

The Term Structure of Inflation Expectations*

Mikhail Chernov

London Business School and CEPR[†]

Philippe Mueller

Columbia Business School[‡]

First Draft: July 2005

This Revision: April 17, 2008

Abstract

We use evidence from the term structure of inflation expectations implicit in the nominal yields and survey forecasts of inflation to address the question of whether or not monetary policy is effective. We construct a model that accommodates forecasts over multiple horizons from multiple surveys and Treasury yields by allowing for differences between risk-neutral, subjective, and objective probability measures. We extract private sector expectations of inflation from this model and establish that they are driven by inflation, real activity and one latent factor, which is correlated with survey forecasts. We show that the interest rate responds to this “survey” factor. The inflation premium and out-of-sample estimates of the inflation long-run mean and persistence suggest that monetary policy became effective over time. As an implication, our model outperforms a standard macro-finance model in inflation and yield forecasting.

*We would like to thank Andrea Buraschi, Stephen Cecchetti, Albert Lee Chun, Darrell Duffie, Eric Ghysels, Marc Giannoni, Francisco Gomes, Don Kim, Sharon Kozicki, Lars Lochstoer, Igor Makarov, Stephen Schaefer, Antoinette Schoar, James Stock, Raman Uppal, Stephen Zeldes and participants at the Bank of England, Barclays Global Investors, Columbia Macro and Finance workshops, the 2007 Financial Econometrics Conference at Imperial College, Goethe University Frankfurt, LBS Finance workshop, McGill University, MIT Sloan, the 2007 NBER Summer Institute, PIMCO, the 2007 SAFE Conference on New Directions in Term Structure Modeling in Verona, State Street Global Advisors. We are especially grateful to Ruslan Bikbov for many insights on term structure modeling.

[†]Finance Area, Sussex Place, Regent’s Park, London NW1 4SA, UK, Phone: (020) 7000 8258, Email: mchernov@london.edu, Web: faculty.london.edu/mchernov

[‡]Division of Finance and Economics, 311 Uris Hall, 3022 Broadway, New York, NY 10027, USA, Email: pm2025@columbia.edu, Web: www.columbia.edu/~pm2025

1 Introduction

One of the most important conclusions in recent research in macroeconomics is that monetary policy matters for the development of the real economy. The main purpose of the paper is to provide yield-curve- and survey-based empirical evidence on how effective the US monetary policy is in handling the economy.

We approach the task by evaluating whether inflation expectations are anchored. The motivation for this line of thinking is Friedman’s (1968) American Economic Association address. He emphasizes that “[the] economic system will work best when producers and consumers, employers and employees, can proceed with full confidence that the average level of prices will behave in a known way in the future – preferably that it will be highly stable.” Friedman further argues that in the US, the effective monetary policy has to be conducted by targeting the price level.

In the rational expectations equilibrium (REE), a monetary policy that reacts to current inflation and output shocks is optimal (see, for example, Clarida, Gali, and Gertler, 1998 for a review). However, if private sector expectations deviate from rational expectations due to imperfect knowledge of the economy’s structure, the REE monetary policy will lead to an unstable economy. The main implication is that monetary policy should respond to either private expectations or their determinants in order to anchor these expectations (e.g., Evans and Honkapohja, 2003, Howitt, 1992, Preston, 2006, among others).

These theoretical developments suggest that the US monetary policy should be evaluated in two stages. First, it should be established whether there is any evidence that the monetary policy responds to the expectations of the private sector. Second, if this is the case, it should be determined whether expectations regarding inflation are anchored or became anchored over time.

There are two principal sources of observations of private sector expectations in the US. First, the expectations might be extracted from the nominal yield curve. However, one needs to disentangle the expectations from the inflation premium and from real yields.¹ Second, surveys, which represent direct estimates of future inflation by individuals, should be useful for learning about the expectations of a marginal economic agent.

