

Asset Pricing with Fading Memory ^{*}

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

Zhengyang Xu[‡]
University of Michigan

April 2018

PRELIMINARY AND INCOMPLETE

Building on recent evidence that lifetime experiences shape individuals' 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 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 growth. Consistent with this theory, we show empirically that a long-run weighted average of past real stock market returns is strongly 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 real returns is also strongly negatively related to subjective expectations errors, in line with the predictions of the model.

^{*}We thank seminar participants at Dartmouth College for comments and The Conference Board for providing data.

[†]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

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 evidence from microdata 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). 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 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) and Malmendier and Nagel (2016) from survey data to pin down the value of this parameter.

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. We avoid this unrealistic outcome by assuming an elasticity of intertemporal substitution of one, which implies a constant consumption-wealth ratio, and by assuming that dividends are cointegrated with consumption.¹ We choose the cointegration parameter to get a realistic equity volatility. The remaining model parameters are similar to common choices in the literature.

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 complications in the endowment process (and the agent’s knowledge of these) in Bansal and Yaron (2004) or time-varying risk aversion built into the agent’s utility as in Campbell and Cochrane (1999). The model further makes predictions about 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 Campbell and Cochrane (1999) and Bansal and Yaron 2004 models, which—as Gormsen (2018) demonstrates—match only the first of these two empirical facts, our model matches both.

The state variable that drives most of the variation in the price-dividend ratio and objec-

1. 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).

tive expected returns in this model is a slow-moving exponentially weighted average of past endowment growth rates, which we label *experienced growth*. Constructing this state variable empirically—for example, to use in return predictability regressions on actual data—is difficult. For example, to construct it at the start of the CRSP sample in 1926, one would need consumption or dividend data stretching back many decades earlier.² Dividend data come with the additional problem that changes in payout policies can generate time-series dynamics that are much more complicated than the simple IID setting that we have in our model. For these reasons, we look for a proxy that sidesteps these issues.

We show within our model that *experienced real returns*—i.e., a slow-moving exponentially weighted average of past real stock returns that is easy to construct empirically—is strongly correlated with experienced growth. As a consequence, future excess stock returns in our model economy are strongly negatively related not only to experienced growth, but also to experienced returns. We show that the same result holds empirically as well. Experienced real stock returns, constructed in the same way as in our model, are a strong predictor 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 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

2. In addition, one would really want stockholder consumption as in Mankiw and Zeldes (1991) and Malloy, Moskowitz, and Vissing-Jørgensen (2009) which is available only for a short period.

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.

With fading memory, our model avoids the arguably unrealistic implication of Bayesian parameter learning models that learning effects disappear deterministically over time as the agent acquires more data. 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.³ 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 is 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. Experienced real returns in our setting also summarize past returns, but with much greater weight on more distant observations in the past. As a consequence, movements of the price-dividend ratio and objective expected returns in our model occur at a lower frequency, consistent with the high persistence of the price-dividend ratio in the data. 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. This creates room for subjective price growth expectations to affect prices. In their model, the

3. See, also, Timmermann (1993) and Lewellen and Shanken (2002) for partial equilibrium models with decreasing-gain learning.

price-dividend ratio moves with an exponentially-weighted average of past price growth, but matching the equity risk premium 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 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 growth rate μ of the log payoffs, c , from stock market investment,

$$\Delta c_t = \mu + \epsilon_t, \tag{1}$$

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

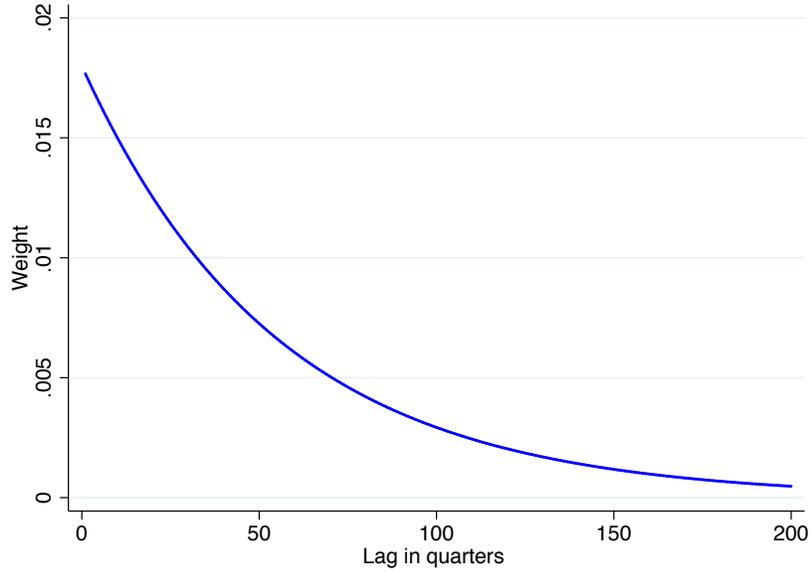


FIGURE I
Weights implied by constant-gain learning

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

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

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

As this expression shows, $\tilde{\mu}$ is updated every period based on the observed surprise $\Delta c_{t+1} - \tilde{\mu}_t$. 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 and

Nagel (2011)). Iterating on (2) one can see that $\tilde{\mu}_t$ is an exponentially-weighted average of past Δc 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.

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

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

$$= \frac{\rho}{1-\rho} (\tilde{\mu}_{t+1} - \tilde{\mu}_t) + \Delta c_{t+1} - \tilde{\mu}_t \quad (4)$$

$$= \left(1 + \frac{\rho\nu}{1-\rho} \right) (\Delta c_{t+1} - \tilde{\mu}_t). \quad (5)$$

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

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

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

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

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

where the term in parentheses times $\mu - \tilde{\mu}_t$ represents the subjective growth-rate expectations revision that the econometrician anticipates, on average, given her knowledge of μ . This expression shows that the econometrician should find returns to be predictable. Specifically, $\tilde{\mu}_t$ 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 $\tilde{\mu}_t$. We can see this by subtracting the subjective equity premium $\tilde{\mathbb{E}}_t r_{t+1} - r_f = \theta$ from (7). We obtain

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

II.A. Experienced returns as proxy for 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 $\tilde{\mu}_t$ and future returns in (7), we want to use a long sample of returns back to the start of the CRSP database in 1926. To compute $\tilde{\mu}_t$ in 1926, we would then want data stretching back at least around 50 years or so, up to the point where the weights become close to negligible. It would be difficult to use consumption data or dividend data for this purpose. Good consumption data is not available going back that far, especially not for stockholder consumption data that we would ideally want to use here (Mankiw and Zeldes

1991; Malloy, Moskowitz, and Vissing-Jørgensen 2009). Dividend data is available, but the time-series dynamics of dividends are influenced by shifts in payout policy that can distort estimates of $\tilde{\mu}$ constructed from dividend growth rates.⁴

For these reasons, we opt for an indirect measurement of experienced growth in stock market fundamentals by taking a weighted average of past real returns instead of some measure of past real payout growth. Expressed recursively, we construct

$$\tilde{\mu}_{r,t+1} = \tilde{\mu}_{r,t} + \nu(r_{t+1} - \tilde{\mu}_{r,t}), \quad (9)$$

where r denotes log real stock market index returns. From the point it becomes available in 1926, we use quarterly returns on the CRSP value-weighted stock market index. Before that, we use data from Shiller (2005) back to 1871 to construct quarterly returns on the S&P Composite index up to 1926. We also use the CPI series in Shiller’s data set to deflate returns.

The experienced return series $\tilde{\mu}_r$ is an imperfect proxy for the experienced fundamentals growth series $\tilde{\mu}$. How well it approximates $\tilde{\mu}$ depends on the gain parameter ν . We simulated dividend growth and returns from (1) and (5) in 1,000 samples of 356 quarters plus 400 quarters as a burn-in period to compute $\tilde{\mu}$ and $\tilde{\mu}_r$ at the start of the estimation sample. For the value $\nu = 0.018$ that we work with here, the correlation is a very high 0.82. Other parameters like μ , 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 $\tilde{\mu}$ should work well. We confirm this again below when we study $\tilde{\mu}_r$ in data simulated from our asset-pricing model.

