Asset Pricing with Fading Memory *

Stefan Nagel[†]
University of Chicago,
NBER, CEPR, and CESIfo

Zhengyang Xu[‡]
University of Michigan

October 2018

Building on recent evidence that lifetime experiences shape individuals' macroeconomic expectations, we study asset prices in an economy in which a representative agent learns with fading memory from experienced endowment growth. The agent updates subjective beliefs with constant gain, which induces memory loss, but is otherwise Bayesian in evaluating uncertainty. The model produces perpetual learning, substantial priced long-run growth rate uncertainty, and, conveniently, a stationary economy. This approach resolves many asset pricing puzzles and it reconciles model-implied subjective belief dynamics with survey data on individual investor return expectations within a simple setting with IID endowment growth, constant risk aversion, and a gain parameter calibrated to microdata estimates. The *objective* equity premium is high and strongly counter-cyclical in the sense of being negatively related to experienced stock market payout growth (a long-run weighted average of past growth rates). In contrast, the *subjective* equity premium is slightly pro-cyclical. As a consequence, subjective expectations errors are predictable and negatively related to past experienced payout growth. Consistent with this theory, we show empirically that experienced payout growth is negatively related to future stock market excess returns. Based on expectations data from individual investor surveys spanning several decades, we show that this measure of experienced growth is also strongly negatively related to subjective expectations errors.

[†]University of Chicago, Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, e-mail: stefan.nagel@chicagobooth.edu

[‡]Ross School of Business, University of Michigan, 701 Tappan St., Ann Arbor, MI 48109, e-mail: zhengyxu@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 countercyclical. If anything, the survey data indicate pro-cyclicality (Vissing-Jorgensen 2003; Amromin and Sharpe 2013; Greenwood and Shleifer 2014).

We show that the behavior of asset prices and survey data 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. The decay of the agent's memory of observations in the distant past is the only modification to an otherwise standard Bayesian 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 of an econometrician examining a large sample of data generated by this economy. For example, after a string of positive growth innovations, the agent is subjectively optimistic about the mean growth rate, the equity price

1

is high, and subsequent returns are low because the agent's expectations are disappointed ex post. Thus, objectively, the equity premium is counter-cyclical, but subjectively it is not. Since fading memory limits the precision of the agent's growth rate estimates, subjective growth rate uncertainty is high, which generates a high unconditional equity premium.

Our fading memory approach is motivated by micro-evidence on household portfolio choice and survey data on expectations 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).[1] Collin-Dufresne, Johannes, and Lochstoer (2017) use an overlapping dynasties approach to introduce such learning from experience and the resulting generational heterogeneity into an asset pricing model. However, even with a starkly simplified demographic structure with only two overlapping dynasties, model solution becomes quite difficult. Moreover, in the two-dynasties setting, risk premiums and risk-free rate in their model jump every 20 years when there is a generational shift. This makes it somewhat difficult to map the model to empirical data.

To obtain a model that is as tractable as the leading rational expectations asset-pricing models in the literature, we abstract from generational heterogeneity. We build on the insight in Malmendier and Nagel (2016) that the dynamics of the *average* individuals' expectation can be approximated very closely by a constant-gain learning scheme. While individual agents learn from experience with decreasing gain—i.e., as individuals age and more data accumulates in their experience set, the sensitivity of expectations to a new piece of information declines—generational turnover implies that older agents with low gain are continuously replaced by younger agents with high gain and, as a consequence, the average of individuals' expectations updates with constant gain. Constant-gain updating implies that an observation's influence on current beliefs gradually fades as it recedes into the past. In our model, we may miss interesting implications for the distribution of wealth and risk-bearing across

---

1. An growing body of evidence also suggest 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.

generations, but risk premiums and risk-free rate move gradually from quarter to quarter in an empirically realistic fashion.

The key parameter in our model is a gain parameter that determines by how much the agent updates beliefs in response to an observed growth rate innovation and how fast memory decays. The volatility and persistence of the price-dividend ratio and the strength of return predictability are strongly influenced by this parameter. 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.

In our model, equity is a levered claim to the endowment stream. Since the value of the equity claim is convex in the growth rate of the equity payoffs, the agent's uncertainty about the unconditional mean growth rate can lead the value of the equity claim to explode. In our baseline model, we avoid this unrealistic outcome by assuming an elasticity of intertemporal substitution (EIS) of one, which implies a constant consumption-wealth ratio, and by assuming that dividends are cointegrated with consumption. We choose the cointegration parameter to get a realistic equity volatility. The remaining model parameters are similar to common choices in the literature. In an extension, we allow for an EIS greater than one, combined with a slightly informative prior that tilts the posterior mean growth rate weakly towards its true mean.[2]

We obtain a high equity premium, excess returns that are predictable by the price-dividend ratio, and a low volatility of the risk-free rate—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

---

2. Alternative methods to deal with this problem include 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).

and Cochrane (1999) (CC).

The state variable that drives most of the variation in the price-dividend ratio and objective expected returns in this model is a slow-moving exponentially weighted average of past endowment growth rates, which we label *experienced growth*. We construct the empirical counterpart with data on experienced dividend growth rates, i.e., a levered version of endowment growth rates, back to the 19th century. Since a shift from dividend payments to repurchases in the later part of the sample distorts the time-series properties of the dividend process, we follow Bansal, Dittmar, and Lundblad (2005) and work with a stock market payout series that includes dividends and repurchases. As an alternative measure, we also construct a series of experienced real returns—an exponentially weighted average of past real stock returns—that should be, according to our model, strongly correlated with experienced growth.

Empirically, experienced payout growth and experienced real stock returns are both strong predictors of excess stock returns. From the agent's subjective viewpoint, the world looks different. Subjective expected excess returns are not negatively related to experienced growth. In fact, since the price of the equity claim is convex in growth rates, the agent perceives equity as somewhat riskier when subjective expected endowment growth and the equity price are high. As a consequence, subjective expected excess returns are slightly positively related to experienced growth. We show that the same is true for subjective expected excess returns in survey data. This wedge between subjective and objective expectations generates a strong negative relationship between subjective expectations errors and experienced payout growth and experienced real returns—which we also find in the empirical data.

While an econometrician can find predictable expectations errors in samples generated from this economy, it would be difficult for the agent to detect any error in real time, even with full memory of existing return history. Standard out-of-sample predictability tests show no out-of-sample return predictability for empirically realistic sample sizes, even though returns are truly predictable under the objective distribution.

4

The time-varying objective equity premium in our model is largely decoupled from conditional equity return volatility. In times of high experienced growth, high price-dividend ratio (and hence low objective equity premium) conditional return volatility is slightly elevated, but only very weakly so. On this dimension, the model's prediction are very different from those of Campbell and Cochrane (1999) and Bansal and Yaron (2004), which both imply that equity valuations are strongly negatively related to future return volatility. In CC, this happens because risk aversion is also very volatile in times when it is high. In BY, the reason is that time-varying endowment growth risk is the driver of the time-varying equity premium. Empirically, however, market return volatility is weakly positively related to lagged experienced growth and equity valuation levels, consistent with the predictions of our model.[3]

The model further makes predictions about the cyclicality of the equity term structure that are in line with the data. Gormsen (2018) shows empirically that the equity term premium is strongly counter-cyclical and negatively related to the slope of the equity yield curve. Unlike the CC and BY models, which—as Gormsen (2018) demonstrates—match only the first of these two empirical facts, our model matches both.

Our model shares many features with full-memory Bayesian parameter learning models, but the fading memory feature avoids the arguably unrealistic implication of these models that learning effects disappear deterministically over time as the agent acquires more data.[4] In our model, learning is perpetual and at every point in time, the agent uses a sample of the same effective size to form expectations about endowment growth. As a consequence, return predictability persists and the economy is conveniently stationary. Our model's predictions are also very different from an alternative model in which the representative agent uses optimal learning to track a drifting growth-rate parameter. For example, if endowment

---

3. As is well known, market crashes are associated with strongly rising conditional volatility. But these effects operate at higher frequencies than those captured by slow-moving variables like the price-dividend ratio or our experienced dividend growth variable.

4. The unknown-mean model of Collin-Dufresne, Johannes, and Lochstoer (2016) shares similarities with ours, but in their case, the agent learns with decreasing gain, return predictability disappears over time, and the economy is nonstationary. See, also, Timmermann (1993) and Lewellen and Shanken (2002) for partial equilibrium models with decreasing-gain learning.

growth had a random walk component, constant-gain learning would actually be optimal, even if full memory is available. However, even though the agent would then use the same updating scheme to form growth-rate expectations as the agent in our model, the model would not produce predictability of returns and subjective expectations errors, because in this alternative model endowment growth does have an actually time-varying component and the agent optimally tracks it, while the growth rate is objectively IID in our model.

Our model is also related to, but 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 expected 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, consistent with the high persistence of the price-dividend ratio 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 growth rate uncertainty which allows us 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. The representative agent in our model also uses exponential weighting of past growth, but

to forecast future fundamentals, not prices. As a consequence, risk premia arise for different reasons. 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 uncertainty.

In addition to the paper by Collin-Dufresne, Johannes, and Lochstoer (2017) that we discussed above, a number of other recent papers take an overlapping generations approach to study learning from experience effects in asset pricing. While this approach can deliver interesting insights into the heterogeneity between cohorts, these models can be solved only with stark simplifications that affect the aggregate asset pricing implications: 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. By abstracting from cross-cohort heterogeneity, we also employ a simplified approach, but one that delivers quantitatively realistic asset-pricing predictions.