However, it is not clear whether information that is derived from yields is identical to or, perhaps, conflicting with that derived from surveys. Indeed, the mechanisms that generate the expectations are different. In contrast to consensus survey forecasts, which average opinions of up to 50 participants, the expectations embedded in the bond prices are formed by thousands of traders who put hundreds of millions of dollars on the line. To date, the literature has been interested in comparing different methods of inflation forecasting, but, to the best of our knowledge, no one has reported whether observed yields and survey forecasts could be rationalized within one model. A joint model of

¹Inflation-linked bonds offer direct observation of real yields, but, until recently, they were not available in the US. However, inflation-linked bonds, even if available, are still contaminated by the inflation premium.

surveys and yields would help in establishing whether the two sources of inflation expectations are compatible by detecting whether there is a common set of factors that explain both.

With these motivations and considerations in mind, we proceed with our task in six steps. First, we build a model that is sufficiently flexible to allow for REE-based and private-expectations-based interest rate rules. Second, we measure the private sector expectations of inflation using this model by simultaneously incorporating information from survey-based inflation forecasts and nominal yields. Third, we establish the determinants of inflation expectations. Fourth, we check whether the interest rate rule implicit in our model responds to these determinants. Fifth, we investigate whether monetary policy is effective, that is, whether inflation expectations are anchored. Sixth, we evaluate the implications of anchoring for out-of-sample forecasting.

Inflation forecasts are available from various surveys and at various horizons (see Figure 1). However, each survey is different in terms of the composition of forecasters, frequency of observations, and forecast horizons. We propose a model of the term structure of inflation forecasts that 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 sets of information or the forecasting objective functions of the surveys. We integrate our model of inflation forecasts together with a no-arbitrage macro-finance term structure model.

In our model, the yields and forecasts are driven by two observed macro variables (real activity and inflation) and by two latent variables. The joint dynamics of these four variables determine inflation expectations at any horizon under the objective probability measure. The yields reflect expectations of inflation under the risk-neutral probability measure. Because forecast surveys may differ from each other systematically, we model the respective expectations as those of heterogeneous agents. This implies that for each survey, the forecasts are computed under a subjective measure that is potentially different from both objective and risk-neutral measures. This flexibility in the modelling allows for state-dependent deviations of the survey forecasts from the optimal objective-measure-based expectations.

For our empirical analysis, we use a panel of eight yields ranging from three months to ten years, with inflation and GDP observed at a quarterly frequency from 1970 to 2004. We combine these with a total of twenty inflation forecasts from the four surveys depicted in Figure 1. The forecasts from these surveys are available for various horizons and at various frequencies.

Allowing for the difference between survey-specific subjective and objective probability measures, we document a rich pattern of survey biases. The surveys overpredict inflation the most and disagree the least at the end of recessions. A model that restricts subjective measures to coincide with the objective one cannot be distinguished statistically from a richer model. Further, the restricted version of the model produces more reasonable inflation expectations than the standard version of a macro-finance model that uses yields only. Thus, we conclude that private sector expectations have to be

extracted from both yields and surveys.

Inflation expectations that are computed from our leading model use inflation and real activity, and load similarly on one of the two latent factors across all forecasting horizons. The second latent factor is barely used in the construction of these expectations. Further, the first latent factor can be explained by a linear combination of contemporaneous real activity, inflation, and one-year inflation forecasts. These properties justify the interpretation of this factor as a “survey” factor. This factor affects the spot interest rate, which suggests that implicit monetary policy responds to the information reflected in this factor.

We use the estimated inflation expectations to infer inflation premia from nominal yields via the Fisher equation. The ten-year inflation premium declines from six to zero per cent during the post-monetary-experiment period. This decline suggests that long-run inflation expectations became more stable over time. Further, we reestimate our model every quarter and find that the long-run expectations have declined over time from 6% to 2%. The inflation persistence declined and the term structure of inflation expectations became flat over time. This evidence suggests that monetary policy became better anchored.

One implication of anchored inflation expectations is that it should be easier to forecast inflation and yields. Consistent with this prediction, we find that the model that incorporates both yields and surveys dominates in out-of-sample forecasting of both inflation and yields. These results lead us to conclude that information in surveys is extremely important for establishing the links between inflation expectations and yields.

