A.32})$$



Plugging the return on the consumption claim into the Euler equation and taking logs,

$$\begin{aligned}\log(\zeta/(\zeta - 1)) &= -\tilde{\mu}_m + \xi\sqrt{1 + \nu}\sigma^2 - \frac{1}{2}(1 + \nu)\sigma^2 - \frac{1}{2}\sigma^2\xi^2 \\ &= -\log \delta,\end{aligned}\tag{A.33}$$

which we can solve for the wealth-consumption ratio

$$\zeta = \frac{1}{1 - \delta}.\tag{A.34}$$

That the consumption-wealth ratio is constant can also be seen by valuing consumption strips. Denoting with  $w_t^1$  the log of the component of time- $t$  wealth that derives from the one-period ahead endowment flow, we have

$$\begin{aligned}w_t^1 - c_t &= \log \tilde{\mathbb{E}}_t \left[ M_{t+1|t} \frac{C_{t+1}}{C_t} \right] \\ &= \log \tilde{\mathbb{E}}_t \left[ \exp(\tilde{\mu}_m + (\sqrt{1 + \nu} - \xi)\sigma\tilde{\varepsilon}_{t+1}) \right] \\ &= \tilde{\mu}_m + \frac{1}{2}(\sqrt{1 + \nu} - \xi)^2\sigma^2,\end{aligned}\tag{A.35}$$

i.e.,  $w_t^1 - c_t$  is constant. It does not vary with  $\tilde{\mu}_t$  because, going from the first to the second line,  $-\tilde{\mu}_t$  in  $m_{t+1|t}$  cancels with  $\tilde{\mu}_t$  in  $\Delta c_{t+1}$ . Working through the valuation equation backwards in time, we obtain the price of an  $n$ -period consumption strip

$$w_t^n - c_t = n\tilde{\mu}_m + \frac{n}{2}(\sqrt{1 + \nu} - \xi)^2\sigma^2.\tag{A.36}$$

Plugging in the solutions for  $\tilde{\mu}_m$  and  $\xi$  from the previous subsection, we get

$$w_t^n - c_t = n \log \delta.\tag{A.37}$$

Summing the value of consumption strips at all horizons strips yields the consumption-wealth ratio in (A.34).

### *E.3. Dividend strip valuation*

By analyzing dividend strips that are claims to single dividends in the future, we can transparently analyze the conditions needed for a finite price. The price of the  $n$ -period dividend strip is

$$P_t^n \equiv \tilde{\mathbb{E}}_t[M_{t+1|t}\tilde{\mathbb{E}}_{t+1}[\cdots\tilde{\mathbb{E}}_{t+n-1}[M_{t+n|t+n-1}D_{t+n}]]].\tag{A.38}$$

As we discussed earlier, when we evaluate these expectations, we do so by iterating backwards from the payoff at  $t + n$ , evaluating one conditional expectation at a time without relying on the Law of Iterated Expectations (LIE).

Taking logs and evaluating (A.38), we obtain

$$p_t^n - d_t = [1 - (1 - \alpha)^n] (c_t - d_t + \mu_{dc} + \frac{\lambda - 1}{\alpha} \tilde{\mu}_t) + n\tilde{\mu}_m + \frac{1}{2}(A_n\sigma^2 + B_n\sigma_d^2), \quad (\text{A.39})$$

where

$$A_n = \sum_{k=0}^{n-1} \left\{ \sqrt{1 + \nu} \left[ \nu(\lambda - 1) \frac{1 - (1 - \alpha)^k}{\alpha} + (\lambda - 1)(1 - \alpha)^k + 1 \right] - \xi \right\}^2, \quad (\text{A.40})$$

and

$$B_n = \frac{1 - (1 - \alpha)^{2n}}{1 - (1 - \alpha)^2}. \quad (\text{A.41})$$

For very large  $n$ , approximately,

$$A_n \approx n \left[ \sqrt{1 + \nu} \left( 1 + \nu \frac{\lambda - 1}{\alpha} \right) - \xi \right]^2, \quad (\text{A.42})$$

