BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel — 2020-14

Last updated: April 7, 2020

Learning, Equilibrium Trend, Cycle, and Spread in Bond Yields

by Guihai Zhao

Financial Markets Department
Bank of Canada, Ottawa, Ontario, Canada K1A 0G9

gzhao@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

I would like to thank Jason Allen, Sharon Kozicki, Jonathan Witmer as well as seminar/conference participants at the Bank of Canada, Rotman School of Management, Haskayne School of Business, CEA 2019, Desautels Faculty of Management, SUFE, PBC School of Finance, and ASSA 2020 for their suggestions. The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Bank of Canada.

Abstract

Some key features in the historical dynamics of U.S. Treasury bond yields—a trend in long-term yields, business cycle movements in short-term yields, and a level shift in yield spreads—pose serious challenges to existing equilibrium asset pricing models. This paper presents a new equilibrium model to jointly explain these key features. The trend is generated by learning from the stable components in GDP growth and inflation, which share similar patterns to the neutral rate of interest (R-star) and trend inflation (Pi-star) estimates in the literature. Cyclical movements in yields and spreads are mainly driven by learning from the transitory components in GDP growth and inflation. The less-frequent inverted yield curves observed after the 1990s are due to the recent secular stagnation and procyclical inflation expectation.

Topics: Asset pricing; Financial markets; Interest rates

JEL codes: G00, G12, E43

1. Introduction

Some basic properties of the stochastic discount factor (SDF) have been established in the literature. Examples include the volatility bound of the SDF (Hansen and Jagannathan, 1991) and the permanent-transitory decomposition of SDF (Alvarez and Jermann, 2005; Hansen and Scheinkman, 2009; Hansen, 2012). Given that the SDF from any equlibrium model has direct implications for yield curves, the dynamics of the Treasury yield curve should tell us, in addition to these basic properties, what a good SDF should look like from a historical perspective. Figure 1 below shows some salient features in the data: (1) a hump-shaped trend in the yields, (2) business cycle (cyclical) movements in short-term yields and in the spreads between long- and short-term yields, (3) more-frequent and deeper inverted curves (accompanied by more-frequent recessions) pre-1990s than post-1990s, and (4) a positive yield spread on average.

Most equilibrium term structure models are designed to interpret and quantify the means, volatilities, and average positive spreads in yields. Given the standard stationarity assumption (which implies stationary short rates), it is hard for these models to generate the hump-shaped trend in yields and match the unconditional volatility in the data. As a result, both the low-frequency variations in the long-term yields and the cyclical movements in the yield spreads are determined to be the risk premia (see, for example, the inflation risk premium in Piazzesi and Schneider 2007). The key assumption for generating a positive inflation risk premium and, hence, an upward-sloping yield curve, is that inflation is bad news for future growth. However recent studies have shown that, following the late 1990s, inflation has switched to a good-news event for future growth.[1] This fact implies a negative inflation risk premium and a downward-sloping nominal curve in equilibrium. Figure 1 shows the opposite: the yield spreads shifted to an even higher level after the 1990s. Furthermore, it is hard to explain business cycle movements in yield spreads using a risk-premium approach – moving from positive out of recessions

[1] See, for example, Burkhardt and Hasseltoft (2012); David and Veronesi (2013); Campbell et al. (2017); Zhao (2020)

to negative in a late expansion stage. Finally, the third fact – the more-frequent and deeper inverted curves for pre-1990s than for post-1990s – has been overlooked in the literature. Given the recent concern over the association between recessions and inverted yield curves, we need an equilibrium interpretation for this fact. In this paper, we show that the historical dynamics of the yield curve can largely be explained by movements in the short-rate expectations and provide a joint-equilibrium understanding of these salient features in the data.

Figure 1: U.S. 10- and 1-year nominal bond yields

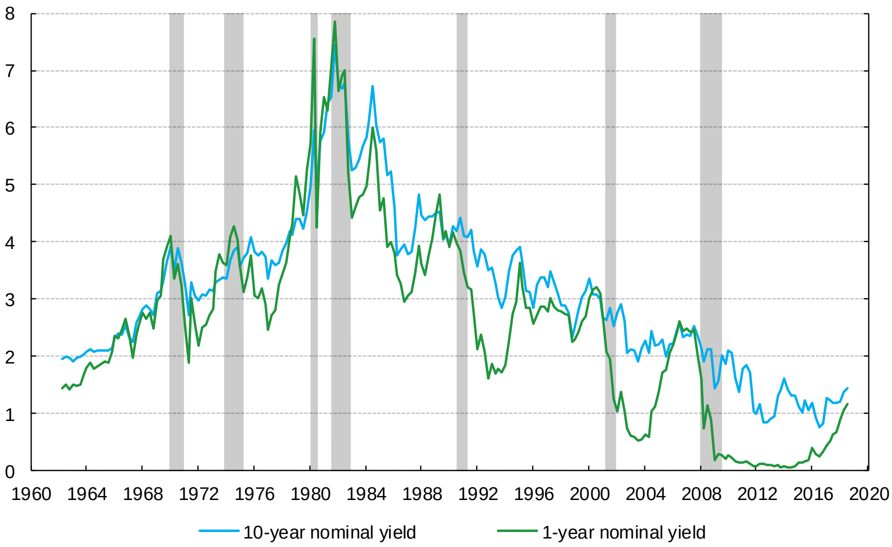
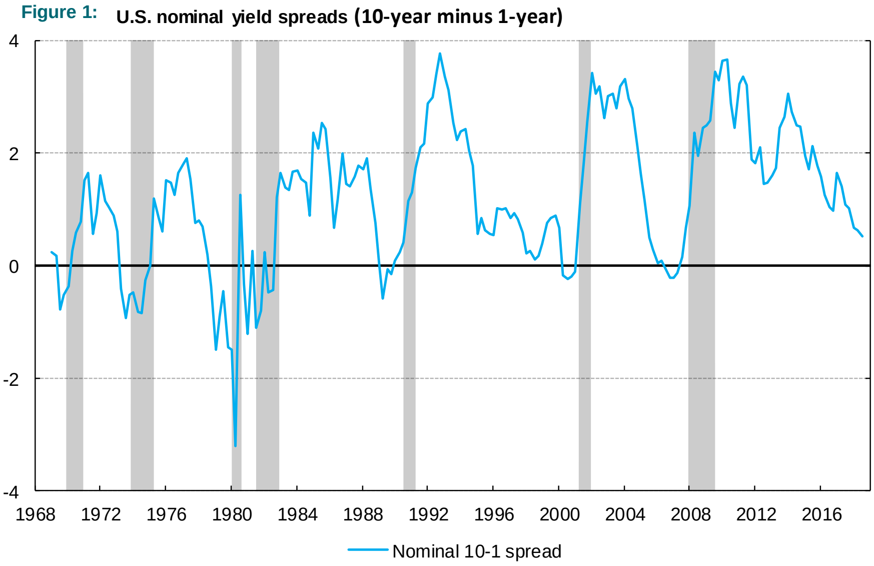


Figure 1: U.S. nominal yield spreads (10-year minus 1-year)



The end-of-quarter 10-year nominal, 1-year nominal yield, and their spreads are obtained from Gürkaynak et al. (2007) from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession defined by the NBER.

It has long been recognized that nominal interest rates contain a slow-moving trend component (Nelson and Plosser, 1982; Rose, 1988). Recent empirical studies propose macro trends as the driving force behind this low-frequency variation. For example, Kozicki and Tinsley (2001) and Cieslak and Povala (2015) document the empirical importance of trend inflation (π_t^*) for explaining the secular decline in Treasury yields since the early 1980s. Bauer and Rudebusch (2019) show that it is crucial to also include the neutral rate of interest (r_t^*), which has driven the downward trend in long term yields over the last 20 years. Both π_t^* and r_t^* (and hence short rates) are modeled as random-walk processes in Bauer and Rudebusch (2019), and the resulting term premium component – the difference between long-term interest rates and the model-implied expectations of average future short-term rates – is relatively small and stationary. In most equilibrium models, however, short rates are assumed to be stationary. Therefore, the large term premia (or residual term) in yields that contain both low-frequency variations and cyclical movements are typically explained as certain types of risk premia.

To illustrate the role of learning in generating the observed low-frequency variations in the yields and in matching the macro trends, we start with a simple equilibrium model (model I) where both the output growth and inflation are exogenous and follow the i.i.d. laws of motion. The representative agent with a constant relative risk aversion (CRRA) utility does not, however, know the mean inflation and mean growth rates. Instead, the agent uses a constant-gain learning scheme as proposed by Nagel and Xu (2019) to learn the unconditional mean inflation and growth rates.[2] This is a modified Bayesian approach, where the learning is perpetual due to the agent's fading memory. The posteriors for the mean inflation and mean growth rates, as state variables, capture the trends in inflation and growth. The bond yields implied by the model as linear functions of the posteriors exhibit a hump-shaped trend.

The model-I implied r_t^*, which is a linear function of the posterior for mean output

[2]The posteriors from standard Bayesian learning (about the fixed parameters) are random-walk processes, which are potentially also consistent with the empirical modeling of the macro trends. However, the posterior variance will decline deterministically to zero and the learning will converge (Collin-Dufresne et al., 2016).

growth, moves closely with the r_t^* estimates in the literature. Not only does the model capture the low-frequency variations, but the model-I-implied r_t^* also exhibits a moderate business cycle component. Furthermore, the posterior of the mean inflation matches the survey-based trend inflation. As a result, the hump-shaped trend in the 10-year Treasury yield (from the late 1960s to the late 1990s) reflects an increase in inflation expectations before the mid-1980s and a secular decline afterwards. More recently, as inflation expectations stabilized, the decline in the posterior for mean output growth and, hence, the 10-year real yield, has been the main driver of the downtrend in nominal yields.[3] However, the posteriors affect the yields of all maturities equally, and the yields for the short- and long-term bonds are almost identical. Therefore, model I cannot be used to explain the cyclical movements in the short-term yield and in the spreads between the long- and short-term yields.

To address these cyclical movements, we consider an extended version (model II), where both the GDP growth and the inflation rates are decomposed into two components: one stable and one transitory/volatile. Using the same learning scheme as above, the representative agent learns about the unconditional mean output growth and inflation rates from the stable component, and they learn about the stationary deviations from the mean by using the transitory/volatile component. As in model I, the posteriors for long-run mean inflation and growth affect short- and long-term bonds equally and, hence, capture the trend in the yields. However, the posteriors for the transitory deviations from the long-run mean are AR(1) processes. And they have larger impacts on the short-term yield than on the long-term yield, which implies a positive (negative) spread when the

[3] In this model, output growth is given exogenously and r_t^* is driven by learning about the mean growth. Alternative interpretations include lower productivity growth, changing demographics, a decline in the price of capital goods, and strong precautionary savings flows from emerging market economies. See, for example, Summers (2014); Kiley (2015); Rachel and Smith (2015); Carvalho et al. (2016); Hamilton et al. (2016); Laubach and Williams (2016); Johannsen and Mertens (2018); Christensen and Rudebusch (2019); Holston et al. (2017); Lunsford and West (2018); Del Negro et al. (2017). Caballero et al. (2008) show that the downward trend in interest rates was due to a shortage of safe assets and the increasing global imbalances. Farhi and Gourio (2019) find that rising market power, rising unmeasured intangibles, and rising risk premia, played a crucial role for the decline in real short rates over the past 30 years.

short-run beliefs are negative (positive). Therefore, model II can be used to generate cyclical movements in the short-term yields and, hence, in the spreads.

The model-II implied 1-year nominal yield and the spread between 10- and 1-year nominal yields closely track their historical movements. Out of recessions, the short-term nominal yield starts to rise when the agent begins to revise their beliefs for short-run growth and inflation upwards towards their long-run means (short-run deviations are still negative and the spread is positive), and the spread starts to shrink as these short-run expectations move towards becoming positive. This pattern continues until the late expansion stage, when the short-run growth and inflation expectations are above their long-run means (the short-run deviations are positive now), implying an inverted yield curve.[4] The agent then begins to revise their short-run beliefs sharply downwards, entering a recession, and the spread switches from negative to positive.

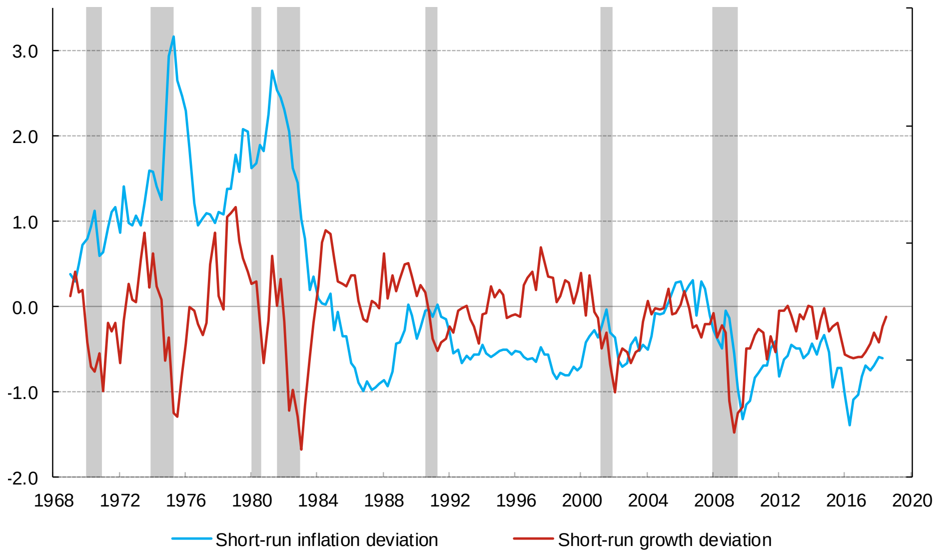
Furthermore, Figure 10 shows that (1) the posteriors for short-run inflation and growth deviations moved in opposite directions before the late 1990s and in the same direction afterwards, and (2) the posteriors for both short-run inflation and growth deviations were persistently negative for most of the post-2000 period (consistent with secular stagnation; see, e.g., Summers 2014). These observations imply that short-run inflation and growth expectations drove the yields in opposite directions pre-2000, and in the same direction afterwards. Moreover, both short-run inflation and growth expectations imply positive spreads for the past two decades. Therefore, consistent with the data, model II generates more-frequent and deeper inverted curves for the period before the late 1990s than for the most-recent period. From a monetary policy point of view, the Federal Reserve faced a trade-off between short-run inflation and growth pre-2000 and no such trade-off afterwards.[5] Hence, we observe more-frequent recessions for the period before the late

[4]The model is consistent with our conventional understanding of business cycles. From trough to peak, when the inflation and output gaps (short-run deviations in this model) are moving from negative to positive, monetary policy turns from accommodative to contractionary (short-term yields increase) according to the Taylor Rule. The difference between the standard Taylor Rule and the model-implied short rate is discussed in Section 3.2.

[5]It is commonly believed that the U.S. economy was mostly hit by supply shocks before the late 1990s, and mostly hit by demand shocks afterwards, and that this generated the change in the correlation

1990s and less-frequent recessions (or longer business cycles) for the period afterwards.

Figure 10: **Short-run deviations**



The model-implied posteriors for the short-run deviations from the long-run mean ($\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$) are from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession defined by the NBER.

Finally, the model-II-implied inflation expectations (posteriors) closely match the survey-based short- and long-run inflation expectations. And as in model I, the model-II-implied r_t^* tracks very closely the r_t^* estimates in the literature; hence, the model-II-implied 10-year nominal yield moves closely with the data. Despite the fact that the model-II-implied nominal spread can stay positive or negative for an extended period of time at different phases of the business cycle, the level is almost in parallel lower than data due to the stationary assumption (mean zero) for the short-run beliefs as well as the CRRA utility.

The common equilibrium explanation for the upward-sloping nominal yield curve is the inflation risk premium (Piazzesi and Schneider, 2007), where inflation is bad news for future growth and the agent prefers an early resolution to the uncertainty. Zhao (2020) shows that this approach was less effective during the past two decades when inflation switched from bad news to good news for future growth, providing an alternative worst-

between short-run inflation and growth expectations.

case belief approach through ambiguity. We show that model II can be extended by incorporating the intuition discussed in Zhao (2020). The ambiguity-averse representative agent (with the recursive multiple priors, or maxmin, preferences in Epstein and Schneider 2003) has in mind a benchmark or reference measure of the economy's dynamics that represents the best estimate of the stochastic process. In model III, the reference measure is the full stochastic environment presented in model II (including the posteriors). But the agent is concerned that the reference measure is misspecified and believes that the true measure is actually within a set of alternative measures that are statistically close to the reference distribution. Using forecast dispersion to quantify the size of the ambiguity (following Ilut and Schneider 2014), the model-III-implied short-rate expectations are upward-sloping under investors' worst-case equilibrium beliefs, which generates upward-sloping nominal and real yield curves, even with a CRRA utility. The expectations hypothesis (EH) roughly holds under the worst-case equilibrium belief; and the ex-post predictability of excess bond returns is due to the difference between investors' worst-case expectations and the reference measure – providing a rational interpretation for expectational errors in Froot (1989), Piazzesi et al. (2015), and Cieslak (2018).