2 Selective empirical literature review

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

The empirical side of the literature on no-arbitrage term structure explores the behaviour of state variables and risk premia in the context of no-arbitrage models. The main focus is to explain and summarize parsimoniously the relationships that are of interest and, therefore, 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 the data that reflects inflation expectations, such as inflation-indexed bonds (Evans (1998)), or survey forecasts of inflation (Chun (2005), Pennacchi (1991)).² The typical research questions that arise are: (i) Does inflation help to explain the yields? (Ang and Piazzesi (2003), Bikbov and Chernov (2005), Duffee

²A recent work by Kim and Orphanides (2005) is related to this literature in methodology, but not in research questions. These authors study the traditional latent factor models and use survey forecasts of yields in addition to contemporaneous yields.

(2005)); (ii) Does inflation help to explain the risk premia? (Ang, Dong, and Piazzesi (2004), Bikbov and Chernov (2005), Duffee (2006)); (iii) How do real interest rates behave? (Ang, Bekaert, and Wei (2004), Evans (1998), Pennacchi (1991)); (iv) What is the size and the nature of variation of the inflation premium? (Ang, Bekaert, and Wei (2004), Evans (1998), Hordahl and Tristani (2007)).

Part of the literature on macroeconomic forecasting focuses on predicting 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, a lot of work focuses on out-of-sample performance. Conceptually, the approach suffers from the following drawback. As any no-arbitrage model would imply, yields simultaneously reflect the market's expectations of the state variables, such as inflation and risk premia. There is formidable evidence that the risk premia are time-varying. As a result, it is virtually impossible to disentangle the yield components in a regression framework. Depending on the variability of the risk premia, yields might introduce a lot 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.

Finally, Ang, Bekaert, and Wei (2007) combine the views contained in the reviewed literature by running a horse-race between the different ways of forecasting one-year inflation.³ 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.

Our work, while having a different focus, is related to the following four 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 (2007) build a no-arbitrage model and estimate it using the nominal yield curve, TIPS, Blue Chip forecasts of one-quarter yields, and one-year and ten-year SPF forecasts of inflation from 1999 to 2007. They document the importance of using TIPS for accurate predictions of inflation. In a similar spirit, Joyce, Lildholdt, and Sorensen (2007) exploit information from the UK real and nominal bonds along with Consensus forecasters' expectations of average inflation from five to 10 years ahead from 1992 to 2007 to construct inflation forecasts and inflation risk premia. Piazzesi and

³They also consider the Phillips curve and pure time-series models. However, discussion of these lies beyond the scope of our paper.

Schneider (2008) focus on using an affine term-structure model to construct a measure of subjective bond risk premia, which they derive from survey data, and specifying a structural model that can explain these premia.

3 The Model

In this section, we develop the theoretical underpinnings of our approach. In section 3.1, we describe our assumptions. Section 3.2 discusses the advantages of our approach and contrasts it with previous methodologies.

3.1 The Setup

3.1.1 State variables

Following Ang and Piazzesi (2003) (AP henceforth), we assume that the state of the economy is captured by the vector $z_t = (m'_t, x'_t)'$. In particular, the vector of macroeconomic variables m_t is equal to $(g_t, \pi_t)'$, where g_t and π_t are quarterly GDP and the inflation rate, respectively. We define the inflation rate π_t as

$$\pi_t = \log \left(\frac{P_t}{P_{t-1}} \right), \quad (3.1)$$

where P_t is the price level. The remaining factors x_t are latent. The latent factors may represent the lags of m_t , other macro variables, or other unknown variables. Importantly, the vector z_t fully reflects all available information at time t , so, for instance, one need not consider lags of x_t .

We assume that the state vector z_t follows a VAR(1) process under the objective probability measure \mathbb{P} :

$$z_t = \mu + \Phi z_{t-1} + \Sigma \epsilon_t \quad (3.2)$$

$$= \begin{bmatrix} \mu^m \\ \mu^x \end{bmatrix} + \begin{bmatrix} \Phi^{mm} & \Phi^{mx} \\ \Phi^{xm} & \Phi^{xx} \end{bmatrix} \begin{bmatrix} m_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \Sigma^{mm} & \Sigma^{mx} \\ \Sigma^{xm} & \Sigma^{xx} \end{bmatrix} \begin{bmatrix} \epsilon_t^m \\ \epsilon_t^x \end{bmatrix}, \quad (3.3)$$

where $\epsilon_t \sim N(0, I)$. We denote the vector of parameters controlling the dynamics of state by $\Theta = (\mu, \Phi, \Sigma)$. The block representation will be useful for later discussions.