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

4. See, e.g., Bansal, Dittmar, and Lundblad (2005) for a discussion of issues with payouts and dividends.

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 2015. 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 2002-2015
- Conference Board Survey, monthly 1987-2015⁵
- 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

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

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.1 provides more detail.⁶

II.C. Return Predictability

Table I presents predictive regressions along the lines suggested by (7), but with $\tilde{\mu}_r$ as proxy for $\tilde{\mu}$. The dependent variable is the quarterly return on the CRSP value-weighted index in excess of the three-month T-bill yield. If we had a direct measurement of $\tilde{\mu}_t$ as predictor variable, then (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 (2000)). For the regressions we run with $\tilde{\mu}_r$ as proxy for $\tilde{\mu}$, the simulations that we described in Section II.A yield an average predictive regression coefficient of -1.36 for $\tilde{\mu}_r$. This average coefficient does not change with parameters except for g .

The first row in Table I reports the OLS slope coefficient for the $\tilde{\mu}_r$ predictive variable. 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. The results show that the coefficient we obtain empirically has the predicted sign, but it is somewhat bigger. For the full sample from 1927-2015, we obtain an OLS estimate of -2.31. While it is larger than the coefficient of -1.36 implied by our simulations, the empirical estimate is within about one

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 AAI survey is conducted among members of the AAI and it records responses of members that self-select into participation. Respondents state whether they are “bullish” or “bearish” about the stock market. We prefer the Conference Board survey for this time period as it is based on a representative sample of the U.S. population. For the early part of their sample starting in the 1960s, Greenwood and Shleifer 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.

TABLE I
PREDICTING RETURNS WITH EXPERIENCED REAL RETURNS

Dependent variable is the log return of the CRSP value-weighted index in quarter $t+1$ in excess of the 3-month T-bill yield at the end of quarter t . Experienced real return denotes a long-run exponentially weighted average of quarterly log returns, deflated using the CPI index, leading up to and including quarter t , with weights implied by constant gain learning with quarterly gain $\nu = 0.018$; 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.

	(1)	(2)	(3)	(4)	(5)
	1927-2015	1927-2015	1927-2015	1946-2015	1946-2015
Experienced real returns	-2.31	-2.49	-2.10	-3.22	-2.39
[bias-adj. coeff.]	[-1.68]	[-1.70]	[-1.90]	[-2.23]	[-2.28]
(p -value)	(0.03)	(0.03)	(0.06)	(0.01)	(0.03)
Inflation		-0.55	-0.52	-2.50	-2.77
[bias-adj. coeff.]		[-0.51]	[-0.50]	[-2.44]	[-2.75]
(p -value)		(0.20)	(0.21)	(0.00)	(0.00)
$p - d$			-0.01		-0.03
[bias-adj. coeff.]			[0.00]		[-0.01]
(p -value)			(0.42)		(0.04)
Observations	356	356	356	280	280
Adj. R^2	0.031	0.033	0.032	0.073	0.087

standard error from it. Bias-adjustment shrinks the point estimate to -1.68. Based on the bootstrapped p -value of 0.03, we can reject the null of no-predictability at conventional levels of confidence.

We motivated these predictive regressions based on a framework in which true dividend growth rates are unpredictable. However, there is evidence from Piazzesi and Schneider (2007) (see, also Kung (2015)) for a strong negative link between lagged inflation and future growth. To the extent that inflation is correlated with experienced real returns, omission of lagged inflation from the predictive regression could bias the coefficient on the experienced real return variable. Column (2) therefore adds the average log CPI inflation rate during quarters $t - 3$ to t to the regression. We obtain a negative coefficient for inflation, but the

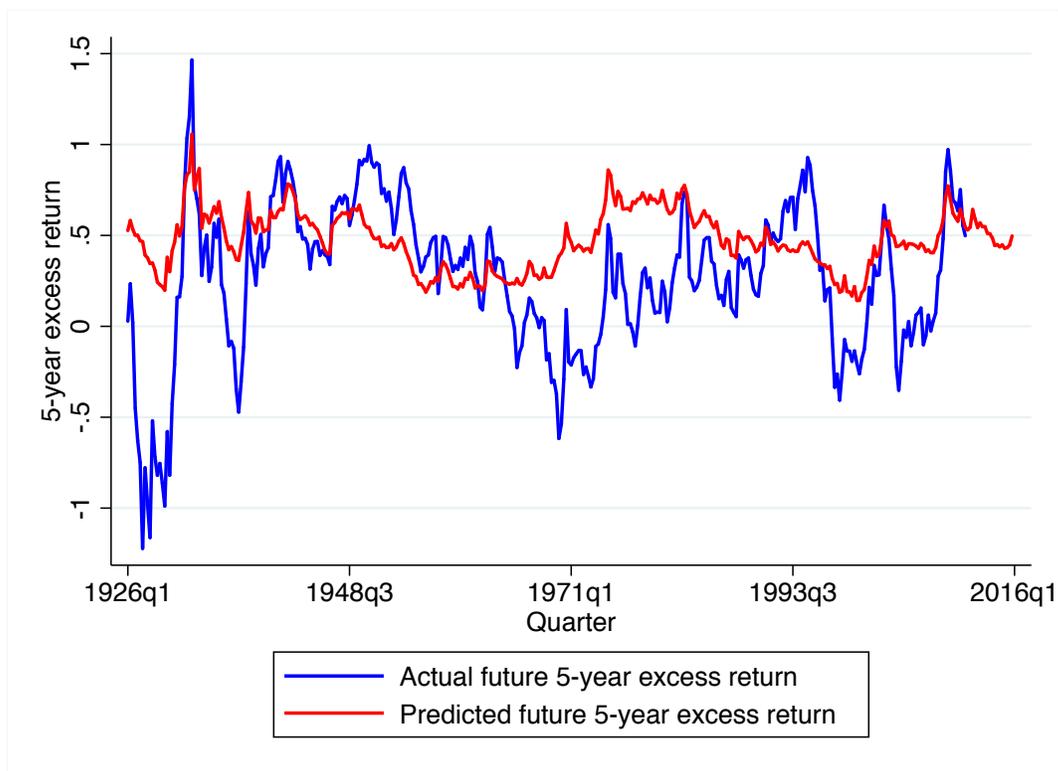


FIGURE II

Predicted five-year excess returns and subsequent actual cumulative five-year excess returns
 Predicted returns are calculated based on bootstrap bias-adjusted coefficients from a predictive regression of quarter $t + 1$ log excess returns on experienced real returns up to quarter t as in column (1) of Table I. The predicted cumulative five-year excess returns refers to the sum of log excess returns on the CRSP value-weighted index in quarters $t + 1$ to $t + 20$, obtained by iterating on the one-quarter forecast using an AR(1) with AR coefficient $1 - \nu = 0.982$.

coefficient for experienced real returns doesn't change much.

In column (3) we add the log price-dividend ratio. It turns out that in the presence of experienced real returns and inflation in the regression, $p - d$ is not a significant predictor and does not raise the R^2 compared with column (2). Columns (4) and (5) re-run the regressions of column (2) and (3) for the post World War II sample to address a potential concern that the results could be driven by the Great Depression period. Comparing the estimates from the two sample periods, there is little change in the estimated coefficient for experienced returns. The effect of experienced returns is consistently negative across both samples.

