Learning, Equilibrium Trend, Cycle, and Spread in Bond Yields[☆]

Guihai Zhao
Bank of Canada

This version: March 2019

Abstract

While the empirical literature has shown the importance of macro trends in modeling the term structure of interest rates, the standard stationarity assumption makes it hard for equilibrium models to capture the trend in bond yields. This paper presents an equilibrium model to explain the trend, cycle, and spread in historical U.S. Treasury bond yields. The trend in yields is generated by learning from the stable components in GDP growth and inflation, which share similar patterns to the neutral rate of interest (r_t^*) and trend inflation (π_t^*) estimates in the literature. Cyclical movements in yields and spread, are mainly driven by learning from the transitory components in GDP growth and inflation. The upward trend in the Treasury yield spread found in the data and the recent secular stagnation are tightly coupled due to persistently negative short-run beliefs. The upward-sloping yield curve is mainly driven by the fact that the amount of Knightian uncertainty that investors face is different for the long run versus the short run.

Keywords: adaptive learning, trend inflation, equilibrium real interest rate, bond yields, term spread, ambiguity
JEL Classification: G00, G12, E43

1. Introduction

It is important for both policy makers and academics to understand the historical dynamics of the Treasury yield curve, because of its key role in the transmission of monetary policy and its tight link with the stochastic discount factor. While the empirical

[☆]Guihai Zhao (gzhao@bankofcanada.ca) is with the Bank of Canada, 234 Wellington Street, Ottawa, ON K1A 0G9. I would like to thank Jason Allen, Jonathan Witmer as well as seminar participants at the Bank of Canada for their suggestions. Any remaining errors are my own. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Canada.

literature has shown the importance of accounting for macro trends in models of the term structure of interest rates,[1] bond yields in equilibrium models are generally modeled as stationary, mean-reverting processes. Hence, it is hard to explain the low-frequency variation in interest rates in such models, and as a result, cyclical movements in short-term yields and in the spread between long- and short-term yields are hard to justify.[2] In this paper, we provide a joint equilibrium for the trend, cycle, and term spread in historical Treasury bond yields.

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 forces 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 (2017) show that it is crucial to include the neutral rate of interest (r_t^*) as well, which has driven the downward trend in long term yields over the last 20 years since inflation expectations have stabilized. Both π_t^* and r_t^* (and hence short rates) are modeled as random walk processes in Bauer and Rudebusch (2017), and their resulting term premium component – the difference between a long-term interest rate and the model-implied expectations of average future short-term rates – is relative small and stationary. In most equilibrium models, however, short rates are stationary. Therefore, the large residual term in yields containing both the low-frequency variation and cyclical movements, is typically explained as certain types of risk premium (for example, the inflation risk premium in Piazzesi and Schneider (2007)).

To illustrate the role of learning in generating the observed low-frequency variation in yields and in matching the macro trends, we start with a simple equilibrium model (model-I) where both output growth and inflation are exogenous and follow i.i.d. laws of motion.

[1] See, for example, Kozicki and Tinsley (2001); Cieslak and Povala (2015); Bauer and Rudebusch (2017).

[2] Most equilibrium term structure models are designed to interpret and quantify the level, volatility, and average positive spread in yields. However, given the stationary short rates, these models typically need some extreme assumptions to match the large volatility in yields.

The representative agent with constant relative risk aversion (CRRA) utility does not, however, know the mean inflation and mean growth rate. The posteriors from standard Bayesian learning will be random walk processes, which are potentially consistent with the empirical modeling of macro trends. However, the posterior variance will decline deterministically to zero and learning will converge (Collin-Dufresne et al., 2016). We assume that the agent uses a constant-gain learning scheme as proposed by Nagel and Xu (2018) based on a weighted-likelihood approach. This is a modified Bayesian approach where learning is perpetual due to the agent's fading memory. The posteriors for mean inflation and mean growth rate, as state variables, capture the trends in inflation and growth. The short rates and yields for long-term bonds implied by the model, as linear functions of the posteriors, exhibit a low frequency movement.

The model-I implied r_t^*, which is a linear function of the posterior for mean output growth, moves closely with r_t^* estimates in the literature.[3] Not only capturing the low-frequency variation, the model-I 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 posterior for mean inflation matches very well the survey-based trend inflation in Bauer and Rudebusch (2017). 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. Over the past two decades when inflation expectations have 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.[4] However, the posteriors impact yields of all maturities equally, and yields for short- and long-term bonds are almost identical in model-I. There-

[3]We follow Bauer and Rudebusch (2017) and use the r_t^* estimates from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015).

[4]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, decline in the price of capital goods, and strong precautionary saving 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 (2017); Del Negro et al. (2017).

fore, model-I cannot explain the cyclical movements in the short-term yield and in the spread between long- and short-term yields.

To address these cyclical movements as well, we consider an extended version (model-II), where both GDP growth and inflation rates are decomposed into two components: one stable component and one transitory/volatile component. Using the same learning scheme, the representative agent learns about the mean output growth and inflation rates from the stable component, and learns about the stationary deviations from mean from the transitory/volatile component. As in model-I, the posteriors for the long-run mean inflation and growth impact short- and long-term bonds equally, and hence capture the trend in yields. However, as AR(1) processes, the posteriors for the transitory deviations from the long-run mean have larger impacts on the short-term yield than on the long-term yield, which implies that the spread would be positive (negative) when the short-run beliefs are negative (positive). Therefore, model-II can generate cyclical movements in the short-term yield, and hence in the spread, mostly due to variations in these short-run beliefs.

The model-II implied 1-year nominal yield and the spread between 10- and 1-year nominal yields match well historical movements in 1-year Treasury yield and the spread between 10- and 1-year Treasury yields.[5] Out of recessions, the short-term nominal yield starts to rise when the agent begins to revise her 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 deviations are turning into positive from negative. This pattern continues until the late expansion stage when 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.[6] The agent

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

[6]The model is consistent with our conventional understanding of business cycle. From trough to peak when inflation gap and output gap (short-run deviations in this model) are moving from negative to positive, monetary policy turns into contractionary from accommodative (short-term yields increase)

then begins to revise her short-run beliefs sharply downwards entering a recession (from positive to negative), and the spread switches from negative to positive.

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

In addition, Figure A1 shows that (1) the posteriors for the short-run inflation and growth deviations move in the opposite directions before 2000, and they move in the same direction afterwards, and (2) the posteriors for both short-run inflation and growth deviations are persistently negative for most of the post 2000 periods. These observations are consistent with the secular stagnation statement in Summers (2014), while the secular decline in r_t^* and 10-year Treasury yield in this model are mainly driven by the decline in the posteriors for the long-run means, and the decline in short- to mid-term Treasury yields are furthermore due to the persistently negative short-run deviations.

Finally, Figure 10 shows that both the nominal spreads in data and in model-II have trended upward over the sample period, and we observe less inverted curves after the 1990s. The reason is that the short-run inflation and growth deviations are strongly postively correlated (hence move the short-term yield and spread in the same direction) and they are persistently negative (hence positive spread) for the past two decades. However, their impacts on the nominal spread cancel each other out before 2000 when they move in the opposite directions. 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

according to Taylor Rule. The difference is that the posteriors for the long-run mean are unknow and keep moving in this model, while they are constants in the standard DSGE models.

the inflation risk premium approach by Piazzesi and Schneider (2007), where inflation is bad news for future growth and the agent prefers early resolution of uncertainty (Epstein and Zin (1989) preference). Zhao (2018) shows that this approach has been less effective during the past two decades when inflation has switched from bad news to good news for future growth, providing an alternative worst-case belief approach through ambiguity. We therefore extend model-II to model-III by incorporating the intuition in Zhao (2018). The ambiguity-averse representative agent (with recursive multiple priors, or maxmin, preferences by 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 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 ambiguity (following Ilut and Schneider 2014), Zhao (2018) finds that 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. Consistent with findings in Froot (1989) and Piazzesi et al. (2015), the expectations hypothesis roughly holds under the subjective equilibrium belief.

Given that the benchmark measure in model-III is the same as in model-II, model-III can still match the trend and cycle in yields as model-II does. However, by comparing the model-implied 1-year nominal yield with data, we observe a recurring pattern that the model-implied short-term yields are higher than data from trough to expansion stage, and they are lower than 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 suggested by the model and there was certain degree of overshoot during the late expansion periods (before recessions). In addition, because the agent's subjective inflation and growth expectations are upward-sloping under her worst-case belief, the model can also generate upward-sloping real and nominal curves

on average.[7] Figure 14 shows that the model-III-implied spread matches the data well. Specifically, compared with the model-II-implied nominal spread in Figure 10, the model-III-implied nominal spread is almost in parallel higher.