and  $B_n$ , which does not grow with  $n$ , becomes very small relative to  $A_n$ . Thus for the price to be well-defined, we need the terms that grow with  $n$  in (A.39) to be (weakly) negative. Using (34), we see that this requires

$$\tilde{\mu}_m + \frac{1}{2} \left[ \sqrt{1 + \nu} \left( 1 + \nu \frac{\lambda - 1}{\alpha} \right) - \xi \right]^2 \sigma^2 \leq 0. \quad (\text{A.43})$$

In our calibration, we will work with a value for  $\alpha$  that satisfies this condition.

As equation (33) shows, the dividend strip log price-dividend ratio is increasing in  $\tilde{\mu}_t$  if  $\lambda > 1$ , i.e., if the dividend claim is levered. With  $\lambda = 1$  the effect of a higher expected growth rate of dividends would be offset by a higher risk-free rate, just like it is for the consumption claim. With leverage, the effect of higher expected dividend growth is stronger than the risk-free rate effect. Since dividends and consumption are co-integrated and hence have the same growth rate in the long-run, the reason why the agent can expect the dividend and consumption growth rates to differ for a substantial period of time may not be immediately obvious. When the agent revises upward her posterior mean  $\tilde{\mu}_t$ , then her expectation of the dividend growth rate in the next period gets revised upward by  $\lambda - 1$  times the revision in  $\tilde{\mu}_t$ . She expects that over the near future dividend growth will exceed consumption growth, leading to a rise in  $d - c$ . Eventually, the higher log dividend-consumption ratio will—through the cointegration relationship in (31)—generate enough negative offset to bring the dividend growth rate back down to  $\tilde{\mu}_t$ . Thus, she expects the process to settle down with similar mean growth rates, but at a higher  $d - c$ .<sup>17</sup> Is it economically plausible that the agent perceives higher unconditional mean economic growth to be associated with a higher dividend-consumption ratio? It is impossible to answer this question within a model with

17. A similar mechanism is at work in Collin-Dufresne, Johannes, and Lochstoer (2017), but there dividends and consumption are not cointegrated, and so  $d - c$  can grow without bound, but the unconditional mean growth rate of consumption (and hence dividends) is truncated.

exogenous endowment flow. In Appendix G we present calculations based on a Ramsey-Cass-Koopmans model with endogenous investment and production that suggests that a positive relationship between growth and the capital income to consumption ratio is indeed plausible.

To get the expected returns of dividend strips, we start with (33) to compute returns. For the one-period claim, we get

$$r_{t+1}^1 = \lambda \Delta c_{t+1} - (\lambda - 1) \tilde{\mu}_t - \tilde{\mu}_m - \frac{1}{2} (\sqrt{1 + \nu \lambda} - \xi)^2 \sigma^2. \quad (\text{A.44})$$

Subtracting  $r_{f,t} = -\tilde{\mu}_m + \tilde{\mu}_t - \frac{1}{2} \xi^2 \sigma^2$  yields

$$r_{t+1}^1 - r_{f,t} = \lambda (\Delta c_{t+1} - \tilde{\mu}_t) - \frac{1}{2} (\sqrt{1 + \nu \lambda} - \xi)^2 \sigma^2 + \frac{1}{2} \xi^2 \sigma^2. \quad (\text{A.45})$$

The subjective conditional variance of  $r_{t+1}^1$  is  $(1 + \nu) \lambda^2 \sigma^2$ , and so, after taking subjective expectations of (A.45), we obtain

$$\log \tilde{\mathbb{E}}_t [R_{t+1}^1] - r_{f,t} = \lambda \xi \sqrt{1 + \nu} \sigma^2. \quad (\text{A.46})$$