Given that the benchmark measure in model III is the same as in model II, model III can still match the trends and cycles in the yields. However, because of ambiguity, the model-III-implied spreads are almost in parallel higher than the model-II-implied spreads and can also match the yield spreads in the data. Furthermore, by comparing the model-implied 1-year nominal yield with the data, we observe a recurring pattern that the model-implied short-term yields are higher than those in the data, from trough to expansion, but they are lower than those in the data during the late expansion periods. Given that short-term yields are controlled by the Federal Reserve, this suggests that the Federal Reserve kept the short rates low for a longer period than the model suggests and there was a certain degree of overshoot during the late expansion periods (before the recessions).

Related literature

This paper is related to a large literature on equilibrium asset/bond pricing models. For example, see Wachter (2006), Lettau and Wachter (2011), David and Veronesi (2013), Bansal and Shaliastovich (2013), Rudebusch and Swanson (2012), Albuquerque et al. (2016), and Berrada et al. (2018). This paper is most closely related to Piazzesi and Schneider (2007), who show the importance of the inflation risk premium in explaining the upward-sloping nominal curve in a stationary state space model (for inflation and growth). While most equilibrium bond-pricing models focus on the first/second moment and the average spread in the yields, we show that some key features in the historical dynamics of U.S. Treasury bond yields pose serious challenges to existing models, and we provide a joint equilibrium understanding of the trends, cycles, and spreads in the data.

This paper differs from previous studies along some important dimensions. First, the agent in this paper takes into consideration the risk of belief updating. Hence, the posteriors for the long-run means are state variables, which move closely with the macro trend estimates (r_t^* and π_t^*) in the literature and explain the hump-shaped trend in the yields. Most importantly, this is the first paper that decomposes both GDP growth and inflation into two components and shows that learning about the long-run mean drives the low-frequency variations in the yields and that the learning about the short-run deviation from the mean drives the business cycle movements in the short-term yield and, hence, in the spreads. Third, this paper provides an interpretation for an important but often overlooked fact – the less-frequent inverted yield curves (and the less-frequent recessions) after the 1990s. Finally, instead of the inflation risk premium (Piazzesi and Schneider, 2007) and the real risk premium (Wachter, 2006; Albuquerque et al., 2016; Berrada et al., 2018), the upward-sloping nominal and real curves in the U.S., at least for the post-2000 periods, were *partially* due to persistently negative short-run inflation and growth expectations and procyclical inflation expectations for most of these periods. However, short-run inflation and growth expectations moved in opposite directions pre-2000, which makes it hard for model II to generate an average upward-sloping nominal curve (the short-run expectations and, hence, the yield spreads are stationary and mean zero). We therefore rely on the worst-case belief approach (Zhao, 2020) to generate upward-sloping nominal and real curves that are consistent with the data.

The paper is also related to a large empirical literature that links macro information and macro trends with yield curve modeling.[6] This paper bridges an important gap between the empirical and equilibrium yield curve literature by interpreting macro trends as posteriors for learning about long-run mean growth and inflation rates. Furthermore, the paper also provides an equilibrium interpretation for the cyclical movements in short-term yields and, hence, in yield spreads.

This paper is related to a number of papers that study the implications of ambiguity and robustness for finance and macroeconomics.[7] Model III incorporates the intuition in Zhao (2020) to generate the upward-sloping nominal and real curves through the upward-sloping short-rate expectations under the representative agent's worst-case belief. Finally, this paper is also related to some recent developments wherein the implications of learning in finance were investigated. For example, Collin-Dufresne et al. (2016) show how a standard Bayesian learning can generate subjective long-run risks when the representative agent prefers an early resolution to uncertainty (Epstein and Zin, 1989). Building on the insight in Malmendier and Nagel (2011) and Malmendier and Nagel (2016) that the dynamics of the average individual's expectations can be approximated closely by a constant-gain learning scheme, Nagel and Xu (2019) show how this constant-gain learning can help separate subjective and objective equity premia and explain the predictability of excess returns. In this paper, we use the constant-gain learning scheme in a different setting to explain bond yield dynamics.

The paper continues as follows. Section 2 outlines and solves model I in closed form

[6]See Singleton (1983), Fama and Bliss (1987), Kozicki and Tinsley (2001), Ang and Piazzesi (2003), Diebold et al. (2006), Rudebusch and Wu (2008), Ludvigson and Ng (2009), Wright (2011), Cieslak and Povala (2015), and Bauer and Rudebusch (2019), among many others.

[7]Recent papers that study the multiple-priors preference and its applications include Epstein and Schneider (2007), Ilut (2012), Ilut and Schneider (2014), Ulrich (2013) , Gagliardini et al. (2009), Bianchi et al. (2017), Ilut et al. (2018), Zhao (2017), and Zhao (2020), among many others. Papers on robustness applications include Anderson, Hansen, and Sargent (2003), Cagetti, Hansen, Sargent, and Williams (2002), Hansen (2007), and Hansen and Sargent (2001, 2010). For a detailed survey, see Epstein and Schneider (2010).

and discusses the model implications. Section 3 undertakes the same steps for model II. Model III is solved in Section 4. Section 5 provides concluding comments.

2. Model I - Learning and trends in bond yields

In this section, we consider an endowment economy with a representative agent who has a CRRA utility function. The agent learns with fading memory about the rates of mean output growth and inflation. Equilibrium prices adjust such that the agent is happy to consume the output as an endowment.

2.1. Learning with fading memory

Both the output growth and inflation follow the i.i.d. laws of motion, as follows:

$$\begin{aligned} \Delta g_{t+1} &= \mu_c + \sigma_c \varepsilon_{c,t+1} \\ \pi_{t+1} &= \mu_\pi + \sigma_\pi \varepsilon_{\pi,t+1}, \end{aligned} \tag{1}$$

where Δg_{t+1} is the growth rate of real output and π_t is inflation. $\varepsilon_{c,t+1}$ and $\varepsilon_{\pi,t+1}$ are the i.i.d. normal shocks. The representative investor knows that both Δg_{t+1} and π_{t+1} are i.i.d., and also knows the values for σ_c and σ_π, but not for μ_c and μ_π. The agent forms expectations about μ_c and μ_π based on the history of the output growth and inflation realizations, $H_t^g \equiv \{\Delta g_0, \Delta g_1, ...\Delta g_t\}$ and $H_t^\pi \equiv \{\pi_0, \pi_1, ...\pi_t\}$.

At each time t, a Bayesian agent updates the agent's prior belief $p(\mu_c)$ or $p(\mu_\pi)$ in a way that assigns each past observation for Δg_{t-j} or π_{t-j} equal weight in the posterior probability. The equal-weighting of the past observations in H_t^g and H_t^π means that there is no decay in memory as the agent uses all of the available data in forming the posterior beliefs.

In this paper, we use the constant-gain learning scheme proposed by Nagel and Xu (2019), based on a weighted-likelihood approach used in the theoretical biology literature (Mangel, 1990). Compared with standard constant-gain learning models, the learning here allows us to derive the full posterior distribution. Taking output growth as an example, with fading memory, the representative agent who has observed an infinite history of past output growth Δg forms their posterior

$$p(\mu_c|H_t^g) \propto p(\mu_c) \prod_{j=0}^{\infty} \left[\exp\left(-\frac{(\Delta g_{t-j} - \mu_c)^2}{2\sigma_c^2} \right) \right]^{(1-\upsilon_c)^j}, \tag{2}$$

where $1-\upsilon_c$ is a positive number close to one and $(1-\upsilon_c)^j$ represents a geometric weight on each observation. The agent assigns more weight to the recent observations than to the observations that are receding into the past. We work with uninformative priors in the model, $\mu_c \sim N(\mu_{c,0}, \sigma_{c,0})$ and $\mu_\pi \sim N(\mu_{\pi,0}, \sigma_{\pi,0})$, with $\sigma_{c,0} \to \infty$ and $\sigma_{\pi,0} \to \infty$.[8] The posteriors are given by the following:

$$\begin{aligned} \mu_c|H_t^g &\sim N\left(\tilde{\mu}_{c,t}, \upsilon_c \sigma_c^2\right) \\ \mu_\pi|H_t^\pi &\sim N\left(\tilde{\mu}_{\pi,t}, \upsilon_\pi \sigma_\pi^2\right), \end{aligned} \tag{3}$$

where

$$\begin{aligned} \tilde{\mu}_{c,t} &= \tilde{\mu}_{c,t-1} + \upsilon_c\left(\Delta g_t - \tilde{\mu}_{c,t-1}\right) = \upsilon_c \sum_{j=0}^{\infty} (1-\upsilon_c)^j \Delta g_{t-j} \\ \tilde{\mu}_{\pi,t} &= \tilde{\mu}_{\pi,t-1} + \upsilon_\pi\left(\pi_t - \tilde{\mu}_{\pi,t-1}\right) = \upsilon_\pi \sum_{j=0}^{\infty} (1-\upsilon_\pi)^j \pi_{t-j}\,. \end{aligned} \tag{4}$$

Unlike standard Bayesian learning, where the variance of the posterior converges to zero, the learning is perpetual here. The variance of the posterior is the same as if the agent had observed and fully retained in memory with equal weight $S^g \equiv \frac{1}{\upsilon_c}$ ($S^\pi \equiv \frac{1}{\upsilon_\pi}$) realized growth rate (inflation) observations. Although the actual number of observations is infinite, the loss of memory induced by the geometric weighted-likelihood implies that the effective sample size is equal to a finite number S^g (S^π). The posterior $\tilde{\mu}_{c,t}$ ($\tilde{\mu}_{\pi,t}$) resulting from this weighted-likelihood approach is identical to the posterior that one obtains from a standard constant-gain updating scheme with gain υ_c (υ_π).

[8] See Nagel and Xu (2019) for a discussion of the informative prior case.

To better understand the stochastic nature of the output growth and inflation from the agent's subjective viewpoint, we further derive the predictive distribution as follows:

$$\begin{aligned}\Delta g_{t+j}|H_t^g &\sim N\left(\tilde{\mu}_{c,t}, (1+\upsilon_c)\,\sigma_c^2\right)\\ \pi_{t+j}|H_t^\pi &\sim N\left(\tilde{\mu}_{\pi,t}, (1+\upsilon_\pi)\,\sigma_\pi^2\right),\end{aligned} \tag{5}$$

where $j = 1, 2, ...$ and the variance of the predictive distribution contains both uncertainty due to future shocks $\varepsilon_{c,t+j}(\varepsilon_{\pi,t+j})$, and uncertainty about $\mu_c(\mu_\pi)$. Denoting the expectations under the predictive distribution by $\tilde{E}_t$, we can rewrite the posteriors as follows:

$$\begin{aligned}&\tilde{\mu}_{c,t+1} = \tilde{\mu}_{c,t} + \upsilon_c\sqrt{1+\upsilon_c}\sigma_c\tilde{\varepsilon}_{c,t+1}\\ &\tilde{\mu}_{\pi,t+1} = \tilde{\mu}_{\pi,t} + \upsilon_\pi\sqrt{1+\upsilon_\pi}\sigma_\pi\tilde{\varepsilon}_{\pi,t+1},\end{aligned} \tag{6}$$

where $\tilde{\varepsilon}_{c,t+1} = \frac{\Delta g_{t+1}-\tilde{\mu}_{c,t}}{\sigma_c\sqrt{1+\upsilon_c}}$ and $\tilde{\varepsilon}_{\pi,t+1} = \frac{\pi_{t+1}-\tilde{\mu}_{\pi,t}}{\sigma_\pi\sqrt{1+\upsilon_\pi}}$. $\tilde{\varepsilon}_{c,t+1}/\tilde{\varepsilon}_{\pi,t+1}$ is $N(0,1)$ distributed and, hence, is unpredictable under the time-t predictive distribution.

2.2. Valuation with fading memory

Unlike in a standard Bayesian environment, where the information structure can be represented by a filtration with posterior beliefs following a martingale, with loss of memory the information structure is not a filtration. The posterior in periods $t+j$ are updated based on different (but not more-informative) information available at time t. Thus, the information structure is not a filtration. Nagel and Xu (2019) show that at time t, the agent knows that the variation in $\tilde{\mu}_{c,t+j}/\tilde{\mu}_{\pi,t+j}$ will be stationary and perceives future increments $\tilde{\varepsilon}_{c,t+j}/\tilde{\varepsilon}_{\pi,t+j}, j = 1, 2, ...$ as negatively serially correlated. However, the agent cannot make use of this serial correlation by using $\tilde{\varepsilon}_{c,t}/\tilde{\varepsilon}_{\pi,t}$ to forecast $\tilde{\varepsilon}_{c,t+1}/\tilde{\varepsilon}_{\pi,t+1}$, because $\tilde{\varepsilon}_{c,t}/\tilde{\varepsilon}_{\pi,t}$ is not observable.

To value the zero-coupon bond under this information structure, we use $M_{t+j|t}$ to denote the one-period SDF from $t+j-1$ to $t+j$, given the agent's predictive distribution at t. The time-t price of a zero-coupon bond that pays one unit of consumption two periods from now is denoted as $P_t^{(2)}$ and satisfies the recursion

$$P_t^{(2)} = \tilde{E}_t[M_{t+1|t}P_{t+1}^{(1)}] = \tilde{E}_t\left[M_{t+1|t}\tilde{E}_{t+1}\left(M_{t+2|t+1}\right)\right], \tag{7}$$

and the valuation at t is based on the anticipation that the value of the asset at date $t+1$ will be determined by an agent – or a future self of the agent – who perceives $\tilde{\varepsilon}_{c,t+1}/\tilde{\varepsilon}_{\pi,t+1}$ as unpredictable.

2.3. Kalman filter alternative

The updating scheme in (4) is similar in spirit to optimal filtering with a latent stochastic trend. For the output growth, if the agent perceives that μ_c follows a random walk ($\mu_{c,t} = \mu_{c,t-1} + \varepsilon_{c,t}^{\mu}$), rather than a constant as in (1), then the resulting posterior distribution from the steady-state Kalman filter is the same as the adaptive learning in (3). With an appropriate choice of the volatility of the $\varepsilon_{c,t}^{\mu}$ shocks, the dynamics of the posterior beliefs from the Kalman filter would be the same as those in the updating scheme with fading memory in (4), however, with the information structure as a filtration. The one-step-ahead predictive distribution and the assets valuation would also be the same. If the true law of motion is (1) with a constant μ_c, then there will be a time-varying wedge $\mu_{c,t} - \tilde{\mu}_{c,t}$ between the subjective and objective beliefs, which could play an important role in generating predictions of excess bond returns. Given that there is no strong empirical evidence of GDP growth predictions, especially for long-run growth, we stick to the fading-memory interpretation in this model.

The same argument applies to the inflation process. However, U.S. inflation was highly persistent before the late 1990s and became less predictable thereafter. It is more reasonable to assume a latent random-walk trend with Kalman filter learning for the period before the late 1990s and a constant mean with adaptive learning for the period after the late 1990s. But, given that the model-implied bond prices and yields are the same, to be consistent, we use the same constant-gain learning.

2.4. Model solutions

Piazzesi and Schneider (2007) show the importance of Epstein and Zin (1989) preferences in generating a sizable inflation risk premium for long-maturity nominal bonds. To illustrate the key role of trend inflation and trend output growth for long maturity bond yields, we assume that investors have recursive preferences with a CRRA utility function (i.e., they are indifferent between an early or late resolution of uncertainty):

$$V_t(C_t) = \tilde{\mathbb{E}}\left(U\left(C_t\right) + \beta V_{t+1}(C_{t+1})\right), \tag{8}$$

where $U\left(C_t\right) = \frac{C_t^{1-\gamma}-1}{1-\gamma}$, γ is the coefficient of the risk aversion and β reflects the investor's time preference. Note that the agent evaluates the continuation value under their subjective expectations.

2.4.1. Bond pricing

Since the representative agent forms expectations under their subjective beliefs when making portfolio choices, the Euler equation holds under these subjective expectations. Given the CRRA utility function, the log nominal pricing kernel or the nominal stochastic discount factor can be written as follows:

$$m^{\$}_{t+1|t} = log\beta - \gamma\Delta g_{t+1} - \pi_{c,t+1} = log\beta - v' z_{t+1}, \tag{9}$$

where $v' = (\gamma, 1)$ and $z_t = (\Delta g_t, \pi_t)^T$. The time-$t$ price of a zero-coupon bond that pays one unit of consumption n periods from now is denoted as $P_t^{(n)}$ and it satisfies the recursion

$$P_t^{(n)} = \tilde{E}[M^{\$}_{t+1|t} P_{t+1}^{(n-1)}] \tag{10}$$

with the initial condition that $P_t^{(0)} = 1$ and $\tilde{E}$ is the expectation operator under the predictive distribution. Given the linear Gaussian framework, we assume that $p_t^{(n)} =$

$log(P_t^{(n)})$ is a linear function of the posteriors $\tilde{\mu}_t = (\tilde{\mu}_{c,t}, \tilde{\mu}_{\pi,t})^T$

$$p_t^{(n)} = -A^{(n)} - C^{(n)}\tilde{\mu}_t. \tag{11}$$

When we substitute $p_t^{(n)}$ and $p_{t+1}^{(n-1)}$ in Euler equation (10), the coefficients in the pricing equation can be solved with $C^{(n)} = C^{(n-1)} + v' = v'n$, and $A^{(n)} = A^{(n-1)} + A^{(1)} - 0.5 * Var_t\left(p_{t+1}^{(n-1)}\right) - Cov_t\left(p_{t+1}^{(n-1)}, m_{t,t+1}^{\$}\right)$ (see the appendix for details). The log holding period return from buying an n period bond at time t and selling it as an $n-1$ period bond at time $t+1$ is defined as $r_{n,t+1} = p_{t+1}^{(n-1)} - p_t^{(n)}$, and the subjective excess return is $er_{n,t+1} = -Cov_t\left(r_{n,t+1}, m_{t,t+1}^{\$}\right) = -v'Cov_t\left(z_{t+1}, \tilde{\mu}_{t+1}\right)C^{(n-1)}$. The n period bond yield is defined as $y_t^{(n)} = log(Y_t^{(n)}) = -\frac{1}{n}p_t^{(n)}$.