Related literature

The paper is related to a large literature in equilibrium asset/bond pricing models.[8] While most equilibrium bond-pricing models focus only on the first/second moment and the average spread in yields, we go further in this paper by providing a joint equilibrium understanding of the trend, cycle, and spread in the historical Treasury bond yields. The most closely related paper is Piazzesi and Schneider (2007) who show the importance of the inflation risk premium for the upward-sloping nominal curve in a stationary state space model (inflation and growth). They also show how recursive estimation (with more weight on recent observations) of the whole state space model can help improve the model's historical performance.

This paper differs from these previous studies along some important dimensions. First, the agent in this paper takes into consideration the risk of belief updating as in Collin-Dufresne et al. (2016) and Nagel and Xu (2018). Hence the posteriors are state variables, which move closely with macro trend estimations (r_t^* and π_t^*) in the literature. Most importantly, this is the first paper that decomposes both GDP growth and inflation

[7]The empirical literature has 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; Champagne et al., 2018). In this model, the reference measure represents the best point estimate from data and agents would use it for forecasting. However, agents are more cautious when they make decisions, and they instead use the worst-case belief for decision making. Therefore, ambiguity in this model provides a rational explanation for the expectation errors.

[8]For example, Bansal and Shaliastovich (2013) show that the long-run risks model of Bansal and Yaron (2004) with time-varying volatility can account for bond return predictability. Rudebusch and Swanson (2012) show that a macroeconomic DSGE model with inflation non-neutrality and Epstein and Zin (1989) preference can also generate sizable inflation risk premium. Wachter (2006) generates an upward-sloping nominal using an external habits model, where innovations to consumption and inflation growth are negatively correlated. Lettau and Wachter (2011) model the term structures of equity and interest rates by specifying a parsimonious stochastic discount factor. Albuquerque et al. (2016) show that the risk to time preference implies a positive real term premium and can generate an upward-sloping real yield curve. Berrada et al. (2018) show the risk to belief updating implies a positive real term premium when agents have beliefs-dependent risk aversion.

into two components; And shows that learning about the long-run mean drives the low-frequency variation in yields and learning about the short-run deviation from the mean drives the business cycle movements in the short-term yield, and hence in the spread. Instead of real risk premium (Wachter, 2006; Albuquerque et al., 2016; Berrada et al., 2018) and inflation risk premium (Piazzesi and Schneider, 2007), the upward-sloping nominal and real curves in the U.S., at least for the post 2000 periods, are *partially* due to the fact that both short-run deviations for inflation and growth are negative for most of these periods. However, the short-run deviations for inflation and growth move in the opposite directions before 2000, which makes it hard for model-II to generate an average upward-sloping nominal curve. We therefore rely on one additional worst-case belief approach (Zhao, 2018) 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.[9] The most closely related paper is Bauer and Rudebusch (2017) who model r_t^* and π_t^* as random walk processes and show the importance of these trends in explaining the low-frequency variation in the long-term yield. 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 the short-term yield, and hence in the spread.

This paper is related to a number of papers that have studied the implications of ambiguity and robustness for finance and macroeconomics.[10] Model-III in this paper incorporates the intuition in Zhao (2018) 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 related to some recent developments that

[9]See Ang and Piazzesi (2003), Diebold et al. (2006), Wright (2011), Kozicki and Tinsley (2001), Cieslak and Povala (2015), and Bauer and Rudebusch (2017), among many others.

[10]See Epstein and Schneider (2003), Epstein and Schneider (2007), Ilut and Schneider (2014), Ulrich (2013) , Drechsler (2013), Gagliardini et al. (2009), Zhao (2017), and Zhao (2018), among many others. For a detailed survey, see Epstein and Schneider (2010).

study the implications of learning in finance. For example, Collin-Dufresne et al. (2016) show how a standard Bayesian learning can generate subjective long-run risks when agent prefers early resolution of uncertainty (Epstein and Zin (1989) preference). Building on the insight in Nagel and Malmendier (2016) that the dynamics of the average individuals' expectation can be approximated closely by a constant-gain learning scheme, Nagel and Xu (2018) show how this constant-gain learning can help separate subjective and objective equity premium and explain the predictability of excess returns. In this paper, we use the constant-gain learning scheme in a different setting to explain bond yields dynamics.

The paper continues as follows. Section 2 outlines and solves model-I in closed form and discusses the model implications. Sections 3 and 4 undertake the same steps for model-II and model-III, respectively. Section 5 provides concluding comments.

2. Model I - Learning and trend in bond yields

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

2.1. Learning with fading memory

Both output growth and inflation follow i.i.d. laws of motion

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

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 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 σ_c and σ_π, but not μ_c and μ_π. The agent forms expectations about μ_c and μ_π based on the history of 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 would update her prior belief $p(\mu_c)$ or $p(\mu_\pi)$ in a way that assigns each past observation Δg_{t-j} or π_{t-j} equal weight in the posterior probability. The equal-weighting of past observations in H_t^g and H_t^π means that there is no decay of memory as the agent uses all available data in forming the posterior beliefs.

In this paper, we use a constant-gain learning scheme proposed by Nagel and Xu (2018) 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 her 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 weights recent observations more than observations 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 $\sigma_{\pi,0} \to \infty$.[11] The posteriors are given by

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

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,

[11] See Nagel and Xu (2018) for a discussion of the informative prior case.

learning is perpetual here. The variance of the posterior is the same as if the agent had observed, and retained fully 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).

To understand better the stochastic nature of the output growth and inflation process from the agent's subjective viewpoint, we further get the predictive distribution

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

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

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

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 unpredictable under the time-t predictive distribution.

2.2. Valuation with fading memory

The information structure for a standard Bayesian agent can be represented by a filtration, and posterior beliefs follow a martingale under this filtration. With loss of memory, however, the posterior in periods $t+j$ will be updated based on information that is different, but not more informative, than the information available to the agent at time t. Thus, the information structure is not a filtration. Nagel and Xu (2018) 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 stochastic discount factor (SDF) from $t+j-1$ to $t+j$ that applies given the agent's predictive distribution at t. The time-t price of a zero-coupon bond that pays one unit of consumption 2 periods from now is denoted $P_t^{(2)}$, and it 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 the optimal filtering with a latent stochastic trend. For output growth, if the agent perceives μ_c to follow a random walk ($\mu_{c,t} = \mu_{c,t-1} + \varepsilon_{c,t}^{\mu}$), rather than a constant as in (1), 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 valuation for assets would be the same as well. If the true law of motion is (1) with constant μ_c, there will be a time-varying wedge $\mu_{c,t} - \tilde{\mu}_{c,t}$ between subjective and objective beliefs, which plays an important role in generating excess bond return predictions. 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 is highly

persistent before the late 1990s and becomes less predictable thereafter. Although the model implied bond prices and yields are the same, it is more reasonable to assume a latent random walk trend with Kalman filter learning before the late 1990s and a constant mean with adaptive learning after the late 1990s.

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 investors have recursive preferences with a CRRA utility function (i.e., they are indifferent between 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 risk aversion, and β reflects the investor's time preference. Note that the agent evaluates the continuation value under her subjective expectations.

2.4.1. Bond pricing

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

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

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 $P_t^{(n)}$, 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 the 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 $er_{n,t+1} = -Cov_t\left(r_{n,t+1}, m_{t,t+1}^{\$}\right) = -v'Cov_t\left(z_{t+1}, \tilde{\mu}_{t+1}\right)C^{(n-1)}$.

As we can see from the solution, the yield parameter for posterior $\tilde{\mu}_t$ is constant over horizon n, and all the variance and covariance terms are relatively small in the data. Hence, given the CRRA utility (i.e., no extra term premium from 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 just replace v' with $v' = (\gamma, 0)$.

2.5. Empirical findings

Using U.S. real GDP growth and the rate of inflation from the GDP deflator, we can calculate the posterior beliefs for 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 historical movements in the data well.

2.5.1. Data

Real output growth and GDP deflator inflation are taken from the Bureau of Economic Analysis from 1947.Q2 to 2018.Q2. The end-of-quarter yields for one- to ten-year bonds

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 parameter values for output growth and inflation processes, and for the constant gain in learning. All parameters are given in quarterly terms. Mean and standard deviation are in percentages.

are from the daily dataset constructed by Gürkaynak et al. (2007) from 1961.Q2 to 2018.Q2. The TIPS yields (2003.Q1 to 2018.Q2) and end-of-quarter yields for three-month and six-month Treasury bills are 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) from 1968.Q3 to 2018.Q2. The r^* and π^* are from Bauer and Rudebusch (2017) from 1971.Q4 to 2017.Q2.

2.5.2. Parameters

The volatility parameters for output growth and inflation are calibrated to match their counterparts in data. The correlation between output growth and inflation in the model is calibrated to match the correlation in data. The constant-gain parameters υ_c and υ_π are calibrated to match variations in r_t^* and π_t^*. We follow the literature and set risk aversion as 3, and time preference β is calibrated to match the level of ten-year nominal yields in the data, which are close to the value in Piazzesi and Schneider (2007).[12] The resulting parameter values are reported 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 interest rate modeling. For example, Kozicki and Tinsley (2001) and Cieslak and Povala (2015) document the empirical importance of the inflation trend for the dynamics of the nominal yields. Bauer and Rudebusch (2017) show that it is crucial to include r_t^* as well, especially for explaining the downtrend in yields over the

[12] A higher time preference β helps to lower bond yield levels. Zhao (2018) shows that ambiguity about output growth can also lower nominal yields. Therefore, we can also set β to be smaller than 1, but with a larger amount of ambiguity to match the bond yield level.

last 20 years as inflation expectations have stabilized, and they model trend inflation and r_t^* as random walk processes.