II. INITIAL EVIDENCE ON SUBJECTIVE AND OBJECTIVE EXPECTED RETURNS

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 μ of log stock market payouts, d,

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

where ϵ is an IID shock. The microdata evidence in Malmendier and Nagel (2011) and
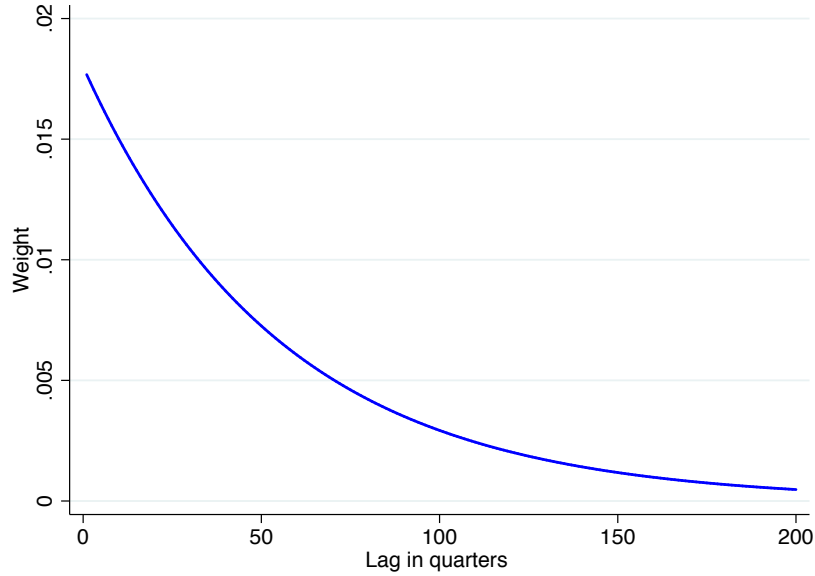
FIGURE I

Weights implied by constant-gain learning

Weights on quarterly past observations implied by constant-gain learning with gain $\nu = 0.018$.

Malmendier and Nagel (2016) suggests that individuals form expectations from data they observe throughout their lifetimes and 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 in this case the belief of the *average* individual can be captured well by a constant-gain learning rule where the perceived growth rate $\tilde{\mu}$ evolves as

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

and where ν 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 ν. Malmendier and Nagel (2016) show that $\nu = 0.018$ for quarterly data fits the dynamics of the average belief in microdata about inflation expectations (and this value is also within the range of estimates obtained from microdata on household investment decisions in Malmendier

8

and Nagel (2011)). Iterating on (2) on can see that $\tilde{\mu}_{d,t}$ is an exponentially-weighted average of past Δd observations, with weights declining more quickly going back in time the higher ν. Figure I shows how the weights decline from 0.018 for the most recent observation to very close to zero for observations dating back to 200 quarters ago or earlier. 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' subjective beliefs are formed through learning with constant gain. Consider a setting in which investors form expectations as in (2). Based on their time-t beliefs, 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 θ and a constant 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.

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

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

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

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

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

9

in investors' subjectively expected component of returns we obtain total returns

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

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

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

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, given her knowledge of μ_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 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

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

II.A. Measurement of experienced growth

To implement the constant-gain learning scheme in (2), we need a long history of past observations on stock market fundamentals. 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. To compute $\tilde{\mu}_{d,t}$ in 1926, we then need data on Δd stretching back at to least around 50 years earlier or so, up to the point where the weights become close to negligible.

Our first empirical measure of $\tilde{\mu}_{d,t}$ uses real per-capita stock market payout growth, Δd. From 1926 onwards, we obtain quarterly observations of Δd on the CRSP value-weighted

10

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. 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, data on aggregate corporate non-farm non-financial dividends from Wright (2004) for the period 1900 to 1913, and per-capita GDP growth rates from Barro and Ursua (2008) as proxy for Δd from 1871 to 1900. Appendix A.1 provides more details on how we construct the real per-capita payout growth series. We then 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 data is of lower quality than the CRSP data. For these reasons, we also construct an alternative measure where we take a weighted average of past log real stock market index returns instead of payout growth. 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. For averages taken over long periods, payout growth and returns should be highly correlated and hence the experienced return series, $\tilde{\mu}_{r,t}$, should capture similar information as $\tilde{\mu}_{d,t}$. To check this, we simulated dividend growth and returns from (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 $\tilde{\mu}_t$ depends on the gain parameter ν. For the value $\nu = 0.018$ that we work with here, the correlation is a very high 0.82. Other parameters like μ_d, the variance of ϵ, θ, r_f, or ρ 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

11

when we study $\tilde{\mu}_r$ in data simulated from our asset-pricing model.

While experienced returns have some advantages as a measure of experienced fundamentals growth, this approach is subject to the potential shortcoming that 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 (that we estimate in our regressions below). For this reason, we use both measures, $\tilde{\mu}_d$ and $\tilde{\mu}_r$, in our tests.

II.B. *Survey data on return expectations*

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 investor expectations from surveys. The changes over time in investor experiences that we focus on are only slowly moving over time. To study their relationship with investor expectations and expectations errors, we need a sufficiently long time series of surveys. For this reason, we put together survey data from several sources that spans the period 1972 to 1977 and 1987 to 2016. We focus largely on surveys that target a representative sample of the U.S. population, supplemented with two surveys of brokerage and investment firm customers.

Several of these surveys 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 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:

- Michigan Survey of Consumers, monthly 2002-2016

12

- Conference Board Survey, monthly 1987-2016[5]

- Roper Center Surveys, annual, 1974-1977

The latter three surveys don't elicit the percentage expected return. Instead, respondents provide the probability of a rise in the stock market over a one-year horizon (Michigan survey) 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, Roper). 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 average reported probability of a rise in the stock market in the periods when the Michigan Survey overlaps with the first set of surveys. We then project the resulting extended series of percentage expectations on the proportion of respondents forecasting an increase in the stock market in the Conference Board and Roper Center surveys. Appendix A.2 provides more detail.[6]

II.C. 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 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. The present-value model in (7) would predict an OLS slope coefficient of -4.70 (based on $\nu = 0.018$ and $\rho = 0.99$, which is the quarterly value implied by the value of $\rho = 0.964$ for annual data reported in Campbell

5. The data was kindly provided by The Conference Board.
6. 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 AAII survey is conducted among members of the AAII 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 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.

(2000)).

As Panel A shows, our estimates in the full 1927-2016 period are close to 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 quite close to the predicted value of -4.7. 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 column (2).

14

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 dividend-price 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) 1927-2016	(2) 1927-2016	(3) 1927-2016	(4) 1946-2016	(5) 1946-2016
Panel A: Predicting returns with experienced real payout growth					
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
Panel B: Predicting returns with experienced returns					
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 (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 of smaller magnitude than the predicted value of -4.70 but it is still roughly in the ballpark of the predictions of the present-value model. 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. The simulations of the present-value model that we described in Section II.A yield an average predictive regression coefficient of -1.36 for $\tilde{\mu}_r$. The results in column (1) show that the coefficient that for the full sample from 1927-2015, we obtain an OLS estimate of -2.36. Bias-adjustment shrinks the point estimate to -1.74. The remaining columns show that adding inflation and $p-d$ and focusing on the post-1945 sample has little effect on the predictive relationship between experienced real stock returns and future excess stock returns.

Figure II 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 tended 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
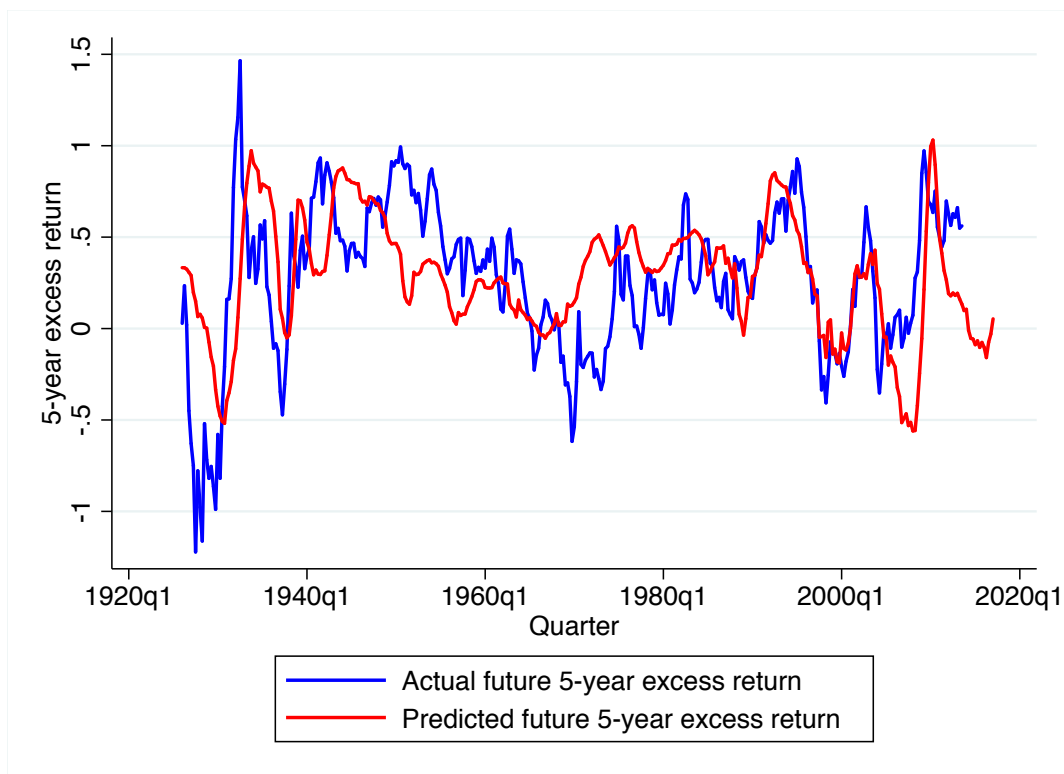
16

Figure II

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 $t + 20$.

pick the ν 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.

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.D. Subjective Expectation Error Predictability

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 we will show now, it is also approximately true empirically at the frequencies that are relevant for our theory.

To check the time-series relationship between experienced growth and subjective expected excess returns, Panel A of Table II presents regressions of quarter t survey expectations of one-year excess returns on experienced real payout growth and experienced real returns. The experienced growth variables are calculated based on observations leading up to the end of quarter $t-1$. We calculate subjective expected excess returns from survey expectations of stock market returns by subtracting the one-year Treasury yield measured at the end of quarter $t-1$. As Panel A shows there is only a weak, and statistically not significant, positive relationship between μ_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 columns (1).