The objective conditional variance of  $r_{t+1}^1$  is only  $\lambda^2 \sigma^2$ , and so taking objective expectations of (A.45) yields,

$$\log \mathbb{E}_t [R_{t+1}^1] - r_{f,t} = \lambda \xi \sqrt{1 + \nu} \sigma^2 - \frac{1}{2} \nu \lambda^2 \sigma^2 + \lambda (\mu - \tilde{\mu}_t). \quad (\text{A.47})$$

For the infinite-horizon claim, again starting from (33), we get

$$r_{t+1}^\infty = \Delta c_{t+1} + \frac{\lambda - 1}{\alpha} (\tilde{\mu}_{t+1} - \tilde{\mu}_t) - \tilde{\mu}_m - \frac{1}{2} \left[ \sqrt{1 + \nu} \left( 1 + \nu \frac{\lambda - 1}{\alpha} \right) - \xi \right]^2 \sigma^2, \quad (\text{A.48})$$

and, after subtracting the risk-free rate,

$$r_{t+1}^\infty - r_{f,t} = \Delta c_{t+1} + \frac{\lambda - 1}{\alpha} (\tilde{\mu}_{t+1} - \tilde{\mu}_t) - \tilde{\mu}_t - \frac{1}{2} \left[ \sqrt{1 + \nu} \left( 1 + \nu \frac{\lambda - 1}{\alpha} \right) - \xi \right]^2 \sigma^2 + \frac{1}{2} \xi^2 \sigma^2. \quad (\text{A.49})$$

The subjective conditional variance of  $r_{t+1}^\infty$  is  $(1 + \nu) \left( 1 + \nu \frac{\lambda - 1}{\alpha} \right)^2 \sigma^2$  and therefore, after taking subjective expectations of (A.49), we obtain

$$\log \tilde{\mathbb{E}}_t [R_{t+1}^\infty] - r_{f,t} = \left[ 1 + \nu \frac{\lambda - 1}{\alpha} \right] \xi \sqrt{1 + \nu} \sigma^2. \quad (\text{A.50})$$

The objective conditional variance of  $r_{t+1}^\infty$  is only  $\left( 1 + \nu \frac{\lambda - 1}{\alpha} \right)^2 \sigma^2$ , and so taking objective expectations of (A.49) yields

$$\log \mathbb{E}_t [R_{t+1}^\infty] - r_{f,t} = \left[ 1 + \nu \frac{\lambda - 1}{\alpha} \right] \xi \sqrt{1 + \nu} \sigma^2 - \frac{1}{2} \nu \left( 1 + \nu \frac{\lambda - 1}{\alpha} \right)^2 \sigma^2 + \left( 1 + \nu \frac{\lambda - 1}{\alpha} \right) (\mu - \tilde{\mu}_t). \quad (\text{A.51})$$

#### E.4. Numerical solution

We use the analytical solutions for dividend strip prices to numerically compute the price,  $P_t$ , of the equity claim to the whole stream of dividends. For  $n > J$  and some big enough  $J$ , equation (33) implies that

$$P_t^n \approx C_t e^{\mu_{dc} + \frac{1}{2} \frac{1}{1-(1-\alpha)^2} \sigma_d^2 + \frac{\lambda-1}{\alpha} \tilde{\mu}_t} \exp(n\tilde{\mu}_m + \frac{1}{2} A_n \sigma^2), \quad n > J, \quad (\text{A.52})$$

where we approximate

$$A_n \approx A_J + (n - J) [\sqrt{1 + \nu} (\nu \frac{\lambda-1}{\alpha} + 1) - \xi]^2, \quad n > J. \quad (\text{A.53})$$