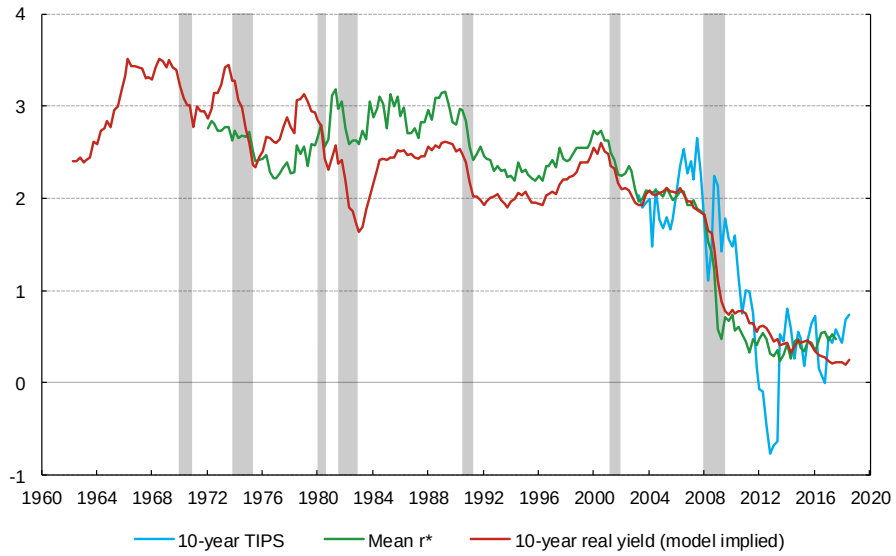
The stationary assumption in leading equilibrium bond pricing models makes it hard for them to generate the historical observed low-frequency variation in interest rates. The posterior mean of output growth (inflation) in Section 2.1 would be a random walk process under standard Bayesian learning; however, the posterior variance would decline deterministically to zero and learning would converge (Collin-Dufresne et al., 2016). In this paper, the representative agent updates her subjective beliefs with constant gain, which induces memory loss, and is otherwise standard Bayesian. However, learning is slow but perpetual in this model, which generates low-frequency variation in posterior beliefs that closely match the r_t^* and π_t^* estimated in the literature and, conveniently, a stationary economy.[13]

Figure 1 shows the model-implied 10-year real yield, which is a linear function of the posterior mean of 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 from Bauer and Rudebusch (2017) and is an average of three macroeconomic estimates of r^* from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015). Figure 2 shows that the model-implied 10-year real yield and the three individual estimates of r^* also co-move closely for most of the sample periods. The three different r^* estimations diverge from each other before the 1980s. Not only capturing the low-frequency variation, the model-I 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 mean inflation matches very well the trend inflation π_t^* from Bauer and Rudebusch (2017). The trend inflation is a survey-based measure, namely, the Federal Reserve's series on the perceived inflation

[13] At each time t, the agent perceives that shocks today are negatively serially correlated with future shocks; see proof in Nagel and Xu (2018).

Figure 1: **Average r* and the model implied 10-year real yield**



The average r^* (quarterly data) are from Bauer and Rudebusch (2017) from 1971:Q4 to 2017:Q2. The model-implied 10 year real yields (quarterly data) are from 1962:Q1 to 2018:Q2. The end-of-quarter 10-year TIPS yields are from the Fed database at the St.Louis Federal Reserve from 2003:Q1 to 2018:Q2. The gray bars represent periods of recession defined by the NBER.

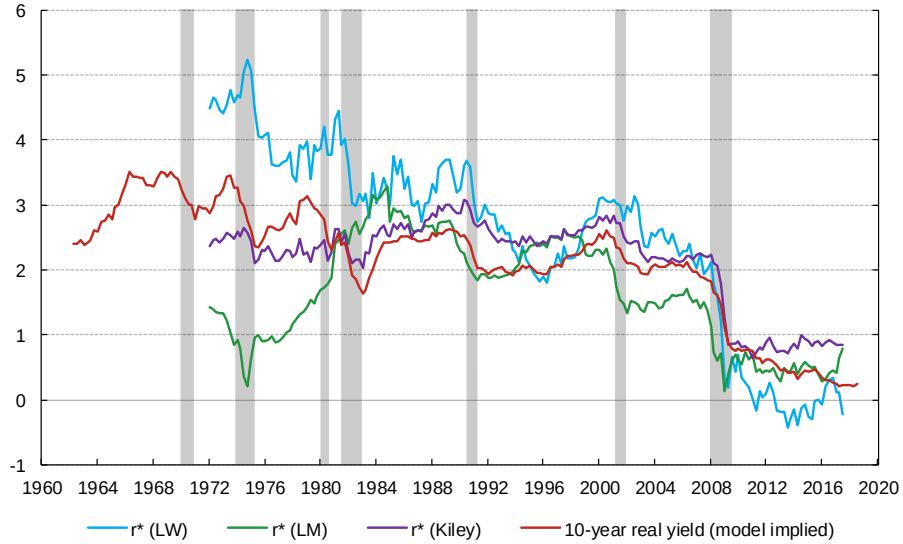
target rate, denoted PTR. It measures long-run expectations of inflation in the price index of personal consumption expenditure (PCE). The result confirms our assumption in the model that the representative agent forms her inflation expectation based on the history of past inflation rates.

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 equilibrium models (and no-arbitrage term structure models in general) are generally modeled as stationary, mean-reverting processes. As a result, low-frequency variation in interest rates is hard to explain in such models and is mostly attributed to the term premium component, which is a residual term in empirical models and usually the inflation risk premium in equilibrium models (Piazzesi and Schneider, 2007).

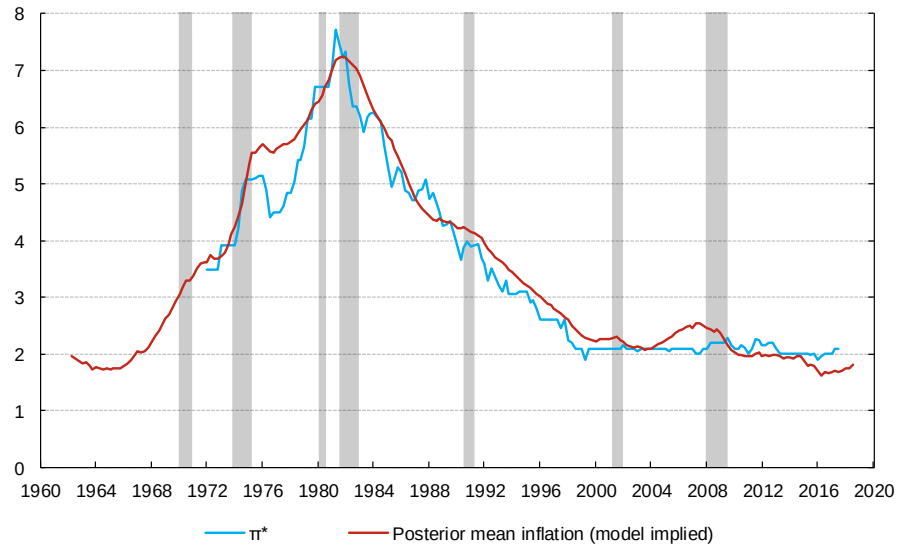
As shown in Figures 1, 2, and 3, the posteriors ($\gamma\tilde{\mu}_{c,t}$ and $\tilde{\mu}_{\pi,t}$) in model-I match well the macro trends (r_t^* and π_t^*). An illustration of the potential importance of these posteriors in the 10-year nominal yield is provided in Figure 4. The hump-shaped 10-year