Figure II shows that the experienced real returns and future excess returns are also strongly correlated at much longer prediction horizons. In this figure, we plot the predicted 5-year excess return based on the bias-adjusted fitted values from column (1) in Table I, and iterating on it using the AR(1) dynamics of the experienced return updating rule (9) with AR coefficient $1 - \nu = 0.982$). For comparison, we then plot the actual future 5-year cumulative excess return 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. Overall, the evidence indicates that there is a negative relationship between experienced real returns and future stock market excess returns.

II.D. Expectation Error Predictability

In the model we have sketched above based on present-value relations, the level of asset prices are 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 later, though not exactly.

To check the time-series relationship between experienced returns and subjective expected excess returns, Panel A of Table II presents regressions of quarter t expected one-year excess returns from surveys on experienced real returns that are calculated based on returns leading up to the end of quarter $t - 1$. As the table shows there is only a weak, and statistically not significant, positive relationship between experienced returns and subjectively expected excess returns. 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

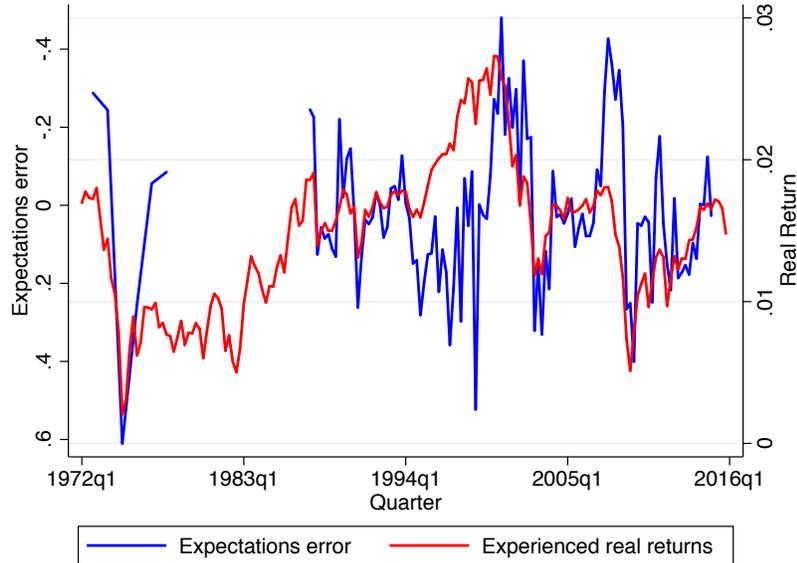


FIGURE III
Experienced real returns and subjective expectation errors

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.

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) shows that when experienced returns and lagged returns are used jointly, the experienced return effect becomes even weaker. Most of what the experienced return captured in column (1) is now absorbed by the lagged one-year return. In terms of the lower frequency movements that we focus on, the subjective equity premium is close to acyclical. The short-run fluctuations in subjective return expectations with one-year lagged returns are also interesting, but not a fact that we try to explain in this paper.

We now turn to the prediction, based on equation (7), that $\tilde{\mu}$, and its proxy $\tilde{\mu}_r$, 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

TABLE II
SURVEY RETURN EXPECTATIONS AND EXPERIENCED REAL RETURNS

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 beginning of quarter t , which we regress on experienced real returns leading up to and including quarter t . Lagged one-year return refers to the return of the CRSP value-weighted index over the four quarters $t - 3$ to t . 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)
<i>Panel A: Subjective expected excess returns</i>			
Experienced real returns	0.95 (0.67)		0.44 (0.80)
Lagged one-year return		0.03 (0.01)	0.03 (0.01)
Constant	0.04 (0.01)	0.05 (0.00)	0.05 (0.01)
Observations	121	121	121
Adjusted R^2	0.041	0.091	0.092
<i>Panel B: Expectation error: Realized - subj. expected</i>			
Experienced real returns	-15.69 (6.48)		-16.01 (7.39)
Lagged one-year return		-0.15 (0.13)	0.02 (0.14)
Constant	0.28 (0.09)	0.04 (0.04)	0.28 (0.10)
Observations	116	116	116
Adjusted R^2	0.138	0.013	0.131

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 returns combined with the fact in Table I that future returns are negatively related to experienced returns implies that the expectations error should be negatively related to experienced returns. However, since the survey data is restricted to the 1970s and 1987-2015, the samples in Table I and II cover very different sample periods. For this reason, it is still useful to check whether there is actually a negative relationship in the sample in which survey data is available.

Panel B of Table II shows that this is the case. There is a strong negative relationship between experienced returns 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 returns and expectations errors (the expectations errors are plotted on a reversed scale).

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

$$\Delta c_{t+1} = \mu + \sigma \varepsilon_{t+1}, \tag{10}$$

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, $H_t \equiv \{\Delta c_0, \Delta c_1, \dots, \Delta c_t\}$, to form an estimate of μ .

After seeing the data H_t , a Bayesian agent 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, but still retains the modeling of the full posterior distribution as in the Bayesian approach. We introduce memory decay using a weighted-likelihood approach that has been used in the theoretical biology literature to model memory decay in organisms (Mangel 1990). The posterior is formed as

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

where $1 - \nu$ is a positive number close to one. Thus, $(1 - \nu)^j$ represents a (geometric) weight on each observation. This weighting scheme assigns smaller weights the more the observation recedes into the past. With a prior

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

we obtain the posterior

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

where

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

The variance of the posterior is the same as if the agent had observed, and retained fully in memory with equal 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. For now, however, we work with an uninformative prior ($\sigma_0 \rightarrow \infty$) and hence the posterior

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

We will return to the informative prior case when we consider versions of the model that generalize our baseline assumption about the elasticity of intertemporal substitution.

With an uninformative prior, $\tilde{\mu}_t$ is the posterior mean and it can be obtained recursively through the updating scheme

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

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

We further get the predictive distribution

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

where the variance of the predictive distribution reflects not only the uncertainty due to future 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 (16) as

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

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

The dynamics of $\tilde{\mu}$ in (18) may look like a (non-stationary) martingale process (e.g., like the posterior mean growth rate in the Bayesian learning model of Collin-Dufresne, Johannes, and Lochstoer (2016)), but it is not. Like in the case of a martingale, we have $\tilde{\mathbb{E}}_t[\tilde{\mu}_{t+1}] = \tilde{\mu}_t$, but unlike in the case of a martingale, the shocks $\tilde{\varepsilon}_{t+j}$, $j = 1, 2, \dots$, are negatively serially correlated under the agent’s predictive distribution at t , as shown in Appendix C. 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 best (in a mean-square sense) forecast. The agent also knows that $\tilde{\mu}_{t+j}$ in future periods will be formed based on information that is different, but no more informative than the information available to the agent at time t . As a consequence, unlike in Bayesian learning without memory loss, $\tilde{\mu}_{t+j}$ in future periods does not have a tendency to converge to the true μ . 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 induced by noise, not convergence to μ . Hence, the agent anticipates mean-reversion towards $\tilde{\mathbb{E}}_t[\tilde{\mu}_{t+j}] = \tilde{\mu}_t$, i.e., towards the current best guess of μ .⁷

The fading memory setting also raises a fundamental issue about asset valuation. Consider

7. Under Bayesian learning with full memory, the agent’s information is represented by a filtration and posterior beliefs follow a martingale under this filtration. In contrast, here memory loss implies that information is not represented by a filtration. For this reason, $\tilde{\mathbb{E}}_t[\tilde{\mu}_{t+1}] = \tilde{\mu}_t$ does not imply that $\{\tilde{\mu}_t\}$ is a martingale. Beliefs are stationary. While $\tilde{\mu}_t$ and hence the agent’s posterior mean of μ varies from period to period, the learning with constant gain (and implied memory loss) implies that the precision of the agent’s information about μ does not change over time. This can also be seen from the fact that the posterior variance in (15) is constant.

the valuation at date t of a claim to consumption at date $t + 2$. One way of valuing the asset would be a “buy-and-hold” valuation, where the agent values the asset based on the stochastically discounted payoff under the time- t predictive distribution

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

where we use $M_{t+j|t}$ to denote the one-period SDF from $t + j - 1$ to $t + j$ based on the predictive distribution at t . An alternative that we could call a “resale” valuation assumes that the agent looks one period ahead at the value and then prices the asset under the time- t predictive distribution of the stochastically discounted next-period asset value. This would lead to the value

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

In the standard Bayesian setting, the law of iterated expectations (LIE) would apply in the valuation equation of $P_{R,t}$ with the result that $P_{R,t} = P_{H,t}$. However, under subjective beliefs formed by learning with loss of memory, the LIE does not generally hold. For example, the agent’s information set at t is not coarser than the agent’s information set at $t + 1$. The two information sets are different, but equal in terms of how coarse they are. In other words, the agent’s information structure is not a filtration and the LIE typically fails to hold.⁸ As a consequence $P_{R,t} \neq P_{H,t}$.