We can show that

$$P_t \approx \left( \sum_{n=1}^J P_t^n \right) + C_t V_J \exp \left( \mu_{dc} + \frac{1}{2} \frac{1}{1-(1-\alpha)^2} \sigma_d^2 + \frac{\lambda-1}{\alpha} \tilde{\mu}_t \right), \quad (\text{A.54})$$

with

$$V_J = \frac{\exp \left( (J+1)\tilde{\mu}_m + \frac{1}{2} A_J \sigma^2 + \frac{1}{2} [\sqrt{1 + \nu} (\nu \frac{\lambda-1}{\alpha} + 1) - \xi]^2 \sigma^2 \right)}{1 - \exp \left( \tilde{\mu}_m + \frac{1}{2} [\sqrt{1 + \nu} (\nu \frac{\lambda-1}{\alpha} + 1) - \xi]^2 \sigma^2 \right)}. \quad (\text{A.55})$$

We implement this by choosing a  $J$  big enough so that the value of  $P_t$  we obtain is not sensitive anymore to further changes in  $J$ . In our calibration, this requires  $J \approx 7,000$ .

We further use numerical methods to solve for the subjective equity premium in the  $\psi = 1$  case. We follow the approach of Pohl, Schmedders, and Wilms (2018). When  $\psi = 1$ , the wealth-consumption ratio is a constant

$$\log \frac{W_t - C_t}{C_t} = \log \frac{\delta}{1 - \delta}, \quad (\text{A.56})$$

and we only need to solve for the log price-dividend ratio. The log P/D ratio should be a function of both  $\tilde{\mu}$  and  $d_t - c_t$ , i.e.

$$\log \frac{P_t}{D_t} = H(\tilde{\mu}_t, d_t - c_t). \quad (\text{A.57})$$

In this case, because there are two state variables, the basis functions are now

$$\psi_{ij}(\tilde{\mu}, d_t - c_t) \equiv \Lambda_i(\tilde{\mu}) \Lambda_j(d_t - c_t), \quad (\text{A.58})$$

where  $\Lambda_i$  denotes the Chebyshev polynomials. We will approximate the log P/D ratio as

$$\hat{H}(\tilde{\mu}, d_t - c_t; \beta_m) = \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} \beta_{m,ij} \psi_{ij}(\tilde{\mu}, d_t - c_t). \quad (\text{A.59})$$

Rewrite the subjective Euler equation

$$\tilde{\mathbb{E}}_t[M_{t+1}R_{m,t+1}] = 1 \quad (\text{A.60})$$

as

$$\begin{aligned} 0 &= I(\tilde{\mu}_t, d_t - c_t) \\ &\equiv \tilde{\mathbb{E}}_t \left[ e^{\tilde{\mu}_m - \tilde{\mu}_t - \xi \sigma \epsilon_{t+1} + \Delta d_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}} \right] - 1 \\ &= e^{\tilde{\mu}_m + (\lambda - 1)\tilde{\mu}_t - \alpha(d_t - c_t - \mu_{dc})} \tilde{\mathbb{E}}_t \left[ e^{(\lambda\sqrt{1+\nu} - \xi)\sigma\tilde{\epsilon}_{t+1} + \sigma_d\eta_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}} \right] - 1. \end{aligned} \quad (\text{A.61})$$

We evaluate the function  $I(\tilde{\mu}_t, d_t - c_t)$  on the two-dimensional grid of  $\tilde{\mu}_t$  and  $d_t - c_t$  and use the two-dimensional Gaussian quadrature approach to calculate the expectation part as an integral. Following Pohl, Schmedders, and Wilms (2018), the numerical solution is implemented by the “fmincon” solver with the SQP algorithm in Matlab. We minimize a constant subject to the nonlinear constraints implied by Equation (A.61). We choose the degree of approximation, i.e.,  $n_1$  and  $n_2$ , such that the log P/D ratio computed using the projection method is closest to the analytically computed log P/D ratio as in Equation (A.54) in terms of the RMSE,

$$RMSE_{pd} = \sqrt{\frac{1}{t} \sum_{j=1}^t (pd_j^{Analytical} - pd_j^{Projection})^2}, \quad (\text{A.62})$$

where  $pd_t^{Analytical}$  is calculated from dividend strip prices as in (A.54). We explore different combinations of  $n_1$  and  $n_2$  up to a maximal degree of 8 and we choose the combination that minimizes  $RMSE_{pd}$ . Table A.I summarizes the parameter choices for this numerical procedure.