Figure 2: **Individual r* and the model implied 10-year real yield**



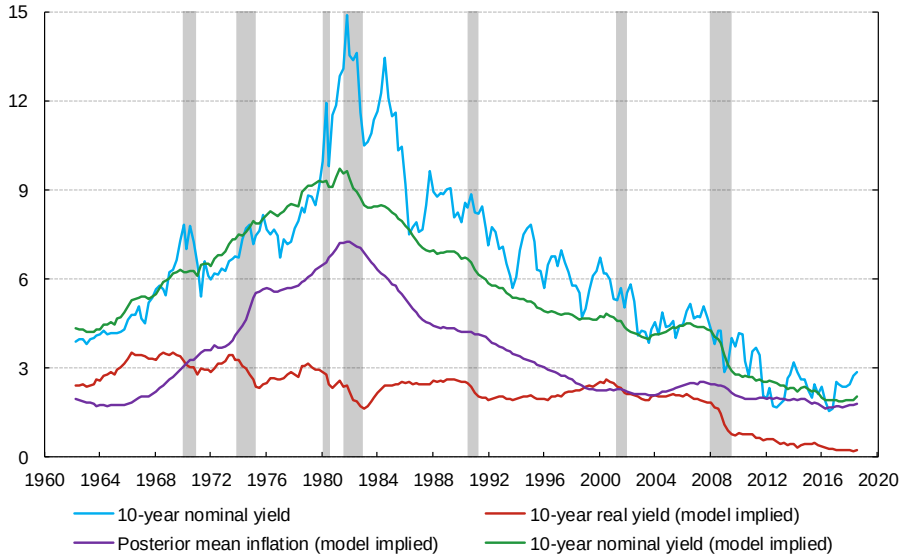
The individual r^*s (quarterly data) are from Bauer and Rudebusch (2017) from 1971:Q4 to 2017:Q2. Three macroeconomic estimates of r^*s are from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015), respectively. The model-implied 10-year real yields (quarterly data) are from 1962:Q1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

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



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

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. Over the past two decades, as inflation expectations have stabilized, the pronounced decline in output growth expectations, and hence the 10-year real yield, is the main driver of the downtrend in nominal yields. As a result, the model-implied 10-year nominal yield captures well the trend movements in the 10-year Treasury yield for the whole sample.

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 track well the macro trends. Therefore, model-I explains well the trend in the 10-year Treasury yield. However, model-I is silent about two other important features in yields - the cycle and term spread.

Historically, short-term bond yields are lower (on average) and much more volatile than long-term bond yields. The spread between long-term and short-term bond yields is

positive in general, shrinks in the late expansion stage of the business cycle, and becomes inverted before recessions. Given the CRRA utility, the term premium is very small in model-I, and so the model-implied yields for 10-year and 1-year bonds are almost identical. Hence the model is not able to explain dynamics in the term spread and cyclical movements in short-term bond yields. We extend model-I in the next session to overcome these shortcomings.

3. Model II - Learning, trend, and cycle in bond yields

In model-II, we still consider an endowment economy with a representative agent who has a CRRA utility function. The agent still learns with fading memory as in model-I. 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 the mean output growth and inflation rates from the stable component, and learns about the stationary deviations from 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 - don't move in lockstep with each other. In fact, their levels of volatility differ greatly. Consumption is highly stable and varies less with the business cycle. In contrast, the other three components are extremely volatile, varying greatly during economic contractions and expansions. For this reason, we assume that there are two unknowns for the agent to learn in the output growth process: the long-run mean and a latent stochastic deviation from the mean (stationary). The agent learns about the long-run mean GDP growth only from the stable component (PCE), and learns about the stationary deviation from mean using only the volatile component (GDP growth excluding PCE). Similarly, for inflation, the agent learns about long-run mean inflation only from core inflation, and learns the transitory and stationary deviation from the mean using only transitory price changes (GDP deflator excluding core inflation). Formally,

output growth and inflation can be decomposed into (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 total real GDP growth and inflation, respectively. Δg^*_{t+1} and π^*_{t+1} are real consumption growth (scaled by total real GDP $\frac{C_{t+1}-C_t}{GDP_t}$) and core inflation (scaled by total price level $\frac{P^{core}_{t+1}-P^{core}_t}{P_t}$), 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 i.i.d. laws of motion

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

where $\varepsilon^*_{c,t+1}$ and $\varepsilon^*_{\pi,t+1}$ are i.i.d. normal shocks. As with model-I, the representative investor knows that both Δg^*_{t+1} and π^*_{t+1} are i.i.d., and also knows 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,

$$\begin{aligned}\mu^*_c | H^{g^*}_t &\sim N\left(\tilde{\mu}^*_{c,t}, \upsilon^*_c \sigma^2_c\right) \\ \mu^*_{\pi} | H^{\pi^*}_t &\sim N\left(\tilde{\mu}^*_{\pi,t}, \upsilon^*_{\pi} \sigma^2_{\pi}\right),\end{aligned} \tag{14}$$

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,

$$\begin{aligned}\Delta g^*_{t+j}|H^{g^*}_t &\sim N\left(\tilde{\mu}^*_{c,t}, (1+\upsilon^*_c)\,\sigma^2_c\right)\\ \pi^*_{t+j}|H^{\pi^*}_t &\sim N\left(\tilde{\mu}^*_{\pi,t}, (1+\upsilon^*_\pi)\,\sigma^2_\pi\right),\end{aligned} \tag{16}$$

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 a latent stationary component as in Bansal and Yaron (2004) and Piazzesi and Schneider (2007):

$$\begin{aligned}Gap^g_{t+1} &= x_{c,t+1} + \sigma^{gap}_c \varepsilon^{gap}_{c,t+1}\\ Gap^\pi_{t+1} &= x_{\pi,t+1} + \sigma^{gap}_\pi \varepsilon^{gap}_{\pi,t+1}\\ x_{c,t+1} &= \rho_c x_{c,t} + \sigma^x_c \varepsilon^x_{c,t+1}\\ x_{\pi,t+1} &= \rho_\pi x_{\pi,t} + \sigma^x_\pi \varepsilon^x_{\pi,t+1},\end{aligned} \tag{17}$$

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 the parameters, but not $x_{c,t+1}$ and $x_{\pi,t+1}$. She forms 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^{gap}_c and υ^{gap}_π. The posteriors are given by

$$\begin{aligned}x_{c,t+1}|H^{gap}_{g,t} &\sim N\left(\rho_c \tilde{x}_{c,t}, \upsilon^{gap}_c \left((\sigma^x_c)^2 + (\sigma^{gap}_c)^2\right)\right)\\ x_{\pi,t+1}|H^{gap}_{\pi,t} &\sim N\left(\rho_\pi \tilde{x}_{\pi,t}, \upsilon^{gap}_\pi \left((\sigma^x_\pi)^2 + (\sigma^{gap}_\pi)^2\right)\right)\\ \tilde{x}_{c,t} &= \rho_c \tilde{x}_{c,t-1} + \upsilon^{gap}_c \left(Gap^g_t - \rho_c \tilde{x}_{c,t-1}\right)\\ \tilde{x}_{\pi,t} &= \rho_\pi \tilde{x}_{\pi,t-1} + \upsilon^{gap}_\pi \left(Gap^\pi_t - \rho_\pi \tilde{x}_{\pi,t-1}\right),\end{aligned} \tag{18}$$

where $H^{gap}_{g,t} \equiv \{Gap^g_0, Gap^g_1, ...Gap^g_t\}$ and $H^{gap}_{\pi,t} \equiv \{Gap^\pi_0, Gap^\pi_1, ...Gap^\pi_t\}$. As discussed in 2.3, the updating is same as for an optimal Kalman filtering with an appropriate choice of parameter values.

To understand better the stochastic nature of the output growth and inflation process

from the agent's subjective viewpoint, we further get the total predictive distribution

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

where $j = 1, 2, ...$ and the variance of the predictive distribution contains both uncertainty due to future shocks and 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 in equation (10).

Model-I has two state variables (the posterior means for output growth and inflation) that explain the trend in 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 the cyclical component from 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 the variance and covariance terms are relatively small in the data. Hence, given 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 long-term and short-term yields are the same, which explains the low frequency movements (trend) in 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 short-term yield. The spread between long-term 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 data, the spread is mean zero because of the stationarity assumption for x_t. To solve the price and yields for real bonds, we can just replace v' with $v' = (\gamma, 0)$.

3.3. Empirical findings

We use the same data sets as in model-I, but real GDP growth and the rate of inflation from the GDP deflator are decomposed into one stable component and one transitory component. 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 10-year real yield closely matches the estimated r star in the literature and the total posterior belief for inflation closely matches the survey-based trend inflation. As a result, the model-implied 10-year nominal bond yields match the historical trend movements in the 10-year Treasury yields well. In addition, the model can also capture the cyclical movements in 1-year Treasury yields.