The valuation discrepancy arises because the agent at t and at $t + 1$ see the statistical properties of the shock $\tilde{\varepsilon}_{t+2}$ differently. To the agent at $t + 1$, after having lost some of the memory that shaped the subjective belief at t , $\tilde{\varepsilon}_{t+2}$ looks unpredictable and she prices the asset accordingly. In contrast, to the agent at t it appears that $\tilde{\varepsilon}_{t+2}$ is negatively serially correlated with $\tilde{\varepsilon}_{t+1}$. The buy-and-hold valuation incorporates this serial correlation. In contrast, the resale valuation is based on the anticipation that the value of the asset at date

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., $\tilde{\mathbb{E}}_t \tilde{\mathbb{E}}_{t+1} [\exp(a + b\Delta c_{t+2})] \neq \tilde{\mathbb{E}}_t \tilde{\mathbb{E}}_{t+1} [\exp(a + b\Delta c_{t+2})]$.

$t + 1$ will be determined by an agent—or a future self of the agent—who perceives $\tilde{\varepsilon}_{t+2}$ as unpredictable.

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.

III.B. Stochastic Discount Factor

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

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

where δ denotes the time discount factor, γ relative risk aversion and ψ the elasticity of intertemporal substitution (EIS). Note that the agent evaluates the continuation value under her subjective expectations $\tilde{\mathbb{E}}_t[\cdot]$.

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,

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

with

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

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

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

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

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

III.C. Pricing the consumption claim

We can now solve for the consumption-wealth ratio and the subjective risk premium for the consumption claim. Let $\zeta \equiv W_t/C_t$ and define the return on the consumption claim $R_{W,t+1} \equiv \frac{W_{t+1}}{W_t - C_t} = \frac{C_{t+1}}{C_t} \frac{\zeta}{\zeta - 1}$. Taking logs of $R_{W,t+1}$, and using (18), we obtain

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

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

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

Thus, similar to the standard rational expectations case, $\psi = 1$ implies a constant and finite consumption-wealth ratio. In the posterior distribution in (15), 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

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

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

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

where the time-varying wedge $\mu - \tilde{\mu}_t$ reflects the disagreement between the econometrician and the agent about the conditional expectation of $R_{w,t+1}$. The wedge is observable to the econometrician who knows 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.D. Pricing the Dividend Claim

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

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

similar to Bansal, Gallant, and Tauchen (2007). We assume that 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

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

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 (32), we obtain

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

where

$$A_n = \sum_{k=0}^{n-1} \left\{ \sqrt{1 + \nu} \left[\nu(\lambda - 1) \frac{1 - (1 - \alpha)^k}{\alpha} + (\lambda - 1)(1 - \alpha)^k + 1 \right] - \xi \right\}^2, \quad (34)$$

and

$$B_n = \frac{1 - (1 - \alpha)^{2n}}{1 - (1 - \alpha)^2}. \quad (35)$$

For very large n , approximately,

$$A_n \approx n \left[\sqrt{1 + \nu} \left(1 + \nu \frac{\lambda - 1}{\alpha} \right) - \xi \right]^2 \quad (36)$$

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 (33) to be (weakly) negative. Using (36), we see that this requires

$$\tilde{\mu}_m + \frac{1}{2} \left[1 + \sqrt{1 + \nu} \nu \frac{\lambda - 1}{\alpha} - \xi \right]^2 \sigma^2 \leq 0. \quad (37)$$

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

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

$$P_t = \sum_{n=1}^{\infty} P_t^n. \quad (38)$$

Having the restriction (37) hold as a strict inequality ensures that this sum remains finite. We compute the sum in (38) numerically using the analytical solution for dividend strips. Details are in Appendix D.4.

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

$$R_{t+1}^{\infty} \equiv \lim_{n \rightarrow \infty} P_{t+1}^{n-1} / P_t^n. \quad (39)$$

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

$$\log \tilde{\mathbb{E}}_t [R_{t+1}^{\infty}] - r_{f,t} = \left[1 + \nu \frac{\lambda - 1}{\alpha} \right] \xi \sqrt{1 + \nu} \sigma^2. \quad (40)$$

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

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

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

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 returns. 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.

For comparison, a short term claim with return $R_{t+1}^1 \equiv D_{t+1}/P_t^1$, has the subjective risk premium

$$\log \tilde{\mathbb{E}}_t[R_{t+1}^1] - r_{f,t} = \lambda \xi \sqrt{1 + \nu \sigma^2}. \quad (42)$$

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

$$\tilde{\theta}_t^{\infty,1} = (\lambda - 1) \left(\frac{\nu}{\alpha} - 1\right) \xi \sqrt{1 + \nu \sigma^2}. \quad (43)$$

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

Based on these results we see that the subjective risk premium is constant at short and long horizons. This does not imply, however, 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 D.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.E. Calibration

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 μ_{dc} such that the mean of $d_t - c_t$ equals the log of the ratio of the sum of 12-month dividends on the CRSP value-weighted index and annual nominal consumption expenditure for nondurables and services (NIPA, obtained from the Federal Reserve Bank of St. Louis FRED database). In our baseline specification, we fix $\psi = 1$ to obtain a solution without state-space truncation, 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 (37) required for a finite price of the dividend claim and it implies that the forces of cointegration 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)).

III.F. 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
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
<i>Belief updating</i>			
Gain	ν	0.018	MN (2016) (survey data)
<i>Endowment process</i>			
Leverage ratio	λ	3	CJL (2017)
Dividend cointegration parameter	α	0.001	
Mean log dividend-consumption ratio	$\mu_{dc} + \frac{(\lambda-1)\mu}{\alpha}$	-4.7	Data 1947 - 2015
Mean consumption growth	μ	0.45%	CJL (2017)
Consumption growth volatility	σ	1.35%	CJL (2017)
Dividend growth volatility	σ_d	5%	CJL (2017)
<i>Preferences</i>			
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 $\tilde{\mu}$ and $\tilde{\mu}_r$ at the start of each sample. The first column shows the corresponding empirical moments from the data for the 1927 to 2015 period (with the exception of consumption growth, where NIPA data starts in 1947 and we compute growth rates starting 1947Q2). We compute $d - c$ from nominal nondurables and services consumption and nominal dividends of the CRSP value-weighted index. For experienced real returns we use data from 1871 to 1926 as pre-sample information to calculate $\tilde{\mu}_r$ at the start of the sample in 1927. 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 $\tilde{\mu}_r$ we multiply means and standard deviations by four. We estimate the empirical moments of Δc from four-quarter changes and those of Δd from one-year changes in the log of trailing 12-months dividends on the CRSP value-weighted index.