In particular, the state dynamics imply that the objective expectation, or \mathbb{P} -expectation, of the future state variables is a linear function of the current state variables:⁴

$$E_t(z_{t+\tau}) = \Psi^\tau \mu + \Phi^\tau z_t, \quad (3.4)$$

⁴In this paper, we use the terms \mathbb{P} -expectation, marginal expectation, optimal expectation, objective expectation, and private sector expectation interchangeably.

where

$$\Psi^\tau \triangleq \sum_{k=0}^{\tau-1} \Phi^k = (I - \Phi)^{-1} (I - \Phi^\tau). \quad (3.5)$$

3.1.2 Spot interest rate and yields

It is customary in the literature on term structure to specify the interest rate as a linear function of the state variables (e.g., Dai and Singleton (2000), or AP):

$$r_t = \delta_0 + \delta'_z z_t = \delta_0 + \delta'_m m_t + \delta'_x x_t. \quad (3.6)$$

The specification of the state variables (3.2), combined with the interest rate specification in (3.6), allows us to complete the usual affine no-arbitrage framework by specifying the stochastic discount factor ξ_t , or, equivalently, the transformation to the risk-neutral probability \mathbb{Q} :

$$\log \xi_t = -r_{t-1} - \frac{1}{2} \Lambda'_{t-1} \Lambda_{t-1} - \Lambda_{t-1} \epsilon_t, \quad (3.7)$$

where the market prices of risk follow the essentially-affine specification (Duffee, 2002):

$$\Lambda_t = \Lambda_0 + \Lambda_z z_t. \quad (3.8)$$

Therefore, yields on zero-coupon bonds are linear in the state variables,

$$\begin{aligned} y_t(\tau) &= -\frac{1}{\tau} \log E_t \left(\prod_{s=t+1}^{t+\tau} \xi_s \right) = a^{\mathbb{Q}}(\tau) + b^{\mathbb{Q}}(\tau)' z_t \\ &\triangleq \underbrace{a^{\mathbb{P}}(\tau) + b^{\mathbb{P}}(\tau)' z_t}_{\text{Short rate expectations}} + \underbrace{a^{TP}(\tau) + b^{TP}(\tau)' z_t}_{\text{Term premium}} \end{aligned} \quad (3.9)$$

where τ is the respective maturity, and $a^{\mathbb{Q}}$ and $b^{\mathbb{Q}}$ solve recursive equations with boundary conditions $a^{\mathbb{Q}}(1) = \delta_0$, and $b^{\mathbb{Q}}(1) = \delta_z$ (see, e.g., Backus, Foresi, and Telmer, 1999). In particular, the one-quarter yield coincides with the short rate, $y_t(1) = r_t$. The last line breaks down the yields into the expectations of the future short rates and the term premium. The former component is equal to the usual factor loadings computed under the assumption of zero market prices of risk.

3.1.3 Heterogeneous Forecasters

We assume that the world is populated by agents who have heterogenous inflation expectations. The heterogeneity might arise from the differential information, or different loss functions used in the forecasts. We treat a consensus forecast from each survey as a forecast of one such agent. A reduced-form representation of this assumption is an equivalent subjective probability measure \mathbb{P}^i ,

that corresponds to beliefs of an agent i .⁵ 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 (3.8):

$$\log \xi_t^i = -\frac{1}{2} \Lambda_{t-1}^{i'} \Lambda_{t-1}^i - \Lambda_{t-1}^i \epsilon_t, \quad (3.10)$$

where Λ_t^i follows the essentially-affine specification

$$\Lambda_t^i = \Lambda_0^i + \Lambda_z^i z_t. \quad (3.11)$$

We assume that the agents diverge only with respect to their beliefs about inflation. Therefore, Λ_t^i should be restricted. We assume that only the second element of the vector Λ_0^i , $\lambda_{0,2}^i$, and the second line of the matrix Λ_z^i , $\lambda_{z,2}^i$, are not equal to zero.⁶ Because \mathbb{P}^i is not a pricing measure, Λ_t^i do not represent prices of risk. These quantities reflect the deviations of individual expectations from the \mathbb{P} -expectations. For this reason, we will refer to Λ_t^i as biases. The resulting dynamics of the state variables corresponding to the individual probability measure \mathbb{P}^i will be denoted by