3.3.1. Parameters

The volatility parameters for 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 the 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 parameter for the model 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 volatility in transitory GDP growth (transitory inflation). In the model, the correlation in stable

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 reports parameter values for output growth and inflation processes, and for the constant gain in learning. All parameters are given in quarterly terms. Mean and standard deviation are in percentages.

and transitory components between output growth and inflation is calibrated to match the correlation in the data

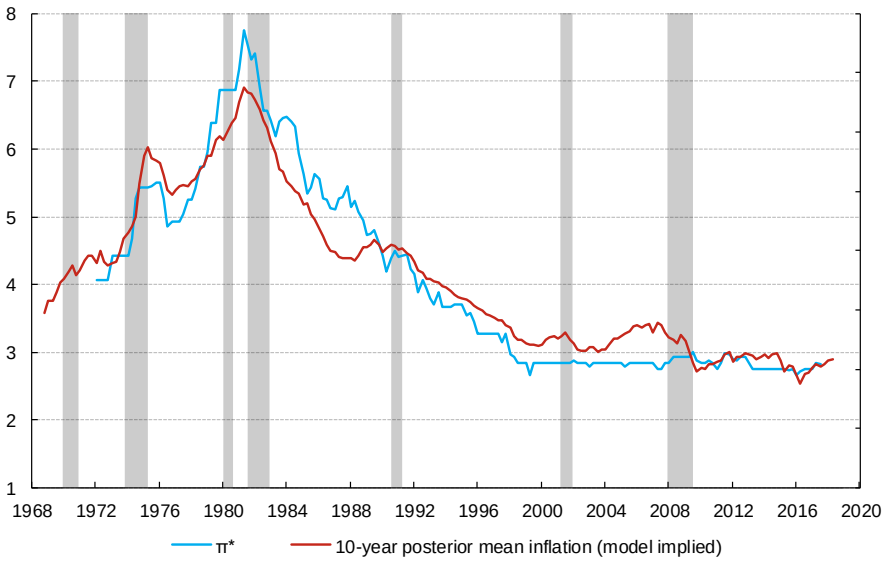
The parameters for learning in inflation υ_π^*, ρ_π, and υ_π^{gap} are calibrated to match variations in π_t^* and the one-quarter-ahead mean survey inflation expectation (persistence and volatility). The parameters for learning in growth υ_c^*, ρ_c, and υ_c^{gap} are calibrated to match variations in r_t^* and one-year nominal yields (persistence and volatility). We follow the literature and set risk aversion as 3, and time preference β is calibrated to match the level of ten-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; in fact, their levels of volatility differ greatly. Hence, it is natural for the agent to learn the long-run mean and a stationary deviation from the mean separately. The agent uses a longer 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 updates her posterior beliefs slowly but perpetually.

Figure 5 shows that the model-implied total posterior belief for 10-year-ahead mean inflation matches very well the survey-based PTR trend inflation π_t^*. Similarly, Figure 6 shows that the model-implied total posterior belief for 1-quarter-ahead mean inflation also tracks very closely the survey mean for 1-quarter-ahead inflation from SPF. The result confirms our assumption that the agent forms her inflation expectations differently

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



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

for long-run versus 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 $\gamma\tilde{\mu}_{c,t}^* + Cov$, and closely tracks the three macroeconomic estimates of r^*.[14]

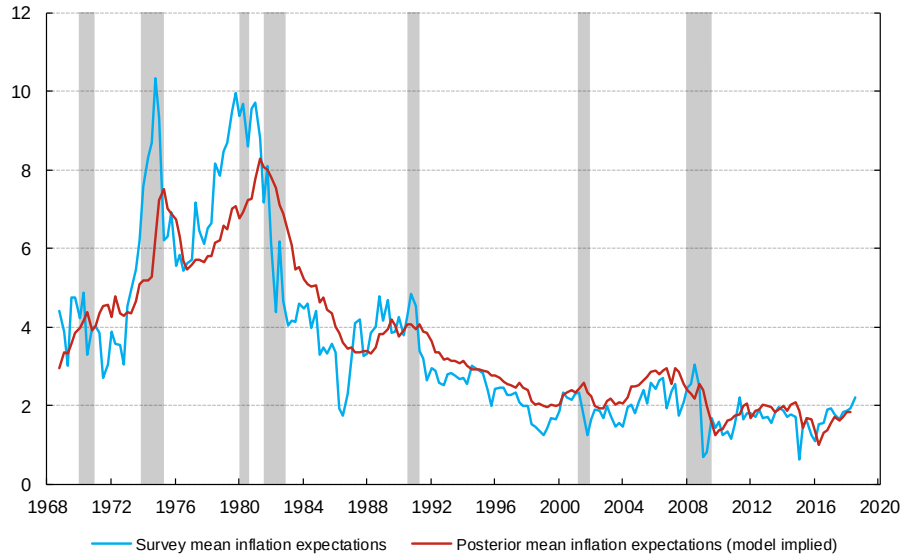
3.3.3. 1-year and 10-year Treasury yields, and their spread

In model-I, we have shown the importance of macro trends in explaining the movements in the long-term nominal yield. For the same reason, the posteriors for long-run mean growth and inflation move slowly and match well the macro trends in model-II, and Figure 8 shows the potential importance of these posteriors in the 10-year nominal yield. The secular decline in the 10-year Treasury yield after the 1980s is 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 explain only the trend in long-term yields. In model-II, however, learning about the deviation

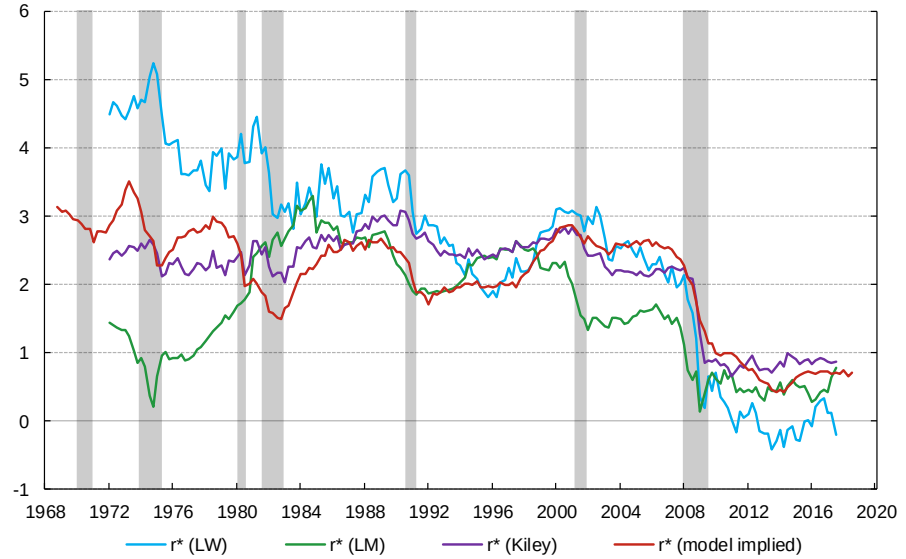
[14]To be consistent with the concept of r^*, the short-run effect from $\tilde{x}_{c,t}$ on the real yield is not included for calculation of the model-implied r^*, because it will eventually vanish. Therefore the model-implied r^* impacts the real yields of any maturity equally.

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



The mean survey one-quarter-ahead inflation (quarterly data) are 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) is from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession defined by the NBER.

Figure 7: Individual r* and model implied r*



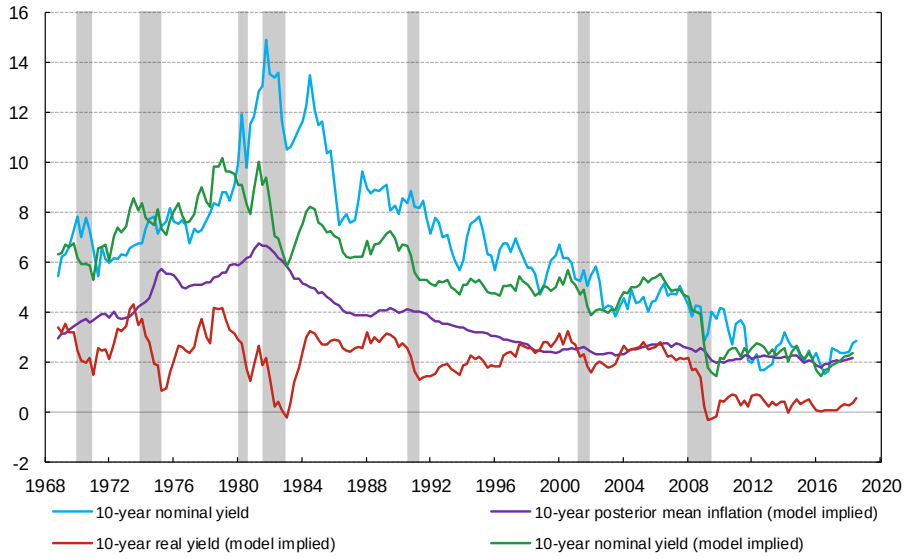
The individual r^*s (quarterly data) are from Bauer and Rudebusch (2017) from 1971:Q4 to 2017:Q2. Three macroeconomic estimates of r^*s are from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015). The model-implied r^* (quarterly) is from 1968:Q3 to 2018:Q2. The gray bars represent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

from mean allows us to also capture the cyclical movements in short-term yields. Figure 9 shows that the model-implied 1-year nominal yield tracks relatively well the historical data for the 1-year Treasury yield, with some exceptions. For the post-global financial crisis period, 1-year Treasury yields are mostly constrained by the zero lower bound, but the model-implied 1-year nominal yields are much more volatile, and these two have started to line up again recently. Also at the beginning of the sample period, the model-implied 1-year nominal yield seems to be 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 long-term and short-term bond yields is almost zero, due to the CRRA utility and the equal impacts of posteriors on real yields of any maturity. We still have the equal impacts of $\tilde{\mu}_t^*$ on yields of all maturities in this model (capturing the long run trends). But, as shown in the model-II solution, the impact of x_t (as AR(1) processes) on the short-term yield is bigger than on the long-term yield, which implies that the spread would be positive (negative) when the short-run beliefs are negative (positive). Hence, model-II can generate cyclical movements in the short-term yield and in the spread, mostly due to variations in x_t. Figure 10 shows that the model-implied 10-year-minus-1-year yield spread correlates well with the spread in the data.