Looking at past returns over a much shorter time window, Greenwood and Shleifer (2014) find that survey 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) shows that when experienced growth variables and lagged returns are used jointly, the experienced growth effect remains very weak. In terms of the lower frequency movements that we focus on in our analysis, the subjective equity premium in the survey data is close to acyclical. That there
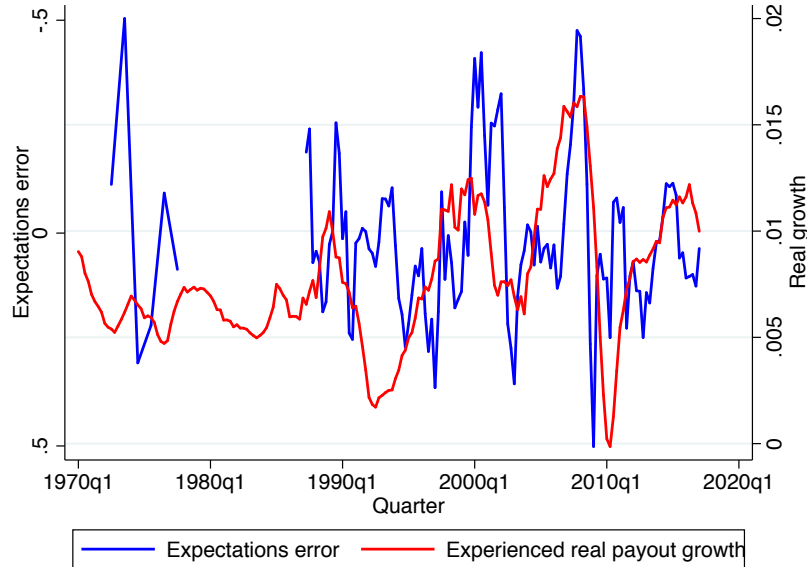
18

FIGURE III

Experienced real payout growth and subjective expectation errors

Expectations error is the one-year realized return on the CRSP value-weighted index minus the survey expectation of stock market returns prior to the return measurement period. Expectations error axis shown with reversed scale.

are short-run fluctuations in subjective return expectations with one-year lagged returns are 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 use the survey return expectations series to calculate the expectations error $r_{t+1} - \tilde{\mathbb{E}}_t r_{t+1}$ on the left-hand side of (7). To be consistent with the one-year time horizon of the survey expectation, we use the simple one-year return from the beginning of quarter $t+1$ to the end of quarter $t+4$ as the return that we compare with the survey expectation in quarter t. 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

19

TABLE II

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. Newey-West standard errors are reported in parentheses (12 lags in Panel A; 6 lags in Panel B).

	(1)	(2)	(3)	(4)	(5)
Panel A: Subjective expected excess returns					
Experienced real payout growth	0.35		0.42		
	(0.99)		(0.93)		
Experienced real returns				0.90	0.42
				(0.65)	(0.77)
Lagged one-year return		0.03	0.03		0.03
		(0.01)	(0.01)		(0.01)
Constant	0.05	0.05	0.05	0.04	0.05
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
Observations	126	126	126	126	126
Adj. R^2	-0.004	0.075	0.074	-0.033	0.075
Panel B: Expectation error: Realized - subj. expected					
Experienced real payout growth	-12.31		-12.55		
	(6.75)		(6.80)		
Experienced real returns				-14.27	-15.48
				(6.48)	(7.08)
Lagged one-year return		-0.10	-0.12		0.07
		(0.14)	(0.15)		(0.13)
Constant	0.12	0.03	0.14	0.25	0.27
	(0.06)	(0.03)	(0.06)	(0.09)	(0.10)
Observations	126	126	126	126	126
Adj. R^2	0.055	0.002	0.059	0.106	0.102

is actually a negative relationship in the part 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. 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 shows that this is approximately true. Figure III provides a visual impression of the time-series relation between experienced payout growth and expectations errors (the expectations errors are plotted on a reversed scale). Even though one-year expectations errors are very noisy, the relationship with experienced real payout growth is quite apparent.

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

$$\Delta c_{t+1} = \mu + \sigma \varepsilon_{t+1}, \tag{9}$$

where $\{\varepsilon_t\}$ is a series of IID standard normal shocks. The agent knows that Δc_{t+1} is IID, and she also knows σ, but not μ. The agent relies on the history of past endowment growth realizations, $H_t \equiv \{\Delta c_0, \Delta c_1, \ldots, \Delta c_t\}$, to form an estimate of μ.

After seeing the data H_t, a Bayesian agent with full memory would modify the prior beliefs $p(\mu)$ she held before seeing H_t in a way that assigns each past observation Δc_{t-j} equal weight in the likelihood. Equal-weighting of data generated from a perceived IID law of motion means that there is no decay of memory as the agent uses all available data in forming posterior beliefs about μ.

Here we develop a constant-gain learning scheme that implies fading memory. Unlike standard constant-gain learning models, however, we retain the modeling of the full posterior distribution—and hence the agent's subjective uncertainty—as in the Bayesian approach. To do so, we use a weighted-likelihood approach 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 Δc, but with fading memory, forms a posterior

$$p(\mu|H_t) \propto p(\mu) \prod_{j=0}^{\infty} \left[\exp\left(-\frac{(\Delta c_{t-j} - \mu)^2}{2\sigma^2} \right) \right]^{(1-\nu)^j}, \tag{10}$$

where $1 - \nu$ is a positive number close to one. 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

$$\mu \sim \mathcal{N}(\mu_0, \sigma_0^2) \tag{11}$$

we then obtain the posterior

$$\mu|H_t \sim \mathcal{N}\left(\frac{\sigma_0^2}{\nu\sigma^2 + \sigma_0^2}\tilde{\mu}_t + \frac{\nu\sigma^2}{\nu\sigma^2 + \sigma_0^2}\mu_0, \left(\frac{1}{\sigma_0^2} + \frac{1}{\nu\sigma^2} \right)^{-1} \right), \tag{12}$$

where

$$\tilde{\mu}_t = \nu \sum_{j=0}^{\infty} (1-\nu)^j \Delta c_{t-j}. \tag{13}$$

The variance of the posterior is the same as if the agent had observed, and retained fully in memory with equal weight, $S \equiv 1/\nu$ 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.

22

For now, however, we work with an uninformative prior ($\sigma_0 \to \infty$) and hence the posterior

$$\mu | H_t \sim \mathcal{N}(\tilde{\mu}_t, \nu\sigma^2). \tag{14}$$

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

$$\tilde{\mu}_t = \tilde{\mu}_{t-1} + \nu(\Delta c_t - \tilde{\mu}_{t-1}). \tag{15}$$

Thus, the $\tilde{\mu}_t$ resulting from this weighted-likelihood approach with an uninformative prior is identical to the perceived μ that one obtains from the constant-gain updating scheme (2) with gain ν. However, in contrast to standard constant-gain learning specifications in macroeconomics that focus purely on the first moment (Evans and Honkapohja 2001), 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

$$\Delta c_{t+j} | H_t \sim \mathcal{N}\left(\tilde{\mu}_t, (1+\nu)\sigma^2\right), \qquad j = 1, 2, ..., \tag{16}$$

where the variance of the predictive distribution reflects not only the uncertainty due to future ϵ_{t+j} shocks, but also the uncertainty about μ. 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 (15) as

$$\tilde{\mu}_{t+1} = \tilde{\mu}_t + \nu\sigma\sqrt{1+\nu}\tilde{\varepsilon}_{t+1}, \qquad \text{where} \quad \tilde{\varepsilon}_{t+1} = \frac{\Delta c_{t+1} - \tilde{\mu}_t}{\sigma\sqrt{1+\nu}}, \tag{17}$$

and $\tilde{\varepsilon}_{t+1}$ 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 μ 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 not posterior means will not be convering to μ. Consistent with stationarity, the time-t agent perceives future increments $\tilde{\varepsilon}_{t+j}$, $j = 1, 2, ...$ in (17) as negatively serially correlated (see Appendix C for more details) and not as martingale differences.[7] When we calculate asset prices in this model, we therefore cannot directly rely on certain laws like the law of iterated expectations that require a filtration or results that presume martingale posteriors. This requires special care in evaluating valuation equations.

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,

$$P_{R,t} = \tilde{\mathbb{E}}_t \left[M_{t+1|t} \tilde{\mathbb{E}}_{t+1} \left[M_{t+2|t+1} C_{t+2} \right] \right], \tag{18}$$

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

---

7. At time t, the agent however cannot make use of this serial correlation by using $\tilde{\varepsilon}_t$ to forecast $\tilde{\varepsilon}_{t+1}$, because $\tilde{\varepsilon}_t$ 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.

based on the stochastically discounted payoff under the time-t predictive distribution

$$P_{H,t} = \tilde{\mathbb{E}}_t[M_{t+1|t} M_{t+2|t} C_{t+2}], \tag{19}$$

In a full-memory Bayesian setting, the law of iterated expectations (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.[8] As a consequence $P_{R,t} \neq P_{H,t}$.

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

The valuation discrepancy between the two valuation approaches arises because the agent at t and at $t+1$ see the statistical properties of the shock $\tilde{\varepsilon}_{t+2}$ differently. The buy-and-hold valuation incorporates this 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 view the one-period ahead $\tilde{\varepsilon}$ shock as unpredictable.

---

8. For subjective expectations of linear functions of Δc, 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 Δc, e.g., $\tilde{\mathbb{E}}_t \tilde{\mathbb{E}}_{t+1}[\exp(a + b\Delta c_{t+2})] \neq \tilde{\mathbb{E}}_t[\exp(a + b\Delta c_{t+2})]$.

III.C. Kalman Filtering Interpretation

The updating scheme in (15) is reminiscent of optimal filtering in the case of a latent stochastic trend. Indeed, if the agent *perceived* μ to follow a random walk—i.e., as $\mu_t = \mu_{t-1} + \zeta_t$, where ζ is an IID-normal shock—rather than a constant as in (9), application of the steady-state Kalman filter yields exactly the same posterior distribution as in (14). With appropriate choice of the volatility of the ζ shocks, and as long as the *actual* law of motion is still (9) with constant μ, 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.[9] 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.

