The term structure of inflation expectations[☆]

Mikhail Chernov[a,b], Philippe Mueller[a,*]

[a] *London School of Economics, London, UK*
[b] *CEPR*

Abstract

We use information in the term structure of survey-based forecasts of inflation to estimate a factor hidden in the nominal yield curve. We construct a model that accommodates forecasts over multiple horizons from multiple surveys and Treasury real and nominal yields by allowing for differences between risk-neutral, subjective, and objective probability measures. We establish that model-based inflation expectations are driven by inflation, output, and one latent factor. We find that this factor affects inflation expectations at all horizons but has almost no effect on the nominal yields; that is, the latent factor is hidden. We show that this hidden factor is not related to either current and past inflation or the standard set of macro variables studied in the literature. Consistent with the theoretical property of a hidden factor, our model outperforms a standard macro-finance model in its forecasting of inflation and yields.

JEL classification: G12, G17, E43, C58

Keywords: Affine term structure models, Macro factors, Hidden factors, Survey forecasts

1. Introduction

Inflation expectations play an important role in policy making and in research on asset pricing. The generalized Fisher equation expresses a nominal yield via a sum of the real yield, the expectation about inflation, and the premium on inflation risk. Therefore, when policy makers attempt to determine how monetary policy affects the real economy, they need to break down the nominal yield into its components. This task prompts them to consider the break-even inflation rate (the difference between the nominal and real yields) and survey-based inflation expectations. Likewise, most theories of asset pricing make predictions about the real economy, while, with few exceptions, researchers observe nominal asset prices. Therefore, it is desirable to have a reliable estimate of expected inflation and the premium on inflation risk to be able to test the real theories on nominal data.

Herein, we highlight another reason why academics and practitioners should be interested in inflation expectations and their role in asset pricing. The recent research on bond predictability points to the importance of considering hidden factors when modeling bond yields (Duffee, 2011; and Joslin, Priebsch and Singleton, 2010). A hidden factor does not itself affect the cross section of yields. However, it predicts factors that do. As a result, this hidden factor helps in

[☆] We are grateful to the editor, Bill Schwert, and the anonymous referee for the exceptional feedback. We would also like to thank Dave Backus, Andrea Buraschi, Stephen Cecchetti, Albert Lee Chun, Pierre Collin-Dufresne, Valentina Corradi, Darrell Duffie, Eric Ghysels, Marc Giannoni, Francisco Gomes, Refet Gürkaynak, Scott Joslin, Don Kim, Sharon Kozicki, Lars Lochstoer, Igor Makarov, Stephen Schaefer, Antoinette Schoar, James Stock, Raman Uppal, Andrea Vedolin, Jonathan Wright, Stephen Zeldes, Stan Zin, and participants at the Bank of England, Barclays Global Investors, College of Queen Mary, Columbia Macro and Finance workshops, Copenhagen Business School, the European Central Bank, the 2007 Financial Econometrics Conference at Imperial College, Goethe University Frankfurt, the London Business School Finance workshop, McGill University, MIT Sloan School of Management, the 2007 National Bureau of Economic Research Summer Institute, PIMCO, the 2007 SAFE Conference on New Directions in Term Structure Modeling in Verona, State Street Global Advisors, and Wharton School of the University of Pennsylvania. We are especially grateful to Ruslan Bikbov for many insights into term structure modeling.

[*] Corresponding author: London School of Economics, Department of Finance, Houghton Street, London WC2A 2AE, Phone: +44 20 7955 7012.

Email addresses: m.chernov@lse.ac.uk (Mikhail Chernov), p.mueller@lse.ac.uk (Philippe Mueller)

forecasting bond-related quantities, such as yields or excess returns. It is important to understand the economic intuition behind these hidden factors. Using macro data in conjunction with yields is one way to achieve this. Moreover, such an approach is likely to lead to better statistical properties of the estimated hidden factor than the entirely yields-based approach of Duffee. We argue that survey-based inflation expectations are the macro variables that are helpful in developing our understanding of hidden factors.

The Duffee (2011) intuition is that a factor is hidden in the nominal yield curve if risk premia vary with the level of this factor in the amount equivalent but opposite to the variation in the expected future short nominal rates. Using this intuition as a basis, the Fisher equation shows us that, if a factor is hidden, the joint variation in inflation risk premia and real risk premia associated with this factor offset the joint variation in the expected future short real rates and the expected inflation that is associated with the same factor. Hence, expected future inflation and expected future short real rates should depend on the hidden factor. Therefore, it should be possible to extract the hidden factor from the cross section of expected future short real rates or of expected inflation. Whereas the former is not readily available, the latter can be inferred from survey-based forecasts of inflation.

If the hidden factor is inflation itself, as in Joslin, Priebsch and Singleton (2010), one need not worry about inflation expectations because it is trivial to retrieve the hidden factor from inflation. Nonetheless, the view that inflation is the hidden factor has exactly the same implications for expectations about future inflation as for other possible hidden factors. Therefore, using inflation expectations to learn about the hidden factor constitutes a more general approach. Moreover, such an approach allows for the possibility that information sets of survey forecasters are richer than allowed for by common term structure models.

These observations highlight a role for survey-based expectations that is complementary to the one pointed out by Kim and Orphanides (2012). These authors argue that incorporating survey data into term structure models alleviates the difficulty of statistical inference about persistent variables in short samples. If there is a hidden factor that is not related to observable macro variables, even an infinite data set consisting of yields and macro variables is not going to help a researcher extract the hidden factor. Adding surveys would help in solving this problem.

The challenge that one faces in implementing this idea is that it is not clear whether information about inflation expectations that is embedded in yields is identical to or in conflict with that derived from surveys. That these two possibilities are a real concern is indicated by the fact that the mechanisms that generate these expectations are different. In contrast to consensus survey forecasts, which average the opinions of up to only about 50 participants, the expectations embedded in bond prices are formed by thousands of traders who invest hundreds of millions of dollars. To date, the literature has been concerned with the comparison of different methods of forecasting inflation, but, to the best of our knowledge, no one has reported whether observed yields and survey forecasts could be rationalized within a single model. A joint model of surveys and yields would help to establish whether the two sources of inflation expectations are compatible by detecting whether a common set of factors explains both.

We specify a no-arbitrage macro-finance model that is sufficiently flexible to accommodate inflation, output, real and nominal yields, and survey-based inflation expectations. The yields and forecasts are driven by two observed macro variables (output and inflation) and by latent variables. The joint dynamics of these variables determine inflation expectations for any horizon under the objective probability measure, i.e., a measure determined by the actual factor dynamics of a model. The yields reflect inflation expectations under the risk-neutral probability measure. We connect the behavior of the real and nominal interest rates via inflation, by relying on the transition from the real to the nominal economy. As a result, we obtain the Fisher equation of the nominal interest rate.

Inflation expectations are available from various surveys and at various horizons (see Fig. 1). However, no two surveys are the same in terms of the composition of forecasters, the frequency of the observations, or the forecast horizons. Our model incorporates these characteristics. The key feature of this model is that forecasts from different surveys and at different horizons enter the model in an internally consistent fashion, taking into account, in a reduced form, potentially different information sets or the different objective functions for forecasting. We model the respective

expectations as those of heterogeneous agents. This implies that, for each survey, the expectations are computed using a subjective measure, which reflects an individual's perception of the factor dynamics of a model. We allow that the subjective measure might differ from both objective and risk-neutral measures. The flexibility inherent in this modeling approach allows for state-dependent deviations of the subjective measure-based expectations from the objective measure-based expectations.

[Insert Fig. 1 near here.]

For our empirical analysis, we use a panel of eight yields ranging from three months to ten years, with inflation and gross domestic product (GDP) observed quarterly from 1971 to 2008. We combine these with a total of 19 inflation forecasts from the three surveys depicted in Fig. 1. As a robustness check, we also use our model with added data on Treasury Inflation-Protected Securities (TIPS) yields (from 2003 to 2008).

We need three latent factors (five in total) in the model to accommodate the joint behavior of yields and survey-based expectations. Allowing for the difference between subjective and objective probability measures, we show a rich pattern of term disagreements, i.e., the differences between model-based subjective and objective inflation expectations for different horizons. However, a model that restricts subjective measures so that they coincide with an objective one cannot be distinguished statistically from a richer model.

Once we settle on the preferred model, we check for hidden factors by inspecting the loadings on the factors that drive the yields. In doing so we find a clear delineation between the roles of the latent factors. The first two, level and slope, have almost no effect on model-based inflation expectations. The third factor has no effect on the nominal yield curve but clearly acts as a level factor for inflation expectations. We label this hidden latent factor as a survey factor. Inflation is not hidden in the nominal yield curve. Output is hidden in the model estimated without real yields. We emphasize that we do not assume that the survey factor is hidden; rather, we evaluate whether it is hidden after estimating the model. Models that were estimated without survey-based expectations do not reveal any hidden factors.

Our findings justify our approach, which puts more economic structure on the yields-only approach to hidden factors of Duffee (2011). He finds a largely hidden factor in nominal yields but can say nothing about its relation to anything fundamental. We provide strong evidence that it is information about the expected path of inflation that is not captured by the history of inflation. Investors have learned something about the macroeconomy that shifts their beliefs about future inflation and simultaneously alters their required compensation for bearing risk.

However, our attempts to make this economic link more explicit fail. We check whether one can associate the uncovered hidden factor with some of the observed variables. Regressing the factor on the Cochrane and Piazzesi (2005) risk premium factor and on the Ludvigson and Ng (2009) macro factors produces very low R^2. The survey factor is significant in multivariate regressions that forecast excess returns on bonds with the Cochrane and Piazzesi (2005) and the Ludvigson and Ng (2009) factors as additional variables. These results leave the issue of what economic interpretation to place on the hidden factor open for future research.

One implication of the presence of a hidden factor in a model is that the survey data allow us to better forecast both expected yields and future inflation. We conduct an out-of-sample forecasting exercise by reestimating our model every quarter and compare its forecasts with standard benchmarks. We find that our model performs well in terms of forecasting both inflation and yields. It also strongly outperforms the model estimated without survey-based expectations.

Our empirical results link the paper to the recent theoretical development on hidden factors. The results are also related to scores of empirical papers that have been dedicated to the study of the interactions between inflation and interest rates. This literature is discussed in Appendix A.

2. The model

In this section we develop a model that accommodates inflation, output, real and nominal yields, and survey-based inflation expectations.

2.1. The Fisher equation

To develop intuition, we start with a generic setup in which all state variables are Gaussian processes. We would like to connect the nominal interest rate to expected inflation by considering the behavior of the real and the nominal economies. The log real pricing kernel is

$$m^r_{t+1} = -r_t - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\epsilon^r_{t+1}, \tag{1}$$

where r_t is the real interest rate, ϵ^r is a $N \times 1$ vector of real shocks, and Λ_t is the price of risk.

A typical discrete-time Gaussian approach to joint modeling of nominal and real yield curves assumes that the shocks to real marginal utility are the same shocks that affect inflation. That could be too restrictive. It rules out the possibility that some shocks to inflation are neutral. To accommodate this possibility, we add a shock that, by construction, affects inflation (and therefore the nominal term structure) but not real marginal utility. We let the data tell us whether such independent shocks are important.

Specifically, assume that the inflation rate is

$$\pi^*_{t+1} = e_t + \sigma\epsilon^\pi_{t+1}, \tag{2}$$

where e_t is the expected inflation and the inflation shock ϵ^π is correlated with the real shocks. That is,

$$\mathrm{Cov}_t(\epsilon^r_{t+1}, \epsilon^\pi_{t+1}) = \mathrm{Corr}_t(\epsilon^r_{t+1}, \epsilon^\pi_{t+1}) = \rho,$$

where ρ is a $N \times 1$ vector or, equivalently,

$$\epsilon^\pi_t = \rho' \cdot \epsilon^r_t + \sqrt{1 - ||\rho||^2} \cdot \epsilon^\perp_t, \tag{3}$$

where $\epsilon^\perp_t$ is independent of ϵ^r_t. Allowing $||\rho|| < 1$, we are assuming a more general nominal yield curve behavior than a typical model does.

The log-nominal pricing kernel is related to the log-real pricing kernel via

$$m^n_{t+1} = m^r_{t+1} - \pi^*_{t+1} = -r_t - e_t - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\epsilon^r_{t+1} - \sigma\epsilon^\pi_{t+1}.$$

The price of a one-period nominal bond is

$$P^n_t(1) = E_t\left(\exp(m^n_{t+1})\right).$$

Therefore, the nominal spot rate is

$$n_t = y^n_t(1) = -\log P^n_t(1) = -E_t(m^r_{t+1} - \pi^*_{t+1}) - \frac{1}{2}\mathrm{Var}_t(m^r_{t+1} - \pi^*_{t+1}) = r_t + e_t - \sigma\Lambda_t'\rho - \frac{1}{2}\sigma^2. \tag{4}$$

This equation implies the generalized Fisher decomposition of the nominal rate into the real rate, expected inflation, the inflation risk premium, and the convexity (Jensen) term.

Eqs. (3) and (4) allow us to rewrite the nominal pricing kernel in terms of the nominal interest rate

$$m^n_{t+1} = -n_t - \frac{1}{2}\tilde{\Lambda}_t'\tilde{\Lambda}_t - \tilde{\Lambda}_t'\epsilon_{t+1}, \tag{5}$$

where

$$\tilde{\Lambda}_t = \begin{pmatrix} \Lambda_t + \sigma\rho \\ \sigma\sqrt{1 - \|\rho\|^2} \end{pmatrix} \tag{6}$$

and the vector of shocks ϵ is equal to ϵ^r stacked on top of $\epsilon^\perp$.

The Fisher Eq. (4) shows that a factor affecting expected inflation could be hidden in the nominal yield curve as long as it affects risk premia in the opposite direction of its effect on expected future inflation. Inflation itself could be such a factor, or it could be a linear function of such a factor (partially hidden inflation). We believe that both possibilities are inconsistent with basic implications of workhorse equilibrium models. Specifically, if the real pricing kernel, that is, r_t and Λ_t, is not a function of current inflation, then the Fisher equation implies that inflation will be (partially) hidden only if the expected inflation is not a function of the current inflation. It is difficult to accept, on both theoretical and empirical grounds, that such a situation is possible.

Is it reasonable to assume that the real pricing kernel is not a function of current inflation? This a mild theoretical restriction that is consistent with most of the existing structural models. The finance literature is dominated by endowment economies, in which the real pricing kernel, that is, the real rate and risk premia, is a function of consumption growth and other state variables, such as habits or changing preferences. The real pricing kernel is never determined by inflation in these models. Some studies make additional assumptions about inflation to derive the corresponding nominal rate. For example, Bansal and Shaliastovich (2010), Piazzesi and Schneider (2006), and Wachter (2006) assume an exogenous process for inflation, which, naturally, leads to a nominal rate that is an explicit function of inflation. This approach does not change the fact that the real pricing kernel does not depend on inflation. As a contrasting example, Gallmeyer, Hollifield, Palomino and Zin (2007, 2009) endogenize inflation by assuming that the nominal rate is determined by the Taylor rule. However, by construction, the solution for inflation is such that the real pricing kernel is not a direct function of inflation. New Keyensian production-based models, such as those of Christiano, Eichenbaum and Evans (2005), Palomino (2012), and Rudebusch and Swanson (2008, 2012), among others, are designed specifically to show that monetary policy is relevant for real quantities. This means, in particular, that the real rate should be affected by inflation via the monetary transmission mechanism. However, even in these models, the real rate is not an explicit function of inflation or output. In equilibrium, inflation is an explicit function of exogenous shocks, and the same shocks are driving the spot real rate. Thus, our reduced-form specification is consistent with this class of model as well, because the real rate and inflation are allowed to be subject to the same shocks.

If inflation is not a hidden factor, then what is? Suppose that a factor represents macroeconomic news other than news about current inflation. Examples are shocks to expected future economic activity in China, which show up in futures prices of oil and other commodities; shocks to domestic monetary policy; and shocks to fiscal policy. By construction, this news is orthogonal to current inflation. Therefore, the factor representing this news also is hidden from current inflation. But if we have data on inflation expectations, we can infer the factor and therefore produce more accurate forecasts of future yields, future inflation, and future excess returns.

2.2. A canonical Gaussian model with a factor hidden in the nominal term structure and inflation

A natural question is whether the intuition developed in Subsection 2.1 can be captured by a Gaussian term structure model. The answer is affirmative, as we show in this subsection. We start with the behavior of the Gaussian state under $\mathbb{Q}$. Assume we have a K–dimensional vector y_t with the following dynamics:

$$y_{t+1} = \mu_y^{\mathbb{Q}} + \Phi_y^{\mathbb{Q}} y_t + \Sigma_y^{1/2} \epsilon_{t+1}^y. \tag{7}$$

Following Joslin (2011), we normalize this system so that $\mu_y^{\mathbb{Q}} = 0$, $\Phi_y^{\mathbb{Q}}$ is diagonal, Σ_y is full of ones in its diagonal, and $\Sigma_y^{1/2}$ is a lower triangular Choleski decomposition of Σ_y. This is an alternative to the canonical form introduced in Dai and Singleton (2000). As Joslin points out, this is not the most flexible representation because it assumes that all

eigenvalues of $\Phi_y^{\mathbb{Q}}$ are distinct and real. We give up some of the potential extra flexibility but gain in tractability and clarity of interpretation.

The nominal spot interest rate is assumed to be a linear function of the state

$$n_t = \delta_{n,0} + \delta_n' y_t, \tag{8}$$

where $\delta_n \geq 0$. This representation ensures that all parameters are identified econometrically on the basis of observed nominal yields. As is well known, many other representations are identified econometrically and produce the same set of yields. Duffee (2011) shows that, given the above representation of the state, if the loading $\delta_{n,i}$ in Eq. (8) is equal to zero, then the corresponding factor $y_{i,t}$ is hidden in the yield curve.

In addition, assume that all macroeconomic variables of interest can be represented as a linear function of the same state variables y_t. In particular, inflation has the form

$$\pi_t^* = \delta_{\pi,0} + \delta_\pi' y_t.$$

If $\delta_{\pi,i} = 0$, then the factor $y_{i,t}$ does not affect inflation either. Such a restriction, if it holds in the data, is the key difference from the model considered by Joslin, Priebsch and Singleton (2010).

It is easy to show that risk-neutral expectations of future inflation are not affected by $y_{i,t}$ in this case:

$$E_t^{\mathbb{Q}}(\pi_{t+\tau}^*) = \delta_{\pi,0} + \delta_\pi' E_t^{\mathbb{Q}}(y_{t+\tau}) = \delta_{\pi,0} + \delta_\pi' (\Phi_y^{\mathbb{Q}})^\tau y_t.$$

Because $\Phi_y^{\mathbb{Q}}$ is diagonal, the loading on $y_{i,t}$ is $\delta_{\pi,i}(\Phi_{y,ii}^{\mathbb{Q}})^\tau$. Thus, if $\delta_{\pi,i} = 0$, the factor $y_{i,t}$ is hidden in the risk-neutral expected future inflation. The factor $y_{i,t}$ is also hidden in the risk-neutral expectation of average future inflation over a given horizon:

$$E_t^{\mathbb{Q}}\left(\frac{1}{\tau}\sum_{j=1}^{\tau}\pi_{t+j}^*\right) = \frac{1}{\tau}\sum_{j=1}^{\tau} E_t^{\mathbb{Q}}\left(\delta_{\pi,0} + \delta_\pi'(\Phi_y^{\mathbb{Q}})^j y_t\right) = \delta_{\pi,0} + \frac{\delta_\pi'}{\tau}\left(I - \Phi_y^{\mathbb{Q}}\right)^{-1}\left(I - (\Phi_y^{\mathbb{Q}})^\tau\right)\Phi_y^{\mathbb{Q}} y_t.$$

Again, because $\Phi_y^{\mathbb{Q}}$ is diagonal, it is easy to see that if $\delta_{\pi,i} = 0$, then the factor $y_{i,t}$ is hidden. The risk-neutral expectation of average future inflation is a more useful variable to consider because it is often interpreted as the break-even inflation that is readily observable as the difference between nominal and real yields. Regardless of the variable being used, $y_{i,t}$ affects the objective expectations of inflation and, therefore, can be inferred from them.

The next question is whether such a structure is supported by the empirical evidence. In the sequel we develop a concrete term structure model to address this issue. We do not impose assumptions leading to a factor being hidden in the nominal yield curve and risk-neutral inflation expectations. Our strategy is to impose an economic restriction that the real pricing kernel is not a function of current inflation that precludes inflation being a (partially) hidden factor, and then we let the data determine if there exist factors that are hidden. To be clear, we are imposing the restriction on the pricing kernel not because we have to — the canonical form of the model shows that we do not — but because we believe that this is a natural property that a term structure model should possess.