Starting from the previous trough, the short-term nominal yield starts to rise when the agent begins to revise her 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 deviations are turning into positive from negative. This pattern continues until the late expansion stage (or peak) when 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, the inflation gap and the output gap move from negative to positive, and monetary policy turns into contractionary from accommodative (short-term yields increase) according to Taylor Rule. The agent then begins to revise her short-run beliefs sharply downwards entering a recession (from positive to negative), and the spread

Figure 8: 10-year nominal rate and macro trends



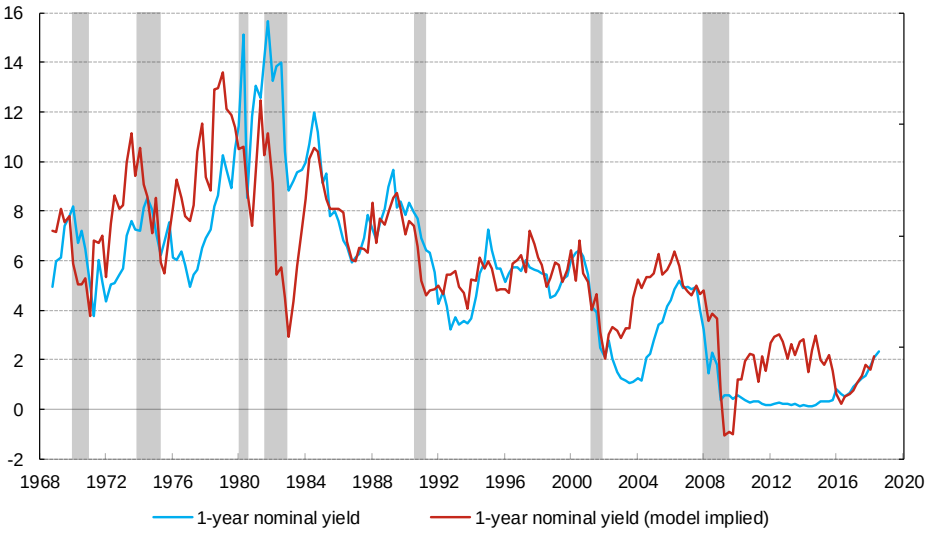
The end-of-quarter 10-year nominal yields are from Gürkaynak et al. (2007) from 1968:Q3 to 2018:Q2. The model-implied 10-year real yield, 10-year nominal yield, and total posterior belief for mean inflation (quarterly data) are from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession defined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

switches from negative to positive.

Furthermore, Figure A1 shows that (1) the posteriors for the short-run deviations ($\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$) move in the opposite directions before 2000, and they move in the same direction afterwards, and (2) $\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$ are persistently negative for most of the post 2000 periods.[15] 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) are likely cancel each other out before 2000, and they simultaneously lower short-term nominal yields (and the nominal spread) afterwards. Therefore, the model can generate an upward trend in nominal spread that is consistent with the data in Figure 10. These observations are also consistent with the secular stagnation statement in Summers (2014) that output gap ($\tilde{x}_{c,t}$ in this model) and inflation gap ($\tilde{x}_{\pi,t}$ in this model) are persistently negative after 2000. The secular decline in r_t^* and 10-year Treasury yield are mainly driven by the decline in the posteriors for the long-run means, and the decline in short- to mid-term Treasury yields are furthermore due to the persistently negative $\tilde{x}_{c,t}$ and $\tilde{x}_{\pi,t}$.

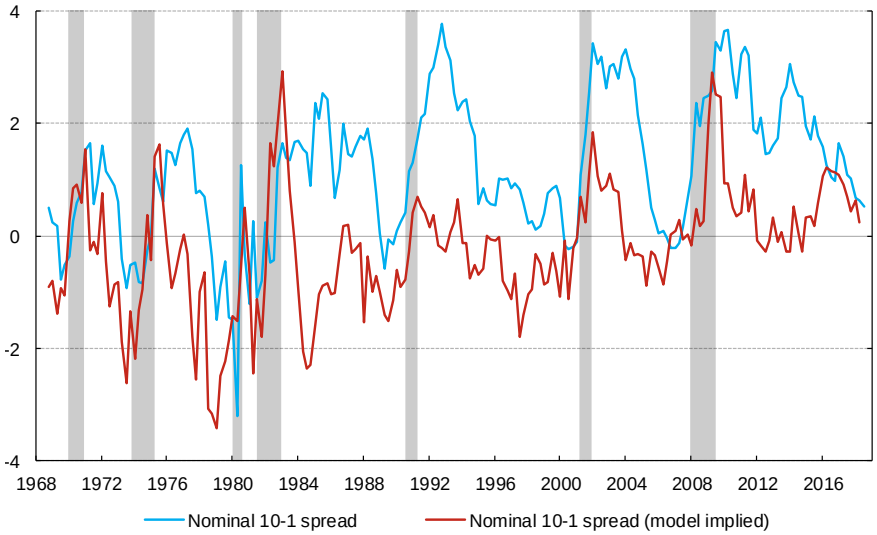
[15]The first observation is consistent with the finding that inflation was bad news for future growth before 2000 (Piazzesi and Schneider, 2007), and switched to good news afterwards (Burkhardt and Hasseltoft, 2012; David and Veronesi, 2013; Campbell et al., 2016; Zhao, 2018).

Figure 9: **1-year nominal yield**



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

Figure 10: **Nominal (10-year - 1-year) spread - actual vs. model**



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

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 (as shown in Figure 10), due to the stationary assumption (mean zero) for x_t, as well as the CRRA utility.

3.3.4. Model-II limitations: upward-sloping yield curve

Given the trend ($\tilde{\mu}_t^*$) and cycle (x_t) posterior beliefs, model-II is able to match both the long run trend and the cyclical variations in yields. But, as shown in Figure 10, the model-implied spread is lower than the data, and in fact, is mean zero because of the stationarity assumption for x_t. However, in the data, the spread between long-term and short-term bond yields is positive in general, and only becomes inverted before recessions.

There are two relevant approaches that can generate an upward-sloping nominal yield curve in an equilibrium model: the inflation risk premium approach by Piazzesi and Schneider (2007) and the worst-case belief approach (through ambiguity) by Zhao (2018). The inflation risk premium approach relies on inflation as bad news for future growth and the assumption that agents prefer early resolution of uncertainty. Given that inflation has switched from bad news to good news for future growth over the past two decades, the inflation risk premium has become negative, which implies a downward-sloping nominal curve. Therefore, we will extend model-II in the next session to incorporate the intuition in Zhao (2018) and generate an upward-sloping yield curve through ambiguity.

4. Model III - Learning, trend, cycle, and spread in bond yields

In model-III, we continue to consider an endowment economy with a representative agent who has a CRRA utility function and the same learning structure as in model-II. Given the learning and posteriors, in addition, the agent is assumed to face a certain amount of Knightian uncertainty regarding GDP growth and inflation. Equilibrium prices adjust such that the ambiguity-averse agent is happy to consume the output as an endowment.

4.1. Learning and ambiguity

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

$$\begin{aligned}\Delta g_{t+1} &= \tilde{\mu}^*_{c,t} + \rho_c \tilde{x}_{c,t} + \varepsilon_{c,t+1} \\ \pi_{t+1} &= \tilde{\mu}^*_{\pi,t} + \rho_\pi \tilde{x}_{\pi,t} + \varepsilon_{\pi,t+1},\end{aligned} \tag{21}$$

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 her 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 $p^{\tilde{a}}$, output growth and inflation are as follows:

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

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.[16]

Using forecast dispersion from the Blue Chip Financial Forecast (BCFF) survey as a measure for the size of ambiguity, Zhao (2018) finds that, before the late 1990s, the size of ambiguity for long-horizon inflation is bigger than those for short horizons, and this pattern is reversed afterwards. However, the size of ambiguity for long-term real

[16]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.

output growth is always smaller than those for the short term. Together with the fact that positive inflation shocks have changed from negative to positive news about future growth in the last twenty years, the ambiguity-averse agent chooses the upper bound ($a_{\pi,t}$) as the worst-case measure for inflation before the late 1990s, and chooses the lower bound ($-a_{\pi,t}$) as the worst-case measure afterwards. While for output growth (as the endowment), the agent always chooses the lower bound ($-a_{c,t}$) as the worst-case measure in equilibrium. Zhao (2018) shows 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 (2018) and model $a_{c,t}$ ($a_{\pi,t}$) as a random walk with drift,

$$\begin{aligned} a_{c,t+1} &= \mu_c^a + a_{c,t} + \sigma_{ac}\varepsilon_{ac,t+1} + \sigma_a^{ac}\varepsilon_{a,t+1} \\ a_{\pi,t+1} &= \mu_\pi^a + a_{\pi,t} + \sigma_a^{a\pi}\varepsilon_{a,t+1}, \end{aligned} \tag{23}$$