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

$$V_t = \left[(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta \tilde{\mathbb{E}}_t [V_{t+1}^{1-\gamma}]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}}, \tag{20}$$

---

9. However, for the asset-pricing predictions to remain the same, it is crucial that (9), with constant μ, 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.

where δ denotes the time discount factor, γ relative risk aversion for static gambles and ψ the elasticity of intertemporal substitution (EIS). Note that the agent evaluates the continuation value under her subjective expectations $\tilde{\mathbb{E}}_t[.]$. 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 stochastic discount factor (SDF) that prices assets under the agent's subjective beliefs,

$$m_{t+1|t} = \tilde{\mu}_m - \tilde{\mu}_t - \xi\sigma\tilde{\varepsilon}_{t+1}, \tag{21}$$

with

$$\tilde{\mu}_m = \log\delta - \frac{1}{2}(1-\gamma)^2(\nu U_v + 1)^2(1+\nu)\sigma^2, \tag{22}$$

$$\xi = [1 - (1-\gamma)(\nu U_v + 1)]\sqrt{1+\nu}, \tag{23}$$

$$U_v = \frac{\delta}{1-\delta}. \tag{24}$$

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

$$r_{f,t} = -\tilde{\mu}_m + \tilde{\mu}_t - \frac{1}{2}\xi^2\sigma^2. \tag{25}$$

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 (17), we can express the log of the return on the consumption claim as

$$r_{w,t+1} = \tilde{\mu}_t + \sqrt{1+\nu}\sigma\tilde{\varepsilon}_{t+1} + \log\left(\frac{\zeta}{\zeta-1}\right). \tag{26}$$

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

$$\zeta = \frac{1}{1-\delta}. \tag{27}$$

Thus, similar to the standard rational expectations case, $\psi = 1$ implies a constant and finite consumption-wealth ratio. In the posterior distribution in (14), extremely large values of μ 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 ζ is constant, we can solve for the subjective risk premium of the consumption claim

$$\log \tilde{\mathbb{E}}_t[R_{w,t+1}] - r_{f,t} = \xi\sqrt{1+\nu}\sigma^2, \tag{28}$$

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 (26), we can calculate the wedge between subjective and objective expectations, and combining with (28), we obtain the objective risk premium

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

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 μ, 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

$$\Delta d_{t+1} = \lambda \Delta c_{t+1} - \alpha(d_t - c_t - \mu_{dc}) + \sigma_d \eta_{t+1}, \quad \alpha > 0, \tag{30}$$

similar to Bansal, Gallant, and Tauchen (2007). We assume that μ_{dc}, λ, and α are known to the agent. The agent's learning problem is focused on the unknown μ.

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 α is very small and so dividends and consumption can drift away from each other quite far, but we keep α sufficiently big to yield a finite price-dividend ratio with empirically reasonable moments.

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

$$P_t^n \equiv \tilde{\mathbb{E}}_t[M_{t+1}|_t\tilde{\mathbb{E}}_{t+1}[\cdots\tilde{\mathbb{E}}_{t+n-1}[M_{t+n}|_{t+n-1}D_{t+n}]]]. \tag{31}$$

As we discussed earlier, when we evaluate these expectations, we do so 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 (31), we obtain

$$p_t^n - d_t = [1 - (1-\alpha)^n]\left(c_t - d_t + \mu_{dc} + \frac{\lambda - 1}{\alpha}\tilde{\mu}_t\right) + n\tilde{\mu}_m + \frac{1}{2}(A_n\sigma^2 + B_n\sigma_d^2), \qquad (32)$$

where

$$A_n = \sum_{k=0}^{n-1}\left\{\sqrt{1+\nu}\left[\nu(\lambda-1)\frac{1-(1-\alpha)^k}{\alpha} + (\lambda-1)(1-\alpha)^k + 1\right] - \xi\right\}^2, \qquad (33)$$

and

$$B_n = \frac{1-(1-\alpha)^{2n}}{1-(1-\alpha)^2}. \qquad (34)$$

For very large n, approximately,

$$A_n \approx n\left[\sqrt{1+\nu}\left(1+\nu\frac{\lambda-1}{\alpha}\right) - \xi\right]^2 \qquad (35)$$

and B_n, which doesn't 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 (32) to be (weakly) negative. Using (35), we see that this requires

$$\tilde{\mu}_m + \frac{1}{2}\left[\sqrt{1+\nu}\left(1+\nu\frac{\lambda-1}{\alpha}\right) - \xi\right]^2\sigma^2 \leq 0. \qquad (36)$$

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

As equation (32) 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

30

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 (30)—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 $d - c$.[10] 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 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.

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

$$R_{t+1}^\infty \equiv \lim_{n \to \infty} P_{t+1}^{n-1}/P_t^n. \tag{37}$$

As we show in Appendix E.3, we can use equation (32) to find the one-period subjective risk premium for this claim

$$\log \tilde{\mathbb{E}}_t[R_{t+1}^\infty] - r_{f,t} = \left[1 + \nu \frac{\lambda - 1}{\alpha} \right] \xi \sqrt{1 + \nu} \sigma^2. \tag{38}$$

For $\gamma \geq 1$, ξ is a positive constant. We observe from the above that lowering α 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 ν

10. 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.

because higher ν implies a smaller effective sample size used to estimate μ and hence higher subjective uncertainty about μ.

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

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

(39)

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 (39) 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 price of the consumption strip is constant and the objective expected growth rate of Δc 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.

We can also use the dividend strips to evaluate the subjective equity term premium. A short term claim with return $R_{t+1}^1 \equiv D_{t+1}/P_t^1$, has the subjective risk premium

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

(40)

The subjective equity term premium $\tilde{\theta}_t^{\infty,1} \equiv \log \tilde{\mathbb{E}}_t[R_{t+1}^\infty] - \log \tilde{\mathbb{E}}_t[R_{t+1}^1]$ is

$$\tilde{\theta}_t^{\infty,1} = (\lambda - 1)\left(\frac{\nu}{\alpha} - 1\right)\xi\sqrt{1+\nu}\sigma^2.$$

(41)

32

With $\alpha < \nu$, the term premium is positive, i.e., the agent perceives long-horizon equity as riskier than short-horizon equity.

The ex-dividend price of the claim to the entire stream of dividends is simply the sum of prices of dividend strips

$$P_t = \sum_{n=1}^{\infty} P_t^n. \tag{42}$$

Having the restriction (36) hold as a strict inequality ensures that this sum remains finite. We compute the sum in (42) 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 contributions 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 it subjectively priced more like long-horizon equity, i.e., with a higher risk premium.

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 μ 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 any point in the future will always have a similarly strong tilt towards the prior mean.

Our approach is similar in spirit to the 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 μ, but this is not essential. Since we work with high prior variance, the prior is still almost uninformative, and setting the prior mean to any value in an economically plausible neighborhood around μ would deliver similar results.

Given the prior in (11) with $\mu_0 = \mu$, we obtain the posterior

$$\mu|H_t \sim \mathcal{N}\left(\phi\tilde{\mu}_t + (1-\phi)\mu, \left(\frac{1}{\sigma_0^2} + \frac{1}{\nu\sigma^2}\right)^{-1}\right), \qquad \text{where} \quad \phi \equiv \frac{\sigma_0^2}{\nu\sigma^2 + \sigma_0^2}. \tag{43}$$

The perceived consumption growth can be represented as

$$\Delta c_{t+1} = \phi\tilde{\mu}_t + (1-\phi)\mu + \sqrt{1+\phi\nu}\sigma\tilde{\varepsilon}_{t+1}, \tag{44}$$

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 III summarizes the parameter values we use in our baseline quantitative analysis. We fix the gain parameter ν 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 σ_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 most important for our purposes. The reason why we cannot match both at the same time is that in our model Δd is IID. In the data, however, Δd 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 Δd in the data, which is influenced by a substantial transitory component.

In our baseline specification, we fix $\psi = 1$ to obtain a solution with diffuse prior, but we relax this below. We choose the remaining parameters γ and α 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 (36) required for a finite price of the dividend claim and it implies that the forces of cointergration are quite weak and dividends can wander quite far away from consumption (but not as far as in models without cointegration, such as, e.g., Bansal, Kiku, and Yaron (2012)).

IV.A. Unconditional moments

We simulate the model at a quarterly frequency. Table IV reports the annualized population moments estimated from an extremely long sample simulated from the model. We also show

TABLE III
Baseline Model Parameters

This table reports the parameters values we use in the baseline calibration of our model for a quarterly frequency. The gain parameter ν 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 many at the same values as in Bansal, Kiku, and Yaron (2012) and Collin-Dufresne, Johannes, and Lochstoer (2017).

Parameter	Symbol	Value	Source
Belief updating			
Gain	ν	0.018	MN (2016) (survey data)
Endowment process			
Leverage ratio	λ	3	CJL (2017)
Dividend cointegration parameter	α	0.001	
Mean consumption growth	μ	0.45%	CJL (2017)
Consumption growth volatility	σ	1.35%	CJL (2017)
Dividend growth volatility	σ_d	1%	
Preferences			
Risk aversion	γ	4	
EIS	ψ	1	
Time discount factor	δ	0.9967	BKY (2012)

TABLE IV
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 the corresponding 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 total 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 Δc from four-quarter changes of quarterly log nondurables and services consumption and those of Δd 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
$\mathrm{corr}(\tilde{\mu}, \tilde{\mu}_d)$	-	0.59	0.59
$\mathrm{corr}(\tilde{\mu}, \tilde{\mu}_r)$	-	0.68	0.68
$\mathbb{E}(R_m - R_f)$	8.11	7.16	7.65
$\sigma(R_m - R_f)$	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

empirical moments for the 1927 to 2016 period for comparison.

As we anticipated, the volatility of Δd 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 Δd in the data is due to transitory components. The persistence of $\tilde{\mu}$ and $\tilde{\mu}_d$ equals $1 - \nu$, i.e., it is pinned down by the gain parameter.

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 model also produces a low subjective real risk-free rate with low volatility. 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). 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 Returns and Subjective Expectations Errors*

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

TABLE V
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 and $\tilde{\mu}_d$ $\tilde{\mu}_r$ 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^2_{adj}	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^2_{adj}	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^2_{adj}	0.01	0.04	0.18	0.01	0.05	0.21

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 V 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. This 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 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 predictive regressions for horizons beyond one quarter, but Figure II 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 $\phi = 0.99$. 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.