$$z_t = \mu_i + \Phi_i z_{t-1} + \Sigma \epsilon_t^i, \quad (3.12)$$

3.1.4 Survey Forecasts

A forecaster from group i (a participant of a forecast survey) has the following estimate of the expected state:

$$E_t^i(z_{t+\tau}) = \Psi_i^\tau \mu_i + \Phi_i^\tau z_t, \quad (3.13)$$

where, by analogy with (3.5),

$$\Psi_i^\tau \triangleq (I - \Phi_i)^{-1} (I - \Phi_i^\tau). \quad (3.14)$$

Because we assumed heterogeneity in expectations only with respect to inflation, the expression (3.4) will differ from the \mathbb{P} -expectation of state only in

$$E_t^i(\pi_{t+\tau}) = e_2' (\Psi_i^\tau \mu_i + \Phi_i^\tau z_t), \quad (3.15)$$

where e_2 is a vector of zeros with a one in the second position.

⁵This modelling approach could be justified in the framework of such models as Detemple and Murthy (1994), Dumas, Kurshev, and Uppal (2005), Harrison and Kreps (1978), Scheinkman and Xiong (2003), among others. See Basak (2005) for a review. Appendix A provides a simple model that motivates our assumptions. Alternatively, Patton and Timmermann (2007) show that asymmetric and different loss functions of forecasters imply unbiased forecasts with iid errors under subjective probability measures.

⁶We do not need to assume that beliefs differ in inflation only. However, because we will be using forecasts of inflation only, most of the parameters controlling Λ_t^i will not be identified. Therefore, the described restrictions would have to be imposed in any case.

In practice, most forecasts represent estimates of inflation over some period (a quarter or a year) beginning at a future date. This amounts to forecasts of averages of quarterly inflation. The forecasts of averages will be denoted and computed as follows:

$$\bar{p}_{t,s}^i(\tau) \triangleq E_t^i(\bar{\pi}_{t,s,\tau}) \triangleq E_t^i\left(\frac{1}{\tau} \sum_{j=1}^{\tau} \pi_{t+s+j}\right). \quad (3.16)$$

Therefore,

$$\begin{aligned} \bar{p}_{t,s}^i(\tau) &= \frac{1}{\tau} \sum_{j=1}^{\tau} E_t^i(E_{t+s}^i(\pi_{t+s+j})) = \frac{1}{\tau} \sum_{j=1}^{\tau} E_t^i(e_2'(\Psi_i^j \mu_i + \Phi_i^j z_{t+s})) \\ &= \frac{e_2'}{\tau} \left(\sum_{j=1}^{\tau} \Psi_i^j + \Psi_i^{\tau} \Phi_i \Psi_i^s \right) \mu_i + \frac{e_2'}{\tau} \Psi_i^{\tau} \Phi_i^{s+1} z_t \triangleq a^i(s, \tau) + b^i(s, \tau)' z_t \\ &\triangleq \underbrace{a^{\mathbb{P}}(s, \tau) + b^{\mathbb{P}}(s, \tau)' z_t}_{\mathbb{P}\text{-forecasts}} + \underbrace{a^{TBi}(s, \tau) + b^{TBi}(s, \tau)' z_t}_{\text{Term bias}}. \end{aligned} \quad (3.17)$$

The last line decomposes the forecasts into the \mathbb{P} -expectations of the future inflation and the survey-specific term biases. The former component is equal to the usual factor loadings computed under the assumption of zero biases Λ_t^i .

3.2 Monetary policy and inflation expectations

Following the insight of Ang, Dong, and Piazzesi (2004) (ADP henceforth), one can use (3.4) or (3.13) to rewrite the interest rate rule (3.6) as a linear function of objective or subjective inflation expectations. In a reduced-form model, such as ours, one generally cannot identify whether subjective or objective expectations and which expectations horizon should be used because of this equivalence between the different representations. Using the direct observations of subjective expectations in conjunction with yields, which reflect objective expectations, should help in solving the identification problem.