2.3. The real yield curve

We deviate from the canonical state y_t to incorporate observable macro variables as factors. Therefore, we distinguish between latent factors x_t and macro variables m_t, which jointly form state $z_t = (x_t', m_t')'$. We consider $m_t = (g_t, \pi_t)$, where g_t is the quarterly log-change in GDP (as reported by the US Bureau of Economic Analysis) and π_t is the quarterly log-change in the consumer price index (CPI) (as reported by the US Bureau of Labor Statistics). The canonical state y_t can be obtained from z_t via a rotation.

We model the spot real rate as a function of $N-1$ latent factors x_t:

$$r_t = \delta_{r,0} + \delta_r' x_t.$$

We combine this expression with the process for x_t under the risk-neutral measure $\mathbb{Q}$:

$$x_{t+1} = \mu_x^{\mathbb{Q}} + \Phi_x^{\mathbb{Q}} x_t + \Sigma_x^{1/2} \epsilon_{t+1}^x.$$

We normalize this system so that $\mu_x^{\mathbb{Q}} = 0$, $\Phi_x^{\mathbb{Q}}$ is diagonal, Σ_x is full of ones in its diagonal, and $\Sigma_x^{1/2}$ is a lower triangular Choleski decomposition of Σ_x. Thus, the real interest rate is represented as a linear function of the first $N-1$ principal components of the real yield curve. We assume that N is selected in such a way that the respective principal components explain a major part of the cross sectional variation of the real curve. N is to be determined empirically.

Further, we assume that inflation π^* and changes in CPI π are identical up to random noise uncorrelated with other shocks:

$$\pi_t = \pi_t^* + \omega_t^\pi,\ \omega_t^\pi \sim \mathcal{N}(0, \sigma_{\omega,\pi}^2). \tag{9}$$

Under $\mathbb{Q}$, z_t follows this process:

$$z_{t+1} = \mu_z^{\mathbb{Q}} + \Phi_z^{\mathbb{Q}} z_t + \Sigma_z^{1/2} \epsilon_{t+1}^z, \tag{10}$$

where the first $N-1$ elements of $\mu_z^{\mathbb{Q}}$ are equal to zero and the remaining two are free; the upper left $(N-1)\times(N-1)$ block of $\Phi_z^{\mathbb{Q}}$ is equal to $\Phi_x^{\mathbb{Q}}$, the upper right $(N-1)\times 2$ block is equal to zero, and all other elements are free; $\Sigma_z^{1/2}$ is a lower triangular Choleski decomposition of Σ_z and the upper left $(N-1)\times(N-1)$ block of Σ_z coincides with Σ_x while all other elements of Σ_z are free.

As before, the restrictions imposed on parameters are dictated by econometric identification. Assuming that $g_t = \delta_{g,0} + \delta_g' y_t$, and rotating the vector y_t (for $K = N+1$) into vector z_t, one can see that the restrictions are equivalent to the restrictions imposed in Subsection 2.2:

$$\begin{aligned} z_t &= \Gamma_0 + \Gamma_1 y_t, \\ \Gamma_0 &= (0_{N-1}', \delta_{g,0}, \delta_{\pi,0})', \end{aligned} \tag{11}$$

and

$$\Gamma_1 = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} = \begin{pmatrix} I_{N-1} & 0 \\ \begin{pmatrix} (\delta_g)_{1:N-1}' \\ (\delta_\pi)_{1:N-1}' \end{pmatrix} & \begin{pmatrix} (\delta_g)_{N:N+1}' \\ (\delta_\pi)_{N:N+1}' \end{pmatrix} \end{pmatrix},$$

where 0_{N-1} is a vector consisting of $N-1$ zeros, I_{N-1} is an identity $(N-1)\times(N-1)$ matrix, and subscripts $I:J$ refer to elements of a vector. Thus, the state z_t has the following dynamics under $\mathbb{Q}$:

$$\begin{aligned} z_{t+1} &= \Gamma_0 + \Gamma_1 \left(\mu_y^{\mathbb{Q}} + \Phi_y^{\mathbb{Q}} \Gamma_1^{-1} (z_t - \Gamma_0) + \Sigma_y^{1/2} \epsilon_{t+1}^{\mathbb{Q}} \right) \\ &= \left(\Gamma_0 + \Gamma_1 \mu_y^{\mathbb{Q}} - \Gamma_1 \Phi_y^{\mathbb{Q}} \Gamma_1^{-1} \Gamma_0 \right) + \Gamma_1 \Phi_y^{\mathbb{Q}} \Gamma_1^{-1} z_t + \Gamma_1 \Sigma_y^{1/2} \epsilon_{t+1}^{\mathbb{Q}}, \end{aligned}$$

and

$$\begin{aligned} \Gamma_1^{-1} &= \begin{pmatrix} \Gamma_{11}^{-1} + \Gamma_{11}^{-1}\Gamma_{12}\Delta^{-1}\Gamma_{21}\Gamma_{11}^{-1} & -\Gamma_{11}^{-1}\Gamma_{12}\Delta^{-1} \\ -\Delta^{-1}\Gamma_{21}\Gamma_{11}^{-1} & \Delta^{-1} \end{pmatrix} = \begin{pmatrix} I_{N-1} & 0 \\ -\Gamma_{22}^{-1}\Gamma_{21} & \Gamma_{22}^{-1} \end{pmatrix}, \\ \Delta &= \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12} = \Gamma_{22}. \end{aligned}$$

These expressions combined with identification restrictions on y_t imply the restrictions we imposed on z_t. See Joslin, Le and Singleton also.

This setup is identical to the one introduced in Joslin, Priebsch and Singleton (2010). However, it is applied to the real, not the nominal, yield curve. As a result, the real rate r_t is unaffected by π_t^*. This setup implies also that changes

in GDP and CPI are hidden in the real curve. In other words, components of these two macro variables are orthogonal to the factors x_t and, therefore, to the real rate r_t. It is important to test whether this construct is empirically valid. However, this would be a difficult undertaking at present because of the limited availability of US real yields. Hence, we rely on structural restrictions, as motivated by the theoretical literature.

Also, it is important to note that changes in CPI being hidden in the real curve is an interesting implication of our assumption that r_t is not a function of π_t^*, but nothing more than that. In particular, it has no bearing on whether there exists a factor hidden in the nominal yield curve. It simply ensures that inflation is not hidden, as we show in Subsection 2.4.

We have already motivated our setup with respect to π_t or, equivalently, due to Eq. (9), π_t^* in Subsection 2.2. Making g_t a hidden factor is motivated by the theoretical literature as well, as long as g_t does not differ from real output by more than an independent noise term. Specifically, in endowment economies, the real spot rate is a function of real output only under the counterfactual assumption that consumption coincides with output. In production-based models, the final goods clearing condition implies that consumption is not equal to output, not only in practice but also in theory, because the difference between the two reflects investment and taxes.

We complete the setup of the real economy by specifying risk premia. Because of the recursive identification of shocks (lower-triangular structure of the matrix $\Sigma_z^{1/2}$), the last element of ϵ_t^z affects π_t only (this is the shock $\epsilon^\perp$ from Subsection 2.1). Thus, there are N shocks (first N elements of ϵ_t^z) affecting the real economy. We combine these shocks into a vector ϵ_t^r. Then, we can define the real pricing kernel as in Eq. (1). Further, following Duffee (2002), we parametrize prices of risk:

$$\Lambda_t = \Lambda_0 + \Lambda_1 \begin{pmatrix} x_t \\ g_t \end{pmatrix}. \tag{12}$$

Eq. (12) implies that real risk premia could share shocks with π_t^* but are not themselves functions of π_t^*. Combining this implication with the fact that the real rate r_t is unaffected by π_t^* in our model, we conclude that the real pricing kernel is not a function of inflation. As we argue in Subsection 2.1, this property is consistent with many economic models and also implies that inflation cannot be a factor hidden in the nominal yield curve.

We also have flexibility in deciding whether g_t affects the real risk premia. For example, if the last column of the $N \times N$ matrix Λ_1 is equal to zero, then GDP risk is priced [per Eq. (1)], but GDP growth does not affect prices of risk. If we allow the element $\Lambda_{1,NN}$ to be nonzero, GDP growth affects the price of its own risk, but not the prices of other risks. The theoretical foundations of prices of risk, especially in the context of yield curve models, are still a matter of debate in the literature. Empirically, the importance of flexible risk premia is well known. Therefore, on balance, we choose a flexible specification and allow g to enter the real pricing kernel directly. In general, there is nothing wrong with allowing GDP to enter the real pricing kernel if it is excluded from the real rate. The spot interest rate is the mean, and risk premium is the volatility of the pricing kernel. The mean and volatility do not have to be affected by an identical set of factors. We believe it is reasonable to assume that components of output other than consumption, such as investment or taxes, could affect the volatility of the pricing kernel.

One can construct an equivalent martingale measure corresponding to a real cash account as a numéraire and thereby value the real bonds. The real yields can be expressed as

$$y_t^r(\tau) = \mathcal{A}^r(\tau) + \mathcal{B}^r(\tau)' x_t,$$

where τ is the respective maturity, and $\mathcal{A}^r$ and $\mathcal{B}^r$ solve recursive equations with boundary conditions $\mathcal{A}^r(1) = \delta_{x,0}$ and $\mathcal{B}^r(1) = \delta_x$, respectively (see, e.g., Backus, Foresi and Telmer, 1999). The dependence of real yields on x_t alone is a manifestation of the assumed hidden nature of macro factors.

Similar to the argument made by Joslin, Priebsch and Singleton (2010) in case of nominal bonds, one concludes that parameters associated with the last element of the vector Λ_t cannot be estimated on the basis of real yields alone. This

element is associated with g_t, which is hidden in the real yield curve. We utilize nominal yields in our estimation also, so this concern does not apply.

2.4. The nominal yield curve

The risk-neutral state dynamics [Eq. (10)] and the assumption about the noise in CPI [Eq. (9)] imply that the shock to inflation in Eq. (2) is

$$\sigma\epsilon_t^\pi = e_\ell'\Sigma_z^{1/2}\epsilon_t^z - \omega_t^\pi,$$

where e_ℓ is a vector of zeros with a 1 in the final position. Therefore, Eqs. (5) and (6) imply that the nominal risk premia can be expressed as

$$\tilde{\Lambda}_t = \begin{pmatrix} \Lambda_0 + \left(e_\ell'\Sigma_z^{1/2}\right)'_{1:N} \\ \left(e_\ell'\Sigma_z^{1/2}\right)'_\ell \end{pmatrix} + \begin{pmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{pmatrix} z_t. \tag{13}$$

The inflation noise ω^π is present in the nominal pricing kernel but does not affect prices of risk because it not correlated with any other shocks in the economy.

This form of risk premia and behavior of the state under $\mathbb{Q}$ determines the drift of z_t under the objective measure $\mathbb{P}$:

$$\mu_z + \Phi_z z_t = \mu_z^{\mathbb{Q}} + \Phi_z^{\mathbb{Q}} z_t + \Sigma_z^{1/2}\tilde{\Lambda}_t.$$

Thus, both μ_z and Φ_z are unconstrained with two exceptions: The last element of μ_z is connected to that of $\mu_z^{\mathbb{Q}}$ via the covariance matrix Σ_z and the last column of Φ_z is the same as that of $\Phi_z^{\mathbb{Q}}$. The resulting dynamics of the state are

$$z_{t+1} = \mu_z + \Phi_z z_t + \Sigma_z^{1/2}\epsilon_{t+1}^z. \tag{14}$$

The dynamics of z_t in Eq. (14) imply that the drift of the inflation dynamics in Eq. (2) is

$$e_t = e_\ell'\left(\mu_z + \Phi_z z_t\right).$$

Then, Eq. (4) combined with all the subsequent assumptions implies the nominal rate in our economy:

$$n_t = \delta_{r,0} + \delta_r' x_t + e_\ell'\left(\mu_z + \Phi_z z_t\right) - \left(e_\ell'\Sigma_z^{1/2}\right)_{1:N}\Lambda_t - \frac{1}{2}(e_\ell'\Sigma_z e_\ell + \sigma_{\omega,\pi}^2) = \tilde{\delta}_{n,0} + \tilde{\delta}_n' z_t, \tag{15}$$

where

$$\tilde{\delta}_{n,0} = \delta_{r,0} + e_\ell'\mu_z - \left(e_\ell'\Sigma_z^{1/2}\right)_{1:N}\Lambda_0 - \frac{1}{2}(e_\ell'\Sigma_z e_\ell + \sigma_{\omega,\pi}^2)$$

and

$$\tilde{\delta}_n' = (\delta_r', 0, 0) + e_\ell'\Phi_z - \left(e_\ell'\Sigma_z^{1/2}\right)_{1:N}\Lambda_1.$$

This expression clarifies the relation between the Taylor rules used in many macro-finance papers and the Fisher Eq. (4). Typically, researchers specify the simple Taylor rule in the final expression of Eq. (15) directly. Our derivation shows that a linear relation between the nominal rate and macro variables holds simply because of the inflation-based connection between the real and nominal economies. Further, we show explicitly how inflation and output enter the nominal rate via expected inflation. As a result, these factors are not hidden in the nominal yield curve.

Which factor is hidden? The Fisher equation tells us that it must be the factor affecting inflation expectations. The canonical Gaussian model in Subsection 2.2 shows which restrictions could make it work. We do not impose any such restrictions and check model implications upon estimation.

The standard arguments imply that yields on zero-coupon nominal bonds are linear in the state variables,

$$\begin{aligned} y_t^n(\tau) &= \mathcal{A}^n(\tau) + \mathcal{B}^n(\tau)' z_t \\ &\triangleq \underbrace{\mathcal{A}^{\mathbb{P}}(\tau) + \mathcal{B}^{\mathbb{P}}(\tau)' z_t}_{\text{Short rate expectations}} + \underbrace{\mathcal{A}^{TP}(\tau) + \mathcal{B}^{TP}(\tau)' z_t}_{\text{Term premium}}, \end{aligned} \tag{16}$$

where τ is the respective maturity, and $\mathcal{A}^n$ and $\mathcal{B}^n$ solve recursive equations with boundary conditions $\mathcal{A}^n(1) = \tilde{\delta}_{z,0}$ and $\mathcal{B}^n(1) = \tilde{\delta}_z$, respectively. Because the CPI noise ω^π does not affect prices of risk, its variance affects yields only through $\tilde{\delta}_{z,0}$. Eq. (16) breaks down the yields into expectations of future short rates and term premia. The first of these two components is equal to the usual factor loadings computed under the assumption of zero market prices of risk.

2.5. *Objective inflation expectations*

As we point out in Subsection 2.1, if a factor is hidden both in the nominal yield curve and inflation, one can use data on inflation expectations to infer such a factor. For this reason, we extend our model so that it could be confronted with such data. Survey-based forecasts of inflation are a natural candidate for observations on inflation expectations. If survey participants form their expectations on the basis of the same objective probability measure $\mathbb{P}$ that an econometrician is trying to uncover, then our model has direct implications for such forecasts. We discuss these in this subsection. If, however, survey participants differ from each other in terms of their information or objective function, then their forecasts would deviate from the objective ones. We allow for such a possibility in Subsection 2.6. In the subsequent empirical work, we test which way of modeling the survey-based expectations is more appropriate.

The functional form of the state's drift in Eq. (14) implies that the model-based objective expectation, or $\mathbb{P}$–expectation, at time t of the future state variables τ periods from now is a linear function of the current state variables:

$$E_t(z_{t+\tau}) = \Psi_z^\tau \mu_z + \Phi_z^\tau z_t,$$

where

$$\Psi_z^\tau \triangleq \sum_{k=0}^{\tau-1} \Phi_z^k = (I - \Phi_z)^{-1} \left(I - \Phi_z^\tau\right).$$

In particular, the objective expectations of changes in CPI are represented as

$$E_t(\pi_{t+\tau}) = e_\ell' \left(\Psi_z^\tau \mu_z + \Phi_z^\tau z_t\right).$$

In what follows, we are interested in forecasts of CPI changes over some period τ beginning at a future date $t+s$. This amounts to forecasts of averages of quarterly changes in CPI. The forecasts of averages are denoted and computed as

$$\overline{p}_{t,s}(\tau) \triangleq E_t(\overline{\pi}_{t,s,\tau}) \triangleq E_t\left(\frac{1}{\tau}\sum_{j=1}^{\tau} \pi_{t+s+j}\right).$$

Therefore,

$$\begin{aligned} \overline{p}_{t,s}(\tau) &= \frac{1}{\tau}\sum_{j=1}^{\tau} E_t\left(E_{t+s}\left(\pi_{t+s+j}\right)\right) = \frac{1}{\tau}\sum_{j=1}^{\tau} E_t\left(e_\ell'\left(\Psi_z^j \mu_z + \Phi_z^j z_{t+s}\right)\right) \\ &= \frac{e_\ell'}{\tau}\left(\sum_{j=1}^{\tau} \Psi_z^j + \Psi_z^\tau \Phi_z \Psi_z^s\right)\mu_z + \frac{e_\ell'}{\tau}\Psi_z^\tau \Phi_z^{s+1} z_t \triangleq \mathcal{A}(s,\tau) + \mathcal{B}(s,\tau)' z_t. \end{aligned}$$

2.6. Heterogeneous forecasters

We assume that the world is populated by agents who have heterogeneous expectations of CPI. The heterogeneity might arise from differential information or the use of different loss functions in forming the forecasts. A reduced-form representation of this assumption is an equivalent subjective probability measure $\mathbb{P}^i$, which corresponds to the beliefs of an agent i. This modeling approach could be justified in the framework of such models as Basak (2005), Detemple and Murthy (1994), Dumas, Kurshev and Uppal (2005), Harrison and Kreps (1978), and Scheinkman and Xiong (2003), among others. (The online Appendix provides a simple model that motivates our assumptions.) Alternatively, Patton and Timmermann (2010) show that the use of asymmetric and different loss functions of forecasters implies unbiased forecasts with independent and identically distributed errors under subjective probability measures.

Agents whose expectations differ from the objective predictions do not pose a problem for no-arbitrage pricing. As the cited literature shows, differing beliefs still result in a unique equilibrium price. Agents with different beliefs demand different risk premia, and as a result, all of them use the same risk-neutral measure or, more precisely, the identical projection of the marginal rate of substitution on the space spanned by traded assets. Our model has the same property. In our model, subjective risk premia are captured, in reduced form, by the transformation from $\mathbb{P}^i$ to $\mathbb{Q}$. All agents end up using the same $\mathbb{Q}$ for pricing; hence, bond prices and realized returns are the same.

We parameterize the transformation of measure $\mathbb{P}$ to measure $\mathbb{P}^i$ similarly to the way it is done with the risk-neutral measure:

$$\log \xi^i_{t+1} = -\frac{1}{2}\mathcal{V}^{i\prime}_t \mathcal{V}^i_t - \mathcal{V}^i_t \epsilon_{t+1},$$

where $\mathcal{V}^i_t$ follows the essentially affine specification

$$\mathcal{V}^i_t = \mathcal{V}^i_0 + \mathcal{V}^i_z z_t. \tag{17}$$

Because $\mathbb{P}^i$ is not a pricing measure, $\mathcal{V}^i_t$ do not represent prices of risk. These quantities reflect the deviations of individual expectations from the model-based objective expectations. The resulting dynamics of the state variables corresponding to the individual probability measure $\mathbb{P}^i$ are denoted by

$$z_{t+1} \quad = \quad \mu^i_z + \Phi^i_z z_t + \Sigma^{1/2}_z \epsilon^{\mathbb{P}^i}_{t+1}.$$

We denote the subjective nominal risk premia associated with the mapping between the $\mathbb{P}^i$ and $\mathbb{Q}$ probability measures by $\tilde{\Lambda}^i_t$ and, similarly, the subjective real risk premia by Λ^i_t. Econometric identification does not impose any constraints on $\mathcal{V}^i_t$. Therefore, $\tilde{\Lambda}^i_t$ is unconstrained as compared with $\tilde{\Lambda}_t$ in Eq. (13). It is sensible to assume that inflation forecasters are aware of the basic economic principles, such as the Fisher equation, so their subjective risk premia have the same functional form as Eq. (13). Thus, we impose the corresponding restrictions on $\tilde{\Lambda}^i_t$. As a result, $\Lambda^i_t = \Lambda_t$ and $\tilde{\Lambda}^i_t = \tilde{\Lambda}_t$ whenever $\mathcal{V}^i_t$ is equal to zero, that is, when subjective and objective expectations coincide.