As we can see from the solution, the yield parameter for the posterior $\tilde{\mu}_t$ is constant over horizon n and all of the variance and covariance terms are relatively small in the data. Hence, given the CRRA utility (i.e., no extra term premium from the agent's time preference, in contrast to the Epstein and Zin (1989) case), the term premium is small in this model, which implies a flat yield curve. To solve the price and yields for real bonds, we can simply replace v' with $v' = (\gamma, 0)$.

2.5. Empirical findings

Using the U.S. real GDP growth and the rate of inflation from the GDP deflator, we can calculate the posterior beliefs for the output growth and inflation. We then show that they closely match the estimated r^* and π^* in the literature. Given the analytical solutions and the posteriors, we can calculate the model-implied 10-year nominal and real bond yields, which match the historical movements in the data well.

2.5.1. Data

Real output growth and GDP deflator inflation are from the Bureau of Economic Analysis for the period 1947.Q2 to 2018.Q2. The end-of-quarter yields for 1- to 10-year bonds are from the daily dataset constructed by Gürkaynak et al. (2007) for the period 1962.Q1 to 2018.Q2. The Treasury inflation protected securities (TIPS) yields (2003.Q1

γ	β	υ_c	υ_π	σ_c	σ_π	$corr$
3	1.019	0.01	0.045	0.94	0.64	-0.072

Table 1: Configuration of model-I parameters
Table 1 reports the parameter values for the output growth and inflation process and for the constant gain in learning. All of the parameters are given in quarterly terms. The mean and standard deviation are in percentages.

to 2018.Q2) and the end-of-quarter yields for the three- and six-month Treasury bills are obtained from the U.S. Department of the Treasury via the Fed database at the St. Louis Federal Reserve (1969.Q4 to 2018.Q2). The forecasts for real output growth and inflation are from the Philadelphia Fed's survey of professional forecasters (SPF) for the period 1968.Q3 to 2018.Q2. The r^* and π^* variables are from Bauer and Rudebusch (2019) for the period 1971.Q4 to 2017.Q2. We use 15 years (1947.Q2 to 1962.Q1) as training periods for the posterior beliefs and show the results starting from 1962.Q2 when the bond yield data first became available.

2.5.2. Parameters

The volatility parameters for the output growth and inflation are calibrated to match their counterparts in the data. The correlation between the output growth and inflation in the model is calibrated to match the correlation in the data. The constant-gain parameters υ_c and υ_π are calibrated to match the variations in r^*_t and π^*_t.[9] We follow the literature and set the risk aversion to 3. The time preference β is calibrated to match the level of the 10-year nominal yields in the data, which is close to the value in Piazzesi and Schneider (2007). The resulting parameter values are shown in Table 1.

2.5.3. Posteriors versus r^*_t and π^*_t

The empirical literature has shown the importance of accounting for macro trends in the term structure of the interest-rate modeling. However, the stationary assumption in the leading equilibrium bond pricing models makes it hard to generate the histor-

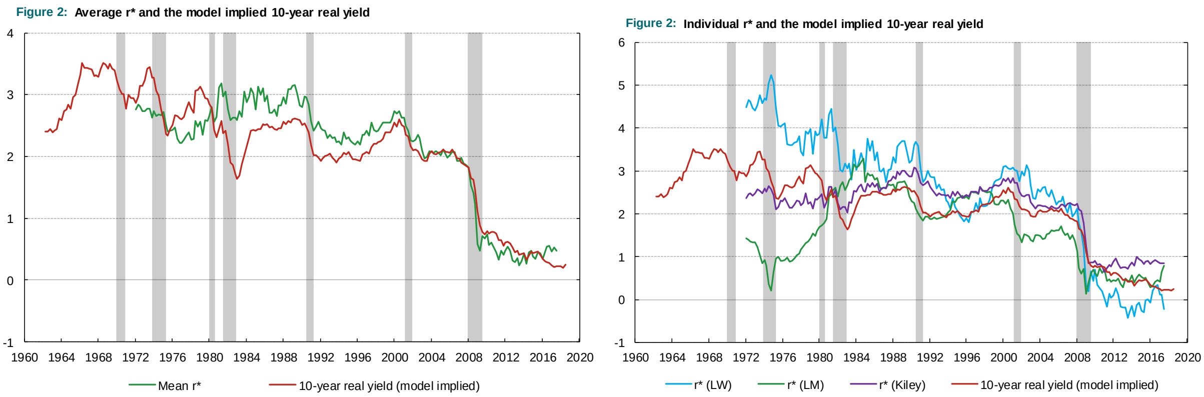
[9] These parameters are within the range of values used in the literature for an optimal Kalman gain. Gilchrist and Saito (2008) use $\upsilon = 0.06138$, Edge et al. (2007) use $\upsilon = 0.11$. Malmendier and Nagel (2016) show that $\upsilon = 0.018$ for the quarterly data that fits the dynamics of the average belief about the inflation expectations in the microdata.

ically observed low-frequency variations in interest rates. The posterior mean of the output growth (inflation) in Section 2.1 would be a random-walk process under a standard Bayesian learning; however, the posterior variance would decline deterministically to zero and the learning would converge (Collin-Dufresne et al., 2016). In this paper, the representative agent updates their subjective beliefs with a constant gain, which induces memory loss and is otherwise a standard Bayesian. However, the learning in this model is slow but perpetual, and it generates low-frequency variations in the posterior beliefs that closely match the r_t^* and π_t^* that were estimated in the literature and, conveniently, also implies a stationary economy.[10]

Figure 2 (left panel) shows the model-implied 10-year real yield (or the model-implied r_t^*), which is a linear function of the posterior mean of the output growth, $\gamma\tilde{\mu}_{c,t} + Cov$, closely tracks the estimated mean r_t^*. Given the CRRA utility, the term premium part Cov is very small in the 10-year real yield, and variations in the model-implied 10-year real yield are mainly driven by variations in $\tilde{\mu}_{c,t}$. The mean r_t^* is taken from Bauer and Rudebusch (2019) and is an average of the three macroeconomic estimates of r^* obtained from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015). Figure 2 (right panel) shows that the model-implied 10-year real yield and the three individual estimates of r^* also closely co-move for most of the sample. The three different r^* estimates diverged from each other before the 1980s. Model I not only captures the low-frequency variations, but its implied r_t^* also exhibits a moderate business cycle component, with dips during recessions and some degree of recovery afterwards.

Figure 3 shows that the model-implied posterior belief of the mean inflation matches very well the trend inflation π_t^* obtained from Bauer and Rudebusch (2019). The trend inflation is a survey-based measure, namely, the Federal Reserve's series on the perceived inflation target rate, denoted as the PTR. It measures the long-run expectations of inflation in the price index of personal consumption expenditures (PCE). The result confirms our assumption in the model that the representative agent forms their inflation

[10]At each time t, the agent perceives that shocks that occur today are negatively serially correlated with future shocks; see proof in Nagel and Xu (2019).



The average and individual r^*s (quarterly data) are obtained from Bauer and Rudebusch (2019) for the period 1971:Q4 to 2017:Q2. The three macroeconomic estimates of r^*s are obtained from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015), respectively. The model-implied 10-year real yields (quarterly data) are from 1962:Q1 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.

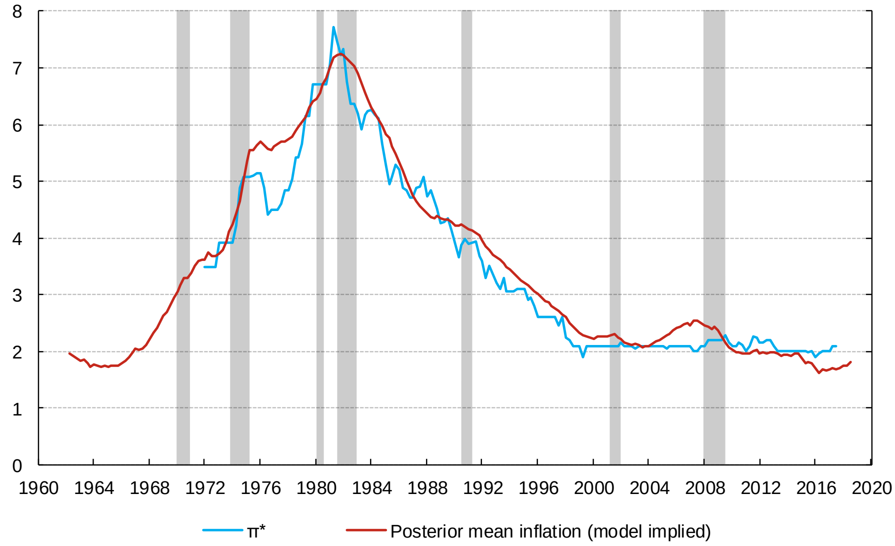
expectations based on past inflation rates.

2.5.4. 10-year Treasury yield and posteriors

While it has long been recognized that nominal interest rates contain a slow-moving trend component (Nelson and Plosser, 1982; Rose, 1988), bond yields in an equilibrium model (and no-arbitrage term structure models in general) are generally modeled as stationary, mean-reverting processes. As a result, it is hard to explain the low-frequency variations in interest rates in such models, and these are mostly attributed to the term-premium component, which is a residual term in empirical models and is usually interpreted as the inflation risk premium in equilibrium models.

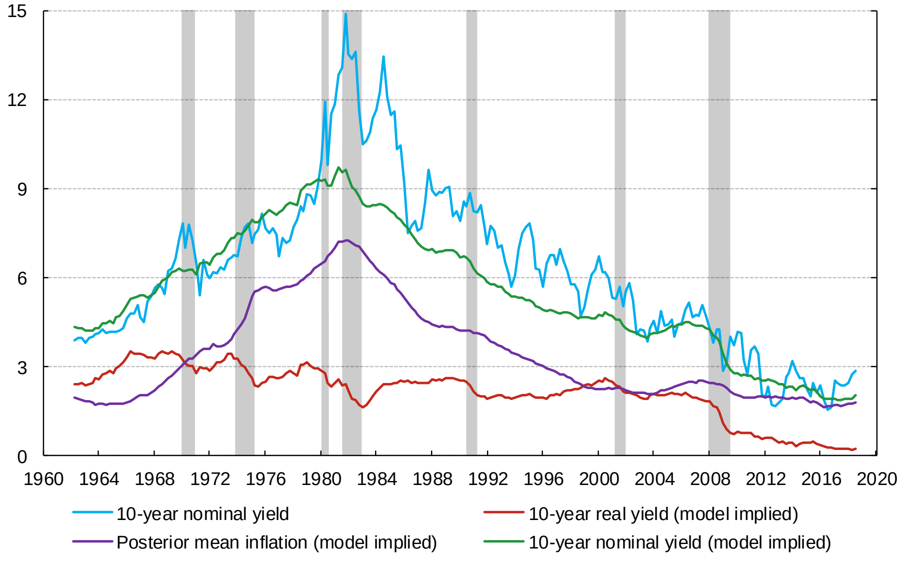
As shown in Figures 2 and 3, the posteriors ($\gamma\tilde{\mu}_{c,t}$ and $\tilde{\mu}_{\pi,t}$) in model I match the macro trends (r_t^* and π_t^*). An illustration of the potential importance of these posteriors in the 10-year nominal yield is provided in Figure 4. The hump-shaped 10-year Treasury yield from the late 1960s to the late 1990s reflects an increase in the inflation expectations before the mid-1980s and a secular decline afterwards. Over the past two decades, as inflation expectations have stabilized, the pronounced decline in the expectations for the output growth and, hence, the 10-year real yield, are what mainly drive the downtrend in nominal yields. As a result, the model-implied 10-year nominal yield captures the trend movements in the 10-year Treasury yield for the whole sample.

Figure 3: π^* **and posterior mean inflation**



The trend inflation π^* (quarterly survey-based PTR measure are obtained from FRB/US data) are obtained from Bauer and Rudebusch (2019) from 1971:Q4 to 2017:Q2. The model-implied posterior belief for the mean inflation (quarterly data) are from 1962:Q1 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.

Figure 4: **10-year nominal yield and macro trends**



The end-of-quarter 10-year nominal yields are from Gürkaynak et al. (2007) for the period 1962:Q1 to 2018:Q2. The model implied 10-year real yield, 10-year nominal yield, and posterior belief for mean inflation (quarterly data) are obtained for the period 1962:Q1 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.

2.5.5. Model-I limitations: cycle and spread in yields

In contrast to the anticipated utility case (Kreps, 1998; Cogley and Sargent, 2008), belief updating is an important risk to the agent, and the posteriors are the only state variables in model I (and, hence, the only source of risk for pricing). We have shown that, with fading memory, the modified Bayesian learning is perpetual, and the resulting posterior beliefs do a good job of tracking the macro trends. Therefore, model I provides a good explanation of the trend in the 10-year Treasury yield. However, model I is silent about two other important features in the yields – the cycle and the term spread.

Historically, short-term bond yields are lower (on average) and much more volatile than long-term bond yields. The spread between long- and short-term bond yields starts to be positive entering recessions, shrinks in the late expansion stage of the business cycle and becomes inverted before recessions. Given the CRRA utility, the term premium is small in model-I, and the model-implied yields for 10- and 1-year bonds are almost identical. Hence, the model is not able to explain the dynamics in the term spread, nor can it explain the cyclical movements in the short-term bond yields. In the next session, we extend model I to overcome these shortcomings.

3. Model II - Learning, trends, and cycles in bond yields

In model II, we still consider an endowment economy with a representative agent who has a CRRA utility function. As in model I, the agent learns with fading memory. But here, both GDP growth and inflation rates are decomposed into two components: one stable component and one transitory/volatile component. The agent learns about mean output growth and inflation rates from the stable component and about the stationary deviations from the mean from the transitory/volatile component. Equilibrium prices adjust such that the agent is happy to consume the output as an endowment.

3.1. Decomposition and learning

The four components of GDP – investment spending, net exports, government spending, and consumption – do not move in lockstep with each other. In fact, their levels of volatility greatly differ. Consumption is highly stable and varies less with the business cycle. In contrast, the other three components vary greatly during economic contractions and expansions. For this reason, we assume that there are two unknowns for the agent to learn about in the output growth process: the long-run mean and a latent stochastic deviation from the mean (stationary). The agent learns about long-run mean GDP growth only from the stable component (PCE), and they learn about the stationary deviation from the mean by using only the volatile component (GDP growth excluding the PCE). Similarly, for inflation, the agent learns about the long-run mean inflation only from core inflation, and they learn about the transitory and stationary deviation from the mean by using only the transitory price changes (the GDP deflator excluding core inflation). Formally, output growth and inflation can be decomposed into the following (account identity):

$$
\begin{aligned}
\Delta g_{t+1} &= \Delta g^*_{t+1} + Gap^g_{t+1} \\
\pi_{t+1} &= \pi^*_{t+1} + Gap^\pi_{t+1},
\end{aligned} \tag{12}
$$

where Δg_{t+1} and π_{t+1} are the total real GDP growth and inflation, respectively. Δg^*_{t+1} and π^*_{t+1} are growth in real consumption (scaled by total real GDP $\frac{C_{t+1}-C_t}{GDP_t}$) and core inflation (scaled by total price level $\frac{P^{core}_{t+1}-P^{core}_t}{P_t}$), respectively. Gap^g_{t+1} and Gap^π_{t+1} are the total GDP growth rate excluding Δg^*_{t+1} and the total inflation rate excluding π^*_{t+1}, respectively.