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 VI 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 μ_d as predictor, is 0.35, while the regression on model-generated data in column (1) of Table VI 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, exactly equal to the empirical estimate of 0.60 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 and so the take-away from this analysis is that subjective expected excess

TABLE VI

SURVEY RETURN EXPECTATIONS AND EXPERIENCED REAL RETURNS IN SIMULATIONS

This table reports the mean estimates from regressing subjective expected excess returns and expectation errors on $\tilde{\mu}_r$ 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)
Panel A: Subjective expected excess returns					
$\tilde{\mu}_d$	0.83		0.83		
$\tilde{\mu}_r$				0.60	0.62
$r_{t-3,t}$		0.01	0.00		0.00
R^2_{adj}	0.93	0.08	0.93	0.83	0.83
Panel B: Expectation error: Realized - subj. expected					
$\tilde{\mu}_d$	-10.75		-11.05		
$\tilde{\mu}_r$				-6.80	-7.24
$r_{t-3,t}$		-0.06	0.02		0.03
R^2_{adj}	0.05	0.01	0.06	0.04	0.04

returns in the data and the model are nearly constant.[11]

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 -12.31 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, but the latter also come with a big standard error of 6.48.

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 μ_d and μ_r, while subjective expected excess returns are largely acyclical.

11. A comparison of the predictive regression coefficients in Tables V and VI 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 9.35 times volatility of $\tilde{\mu}_d$ based on quarterly Δd of $0.39\% \approx 3.65\%$) as the volatility of subjective expected returns (regression coefficient of 0.83 times volatility of $\tilde{\mu}_d \approx 0.32\%$).

IV.C.   Lack of Out-of-Sample Return Predictability

Goyal and Welch (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 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 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 $d-p$ or $\tilde{\mu}_d$ as predictor. The blue line in the upper half of each plot in Figure IV 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 $d-p$ as a predictor (top) and $\tilde{\mu}_r$ (bottom), indicates that the in-sample R^2 is greater than zero.

To assess the out-of-sample performance, we start after 80 quarters into each 360-quarter sample, use the backward looking history from the start of the 360-quarter sample to compute the sample mean and the predictive regression and calculate squared prediction errors of the mean as a forecaster minus the cumulative squared prediction error of the fitted predictive regression, which we then cumulate forward. We then average the resulting path of the cumulative squared error differences across all simulations. The red lines in the lower half of each plot in Figure IV show that this path is in negative territory, which means that the
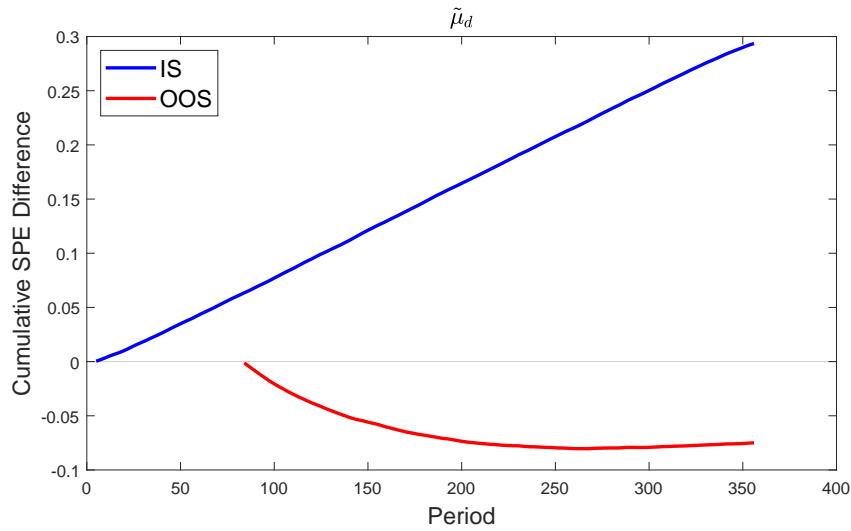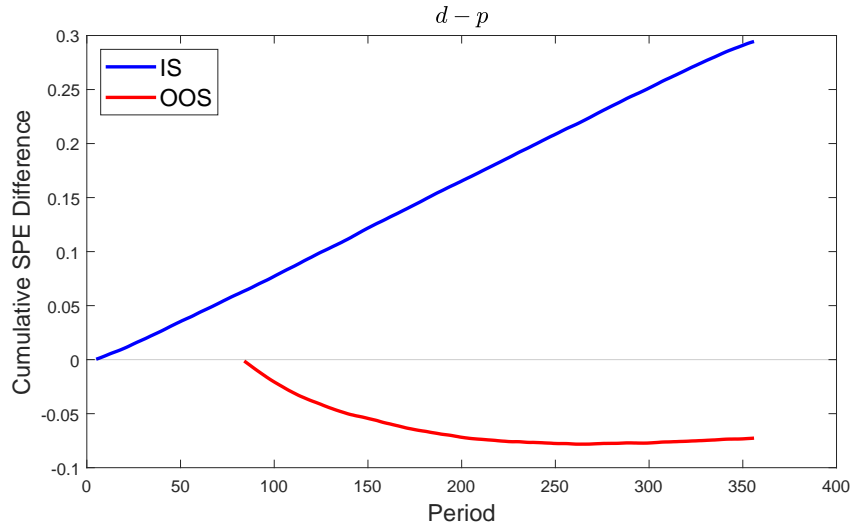
42

FIGURE IV

Out-of-sample predictive performance

In-sample and out-of-sample performance of predictors $d-p$ 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.

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 Goyal and Welch (2008).

The out-of-sample exercise is 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.

IV.D.   Conditional return volatility

In the model, subjective and objective conditional volatility of the dividend strip returns are constant over time. However, for the same reason that the subjective equity claim risk premium is slightly increasing in $\tilde{\mu}_t$, subjective and objective conditional volatility of the equity claims return also weakly increase with $\tilde{\mu}_t$: higher subjectively expected growth implies higher weight of riskier longer-horizon dividend strips in the equity claim's value. Thus, the objective equity premium, which is decreasing in $\tilde{\mu}_t$, is negatively related to conditional equity return volatility.

On this dimension, the model predictions are very different from those of the Bansal and Yaron (2004) (BY) and Campbell and Cochrane (1999) (CC) models. In both of these models, the equity premium is approximately linearly and positively related to conditional equity return volatility.[12]

Table VII shows that empirical data is broadly consistent with the predictions of our

12. See Beeler and Campbell (2012) for the BY model. For the CC model one can infer the positive and close to linear relationship by comparing Figures 4 and 5 in CC.

44

model. The dependent variable in these regressions is the square root of the sum of squared daily log returns of the CRSP value-weighted index in quarter $t + 1$, i.e., an estimate of quarterly realized volatility. As columns (1) and (3) show, there is a weakly positive, but not statistically significant relationship between $\tilde{\mu}_{d,t}$ and next-quarter volatility (and hence a negative relationship between the conditional equity premium captured by $\tilde{\mu}_d$ and conditional volatility) both in the full sample and the post-WW II sample. The magnitude of the conditional volatility movement implied by the point estimates is miniscule: a one standard deviation increase in $\tilde{\mu}_d$ is associated with an increase of 0.52 percentage points in conditional quarterly volatility. From our simulations we calculate that the corresponding increase in conditional volatility in the model would be a similarly small 0.11 percentage points.

Using $p_t - d_t$ as a predictor in columns (2) and (3), the picture is mixed. In the full sample, $p_t - d_t$ 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 BY and CC models.

Thus, the decoupling objective and subjective beliefs allows the model to also decouple conditional equity return volatility and the objective conditional equity premium. In this way the model is able to capture the fact that at low frequencies captured by the slow-moving predictors $\tilde{\mu}_d$ and $p - d$, there isn't much co-movement between equity premium and volatility.

There is, of course, a substantial body of evidence showing that volatility rises strongly after market crashes. Our model is not able to explain this fact without adding further elements to the model, but neither are the BY and CC models where conditional volatility varies with slow-moving state variables that are tied to the conditional equity premium.

IV.E. Objective Equity Term Structure

As Van Binsbergen, Brandt, and Koijen (2012) and Van Binsbergen and Koijen (2017) demonstrate, the equity term structure is another useful dimension of the data that can help discriminate between asset pricing models. As we show in Appendix E.3, the objective conditional

TABLE VII
Predicting Volatility

Dependent variable is the square root of 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 dividend-price 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) 1927-2016	(2) 1927-2016	(3) 1946-2016	(4) 1946-2016
Experienced real payout growth	1.58 (1.65)		1.39 (1.41)	
$(p - d)/100$		-0.67 (1.56)		2.20 (0.66)
Constant	0.06 (0.01)	0.10 (0.06)	0.05 (0.01)	-0.01 (0.02)
Observations	360	360	284	284
Adjusted R^2	0.010	0.001	0.010	0.068

equity term premium earned by the infinite-horizon claim in our model is

$$
\log \mathbb{E}_t[R_{t+1}^\infty] - \mathbb{E}_t[R_{t+1}^1] = (\lambda - 1) \left[\frac{\nu}{\alpha} - 1 \right] \xi \sqrt{1+\nu} \sigma^2 - \frac{1}{2} \nu \left[\left(1 + \nu \frac{\lambda - 1}{\alpha} \right)^2 - \lambda^2 \right] \sigma^2
$$
$$
+ (\lambda - 1) \left[\frac{\nu}{\alpha} - 1 \right] (\mu - \tilde{\mu}_t). \tag{45}
$$

On average, the slope is positive in our baseline calibration with $\psi = 1$ and diffuse prior, as shown by the solid line in Figure V (with $\psi = 1.5$ it is slightly flatter). This prediction seems to be in conflict with the empirical evidence of an unconditionally negative slope in Van Binsbergen, Brandt, and Koijen (2012) and Van Binsbergen and Koijen (2017). However, in terms of statistical significance, the existing empirical evidence for an unconditionally positive slope is quite weak. Moreover, as Gormsen (2018) highlights, there is a large amount of predictable counter-cyclical variation in the slope around the slightly negative unconditional slope in the data.