In parallel with the model-based objective expectations of changes in CPI, we can now construct model-based subjective expectations of changes in CPI:

$$E^i_t(\pi_{t+\tau}) = e'_\ell \left(\Psi^{i\tau}_z \mu^i_z + (\Phi^i_z)^\tau z_t\right),$$

where

$$\Psi^{i\tau}_z \triangleq \left(I - \Phi^i_z\right)^{-1}\left(I - (\Phi^i_z)^\tau\right).$$

Model-based subjective expectations of average changes in CPI are denoted and computed as

$$\overline{p}^i_{t,s}(\tau) \triangleq E^i_t(\overline{\pi}_{t,s,\tau}) \triangleq E^i_t\left(\frac{1}{\tau}\sum_{j=1}^{\tau} \pi_{t+s+j}\right).$$

Therefore,

$$
\begin{aligned}
\overline{p}^i_{t,s}(\tau) &= \frac{1}{\tau}\sum_{j=1}^{\tau} E^i_t\left(E^i_{t+s}(\pi_{t+s+j})\right) = \frac{1}{\tau}\sum_{j=1}^{\tau} E^i_t\left(e'_\ell\left(\Psi^{ij}_z\mu^i_z + (\Phi^i_z)^j z_{t+s}\right)\right) \\
&= \frac{e'_\ell}{\tau}\left(\sum_{j=1}^{\tau}\Psi^{ij}_z + \Psi^{i\tau}_z\Phi^i_z\Psi^{is}_z\right)\mu^i_z + \frac{e'_\ell}{\tau}\Psi^{i\tau}_z(\Phi^i_z)^{s+1}z_t \triangleq \mathcal{A}^i(s,\tau) + \mathcal{B}^i(s,\tau)'z_t \\
&\triangleq \underbrace{\mathcal{A}(s,\tau) + \mathcal{B}(s,\tau)'z_t}_{\mathbb{P}\text{–expectations}} + \underbrace{\mathcal{A}^{TDi}(s,\tau) + \mathcal{B}^{TDi}(s,\tau)'z_t}_{\text{Term expectation disagreement}}.
\end{aligned}
$$

In the last line, we decompose the forecasts into the $\mathbb{P}$–expectations of future π and the agent-specific term expectation disagreement TD^i. The former component is equal to the usual factor loadings computed under the assumption of zero $\mathcal{V}^i_t$.

3. Empirical approach

This section outlines all the steps we take to implement and test the model described in Section 2. We also provide technical details of the implementation and how we make inferences via the parametric bootstrap.

3.1. Implementation

In this subsection, we describe the data, how the model is related to the data via the state-space framework, and which versions of our model we estimate.

3.1.1. Data

We use three types of data in this paper. Direct measures of macro variables (changes in the GDP and the CPI) represent two observable state variables in our model. Treasury yields, TIPS yields, and survey forecasts are the observable data that help us understand the model parameters and latent state variables. Using these data, we construct a quarterly panel from 1971 to 2008 with a total of $T = 148$ time series observations. We describe the details of the data in Appendix B.

3.1.2. Connecting real yields to TIPS

In the absence of observation of the true price level, the TIPS are indexed to the CPI. Thus, one could represent a yield on a zero-coupon TIPS via a yield on a zero-coupon nominal Treasury bond and risk-neutral CPI-based inflation of matching horizon:

$$
\begin{aligned}
TIPS_t(\tau) &= Treas_t(\tau) - E^{\mathbb{Q}}_t(\overline{\pi}_{t,0,\tau}) + \frac{\tau}{2}\mathrm{Var}_t(\overline{\pi}_{t,0,\tau}) \\
&= Treas_t(\tau) - E^{\mathbb{Q}}_t(\overline{\pi}^*_{t,0,\tau}) + \frac{\tau}{2}\mathrm{Var}_t(\overline{\pi}^*_{t,0,\tau}) + \frac{1}{2}\sigma^2_{\omega,\pi},
\end{aligned} \tag{18}
$$

where $Treas$ refers to the observed nominal yields, $TIPS$ refers to the observed yields on inflation-linked bonds, and the last term is the convexity adjustment, as in Eqs. (4) and (15). The second equation follows from the assumed relation between the true inflation and changes in the CPI in Eq. (9). Thus, the only difference between the theoretical real yield and TIPS is the convexity component that is related to the variance of the CPI noise $\sigma^2_{\omega,\pi}$. So, the connection between the observed TIPS yields and theoretical real yields is affected by our assumption about $\sigma_{\omega,\pi}$. We discuss this next as a part of the overall estimation of the model.

3.1.3. Observation equations

We estimate our term structure model via maximum likelihood with the Kalman filter, following Duffee and Stanton (2004) and de Jong (2000), among others.[1] We attach measurement errors to all yields and survey forecasts so that the latent factors are not associated with prespecified observables. Duffee (2011) emphasizes the importance of taking measurement errors into account when detecting a hidden factor. However, we assume that the macro variables are observed without error because we have assumed explicitly that our macro factors are the variables reported by US agencies. In other words, we do not associate the macro factors with the underlying equilibrium concepts that could be measured imprecisely.

As a result, we have the following set of measurement equations:

$$\begin{aligned} Treas_t(\tau) &= y_t^n(\tau) + \omega_t^n, \\ TIPS_t(\tau) &= y_t^r(\tau) + \omega_t^r \end{aligned} \tag{19}$$

and

$$Surv_{t,s}^i(\tau) = \overline{p}_{t,s}^i(\tau) + \chi_{t,s}^i(\tau),$$

where $Surv^i$ refers to the observed inflation forecast from survey i.

The errors in measurement of the yields are denoted by ω. We assume the simplest possible structure of the errors—that they are independent and normally distributed with zero mean and standard deviation σ_ω (for each individual element of the vector ω). We need not specify a more flexible error structure because these variables are introduced in addition to the VAR shocks that we considered earlier.

Within each survey, we introduce two types of measurement error. The forecasts with shorter horizons, $\tau \leq 6$ quarters, have many observations and are allowed to have an unrestricted error χ. The longer term forecasts have few observations. We do not want these observations to influence the estimation results unduly, yet we do not want to ignore them completely, because long-term forecasts could be important in providing the answers to our questions. Therefore, for each survey, we restrict the errors in the measurement of the long-term forecasts to be no less than those of the short-term forecasts and those of the yields.[2] Similarly, when we include real yields in the estimation we do not allow the corresponding error ω^r to be less than that of the nominal yields. While such a flexible specification runs the risk of overparameterization, we choose to use it because of the divergent properties of forecasts and yields. As per the description in Appendix B, the forecasts within one survey but at different horizons might be available for different data spans, different quarters, and different frequencies. The real bonds are available only for a short span and the market is less liquid than that of the nominal bonds, so we aim to avoid overfitting the real yields.

We assume that $\sigma_{\omega,\pi} = 0$ in Eq. (9). This choice is motivated by a desire for a more streamlined interpretation of the model. When we discuss how a factor affects the yield curve or whether it is hidden or not, we know that we are talking specifically about changes in the CPI. The costs of doing so are minimal. As we have shown, the noise term affects the nominal yields only via the Jensen convexity term in Eq. (15). Omitting this term could introduce bias into the estimates of some elements of Σ_z. In practice, the convexity term is tiny, in the order of 1 basis point per year for ten-year bonds, so the potential bias is negligible. Further, ω^π appears in our computations because we relate nominal yields to the true real yields in our model, as in Eq. (18). Thus, our simplification justifies Eq. (19). Finally, because of this assumption, we use the terms "inflation" and "changes in the CPI" interchangeably.[3]

[1] Other important estimation strategies applied to term structure models include, but are not limited to, the exact inversion likelihood of Chen and Scott (1993), the closed-form approximate likelihood of Aït-Sahalia and Kimmel (2010), the simulated maximum likelihood of Brandt and He (2006), and the Bayesian Markov chain Monte Carlo of Collin-Dufresne, Goldstein and Jones (2008).

[2] Long-term forecasts are available only for the Livingston Survey and the Survey of Professional Forecasters (SPF).

[3] By assuming $\sigma_{\omega,\pi} = 0$, we ignore potentially interesting testable implications of our model. In particular, one could test whether the true

3.1.4. The impact of survey data on estimation

Kim and Orphanides (2012) study the traditional latent factor models and use survey forecasts of yields in addition to contemporaneous yields. They argue for the use of surveys as a substitute for long samples to achieve precision in their estimates. They show that a short sample period starting in 1990 combined with surveys produces results similar to a long sample period starting in 1965 without surveys. We use a long sample (starting in 1971); hence, our focus is on extracting expectations from whatever data are necessary. Specifically, if there is a hidden factor, even an infinite data set of yields is not going to help a researcher to extract it. Adding survey-based expectations would help in solving this problem.

A related point arises in our model. If there are no hidden factors, using survey forecasts in estimation cannot have a material effect on inflation forecasts if model-based subjective expectations are allowed to diverge from model-based objective expectations in an arbitrary fashion.[4] Intuitively, ignoring the measurement error in yields and the stochastic singularity problem, one can express all the state variables in terms of yields using the objective and risk-neutral parameters. Therefore, the likelihood for the joint yields–surveys data set can be factored into two components. The first component is determined by the yield data and objective and risk-neutral parameters. The second component depends on survey and yield data and subjective parameters. Thus, the maximum likelihood methodology does not use information from surveys to draw inferences about model-based objective expectations. One needs both yields and surveys to infer state variables in the presence of hidden factors. Therefore, surveys are helpful in estimating objective parameters even if objective and subjective expectations diverge. (The online Appendix provides the details.)

3.1.5. Model nomenclature

We estimate eight versions of our model. The full model that uses the survey and nominal yield data is called AS (All data, Subjective expectations). A model that uses the same data, but restricts the subjective measures to coincide with the objective measure, is called AO (All data, Objective expectations). A model that uses only nominal yields for estimation is called NF (No Forecasts). In this implementation, the subjective measures cannot be estimated because they are not identified. We also estimate a model that uses only survey-based forecasts and call it OF (Only Forecasts). In this implementation, the risk-neutral measure cannot be estimated. We consider four- and five-factor versions of these models ($N = 3$ or 4, respectively). To implement various robustness checks, we estimate four more models. AST, AOT and NFT are versions of AS, AO, and NF, respectively, complemented with TIPS yields. OFO (Only Forecasts, Objective expectations) is a version of OF that restricts the subjective measures to coincide with the objective measure.[5]

3.2. Additional considerations

We now review briefly some additional matters that we take into account at the estimation stage.

3.2.1. Missing observations

We determine the main span of our data set according to the availability of the nominal yield data. Thus, we end up with 148 quarters from 1971:3 to 2008:2. However, many forecast observations are missing, because some surveys were available only from a date later than 1971 and some are available less frequently than every quarter. Similarly, real yields data are used starting in 2003. Despite this, missing observations do not cause a problem because they can easily be handled in the Kalman filter framework. We simply do not update, or only partially update, the state vector

inflation is an important determinant of the yield curve. Our view is that such tests are informative if one takes a stand on what the true inflation is, e.g., the first principal component of multiple measures of inflation or a variable estimated from a general equilibrium model. In our setting, a large $\sigma_{\omega,\pi}$ could mean one of two things: an incorrect link between the true inflation and changes in the CPI or a misspecified model of changes in the CPI. It would be difficult to distinguish between these possibilities in our setup. Given that these issues are not central to our focus, we ensure that our model fits changes in the CPI without error.

[4] We are grateful to the referee for this point.

[5] OFO is similar to the time series model of Kozicki and Tinsley (2006).

when observations are missing. This procedure is automatic when one uses a new forecast vector (a new yield vector) that is an old forecast vector $Surv$ (old yield vector $TIPS$) scaled by a matrix that has zeros in place of the missing observations and ones in place of the available ones (see Harvey, 1989).[6]

3.2.2. Latent factor indeterminacy

Dai and Singleton (2000) point out that the restrictions imposed at the estimation stage are not necessarily unique. There are many sets of restrictions, or invariant transformations of the model, such that the yields or inflation expectations are left unchanged. Naturally, when a parameter configuration changes, the respective latent variables change as well by rotating. It is sensible to rotate the factors to identify x with observable variables. We use the invariant affine transformation, which scales factors by a matrix. Appendix A of Dai and Singleton (2000) describes how such a transformation affects the model parameters.

Our first implicit rotation is discussed in the setup of the state space. We assume that our measures of inflation and output are the observable factors. In addition to associating two factors with macro variables, we perform two types of rotation on the remaining $(N-1)$ factors. The first rotation is based on the idea of Bikbov and Chernov (2010) that latent factors are related to macro variables and, therefore, one has to construct projection residuals to extract information that is not in the macro variables. Using the objective measure dynamics [Eq. (14)], we can break down each latent variable x_t into a component explained by changes in GDP and CPI and a residual f that is orthogonal to the entire history $m^t = \{m_t, m_{t-1}, \ldots, m_0\}$ of the two macro variables. We could thus write

$$x_t = \hat{x}(m^t) + f_t$$

and

$$\hat{x}(m^t) = c(\psi) + \sum_{j=0}^{t} c_{t-j}(\psi) m_{t-j} \equiv c(\psi) + c_t(\psi) m_t + c(\psi, L) m_{t-1}, \tag{20}$$

where the matrices c are the functions of the parameters ψ that control the dynamics of the state variables and $c(\cdot, L)$ emphasizes the lag-polynomial structure of the expression. (The online Appendix provides the details.) We define a rotation $\mathcal{O} = Rx_t$, such that the variance-covariance matrix of f becomes diagonal. Such a rotation aids in our interpretation of factors f as shocks. The matrix R is not unique; i.e., the rotation of type $\mathcal{O}$ can generate many sets of orthogonal factors f. Our second proposed rotation, $\mathcal{M}$, can be applied after any of the rotations from class $\mathcal{O}$ and resolves this type of indeterminacy. We define $\mathcal{M} = Ux_t$, where the matrix U is the orthogonal matrix; i.e., $UU' = I$, which preserves the structure of the correlation between the factors. The matrix U is determined by $N - 2 = 1$ or 2 parameters. To determine these parameters, we follow Mueller (2008) and rotate them so that x_1 is interpreted as the factor that is driving the level of the nominal yield curve, and, in the case of a five-factor model, x_2 is interpreted as the factor that is driving the slope, measured as the difference between the ten-year and three-month yields, of the nominal yield curve. These properties are achieved by maximizing the three-month yield's loading on x_1 and the slope's loading on x_2. Appendix C discusses the mechanics of this procedure.

3.2.3. Risk premia and survey-specific disagreements

The elaborate risk-premia specification combined with forecast-specific disagreements leads to concerns about overfitting. We follow Bikbov and Chernov (2010) and augment the standard log-likelihood function, $\mathcal{L}$, with a penalization term that is proportional to the variation of the term premium in Eq. (16):

$$\begin{aligned}\mathcal{L}_p &= \mathcal{L} - \frac{1}{2\sigma_p^2} \sum_{\tau} (\mathcal{A}^{TP}(\tau))^2 + \mathcal{B}^{TP}(\tau)' \cdot \mathrm{Diag}(\mathrm{Var}(z_t)) \cdot \mathcal{B}^{TP}(\tau) \\ &- \frac{1}{2\sigma_p^2} \sum_{i,s,\tau} (\mathcal{A}^{TDi}(s,\tau))^2 + \mathcal{B}^{TDi}(s,\tau)' \cdot \mathrm{Diag}(\mathrm{Var}(z_t)) \cdot \mathcal{B}^{TDi}(s,\tau),\end{aligned}$$

[6] This approach was also used by Kim and Orphanides (2012), Kozicki and Tinsley (2006), Lu and Wu (2005), and Pennacchi (1991).

where σ_p controls the importance of the penalization term and the Diag operator creates a diagonal matrix out of a regular one. If market prices of risk and disagreements are equal to zero, the term premium and term disagreements are equal to zero as well. Therefore, $\mathcal{L}_p$ imposes an extra burden on the model to use the risk premia and survey disagreements as a last resort in fitting the yields. In practice, we take $\sigma_p = 300$, which introduces a modest modification to the original log likelihood. Nonetheless, it helps to stabilize the likelihood and simplifies the search for the global optimum. In particular, this setup helps us to avoid very large values of risk premia.

3.2.4. Optimization

We need to estimate 24 (OF with four factors) to 100 (AS and AST with with five factors) parameters, depending on the specific version of the model used. We have a large cross section of observations, which helps in pinning these parameters down. However, with a time series of 148 observations, a concern remains as to whether or not a global optimum can be found. We use a large and efficient set of starting values to search for the global optimum. The grid search is extremely costly in a multidimensional space and, in practice, limits the extent of the global search. We reduce the computational costs by using Sobol' quasi-random sequences to generate the starting points (see, e.g., Press, Teukovsky, Vetterling and Flannery, 1992). We evaluate the likelihood in two billion Sobol' points and then optimize the likelihood using the best 20 thousand points as starting values. We optimize by alternating between simplex and sequential quadratic programming algorithms, eliminating half of the likelihoods at each stage.

3.2.5. Inference

In our case, it is difficult to make inferences on the basis of the asymptotic distribution of the estimated parameters and test statistics. First, the persistence of the interest rate can affect the asymptotic properties of our tests (Conley, Hansen and Liu, 1997). Second, some of the test statistics that we consider do not have known asymptotic distributions [e.g., root mean squared error (RMSE) ratios]. A nonparametric block bootstrap approach is commonly employed in such cases. However, due to the high persistence of interest rates, the optimal size of a block could be not much less than the size of our whole sample. Therefore, throughout the paper, we use a parametric bootstrap instead. We follow the general principles outlined in Davison and Hinkley (1997) and specific term structure applications, such as Bikbov and Chernov (2011) and Duffee (2007).

Consider the null and alternative models $M_i(\Theta_i)$, $i = 0, A$, where Θ_i denotes the corresponding sets of parameters. We simulate $G = 1000$ artificial samples of size $T = 148$ from the null model $M_0(\hat{\Theta}_0)$ that was estimated on the basis of the observed data set. For each artificial sample $g = 1, \ldots, G$ we compute the estimates $\hat{\Theta}_i^{(g)}$ for both models by maximizing the respective likelihoods in precisely the same way as we do it with actual data. When we want to evaluate the significance of parameters of a certain model, we construct 95% confidence bounds by selecting the 2.5th and 97.5th percentiles from G estimated parameters $\hat{\Theta}_0^{(g)}$. Here, subscript 0, which matches the model used for simulation, emphasizes that the simulation and reestimation are applied to one and the same model. When we want to test one model against the other, we use the estimated parameters to construct a test statistic of interest. For example, in the case of the likelihood ratio test, we compute $2\left(\mathcal{L}_A(\hat{\Theta}_A^{(g)}) - \mathcal{L}_0(\hat{\Theta}_0^{(g)})\right)$, where $\mathcal{L}_i$ is the log-likelihood of model i. The set of G test statistics forms the finite-sample distribution, which we use to compute the p-value for the respective test statistic computed on the basis of the actual data. In the case of a forecasting test, we compute $\mathrm{RMSE}_0(\hat{\Theta}_0^{(g)})/\mathrm{RMSE}_A(\hat{\Theta}_A^{(g)})$ and test whether it is significantly different from one by computing the 95% confidence bound on the basis of the finite-sample distribution formed by all G RMSE ratios.

4. Findings

The main objective of this section is to characterize the information contained in survey forecasts and to establish whether they help to uncover and interpret hidden factors. To do so, we have to establish the preferred specification of our model. Specifically, we test for the appropriate number of factors and whether the distinction between subjective

and objective expectations is warranted in our model. Having settled on a specification, we study its implications for hidden factors, their interpretation, and their impact on the model's forecasting properties.

4.1. Factor structure

In this subsection, we aim to establish whether the yield- and survey-based evidence can be accommodated within a common factor structure. A preliminary principal component analysis suggests that at least four factors are required to explain the joint variation in the macro variables, yields, and survey-based expectations. However, a separate analysis of real bonds implies that three factors might be needed to capture variation in real yields.[7] If correct, this finding implies that at least five factors are required for the joint data set.

This preliminary analysis prompts us to consider four- and five-factor versions of our model ($N = 3$ or 4). Our strategy is to estimate our model using either data excluding survey-based expectations, NF or NFT; data excluding yields, OF or OFO; or all data, AS, AST, AO, or AOT. We compare the models using mean absolute fitting errors and in-sample inflation forecasting performance. We add a number 4 or 5 to the two- or three-letter descriptor of each model to indicate the number of factors involved. For example, AOT5 refers to a five-factor version of the AOT model. Tables 1 and 2 display the respective results.