Both real consumption growth and core inflation follow the i.i.d. laws of motion as follows:

$$
\begin{aligned}
\Delta g^*_{t+1} &= \mu^*_c + \sigma_c \varepsilon^*_{c,t+1} \\
\pi^*_{t+1} &= \mu^*_\pi + \sigma_\pi \varepsilon^*_{\pi,t+1},
\end{aligned} \tag{13}
$$

where $\varepsilon^*_{c,t+1}$ and $\varepsilon^*_{\pi,t+1}$ are the i.i.d. normal shocks. As with model I, the representative agent knows that both Δg^*_{t+1} and π^*_{t+1} are i.i.d., and they also know σ_c and σ_π but not the long-run mean μ_c^* and μ_π^*. The agent forms expectations about μ_c^* and μ_π^*, based on the same learning scheme as in model I, with the same posteriors as follows:

$$\begin{aligned}
\mu_c^*|H_t^{g^*} &\sim N\left(\tilde{\mu}_{c,t}^*, \upsilon_c^*\sigma_c^2\right) \\
\mu_\pi^*|H_t^{\pi^*} &\sim N\left(\tilde{\mu}_{\pi,t}^*, \upsilon_\pi^*\sigma_\pi^2\right),
\end{aligned} \tag{14}$$

where

$$\begin{aligned}
\tilde{\mu}_{c,t}^* &= \tilde{\mu}_{c,t-1}^* + \upsilon_c^*\left(\Delta g_t^* - \tilde{\mu}_{c,t-1}^*\right) \\
\tilde{\mu}_{\pi,t}^* &= \tilde{\mu}_{\pi,t-1}^* + \upsilon_\pi^*\left(\pi_t^* - \tilde{\mu}_{\pi,t-1}^*\right),
\end{aligned} \tag{15}$$

and the same predictive distribution,

$$\begin{aligned}
\Delta g_{t+j}^*|H_t^{g^*} &\sim N\left(\tilde{\mu}_{c,t}^*, (1+\upsilon_c^*)\,\sigma_c^2\right) \\
\pi_{t+j}^*|H_t^{\pi^*} &\sim N\left(\tilde{\mu}_{\pi,t}^*, (1+\upsilon_\pi^*)\,\sigma_\pi^2\right),
\end{aligned} \tag{16}$$

where $j = 1, 2, ..., H_t^{g^*} \equiv \{\Delta g_0^*, \Delta g_1^*, ...\Delta g_t^*\}$, and $H_t^{\pi^*} \equiv \{\pi_0^*, \pi_1^*, ...\pi_t^*\}$. Both Gap_{t+1}^g and Gap_{t+1}^π are assumed to contain a latent stationary component as in Bansal and Yaron (2004) and Piazzesi and Schneider (2007):

$$\begin{aligned}
Gap_{t+1}^g &= x_{c,t+1} + \sigma_c^{gap}\varepsilon_{c,t+1}^{gap} \\
Gap_{t+1}^\pi &= x_{\pi,t+1} + \sigma_\pi^{gap}\varepsilon_{\pi,t+1}^{gap} \\
x_{c,t+1} &= \rho_c x_{c,t} + \sigma_c^x\varepsilon_{c,t+1}^x \\
x_{\pi,t+1} &= \rho_\pi x_{\pi,t} + \sigma_\pi^x\varepsilon_{\pi,t+1}^x,
\end{aligned} \tag{17}$$

where $\varepsilon_{c,t+1}^{gap}$, $\varepsilon_{\pi,t+1}^{gap}$, $\varepsilon_{c,t+1}^x$, and $\varepsilon_{\pi,t+1}^x$ are i.i.d. normal shocks. The representative agent knows all of the parameters but not $x_{c,t+1}$ and $x_{\pi,t+1}$. They form expectations about $x_{c,t+1}$ and $x_{\pi,t+1}$, based on the same learning scheme as for the long-run mean but with potentially different geometric weighting parameters, υ_c^{gap} and υ_π^{gap}. The posteriors are given by the following:

$$
\begin{aligned}
x_{c,t+1}|H_{g,t}^{gap} &\sim N\left(\rho_c \tilde{x}_{c,t}, \upsilon_c^{gap}\left((\sigma_c^x)^2 + (\sigma_c^{gap})^2\right)\right)\\
x_{\pi,t+1}|H_{\pi,t}^{gap} &\sim N\left(\rho_\pi \tilde{x}_{\pi,t}, \upsilon_\pi^{gap}\left((\sigma_\pi^x)^2 + (\sigma_\pi^{gap})^2\right)\right)\\
\tilde{x}_{c,t} &= \rho_c \tilde{x}_{c,t-1} + \upsilon_c^{gap}\left(Gap_t^g - \rho_c \tilde{x}_{c,t-1}\right)\\
\tilde{x}_{\pi,t} &= \rho_\pi \tilde{x}_{\pi,t-1} + \upsilon_\pi^{gap}\left(Gap_t^\pi - \rho_\pi \tilde{x}_{\pi,t-1}\right),
\end{aligned} \tag{18}
$$

where $H_{g,t}^{gap} \equiv \{Gap_0^g, Gap_1^g, ...Gap_t^g\}$ and $H_{\pi,t}^{gap} \equiv \{Gap_0^\pi, Gap_1^\pi, ...Gap_t^\pi\}$. As discussed in 2.3, the updating is the same as for an optimal Kalman filtering with an appropriate choice of parameter values.

To better understand the stochastic nature of the output growth and inflation processes from the agent's subjective viewpoint, we calculate the total predictive distribution

$$
\begin{aligned}
\Delta g_{t+j}|H_t^g &\sim N\left(\tilde{\mu}_{c,t}^* + \rho_c \tilde{x}_{c,t}, \left(1+\upsilon_c^*\right)\sigma_c^2 + \left(1+\upsilon_c^{gap}\right)\left((\sigma_c^x)^2 + (\sigma_c^{gap})^2\right)\right)\\
\pi_{t+j}|H_t^\pi &\sim N\left(\tilde{\mu}_{\pi,t}^* + \rho_\pi \tilde{x}_{\pi,t}, \left(1+\upsilon_\pi^*\right)\sigma_\pi^2 + \left(1+\upsilon_\pi^{gap}\right)\left((\sigma_\pi^x)^2 + (\sigma_\pi^{gap})^2\right)\right),
\end{aligned} \tag{19}
$$

where $j = 1, 2, ...$ and the variance of the predictive distribution contains both the uncertainty due to future shocks and the uncertainty about μ_c^*/μ_π^* and $x_{c,t+1}/x_{\pi,t+1}$. H_t^g contains both $H_t^{g^*}$ and $H_{g,t}^{gap}$, and H_t^π contains both $H_t^{\pi^*}$ and $H_{\pi,t}^{gap}$.

3.2. Bond pricing

The Euler equation holds under the representative agent's subjective expectations, and the log nominal pricing kernel is the same as in model I. The time-t price of a zero-coupon bond satisfies the same recursion as in equation (10).

Model I has two state variables (the posterior means for both the output growth and inflation) that explain the trend in the long-term yields. However, the different GDP growth and inflation components appear to have very different dynamics in the data; hence, we allow the agent to learn the long-run mean and cyclical components from the data separately in this model. Model II has four state variables: $\tilde{\mu}_{c,t}^*$, $\tilde{\mu}_{\pi,t}^*$, $\tilde{x}_{c,t}$, and $\tilde{x}_{\pi,t}$.

Given the linear Gaussian framework, we assume that $p_t^{(n)} = log(P_t^{(n)})$ is a linear function of these state variables $\tilde{\mu}_t^* = (\tilde{\mu}_{c,t}^*, \tilde{\mu}_{\pi,t}^*)^T$ and $x_t = (\tilde{x}_{c,t}, \tilde{x}_{\pi,t})^T$:

$$p_t^{(n)} = -A^{(n)} - B^{(n)} x_t - C^{(n)} \tilde{\mu}_t^*. \tag{20}$$

When we substitute $p_t^{(n)}$ and $p_{t+1}^{(n-1)}$ in the Euler equation (10), the coefficients in the pricing equation can be solved with $B^{(n)} = B^{(n-1)}\rho + v'\rho$, $C^{(n)} = C^{(n-1)} + v' = v'n$, and $A^{(n)} = A^{(n-1)} + A^{(1)} - 0.5 * Var_t \left(p_{t+1}^{(n-1)}\right) - Cov_t \left(p_{t+1}^{(n-1)}, m_{t,t+1}^{\$}\right)$ (see the appendix for details). The subjective excess return is $er_{n,t+1} = -Cov_t \left(r_{n,t+1}, m_{t,t+1}^{\$}\right) = -v'Cov_t \left(z_{t+1}, \tilde{\mu}_{t+1}^*\right) C^{(n-1)} - v'Cov_t \left(z_{t+1}, x_{t+1}\right) B^{(n-1)}$. All of the variance and covariance terms are relatively small in the data. Hence, given the CRRA utility, the term premium is small in this model.

As we can see from the solution, the yield parameter ($\frac{C^{(n)}}{n}$) for the posterior mean $\tilde{\mu}_t^*$ is constant over horizon n; therefore, the impacts of $\tilde{\mu}_t^*$ on the long- and short-term yields are the same, which explains the low-frequency movements (trend) in the yields. However, the yield parameter for x_t, $\frac{B^{(n)}}{n}$, is decreasing over horizon n. Hence, the impact of x_t on the short-term yield is bigger than on the long-term yield, which captures the cyclical movements in the short-term yields. The spread between the long- and short-term yields is mainly driven by the cyclical component x_t, which could be positive or negative for many periods (depending on the persistence parameter ρ). Still, in contrast with the data, the model-II-implied spread is mean zero because of the stationarity assumption for x_t. To solve the price and yields for real bonds, we can simply replace v' with $v' = (\gamma, 0)$.

3.2.1. Taylor Rule interpretation

The one-quarter-ahead nominal yield from equation (20) is a market-based short rate that reflects investors' expectations (of growth and inflation). Given its performance in matching short-term interest rates in the data (see Section 3.3), one natural question is: What is the connection with the Taylor Rule (Taylor, 1993)? To answer this question, the one-quarter-ahead rate from equation (20) can be rewritten as follows:

$$
\begin{aligned}
y_t^{(1)} &= A^{(1)} + \gamma\tilde{\mu}^*_{c,t} + \tilde{\mu}^*_{\pi,t} + \rho_\pi \tilde{x}_{\pi,t} + \gamma\rho_c \tilde{x}_{c,t} \\
&= r_t^* + \tilde{\mu}^*_{\pi,t} + \rho_\pi \tilde{x}_{\pi,t} + \gamma\rho_c \tilde{x}_{c,t} \\
&= i_t^* + \rho_\pi \tilde{x}_{\pi,t} + \gamma\rho_c \tilde{x}_{c,t},
\end{aligned}
\tag{21}
$$

where the neutral real rate of interest $r_t^* = A^{(1)} + \gamma\tilde{\mu}^*_{c,t}$ and the neutral nominal rate $i_t^* = r_t^* + \tilde{\mu}^*_{\pi,t}$. $\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$ are the short-run growth rate and inflation expecations, respectively. The nominal short rate in equation (21) is similar to a Taylor Rule specification in the literature, with a few exceptions. The first difference is that r_t^* and i_t^* move endogenously, instead of as constants. Secondly, $\tilde{x}_{c,t}$ is the short-run growth-rate deviation from its long-run mean and not an output deviation from its potential in level. Finally, $\tilde{x}_{\pi,t}$ is the short-run inflation deviation from its long-run mean and not a realized inflation deviation from a constant target.

3.3. Empirical findings

We use the same data sets as in model I, but the real GDP growth and rate of inflation are decomposed into one stable component and one transitory component, respectively.[11] We can then calculate the posterior beliefs for the long-run mean and the stationary deviation from the mean. Again, we show that the model-implied r_t^* closely matches the estimated r_t^* in the literature and the total posterior belief for inflation closely matches the survey-based trend in inflation. As a result, the model-implied 10-year nominal bond yields closely match the historical trend movements in the 10-year Treasury yields. In addition, the model can also capture the cyclical movements in 1-year Treasury yields.

[11] Note that the core inflation data are available starting from 1959.Q2. Therefore, we use about 10 years (1959.Q2 - 1968.Q3) of data as a training period for the posterior beliefs and show the results starting from 1968.Q3.

γ	β	υ_c^*	υ_π^*	υ_c^{gap}	υ_π^{gap}	$corr^*$
4	1.0245	0.015	0.05	0.12	0.2	-0.14
ρ_c	ρ_π	σ_c	σ_π	σ_{cx}	$\sigma_{\pi x}$	$corr^{gap}$
0.92	0.98	0.42	0.33	0.64	0.35	-0.04

Table 2: Configuration of model-II parameters
Table 2 shows the parameter values for the output growth and inflation processes and for the constant gain in learning. All of the parameters are given in quarterly terms. The means and standard deviations are in percentages.

3.3.1. Parameters

The volatility parameters for the consumption growth and core inflation are calibrated to match their counterparts in the data. As shown in the appendix, even though we have two different parameters for growth gap (inflation gap) volatility, σ_c^x and σ_c^{gap} (σ_π^x and σ_π^{gap}), these parameters can be considered as one for the model's solution; hence, $\sigma_{cx} = \sqrt{(\sigma_c^x)^2 + (\sigma_c^{gap})^2}$ ($\sigma_{\pi x} = \sqrt{(\sigma_\pi^x)^2 + (\sigma_\pi^{gap})^2}$) is calibrated to match the volatility in the transitory GDP growth (transitory inflation). In the model, the correlations between output growth and inflation (for both the stable and transitory components) are calibrated to match the correlations in the data.

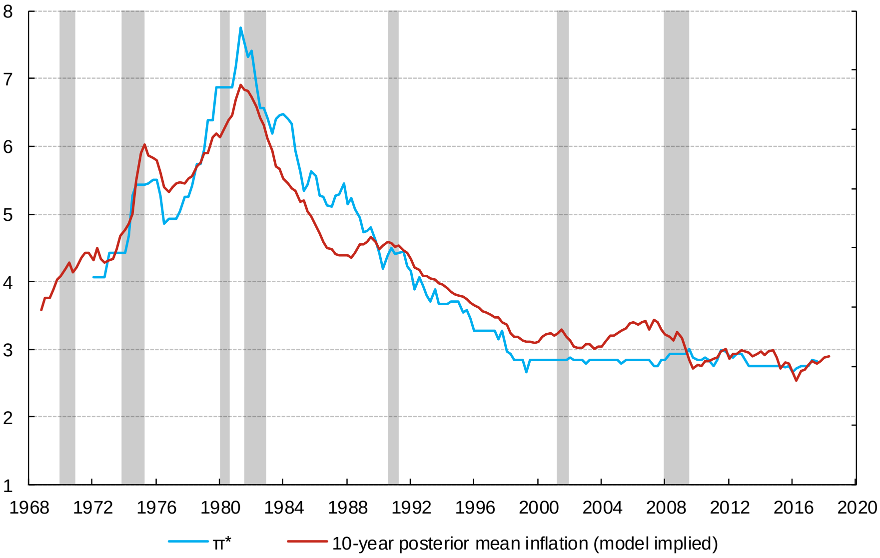
The parameters for learning in inflation υ_π^*, ρ_π, and υ_π^{gap} are calibrated to match the variations in π_t^* and the one-quarter-ahead mean survey inflation expectation (persistence and volatility). The parameters for learning in growth υ_c^*, ρ_c, and υ_c^{gap} are calibrated to match the variations in r_t^* and in the 1-year nominal yields (persistence and volatility).[12] We follow the literature and set the risk aversion as 4, and the time preference, β, is calibrated to match the level of the 10-year nominal yields in the data. The resulting parameter values are reported in Table 2.

3.3.2. Posteriors versus r_t^* and π_t^*

It is well understood by investors that some components of inflation and GDP growth are more volatile than others. Hence, it is natural for the agent to separately learn the long-run mean and the stationary deviation from the mean. The agent uses a longer

[12]These parameter values imply that investors learn the long-run mean growth (inflation rate) through an effective sample size of 16 years (5 years), and they learn the business cycle deviations through an effective sample size of 1-2 years.

Figure 5: π^* and 10-year posterior mean inflation



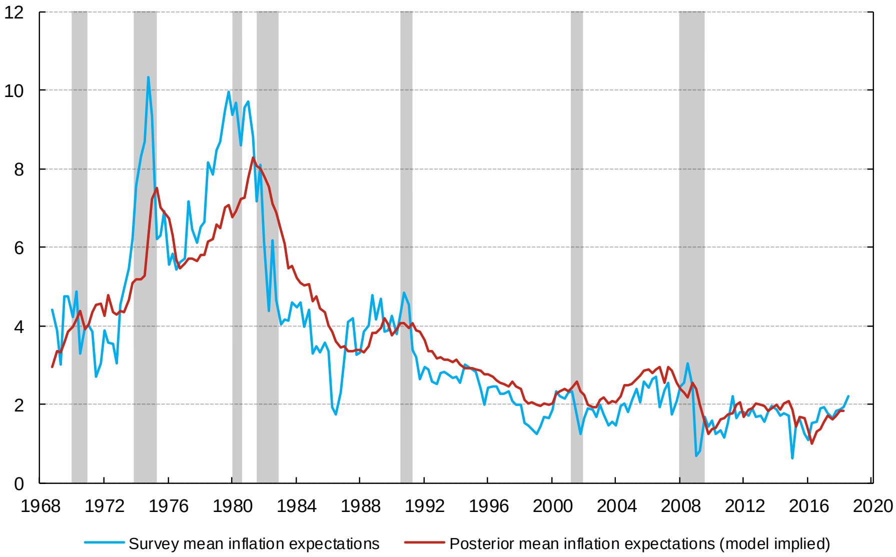
The trend inflation π^* (the quarterly survey-based PTR measure obtained from FRB/US data) are obtained from Bauer and Rudebusch (2019) for the period from 1971:Q4 to 2017:Q2. The model-implied total posterior belief for the 10-year-ahead mean inflation (quarterly) was obtained for the period from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.

effective sample size to learn the long-run mean and a shorter effective sample size to learn the short-run deviation ($\upsilon_c^* < \upsilon_c^{gap}$ and $\upsilon_\pi^* < \upsilon_\pi^{gap}$). With the same learning with memory loss, the representative agent slowly but perpetually updates their posterior beliefs.