	Data 1927-2015	Model
$\mathbb{E}(\Delta c)^*$	1.91	1.80
$\sigma(\Delta c)^*$	1.37	2.70
$\mathbb{E}(\Delta d)$	2.83	1.80
$\sigma(\Delta d)$	12.66	12.87
$\mathbb{E}(d - c)^*$	-4.70	-4.70
$\mathbb{E}(R_m - R_f)$	7.98	7.27
$\sigma(R_m - R_f)$	22.54	16.52
$SR(R_m - R_f)$	0.35	0.44
$\mathbb{E}(p - d)$	3.40	2.87
$\sigma(p - d)$	0.44	0.16
$\rho(p - d)$	0.97	0.90
$\mathbb{E}(r_f)$	0.46	1.64
$\sigma(r_f)$	2.81	0.51
$\sigma(\tilde{\mu})$	-	0.51
$\rho(\tilde{\mu})$	-	0.98
$\sigma(\tilde{\mu}_r)$	2.29	2.56
$\rho(\tilde{\mu}_r)$	0.94	0.97
$\text{corr}(\tilde{\mu}, \tilde{\mu}_r)$	-	0.62

* Empirical moments from 1947 - 2015 data.

empirical moments for the 1927 to 2015 period for comparison. The model produces a high equity premium (7.27%), equity volatility (16.52%), and Sharpe Ratio (0.44), close to the empirical estimates in the first column. The log price-dividend ratio is also quite volatile, although its standard deviation (0.16) is not quite as high as in the empirical data (0.44)

The model also produces a low real risk-free rate with low volatility. For the purpose of this moments comparison, we calculate the real risk-free rate in the data using as inflation expectation the mean forecast from the Michigan Survey of Consumers from 1978 to 2015. Prior to 1978 when the Michigan Survey data is not available, we use 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 the Michigan Survey quite well. The volatility of r_f in the data (2.81%) is higher than in the model (0.51%), but one should keep in mind that the inflation expectations are measured 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) specifically reverse-engineer their model to produce a constant risk-free rate). Overall, the model provides a reasonable fit to the main asset pricing moments.

The standard deviation of $\tilde{\mu}$ is the same as that of r_f . The persistence of $\tilde{\mu}$ equals $1 - \nu$, i.e., it is pinned down by the gain parameter. The bottom of the table shows the correlation of $\tilde{\mu}$ and $\tilde{\mu}_r$, the experienced real return. In line with our conjecture in Section II.A, we find a strong positive correlation of $\tilde{\mu}$ with $\tilde{\mu}_r$ of 0.62 and we use $\tilde{\mu}_r$ as a proxy for $\tilde{\mu}$ below when we compare the predictability predictions of the model to the data.

III.G. Predictability of Returns and Subjective Expectations Errors

We now evaluate the time-variation in the objective equity premium implied by the model. Time-variation in the objective equity premium due to subjective belief dynamics is one of the most important features of the model.

Table V presents the results from predictive regressions of log excess returns on the equity

TABLE V
PREDICTIVE REGRESSIONS IN SIMULATED DATA

This table reports return predictability regression coefficients obtained as means and medians across 10,000 simulations of the model for 356 quarters plus a 400-quarter burn-in period to compute $\tilde{\mu}$ and $\tilde{\mu}_r$ at the start of each simulated sample. The dependent variable is the log excess return on the equity claim and $\tilde{\mu}_r$ is constructed as the exponentially-weighted average of experienced log returns, with the gain parameter $\nu = 0.018$.

		1Q			1Y			5Y		
	coef.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\tilde{\mu}_r$	mean	-1.23		-1.08	-4.75		-4.18	-19.82		-17.35
	median	-1.09		-0.93	-4.23		-3.62	-18.59		-16.35
$p - d$	mean		-0.04	-0.02		-0.16	-0.08		-0.66	-0.31
	median		-0.04	-0.02		-0.15	-0.08		-0.65	-0.34

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. Columns (1) to (3) present mean and median slope coefficients on the predictors $\tilde{\mu}_r$ and $p - d$ from regressions with a one-quarter prediction horizon. The means and medians for $\tilde{\mu}_r$ are somewhat higher than in Table I, those for $p - d$ are about the same. Overall, given the substantial statistical uncertainty in the empirical regression estimates, the model and data are remarkably consistent.

Columns (4) to (9) show results for a one-year and five-year prediction horizon, respectively. The results show that the predictive relationship is very persistent: the long-run regression coefficients are only slightly smaller than the forecast horizon (in quarters) times the coefficients in column (1) to (3). This is a consequence of the persistence of the state variable $\tilde{\mu}$ (and its proxy $\tilde{\mu}_r$).

To replicate the subjective expectations error regressions that we ran on the empirical data, we need to calculate the subjective equity premium in our model. As we showed in Section III.D, the subjective premia on dividend strips are constant, but the subjective premium for the claim to the whole stream of dividends is not. We compute the subjective equity premium numerically, following the method of Pohl, Schmedders, and Wilms (2018).

TABLE VI
SURVEY RETURN EXPECTATIONS AND EXPERIENCED REAL RETURNS IN SIMULATIONS

This table reports the results of regressing subjective expected excess returns and expectation errors on $\tilde{\mu}_r$ and 1-year lagged log returns from 10,000 simulations of 356 quarters with a 400-quarter burn-in period. 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 $\prod_{i=1}^4 R_{m,t+i} - (\tilde{\mathbb{E}}_t[R_{m,t+1}]^4)$.

	coef.	(1)	(2)	(3)
<i>Panel A: Subjective expected excess returns</i>				
$\tilde{\mu}_r$	mean	0.34		0.35
	median	0.34		0.35
$r_{t-3,t}$	mean		0.004	-0.001
	median		0.004	-0.001
R_{adj}^2	mean	0.44	0.05	0.44
	median	0.45	0.05	0.46
<i>Panel B: Expectation error: Realized - subj. expected</i>				
$\tilde{\mu}_r$	mean	-5.25		-5.63
	median	-4.66		-5.01
$r_{t-3,t}$	mean		-0.042	0.029
	median		-0.041	0.029
R_{adj}^2	mean	0.03	0.01	0.03
	median	0.02	0.00	0.02

Appendix D.4 provides more detail.

Table VI is the model-implied 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 (3) of Table II is 0.44, while the regression on model-generated data in column (3) of Table VI yields a coefficient of 0.35. In Panel B, with subjective expectations errors as dependent variable, the coefficients on experienced returns are strongly negative, just like in the empirical data in Table VI. In terms of magnitude, the model-implied coefficients are not quite as big as the empirical ones, however, but the latter also come with a fairly big standard error.

III.H. 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 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 356-quarter sample period with a 400-quarter burn-in period to compute $\tilde{\mu}$ and $\tilde{\mu}_r$ at the start of each sample. Within each 356-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 $\tilde{\mu}_r$ 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 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 356-quarter sample, use the backward looking history from the start of the 356-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 predictive regression forecast underperforms the trailing sample mean as a forecaster. That