After we obtain the coefficients for  $H(\tilde{\mu}_t, d_t - c_t)$ , we can calculate the subjective equity return as

$$\begin{aligned} \tilde{\mathbb{E}}_t[R_{m,t+1}] &= \tilde{\mathbb{E}}_t \left[ e^{\Delta d_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}} \right] \\ &= e^{\lambda\tilde{\mu}_t - \alpha(d_t - c_t - \mu_{dc})} \tilde{\mathbb{E}}_t \left[ e^{\lambda\sqrt{1+\nu}\sigma\tilde{\epsilon}_{t+1} + \sigma_d\eta_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}} \right]. \end{aligned} \quad (\text{A.63})$$

## F. MODEL SOLUTION FOR $\psi \neq 1$

### F.1. Existence

Hansen and Scheinkman (2012) provide sufficient conditions for existence of equilibrium in a Markovian setting with Epstein-Zin preferences. As we show in Appendix D, our fading memory model can be mapped into an equivalent full memory model in which the informa-

TABLE A.I  
PARAMETER VALUES FOR PROJECTION METHOD WITH  $\psi = 1$

Parameter	Symbol	Value
Grid for $\tilde{\mu}_t$	$[\mu - 4\sigma(\tilde{\mu}), \mu + 4\sigma(\tilde{\mu})]$	[0.00 , 0.01]
Grid for $d_t - c_t$	$[E(dc) - 4\sigma(dc), E(dc) + 4\sigma(dc)]$	[-5.5 , -3.9]
# of points for Gaussian-Quad sampling	v	200
Degrees of approximation for $\tilde{\mu}_t$	$n_1$	3
Degrees of approximation for $d_t - c_t$	$n_2$	3
$RMSE_{pd}$ across simulations	Mean	0.008
	Median	0.008
	Max	0.038

tion structure is a filtration and Markovian. This allows us to use the results in Hansen and Scheinkman (2012) to derive parameter restrictions sufficient to ensure existence of equilibrium.

In our equivalent “Kalman filtering” economy, we only have one state variable, which is  $\{\hat{\mu}_{t+1|t}\}$  as in Equation (A.17). This equation also shows that  $\Delta c_{t+1}$  can be written as a function of  $\hat{\mu}_{t+2|t+1}$  and  $\hat{\mu}_{t+1|t}$ . In addition, given the Markov property of  $\{\hat{\mu}_{t+1|t}\}$ , Assumption 1 in Hansen and Scheinkman (2012) is satisfied.

To distinguish the perceived time- $t$  predictive distribution for  $\Delta c_{t+j}$  in this “Kalman filtering” economy here from the fading memory economy, we denote the subjective expectation here as  $\tilde{\mathbb{E}}^*$ . The Perron-Frobenius eigenvalue equation of interest is

$$\mathcal{T}v(x) = \exp(\eta)v(x), \quad v(\cdot) > 0, \quad (\text{A.64})$$

where

$$\mathcal{T}f(x) = \tilde{\mathbb{E}}^* [f(\hat{\mu}_{t+2|t+1}) \exp[(1 - \gamma)\Delta c_{t+1}] | \hat{\mu}_{t+1|t} = x]. \quad (\text{A.65})$$

Another random variable of interest is

$$N_{t+1} = \frac{e^{(1-\gamma)\Delta c_{t+1}} v(\hat{\mu}_{t+2|t+1})}{\exp(\eta)v(\hat{\mu}_{t+1|t})}. \quad (\text{A.66})$$

Hansen and Scheinkman (2012) show that solutions exist for our model if the following additional assumptions are met:

**Assumption 1.**

$$\log \delta + \frac{\eta}{\theta} < 0. \quad (\text{A.67})$$

**Assumption 2.**

$$\lim_{t \rightarrow \infty} \tilde{\mathbb{E}}^* [N_{t+1} v(\hat{\mu}_{t+2|t+1})^{-1/\theta} | \hat{\mu}_{1|0} = x] < \infty. \quad (\text{A.68})$$