Figure 5 shows that the model-implied total posterior belief for the 10-year-ahead mean inflation matches the survey-based PTR trend inflation π_t^*. Similarly, Figure 6 shows that the model-implied total posterior belief for the 1-quarter-ahead mean inflation also closely tracks the survey mean for the 1-quarter-ahead inflation from the SPF. The result confirms our assumption that the agent forms their inflation expectations differently for the long-run versus the short-run. Figure 7 shows the model-implied r_t^*, which is a linear function of only the posterior for the long-run mean growth $\gamma\tilde{\mu}_{c,t}^* + Cov$ closely tracks the three macroeconomic estimates of r^*.[13]

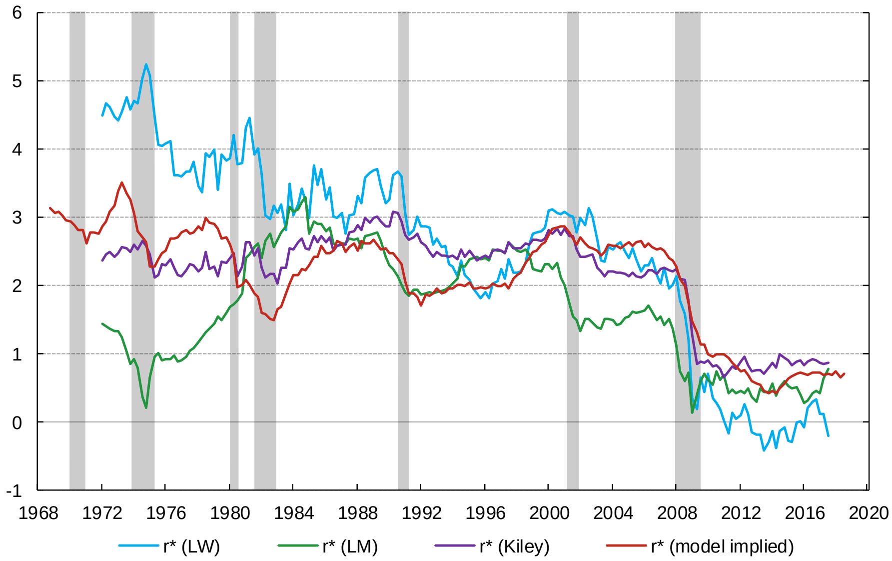
[13]To be consistent with the concept of r^*, the short-run effect from $\tilde{x}_{c,t}$ on the real yield is not included for our calculation of the model-implied r^* because it will eventually vanish. Therefore, the model-implied r^* equally affects the real yields of any maturity.

Figure 6: One-quarter-ahead inflation expectations - posterior vs. survey



The mean survey one-quarter-ahead inflation (quarterly data) are obtained from the Philadelphia Fed's SPF from 1968:Q3 to 2018:Q2. The model-implied total posterior belief for 1-quarter-ahead mean inflation (quarterly) covers the period from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.

Figure 7: Individual r* and model implied r*



The individual r^*s (quarterly data) are obtained from Bauer and Rudebusch (2019) for the period from 1971:Q4 to 2017:Q2. The three macroeconomic estimates of r^*s are obtained from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015). The model-implied r^* (quarterly) covers the period from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.

3.3.3. One-year and 10-year Treasury yields, and their spread

In model I, we show the importance of macro trends in explaining the movements in the long-term nominal yield. For the same reason, the posteriors for the long-run mean growth and inflation move slowly and match the macro trends in model II, and Figure 8 shows the importance of these posteriors in the 10-year nominal yield.[14] The secular decline in the 10-year Treasury yield after the 1980s was mainly driven by a combination of two phenomena: first, a downtrend in inflation expectations and then a steady decline in r_t^*.

Due to the lack of cyclical movements in the posteriors, model I is limited to explaining only the trend in the long-term yields. In model II, however, learning about the deviation from the mean allows the model to also capture the cyclical movements in the short-term yields. Figure 9 (left panel) shows that the model-implied 1-year nominal yield tracks the historical data for the 1-year Treasury yield relatively well, with some exceptions. For the post-global financial crisis period, the 1-year Treasury yields were mostly constrained by the zero lower bound, but the model-implied 1-year nominal yields were much more volatile, and these two bonds recently began to line up again.[15] Also at the beginning of the sample period, the model-implied 1-year nominal yield seems to have been more volatile, which suggests that the agent could have been using different constant-gain parameters for these periods.

Another limitation of model I is that the spread between the long- and short-term bond yields is almost zero, due to the CRRA utility and the equal impacts of the posteriors on the yields of any maturity. In model II, it is still true that the impacts of $\tilde{\mu}_t^*$ on the yields of all maturities are the same (capturing the long-run trends). But, as shown in the solution, the impact of x_t (as AR(1) processes) on the short-term yield is bigger

[14]The long-term yield is less sensitive to the short-run belief movements, and its variations (trend and business cycle) are mainly driven by updating the beliefs in the long-run mean growth and inflation rates (as in model I).

[15]Note that the whole nominal yield curve for the post-global financial crisis period is distorted by the zero lower bound, quantitative easing, and other unconventional monetary policies. The model-implied yield curve for these periods is very different from the data, which can be used as shadow rates, and they started to line up again after 2015.

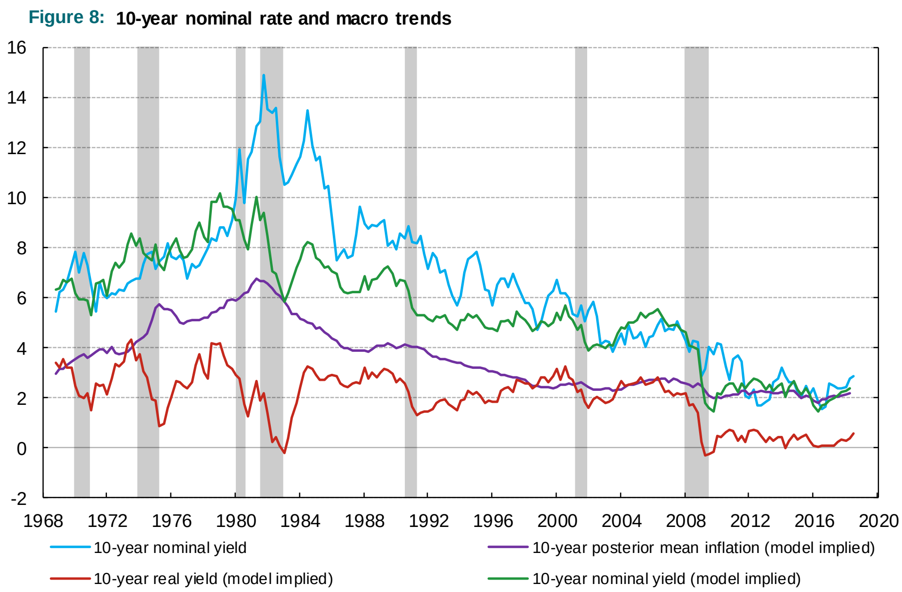
than on the long-term yield, which implies that when the short-run beliefs are negative (positive), the spread would be positive (negative). Hence, model II can generate cyclical movements in the short-term yield and in the spread, mostly due to variations in x_t. Figure 9 (right panel) shows that the model-implied 10-year-minus-1-year yield spread tracks the business cycle movements in Treasury yield spreads quite well.

Starting from the previous trough, the short-term nominal yield starts to rise when the agent begins to revise their beliefs about the short-run growth and inflation rates upwards towards their long-run means (the short-run deviations are still negative and the spread is positive), and the spread starts to shrink as these short-run deviations are turning from negative to positive. This pattern continues until the late expansion stage (or peak) when both the short-run growth rate and inflation expectations are above their long-run means (short-run deviations are positive now), which implies an inverted yield curve. From the previous trough to peak, both the inflation gap and the output gap move from negative to positive, and monetary policy turns from accommodative to contractionary (short-term yields increase), similar to the Taylor Rule. The agent then begins to revise their short-run beliefs sharply downwards, entering a recession (from positive to negative), and the spread switches from negative to positive.

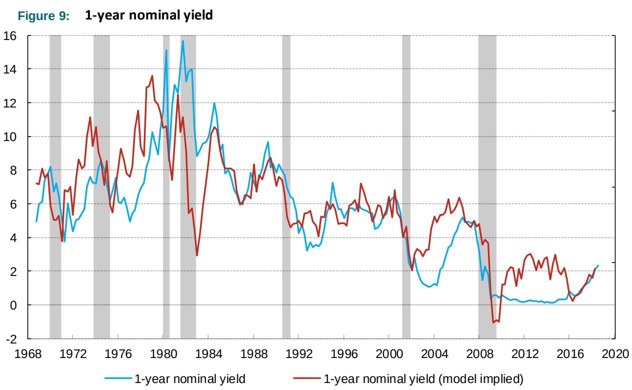
3.3.4. Inverted curves, secular stagnation, and recessions

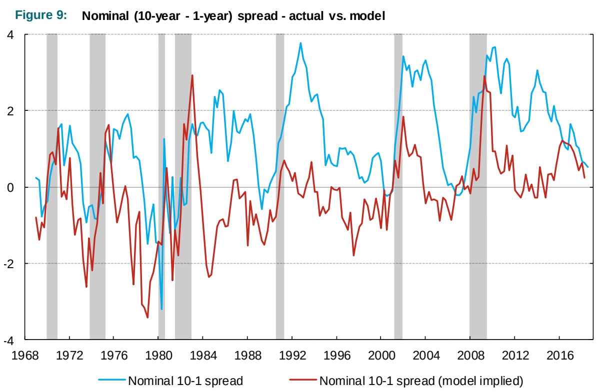
Figure 10 shows that (1) the posteriors for the short-run deviations ($\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$) moved in opposite directions before 2000 and in the same direction afterwards, and (2) $\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$ were persistently negative for most of the post-2000 period.[16] Given the bond yield solution in Section 3.2, the impacts of $\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$ on short-term nominal yields (and hence on the nominal spread) likely canceled each other out before 2000, and they simultaneously lowered the short-term nominal yields (and hence increased the nominal spread) afterwards. Therefore, the model can generate an upward shift in the nominal spread that is consistent with the data in Figure 1. These observations are also consistent

[16] The first observation is consistent with the finding that inflation was bad news for future growth before 2000 (Piazzesi and Schneider, 2007) and switched to good news afterwards (Burkhardt and Hasseltoft, 2012; David and Veronesi, 2013; Campbell et al., 2017; Zhao, 2020).



The end-of-quarter 10-year nominal yields are obtained from Gürkaynak et al. (2007) for the period from 1968:Q3 to 2018:Q2. The model-implied 10-year real yield, 10-year nominal yield, and the total posterior belief for the mean inflation (quarterly data) are for the period from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.





The end-of-quarter 1-year nominal yields and the 10-year nominal - 1-year nominal yield spreads are obtained from Gürkaynak et al. (2007) for the period from 1968:Q3 to 2008:Q2. The model-implied 1-year nominal yields and the 10-year nominal - 1-year nominal yield spreads are for the period from 1968:Q3 to 2018:Q3. The gray bars represent periods of recession as defined by the NBER.

with the secular stagnation statement in Summers (2014) that the output gap ($\tilde{x}_{c,t}$ in this model) and the inflation gap ($\tilde{x}_{\pi,t}$ in this model) were persistently negative after 2000. The secular decline in r_t^* and the 10-year Treasury yield were mainly driven by the decline in the posteriors for the long-run means, and the declines in short- to mid-term Treasury yields were further due to the persistently negative $\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$.

From a monetary policy point of view, the Federal Reserve faced a trade-off between short-run inflation and growth pre-2000 (the U.S. economy was mostly hit by supply shocks); the Fed had to raise the short-rate to either lower inflation while the output gap was negative (1970-1985) or to lower the output gap while the inflation gap was negative (1985-2000). However, starting in 2000, the U.S. economy was mostly hit by demand shocks, and the Fed could stay low for a longer period of time without facing a trade-off between the inflation and output gaps. Hence, we observed more-frequent recessions for the period before the late 1990s and less-frequent recessions (or longer business cycles) for the period afterwards.

3.3.5. Model II limitations: upward-sloping yield curve

Given the trend ($\tilde{\mu}_t^*$) and cycle (x_t) as the posterior beliefs, model II is able to match both the long-run trend and the cyclical variations in the yields. However, despite the fact that the model-II-implied nominal spread can stay positive or negative for an extended period of time at different phases of the business cycle, the level was almost in parallel lower than data (as shown in Figure 9) due to the stationary assumption for x_t and the CRRA utility (hence, the model-implied spread is stationary and mean zero). But, in the data, the spread between the long- and short-term bond yields was positive, in general, and only became inverted before recessions.

The standard equilibrium mechanism used to generate an upward-sloping nominal yield curve is the inflation-risk-premium approach by Piazzesi and Schneider (2007), which relies on inflation as bad news for future growth and the assumption that agents prefer an early resolution to uncertainty. Given that inflation switched from bad to good news for future growth over the past two decades, we extend model II in the next section by incorporating an alternative worst-case belief approach (Zhao, 2020). The model-III-implied short-rate expectations (both nominal and real) are upward-sloping under the agent's equilibrium worst-case belief and, hence, is consistent with the data; the model-implied spread is positive on average.

4. Model III - Learning, trends, cycles, and spreads in bond yields

In model III, we continue to consider an endowment economy with a representative agent who has a CRRA utility function. However, the agent is assumed to have limited information about the stochastic environment and, hence, faces both risk and ambiguity. Here, the risk refers to the situation where there is a probability law that guides the choice. At the same time, the ambiguity-averse agent (with recursive multiple priors or a maxmin preference by Epstein and Schneider 2003) lacks the confidence to assign probabilities to all of the relevant events. Instead, they act as if they are evaluating future prospects using a worst-case probability drawn from a set of multiple distributions.

Investors in this economy have in mind a benchmark or reference measure of the economy's dynamics that represents the best estimate of the stochastic process. In model III, the reference measure is the full stochastic environment in model II (including the posteriors). But the agent is concerned that the reference measure is misspecified and believes that the true measure is actually within a set of alternative measures that are statistically close to the reference distribution.[17] Equilibrium prices adjust such that the ambiguity-averse agent is happy to consume the output as an endowment.

[17] The empirical literature found that professional forecasters, and even central banks, make systematic forecast errors by comparing the mean forecasts and the subsequent realized values (Faust and Wright, 2009; Cieslak, 2018). In this model, the reference measure represents the best point estimate from the data and the agents would use it for forecasting. However, the agents are more cautious when they make decisions and, instead, they use the worst-case belief for decision making. Therefore, ambiguity in this model provides a rational explanation for the expectation errors.