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 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:

$$V_t(C_t) = \min_{p_t \in \mathcal{P}_t} \mathbb{E}_{p_t}\left(U\left(C_t\right) + \beta V_{t+1}(C_{t+1})\right), \tag{24}$$

where $U\left(C_t\right) = \frac{C_t^{1-\gamma}-1}{1-\gamma}$, γ is the coefficient of 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 trend and cycle in 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 $a_t = (a_{c,t}, a_{\pi,t})^T$:

$$p_t^{(n)} = -A^{(n)} - B^{(n)} x_t - C^{(n)} \tilde{\mu}^*_t - D^{(n)} a_t. \tag{25}$$

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 bound, equal to -1 or $+1$ on the diagonal. All 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 yields, and x_t captures the cyclical movements in 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 yields for the whole sample periods ($\phi^a_c = -1$), and $a_{\pi,t}$ lowers (raises) yields for the second subperiod when $\phi^a_\pi = -1$ (the first subperiod when $\phi^a_\pi = 1$). The impacts of a_t on long-term 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 just replace v' with $v' = (\gamma, 0)$.

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 parameter values for output growth and inflation processes, the constant gain in learning, and the ambiguity process. All parameters are given in quarterly terms. Mean and standard deviation are in percentages.

4.4. Empirical findings

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

4.4.1. Parameters

All parameters (excluding the ambiguity parameters) are estimated the same way as in model-II. The ambiguity parameters are the same as in Zhao (2018), where the whole sample period is split into two subperiods and the parameters are different for each subperiod (mainly the trend parameter for inflation ambiguity μ_π^a). The realized size of ambiguity is measured by the past one-year average of SPF forecast dispersions that are calculated by the 60th percentile minus the 40th percentile of individual forecasts. The resulting parameter values are reported in Table 3.

4.4.2. Trend, cycle, and spread in 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 short- and long-term inflation expectations in this model are also exactly the same as shown in Figure 5 and Figure 6. Both the learning

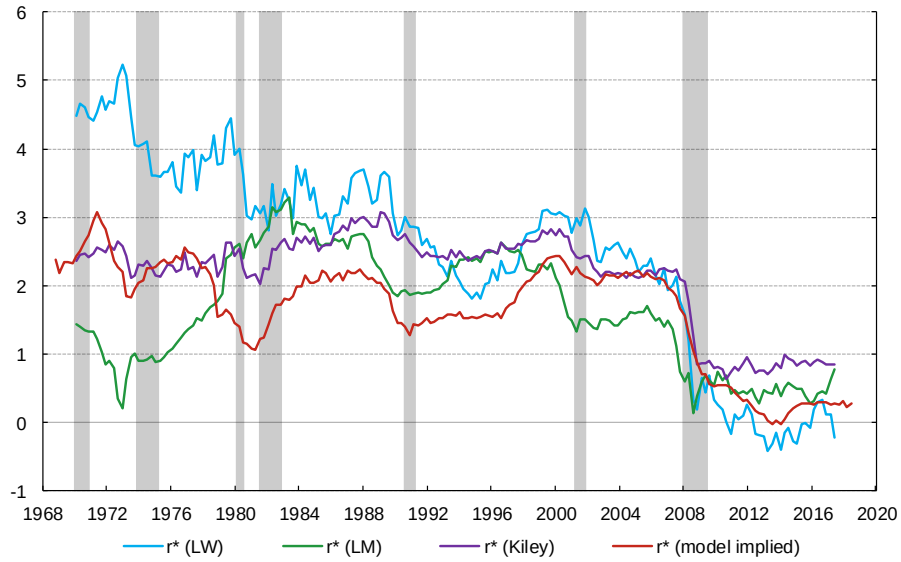
parameters and the posterior beliefs for short- and long-term output growth are also the same as in model-II. But the time preference is slightly different from the previous value; therefore, as shown in Figure 11, the model-implied r_t^* is almost the same as before.[17]

As in model-I and model-II, Figure 12 shows the importance of the posteriors for long-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 short-run deviation from the long-run mean in explaining the movements in the one-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, 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.

Finally, by comparing the model-implied 1-year nominal yield with data in Figures 13&9, we observe a recurring pattern that the model-implied short-term yields are higher than data from trough to expansion stage, and they are lower than 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 suggested by the model (behind the curve) and there was certain degree of overshoot (ahead of the curve) during the late expansion periods (before recessions). This is consistent with Taylor (2018) who argues that the over accommodative policy between 2003-2005 was a source of the housing bubble.

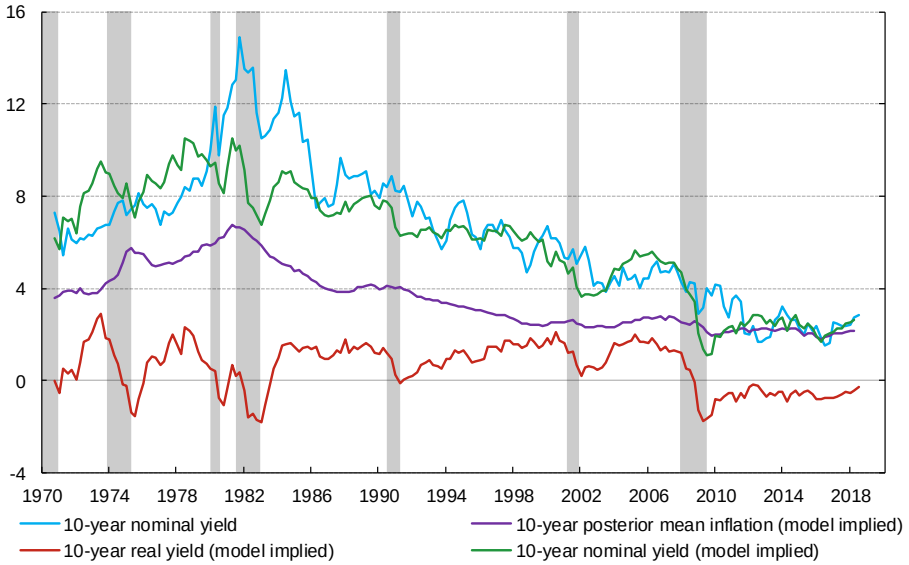
[17]To be consistent with the concept of r^*, neither the short-run effect from $\tilde{x}_{c,t}$ on the real yield, nor the effect of ambiguity on yields, are included for calculation of the model-implied r^*.

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



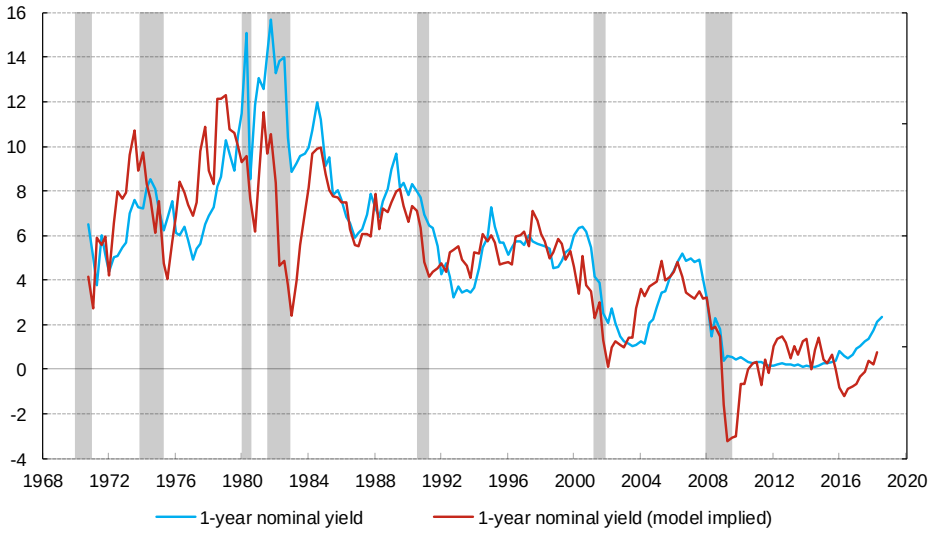
The individual r^*s (quarterly data) are from Bauer and Rudebusch (2017) from 1971:Q4 to 2017:Q2. Three macroeconomic estimates of r^*s are from Laubach and Williams (2003), Lubik and Matthes (2015), and Kiley (2015). The model-implied r^* (quarterly) is from 1968:Q3 to 2018:Q2. The gray bars represent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