[Insert Tables 1 and 2 near here.]

In the four-factor case, models NF4 and OF4 produce a good fit to nominal yields and survey-based inflation forecasts, respectively. However, the AS4 and AO4 models (unreported) have some difficulty in matching both. In particular, yield errors increase three- to fivefold. At this point, we should either increase the number of factors in the AS and AO models or drop the survey data. The latter is a reasonable step, if we assume that market prices of bonds reflect all the available information, in particular, survey-based inflation expectations.

However, the difficulty in fitting both data sources together indicates that survey-based expectations could posses some incremental information. To investigate this possibility, we compute in-sample inflation forecasts from the NF4 and OF(O)4 models. As mentioned in Subsection 3.2, all our inferences are made via the parametric bootstrap. Here, the null model is NF4 and the alternative is either OF4 or OFO4.[8] We use RMSE ratios of inflation forecasts as our test statistics.

Table 2 shows that OF(O)-based forecasts dominate or are similar to the NF-based ones, depending on the forecasting horizon. To ensure that there is survey-specific information that is critical for forecasting inflation, we also evaluate NFT4. The rationale is that TIPS, combined with nominal yields, could help in pinning down expected inflation without using survey-based expectations. We implement the tests in exactly the same way as before. Compared with NF4, this model has a slightly worse, but similar, fit for nominal yields and it does not perform as well as OFO4 in inflation forecasting. This evidence suggests that we should make a further attempt to accommodate survey-based information in our model. Therefore, we proceed with estimating the five-factor models.

The yield and survey-based expectation fitting errors in AO(T)5 and AS(T)5 are similar to those in NF(T)4 and OF4, respectively. Thus, adding a fifth factor resolves the tension in fitting the two sources of data. We formally test four-factor models (null) against the five-factor (alternative) models via the likelihood ratio test. The likelihood ratio tests always select five-factor versions of the model when both sources of data are used. The likelihood ratio tests select four-factor models when either only yields or only surveys are involved. This result raises the need to understand why bond prices do not reflect all the information that pertains to inflation forecasts.

[7] This analysis was performed using UK inflation-linked bonds because of their longer history and their wider range of available maturities.

[8] If one of the OF models is a null hypothesis, we have to reestimate NF using the samples simulated from OF. However, OF is silent about yields. Therefore, we selected NF as the null model.

4.2. Expectation structure

Having established an appropriate factor structure, we have to understand the best way of accommodating the survey-based expectations in our modeling framework. Specifically, we need to test whether the extra flexibility afforded by the distinction between the objective and subjective model-based expectations is needed to capture the joint dynamics of yields, macro variables, and survey-based expectations. Here, we are not interested in testing whether the observed expectations from different surveys are the same or if one of the survey-based expectations is the same as a model-based objective expectation. If we have two sets of expectations that do not line up point by point, we reject the null hypothesis that they must be the same if we assume zero measurement error. However, such a rejection would not be informative about the properties of our models.

We want to test whether a simple model [AO(T)], where model-based expectations are the same across all surveys and are equal to the objective model-based expectations, is no worse than the corresponding AS(T) model in its ability to replicate the salient features of our full data set. Formally, the null hypothesis is that $\mathcal{V}_0^i$ and $\mathcal{V}_z^i$ in Eq. (17) are equal to zero for all i.

To illustrate the intuition behind the test, we plot the term structure of term disagreements in Fig. 2. We consider the term disagreements about inflation expectations formed today over the horizons matching the maturity of the bonds in our sample; that is,

$$TD_t^i(\tau) = E_t^i(\pi_{t+\tau}) - E_t(\pi_{t+\tau}),\ \tau = 1, 2, 4, 8, 12, 20, 28, 40 \text{ quarters}.$$

We use the AST5 model for the computations. We report the unconditional term structure, which measures the average term disagreement for each horizon, and the time series of conditional term disagreements for the horizons of one ($\tau = 4$) and ten ($\tau = 40$) years.

[Insert Fig. 2 near here.]

The first panel shows the unconditional term structure of the disagreements. All the surveys exhibit distinct and intriguing patterns. However, the magnitudes are very small. The largest absolute value of a disagreement is 0.3% for the short-term Livingston Survey. As to the conditional disagreements in Fig. 2, at certain times the disagreement could be large, up to 4% in the case of ten-year expectations.

However, these figures do not account for the parameter sampling uncertainty. Moreover, they show deviations between objective and subjective expectations within one and the same model (AST5). We are interested in establishing whether the same set of survey-based expectations can be accommodated in a model that does not make a distinction between the subjective and objective probability measures. Finally, the figures do not characterize the full picture because our models are designed to capture the joint dynamics of yields, macro variables, and survey-based expectations. Therefore, we cannot study the models' implications for expectations without regard for other variables. A likelihood ratio (LR) test is a natural framework that automatically addresses all of these points.

Under the null hypothesis that subjective and objective probability measures are the same, that is, $(\mathcal{V}_0^i, \mathcal{V}_z^i)$ are zero for all i, the likelihood is constructed from the null model [AO(T)] using all the data. Under the alternative, the likelihood is constructed from all the data taking into account that the unrestricted model [AS(T)] allows for a different subjective measure for each of the surveys. As before, the inference is implemented via the parametric bootstrap. The LR test statistics are equal to 2.19 (AO versus AS) and 2.14 (AOT versus AST), both of which are insignificant with both p-values equal to 0.6. Thus, there is no statistical difference between the two models.

An omnibus test, such as LR, might fail to reject the null hypothesis, but there could still be an important dimension where the models differ. This is a common issue with all likelihood-based estimation. One uses full information to estimate and test the model, but then has to come up with additional less formal diagnostics to see how the model fares vis-a-vis data along the dimensions that matter for the research topic. The reason we introduce subjective expectations

into our models is to give them more flexibility in capturing expectations. Thus, it is natural to test whether objective and subjective measures produce the same expectations.

A two-sample t-test allows us to test such a hypothesis for a pair of expectations: the objective versus the subjective from survey i. However, because we have multiple groups, the question of interest is whether there is at least one average expectation that is significantly different from another average expectation. One-way analysis of variance (ANOVA) is designed to address such questions. Formally, the null hypothesis [under the AO(T) model] is that all average subjective expectations are the same and are equal to the average objective expectation against the alternative [under the AS(T) model] that at least one of the subjective expectations is different. The F-test employed in ANOVA requires equality of variances across the different groups. We use the Levene (1960) test, which does not require normality of the data, to establish whether the variances are the same. The test statistic has an F-distribution in large samples under the null hypothesis as well.[9]

Both tests assume independence between the groups and in time (in fact, there is no time dimension in the classical ANOVA). This is why we again rely on parametric bootstrapping instead of the F-distribution to generate p-values for both tests. We re-use the AS(T) model estimated on the basis of one thousand artificial histories simulated from AO(T). Along each history, we compute the ANOVA F-statistic, which is equal to the ratio of variation between the groups to variation within groups. In the context of our models, we have three groups of subjective expectations for $i = 1, 2, 3$. We compute the Levene test statistic in a similar fashion. These statistics form the finite-sample distribution, which we compare with the respective F-statistic computed on the basis of the actual data.

Table 3 reports the results. The Levene test fails to reject the null hypothesis of equal variances across all forecasting horizons. Therefore, it is appropriate to use ANOVA to test for the equality of expectations. Again, we fail to reject the null hypothesis of equal average expectations across all the forecasting horizons. We conclude that modeling survey-based expectations does not require a distinction between objective and subjective probability measures in our case.

[Insert Table 3 near here.]

4.3. The term structure of inflation expectations

We conclude that we should be using the AO-type models, by imposing $\mathbb{P}^i = \mathbb{P}$, to produce a model-based measure of inflation expectations. Thus, in what follows, we focus on this class of models. Fig. 3 shows the time series of the model-based inflation expectations at multiple horizons. These expectations are computed from AOT5. In contrast to the survey forecasts shown in Fig. 1, these expectations can be computed each period for any horizon.

[Insert Fig. 3 near here.]

The term structure effects are pronounced. The inflation curve becomes inverted in 1973, just before the recession, and continues to be inverted until early 1982. This period coincides with the unstable period of monetary policy during the Arthur Burns and G. William Miller chairmanships of the US Federal Reserve System and the monetary policy experiment under Paul A. Volcker's chairmanship. The curve became inverted again during the early part of Alan Greenspan's tenure from 1987 to 1991. Afterward, it had a normal, nearly flat, shape. The most recent period, starting in 2005, is characterized by relatively volatile short-term inflation expectations. As a result, the slope of the inflation curve switches sign frequently.

[9] We do not provide explicit expressions for the ANOVA and Levene test statistics in the interest of saving space. These are sufficiently standard statistics, which are available in statistics textbooks and implemented in Matlab.

4.4. *The information in survey-based forecasts*

Having selected the appropriate model, we study its implications vis-à-vis models studied in the past. The Fisher equation is an important source of our intuition about how to look for a hidden factor. Therefore, we describe what the model implies for the decomposition of nominal yields on the basis of this equation. Next, we determine whether we can uncover and characterize a hidden factor. We do find an approximately hidden factor and we attempt to provide an economic interpretation of it. We conclude the section by describing the out-of-sample forecasting performance of the model.

4.4.1. *Inflation risk premia*

To get a sense of how survey-based expectations contribute to our models, we describe the models' implication for the Fisher Eq. (4). Fig. 4 displays the time series of the decomposition of ten-year nominal yields into real yields, inflation expectation, and inflation risk premium using the equation. Table 4 reports summary statistics associated with this decomposition. We have settled on the AOT5 model, following the earlier analysis. However, we also report the decomposition for the NF4, NFT4, and AO5 models. We do so to highlight the role of survey-based inflation expectations and TIPS in model-based measures of the inflation premium.

[Insert Fig. 4 and Table 4 near here.]

The following five broad lessons emerge. First, the use of TIPS for estimation leads to more volatile estimates of real yields. We observe a doubling of the time series volatility of the ten-year real yield as we move from NF4 to NFT4 and nearly a quadrupling as we move from AO5 to AOT5. Second, the presence of survey-based inflation expectations leads to similar estimates of model-based objective expectations, regardless of whether or not TIPS are used for estimation. Third, in the absence of the survey data, adding TIPS leads to a trade-off between the levels of the real rate and the inflation premium. Adding TIPS data to the NF model leads to a 140% increase in the average real yield and to a 65% decline in the average inflation premium. In fact, the inflation premium level drops so much that the premium becomes negative after 1995. Hördahl and Tristani (2007) provide a thorough discussion of the debate regarding the correct sign of the inflation premium. Fourth, in the presence of the survey data, adding TIPS leads to a trade-off between the volatilities of the real rate and the inflation premium. Adding TIPS data to the AO model leads to a 280% increase in the volatility of the real rate and to a 55% decline in the volatility of the inflation premium. Fifth, for three out of the four models, the inflation premium gradually declines from the late 1980s onward and reaches zero by the late 1990s. The only exception is the NFT model, in which the premium becomes negative in the mid-1990s. This is a direct result of the low overall level of the premium in this model.

4.4.2. *Hidden factor*

We start our analysis of hidden factors by characterizing how the state variables affect yields in our models. To save space, we show factor loadings for representative models only. Figs. 5 and 6 display the factor loadings for NF4 and NFT4, and AO5 and AOT5, respectively. Fig. 5 shows clearly that the same factors affect both inflation expectations and nominal yields in NF models; that is, there are no hidden factors. Fig. 6 shows that, in the AO(T) models, the level factor x_1 has almost no effect on model-based inflation expectations. In contrast, factor x_3 has no effect on the nominal yield curve but clearly acts as a level factor for inflation expectations (and real yields). We label x_3 as a survey factor because it cannot be identified without the survey data. The fact that the survey factor x_3 appears to be hidden in the nominal yield curve is not an assumption, but an empirical implication of the models.

[Insert Figs. 5 and 6 near here.]

One notable difference between AO and AOT is that output g is hidden in the nominal curve in AO, but not in AOT. This result has a potential bearing on the finding that output has predictive power in excess of yields (e.g.,

Ludvigson and Ng, 2009). If TIPS data were available for a long time span, the addition of such data to the regressions implemented in these studies could have mitigated the need to use output as an additional predictor. More generally, it appears that whether or not a factor is hidden depends on the data set used in the estimation.

We find, on the basis of the Wald test, that loadings on x_3 that correspond to the spot interest rate and to maturities beyond five years are statistically significantly different from zero. Nonetheless, the magnitude of the factor's contribution to the yield curve is smaller than the measurement errors. Therefore, we conclude that the factor x_3 is approximately hidden in the nominal yield curve. Our findings connect with the Duffee (2011) conjecture that inflation expectations might be hidden in the nominal yield curve. He estimates a five-latent-factor model and finds evidence that the fifth factor is approximately hidden in the yield curve. Duffee shows a nonzero correlation between this factor and survey-based inflation expectations. The difference is that we find that inflation expectations are partially hidden because x_3 is only one of the factors affecting them.

Inflation is not a hidden factor because of our assumption that it does affect the real pricing kernel. Whereas inflation is not a hidden factor, it could be partially hidden. We can check directly whether there is a common source of variation driving x_3 and macro factors by representing yields, survey-based forecasts, and macro variables as linear functions of common unobservable factors y as in Eq. (7). If x_3 and inflation load on the same hidden factor y_i, then x_3 can be backed out of inflation; that is, inflation is partially hidden.

The task of switching into the canonical representation of Eq. (7) is accomplished by a simple factor rotation on the basis of Eq. (11). We find that, in all models except for AO5, none of the latent factors y is hidden. This finding contrasts with that of Duffee (2011) and could be explained by the fact that we are using macro variables in addition to yields or that we use yields with maturities beyond five years. Nonetheless, in some of the models (AOT5, AS5, AST5) a linear combination of y's, a.k.a. the factor x_3, is approximately hidden. The model AO5 is an exception in that one of the factors, y, is hidden. Inflation and break-even inflation do not load on this factor in this model. Thus, AO5 is consistent with the canonical form described in Subsection 2.2. In summary, the evidence supports our initial argument that inflation is unlikely to be partially hidden.

4.4.3. Economic interpretation of the hidden factor

Subsection 2.1 provides intuition on how a factor can be hidden in the nominal yield curve and inflation and yet affect expectations of future inflation. In this subsection, we investigate whether one can assign a specific economic interpretation to the factor x_3. Here, we are not shy about data snooping because we want to see if we can find some observable variable that could be related to x_3, even if spuriously. We do not find any. In this regard, our conclusions are similar to those of Duffee (2011) with respect to his factor. This finding leads us to conclude that survey forecasters rely on a rich and potentially varying information sets when producing their expectations.

We start our analysis by projecting x_3 on the macro variables that we use in our model via Eq. (20). The projection residual, which we call f_3, retains the key properties of x_3: It is hidden in the nominal yield curve and acts as a level factor for the inflation expectations. We should be able to relate f_3 to observable macro variables if π_t and g_t are imperfect measures of inflation and real output, respectively. Thus, other variables could be related to inflation and output that simultaneously explain variation in f_3. Alternatively, f_3 could be related to macroeconomic variables other than the ones related to inflation and real activity. To address both possibilities, we should regress f_3 on a comprehensive set of observable variables.

Ludvigson and Ng (2009) suggest such a set of variables. Specifically, they construct eight common factors from 132 measures of economic activity. We regress f_3 on all these factors and find that none of them is significant with a joint adjusted R^2 of 9% (Table 5, Panel A). Dropping some of these factors does not help with the R^2 and does not help with making some of the remaining factors significant.

[Insert Table 5 near here.]

Given these results, we try to take a slightly different angle and investigate how f_3 affects forecasts of bond risk premia. The hope is that observing the interaction with other variables would help us to gain insight into the nature of f_3. We start by checking how the factor affects expected excess returns in our model by plotting the loadings on the factors for various expectation horizons (see Fig. 7). We intentionally include both AO(T) and AS(T) models to illustrate that the implications of these models for the role of x_3 and, therefore, f_3 are robust. An increase of 1 standard deviation in the factor (and, by a sign convention, an increase of 1 standard deviation in expectations about inflation) results in a decline of almost 1 standard deviation in expected excess returns across all maturities.

[Insert Fig. 7 near here.]

Next, we check how x_3 or, more precisely, f_3 affects the forecasting of realized excess returns. We run the variations of the Cochrane and Piazzesi (2005) regression, which include combinations of variables. Table 5 reports the results. We start by reproducing the Cochrane and Piazzesi (2005) regression in our sample (Panel B). The CP factor is highly significant with adjusted R^2 ranging from 29% to 35%, depending on a bond's maturity. Then we add the Ludvigson and Ng (2009) factors (Panel C). We use all eight of them, but the table reports the loadings for the LN factors 1, 6, 7, and 8 only (because they show up in subsequent regressions). The R^2 increases to 42–47%, but not all of the LN factors are significant.

Alternatively, we can check what our model implies for excess returns. We construct the projection residuals for the remaining two latent factors x_1 and x_2 and regress excess returns on f_1, f_2, and f_3 (Panel D). The first factor is significant at short maturities only. The other two are significant for all excess returns and the adjusted R^2 is between 33% and 36%, on par with the CP factor. So, do f_2 and f_3 explain the CP factor? We add the CP factor to the regression and find that it is significant at all bond maturities and drives out f_2 (Panel E). Thus, f_3 contains information that is complimentary to CP.

Fig. 8 displays the time series of the standardized CP and $-f_3$ factors. We change the sign of the hidden factor to make it clear that the two factors line up only during the period of the monetary experiment. Both factors have a strong business-cycle component, but they differ in higher frequency fluctuations. What could account for these differences?

[Insert Fig. 8 near here.]

We know that LN factors do not replace CP in these regressions. It might be that these factors capture some of the higher-frequency movements in f_3. To determine whether this is so, we go through a comprehensive factor search in which we gradually eliminate the insignificant ones. We end up with the regression reported in Panel F, where the surviving factors are CP; LN factors 1, 6, 7, and 8; f_1; and f_3. It appears that survey expectations contain information about the bond risk premium that is complimentary to that contained in both the CP and LN factors.

4.4.4. Forecasting

Table 6 shows the out-of-sample results for the forecasting of inflation. Ang, Bekaert and Wei (2007) find that raw surveys dominate term-structure models for inflation forecasts that are made one year in advance. Our results should complement their findings in two dimensions. First, we provide evidence for a term structure of inflation forecasts. Second, we evaluate the performance of term-structure models that are estimated using information from the surveys.

[Insert Table 6 near here.]

We use RMSE ratios as our test statistic. The bootstrap procedure that we use is complicated in this case, because the model for survey-based forecasts is not known. Each survey participant likely uses her own model to produce an individual forecast and then these forecasts are aggregated into the consensus forecast. If we knew all of these models, in theory, we could replicate the entire procedure under the null of the AO or NF model. However, we do not know what each of the forecasters is doing.

We can treat forecasts from a model and from the surveys as raw inputs (ignoring where they are coming from) and use various asymptotic tests to evaluate the difference (e.g., Diebold and Mariano, 1995). Perhaps, the methodology that is the closest to our setup is an asymptotic test of forecasting methods developed in Giacomini and White (2006). However, to achieve asymptotically nonvanishing estimation uncertainty, these authors assume a fixed rolling window for model reestimation, whereas we use an expanding window. Alternatively, we could have used nonparametric block bootstrap. We express our reservations regarding both asymptotic and nonparametric bootstrap tests in Subsection 3.2.

We use a different procedure that has its own shortcomings but has the virtue of being in line with our overall testing strategy. We simulate one thousand artificial samples from the null model (NF or AO) and reestimate the null model on the basis of each sample. Then we construct the ratio $\text{RMSE}_0(\hat{\Theta}_0^{(g)})/\text{RMSE}_{Surv}$. Thus, we implicitly assume that the survey-based forecasts are data, that is, they are not generated by aggregating multiple models, but by the economic data–generating process directly. Our hope is that, because of our focus on the first two moments and not the entire distribution, this implicit assumption is not going to affect the inference too much. The lack of extra variation in the test statistic results in tighter confidence bounds and biases the results toward Type I error (over-rejection).

Despite these caveats, the reported RMSE ratios send a clear message. Panel A shows that AO performs on a par with raw surveys. The ratio of AO to surveys is close to one, and the difference is statistically insignificant. The same cannot be said about the NF model. The ratios are much higher than the one in the data and the difference from one is statistically significant. To be supremely conservative, given the heightened probability of Type I error, the conclusion would be that NF does not outperform the surveys.