This is also true in our model. The dashed lines in Figure V show the slope based on
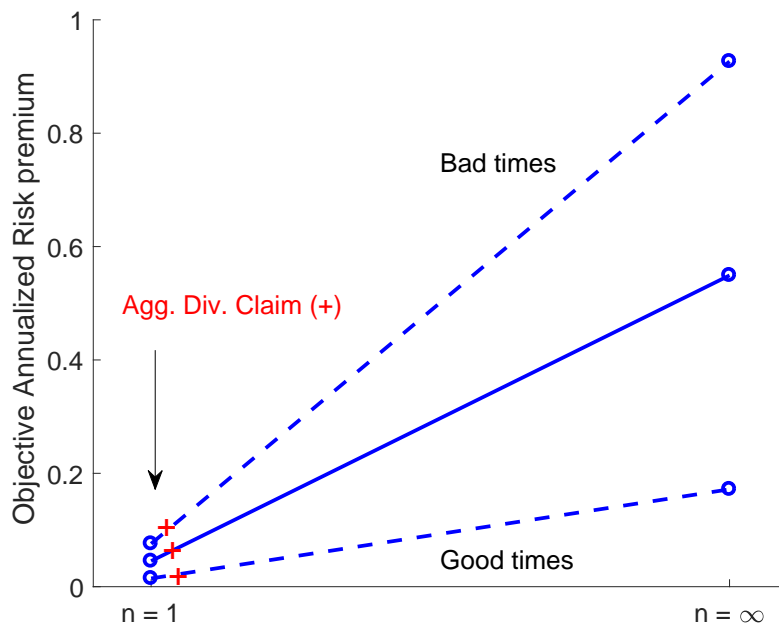
FIGURE V
Objective Equity Term Structure

The figure shows the annualized log conditional expected excess returns for the one-period dividend strip and the infinite-horizon dividend strip with $\tilde{\mu} = \mu$ (solid line) and $\tilde{\mu} = \mu \pm 2$ S.D. (dashed lines).

(45) in good times (when $\tilde{\mu}$ is two standard deviations above μ) and bad times (when $\tilde{\mu}$ is two standard deviations below μ). In line with Gormsen's empirical evidence, the equity term structure in our model is countercyclical: the slope is steeper in bad times. While the slope can turn negative in our model, in our calibration it does so less frequently than in Gormsen's empirical data. It takes an about 2.5 standard deviation positive shock to $\tilde{\mu}$ to get the conditional term premium to turn negative.

Figure V also shows the conditional equity premium of the equity claim to the aggregate stream of dividends. It shows that the equity claim in bad times behaves somewhat more like a claim to short-horizon dividends, consistent with our discussion at the end of Section III.F.

As Gormsen (2018) shows, the Campbell and Cochrane (1999) and Bansal and Yaron

(2004) models produce a countercyclical equity term structure. Our model shares this prediction with these earlier models. However, in our model this effect arises for a very different reason: All variation in the equity term structure is driven by subjective beliefs. Following a string of positive endowment growth surprises, the agent in our model is subjectively too optimistic about long-run growth relative to the objective distribution, i.e., $\tilde{\mu} - \mu > 0$. This optimism about long-run growth leads to overpricing especially of long-horizon levered claims. To the econometrician, the correction of this overpricing in the future is predictable, leading to low objective expected excess returns for long-horizon claims. In bad times, the opposite applies.

Moreover, in Campbell and Cochrane (1999) and Bansal and Yaron (2004), the slope of the equity term premium can never be negative. In contrast, in our model it can turn negative, albeit only rarely. In addition, as we show now, our model differs strongly from these earlier models in terms of the predictability of the relative returns of long and short-horizon equity claims. Specifically, we follow Gormsen (2018) and look at predicting relative returns with the equity yield spread. Equity yields are defined as

$$e_t^n = \frac{1}{n}(d_t - f_t^n), \tag{46}$$

where $f_t^n = p_t^n + n y_t^n$ is the forward price of a claim to the time $t + n$ dividend, y_t^n is the yield on a zero-coupon bond maturity at $t + n$, and the equity yield spread is

$$s_t^n \equiv e_t^n - e_t^1 = \frac{1-n}{n}d_t - \frac{1}{n}p_t^n + p_t^1 - y_t^n + y_t^1, \tag{47}$$

where $y_t^1 = r_{f,t}$. The bond yield spread $y_t^n - y_t^1$ is constant in our model. Using the definition of s_t^n combined with equation (32), we obtain, by taking the appropriate limit as $n \to \infty$,

$$s_t^\infty = \text{const.} - \alpha(d_t - c_t) + (\lambda - 1)\tilde{\mu}_t. \tag{48}$$

Here a rise in $\tilde{\mu}_t$ has a positive effect on p_t^1 and, accordingly, a negative one on e_t^1. There is also a positive effect on p_t^n for $n > 1$, but, because dividends slowly revert back towards consumption, the magnitude is less than n times the effect on p_t^1, so the effect on e_t^n is smaller. In the limit, the effect on e_t^∞ is nil, leading to a positive relationship between $\tilde{\mu}_t$ and s_t^∞. In contrast, the objective equity term premium in (45) is negatively related to $\tilde{\mu}_t$. Thus, our model produces two predictions that are in line with key findings in Gormsen (2018): (i) the equity yield spread is negatively related to the objective equity term premium; (ii) the pro-cyclical variation in the equity yield spread in our model comes from the short end of the curve.

As Gormsen (2018) points out, the habit model of Campbell and Cochrane (1999) and the long-run risks model of Bansal and Yaron (2004)) produce a positive relationship between the equity yield spread and the objective term premium, which is not consistent with his empirical findings.

V. 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 and risk premia are high because subjective 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. 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.

<center>REFERENCES</center>

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.

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.

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.

<center>51</center>

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, 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).

Gormsen, Niels, 2018, "Time-Variation in the Equity Term Structure," Working paper, Copenhagen Business School.

Goyal, Amit, and Ivo Welch, 2008, "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction," *Review of Financial Studies* 21, 1455–1508.

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, 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.

Lease, Ronald C, Wilbur G Lewellen, and Gary G Schlarbaum, 1974, "The individual investor: attributes and attitudes," *Journal of Finance* 29, 413–433.

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.

Schraeder, Stefanie, 2015, "Information Processing and Non-Bayesian Learning in Financial Markets," *Review of Finance* 20, 823–853.

Shiller, Robert J. *Irrational Exuberance* (second edition). (Princeton University Press, Princeton, NJ, 2005).

Timmermann, Allan G, 1993, "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *Quarterly Journal of Economics*, 1135–1145.

Van Binsbergen, Jules, Michael Brandt, and Ralph Koijen, 2012, "On the Timing and Pricing of Dividends," *American Economic Review* 102, 1596–1618.

Van Binsbergen, Jules H, and Ralph SJ Koijen, 2017, "The term Structure of Returns: Facts and Theory," *Journal of Financial Economics* 124, 1–21.

Vissing-Jorgensen, Annette, "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions," in *NBER Macroeconomics Annual*.

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 described 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 when 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, it's advantage is that it is based on high-quality administrative data. As long as the share of total aggregate dividends received by households doesn't 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 per-capita GDP growth rates from Barro and Ursua (2008) (which are in turn based on Balke and Gordon (1989)) as proxy for Δd from 1871 to 1900.
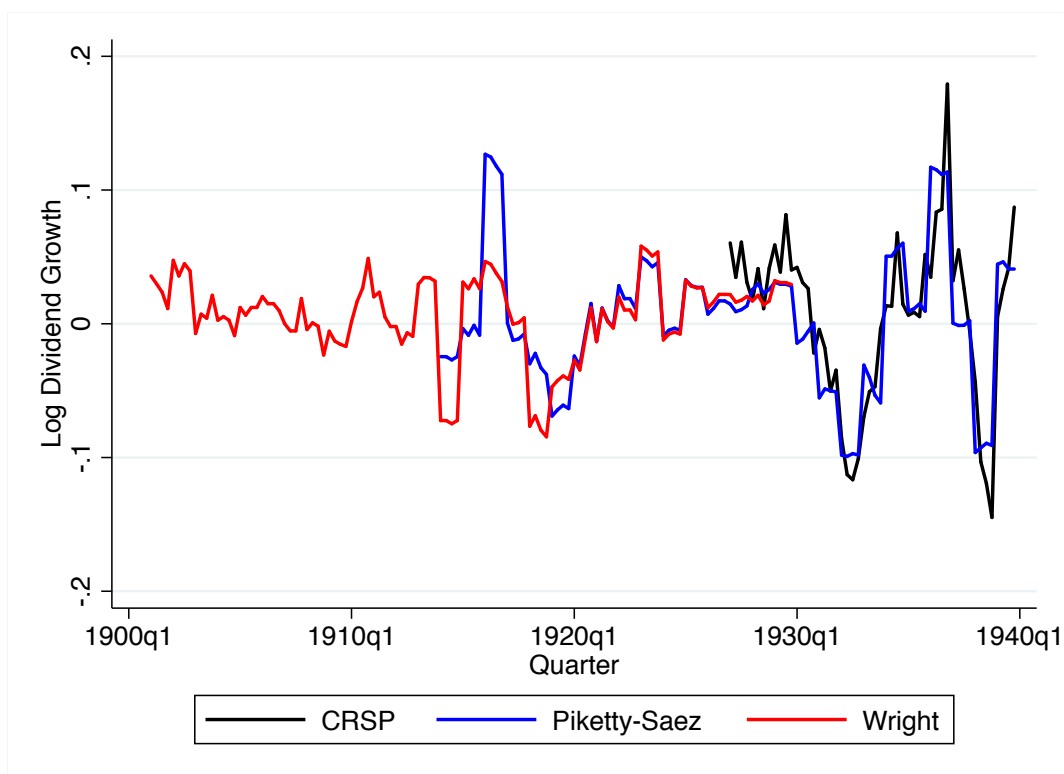
FIGURE A.1
Dividend growth from three data sources

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. Subjective 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 Michigan Survey of Consumers (MSC). The MSC elicits the perceived probability that an investment in a diversified 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 stock (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 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.