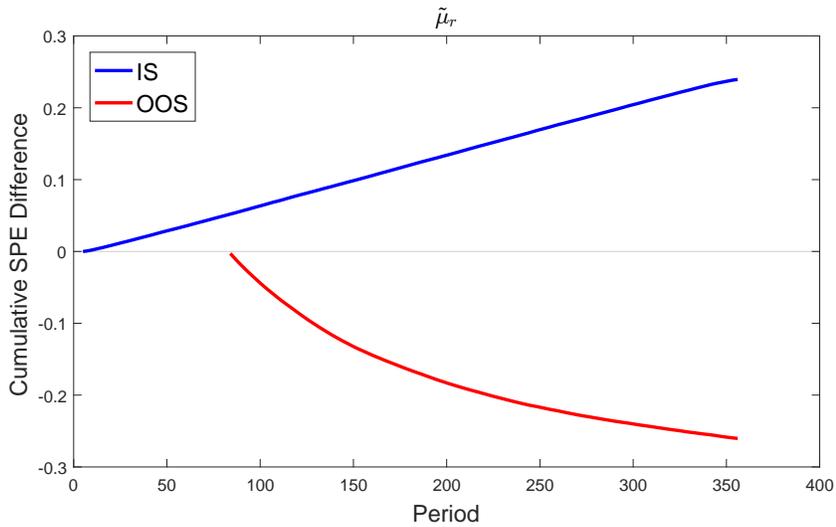
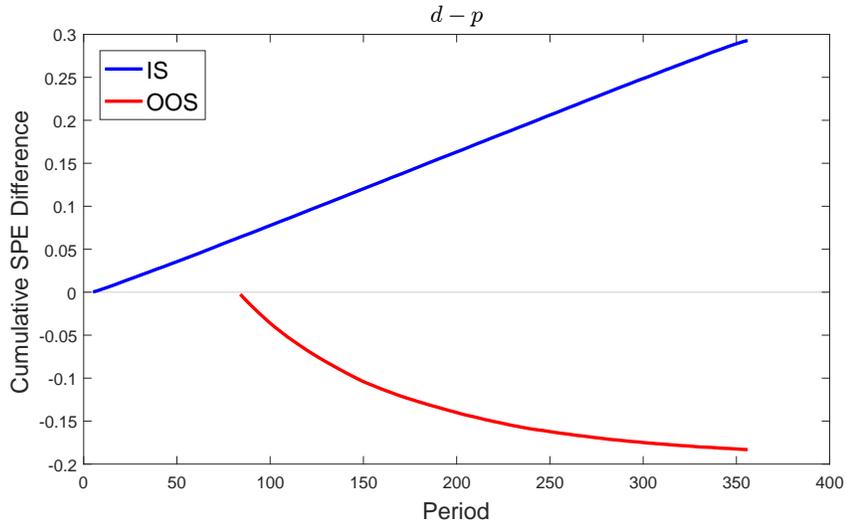


FIGURE IV
Out-of-sample predictive performance

In-sample and out-of-sample performance of predictors $d-p$ and $\tilde{\mu}_r$ from 10,000 simulations of 356 quarters with a 400-quarter burn-in period to compute $\tilde{\mu}_r$ 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.

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 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 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.

III.I. Objective Equity Term Structure

As Van Binsbergen, Brandt, and Kojien (2012) and Van Binsbergen and Kojien (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 D.3, the objective conditional equity term premium earned by the infinite-horizon claim in our model is

$$\begin{aligned} \log \mathbb{E}_t[R_{t+1}^\infty] - \mathbb{E}_t[R_{t+1}^1] &= (\lambda - 1) \left[\frac{\nu}{\alpha} - 1 \right] \xi \sqrt{1 + \nu} \sigma^2 - \frac{1}{2} \nu \left[\left(1 + \nu \frac{\lambda - 1}{\alpha} \right)^2 - \lambda^2 \right] \sigma^2 \\ &\quad + (\lambda - 1) \left[\frac{\nu}{\alpha} - 1 \right] (\mu - \tilde{\mu}_t). \end{aligned} \quad (44)$$

On average, the slope is positive in our baseline calibration, as shown by the solid line in Figure V. This prediction seems to be in conflict with the empirical evidence of an unconditionally negative slope in Van Binsbergen, Brandt, and Kojien (2012) and Van Binsbergen and Kojien (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

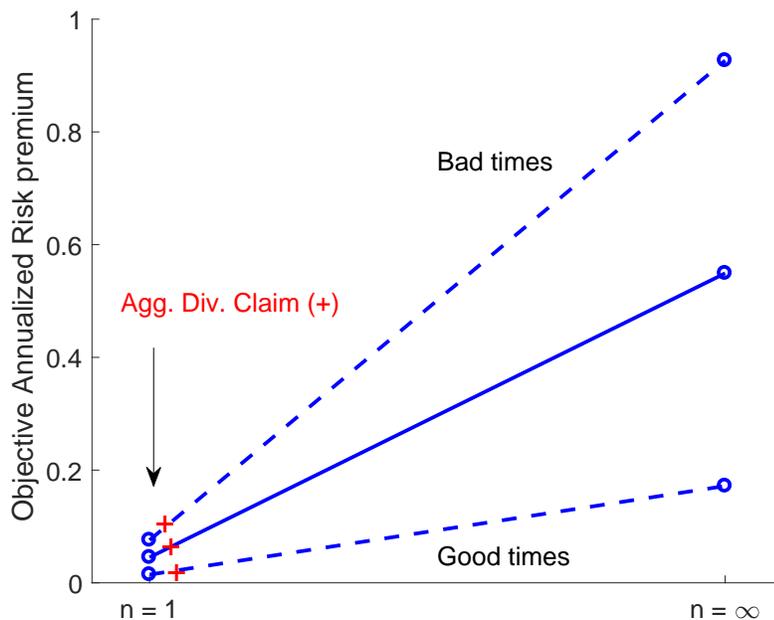


FIGURE V
Objective Equity Term Structure

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).

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 (44) 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.D.

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 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

$$e_t^n = \frac{1}{n}(d_t - f_t^n), \quad (45)$$

where $f_t^n = p_t^n + ny_t^n$ 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

$$s_t^n \equiv e_t^n - e_t^1 = \frac{1-n}{n}d_t - \frac{1}{n}p_t^n + p_t^1 - y_t^n + y_t^1, \quad (46)$$

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 (33), we obtain, by taking the appropriate limit as $n \rightarrow \infty$,

$$s_t^\infty = \text{const.} - \alpha(d_t - c_t) + (\lambda - 1)\tilde{\mu}_t. \quad (47)$$

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 (44) 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.

III.J. $\psi \neq 1$

TO BE COMPLETED

IV. 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 price 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, a long-run weighted average of past real stock market returns is a good empirical predictor of excess returns and subjective expectations errors.

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.

REFERENCES

- Adam, Klaus, Albert Marcet, and Johannes Beutel, 2017, “Stock price booms and expected capital gains,” *American Economic Review* 107, 2352–2408.
- Ameriks, John, Gábor Kézdi, Minjoon Lee, and Matthew D Shapiro, 2016, “Heterogeneity in Expectations, Risk Tolerance, and Household Stock Shares,” Working paper, University of Michigan.
- Amromin, Gene, and Steven A Sharpe, 2013, “From the Horse’s Mouth: Economic Conditions and Investor Expectations of Risk and Return,” *Management Science*.
- 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.
- Campbell, John Y, 1991, “A Variance Decomposition for Stock Returns,” *Economic Journal* 101, 157–179.

- Campbell, John Y., 2000, “Asset Pricing at the Millenium,” *Journal of Finance* 55, 1515–1567.
- Campbell, John Y., and John H. Cochrane, 1999, “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy* 107, 205–251.
- Campbell, John Y., and Robert J. Shiller, 1988, “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors,” *Review of Financial Studies* 1, 195–228.
- Choi, 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 Andrei Shleifer, 2014, “Expectations of returns and expected returns,” *Review of Financial Studies* 27, 714–746.
- 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.
- Hirshleifer, David, Jun Li, and Jianfeng Yu, 2015, “Asset Pricing in Production Economies with Extrapolative Expectations,” *Journal of Monetary Economics* 76, 87–106.
- 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.
- Kung, Howard, 2015, “Macroeconomic linkages between monetary policy and the term structure of interest rates,” *Journal of Financial Economics* 115, 42–57.
- 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.
- Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-Jørgensen, 2009, “Long-Run Stockholder Consumption Risk and Asset Returns,” *Journal of Finance* 64, 2427–2479.
- Malmendier, Ulrike, and Stefan Nagel, 2011, “Depression Babies: Do Macroeconomic Ex-

- periences 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, 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.
- Mankiw, N. Gregory, and Stephen Zeldes, 1991, “The Consumption of Stockholders and Nonstockholders,” *Journal of Financial Economics* 29, 97–112.
- Nakov, Anton, and Galo Nuño, 2015, “Learning from Experience in the Stock Market,” *Journal of Economic Dynamics and Control* 52, 224–239.
- Pástor, L’uboš, and Pietro Veronesi, 2003, “Stock Valuation and Learning About Profitability,” *Journal of Finance* 58, 1749–1790.
- Pástor, L’uboš, and Pietro Veronesi, 2006, “Was there a Nasdaq bubble in the late 1990s?,” *Journal of Financial Economics* 81, 61–100.
- Piazzesi, M., and M. Schneider, “Equilibrium Yield Curves,” in *NBER Macroeconomics Annual*, D. Acemoglu, K. Rogoff, and M. Woodford, eds., (2007).
- 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*.