We implement a direct RMSE-based comparison of AO and NF to clarify the difference in performance of the two models. Panel B displays the RMSE ratio test of NF (null) against AO (alternative) via parametric bootstrap. NF is overwhelmingly rejected. The improvement associated with using AO ranges from 23% to 55%, depending on the sample and horizon.

That essentially no difference exists between the results of the raw surveys and those of a complicated AO model could be disappointing. It is important to remember that a big limitation of raw surveys is that an end user cannot select the frequency of forecasts and the forecast horizon. We present and defend a model that can produce survey-quality forecasts, but at any time and at any horizon. This conclusion is particularly striking in the light of the findings in Ang, Bekaert and Wei (2007), where, after searching through a list of 30 different models, they could not find any such model even for one forecasting horizon.

Table 7 displays out-of-sample results for the forecasting of yields. We use the random walk model (RW) as an alternative model and report all the results in the form of the RMSE ratios of a null model (AO or NF) to RW. We use bootstrap once again. In this case, estimating the alternative model along the artificial samples simulated from the null model is particularly easy as there are no parameters to estimate for RW. We simply use yields simulated from the relevant null model.

[Insert Table 7 near here.]

To the best of our knowledge, this is the first analysis of this kind on such a scale. Typically, authors leave about five years for out-of-sample exercises. The parameters for out-of-sample analysis are either estimated from an earlier long subsample or reestimated for each period over the short subsample. The range of forecasting horizons is typically narrow, from the next period to one year ahead. Finally, formal statistical analysis is not conducted, that is, the numerical values of RMSE are simply compared with each other.

On the basis of point values of RMSE ratios, we see that RW dominates or is indistinguishable from NF. The AO model dominates RW at longer forecasting horizons (eight and 20 quarters ahead). We conclude that the incorporation of forecasts of inflation is helpful in the forecasting of yields. However, there remains room for improvement, because the results for yield forecasts are not as clear as for inflation forecasts. The analysis provided herein suggests that survey forecasts are very useful, and the incorporation of forecasts of real activity or yields should help when forecasting yields and can reduce the statistical uncertainty.

5. Conclusion

We have built a dynamic macro-finance model that incorporates the behavior of inflation, real activity, nominal yields, and survey-based forecasts of inflation. The model prohibits arbitrage opportunities and allows for the heterogeneity of survey forecasters via a subjective probability measure. We find that observed yields and survey forecasts are internally consistent with each other, that is, their joint behavior can be accommodated within a term structure model without expanding the number of state variables. Moreover, both yields and forecasts are important for producing realistic expectations about future inflation and yields. Information other than yields is required because a factor that drives model-based inflation expectations has almost no effect on the cross section of yields; that is, it is hidden.

We make a case using both theoretical and empirical analysis that the hidden factor is not inflation itself. We present a canonical Gaussian term structure model that suggests a mechanism achieving the same result. However, such a model still leaves an important question: What forces drive survey-based expectations about future inflation that are correlated with bond yields, yet do not affect them cross sectionally and do not affect current inflation? Our attempts to relate this factor to other macro variables led to nothing. In this regard, our conclusions are consistent with those of Duffee (2011). When forecasting bond risk premia, our hidden factor is not driven out by the Cochrane and Piazzesi (2005) factor or the Ludvigson and Ng (2009) macro factors.

Appendix A. Selective empirical literature review

Scores of empirical papers have been dedicated to the study of the interactions between inflation and interest rates. Discussing all of these in this brief review is not possible. Two main strands of the literature that study this subject are those on term structure and macro forecasting.

The empirical side of the literature on no-arbitrage term structure explores the behavior of state variables and risk premia. The main focus is to explain and summarize parsimoniously the relations that are of interest, and for this reason most of the analysis is conducted in-sample. In addition to the data on yields, various authors utilize either inflation itself (the no-arbitrage macro literature beginning with Ang and Piazzesi, 2003) or data that reflects inflation expectations, such as inflation-indexed bonds (Evans, 1998), or survey forecasts of inflation (Chun, 2011; and Pennacchi, 1991). The typical research questions that arise are: Does inflation help to explain the yields? (See Ang and Piazzesi, 2003; Bikbov and Chernov, 2010; and Duffee, 2006.) Does inflation help to explain the risk premia? (See Ang, Dong and Piazzesi, 2004; Bikbov and Chernov, 2010; and Duffee, 2007.) How do real interest rates behave? (See Ang, Bekaert and Wei, 2008; Evans, 1998; and Pennacchi, 1991.) What is the size and the nature of the variation in the inflation premium? (See Ang, Bekaert and Wei, 2008 and Evans, 1998; Hördahl and Tristani, 2007.)

Part of the literature on macroeconomic forecasting focuses on the prediction of inflation from the prices of financial assets, notably the short rate and the term spread, in the regression framework (see Stock and Watson, 2003, for a comprehensive survey). The natural motivation for such an approach is that, because of their forward-looking nature, yields should serve as good predictors of macroeconomic activity. In contrast to the literature on term structure, much of this work focuses on out-of-sample performance. Conceptually, the approach suffers from the drawback that, as any no-arbitrage model would imply, the yields simultaneously reflect the market's expectations of the state variables, such as inflation, and risk premia. A formidable body of evidence shows that risk premia are time-varying. As a result, disentangling the yield components in a regression framework is virtually impossible. Depending on the variability of the risk premia, yields might introduce a fair degree of noise, which would make the inference unreliable.

A separate strand of the literature on forecasting investigates the predictive ability of survey forecasts. As with yields, the analysis is performed in a regression framework. The analysis focuses on the forecasts' rationality and efficiency and typically concentrates on a one-year horizon. The key finding, based on considering roughly pre- and post-Volcker subsamples, is that inflation is underestimated when it is high and vice versa (see Thomas, 1999, among many others). However, a more refined conditional analysis is not feasible because of the inherently unconditional nature of the regression-based analysis. In contrast, Ghysels and Wright (2009) use high-frequency financial variables to forecast the survey forecasts.

Finally, Ang, Bekaert and Wei (2007) combine the views contained in the reviewed literature by running a horse race between the different methods of forecasting one-year inflation.[10] They use the different forecasts separately and also combine them via various weighted-averaging schemes. The authors find that surveys forecast inflation better than other models and approaches.

While having a different focus, our work is related to the following concurrent studies. Kozicki and Tinsley (2006) use a descriptive time series model of inflation to construct the term structure of inflation expectations using inflation forecasts from the Livingston Survey. D'Amico, Kim and Wei (2008) build a no-arbitrage model and estimate it using the nominal yield curve, TIPS, Blue Chip Economic Indicators forecasts of one-quarter yields, and one- and ten-year SPF forecasts of inflation from 1999 to 2007. They show the importance of using TIPS for accurate predictions of inflation. In a similar spirit, Joyce, Lildholdt and Sorensen (2009) exploit information from UK real and nominal bonds along with consensus forecasters' expectations of average inflation from five to ten years ahead from 1992 to 2007 to construct inflation forecasts and inflation risk premia. Piazzesi and Schneider (2008) focus on the use of an affine term-structure model to construct a measure of subjective bond risk premia, which they derive from survey data,

[10]They also consider the Phillips curve and pure time series models. However, discussion of these lies beyond the scope of our paper.

and they also specify a structural model that can explain these premia. Adrian and Wu (2009) and Haubrich, Pennacchi and Ritchken (2011) highlight the role of heteroscedastic shocks in the context of building inflation expectations with no-arbitrage models. Faust and Wright (2011) emphasize that subjective survey-based forecasts of inflation tend to outperform model-based forecasts and investigate the reasons for this. Cieslak and Povala (2011) propose a novel decomposition of the nominal yield curve into three economic frequencies. One of them is related to the smoothed CPI that can be interpreted as a long-run inflation forecasts. The residual from the projection of this factor onto yields explains a high fraction of variation in the bond risk premia. The authors do not take a stand on whether this residual is a hidden factor or not. Our survey factor is constructed from survey-based inflation expectations, it is orthogonal to factors driving yields, to current and past inflation, and is approximately hidden in the nominal yield curve.

Appendix B. Details of the data set

Here, we discuss the details of our data set that consists of inflation, output, real and nominal yields, and survey forecasts.

B.1. Nominal yields

We use quarterly time series of zero-coupon bond yields from 1971 to 2008 constructed in Gurkaynak, Sack and Wright (2007). We rely on bond maturities of three and six months and one, two, three, five, seven, and ten years. It is important to measure the full yield curve because its slope is correlated with the macro environment (Estrella and Hardouvelis, 1991 and Estrella and Mishkin, 1998). Given this constraint, we do not consider earlier years, because the longest maturity available was five years. In addition, the use of rich yield data helps to identify the risk premia.

B.2. Real yields

We use quarterly time series of zero-coupon yields on TIPS from 2003 to 2008 constructed in Gurkaynak, Sack and Wright (2010). We rely on bond maturities of three, five, seven and ten years. TIPS were introduced in 1997, but initially few maturities were available for trading and the trading volume was thin (D'Amico, Kim and Wei, 2008). Many observers argue that the bond pricing became reliable circa 2003.[11] An important related question is whether yields on TIPS can be used as yields on real bonds. D'Amico, Kim and Wei (2008) explain the differences between the TIPS and real yields by a liquidity premium. They find a strong trend in a liquidity component that dissipates by 2003, which is consistent with the above remarks. We cannot distinguish the true real yields from TIPS in our framework. This is why we choose to drop the data prior to 2003 and allow for a measurement error that is larger than that for nominal yields at the estimation stage.

B.3. Macro variables

We use quarterly time series of log changes in seasonally adjusted CPI and real GDP to proxy for π_t, and g_t, respectively. GDP and CPI numbers are available from FRED (Federal Reserve Economic Data). CPI is the consumer price index for all urban consumers (all items, seasonally adjusted) and real GDP is a three decimal time series in billions of chained year 2000 USD (seasonally adjusted annual rate).

All the inflation forecasts that we use are released some time during the third month of a quarter. Therefore, to avoid a look-ahead bias in how we construct and use the state variable π, we use the price level reported in the second month of the quarter, which corresponds to price level in the first month of a quarter.

[11] We are grateful to Refet Gürkaynak and Jonathan Wright for conversations regarding this issue.

B.4. Survey forecasts

Livingston Survey (LS) is a semiannual survey of economists from industry, government, banking, and academia. The forecasts of the price level $P_{t+\tau}$, specifically of nonseasonally adjusted CPI, are released in the last month of the second or fourth quarter and are based on the CPI information released in the first month of a quarter. Because of the timing of this survey, Carlson (1977) argues that the six-month ahead and 12-month ahead level forecasts should be converted to the inflation rate using the eight- and 14-month horizons, respectively, as a basis. As a result, we have observations for $\overline{p}^1_{t,0}(\tau)$, $\tau = 2, 4$ (available from 1947). We also have annual forward forecasts available from 1974: $\overline{p}^1_{t,2}(4)$ (released in the second quarter) and $\overline{p}^1_{t,4}(4)$ (released in the fourth quarter). Finally, ten-year forecasts, $\overline{p}^1_{t,0}(40)$, are available semiannually from 1991, and two-year forecasts , $\overline{p}^1_{t,0}(8)$, are available annually in the fourth quarter from 1992.

The fact that the respondents forecast seasonally unadjusted CPI, while our state variable is seasonally adjusted, matters only for the six-month forecasts, $\overline{p}^1_{t,0}(2)$. We perform a simple seasonal adjustment of the forecast. We compute the average annual inflation, as well as the average inflation over the first and second half-year in our sample. Then we adjust the six-month forecasts by the respective differences in annual and semiannual annualized average inflations. Ghysels and Osborn (2001) provide the details of, and a justification for, this procedure. This adjustment involves some look-ahead bias because we are performing the adjustment on the whole sample. However, this bias should be very small, because the seasonal adjustment is tiny.

Survey of Professional Forecasters (SPF) is a quarterly survey available from the third quarter of 1981. The forecasts of the annualized percentage price change $P_{t+\tau}/P_t - 1$, specifically of the changes in seasonally adjusted CPI, are released in the middle of the second month of the quarter and are based on the CPI information released in the previous month. We have observations for $\overline{p}^2_{t,s}(1)$, $s = 0, 1, 2, 3$; $\overline{p}^2_{t,0}(\tau)$, $\tau = 4, 20, 40$. The five-year forecast is available from the third quarter of 2005, and the ten-year forecast is available from the fourth quarter of 1991.

Blue Chip Economic Indicators (BC) is a monthly survey of economic forecasters at approximately 50 banks, corporations, and consulting firms.[12] It is available from 1981. The forecasts of the annualized percentage price change $P_{t+\tau}/P_t - 1$, specifically of the changes in seasonally adjusted CPI, are released in the beginning of the third month of the quarter and are based on the CPI information released in the previous month. We have observations for $\overline{p}^3_{t,s}(1)$, $s = 0 - 6$. The forecasts with $s = 4, 5, 6$ are available in the first three, two, and one quarters of a year, respectively.

Appendix C. Factor rotations

We wish to rotate vector x_t (a subvector of vector z_t) so that these latent factors could become interpretable. The first rotation, $\mathcal{O}$, ensures that the three factors are orthogonal to each other. We define a rotation $\mathcal{O} = Rx_t$, so that the variance-covariance matrix of f becomes diagonal. Then, define $\mathcal{M} = Ux_t$, where the matrix U is the orthogonal matrix; i.e., $UU' = I$, which preserves the correlation structure between the factors. We use the following matrix U, governed by the two parameters α and β:

$$U = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 & 0 & 0 \\ -\sin\alpha & \cos\alpha & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

[12] We are grateful to Randell Moore for providing us with the data.

$$= \begin{pmatrix} \cos\alpha\cos\beta & \sin\alpha & \cos\alpha\sin\beta & 0 & 0 \\ -\sin\alpha\cos\beta & \cos\alpha & -\sin\alpha\sin\beta & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The first two matrixes represent clockwise rotations about one of the axes, specifically, x_3 and x_2, respectively.

The $\mathbb{Q}$-parameters are affected by the rotation as follows:

$$\begin{aligned} \mu_U^{\mathbb{Q}} &= U\mu^{\mathbb{Q}} \\ \Phi_U^{\mathbb{Q}} &= U\Phi^{\mathbb{Q}}U^{-1} \end{aligned}$$

and

$$\Sigma_U = U\Sigma$$

The parameter α is chosen such that $e_1'\mathcal{B}^n(3)$ is maximized, and β is chosen so that $e_2'\left(\mathcal{B}^n(40) - \mathcal{B}^n(3)\right)$ is maximized.

References

Adrian, T., Wu, H., 2009. The term structure of inflation expectations. Federal Reserve Bank of New York Staff Reports, No. 362.

Aït-Sahalia, Y., Kimmel, R., 2010. Estimating affine multifactor term structure models using closed-form likelihood expansions. Journal of Financial Economics 98, 113–144.

Ang, A., Bekaert, G., Wei, M., 2007. Do macro variables, asset markets, or surveys forecast inflation better? Journal of Monetary Economics 54, 1163–1212.

Ang, A., Bekaert, G., Wei, M., 2008. The term structure of real rates and expected inflation. Journal of Finance 63, 797–849.

Ang, A., Dong, S., Piazzesi, M., 2004. No-arbitrage Taylor rules. Unpublished working paper. Columbia University, New York.

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.

Backus, D., Foresi, S., Telmer, C., 1999. Discrete time models of bond pricing, in: Jegadeesh, N., Tuckman, B. (Eds.), Advanced Fixed-Income Valuation Tools. Wiley and Sons, New York, pp. 87–127.

Bansal, R., Shaliastovich, I., 2010. A long-run risks explanation of predictability puzzles in bond and currency markets. Unpublished working paper. Duke University Fuqua School of Business, Durham, NC.

Basak, S., 2005. Asset pricing with heterogeneous beliefs. Journal of Banking and Finance 29, 2849–2881.

Bikbov, R., Chernov, M., 2010. No-arbitrage macroeconomic determinants of the yield curve. Journal of Econometrics 159, 166–182.

Bikbov, R., Chernov, M., 2011. Yield curve and volatility: Lessons from eurodollar futures and options. Journal of Financial Econometrics 9, 66–105.

Brandt, M.W., He, P., 2006. Simulated likelihood estimation of affine term structure models from panel data. Unpublished working paper. Duke University Fuqua School of Business, Durham, NC.

Carlson, J.A., 1977. A study of price forecasts. Annals of Economic and Social Measurement 6, 27–56.

Chen, R.R., Scott, L.O., 1993. Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. Journal of Fixed Income 3, 14–31.

Christiano, L., Eichenbaum, M., Evans, C., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. Journal of Political Economy 113, 1–45.

Chun, A.L., 2011. Expectations, bond yields, and monetary policy. Review of Financial Studies 24, 208–247.

Cieslak, A., Povala, P., 2011. Understanding bond risk premia. Unpublished working paper. Kellogg School of Management, Evanston, IL.

Cochrane, J., Piazzesi, M., 2005. Bond risk premia. American Economic Review 95, 138–160.

Collin-Dufresne, P., Goldstein, R., Jones, C., 2008. Identification and estimation of maximal affine term structure models. Journal of Finance 63, 743–795.

Conley, T., Hansen, L.P., Liu, W., 1997. Bootstrapping the long run. Macroeconomic Dynamics 1, 279–311.

Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. Journal of Finance 55, 1943–1978.

D'Amico, S., Kim, D.H., Wei, M., 2008. Tips from TIPS: the informational content of Treasury inflation-protected security prices. Unpublished working paper 2008-30. Federal Reserve Board, Washington, D.C.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and Their Application. Cambridge University Press, Cambridge, UK.

Detemple, J., Murthy, S., 1994. Intertemporal asset pricing with heterogeneous beliefs. Journal of Economic Theory 62, 294–320.

Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253–263.

Duffee, G.R., 2002. Term premia and interest rate forecasts in affine models. Journal of Finance 57, 405–443.

Duffee, G.R., 2006. Term structure estimation without using latent factors. Journal of Financial Economics 79, 507–536.

Duffee, G.R., 2007. Are variations in term premia related to the macroeconomy? Unpublished working paper. University of California, Berkeley, CA.

Duffee, G.R., 2011. Information in (and not in) the term structure. Review of Financial Studies 24, 2895–2934.

Duffee, G.R., Stanton, R., 2004. Estimation of dynamic term structure models. Unpublished working paper. University of California, Berkeley, CA.

Dumas, B., Kurshev, A., Uppal, R., 2005. What can rational investors do about excessive volatility and sentiment fluctuations? Unpublished working paper. INSEAD, Fontainebleau, France.

Estrella, A., Hardouvelis, G.A., 1991. The term structure as a predictor of real economic activity. Journal of Finance 46, 555–576.

Estrella, A., Mishkin, F., 1998. Predicting US recessions: Financial variables as leading indicators. Review of Economics and Statistics 80, 45–61.

Evans, M.D.D., 1998. Real rates, expected inflation, and inflation risk premia. Journal of Finance 53, 187–218.

Faust, J., Wright, J.H., 2011. Forecasting inflation. Unpublished working paper. Johns Hopkins University, Baltimore, MD.

Gallmeyer, M., Hollifield, B., Palomino, F., Zin, S., 2007. Arbitrage free bond pricing with dynamic macroeconomic models. Federal Reserve Bank of St. Louis Review 89, 305–326.

Gallmeyer, M., Hollifield, B., Palomino, F., Zin, S., 2009. Term premium dynamics and the Taylor rule. Unpublished working paper. Carnegie Mellon University, Pittsburgh, PA.

Ghysels, E., Osborn, D., 2001. The Econometric Analysis of Seasonal Time Series. Cambridge University Press, Cambridge, UK.

Ghysels, E., Wright, J., 2009. Forecasting professional forecasters. Journal of Business and Economic Statistics 27, 504–516.

Giacomini, R., White, H., 2006. Tests of conditional predicitive ability. Econometrica 74, 1545–1578.

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

Gurkaynak, R., Sack, B., Wright, J., 2010. The TIPS yield curve and inflation compensation. American Economic Journal: Macroeconomics 2, 70–92.

Harrison, J., Kreps, D., 1978. Speculative investor behavior in a stock market with heterogenoius expectations. Quarterly Journal of Economics 92, 323–336.

Harvey, A.C., 1989. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge, UK.

Haubrich, J., Pennacchi, G., Ritchken, P., 2011. Inflation expectations, real rates, and risk premia: Evidence from inflation swaps. Review of Financial Studies, forthcoming.