4.1. Learning and ambiguity

We assume the total predictive distribution from model II as the reference measure, which can be rewritten as follows:

$$\begin{aligned}\Delta g_{t+1} &= \tilde{\mu}^*_{c,t} + \rho_c \tilde{x}_{c,t} + \varepsilon_{c,t+1} \\ \pi_{t+1} &= \tilde{\mu}^*_{\pi,t} + \rho_\pi \tilde{x}_{\pi,t} + \varepsilon_{\pi,t+1},\end{aligned} \tag{22}$$

where $\varepsilon_{c,t+1}$ is a combination of $\tilde{\varepsilon}^*_{c,t+1}$, $\tilde{\varepsilon}^{gap}_{c,t+1}$, and $\tilde{\varepsilon}^x_{c,t+1}$. $\varepsilon_{\pi,t+1}$ is a combination of $\tilde{\varepsilon}^*_{\pi,t+1}$, $\tilde{\varepsilon}^{gap}_{\pi,t+1}$, and $\tilde{\varepsilon}^x_{\pi,t+1}$. $\tilde{\mu}^*_{c,t}$, $\tilde{\mu}^*_{\pi,t}$, $\tilde{x}_{c,t}$, and $\tilde{x}_{\pi,t}$ have exactly the same dynamics as in model II. However, the agent is concerned that their reference measure is misspecified and that the true measure is actually within a set of alternative measures that are statistically close to the reference measure. The set of alternative measures is generated by a set of different mean output growth (inflation) rates around the reference mean value $\tilde{\mu}^*_{c,t} + \rho_c \tilde{x}_{c,t}$ ($\tilde{\mu}^*_{\pi,t} + \rho_\pi \tilde{x}_{\pi,t}$). Specifically, under alternative measure $p^{\tilde{a}}$, the output growth and inflation rates are as follows:

$$\begin{aligned}\Delta g_{t+1} &= \tilde{a}_{c,t} + \tilde{\mu}^*_{c,t} + \rho_c \tilde{x}_{c,t} + \tilde{\varepsilon}_{c,t+1} \\ \pi_{t+1} &= \tilde{a}_{\pi,t} + \tilde{\mu}^*_{\pi,t} + \rho_\pi \tilde{x}_{\pi,t} + \tilde{\varepsilon}_{\pi,t+1},\end{aligned} \tag{23}$$

where $\tilde{a}_{c,t} \in A_{c,t} = [\tilde{\mu}^*_{c,t} + \rho_c \tilde{x}_{c,t} - a_{c,t}, \tilde{\mu}^*_{c,t} + \rho_c \tilde{x}_{c,t} + a_{c,t}]$ and $\tilde{a}_{\pi,t} \in A_{\pi,t} = [\tilde{\mu}^*_{\pi,t} + \rho_\pi \tilde{x}_{\pi,t} - a_{\pi,t}, \tilde{\mu}^*_{\pi,t} + \rho_\pi \tilde{x}_{\pi,t} + a_{\pi,t}]$ with both $a_{c,t}$ and $a_{\pi,t}$ being positive. Each trajectory of $\tilde{a}_t$ will yield an alternative measure $p^{\tilde{a}}$ for the joint process. A larger $a_{c,t}(a_{\pi,t})$ implies that investors are less confident about the reference distribution.[18]

Using the forecast dispersion from the Blue Chip Financial Forecast (BCFF) survey as a measure for the size of the ambiguity, Zhao (2020) finds that, before the late 1990s, the size of the ambiguity for long-horizon inflation was bigger than that for short horizons and that this pattern was reversed afterwards. However, the size of the ambiguity for

[18] Ilut and Schneider (2014) link the size of ambiguity with the observed volatility under the reference measure and provide a detailed discussion for the source of ambiguity.

long-term real output growth was always smaller than those for the short term. Together with the fact that the positive inflation shocks changed from negative to positive news about future growth in the past 20 years, the ambiguity-averse agent chose the upper bound ($a_{\pi,t}$) as the worst-case measure for inflation before the late 1990s and the lower bound ($-a_{\pi,t}$) as the worst-case measure afterwards, while for output growth (as the endowment), the agent always chose the lower bound ($-a_{c,t}$) as the worst-case measure in equilibrium. Zhao (2020) showed that the expectations hypothesis roughly holds under investors' worst-case beliefs and the upward-sloping nominal and real yield curves are mainly driven by the upward-sloping nominal and real short-rate expectations.

We follow Zhao (2020) and model $a_{c,t}$ ($a_{\pi,t}$) as a random walk with drift as follows:

$$\begin{aligned} a_{c,t+1} &= \mu_c^a + a_{c,t} + \sigma_{ac}\varepsilon_{ac,t+1} + \sigma_a^{ac}\varepsilon_{a,t+1} \\ a_{\pi,t+1} &= \mu_\pi^a + a_{\pi,t} + \sigma_a^{a\pi}\varepsilon_{a,t+1}, \end{aligned} \tag{24}$$

where μ_c^a and μ_π^a are the drift parameters, which can be positive or negative. $a_{c,t}$ and $a_{\pi,t}$ are driven by a common exogenous shock $\varepsilon_{a,t+1}$, where the coefficients σ_a^{ac} and $\sigma_a^{a\pi}$ capture the correlation between them. $\varepsilon_{ac,t+1}$ is an $a_{c,t}$ specific shock that captures the difference of these two.

4.2. Preference: recursive multiple priors

Piazzesi and Schneider (2007) show the importance of the Epstein and Zin (1989) preference in generating a sizable inflation risk premium. To illustrate the key role of the ambiguity yields, we assume investors have a recursive multiple-priors preference axiomatized by Epstein and Schneider (2003) but with the same CRRA utility function as in model I:

$$V_t(C_t) = \min_{p_t \in \mathcal{P}_t} \mathbb{E}_{p_t} \left(U\left(C_t\right) + \beta V_{t+1}(C_{t+1}) \right), \tag{25}$$

where $U\left(C_t\right) = \frac{C_t^{1-\gamma}-1}{1-\gamma}$, γ is the coefficient of the risk aversion and β reflects the investor's time preference.

4.3. Bond pricing

The Euler equation holds under the representative agent's worst-case belief, and the log nominal pricing kernel is the same as in model I. The time-t price of a zero-coupon bond satisfies the same recursion in equation (10), but the expectation is under the worst-case belief.

Model II has four state variables (the posterior beliefs $\tilde{\mu}^*_{c,t}$, $\tilde{\mu}^*_{\pi,t}$, $\tilde{x}_{c,t}$, and $\tilde{x}_{\pi,t}$) that explain the trends and cycles in the yields. To generate upward-sloping nominal and real yield curves, we add two more state variables in Model III: $a_{c,t}$ and $a_{\pi,t}$. Given the linear Gaussian framework, we assume that $p_t^{(n)} = log(P_t^{(n)})$ is a linear function of these state variables $\tilde{\mu}^*_t = (\tilde{\mu}^*_{c,t}, \tilde{\mu}^*_{\pi,t})^T$, $x_t = (\tilde{x}_{c,t}, \tilde{x}_{\pi,t})^T$ and $a_t = (a_{c,t}, a_{\pi,t})^T$:

$$p_t^{(n)} = -A^{(n)} - B^{(n)} x_t - C^{(n)} \tilde{\mu}^*_t - D^{(n)} a_t. \tag{26}$$

When we substitute $p_t^{(n)}$ and $p_{t+1}^{(n-1)}$ in the Euler equation (10), the coefficients in the pricing equation can be solved with $B^{(n)} = B^{(n-1)}\rho + v'\rho$, $C^{(n)} = C^{(n-1)} + v' = v'n$, $D^{(n)} = D^{(n-1)} + v'\phi_a = v'n\phi_a$, and $A^{(n)} = A^{(n-1)} + A^{(1)} - 0.5 * Var_t\left(p_{t+1}^{(n-1)}\right) - Cov_t\left(p_{t+1}^{(n-1)}, m^{\$}_{t,t+1}\right) + D^{(n-1)}\mu_a$ (see the appendix for details), where ϕ_a represents the equilibrium choice of the upper or lower bounds, which are equal to -1 or $+1$ on the diagonal. All of the variance and covariance terms are relatively small in the data. Hence, given the CRRA utility, the subjective excess return is small in this model.

As in model II, $\tilde{\mu}^*_t$ explains the low-frequency movements (trend) in the yields and x_t captures the cyclical movements in the short-term yields. The yield parameter ($\frac{D^{(n)}}{n}$) for ambiguity a_t is $v'\phi_a$ (constant over horizon n); hence $a_{c,t}$ lowers the yields for the whole sample period ($\phi^a_c = -1$), and $a_{\pi,t}$ lowers (raises) the yields for the second subperiod when $\phi^a_\pi = -1$ (the first subperiod when $\phi^a_\pi = 1$). The impacts of a_t on the long- and short-term yields are the same, and the upward-sloping nominal and real yield curves are mainly driven by increases in $\frac{A^{(n)}}{n}$ over horizon n due to μ_a. To solve the price and yields for real bonds, we can simply replace v' with $v' = (\gamma,\, 0)$.

4.3.1. Model intuition

Given the closed-form solution, the intuition of the model follows directly from the fact that interest rates reflect investors' worst-case expectations.

First, the agent chooses the lowest growth rate as their worst-case belief in equilibrium; thus, the ambiguity about growth pushes down real yields. Given that the size of the ambiguity for growth is higher for short horizons, short-term real rates are pushed down by more than long-term real rates. Therefore, the real yield curve slopes upward.

Second, the ambiguity about the inflation rate contributes to an upward-sloping nominal yield curve but for different reasons in the two regimes. Pre-2000, positive inflation shocks were bad news for future growth – the worst-case inflation was the highest rate; thus, the ambiguity about inflation pushed up nominal yields. Since there was more ambiguity about long-run inflation, the long-term nominal yields were pushed up by more than just the short-term nominal yields. Post-2000, positive inflation shocks were good news for future growth – the worst-case inflation was the lowest rate; thus, the ambiguity about inflation pushed down the nominal yields. But now there is more ambiguity about the short-run inflation, and the short-term nominal yields are pushed down farther than the long-term nominal yields are. In both cases, the model implies an upward-sloping and steeper nominal yield curve.

4.4. Empirical findings

We use the same data sets as in model II. We also use the forecast dispersions for the real output growth and inflation obtained from the Philadelphia Fed's SPF as a measure for the realized size of the ambiguity. Then we can calculate the realized values for all of the state variables and, hence, the model-implied yields. Since the only change in model III is the ambiguity, the model can still match the historical trends and cycles in the yields as in model II. In addition, the model generates upward-sloping nominal and real yield curves as in the data.

γ	β	υ_c^*	υ_π^*	υ_c^{gap}	υ_π^{gap}	$corr^*$
4	1.0255	0.015	0.05	0.12	0.2	-0.14
ρ_c	ρ_π	σ_c	σ_π	σ_{cx}	$\sigma_{\pi x}$	$corr^{gap}$
0.92	0.98	0.42	0.33	0.64	0.35	-0.04
	μ_c^a	μ_π^a	σ_{ac}	σ_a^{ac}	$\sigma_a^{a\pi}$	
Period 1	-0.0044	0.0029	0.012	-0.009	0.010	
Period 2	-0.0043	-0.0063	0.009	0.0047	0.015	

Table 3: Configuration of model-III parameters

Table 3 reports the parameter values for the output growth and inflation processes, the constant gain in learning, and the ambiguity process. All of the parameters are given in quarterly terms. The means and standard deviations are in percentages.

4.4.1. Parameters

All of the parameters (excluding the ambiguity parameters) are estimated the same way as in model II. The ambiguity parameters are the same as in Zhao (2020), where the whole sample period is split into two subperiods and the parameters are different for each subperiod (mainly the trend parameter for the inflation ambiguity μ_π^a). The realized size of the ambiguity is measured by the past one-year average of the SPF forecast dispersions that are calculated by the 60th percentile minus the 40th percentile of the individual forecasts. The resulting parameter values are reported in Table 3.

4.4.2. Trends, cycles, and spreads in the yields, and r_t^*

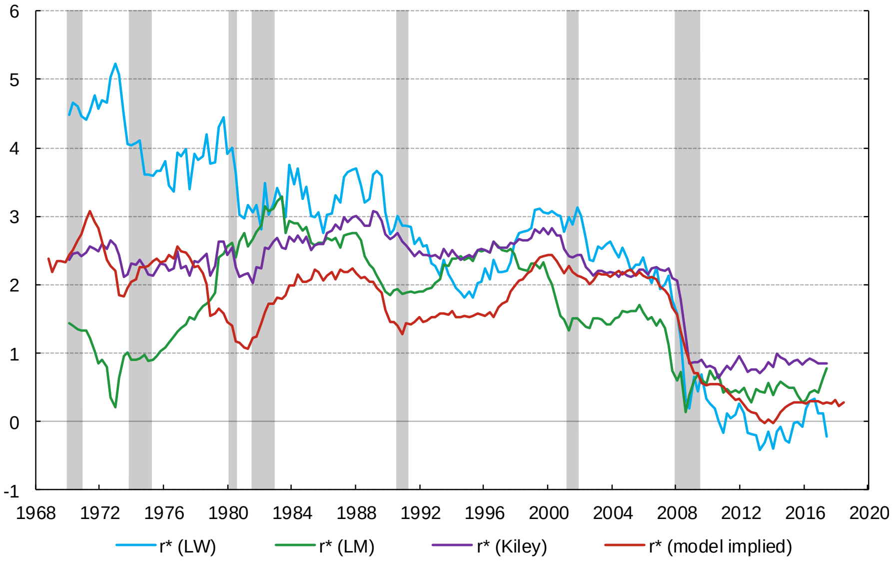
As shown in Table 3, the parameters for the inflation process are exactly the same as in model II; hence, the posterior beliefs for the short- and long-term inflation expectations in this model are also exactly the same as shown in Figures 5 and 6. Both the learning parameters and the posterior beliefs for the short- and long-term output growth are also the same as in model II. However, the time preference is slightly different from the previous value; therefore, as shown in Figure 11, the model-implied r_t^* is almost the same as before.[19]

As in models I and II, Figure 12 shows the importance of the posteriors for the long-

[19] To be consistent with the concept of r^*, neither the short-run effect from $\tilde{x}_{c,t}$ on the real yield, nor the effect of ambiguity on the yields, is included for calculation of the model-implied r^*.

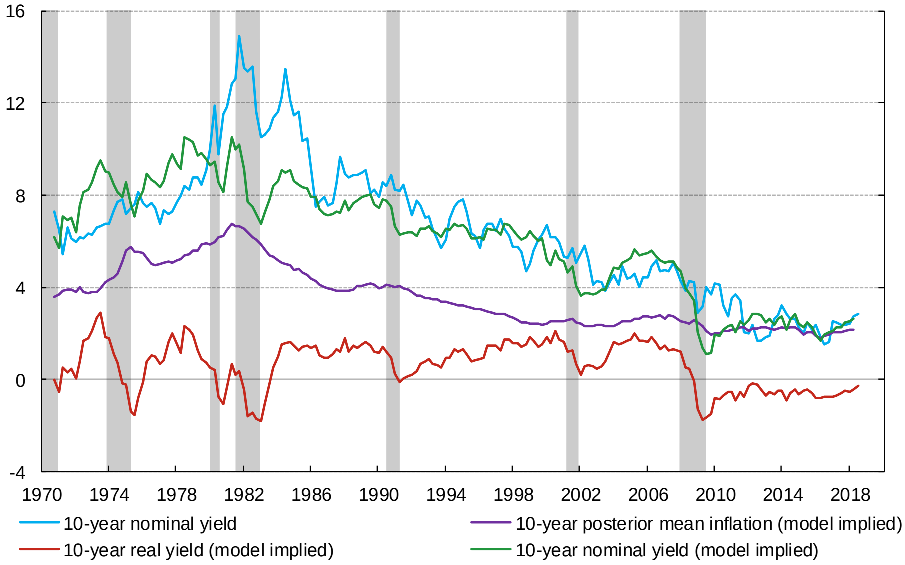
run mean growth and inflation in explaining the movements in the long-term nominal yield, and Figure 13 shows the importance of the posteriors for the short-run deviation from the long-run mean in explaining the movements in the 1-year nominal yield. The limitation of model II is that the model-implied spread is mean zero because of the stationarity assumption for x_t. The short-rate expectations (both nominal and real) are upward-sloping under the agent's equilibrium worst-case belief in this model and, hence, are consistent with the data. The model-implied spread is positive, on average. Furthermore, Figure 14 shows that the dynamics of the model-implied spread match the data well.

Figure 11: **Individual r* and model implied r***



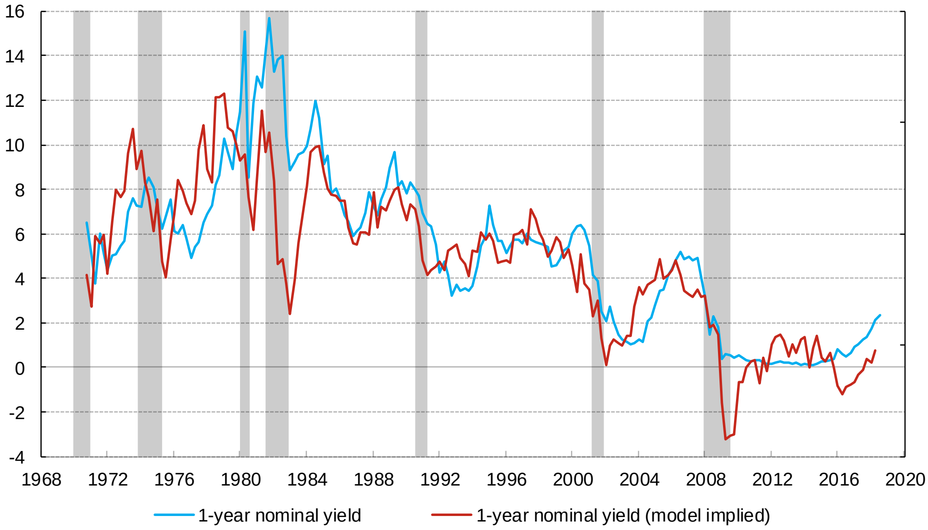
The individual r^*s (quarterly data) are obtained from Bauer and Rudebusch (2019) and cover the period from 1971:Q4 to 2017:Q2. The three macroeconomic estimates of r^*s are obtained from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015). The model-implied r^* (quarterly) represents the period from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.