**Assumption 3.**

$$\lim_{t \rightarrow \infty} \tilde{\mathbb{E}}^* [N_{t+1} v(\hat{\mu}_{t+2|t+1})^{-1} | \hat{\mu}_{1|0} = x] < \infty. \quad (\text{A.69})$$

To derive the explicit expressions of constraints in our model, we first calculate the Perron-Frobenius eigenvalue function  $v$ . We can show that (one of) the solution is

$$v(x) = \exp\left(\frac{1-\gamma}{\nu-\phi\nu}x\right), \quad (\text{A.70})$$

$$\eta = (1-\gamma)\mu + \frac{1}{2}\left(\frac{1-\gamma}{1-\phi}\right)^2(1+\phi\nu)\sigma^2. \quad (\text{A.71})$$

With some algebra, both Assumption 2 and Assumption 3 can be reduced to the form

$$\lim_{t \rightarrow \infty} \tilde{\mathbb{E}}^* \left[ \exp(k_1 \Delta c_{t+1} + k_2 \hat{\mu}_{t+1|t}) \middle| \hat{\mu}_{1|0} = x \right] < \infty, \quad (\text{A.72})$$

for some corresponding pairs of constants  $(k_1, k_2)$ .

With the Wold representation, we have

$$\hat{\mu}_{t+1|t} = [1 - (h - K)L]^{-1} K \Delta c_t, \quad (\text{A.73})$$

$$\Delta c_{t+1} = [1 - (h - K)L]^{-1} K \Delta c_t + \tau_{t+1}, \quad (\text{A.74})$$

$$\Delta c_t = [1 + (1 - hL)^{-1} KL] \tau_t, \quad (\text{A.75})$$

where  $\{\tau_{t-j}\}$  are uncorrelated with variance  $(1 + \phi\nu)\sigma^2$ . It suffices to show that

$$\lim_{t \rightarrow \infty} \tilde{\mathbb{E}}^* \left[ \exp(k [1 - (h - K)L]^{-1} K [1 + (1 - hL)^{-1} KL] \tau_t) \right] < \infty, \quad (\text{A.76})$$

or

$$\lim_{t \rightarrow \infty} \tilde{\mathbb{E}}^* \left[ \exp(kK(1 - hL)^{-1} \tau_t) \right] < \infty. \quad (\text{A.77})$$

As long as  $h < 1$ , we have

$$\lim_{t \rightarrow \infty} \tilde{\mathbb{E}}^* \left[ \exp(kK(1 - hL)^{-1} \tau_t) \right] \quad (\text{A.78})$$

$$= \lim_{t \rightarrow \infty} \tilde{\mathbb{E}}^* \left[ \exp\left(kK \sum_{j=0}^{\infty} h^j L^j \tau_t\right) \right] \quad (\text{A.79})$$

$$= \exp\left(\frac{1}{2}k^2 K^2 \sum_{j=0}^{\infty} h^{2j} (1 + \phi\nu)\sigma^2\right) < \infty. \quad (\text{A.80})$$

Finally, Assumption 1 translates to

$$\log \delta + \left(1 - \frac{1}{\psi}\right) \left[ \mu + \frac{1}{2} \frac{1-\gamma}{(1-\phi)^2} (1 + \phi\nu)\sigma^2 \right] < 0, \quad (\text{A.81})$$

and this is the only parameter constraint we apply to our model to ensure existence of

equilibrium.

### F.2. Log-linearized solution

We solve the model for  $\psi \neq 1$  using log-linearization along similar lines as, e.g., in Beeler and Campbell (2012). We can write the Epstein-Zin log SDF as

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1)r_{w,t+1}, \quad (\text{A.82})$$

with  $\theta = \frac{1-\gamma}{1-1/\psi}$ .