Hördahl, P., Tristani, O., 2007. Inflation risk premia in the term structure of interest rates. Unpublished working paper 228. BIS, Basel, Switzerland.

de Jong, F., 2000. Time series and cross-section information in affine term-structure models. Journal of Business and Economic Statistics 18, 300–314.

Joslin, S., 2011. Can unspanned stochastic volatility models explain the cross section of bond volatilities? Unpublished working paper. USC Marshall School of Business, Los Angeles, CA.

Joslin, S., Le, A., Singleton, K.J., . Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs. Journal of Financial Economics, forthcoming.

Joslin, S., Priebsch, M., Singleton, K.J., 2010. Risk premiums in dynamic term structure models with unspanned macro risks. Unpublished working paper. Stanford Graduate School of Business, Stanford, CA.

Joyce, M., Lildholdt, P., Sorensen, S., 2009. Extracting inflation expectations and inflation risk premia from the term structure: a joint model of the UK nominal and real yield curves. Unpublished working paper 360. Bank of England, London, UK.

Kim, D.H., Orphanides, A., 2012. Term structure estimation with survey data on interest rate forecasts. Journal of Financial and Quantitative Analysis 47, 241–271.

Kozicki, S., Tinsley, P., 2006. Survey-based estimates of the term structure of expected US inflation. Unpublished working paper 06-46. Bank of Canada, Ottawa, Ontario, Canada.

Levene, H., 1960. Robust tests for equality of variances, in: Olkin, I. (Ed.), Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press, Stanford, CA, pp. 278–292.

Lu, B., Wu, L., 2005. Systematic movements in macroeconomic releases and the term structure of interest rates. Unpublished working paper. Baruch College, New York, NY.

Ludvigson, S.C., Ng, S., 2009. Macro factors in bond risk premia. Review of Financial Studies 22, 5027–5067.

Mueller, P., 2008. Credit spreads and real activity. Unpublished working paper. London School of Economics, London, UK.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Palomino, F., 2012. Bond risk premiums and optimal monetary policy. Review of Economic Dynamics 15, 19–40.

Patton, A., Timmermann, A., 2010. Generalized forecast errors, a change of measure, and forecast optimality conditions, in: Bollerslev, T., Russell, J., Watson, M. (Eds.), Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle. Oxford University Press, Oxford, UK, pp. 194–212.

Pennacchi, G.G., 1991. Identifying the dynamics of real interest rates and inflation: Evidence using survey data. Review of Financial Studies 4, 53–86.

Piazzesi, M., Schneider, M., 2006. Equilibrium yield curves, in: Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), NBER Macroeconomics Annual 2006. MIT Press, Cambridge MA, pp. 389–496.

Piazzesi, M., Schneider, M., 2008. Bond positions, expectations, and the yield curve. Unpublished working paper. University of Chicago Booth School of Business, Chicago, IL.

Press, W.H., Teukovsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipies in C. Cambridge University Press, Cambridge, UK. second edition.

Rudebusch, G., Swanson, E., 2008. Examining the bond premium puzzle with a DSGE model. Journal of Monetary Economics 55, 111–126.

Rudebusch, G., Swanson, E., 2012. The bond premium in a DSGE model with long-run real and nominal risks. American Economic Journal: Macroeconomics 4, 105–43.

Scheinkman, J., Xiong, W., 2003. Overconfidence and speculative bubbles. Journal of Political Economy 111, 1183–1219.

Stock, J.H., Watson, M.W., 2003. Forecasting output and inflation: The role of asset prices. Journal of Economic Literature 41, 788–829.

Thomas, L.B., 1999. Survey measures of expected US inflation. Journal of Economic Perspectives 13, 125–144.

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

Table 1

Mean absolute errors.

We report mean absolute fitting errors for eight versions of our model. The full model that uses all survey and yield data is labeled AST (All data, Subjective expectations, TIPS). The model that does not use Treasury Inflation-Protected Securities is labeled AS. A model that uses all the data, but restricts subjective measures to coincide with objective measure is labeled AOT (All data, Objective expectations, TIPS) and AO without TIPS. The two models that use only yields in the estimation are NFT (No Forecasts, TIPS) and NF (without TIPS). Finally, we estimate a model that uses only forecasts and label it OF (Only Forecasts), and a model that restricts subjective measures to coincide with objective measure is labeled OFO (Only Forecasts, Objective expectations). The number at the end of mnemonic refers to a total number of factors in the model. The notation for surveys is as follows: LS, Livingston Survey; SPF, Survey of Professional Forecasters; BC, Blue Chip Economic Indicators. The row $\mathcal{L}$ reports log-likelihood values. The log-likelihoods corresponding to the unreported models' counterparts are displayed in parentheses. A counterpart is the same model with a different number of factors. Thus, for four-factor models, the five-factor versions would be in parenthesis, and vice versa. The bootstrapped p-values for the null hypothesis of a four-factor model are provided in brackets.

	Number	Horizon (in quarters)		Model							
Survey	of observations	s	τ	NF4	NFT4	OF4	OFO4	AO5	AOT5	AS5	AST5
LS	74	0	2			0.15	0.21	0.22	0.22	0.16	0.16
	74	0	4			0.11	0.13	0.14	0.14	0.12	0.12
	16	0	8			0.07	0.10	0.13	0.12	0.11	0.11
	36	0	40			0.03	0.05	0.12	0.11	0.11	0.12
	36	2	4			0.11	0.13	0.14	0.14	0.11	0.12
	34	4	4			0.14	0.16	0.20	0.21	0.18	0.18
SPF	108	0	1			0.09	0.10	0.14	0.14	0.13	0.13
	12	0	20			0.03	0.07	0.13	0.11	0.05	0.05
	67	0	40			0.04	0.04	0.10	0.09	0.09	0.09
	108	1	1			0.06	0.07	0.08	0.08	0.07	0.07
	108	2	1			0.08	0.09	0.09	0.09	0.08	0.08
	108	3	1			0.10	0.11	0.10	0.11	0.10	0.10
BCEI	114	0	1			0.19	0.19	0.20	0.20	0.20	0.21
	114	1	1			0.14	0.14	0.13	0.13	0.14	0.14
	114	2	1			0.16	0.18	0.19	0.19	0.21	0.21
	113	3	1			0.11	0.13	0.15	0.15	0.13	0.13
	84	4	1			0.11	0.13	0.15	0.15	0.15	0.15
	56	5	1			0.10	0.13	0.17	0.17	0.17	0.17
	28	6	1			0.10	0.13	0.19	0.20	0.19	0.19
Nominal yields	148	0	1	0.16	0.16			0.16	0.16	0.17	0.16
	148	0	2	0.06	0.06			0.06	0.06	0.06	0.06
	148	0	4	0.10	0.10			0.09	0.09	0.09	0.09
	148	0	8	0.12	0.13			0.12	0.12	0.12	0.12
	148	0	12	0.10	0.10			0.10	0.10	0.10	0.10
	148	0	20	0.05	0.05			0.05	0.05	0.05	0.05
	148	0	28	0.05	0.06			0.06	0.05	0.06	0.05
	148	0	40	0.13	0.13			0.13	0.13	0.13	0.13
Real yields	18	0	12		0.24				0.20		0.20
	22	0	20		0.23				0.16		0.16
	22	0	28		0.22				0.15		0.15
	22	0	40		0.20				0.16		0.16
$\mathcal{L}$				53.89	53.90	56.32	55.23	101.45	101.65	102.55	102.72
				(57.22)	(57.28)	(57.10)	(55.51)	(92.29)	(86.54)	(93.23)	(93.93)
p-values				[0.31]	[0.30]	[0.96]	[0.95]	[0.02]	[0.00]	[0.04]	[0.05]

Table 2
In-sample inflation forecast root mean squared error (RMSE) ratios for NF (No Forecasts) and OF (Only Forecasts) models.
We compare the in-sample inflation forecasting performance of the OF and NF models. We select forecasting horizons that are not necessarily available in the actual surveys. This is done to emphasize that the advantage of using a model is that a forecasting horizon or forecasting date can be arbitrary. We report the RMSE ratios of NF to OF and other variations of these two models. The bootstrapped 95% confidence bounds are provided in brackets. The letter T in the model indicates that we include Treasury Inflation-Protected Securities (TIPS) in the estimation. All models have four factors.

Horizon (in quarters)	NF4/OFO4	NFT4/OFO4	NF4/OF4	NFT4/OF4
1	0.98	1.01	1.04	1.07
	[0.98, 1.01]	[0.99, 1.01]	[0.89, 1.04]	[0.93, 1.04]
2	1.01	1.05	1.08	1.12
	[0.97, 1.02]	[0.99, 1.01]	[0.83, 1.07]	[0.88, 1.07]
4	1.04	1.09	1.09	1.15
	[0.96, 1.02]	[0.98, 1.01]	[0.78, 1.10]	[0.84, 1.12]
8	1.12	1.13	1.12	1.13
	[0.96, 1.02]	[0.97, 1.02]	[0.74, 1.18]	[0.75, 1.22]
12	1.16	1.13	1.11	1.08
	[0.96, 1.02]	[0.96, 1.02]	[0.70, 1.27]	[0.71, 1.29]
20	1.15	1.12	1.09	1.05
	[0.94, 1.03]	[0.95, 1.04]	[0.65, 1.45]	[0.66, 1.46]
28	1.06	1.04	1.02	1.00
	[0.93, 1.05]	[0.93, 1.06]	[0.57, 1.66]	[0.59, 1.68]
40	0.96	0.93	0.95	0.92
	[0.92, 1.06]	[0.91, 1.08]	[0.47, 2.07]	[0.47, 2.00]

Table 3
Inflation expectations for AO (All data, Objective expectations) and AS (All data, Subjective expectations) models.
We compare the average inflation expectations from the AO model with the subjective expectations from the AS model. One-way analysis of variance (ANOVA) tests whether at least one average expectation that is significantly different from another average expectation. The Levene statistic is used to test the equality of variances. The bootstrapped *p*-values are for the null hypothesis of a model under the objective measure (AO5 and AOT5) are provided in brackets. The letter T in the model indicates that we include Treasury Inflation-Protected Securities (TIPS) in the estimation. All models have five factors.

Horizon	AO5 vs AS5		AOT5 vs AST5	
(in quarters)	ANOVA	Levene	ANOVA	Levene
1	0.44	0.75	0.97	0.37
	[0.50]	[0.25]	[0.30]	[0.25]
2	1.14	0.92	1.30	0.26
	[0.29]	[0.13]	[0.23]	[0.30]
4	0.38	0.71	0.54	0.14
	[0.46]	[0.16]	[0.43]	[0.39]
8	0.33	0.54	0.47	0.09
	[0.47]	[0.20]	[0.45]	[0.50]
12	0.23	0.50	0.41	0.11
	[0.56]	[0.21]	[0.51]	[0.47]
20	0.23	0.60	0.38	0.25
	[0.63]	[0.21]	[0.60]	[0.37]
28	0.73	0.69	0.45	0.40
	[0.55]	[0.27]	[0.62]	[0.33]
40	2.56	0.72	0.68	0.42
	[0.44]	[0.41]	[0.61]	[0.43]

Table 4
Sample statistics for nominal yields decomposition.
We report the annualized mean and standard deviation of the ten-year nominal yield components in the decomposition

$$y_t^n(40) = y_t^r(40) + \overline{p}_{t,0}(40) + IP_t(40) + \text{convexity term},$$

where y_t^r is the real yield, $\overline{p}_{t,0}$ is the inflation expectation under the $\mathbb{P}$ measure, and IP_t is the inflation premium. The notation for models is as follows: AO5 and AOT5 (All data, Objective expectations, five factors, with and without Treasury Inflation-Protected Securities, respectively); NF4 and NFT4 (No Forecasts, four factors, with and without Treasury Inflation-Protected Securities, respectively).

Model	Statistic	y_t^r	$\overline{p}_{t,0}$	IP_t	Convexity
NF4	Mean	1.11	4.34	2.01	0.02
	Standard deviation	0.26	0.95	1.47	
NFT4	Mean	2.66	4.08	0.73	0.01
	Standard deviation	0.59	0.56	1.36	
AO5	Mean	1.93	3.96	1.57	0.01
	Standard deviation	0.24	1.38	1.37	
AOT5	Mean	2.75	4.05	0.67	0.01
	Standard deviation	0.92	1.50	0.63	

Table 5
Excess returns and the survey factor.
Panel A reports regression results of the survey factor f_3 on the Cochrane and Piazzesi (2005) factor (CP) and the macro factors $\widehat{F}^{(j)}$, $j = 1\ldots, 8$ from Ludvigson and Ng (2009) (LN). Panels B through F report regression results of one-year excess returns on a nominal Treasury bond with maturity τ, $rx^{(\tau)}$, on the CP and LN factors, and the projection residuals f_1, f_2, and f_3 from the AOT5 model. The regressions in Panels A and C were implemented with all LN factors, but we are only reporting the selected ones. The regression in Panel F is a result of a factor selection procedure, so only the reported factors were used. All time series are standardized, and *t*-statistics are calculated using Newey and West (1987) standard errors and reported in parentheses. GDP is the gross domestic product and CPI is the consumer price index.

Regressand	CP	GDP	CPI	$\widehat{F}^{(1)}$	$\widehat{F}^{(6)}$	$\widehat{F}^{(7)}$	$\widehat{F}^{(8)}$	f_1	f_2	f_3	Adj. R^2
Panel A: Survey factor regressions											
f_3				-0.24	0.07	0.01	-0.12				0.09
				(-1.59)	(0.73)	(0.12)	(-1.20)				
f_3	-0.22			-0.16	0.04	-0.02	0.04				0.10
	(-1.34)			(-1.35)	(0.33)	(-0.25)	(0.45)				
Panel B: CP factor											
$rx^{(2)}$	0.54										0.29
	(5.49)										
$rx^{(3)}$	0.57										0.32
	(6.06)										
$rx^{(4)}$	0.59										0.35
	(6.41)										
$rx^{(5)}$	0.56										0.31
	(6.11)										
Panel C: CP factor and macro variables											
$rx^{(2)}$	0.45	0.09	-0.21	0.28	-0.12	-0.10	0.04				0.47
	(4.15)	(1.07)	(-1.80)	(4.71)	(-1.85)	(-1.28)	(0.80)				
$rx^{(3)}$	0.47	0.11	-0.23	0.23	-0.12	-0.11	0.02				0.46
	(0.49)	(1.44)	(-1.99)	(3.64)	(-1.97)	(-1.39)	(0.45)				
$rx^{(4)}$	0.49	0.12	-0.22	0.20	-0.13	-0.13	0.02				0.47
	(4.25)	(1.64)	(-1.93)	(3.16)	(-2.19)	(-1.69)	(0.47)				
$rx^{(5)}$	0.45	0.14	-0.23	0.18	-0.13	-0.14	0.04				0.42
	(3.73)	(1.74)	(-1.87)	(2.57)	(-2.19)	(-1.68)	(0.63)				
Panel D: Projection residuals											
$rx^{(2)}$								0.20	0.33	-0.35	0.36
								(2.38)	(3.24)	(-4.22)	
$rx^{(3)}$								0.13	0.39	-0.32	0.33
								(1.47)	(3.34)	(-3.77)	
$rx^{(4)}$								0.08	0.43	-0.32	0.33
								(0.95)	(3.60)	(-3.55)	
$rx^{(5)}$								0.05	0.46	-0.32	0.34
								(0.62)	(3.93)	(-3.69)	
Panel E: CP factor and projection residuals											
$rx^{(2)}$	0.38							0.26	0.08	-0.25	0.45
	(3.34)							(2.94)	(0.70)	(-3.46)	
$rx^{(3)}$	0.41							0.19	0.11	-0.21	0.43
	(3.56)							(2.09)	(0.88)	(-2.87)	
$rx^{(4)}$	0.43							0.14	0.14	-0.21	0.44
	(4.14)							(1.73)	(1.17)	(-2.71)	
$rx^{(5)}$	0.35							0.11	0.22	-0.22	0.35
	(3.29)							(1.33)	(1.74)	(-2.85)	
Panel F: All regressors											
$rx^{(2)}$	0.23		-0.20	0.20	-0.19	-0.11	0.10	0.32		-0.24	0.57
	(2.11)		(-2.68)	(3.40)	(-2.25)	(-2.33)	(1.75)	(4.18)		(-3.38)	
$rx^{(3)}$	0.28		-0.24	0.15	-0.20	-0.11	0.11	0.27		-0.22	0.55
	(2.34)		(-2.96)	(2.58)	(-2.32)	(-2.02)	(1.87)	(3.70)		(-3.03)	
$rx^{(4)}$	0.32		-0.22	0.11	-0.20	-0.12	0.12	0.24		-0.21	0.55
	(2.77)		(-2.80)	(1.92)	(-2.46)	(-2.17)	(2.38)	(3.44)		(-2.90)	
$rx^{(5)}$	0.29		-0.24	0.07	-0.22	-0.13	0.11	0.22		-0.22	0.52
	(2.34)		(-2.98)	(1.29)	(-2.61)	(-2.10)	(1.97)	(3.35)		(-3.21)	

Table 6
Out-of-sample inflation forecast root mean squared error (RMSE) ratios.
We compare the out-of-sample inflation forecasting performance of the AO5 (All data, Objective expectations, five factors) and NF4 (No Forecasts, four factors) models and the observed survey forecasts. Panel A compares the models to the actual surveys by reporting the RMSE ratio of a model to a survey. The notation for surveys is as follows: LS, Livingston Survey; SPF, Survey of Professional Forecasters; BC, Blue Chip Economic Indicators. In Panel B we select alternative forecasting horizons. We report RMSE ratios of AO5 to NF4 as the surveys are not available for most of the selected horizons. Actual survey forecasts have missing observations. Therefore, the RMSE ratios are different for the Panel B forecasts even if a horizon matches the one from a survey in Panel A. The RMSE ratios are reported for two samples. The first sample represents the full out-of-sample period from 1983 to 2008. The second sample begins in 1990 after the long-run forecasts were incorporated into the surveys. The bootstrapped 95% confidence bounds are provided in brackets.

Panel A: Survey-selected horizons

Survey	Horizon (in quarters)		1983–2008		1990–2008	
	s	τ	AO5/Survey	NF4/Survey	AO5/Survey	NF4/Survey
LS	0	2	0.98 [0.97, 0.99]	1.44 [1.17, 1.80]	0.97 [0.96, 0.97]	1.41 [1.16, 1.79]
	0	4	0.98 [0.97, 0.99]	1.54 [1.19, 2.13]	0.99 [0.98, 1.00]	1.76 [1.28, 2.49]
	0	8	1.00 [0.99, 1.01]	2.61 [1.76, 4.46]	1.00 [0.99, 1.01]	2.61 [1.76, 4.46]
	0	40	1.04 [0.85, 1.40]	2.22 [0.44, 4.25]	1.04 [0.85, 1.40]	2.22 [0.44, 4.25]
	2	4	1.01 [0.99, 1.04]	1.43 [1.08, 2.27]	1.02 [1.00, 1.04]	1.87 [1.33, 3.08]
	4	4	1.02 [0.98, 1.10]	1.52 [1.02, 2.56]	1.07 [1.04, 1.11]	1.99 [1.26, 3.66]
SPF	0	1	1.01 [1.00, 1.01]	1.36 [1.20, 1.60]	1.00 [0.99, 1.00]	1.30 [1.17, 1.52]
	0	40	1.09 [0.84, 1.41]	2.51 [0.48, 4.89]	1.09 [0.84, 1.41]	2.51 [0.48, 4.89]
	1	1	1.01 [1.01, 1.01]	1.21 [1.08, 1.44]	1.00 [1.00, 1.00]	1.18 [1.06, 1.42]
	2	1	1.01 [1.01, 1.01]	1.13 [1.00, 1.44]	1.00 [0.99, 1.01]	1.14 [1.02, 1.49]
	3	1	1.01 [1.01, 1.02]	1.21 [1.07, 1.59]	1.00 [1.00, 1.01]	1.26 [1.09, 1.68]
BC	0	1	1.04 [1.04, 1.05]	1.40 [1.24, 1.65]	1.03 [1.02, 1.04]	1.34 [1.21, 1.57]
	1	1	1.00 [1.00, 1.01]	1.21 [1.07, 1.43]	1.00 [1.00, 1.01]	1.19 [1.06, 1.42]
	2	1	0.97 [0.97, 0.97]	1.09 [0.97, 1.38]	1.01 [1.01, 1.01]	1.16 [1.03, 1.51]
	3	1	1.01 [1.00, 1.02]	1.21 [1.06, 1.58]	1.01 [1.01, 1.02]	1.27 [1.10, 1.69]
	4	1	1.01 [1.00, 1.02]	1.14 [1.01, 1.47]	1.02 [1.01, 1.02]	1.20 [1.07, 1.62]
	5	1	1.00 [0.99, 1.02]	1.17 [1.05, 1.52]	1.02 [1.01, 1.02]	1.22 [1.09, 1.59]
	6	1	1.00 [0.99, 1.02]	1.18 [1.03, 1.51]	1.00 [1.00, 1.02]	1.18 [1.02, 1.52]

Panel B: Arbitrary horizons

Horizon (in quarters)		1983–2008	1990–2008
s	τ	NF4/AO5	NF4/AO5
0	1	1.35 [0.96, 1.00]	1.30 [0.93, 1.03]
0	2	1.50 [0.94, 1.01]	1.46 [0.90, 1.04]
0	4	1.57 [0.91, 1.02]	1.71 [0.85, 1.08]
0	8	1.69 [0.86, 1.04]	2.18 [0.79, 1.15]
0	12	1.65 [0.82, 1.07]	2.21 [0.73, 1.24]
0	20	1.57 [0.74, 1.12]	2.09 [0.62, 1.41]
0	28	1.49 [0.65, 1.20]	2.03 [0.55, 1.57]
0	40	1.32 [0.53, 1.37]	1.97 [0.47, 1.72]

Table 7
Out-of-sample yield forecast root mean squared error (RMSE) ratios.
We compare the out-of-sample yield forecasting performance of the AO5 (All data, Objective expectations, five factors), NF4 (No Forecasts, four factors), and RW (random walk) models. We report the RMSE ratios of a no-arbitrage model (AO or NF) to RW. The RMSE ratios are reported for two samples. The first sample represents the full out-of-sample period from 1983 to 2008. The second sample begins in 1990 after the long-run forecasts were incorporated into the surveys. The bootstrapped 95% confidence bounds are provided in brackets.