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



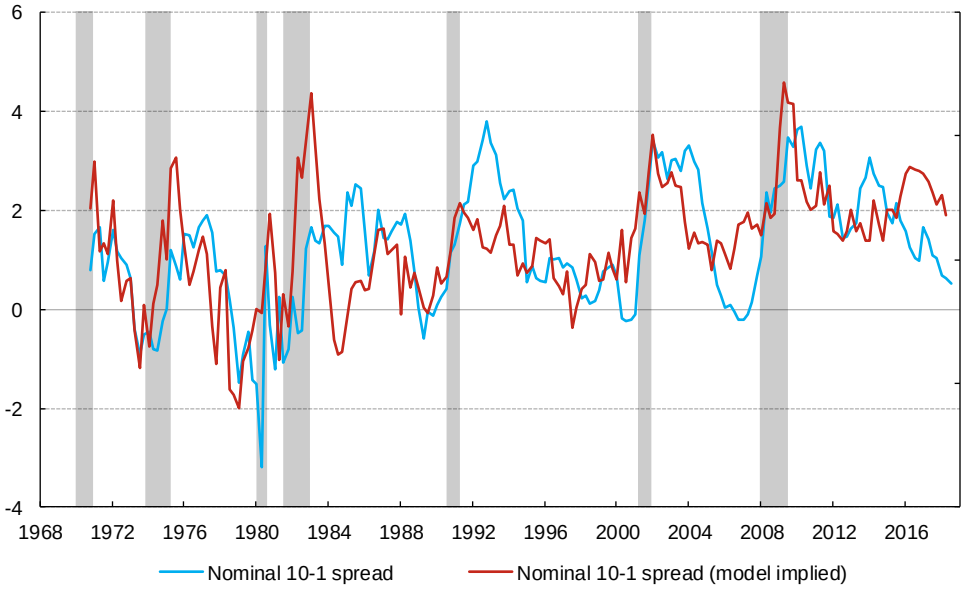
The end-of-quarter 10-year nominal yields are from Gürkaynak et al. (2007) from 1968:Q3 to 2018:Q2. The model-implied 10-year real yield, 10-year nominal yield, and total posterior belief for mean inflation (quarterly data) are from 1968:Q3 to 2018:Q2. The gray bars represent periods of recession defined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

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



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

Figure 14: **Nominal (10-year - 1-year) spread - actual vs. model**



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

5. Conclusion

This paper bridges the gap between empirical and equilibrium yield curve studies by providing an equilibrium interpretation for the trend, cycle, and spread in historical

Treasury bond yields. The representative agent learns the long-run mean and a short-run deviation from mean separately from different components of GDP growth and inflation rates. Instead of using all available data in forming the posterior beliefs, past data gradually loses relevance for learning, either because of fading memory or because it is perceived as irrelevant. The slow-moving trend component in yields is driven by the posteriors for the long-run mean inflation and growth rates, which also moves closely with r_t^* and π_t^* estimations in the literature. The cyclical movements in short-term yields and in the spread between 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 trending in the Treasury yield spread are tightly coupled because both of them are driven by the persistently negative short-run deviations for both inflation and growth. At each point in time, the amount of Knightian uncertainty that the ambiguity-averse agent faces is different for long run versus short run, which gives rise to upward-sloping short-rate expectations under her 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 is to extend the model and incorporate the inflation risk premium in Piazzesi and Schneider (2007), as well as the potentially stochastic volatility in Bansal and Shaliastovich (2013). Given the framework in this paper, some future research avenues seem particularly promising: First, extend the model into a New Keynesian Dynamic Stochastic General Equilibrium (DSGE) model taking into consideration the risk of belief updating. Second, investigate the source of excess bond return predictability and the expectations hypothesis. Third, apply the model to international data. Fourth, investigate the model's implication for long run price-dividend ratio movements.

References

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

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., 2017. Interest Rates Under Falling Stars. Working Paper .

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

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

Campbell, J. Y., Sunderam, A., Viceira, L. M., 2016. Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds. Harvard University Working paper.

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

Champagne, J., Poulin-Bellisle, G., Sekkel, R., 2018. Evaluating the Bank of Canada Staff Economic Projections Using a New Database of Real-Time Data and Forecasts. Staff Working Paper 2018-52, The Bank of Canada .

Christensen, J. H. E., Rudebusch, G. D., 2019. A New Normal for Interest Rates? Evidence from Inflation-Indexed Debt. Working Paper 2017-07, Federal Reserve Bank of San Francisco .

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

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 .

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.

Drechsler, I., 2013. Uncertainty, Time-Varying Fear, and Asset Prices. The Journal of Finance 68, 1843–1889.

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

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 .

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.

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.

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. 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 .

Kiley, M. T., 2015. What Can the Data Tell Us About the Equilibrium Real Interest Rate? FEDS Working Paper No. 2015-077 .

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 .

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.

Lunsford, K. G., West, K. D., 2017. Some Evidence on Secular Drivers of U.S. Safe Real Rates. Working Paper 1723, Federal Reserve Bank of Cleveland .

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

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

Nagel, S., Xu, Z., 2018. Asset pricing with fading memory. Working paper .

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.

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 571 .

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.

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

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" .

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 .

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

Zhao, G., 2018. Ambiguity, Nominal Bond Yields, and Real Bond Yields. Working paper .

Appendix

Since model-III is a comprehensive model and the solutions to model-I and model-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 total real GDP growth and inflation, respectively. Δg_{t+1}^* and π_{t+1}^* are 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 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 she also knows 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

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

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

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

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 latent stationary components

$$Gap_{t+1}^g = x_{c,t+1} + \sigma_c^{gap}\varepsilon_{c,t+1}^{gap}$$
$$Gap_{t+1}^\pi = x_{\pi,t+1} + \sigma_\pi^{gap}\varepsilon_{\pi,t+1}^{gap}$$

and

$$x_{c,t+1} = \rho_c x_{c,t} + \sigma_c^x\varepsilon_{c,t+1}^x$$
$$x_{\pi,t+1} = \rho_\pi x_{\pi,t} + \sigma_\pi^x\varepsilon_{\pi,t+1}^x,$$

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 the parameters, but not $x_{c,t+1}$ and $x_{\pi,t+1}$. She forms 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

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

and

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

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

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

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:

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

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.

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

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

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

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

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 $P_t^{(n)}$, 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:

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

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

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

and

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

Given the small impacts of the covariance terms, we assume all the covariance 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.

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 $er_{n,t+1} = -Cov_t\left(r_{n,t+1}, m^{\$}_{t,t+1}\right) = -Cov_t\left(p^{(n-1)}_{t+1}, m^{\$}_{t,t+1}\right)$. To solve the price and yields for real bonds, we can just replace v' with $v' = (\gamma,\, 0)$.

D. Kalman filter alternative

In this session, 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, 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

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

where $\epsilon_{t+1} \sim N\left(0, \sigma^2_\epsilon\right)$ and $\varsigma_{t+1} \sim N\left(0, \sigma^2_\varsigma\right)$. 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^g_t, of observed data on Δg, the steady-state optimal forecast $\tilde{\mu}_{c,t+1|t}$ is updated as (see, e.g., Edge et al. (2007) and Gilchrist and Saito (2008))

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

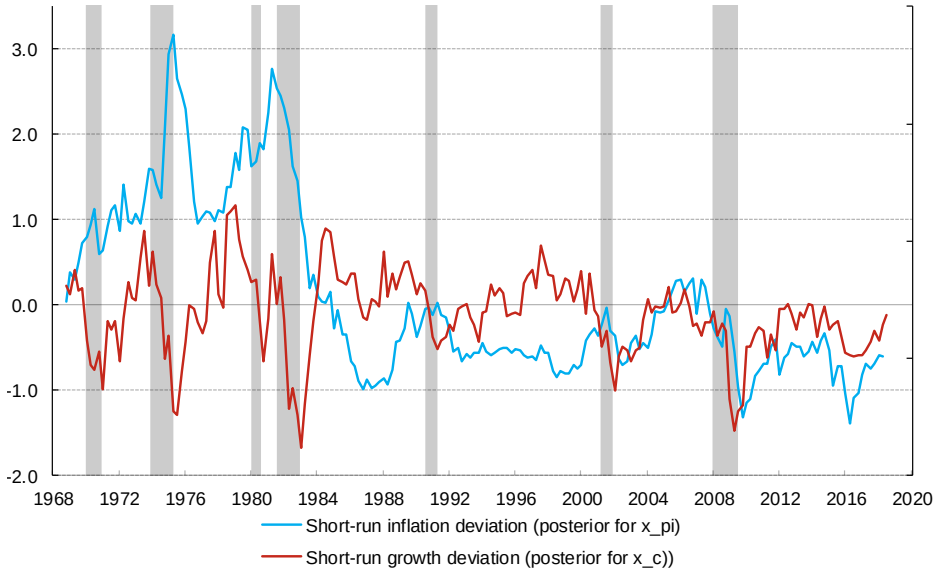
with

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

The steady-state Kalman filtering is equivalent to the 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.

E. Additional figures

Figure A1: Posteriors for 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.