We log-linearize the return on wealth and the return on the equity claim as

$$r_{w,t+1} = k_0 + k_1 z_{t+1} - z_t + \Delta c_{t+1}, \quad (\text{A.83})$$

$$r_{m,t+1} = k_{d,0} + k_{d,1} z_{d,t+1} - z_{d,t} + \Delta d_{t+1}, \quad (\text{A.84})$$

where  $z_t \equiv \log((W_t - C_t)/C_t)$  and  $z_{d,t} \equiv \log(P_t/D_t)$ . We then conjecture  $z_t$  and  $z_{d,t}$  are linear in the state variables

$$z_t = A_0 + A_1 \phi \tilde{\mu}_t, \quad (\text{A.85})$$

$$z_{d,t} = A_{d,0} + A_{d,1} \phi \tilde{\mu}_t + A_{d,2}(d_t - c_t). \quad (\text{A.86})$$

By applying the subjective pricing equation to both  $r_{w,t+1}$  and  $r_{m,t+1}$ , we can show that

$$A_0 = \frac{\log \delta + k_0 + (1 - \phi)(1 - k_1 + \nu k_1)\mu A_1 + \frac{1}{2}(1 - 1/\psi + \phi \nu k_1 A_1)^2 \theta (1 + \phi \nu) \sigma^2}{1 - k_1}, \quad (\text{A.87})$$

$$A_1 = \frac{1 - 1/\psi}{1 - (1 - \nu + \phi \nu)k_1}, \quad (\text{A.88})$$

and

$$A_{d,0} = \frac{\theta \log \delta + (\theta - 1)k_0 + (\theta - 1)(k_1 - 1)A_0 + k_{d,0} + (1 - \phi)(1 - k_{d,1} + \nu k_{d,1})\mu A_{d,1}}{1 - k_{d,1}} \quad (\text{A.89})$$

$$+ \frac{(\theta - 1)k_1 \nu (1 - \phi)\mu A_1 + (k_{d,1} A_{d,2} + 1)\alpha \mu_{dc} + \frac{1}{2}(k_{d,1} A_{d,2} + 1)^2 \sigma_d^2}{1 - k_{d,1}} \quad (\text{A.90})$$

$$+ \frac{\frac{1}{2}[\theta - 1 - \theta/\psi + \lambda + (\lambda - 1)k_{d,1} A_{d,2} + \phi \nu (k_{d,1} A_{d,1} + (\theta - 1)k_1 A_1)]^2 (1 + \phi \nu) \sigma^2}{1 - k_{d,1}}, \quad (\text{A.91})$$

$$A_{d,1} = \frac{\lambda - 1/\psi + (\lambda - 1)k_{d,1} A_{d,2}}{1 - (1 - \nu + \phi \nu)k_{d,1}}, \quad (\text{A.92})$$

$$A_{d,2} = \frac{\alpha}{(1 - \alpha)k_{d,1} - 1}. \quad (\text{A.93})$$



Applying the subjective pricing equation to the risk-free payoff, we obtain

$$r_{f,t} = -\log \delta + \frac{1}{\psi} [\phi \tilde{\mu}_t + (1 - \phi)\mu] \quad (\text{A.94})$$

$$+ \frac{1}{2} (1 - \gamma) \left( \frac{1}{\psi} - \gamma \right) \left( 1 + \frac{k_1 \phi \nu}{1 - (1 - \nu + \phi \nu) k_1} \right)^2 (1 + \phi \nu) \sigma^2 \quad (\text{A.95})$$

$$- \frac{1}{2} \left[ \left( \frac{1}{\psi} - \gamma \right) \left( 1 + \frac{k_1 \phi \nu}{1 - (1 - \nu + \phi \nu) k_1} \right) - \frac{1}{\psi} \right]^2 (1 + \phi \nu) \sigma^2. \quad (\text{A.96})$$