Yield maturities	Forecast horizon	1983–2008		1990–2008	
(in quarters)	(in quarters)	AO/RW	NF/RW	AO/RW	NF/RW
1	1	1.08 [1.00, 1.41]	0.95 [0.92, 1.08]	1.08 [0.89, 1.62]	0.89 [0.83, 1.12]
	4	0.93 [0.85, 1.04]	1.58 [1.03, 2.24]	0.89 [0.78, 1.17]	0.99 [0.83, 1.28]
	8	0.72 [0.69, 0.75]	1.42 [0.97, 1.88]	0.75 [0.67, 0.95]	0.99 [0.77, 1.32]
	20	0.62 [0.62, 0.73]	1.52 [0.97, 1.76]	0.57 [0.56, 0.62]	1.29 [0.81, 2.13]
2	1	1.13 [1.06, 1.46]	1.01 [0.97, 1.13]	1.17 [0.99, 1.67]	1.00 [0.93, 1.23]
	4	0.94 [0.87, 1.05]	1.52 [1.03, 2.05]	0.91 [0.81, 1.16]	0.98 [0.85, 1.25]
	8	0.72 [0.70, 0.75]	1.37 [0.95,1.69]	0.74 [0.68, 0.93]	0.95 [0.76, 1.26]
	20	0.64 [0.62, 0.75]	1.45 [0.94, 1.82]	0.57 [0.56, 0.62]	1.21 [0.77, 2.02]
4	1	1.15 [1.07, 1.46]	1.04 [1.01, 1.15]	1.21 [1.05, 1.62]	1.09 [1.02, 1.30]
	4	0.96 [0.88, 1.08]	1.42 [1.02, 1.75]	0.93 [0.83, 1.16]	0.96 [0.87, 1.20]
	8	0.72 [0.71, 0.82]	1.28 [0.90, 1.45]	0.74 [0.68, 0.91]	0.90 [0.75, 1.20]
	20	0.65 [0.62, 0.84]	1.34 [0.89, 2.28]	0.57 [0.55, 0.63]	1.11 [1.03, 1.29]
8	1	1.12 [1.04, 1.37]	1.04 [1.01, 1.16]	1.17 [1.05, 1.50]	1.10 [1.03, 1.29]
	4	0.97 [0.88, 1.11]	1.29 [0.98, 1.91]	0.93 [0.84, 1.16]	0.92 [0.87, 1.14]
	8	0.72 [0.70, 0.85]	1.18 [0.88, 2.10]	0.73 [0.68, 0.92]	0.84 [0.75, 1.12]
	20	0.66 [0.61, 0.88]	1.23 [0.81, 3.11]	0.57 [0.54, 0.65]	0.99 [0.70, 1.94]
12	1	1.09 [1.02, 1.28]	1.02 [0.99, 1.15]	1.11 [1.01, 1.39]	1.06 [1.01, 1.25]
	4	0.98 [0.88, 1.13]	1.22 [0.97, 2.37]	0.91 [0.83, 1.17]	0.89 [0.85, 1.02]
	8	0.73 [0.69, 0.89]	1.14 [0.86, 2.66]	0.73 [0.69, 0.94]	0.81 [0.75, 1.01]
	20	0.68 [0.60, 0.91]	1.19 [0.77, 3.76]	0.55 [0.52, 0.67]	0.91 [0.67, 2.01]
20	1	1.06 [1.00, 1.22]	1.01 [0.99, 1.15]	1.03 [0.96, 1.24]	1.01 [0.96, 1.20]
	4	1.00 [0.89, 1.17]	1.17 [0.95, 3.00]	0.89 [0.82, 1.17]	0.85 [0.82, 1.18]
	8	0.76 [0.70, 0.94]	1.13 [0.83, 3.51]	0.73 [0.69, 0.98]	0.77 [0.74, 1.26]
	20	0.73 [0.59, 0.93]	1.16 [0.74, 4.62]	0.50 [0.46, 0.67]	0.80 [0.61, 1.78]
28	1	1.07 [1.01, 1.24]	1.03 [1.01, 1.17]	1.00 [0.96, 1.19]	1.03 [0.97, 1.23]
	4	1.03 [0.91, 1.19]	1.15 [0.94, 3.35]	0.87 [0.81, 1.13]	0.85 [0.84, 1.27]
	8	0.81 [0.70, 0.98]	1.14 [0.81, 4.03]	0.73 [0.69, 1.00]	0.76 [0.76, 1.40]
	20	0.78 [0.58, 0.93]	1.18 [0.74, 5.06]	0.44 [0.40, 0.63]	0.72 [0.55, 1.42]
40	1	1.13 [1.03, 1.27]	1.09 [1.05, 1.27]	1.05 [1.01, 1.26]	1.12 [1.06, 1.33]
	4	1.08 [0.92, 1.20]	1.16 [0.95, 3.57]	0.87 [0.81, 1.14]	0.88 [0.87, 1.37]
	8	0.87 [0.72, 1.00]	1.17 [0.81, 4.44]	0.75 [0.69, 1.09]	0.78 [0.77, 1.61]
	20	0.86 [0.59, 0.94]	1.26 [0.75, 5.28]	0.39 [0.35, 0.65]	0.66 [0.49, 1.31]

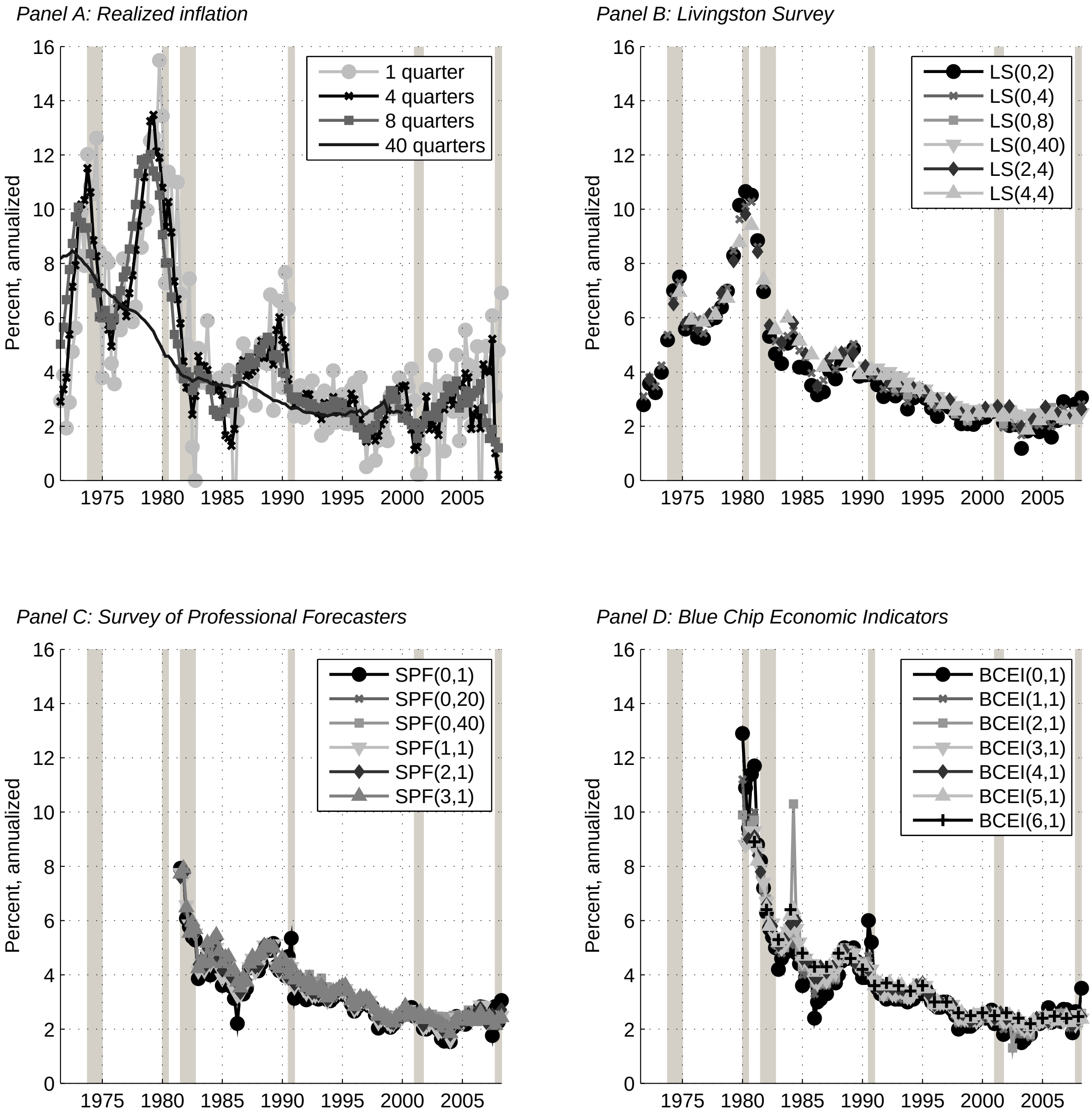

Fig. 1. The term structure of survey forecasts. We plot the realized inflation and survey-based inflation expectations that we use as inputs in our model. Our sample period is 1971 to 2008 at a quarterly frequency. Some forecasts do not start until a later date. Some are reported at a frequency lower than quarterly. The forecast horizon is indicated in the legend in quarters. (5,1), for example, indicates a one-quarter long forecasting horizon starting in five quarters from today. The notation for surveys is as follows: LS, Livingston Survey; SPF, Survey of Professional Forecasters; BC, Blue Chip Economic Indicators. The shaded regions show the National Bureau of Economic Research recessions.

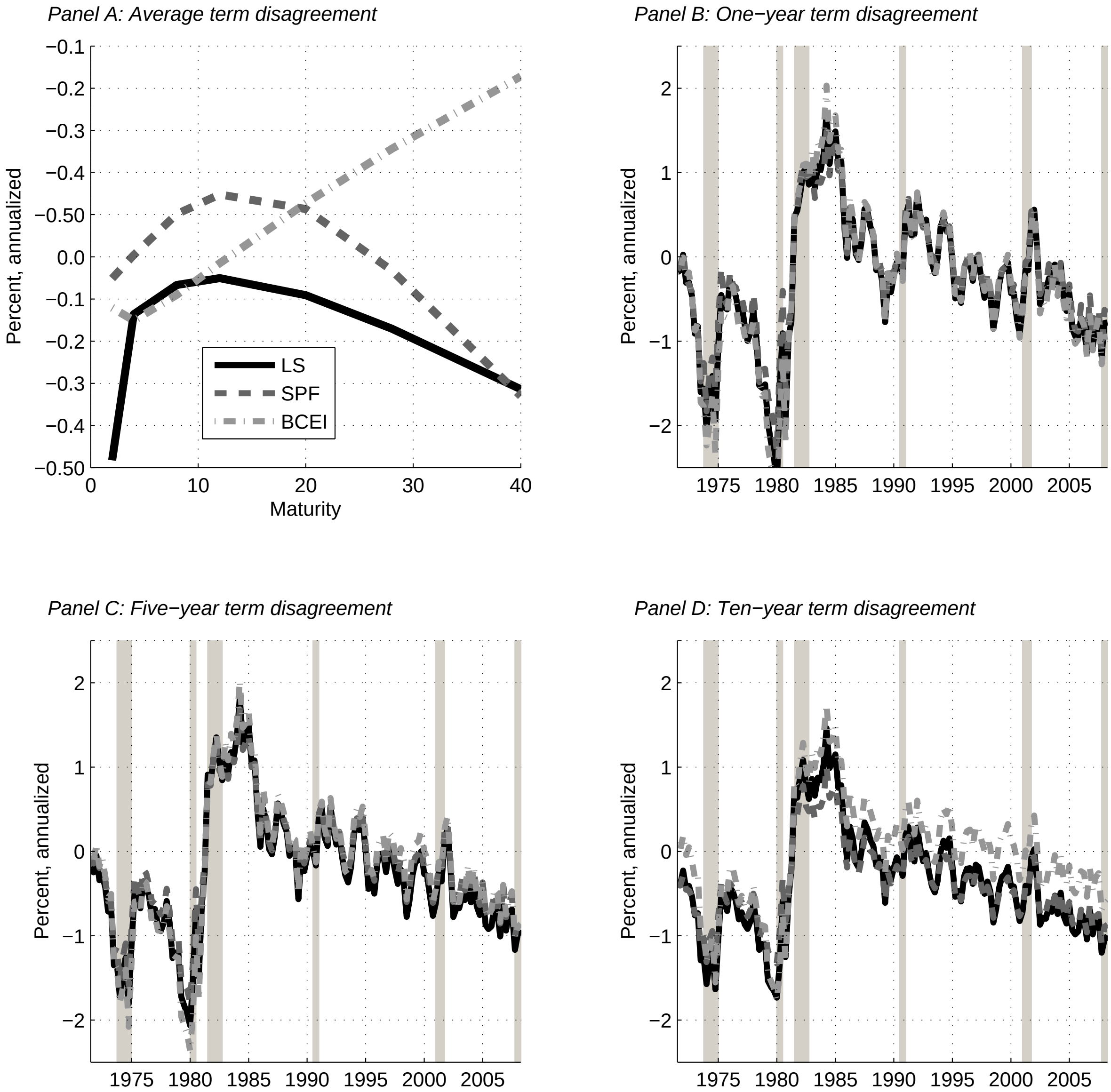

Fig. 2. Term disagreement from AST5 (All data, Subjective expectations, including Treasury Inflation-Protected Securities, five factors). The figure shows unconditional and conditional deviations of model-based subjective inflation expectations from the objective ones. The notation for surveys is as follows: LS, Livingston Survey; SPF, Survey of Professional Forecasters; BC, Blue Chip Economic Indicators. The shaded regions show the National Bureau of Economic Research recessions.

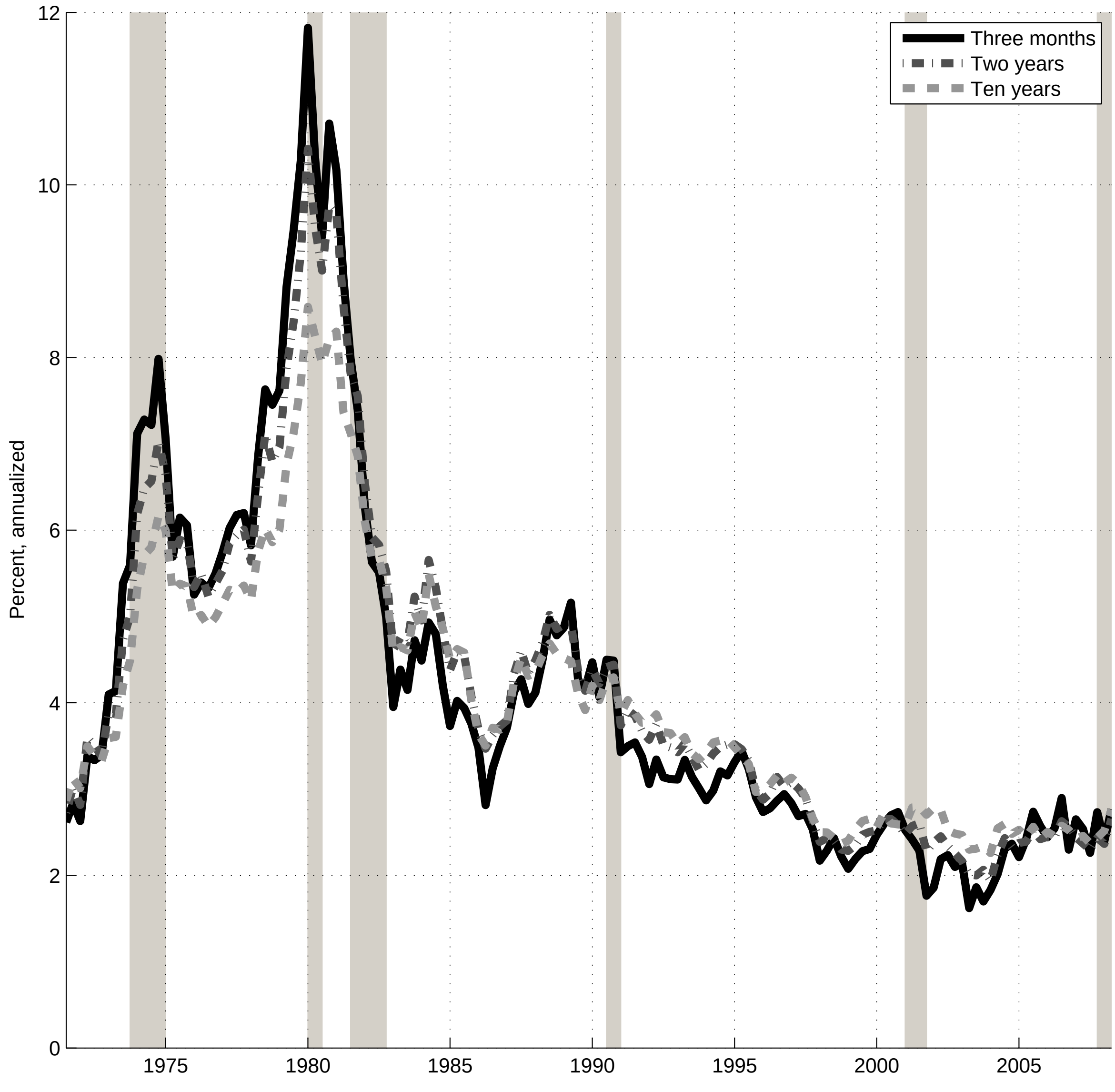

Fig. 3. The term structure of inflation expectations. The figure displays the model-based objective inflation expectations computed from the AOT5 model (All data, Objective expectations, including Treasury Inflation-Protected Securities, five factors). The shaded regions show the National Bureau of Economic Research recessions.