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.
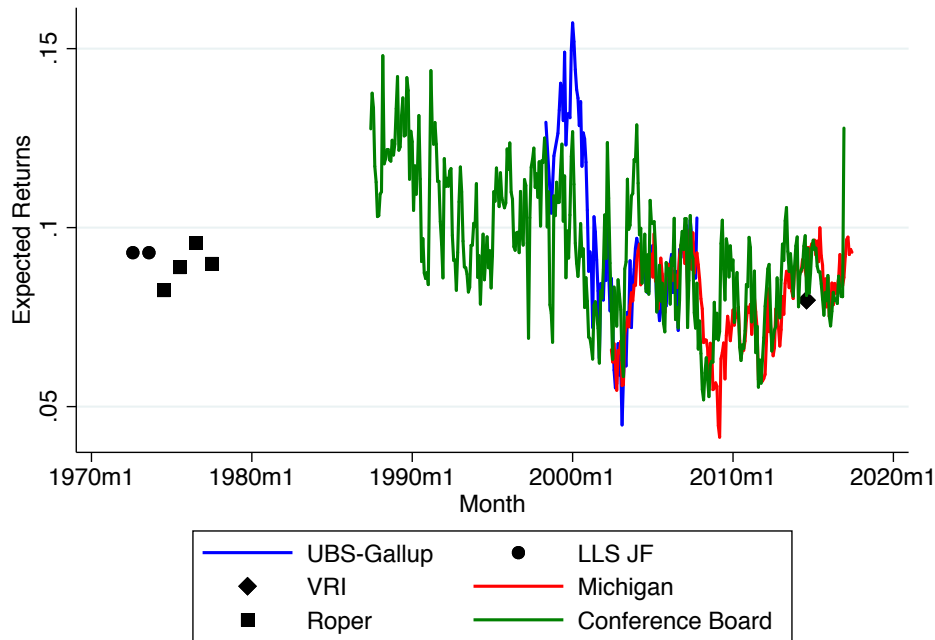
FIGURE A.2
Expected return imputation

.

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 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 $\tilde{\varepsilon}_{t+j}$, $j = 1, 2, ...$ 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 autocovari-

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

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

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

Because of the constant autocovariance structure of perceived consumption growth,

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

and

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

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 $\tilde{\varepsilon}_{t+i}$,

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

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

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

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

where the agent knows σ_ξ^2 and σ_ζ^2, but not μ_t. With diffuse prior and an infinite history, H_t, of observed data on Δc, the predictive distribution can be obtained from the steady-state Kalman filter (see, e.g., Hamilton (1994)) as

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

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

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

with

$$K = \frac{\omega^2 + \sigma_\zeta^2}{\omega^2 + \sigma_\zeta^2 + \sigma_\xi^2} \tag{A.11}$$

and

$$\omega^2 = K\sigma_\xi^2. \tag{A.12}$$

Thus the predictive distribution of Δc_{t+1} at time t is

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

To map into our fading memory setup, we choose

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

The time-t predictive distribution for Δc_{t+1} then is exactly the same as in our fading memory setting. The time-t predictive distribution for Δc_{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 the

predictive distribution converges to a stationary one at long horizons. For pricing, however, this difference in the perceived distribution for $j > 1$ doesn't 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

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

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

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

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

with

$$K = h\frac{\sigma_\zeta^2 + h^2\omega^2}{\sigma_\zeta^2 + \sigma_\xi^2 + h^2\omega^2}, \tag{A.18}$$

$$\omega^2 = K\sigma_\xi^2/h. \tag{A.19}$$

Iterating yields

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

The predictive distributions are

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

and

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

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

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

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

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

where $v_t^1 = \log(V_t/C_t)$ and V_t^1 is the continuation value. We conjecture the solution to be linear in the state variable, i.e.

$$v_t^1 = \mu_v + U_v \tilde{\mu}_t. \tag{A.25}$$

Plugging in the conjectured solution we get

$$U_v = \frac{\delta}{1-\delta}, \tag{A.26}$$

and

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

We obtain the log SDF

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

where

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

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

E.2. Consumption claim valuation

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

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

and in logs,

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

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

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

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

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

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 Δc_{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 ξ 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

To get the expected returns of dividend strips, we start with (32) to compute returns. For the one-period claim, we get

$$r_{t+1}^1 = \lambda\Delta c_{t+1} - (\lambda - 1)\tilde{\mu}_t - \tilde{\mu}_m - \frac{1}{2}\left(\sqrt{1+\nu}\lambda - \xi\right)^2\sigma^2. \tag{A.38}$$

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

$$r_{t+1}^1 - r_{f,t} = \lambda(\Delta c_{t+1} - \tilde{\mu}_t) - \frac{1}{2}\left(\sqrt{1+\nu}\lambda - \xi\right)^2\sigma^2 + \frac{1}{2}\xi\sigma^2. \tag{A.39}$$

The subjective conditional variance of r_{t+1}^1 is $(1+\nu)\lambda^2\sigma^2$, and so, after taking subjective expectations of (A.39), we obtain

$$\log\tilde{\mathbb{E}}_t[R_{t+1}^1] - r_{f,t} = \lambda\xi\sqrt{1+\nu}\sigma^2. \tag{A.40}$$

The objective conditional variance of r_{t+1}^1 is only $\lambda^2\sigma^2$, and so taking objective expectations of (A.39) yields,

$$\log\mathbb{E}_t[R_{t+1}^1] - r_{f,t} = \lambda\xi\sqrt{1+\nu}\sigma^2 - \frac{1}{2}\nu\lambda^2\sigma^2 + \lambda(\mu - \tilde{\mu}_t). \tag{A.41}$$

For the infinite-horizon claim, again starting from (32), we get

$$r_{t+1}^\infty = \Delta c_{t+1} + \frac{\lambda-1}{\alpha}(\tilde{\mu}_{t+1} - \tilde{\mu}_t) - \tilde{\mu}_m - \frac{1}{2}\left[\sqrt{1+\nu}\left(1+\nu\frac{\lambda-1}{\alpha}\right) - \xi\right]^2\sigma^2, \tag{A.42}$$

and, after subtracting the risk-free rate,

$$r_{t+1}^\infty - r_{f,t} = \Delta c_{t+1} + \frac{\lambda-1}{\alpha}(\tilde{\mu}_{t+1} - \tilde{\mu}_t) - \tilde{\mu}_t - \frac{1}{2}\left[\sqrt{1+\nu}\left(1+\nu\frac{\lambda-1}{\alpha}\right) - \xi\right]^2\sigma^2 + \frac{1}{2}\xi^2\sigma^2. \tag{A.43}$$

The subjective conditional variance of r_{t+1}^1 is $(1+\nu)\left(1+\nu\frac{\lambda-1}{\alpha}\right)^2\sigma^2$ and therefore, after taking subjective expectations of (A.43), we obtain

$$\log\tilde{\mathbb{E}}_t[R_{t+1}^\infty] - r_{f,t} = \left[1+\nu\frac{\lambda-1}{\alpha}\right]\xi\sqrt{1+\nu}\sigma^2. \tag{A.44}$$

The objective conditional variance of r_{t+1}^∞ is only $\left(1+\nu\frac{\lambda-1}{\alpha}\right)^2\sigma^2$, and so taking objective expectations of (A.43) yields

$$\log\mathbb{E}_t[R_{t+1}^\infty] - r_{f,t} = \left[1+\nu\frac{\lambda-1}{\alpha}\right]\xi\sqrt{1+\nu}\sigma^2 - \frac{1}{2}\nu\left(1+\nu\frac{\lambda-1}{\alpha}\right)^2\sigma^2 + \left(1+\nu\frac{\lambda-1}{\alpha}\right)(\mu-\tilde{\mu}_t). \tag{A.45}$$

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 (32) implies that

$$P_t^n \approx C_t e^{\mu_{dc} + \frac{1}{2}\frac{1}{1-(1-\alpha)^2}\sigma_d^2 + \frac{\lambda-1}{\alpha}\tilde{\mu}_t} \exp(n\tilde{\mu}_m + \frac{1}{2}A_n\sigma^2), \quad n > J, \tag{A.46}$$

where we approximate

$$A_n \approx A_J + (n-J)[\sqrt{1+\nu}(\nu\frac{\lambda-1}{\alpha} + 1) - \xi]^2, \quad n > J. \tag{A.47}$$

64

We can show that

$$P_t \approx \left(\sum_{n=1}^{J} P_t^n \right) + C_t V_J \exp\left(\mu_{dc} + \frac{1}{2} \frac{1}{1-(1-\alpha)^2} \sigma_d^2 + \frac{\lambda-1}{\alpha} \tilde{\mu}_t \right) \tag{A.48}$$

with

$$V_J = \frac{\exp\left((J+1)\tilde{\mu}_m + \frac{1}{2} A_J \sigma^2 + \frac{1}{2}[\sqrt{1+\nu}(\nu\frac{\lambda-1}{\alpha}+1) - \xi]^2 \sigma^2 \right)}{1 - \exp\left(\tilde{\mu}_m + \frac{1}{2}[\sqrt{1+\nu}(\nu\frac{\lambda-1}{\alpha}+1) - \xi]^2 \sigma^2 \right)}. \tag{A.49}$$

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 $J \approx 5,000$.

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

$$\log \frac{W_t - C_t}{C_t} = \log \frac{\delta}{1-\delta}, \tag{A.50}$$

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.

$$\log \frac{P_t}{D_t} = H(\tilde{\mu}_t, d_t - c_t). \tag{A.51}$$

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

$$\psi_{ij}(\tilde{\mu}, d_t - c_t) \equiv \Lambda_i(\tilde{\mu})\Lambda_j(d_t - c_t), \tag{A.52}$$

where Λ_i denotes the Chebyshev polynomials. We will approximate the log P/D ratio as

$$\hat{H}(\tilde{\mu}, d_t - c_t; \boldsymbol{\beta}_m) = \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} \beta_{m,ij} \psi_{ij}(\tilde{\mu}, d_t - c_t) \tag{A.53}$$

Rewrite the subjective Euler equation

$$\tilde{\mathbb{E}}_t[M_{t+1} R_{m,t+1}] = 1 \tag{A.54}$$

as

$$
\begin{aligned}
0 &= I(\tilde{\mu}_t, d_t - c_t) \\
&\equiv \tilde{\mathbb{E}}_t[e^{\tilde{\mu}_m - \tilde{\mu}_t - \xi\sigma\epsilon_{t+1} + \Delta d_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}}] - 1 \\
&= e^{\tilde{\mu}_m + (\lambda-1)\tilde{\mu}_t - \alpha(d_t - c_t - \mu_{dc})} \tilde{\mathbb{E}}_t[e^{(\lambda\sqrt{1+\nu}-\xi)\sigma\tilde{\varepsilon}_{t+1} + \sigma_d\eta_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}}] - 1. \quad (A.55)
\end{aligned}
$$