Appendix

A. DATA

A.1. 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.1 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.1 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.1 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.

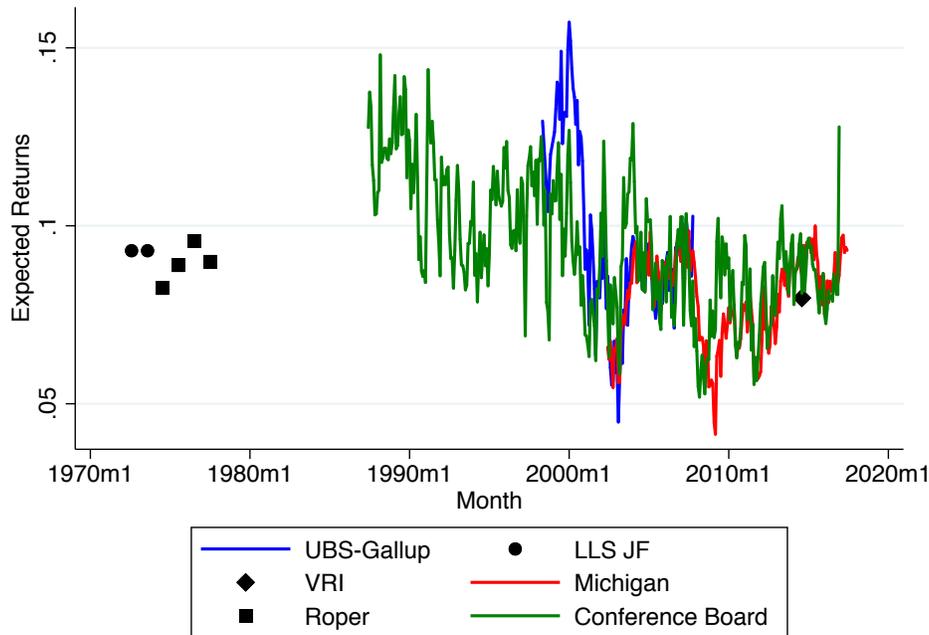


FIGURE A.1
Expected return imputation

B. BOOTSTRAP SIMULATIONS FOR PREDICTIVE REGRESSIONS

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

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

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

We use the described approach to simulate 10,000 bootstrap samples of predictors the

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

C. PROPERTIES OF THE PREDICTIVE DISTRIBUTION

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

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

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

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

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

Because of the constant autocovariance structure of perceived consumption growth,

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

and

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

This leads to

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

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

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

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

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

D. MODEL SOLUTION FOR $\psi = 1$

D.1. SDF

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

$$v_t^1 = \frac{\delta}{1-\gamma} \log \tilde{\mathbb{E}}_t[e^{(1-\gamma)(v_{t+1}^1 + \Delta c_{t+1})}], \tag{A.7}$$

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.

$$v_t^1 = \mu_v + U_v \tilde{\mu}_t. \tag{A.8}$$

Plugging in the conjectured solution we get

$$U_v = \frac{\delta}{1-\delta}, \tag{A.9}$$

and

$$\mu_v = \frac{1}{2}(1-\gamma)U_v(\nu U_v + 1)^2(1+\nu)\sigma^2. \tag{A.10}$$

We obtain the log SDF

$$\begin{aligned}
m_{t+1|t}^1 &= \log\left(\delta \frac{C_t}{C_{t+1}} \frac{(V_{t+1}^1)^{1-\gamma}}{\tilde{\mathbb{E}}_t[(V_{t+1}^1)^{1-\gamma}]}\right) \\
&= \log \delta - \Delta c_{t+1} + (1-\gamma) \log(V_{t+1}^1) - \log \tilde{\mathbb{E}}_t[(V_{t+1}^1)^{1-\gamma}] \\
&= \log \delta - \Delta c_{t+1} + (1-\gamma)(v_{t+1}^1 + c_{t+1}) - \log \tilde{\mathbb{E}}_t(e^{(1-\gamma)(v_{t+1}^1 + c_{t+1})}) \\
&= \tilde{\mu}_m - \tilde{\mu}_t - \xi \sigma \tilde{\varepsilon}_{t+1},
\end{aligned} \tag{A.11}$$

where

$$\tilde{\mu}_m = \log \delta - \frac{1}{2}(1 - \gamma)^2(\nu U_v + 1)^2(1 + \nu)\sigma^2, \quad (\text{A.12})$$

$$\xi = [1 - (1 - \gamma)(\nu U_v + 1)]\sqrt{1 + \nu}. \quad (\text{A.13})$$

D.2. Consumption claim valuation

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

$$R_{W,t+1} \equiv \frac{W_{t+1}}{W_t - C_t} = \frac{C_{t+1}}{C_t} \frac{\zeta}{\zeta - 1}, \quad (\text{A.14})$$

and in logs,

$$\begin{aligned} r_{w,t+1} &= \Delta c_{t+1} + \log(\zeta/(\zeta - 1)) \\ &= \tilde{\mu}_t + \sqrt{1 + \nu}\sigma\tilde{\varepsilon}_{t+1} + \log(\zeta/(\zeta - 1)). \end{aligned} \quad (\text{A.15})$$

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

$$\begin{aligned} \log(\zeta/(\zeta - 1)) &= -\tilde{\mu}_m + \xi\sqrt{1 + \nu}\sigma^2 - \frac{1}{2}(1 + \nu)\sigma^2 - \frac{1}{2}\sigma^2\xi^2 \\ &= -\log \delta, \end{aligned} \quad (\text{A.16})$$

which we can solve for the wealth-consumption ratio

$$\zeta = \frac{1}{1 - \delta}. \quad (\text{A.17})$$

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

$$\begin{aligned} w_t^1 - c_t &= \log \tilde{\mathbb{E}}_t \left[M_{t+1|t} \frac{C_{t+1}}{C_t} \right] \\ &= \log \tilde{\mathbb{E}}_t \left[\exp(\tilde{\mu}_m + (\sqrt{1 + \nu} - \xi)\sigma\tilde{\varepsilon}_{t+1}) \right] \\ &= \tilde{\mu}_m + \frac{1}{2}(\sqrt{1 + \nu} - \xi)^2\sigma^2, \end{aligned} \quad (\text{A.18})$$

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

$$w_t^n - c_t = n\tilde{\mu}_m + \frac{n}{2}(\sqrt{1 + \nu} - \xi)^2\sigma^2, \quad (\text{A.19})$$

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

$$w_t^n - c_t = n \log \delta \quad (\text{A.20})$$

Summing the value of consumption strips at all horizons strips yields the consumption-wealth ratio in (A.17).