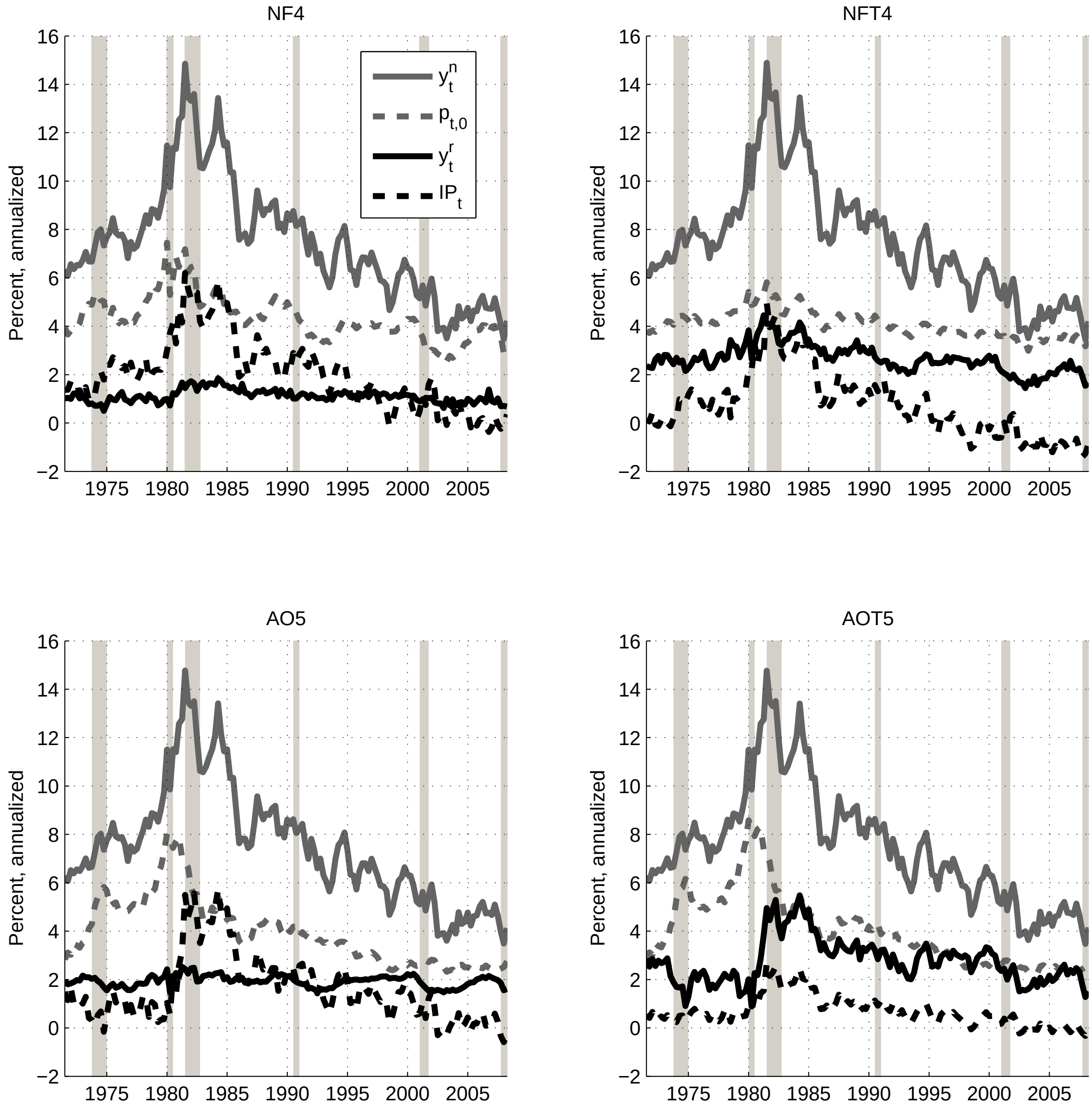

Fig. 4. The Fisher equation. The figure plots the time series of the ten-year nominal yield components in the decomposition

$$y_t^n(40) = y_t^r(40) + \overline{p}_{t,0}(40) + IP_t(40) + \text{convexity term},$$

where y_t^r is the real yield, $\overline{p}_{t,0}$ is the inflation expectation under the $\mathbb{P}$ measure, and IP_t is the inflation premium. The convexity term is tiny at 1 basis point and constant, so we do not display it. The notation for models is as follows: AO5 and AOT5 (All data, Objective expectations, five factors, with and without Treasury Inflation-Protected Securities, respectively); NF4 and NFT4 (No Forecasts, four factors, with and without Treasury Inflation-Protected Securities, respectively). The shaded regions show the National Bureau of Economic Research recessions.

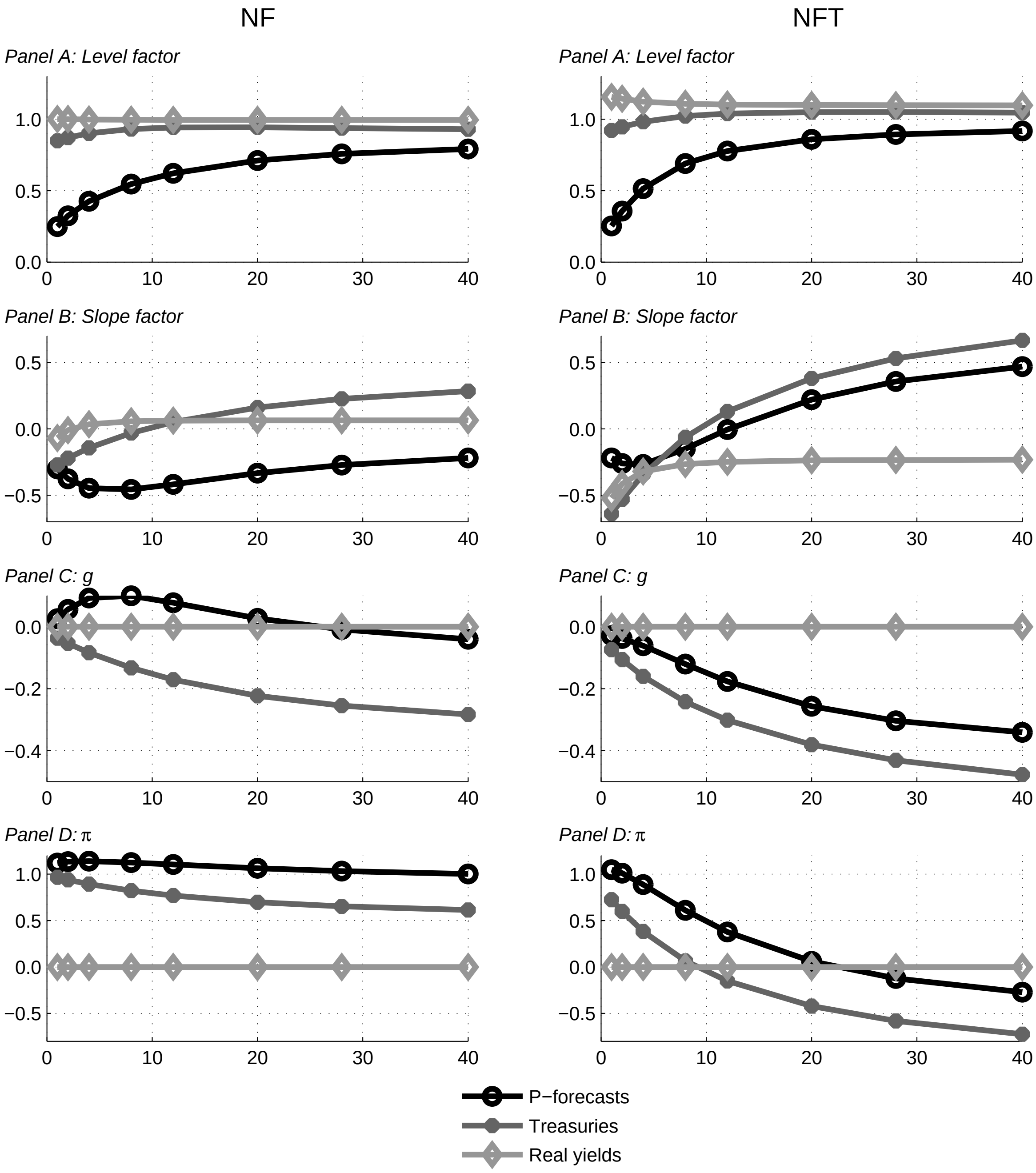

Fig. 5. Factor loadings for NF (No Forecasts) models. We plot factor loadings that are used by models NF4 and NFT4 (No Forecasts, four factors, with and without Treasury Inflation-Protected Securities, respectively) to establish model-based objective inflation expectations, nominal, and real yields for multiple horizons (zero to 40 quarters, x-axis). The loadings measure a response of a variable, in terms of a number of its standard deviations, to a 1 standard deviation change in a factor (y-axis).

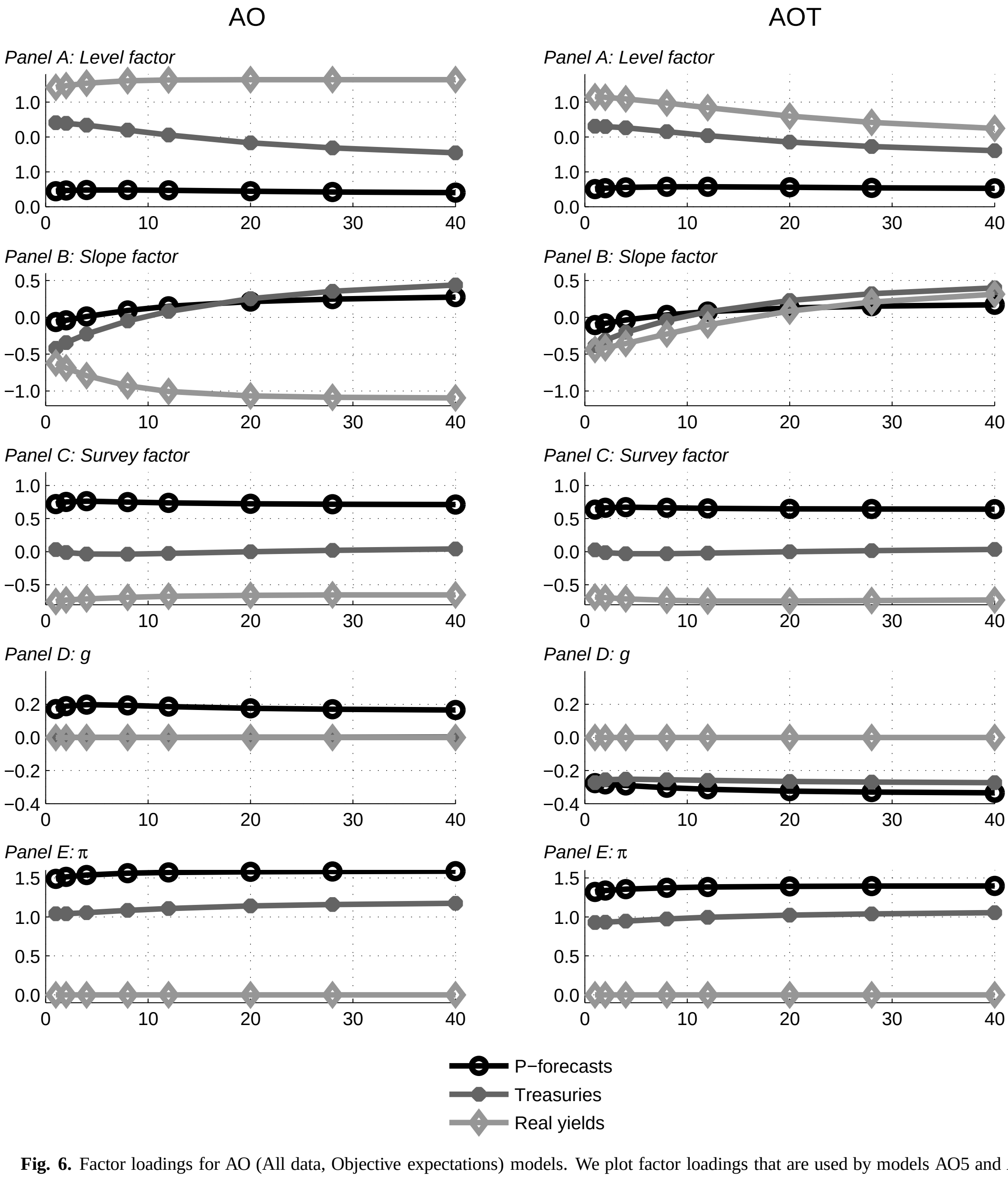

Fig. 6. Factor loadings for AO (All data, Objective expectations) models. We plot factor loadings that are used by models AO5 and AOT5 (All data, Objective expectations, five factors, with and without Treasury Inflation-Protected Securities, respectively) to establish model-based objective inflation expectations, nominal, and real yields for multiple horizons (zero to 40 quarters, x-axis). The loadings measure a response of a variable, in terms of a number of its standard deviations, to a 1 standard deviation change in a factor (y-axis).

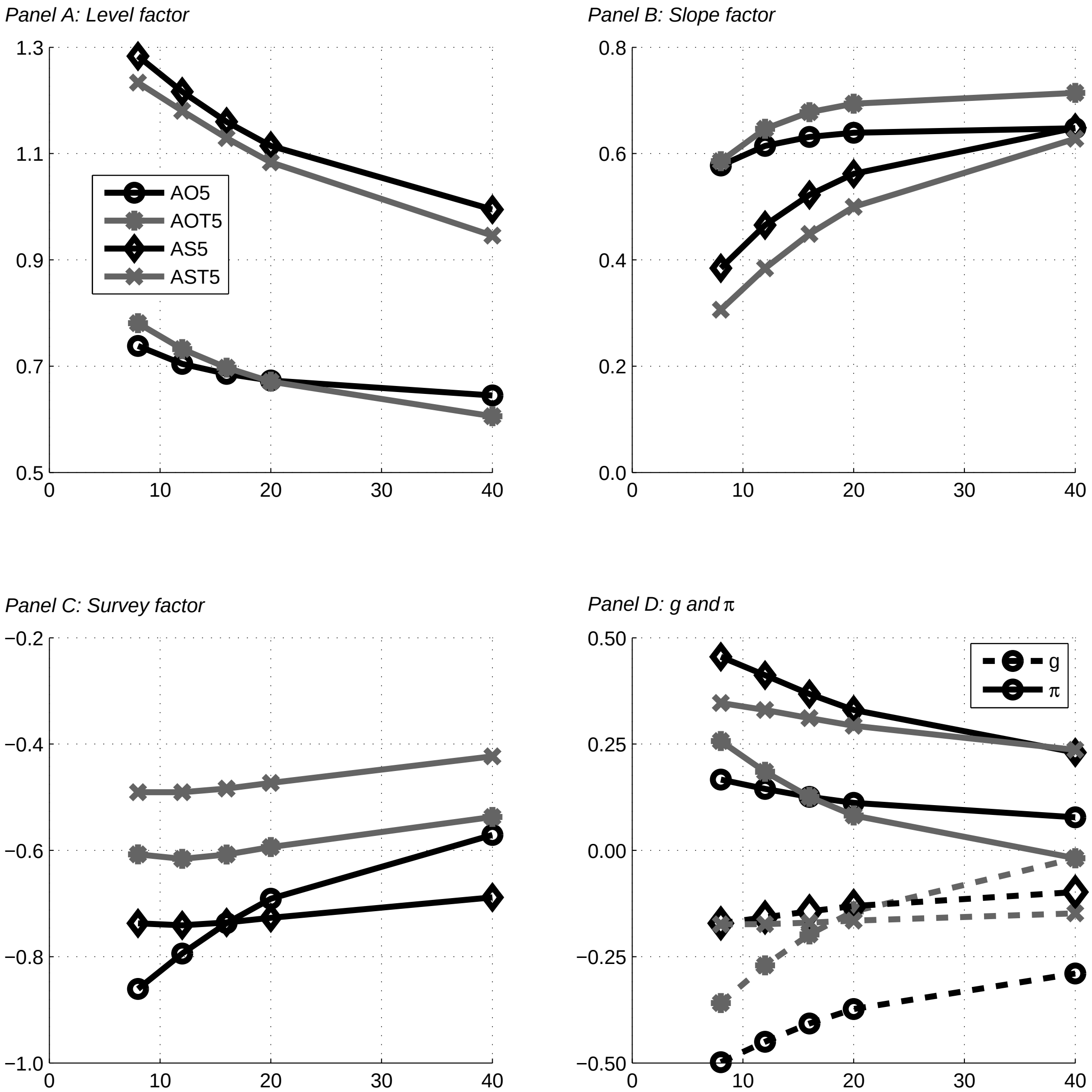

Fig. 7. Factor loadings for bond risk premia. We plot factor loadings that are used by models AO5 and AOT5 (All data, Objective expectations, five factors, with and without Treasury Inflation-Protected Securities, respectively) and AS5 and AST5 (All data, Subjective expectations, five factors, with and without Treasury Inflation-Protected Securities, respectively) to establish expected excess bond returns for multiple horizons (zero to 40 quarters, x-axis). The loadings measure a response of a variable, in terms of a number of its standard deviations, to a 1 standard deviation change in a factor (y-axis).

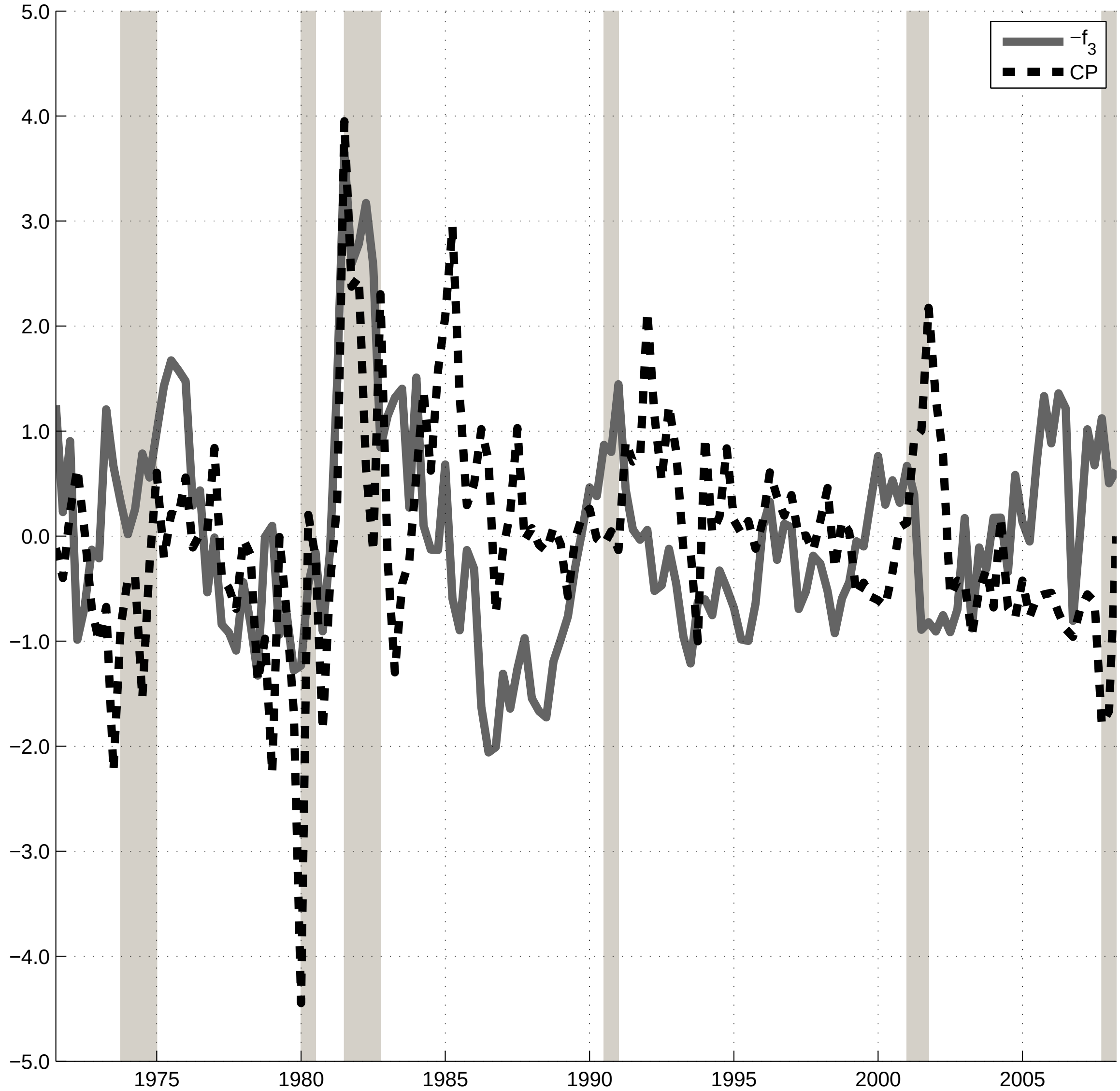

Fig. 8. The time series of the Cochrane and Piazzesi (2005) and the hidden factors. We display the time series of the standardized Cochrane and Piazzesi (2005) factor (CP) and the negative of the standardized hidden factor ($-f_3$) extracted from the AOT5 model (All data, Objective expectations, five factors, with Treasury Inflation-Protected Securities). By the sign convention, an increase in $-f_3$ corresponds to a decline in inflation expectations. The shaded regions show the National Bureau of Economic Research recessions.