Figure 12: **10-year nominal yield and macro trends**

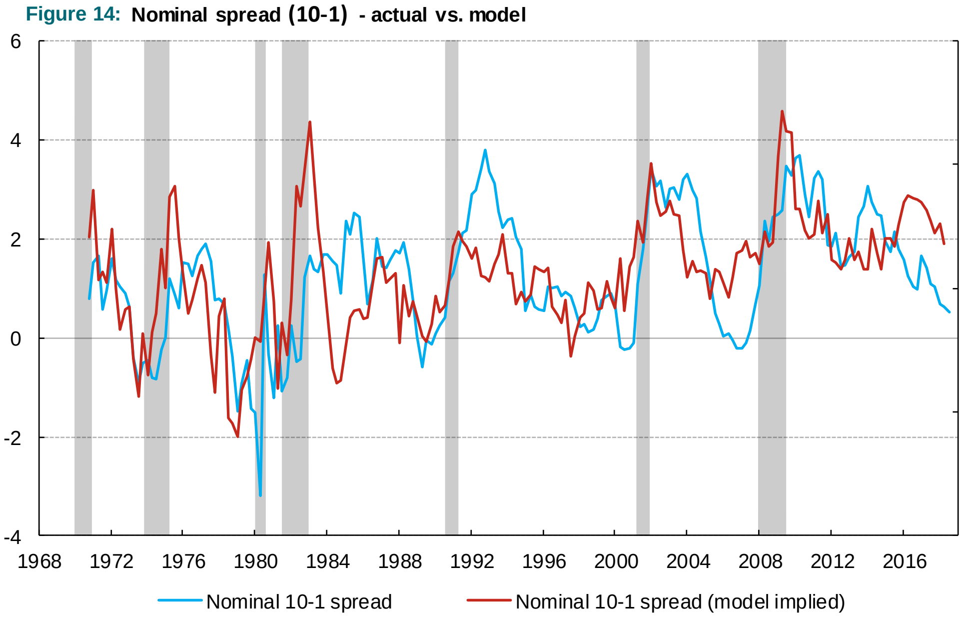


The end-of-quarter 10-year nominal yields are obtained from Gürkaynak et al. (2007) for the period from 1968:Q3 to 2018:Q2. The model-implied 10-year real yield, the 10-year nominal yield, and total posterior belief for the mean inflation (quarterly data) represent the period from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession as defined by the NBER.

Figure 13: **1-year nominal yield**



The end-of-quarter 1-year nominal yields are obtained from Gürkaynak et al. (2007) for the period from 1968:Q3 to 2018:Q2. The model-implied 1-year nominal yield represents the period from 1968:Q3 to 2008:Q4. The gray bars represent periods of recession as defined by the NBER.



The end-of-quarter 10-year minus 1-year nominal yield spreads are obtained from Gürkaynak et al. (2007) for the period from 1968:Q3 to 2018:Q2. The model-implied 10-year minus 1-year nominal yield spread represents the period from 1968:Q3 to 2008:Q4. The gray bars represent periods of recession as defined by the NBER.

Finally, by comparing the model-implied (both models II and III) 1-year nominal yields with the data in Figures 13 and 9, we observe a recurring pattern that the model-implied short-term yields are higher than the data from the trough to the expansion, and they are lower than those in the data for the late expansion periods. Given that the short-term yields are controlled by the Federal Reserve, this suggests that the Federal Reserve kept the short rates low for a longer period than suggested by the model (behind the curve) and there was a certain degree of overshoot during the late expansion periods (before recessions). This is consistent with Taylor (2018), who argued that the over-accommodative policy between 2003 and 2005 was a source of the housing bubble.

4.4.3. Expectations hypothesis and predictability of bond returns

Investors' subjective nominal and real short-rate expectations, reflecting their worst-case beliefs in growth and inflation, are upward sloping. And given the CRRA utility (the subjective bond premium $er_{n,t+1}$ is close to zero), the model-implied yields for long-term bonds are roughly equal to the average of the expected (worst-case) future short rates. Thus, the EH roughly holds under the subjective equilibrium belief. However, because long-run growth and inflation evolve over time under a true distribution that is different from their worst-case beliefs, investors' ambiguity about long-run inflation or GDP growth does not materialize when the time arrives. At each time t, the realized one-step-ahead ambiguity ($a_{1c,t}$ or $a_{1\pi,t}$) contains only the random walk with no trend (the trend has not materialized). Hence, the realized one-step-ahead ambiguity does not become larger or smaller as investors had perceived in the past, and the realized short rates (nominal and real) are lower than expected under their worst-case beliefs. These differences and the current yield spreads/forward rates are both driven by a trend component in the ambiguity process. Hence, consistent with the empirical evidence, the realized excess bond returns are predictable. To an observer outside the model, the difference between the worst-case expectation and the realized short rate (governed by the benchmark measure) looks like expectational error but is rational for agents inside the model who face ambiguity, thus, providing a rational interpretation for the expectational errors in Froot (1989), Piazzesi et al. (2015), and Cieslak (2018). A detailed discussion and formal tests of the EH (Campbell and Shiller, 1991; Cochrane and Piazzesi, 2005) are provided in Zhao (2020).

5. Conclusion

This paper bridges the gap between empirical and equilibrium yield curve studies by providing an equilibrium interpretation for the trends, cycles, and spreads in historical Treasury bond yields. The representative agent learns the long-run mean and the short-run deviation from the mean separately from the different components of the GDP growth and inflation rates. Instead of using all of the available data in forming the posterior beliefs, the past data gradually lose relevance for learning, either because of fading memory or because it is perceived as irrelevant. The slow-moving trend component in the yields is driven by the posteriors for the long-run mean inflation and growth rates, which also move closely with the r_t^* and π_t^* estimations in the literature. The cyclical movements in the short-term yields and in the spreads between the long- and short-term yields are mostly driven by belief-updating regarding the short-run deviation from the mean. The secular stagnation and the upward trend in the Treasury yield spread are tightly coupled because both are driven by persistently negative short-run deviations for both inflation and growth. At each point in time, the amount of Knightian uncertainty the ambiguity-averse agent faces is different for the long run versus the short run, which gives rise to upward-sloping short-rate expectations under the agent's worst-case beliefs and, hence, upward-sloping nominal and real yield curves.

Empirical yield curve modeling has been widely used by central banks and practitioners. However, equilibrium yield curve models are rarely used because of their inaccuracy. Given the historical performance of the model and its closed-form solutions, the model can be used for real-time interest-rate forecasting, using survey forecasts or central bank projections for GDP growth and inflation as inputs. To increase its performance, one important future research avenue would be to extend the model and incorporate the inflation risk premium in Piazzesi and Schneider (2007) as well as the potentially stochastic volatility.

References

Albuquerque, R., Eichenbaum, M., Luo, V., Rebelo, S., 2016. Valuation Risk and Asset Pricing. The Journal of Finance 71, 2861–2904.

Alvarez, F., Jermann, U. J., 2005. Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth. Econometrica 73, 1977–2016.

Anderson, E. W., Hansen, L. P., Sargent, T. J., 2003. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. Journal of the European Economic Association 1, 68–123.

Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics 50, 745–787.

Bansal, R., Shaliastovich, I., 2013. A long-run risks explanation of predictability puzzles in bond and currency markets. Review of Financial Studies 26, 1–33.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. The Journal of Finance 59, 1481–1509.

Bauer, M. D., Rudebusch, G. D., 2019. Interest Rates Under Falling Stars. the American Economic Review Forthcoming.

Berrada, T., Detemple, J., Rindisbacher, M., 2018. Asset pricing with beliefs-dependent risk aversion and learning. Journal of Financial Economics 128, 504–534.

Bianchi, F., Ilut, C. L., Schneider, M., 2017. Uncertainty Shocks, Asset Supply and Pricing over the Business Cycle. The Review of Economic Studies 85, 810–854.

Burkhardt, D., Hasseltoft, H., 2012. Understanding asset correlations. University of Zurich and Swiss Finance Institute Manuscript.

Caballero, R. J., Farhi, E., Gourinchas, P.-O., 2008. An Equilibrium Model of "Global Imbalances" and Low Interest Rates. American Economic Review 98, 358–93.

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

Campbell, J. Y., Shiller, R. J., 1991. Yield spreads and interest rate movements: A bird's eye view. The Review of Economic Studies 58, 495–514.

Campbell, J. Y., Sunderam, A., Viceira, L. M., 2017. Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds. Critical Finance Review 6, 263–301.

Carvalho, C., Ferrero, A., Nechio, F., 2016. Demographics and real interest rates: Inspecting the mechanism. European Economic Review 88, 208–226.

Christensen, J. H. E., Rudebusch, G. D., 2019. A New Normal for Interest Rates? Evidence from Inflation-Indexed Debt. The Review of Economics and Statistics 101, 933–949.

Cieslak, A., 2018. Short-Rate Expectations and Unexpected Returns in Treasury Bonds. Review of Financial Studies 31, 3265–3306.

Cieslak, A., Povala, P., 2015. Expected Returns in Treasury Bonds. The Review of Financial Studies 28, 2859–2901.

Cochrane, J. H., Piazzesi, M., 2005. Bond risk premia. The American economic review 95, 138–160.

Cogley, T., Sargent, T. J., 2008. Anticipated Utility and Rational Expectations as Approximations of Bayesian Decision Making. International Economic Review 49, 185–221.

Collin-Dufresne, P., Johannes, M., Lochstoer, L. A., 2016. Parameter Learning in General Equilibrium: The Asset Pricing Implications. American Economic Review 106, 664–98.

David, A., Veronesi, P., 2013. What Ties Return Volatilities to Fundamentals and Price Valuations? Journal of Political Economy 121, 682–746.

Del Negro, M., Giannone, D., Giannoni, M. P., Tambalotti, A., 2017. Safety, liquidity, and the natural rate of interest. Brookings Papers on Economic Activity Spring.

Diebold, F. X., Rudebusch, G. D., Aruoba, S. B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. Journal of Econometrics 131, 309–338.

Edge, R. M., Laubach, T., Williams, J. C., 2007. Learning and shifts in long-run productivity growth. Journal of Monetary Economics 54, 2421–2438.

Epstein, L. G., Schneider, M., 2003. Recursive multiple-priors. Journal of Economic Theory 113, 1–31.

Epstein, L. G., Schneider, M., 2007. Learning under ambiguity. The Review of Economic Studies 74, 1275–1303.

Epstein, L. G., Schneider, M., 2010. Ambiguity and asset markets. Annual Review of Financial Economics 2, 315–346.

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

Fama, E. F., Bliss, R. R., 1987. The information in long-maturity forward rates. The American Economic Review 77, 680–692.

Farhi, E., Gourio, F., 2019. Accounting for Macro-Finance Trends: Market Power, Intangibles, and Risk Premia. Brookings Papers on Economic Activity Fall, 147–250.

Faust, J., Wright, J. H., 2009. Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset. Journal of Business & Economic Statistics 27, 468–479.

Froot, K., 1989. New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates. Journal of Finance 44, 283–305.

Gagliardini, P., Porchia, P., Trojani, F., 2009. Ambiguity Aversion and the Term Structure of Interest Rates. The Review of Financial Studies 22, 4157–4188.

Gilchrist, S., Saito, M., 2008. Expectations, asset prices, and monetary policy: The role of learning. University of Chicago Press.

Gürkaynak, R. S., Sack, B., Wright, J. H., 2007. The US Treasury yield curve: 1961 to the present. Journal of Monetary Economics 54, 2291–2304.

Hamilton, J. D., Harris, E. S., Hatzius, J., West, K. D., 2016. The Equilibrium Real Funds Rate: Past, Present, and Future. IMF Economic Review 64, 660–707.

Hansen, L. P., 2007. Beliefs, Doubts and Learning: The Valuation of Macroeconomic Risk. American Economic Review, Papers and Proceedings 97, 1–30.

Hansen, L. P., 2012. Dynamic Valuation Decomposition Within Stochastic Economies. Econometrica 80, 911–967.

Hansen, L. P., Jagannathan, R., 1991. Implications of Security Market Data for Models of Dynamic Economies. Journal of Political Economy 99, 225–262.

Hansen, L. P., Sargent, T. J., 2001. Robust control and model uncertainty. The American Economic Review 91, 60–66.

Hansen, L. P., Sargent, T. J., 2010. Fragile Beliefs and the Price of Model Uncertainty. Quantitative Economics 1, 129–162.

Hansen, L. P., Scheinkman, J. A., 2009. Long-Term Risk: An Operator Approach. Econometrica 77, 177–234.

Holston, K., Laubach, T., Williams, J. C., 2017. Measuring the natural rate of interest: International trends and determinants. Journal of International Economics 108, S59–S75.

Ilut, C., 2012. Ambiguity aversion: Implications for the uncovered interest rate parity puzzle. American Economic Journal: Macroeconomics 4, 33–65.

Ilut, C., Kehrig, M., Schneider, M., 2018. Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to News. Journal of Political Economy 126.

Ilut, C. L., Schneider, M., 2014. Ambiguous business cycles. The American Economic Review 104, 2368–2399.

Johannsen, B. K., Mertens, E., 2018. A time series model of interest rates with the effective lower bound. BIS Working Papers Manuscript.

Kiley, M. T., 2015. What Can the Data Tell Us About the Equilibrium Real Interest Rate? FEDS Working Paper Manuscript.

Kozicki, S., Tinsley, P., 2001. Shifting endpoints in the term structure of interest rates. Journal of Monetary Economics 47, 613–652.

Kreps, D. M., 1998. Anticipated Utility and Dynamic Choice. Frontiers of Research in Economic Theory p. 242074.

Laubach, T., Williams, J. C., 2003. Measuring the Natural Rate of Interest. The Review of Economics and Statistics 85, 1063–1070.

Laubach, T., Williams, J. C., 2016. Measuring the Natural Rate of Interest Redux. Business Economics 51, 57–67.

Lettau, M., Wachter, J. A., 2011. The term structures of equity and interest rates. Journal of Financial Economics 101, 90–113.

Lubik, T. A., Matthes, C., 2015. Calculating the Natural Rate of Interest: A Comparison of Two Alternative Approaches. Richmond Fed Economic Brief pp. 1–6.

Ludvigson, S. C., Ng, S., 2009. Macro Factors in Bond Risk Premia. The Review of Financial Studies 22, 5027–5067.

Lunsford, K. G., West, K. D., 2018. Some Evidence on Secular Drivers of U.S. Safe Real Rates. NBER Working Paper Manuscript.

Malmendier, U., Nagel, S., 2011. Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? The Quarterly Journal of Economics 126, 373–416.

Malmendier, U., Nagel, S., 2016. Learning from Inflation Experiences. The Quarterly Journal of Economics 131, 53–87.

Mangel, M., 1990. Dynamic information in uncertain and changing worlds. Journal of Theoretical Biology 146.

Nagel, S., Xu, Z., 2019. Asset pricing with fading memory. NBER Working Paper Manuscript.

Nelson, C. R., Plosser, C. R., 1982. Trends and random walks in macroeconmic time series: Some evidence and implications. Journal of Monetary Economics 10, 139–162.

Piazzesi, M., Salomao, J., Schneider, M., 2015. Trend and cycle in bond premia. Working paper.

Piazzesi, M., Schneider, M., 2007. Equilibrium yield curves. In: *NBER Macroeconomics Annual 2006, Volume 21*, MIT Press, pp. 389–472.

Rachel, L., Smith, T. D., 2015. Secular drivers of the global real interest rate. Bank of England Working Paper Manuscript.

Rose, A. K., 1988. Is the Real Interest Rate Stable? The Journal of Finance 43, 1095–1112.

Rudebusch, G. D., Swanson, E. T., 2012. The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks. American Economic Journal: Macroeconomics 4, 105–43.

Rudebusch, G. D., Wu, T., 2008. A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy. The Economic Journal 118, 906–926.

Singleton, K. J., 1983. Real and Nominal Factors in the Cyclical Behavior of Interest Rates, Output, and Money. Journal of Economic Dynamics and Control 5.

Summers, L. H., 2014. US economic prospects: Secular stagnation, hysteresis, and the zero lower bound. Business Economics 49, 65–73.

Taylor, J., 1993. Discretion versus policy rules in practice. Carnegie-Rochester Conference Series on Public Policy 39.

Taylor, J. B., 2018. Government as a Cause of the 2008 Financial Crisis: A Reassessment After 10 Years. Remarks prepared for the "Causes" Session "Workshop Series on the 2008 Financial Crisis: Causes, The Panic, The Recession, Lessons" Manuscript.

Ulrich, M., 2013. Inflation ambiguity and the term structure of US Government bonds. Journal of Monetary Economics 60, 295–309.

Wachter, J. A., 2006. A consumption-based model of the term structure of interest rates. Journal of Financial Economics 79, 365–399.

Wright, J. H., 2011. Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset. American Economic Review 101, 1514–34.

Zhao, G., 2017. Confidence, Bond Risks, and Equity Returns. Journal of Financial Economics 126, 668–688.

Zhao, G., 2020. Ambiguity, Nominal Bond Yields, and Real Bond Yields. American Economic Review: Insights Forthcoming.

Appendix - Model Solution

Since model III is a comprehensive model and the solutions to models I and II are embedded in the model-III solution, we will only provide the model III solution in this appendix.

A. Forcing process

Under the worst-case measure, the economic dynamics follow

$$\begin{aligned}\Delta g_{t+1} &= \phi_c^a a_{c,t} + \Delta g_{t+1}^* + Gap_{t+1}^g \\ \pi_{t+1} &= \phi_\pi^a a_{\pi,t} + \pi_{t+1}^* + Gap_{t+1}^\pi,\end{aligned}$$

where Δg_{t+1} and π_{t+1} are the total real GDP growth and inflation, respectively. Δg_{t+1}^* and π_{t+1}^* are the real consumption growth (scaled by total real GDP $\frac{C_{t+1}-C_t}{GDP_t}$) and core inflation (scaled by total price level $\frac{P_{t+1}^{core}-P_t^{core}}{P_t}$), respectively. Gap_{t+1}^g and Gap_{t+1}^π are the total GDP growth rate excluding Δg_{t+1}^* and the total inflation rate excluding π_{t+1}^*, respectively. ϕ_c^a and ϕ_c^a represent the equilibrium choice of the upper or lower bound, respectively, equal to -1 or $+1$.