We solve for the log-linearization coefficients by iterating on

$$\bar{z} = A_0 + A_1 \phi \mu, \quad (\text{A.97})$$

$$k_1 = \frac{\exp(\bar{z})}{1 + \exp(\bar{z})}, \quad (\text{A.98})$$

$$k_0 = \log(1 + e^{\bar{z}}) - \bar{z} k_1, \quad (\text{A.99})$$

and

$$\bar{z}_d = A_{d,0} + A_{d,1} \phi \mu + A_{d,2} \mathbb{E}[d_t - c_t], \quad (\text{A.100})$$

$$k_{d,1} = \frac{\exp(\bar{z}_d)}{1 + \exp(\bar{z}_d)}, \quad (\text{A.101})$$

$$k_{d,0} = \log(1 + e^{\bar{z}_d}) - \bar{z}_d k_{d,1}, \quad (\text{A.102})$$

until we reach fixed points of  $k_0$ ,  $k_1$ ,  $k_{d,0}$ , and  $k_{d,1}$ , determined by a difference of less than  $10^{-6}$ .

We then simulate the model and construct  $\tilde{\mu}_t$ . To calculate the objective risk premium, we directly use the return on equity claim as in (A.84) and the risk-free rate as in (A.96). To calculate subjective expected return, we take the subjective expectation of Equation (A.84) to yield

$$\tilde{\mathbb{E}}_t[r_{m,t+1}] = \frac{\phi}{\psi} \tilde{\mu}_t + \tilde{A}_{d,0}, \quad (\text{A.103})$$

where

$$\tilde{A}_{d,0} = (k_{d,1} - 1)A_{d,0} + k_{d,0} + (k_{d,1}A_{d,2} + 1)\alpha\mu_{dc} \quad (\text{A.104})$$

$$+ (1 - \phi)\mu [k_{d,1}A_{d,1}\phi\nu + (\lambda - 1)k_{d,1}A_{d,2} + \lambda]. \quad (\text{A.105})$$

As a result, with log-linearization, the model implies a constant subjective premium.

## G. CAPITAL INCOME TO CONSUMPTION RATIO IN RAMSEY-CASS-KOOPMANS MODEL

To illustrate plausible properties of the dividend-consumption ratio in a model in which investment and production is endogenous, this section presents a calculation of the capital

income to consumption ratio in the Ramsey-Cass-Koopmans model with Cobb-Douglas technology based on Barro and Sala-i-Martin (2004), Chapter 2. For ease of comparison, we use their notations here: IES  $1/\theta$ , interest rate  $r$ , capital and consumption in efficiency units  $\hat{k}$  and  $\hat{c}$ , productivity growth rate  $x$ , population growth rate  $n$ , time discount rate  $\rho$ , capital share  $\alpha$  and depreciation rate  $\delta$ . Consistent with our baseline calibration, we set  $1/\theta = 1$ .

Solving their equations (2.24) and (2.25) for  $\hat{c}$  and  $\hat{k}$  with the left-hand side equal to zero in the steady state, and using the property of Cobb-Douglas technology that  $f'(\hat{k})/\alpha = f(\hat{k})/\hat{k}$ , we use this solution to calculate the ratio of capital income to consumption as

$$r\hat{k} = (x + \rho\hat{k})/\hat{c}. \quad (\text{A.106})$$

Taking the derivative with respect to the productivity growth rate  $x$ , we obtain a positive derivative if

$$n - \rho < \left( \frac{1}{\alpha} - 1 \right) \delta. \quad (\text{A.107})$$

With an annual population growth rate of  $n = 0.01$ , and  $\rho$  slightly bigger than 1% as in our calibration, the left-hand side is negative and so (since  $0 \leq \alpha \leq 1$  and  $\delta > 0$ ) this inequality always holds. With typical values of  $\delta = 0.05$  and  $\alpha = 0.4$ , the inequality holds unless population growth rates are implausibly high (more than 8% for  $\rho = 0.01$ ) or the time discount rate implausibly low.