D.3. Dividend strip valuation

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

$$r_{t+1}^1 = \lambda \Delta c_{t+1} - (\lambda - 1) \tilde{\mu}_t - \tilde{\mu}_m - \frac{1}{2} (\sqrt{1 + \nu \lambda} - \xi)^2 \sigma^2. \quad (\text{A.21})$$

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

$$r_{t+1}^1 - r_{f,t} = \lambda (\Delta c_{t+1} - \tilde{\mu}_t) - \frac{1}{2} (\sqrt{1 + \nu \lambda} - \xi)^2 \sigma^2 + \frac{1}{2} \xi^2 \sigma^2. \quad (\text{A.22})$$

The subjective conditional variance of r_{t+1}^1 is $(1 + \nu) \lambda^2 \sigma^2$, and so, after taking subjective expectations of (A.22), we obtain

$$\log \tilde{\mathbb{E}}_t [R_{t+1}^1] - r_{f,t} = \lambda \xi \sqrt{1 + \nu} \sigma^2. \quad (\text{A.23})$$

The objective conditional variance of r_{t+1}^1 is only $\lambda^2 \sigma^2$, and so taking objective expectations of (A.22) yields,

$$\log \mathbb{E}_t [R_{t+1}^1] - r_{f,t} = \lambda \xi \sqrt{1 + \nu} \sigma^2 - \frac{1}{2} \nu \lambda^2 \sigma^2 + \lambda (\mu - \tilde{\mu}_t). \quad (\text{A.24})$$

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

$$r_{t+1}^\infty = \Delta c_{t+1} + \frac{\lambda - 1}{\alpha} (\tilde{\mu}_{t+1} - \tilde{\mu}_t) - \tilde{\mu}_m - \frac{1}{2} \left[\sqrt{1 + \nu} \left(1 + \nu \frac{\lambda - 1}{\alpha} \right) - \xi \right]^2 \sigma^2, \quad (\text{A.25})$$

and, after subtracting the risk-free rate,

$$r_{t+1}^\infty - r_{f,t} = \Delta c_{t+1} + \frac{\lambda - 1}{\alpha} (\tilde{\mu}_{t+1} - \tilde{\mu}_t) - \tilde{\mu}_t - \frac{1}{2} \left[\sqrt{1 + \nu} \left(1 + \nu \frac{\lambda - 1}{\alpha} \right) - \xi \right]^2 \sigma^2 + \frac{1}{2} \xi^2 \sigma^2. \quad (\text{A.26})$$

The subjective conditional variance of r_{t+1}^∞ is $(1 + \nu) \left(1 + \nu \frac{\lambda - 1}{\alpha} \right)^2 \sigma^2$ and therefore, after taking subjective expectations of (A.26), we obtain

$$\log \tilde{\mathbb{E}}_t [R_{t+1}^\infty] - r_{f,t} = \left[1 + \nu \frac{\lambda - 1}{\alpha} \right] \xi \sqrt{1 + \nu} \sigma^2. \quad (\text{A.27})$$

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.26) yields

$$\log \mathbb{E}_t[R_{t+1}^\infty] - r_{f,t} = \left[1 + \nu \frac{\lambda - 1}{\alpha}\right] \xi \sqrt{1 + \nu} \sigma^2 - \frac{1}{2} \nu \left(1 + \nu \frac{\lambda - 1}{\alpha}\right)^2 \sigma^2 + \left(1 + \nu \frac{\lambda - 1}{\alpha}\right) (\mu - \tilde{\mu}_t). \quad (\text{A.28})$$

D.4. Numerical solution

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

$$P_t^n \approx C_t e^{\mu_{dc} + \frac{1}{2} \frac{1}{1 - (1 - \alpha)^2} \sigma_d^2 + \frac{\lambda - 1}{\alpha} \tilde{\mu}_t} \exp(n \tilde{\mu}_m + \frac{1}{2} A_n \sigma^2), \quad n > J, \quad (\text{A.29})$$

where we approximate

$$A_n \approx A_J + (n - J) [\sqrt{1 + \nu} (\nu \frac{\lambda - 1}{\alpha} + 1) - \xi]^2, \quad n > J. \quad (\text{A.30})$$

We can show that

$$P_t \approx \left(\sum_{n=1}^J P_t^n \right) + C_t V_J \exp \left(\mu_{dc} + \frac{1}{2} \frac{1}{1 - (1 - \alpha)^2} \sigma_d^2 + \frac{\lambda - 1}{\alpha} \tilde{\mu}_t \right) \quad (\text{A.31})$$

with

$$V_J = \frac{\exp \left((J + 1) \tilde{\mu}_m + \frac{1}{2} A_J \sigma^2 + \frac{1}{2} [\sqrt{1 + \nu} (\nu \frac{\lambda - 1}{\alpha} + 1) - \xi]^2 \sigma^2 \right)}{1 - \exp \left(\tilde{\mu}_m + \frac{1}{2} [\sqrt{1 + \nu} (\nu \frac{\lambda - 1}{\alpha} + 1) - \xi]^2 \sigma^2 \right)}. \quad (\text{A.32})$$

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

$$\log \frac{W_t - C_t}{C_t} = \log \frac{\delta}{1 - \delta}, \quad (\text{A.33})$$

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

$$\log \frac{P_t}{D_t} = H(\tilde{\mu}_t, d_t - c_t). \quad (\text{A.34})$$

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

$$\psi_{ij}(\tilde{\mu}, d_t - c_t) \equiv \Lambda_i(\tilde{\mu}) \Lambda_j(d_t - c_t), \quad (\text{A.35})$$

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

$$\hat{H}(\tilde{\mu}, d_t - c_t; \beta_m) = \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} \beta_{m,ij} \psi_{ij}(\tilde{\mu}, d_t - c_t) \quad (\text{A.36})$$

Rewrite the subjective Euler equation

$$\tilde{\mathbb{E}}_t[M_{t+1}R_{m,t+1}] = 1 \quad (\text{A.37})$$

as

$$\begin{aligned} 0 &= I(\tilde{\mu}_t, d_t - c_t) \\ &\equiv \tilde{\mathbb{E}}_t \left[e^{\tilde{\mu}_m - \tilde{\mu}_t - \xi \sigma \epsilon_{t+1} + \Delta d_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}} \right] - 1 \\ &= e^{\tilde{\mu}_m + (\lambda-1)\tilde{\mu}_t - \alpha(d_t - c_t - \mu_{dc})} \tilde{\mathbb{E}}_t \left[e^{(\lambda\sqrt{1+\nu} - \xi)\sigma \tilde{\epsilon}_{t+1} + \sigma_d \eta_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}} \right] - 1. \end{aligned} \quad (\text{A.38})$$

We evaluate the function $I(\tilde{\mu}_t, d_t - c_t)$ on the two-dimensional grid of $\tilde{\mu}_t$ and $d_t - c_t$ and use the two-dimensional Gaussian quadrature approach to calculate the expectation part as an integral. Following Pohl, Schmedders, and Wilms (2018), the numerical solution is implemented by the “fmincon” solver with the SQP algorithm in Matlab. We minimize a constant subject to the nonlinear constraints implied by Equation (A.38). 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.31) in terms of the RMSE,

$$RMSE_{pd} = \sqrt{\frac{1}{t} \sum_{j=1}^t (pd_j^{Analytical} - pd_j^{Projection})^2}, \quad (\text{A.39})$$

where $pd_t^{Analytical}$ is calculated from dividend strip prices as in (A.31). We explore different combinations of n_1 and n_2 up to a maximal degree of 8 and we choose the combination that minimizes $RMSE_{pd}$. Table A.I summarizes the parameter choices for this numerical procedure.

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

$$\begin{aligned} \tilde{\mathbb{E}}_t[R_{m,t+1}] &= \tilde{\mathbb{E}}_t \left[e^{\Delta d_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}} \right] \\ &= e^{\lambda \tilde{\mu}_t - \alpha(d_t - c_t - \mu_{dc})} \tilde{\mathbb{E}}_t \left[e^{\lambda\sqrt{1+\nu}\sigma \tilde{\epsilon}_{t+1} + \sigma_d \eta_{t+1}} \frac{e^{H(\tilde{\mu}_{t+1}, d_{t+1} - c_{t+1})} + 1}{e^{H(\tilde{\mu}_t, d_t - c_t)}} \right]. \end{aligned} \quad (\text{A.40})$$

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

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

E. MODEL SOLUTION FOR $\psi \neq 1$

TO BE COMPLETED