We evaluate the function $I(\tilde{\mu}_t, d_t - c_t)$ on the two-dimensional grid of $\tilde{\mu}_t$ and $d_t - c_t$

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
	Mean	0.008
$RMSE_{pd}$ across simulations	Median	0.008
	Max	0.038

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.55). 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.48) in terms of the RMSE,

$$RMSE_{pd} = \sqrt{\frac{1}{t}\sum_{j=1}^{t}(pd_j^{Analytical} - pd_j^{Projection})^2}, \tag{A.56}$$

where $pd_t^{Analytical}$ is calculated from dividend strip prices as in (A.48). 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

$$\tilde{\mathbb{E}}_t[R_{m,t+1}] = \tilde{\mathbb{E}}_t[e^{\Delta d_{t+1}}\frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}}]$$

$$= e^{\lambda\tilde{\mu}_t - \alpha(d_t - c_t - \mu_{dc})}\tilde{\mathbb{E}}_t[e^{\lambda\sqrt{1+\nu}\sigma\tilde{\varepsilon}_{t+1} + \sigma_d\eta_{t+1}}\frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}}]. \tag{A.57}$$

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 information 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 Δc_{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 Δc_{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

$$\mathcal{T}v(x) = \exp(\eta)v(x), \quad v(\cdot) > 0 \tag{A.58}$$

where

$$\mathcal{T}f(x) = \tilde{\mathbb{E}}^*\big[f(\hat{\mu}_{t+2|t+1}) \exp\left[(1-\gamma)\Delta c_{t+1}\right] |\hat{\mu}_{t+1|t} = x\big]. \tag{A.59}$$

Another random variable of interest is

$$N_{t+1} = \frac{e^{(1-\gamma)\Delta c_{t+1}}v(\hat{\mu}_{t+2|t+1})}{\exp(\eta)v(\hat{\mu}_{t+1|t})}. \tag{A.60}$$

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

Assumption 1.

$$\log \delta + \frac{\eta}{\theta} < 0. \tag{A.61}$$

Assumption 2.

$$\lim_{t\to\infty} \tilde{\mathbb{E}}^*[N_{t+1}v(\hat{\mu}_{t+2|t+1})^{-1/\theta}|\hat{\mu}_{1|0} = x] < \infty. \tag{A.62}$$

Assumption 3.

$$\lim_{t\to\infty} \tilde{\mathbb{E}}^*[N_{t+1}v(\hat{\mu}_{t+2|t+1})^{-1}|\hat{\mu}_{1|0} = x] < \infty. \tag{A.63}$$

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

$$v(x) = \exp\left(\frac{1-\gamma}{\nu - \phi\nu}x\right), \tag{A.64}$$

$$\eta = (1-\gamma)\mu + \frac{1}{2}\left(\frac{1-\gamma}{1-\phi}\right)^2 (1+\phi\nu)\sigma^2. \tag{A.65}$$

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

$$\lim_{t\to\infty} \tilde{\mathbb{E}}^*\left[\exp\left(k_1\Delta c_{t+1} + k_2\hat{\mu}_{t+1|t}\right)\Big|\hat{\mu}_{1|0} = x\right] < \infty \tag{A.66}$$

for some corresponding pair of constants (k_1, k_2).

With Wold representation, we have

$$\hat{\mu}_{t+1|t} = [1 - (h - K)L]^{-1} K\Delta c_t, \tag{A.67}$$

$$\Delta c_{t+1} = [1 - (h - K)L]^{-1} K\Delta c_t + \tau_{t+1}, \tag{A.68}$$

$$\Delta c_t = \left[1 + (1 - hL)^{-1}KL\right]\tau_t \tag{A.69}$$

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

$$\lim_{t\to\infty} \tilde{\mathbb{E}}^*\left[\exp\left(k\left[1 - (h - K)L\right]^{-1} K\left[1 + (1 - hL)^{-1}KL\right]\tau_t\right)\right] < \infty \tag{A.70}$$

or

$$\lim_{t\to\infty} \tilde{\mathbb{E}}^*\left[\exp\left(kK(1 - hL)^{-1}\tau_t\right)\right] < \infty. \tag{A.71}$$

As long as $h < 1$, we have

$$\lim_{t\to\infty} \tilde{\mathbb{E}}^*\left[\exp\left(kK(1 - hL)^{-1}\tau_t\right)\right] \tag{A.72}$$

$$= \lim_{t\to\infty} \tilde{\mathbb{E}}^*\left[\exp\left(kK\sum_{j=0}^{\infty} h^j L^j \tau_t\right)\right] \tag{A.73}$$

$$= \exp\left(\frac{1}{2}k^2 K^2 \sum_{j=0}^{\infty} h^{2j}(1 + \phi\nu)\sigma^2\right) < \infty. \tag{A.74}$$

Finally, Assumption 1 translates to

$$\log\delta + \left(1 - \frac{1}{\psi}\right)\left[\mu + \frac{1}{2}\frac{1 - \gamma}{(1 - \phi)^2}(1 + \phi\nu)\sigma^2\right] < 0 \tag{A.75}$$

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

$$m_{t+1} = \theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1)r_{w,t+1} \tag{A.76}$$

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

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

$$r_{w,t+1} = k_0 + k_1 z_{t+1} - z_t + \Delta c_{t+1} \tag{A.77}$$

$$r_{m,t+1} = k_{d,0} + k_{d,1} z_{d,t+1} - z_{d,t} + \Delta d_{t+1} \tag{A.78}$$

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 variable $\tilde{\mu}_t$

$$z_t = A_0 + A_1 \phi \tilde{\mu}_t, \tag{A.79}$$

$$z_{d,t} = A_{d,0} + A_{d,1} \phi \tilde{\mu}_t + A_{d,2}(d_t - c_t). \tag{A.80}$$

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

$$A_0 = \frac{\log \delta + k_0 + (1-\phi)(1 - k_1 + \nu k_1)\mu A_1 + \frac{1}{2}(1 - 1/\psi + \phi \nu k_1 A_1)^2 \theta(1+\phi\nu)\sigma^2}{1 - k_1}, \tag{A.81}$$

$$A_1 = \frac{1 - 1/\psi}{1 - (1 - \nu + \phi\nu)k_1}, \tag{A.82}$$

and

$$A_{d,0} = \frac{\theta \log \delta + (\theta - 1)k_0 + (\theta - 1)(k_1 - 1)A_0 + k_{d,0} + (1-\phi)(1 - k_{d,1} + \nu k_{d,1})\mu A_{d,1}}{1 - k_{d,1}} \tag{A.83}$$

$$+ \frac{(\theta - 1)k_1 \nu(1-\phi)\mu A_1 + (k_{d,1}A_{d,2} + 1)\alpha\mu_{dc} + \frac{1}{2}(k_{d,1}A_{d,2} + 1)^2 \sigma_d^2}{1 - k_{d,1}} \tag{A.84}$$

$$+ \frac{\frac{1}{2}[\theta - 1 - \theta/\psi + \lambda + (\lambda - 1)k_{d,1}A_{d,2} + \phi\nu(k_{d,1}A_{d,1} + (\theta - 1)k_1 A_1)]^2(1+\phi\nu)\sigma^2}{1 - k_{d,1}}, \tag{A.85}$$

$$A_{d,1} = \frac{\lambda - 1/\psi + (\lambda - 1)k_{d,1}A_{d,2}}{1 - (1 - \nu + \phi\nu)k_{d,1}}, \tag{A.86}$$

$$A_{d,2} = \frac{\alpha}{(1-\alpha)k_{d,1} - 1}. \tag{A.87}$$

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

$$r_{f,t} = -\log \delta + \frac{1}{\psi}\left[\phi\tilde{\mu}_t + (1-\phi)\mu\right] \tag{A.88}$$

$$+ \frac{1}{2}(1-\gamma)\left(\frac{1}{\psi} - \gamma\right)\left(1 + \frac{k_1\phi\nu}{1 - (1-\nu+\phi\nu)k_1}\right)^2 (1+\phi\nu)\sigma^2 \tag{A.89}$$

$$- \frac{1}{2}\left[\left(\frac{1}{\psi} - \gamma\right)\left(1 + \frac{k_1\phi\nu}{1 - (1-\nu+\phi\nu)k_1}\right) - \frac{1}{\psi}\right]^2 (1+\phi\nu)\sigma^2. \tag{A.90}$$

We solve for the log-linearization coefficients by iterating on

$$\bar{z} = A_0 + A_1\phi\mu, \tag{A.91}$$

$$k_1 = \frac{\exp(\bar{z})}{1 + \exp(\bar{z})}, \tag{A.92}$$

$$k_0 = \log(1 + e^{\bar{z}}) - \bar{z}k_1, \tag{A.93}$$

and

$$\bar{z}_d = A_{d,0} + A_{d,1}\phi\mu + A_{d,2}\mathbb{E}[d_t - c_t], \tag{A.94}$$

$$k_{d,1} = \frac{\exp(\bar{z}_d)}{1 + \exp(\bar{z}_d)}, \tag{A.95}$$

$$k_{d,0} = \log(1 + e^{\bar{z}_d}) - \bar{z}_d k_{d,1}. \tag{A.96}$$

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.78) and the risk-free rate as in (A.90). To calculate subjective expected return, we take the subjective expectation of Equation (A.78) to yield

$$\tilde{\mathbb{E}}_t[r_{m,t+1}] = \frac{\phi}{\psi}\tilde{\mu}_t + \tilde{A}_{d,0} \tag{A.97}$$

where

$$\tilde{A}_{d,0} = (k_{d,1} - 1)A_{d,0} + k_{d,0} + (k_{d,1}A_{d,2} + 1)\alpha\mu_{dc} \tag{A.98}$$

$$+ (1 - \phi)\mu\left[k_{d,1}A_{d,1}\phi\nu + (\lambda - 1)k_{d,1}A_{d,2} + \lambda\right]. \tag{A.99}$$

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

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 notation 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 ρ, capital share α and depreciation rate δ. 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

$$r\hat{k} = (x + \rho\hat{k})/\hat{c}. \tag{A.100}$$

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

$$n - \rho < \left(\frac{1}{\alpha} - 1\right)\delta. \tag{A.101}$$

With an annual population growth rate of $n = 0.01$, and ρ 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.