Both real consumption growth and core inflation follow the i.i.d. laws of motion

$$\begin{aligned}\Delta g_{t+1}^* &= \mu_c^* + \sigma_c \varepsilon_{c,t+1}^* \\ \pi_{t+1}^* &= \mu_\pi^* + \sigma_\pi \varepsilon_{\pi,t+1}^*,\end{aligned}$$

where $\varepsilon^*_{c,t+1}$ and $\varepsilon^*_{\pi,t+1}$ are i.i.d. normal shocks. The agent knows that both Δg^*_{t+1} and π^*_{t+1} are i.i.d., and they also know σ_c and σ_π but not the long-run mean μ^*_c and μ^*_π. The agent forms expectations about μ^*_c and μ^*_π based on the constant-gain learning scheme, with the posteriors

$$\mu^*_c|H^{g^*}_t \sim N\left(\tilde{\mu}^*_{c,t}, \upsilon^*_c\sigma^2_c\right)$$
$$\mu^*_\pi|H^{\pi^*}_t \sim N\left(\tilde{\mu}^*_{\pi,t}, \upsilon^*_\pi\sigma^2_\pi\right),$$

where

$$\tilde{\mu}^*_{c,t} = \tilde{\mu}^*_{c,t-1} + \upsilon^*_c\left(\Delta g^*_t - \tilde{\mu}^*_{c,t-1}\right)$$
$$\tilde{\mu}^*_{\pi,t} = \tilde{\mu}^*_{\pi,t-1} + \upsilon^*_\pi\left(\pi^*_t - \tilde{\mu}^*_{\pi,t-1}\right)$$

and the predictive distribution

$$\Delta g^*_{t+j}|H^{g^*}_t \sim N\left(\tilde{\mu}^*_{c,t}, (1+\upsilon^*_c)\,\sigma^2_c\right)$$
$$\pi^*_{t+j}|H^{\pi^*}_t \sim N\left(\tilde{\mu}^*_{\pi,t}, (1+\upsilon^*_\pi)\,\sigma^2_\pi\right),$$

where $j = 1, 2, ...,$ $H^{g^*}_t \equiv \{\Delta g^*_0, \Delta g^*_1, ...\Delta g^*_t\}$, and $H^{\pi^*}_t \equiv \{\pi^*_0, \pi^*_1, ...\pi^*_t\}$.

Both Gap^g_{t+1} and Gap^π_{t+1} are assumed to contain latent stationary components

$$Gap^g_{t+1} = x_{c,t+1} + \sigma^{gap}_c\varepsilon^{gap}_{c,t+1}$$
$$Gap^\pi_{t+1} = x_{\pi,t+1} + \sigma^{gap}_\pi\varepsilon^{gap}_{\pi,t+1}$$

and

$$x_{c,t+1} = \rho_c x_{c,t} + \sigma^x_c\varepsilon^x_{c,t+1}$$
$$x_{\pi,t+1} = \rho_\pi x_{\pi,t} + \sigma^x_\pi\varepsilon^x_{\pi,t+1},$$

where $\varepsilon^{gap}_{c,t+1}$, $\varepsilon^{gap}_{\pi,t+1}$, $\varepsilon^x_{c,t+1}$, and $\varepsilon^x_{\pi,t+1}$ are i.i.d. normal shocks. The representative agent knows all of the the parameters but not $x_{c,t+1}$ and $x_{\pi,t+1}$. They form expectations about $x_{c,t+1}$ and $x_{\pi,t+1}$, based on the same learning scheme as for the long-run mean but with potentially different geometric weighting parameters, υ_c^{gap} and υ_π^{gap}. The posteriors are given by

$$
\begin{aligned}
x_{c,t+1}|H_{g,t}^{gap} &\sim N\left(\rho_c \tilde{x}_{c,t}, \upsilon_c^{gap}\left((\sigma_c^x)^2 + (\sigma_c^{gap})^2\right)\right) \\
x_{\pi,t+1}|H_{\pi,t}^{gap} &\sim N\left(\rho_\pi \tilde{x}_{\pi,t}, \upsilon_\pi^{gap}\left((\sigma_\pi^x)^2 + (\sigma_\pi^{gap})^2\right)\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{x}_{c,t} &= \rho_c \tilde{x}_{c,t-1} + \upsilon_c^{gap}\left(Gap_t^g - \rho_c \tilde{x}_{c,t-1}\right) \\
\tilde{x}_{\pi,t} &= \rho_\pi \tilde{x}_{\pi,t-1} + \upsilon_\pi^{gap}\left(Gap_t^\pi - \rho_\pi \tilde{x}_{\pi,t-1}\right)
\end{aligned}
$$

where $H_{g,t}^{gap} \equiv \{Gap_0^g, Gap_1^g, ...Gap_t^g\}$ and $H_{\pi,t}^{gap} \equiv \{Gap_0^\pi, Gap_1^\pi, ...Gap_t^\pi\}$, and the total predictive distribution is given by

$$
\begin{aligned}
\Delta g_{t+j}|H_t^g &\sim N\left(\tilde{\mu}_{c,t}^* + \rho_c \tilde{x}_{c,t}, (1+\upsilon_c^*)\,\sigma_c^2 + (1+\upsilon_c^{gap})\left((\sigma_c^x)^2 + (\sigma_c^{gap})^2\right)\right) \\
\pi_{t+j}|H_t^\pi &\sim N\left(\tilde{\mu}_{\pi,t}^* + \rho_\pi \tilde{x}_{\pi,t}, (1+\upsilon_\pi^*)\,\sigma_\pi^2 + (1+\upsilon_\pi^{gap})\left((\sigma_\pi^x)^2 + (\sigma_\pi^{gap})^2\right)\right)
\end{aligned}
$$

where $j = 1, 2, ...$ and the variance of the predictive distribution contains both the uncertainty due to future shocks and the uncertainty about μ_c^*/μ_π^* and $x_{c,t+1}/x_{\pi,t+1}$. H_t^g contains both $H_t^{g^*}$ and $H_{g,t}^{gap}$, and H_t^π contains both $H_t^{\pi^*}$ and $H_{\pi,t}^{gap}$.

The size of ambiguity $a_{c,t}$ and $a_{\pi,t}$ are modeled as a random walk with drift as follows:

$$
\begin{aligned}
a_{c,t+1} &= \mu_c^a + a_{c,t} + \sigma_{ac}\varepsilon_{ac,t+1} + \sigma_a^{ac}\varepsilon_{a,t+1} \\
a_{\pi,t+1} &= \mu_\pi^a + a_{\pi,t} + \sigma_a^{a\pi}\varepsilon_{a,t+1},
\end{aligned}
$$

where μ_c^a and μ_π^a are the drift parameters, which can be positive or negative. $a_{c,t}$ and $a_{\pi,t}$ are driven by a common exogenous shock $\varepsilon_{a,t+1}$, where the coefficients σ_a^{ac} and $\sigma_a^{a\pi}$ capture the correlation between them. $\varepsilon_{ac,t+1}$ is an $a_{c,t}$ specific shock that captures the difference between these two.

To solve the model, we first rewrite the dynamics of the whole economy in vector forms as follows:

$$
\begin{aligned}
z_{t+1} &= \phi_a a_t + \tilde{\mu}_t^* + \rho_x x_t + \sigma^z \tilde{\varepsilon}_{z,t+1} \\
x_{t+1} &= \rho_x x_t + \upsilon^{gap} \left(Gap_{t+1} - \rho_x x_t\right) \\
\tilde{\mu}_{t+1}^* &= \tilde{\mu}_t^* + \upsilon^* \left(\Delta z_{t+1}^* - \tilde{\mu}_t^*\right) \\
a_{t+1} &= \mu_a + a_t + \sigma^a \varepsilon_{t+1}^a,
\end{aligned}
$$

where $z_{t+1} = (\Delta g_{t+1}, \pi_{t+1})^T$, $z_{t+1}^* = (\Delta g_{t+1}^*, \pi_{t+1}^*)^T$, $Gap_{t+1} = (Gap_{t+1}^g, Gap_{t+1}^\pi)^T$, $x_{t+1} = (\tilde{x}_{c,t+1},\, \tilde{x}_{\pi,t+1})^T$, $a_{t+1} = (a_{c,t+1},\, a_{\pi,t+1})^T$, $\tilde{\mu}_t^* = (\tilde{\mu}_{c,t}^*,\, \tilde{\mu}_{\pi,t}^*)^T$, $\mu_a = (\mu_c^a,\, \mu_\pi^a)^T$, $\upsilon^{gap} = \begin{pmatrix} \upsilon_c^{gap} & 0 \\ 0 & \upsilon_\pi^{gap} \end{pmatrix}$, $\upsilon^* = \begin{pmatrix} \upsilon_c^* & 0 \\ 0 & \upsilon_\pi^* \end{pmatrix}$, $\rho_x = \begin{pmatrix} \rho_c & 0 \\ 0 & \rho_\pi \end{pmatrix}$, $\phi_a = \begin{pmatrix} \phi_c^a & 0 \\ 0 & \phi_\pi^a \end{pmatrix}$, $\sigma^z = \begin{pmatrix} \sigma_c & 0 \\ 0 & \sigma_\pi \end{pmatrix}$, $\sigma^a = \begin{pmatrix} \sigma_{ac} & \sigma_a^{ac} \\ 0 & \sigma_a^{a\pi} \end{pmatrix}$, $\tilde{\varepsilon}_{z,t+1} = (\tilde{\varepsilon}_{c,t+1}, \tilde{\varepsilon}_{\pi,t+1})^T$, and $\varepsilon_{t+1}^a = (\varepsilon_{ac,t+1}, \varepsilon_{a,t+1})^T$. The shocks $\tilde{\varepsilon}_{c,t+1}$, $\tilde{\varepsilon}_{\pi,t+1}$, $\varepsilon_{d,t+1}$, $\varepsilon_{ac,t+1}$, and $\varepsilon_{a,t+1} \sim i.i.d.\ N(0,1)$.

B. Stochastic discount factor

Given the CRRA utility, the nominal stochastic discount factor can be written as follows:

$$
m_{t,t+1}^{\$} = log\beta - \gamma \Delta g_{t+1} - \pi_{c,t+1} = log\beta - v' z_{t+1},
$$

where $v' = (\gamma,\, 1)$. For the real stochastic discount factor, we can replace v' with $v' = (\gamma,\, 0)$.

C. Bond yields

The time-t price of a zero-coupon bond that pays one unit of consumption n periods from now is denoted as $P_t^{(n)}$ and it satisfies the recursion,

$$P_t^{(n)} = E_{p_t^o}[M_{t,t+1}^{\$} P_{t+1}^{(n-1)}]$$

with the initial condition that $P_t^{(0)} = 1$ and $E_{p_t^o}$ is the expectation operator for the worst-case measure. Given the linear Gaussian framework, we assume that $p_t^{(n)} = log(P_t^{(n)})$ is a linear function of $\tilde{\mu}_t^*$, x_t, and a_t as follows:

$$p_t^{(n)} = -A^{(n)} - B^{(n)} x_t - C^{(n)} \tilde{\mu}_t^* - D^{(n)} a_t.$$

When we substitute $p_t^{(n)}$ and $p_{t+1}^{(n-1)}$ in the Euler equation, the coefficients in the pricing equation can be solved with $B^{(n)} = B^{(n-1)}\rho + v'\rho$, $C^{(n)} = C^{(n-1)} + v' = v'n$, $D^{(n)} = D^{(n-1)} + v'\phi_a = v'n\phi_a$, and $A^{(n)} = A^{(n-1)} + A^{(1)} - 0.5 * Var_t\left(p_{t+1}^{(n-1)}\right) - Cov_t\left(p_{t+1}^{(n-1)}, m_{t,t+1}^{\$}\right) + D^{(n-1)}\mu_a$, where

$$\begin{aligned}
Var_t\left(p_{t+1}^{(n-1)}\right) &= \left(B^{(n-1)}\right) Var_t\left(x_{t+1}\right) \left(B^{(n-1)}\right)' \\
&+ \left(C^{(n-1)}\right) Var_t\left(\tilde{\mu}_{t+1}^*\right) \left(C^{(n-1)}\right)' \\
&+ \left(D^{(n-1)}\right) Var_t\left(a_{t+1}\right) \left(D^{(n-1)}\right)' \\
&+ 2\left(B^{(n-1)}\right) Cov_t\left(x_{t+1}, \tilde{\mu}_{t+1}^*\right) \left(C^{(n-1)}\right)' \\
&+ 2\left(B^{(n-1)}\right) Cov_t\left(x_{t+1}, a_{t+1}\right) \left(D^{(n-1)}\right)' \\
&+ 2\left(C^{(n-1)}\right) Cov_t\left(\tilde{\mu}_{t+1}^*, a_{t+1}\right) \left(D^{(n-1)}\right)',
\end{aligned}$$

$$\begin{aligned}
Cov_t\left(p_{t+1}^{(n-1)}, m_{t,t+1}^{\$}\right) &= v' Cov_t\left(x_{t+1}, z_{t+1}\right) \left(B^{(n-1)}\right)' \\
&+ v' Cov_t\left(\tilde{\mu}_{t+1}^*, z_{t+1}\right) \left(C^{(n-1)}\right)' \\
&+ v' Cov_t\left(a_{t+1}, z_{t+1}\right) \left(D^{(n-1)}\right)',
\end{aligned}$$

and

$$A^{(1)} \quad = \quad -log\beta - 0.5 * v'Var_t\left(z_{t+1}\right)v.$$

Given the small impacts of the covariance terms, we assume all of the covariances in the $Var_t\left(p_{t+1}^{(n-1)}\right)$ and $Cov_t\left(a_{t+1}, z_{t+1}\right)$ to be zero for simplicity.

The nominal bond yields can be calculated as $y_t^{(n)} = -\frac{1}{n}p_t^{(n)} = \frac{A^{(n)}}{n} + \frac{B^{(n)}}{n}x_t + \frac{C^{(n)}}{n}\tilde{\mu}_t^* + \frac{D^{(n)}}{n}a_t$. The log holding period return from buying an n periods bond at time t and selling it as an $n-1$ periods bond at time $t-1$ is defined as $r_{n,t+1} = p_{t+1}^{(n-1)} - p_t^{(n)}$, and the subjective excess return is $er_{n,t+1} = -Cov_t\left(r_{n,t+1}, m_{t,t+1}^{\$}\right) = -Cov_t\left(p_{t+1}^{(n-1)}, m_{t,t+1}^{\$}\right)$. To solve the price and yields for the real bonds, we can simply replace v' with $v' = (\gamma, 0)$.

D. Kalman filter alternative

In this section, we show that, in terms of the subjective belief dynamics and bond prices, the adaptive learning scheme in our model is equivalent to a full-memory optimal-learning model. Taking output growth as an example, in this full-memory model the agent perceives a latent AR(1) trend growth rate and uses the Kalman filter to optimally track this latent trend, while objectively the trend growth rate is constant. The information structure is a filtration and Markovian. In the agent's subjective view, the past data gradually loses relevance for forecasting, not because of fading memory but because it is perceived as irrelevant given the perceived stochastic drift over time in the trend growth rate.

Suppose the agent's perceived law of motion for output growth is as follows:

$$\Delta g_{t+1} = \mu_{c,t} + \epsilon_{t+1}$$
$$\mu_{c,t+1} = \rho_\mu \mu_{c,t} + \varsigma_{t+1},$$

where $\epsilon_{t+1} \sim N\left(0, \sigma_\epsilon^2\right)$ and $\varsigma_{t+1} \sim N\left(0, \sigma_\varsigma^2\right)$. The agent knows σ_ϵ, σ_ς, and ρ_μ but perceives $\mu_{c,t}$ as a latent AR(1) process. Given a diffuse prior and an infinite history, H_t^g, of the observed data on Δg, the steady-state optimal forecast $\tilde{\mu}_{c,t+1|t}$ is updated as follows (see, e.g., Edge et al. (2007) and Gilchrist and Saito (2008))

$$\tilde{\mu}_{c,t+1|t} = \rho_\mu \tilde{\mu}_{c,t|t-1} + K(\Delta g_t - \rho_\mu \tilde{\mu}_{c,t|t-1})$$

with

$$\begin{aligned}
K &= \frac{\Sigma}{\Sigma + \sigma_\epsilon^2} \\
\Sigma &= \frac{\sigma_\epsilon^2}{2}\left(-(1-\rho_\mu^2-\phi) + \sqrt{(1-\rho_\mu^2-\phi)^2 + 4\phi}\right) \\
\phi &= \frac{\sigma_\varsigma^2}{\sigma_\epsilon^2}.
\end{aligned}$$

The steady-state Kalman filtering is equivalent to adaptive learning with appropriately chosen parameter values. Hence, bond pricing in this perceived stochastic trend setting is the same as in the fading-memory setting.