To aid the interpretation of our empirical results, we rewrite the spot interest rate in a form that is similar to the theoretical work on optimal monetary policy (e.g., Evans and Honkapohja, 2003). First, we use the ADP approach to express the spot interest rate as a function of one-period inflation expectations from survey i (omitted i corresponds to objective expectations):

$$r_t = \gamma_0^i + \gamma_g^i g_t + \gamma_{\pi}^i E_t^i(\pi_{t+1}) + \gamma_x^i x_t, \quad (3.18)$$

where the coefficients γ are related to the coefficients δ in the original rule (3.6) via

$$\delta_0 = \gamma_0^i + \gamma_{\pi}^i e_2' \mu_i, \quad (3.19)$$

$$\delta_z = \gamma_g^i e_1 + \Phi_i' \gamma_{\pi}^i e_2 + \gamma_{x,1}^i e_3 + \gamma_{x,2}^i e_4, \quad (3.20)$$

where e_i is a vector of zeros with a one in position i . A similar representation can be obtained for any forecasting horizon τ (see ADP for details).

Further, following Bikbov and Chernov (2005), we break down each latent variable x into a component explained by inflation and real activity, and into a residual f that is orthogonal to the entire history $M_t = \{m_t, m_{t-1}, \dots, m_0\}$ of the two macro variables:

$$x_t = \hat{x}(M_t) + f_t, \quad (3.21)$$

$$\hat{x}(M_t) = c(\Theta) + \sum_{j=0}^t c_{t-j}(\Theta)m_{t-j} \equiv c(\Theta) + c_t(\Theta)m_t + c(\Theta, L)m_{t-1}, \quad (3.22)$$

where the matrices c are functions of parameters Θ that control the dynamics of the state variables and $c(\cdot, L)$ emphasizes the lag-polynomial structure of the expression.⁷ In particular, the results in Bikbov and Chernov (2005) imply that under the standard identification assumptions, the joint dynamics of m_t and f_t can be described as

$$\begin{bmatrix} m_t \\ f_t \end{bmatrix} = \begin{bmatrix} \mu^m \\ 0 \end{bmatrix} + \begin{bmatrix} \Phi^{mm}(L) & \Phi^{mx} \\ 0 & \Phi^{ff} \end{bmatrix} \begin{bmatrix} m_{t-1} \\ f_{t-1} \end{bmatrix} + \begin{bmatrix} \Sigma^{mm} & 0 \\ \Sigma^{fm} & \Sigma^{xx} \end{bmatrix} \begin{bmatrix} \epsilon_t^m \\ \epsilon_t^x \end{bmatrix}. \quad (3.23)$$

This is a stylized representation. In order to focus on the most important issues, we do not reproduce full expressions for some matrices (e.g., for the lag-polynomial matrix $\Phi^{mm}(L)$ or the covariance matrix Σ^{fm}). Therefore, we can think of an economy that is driven by output, inflation, and persistent shocks. This is consistent with typical New-Keynesian models of the private sector (see Clarida, Gali, and Gertler, 1998 for details).

Substituting (3.21) and (3.22) into (3.18) we obtain

$$r_t = \tilde{\gamma}_0^i + \tilde{\gamma}_g^i g_t + \tilde{\gamma}_\pi^i \pi_t + \gamma_\pi^i E_t^i(\pi_{t+1}) + \tilde{\gamma}_x^{i'}(L)m_{t-1} + \gamma_x^{i'} f_t, \quad (3.24)$$

where coefficients $\tilde{\gamma}$ represent a convolution of the coefficients γ in (3.18) and $c(\Theta)$ in (3.22). The combination of this particular form of the interest rate rule and state dynamics in (3.23) is interesting because it nests, in a reduced form, the private sector dynamics and optimal rule prescription in a world in which the private sector deviates from rational expectations and the policy maker is learning about parameters that describe the economy (see, e.g. Evans and Honkapohja, 2003, or Primiceri, 2006). In this world, monetary policy should respond to the private sector expectations of inflation or to their determinants.

As the expression (3.24) demonstrates, our framework is sufficiently flexible to allow for this possibility. Moreover, by taking the model to the data, we can establish whether we should be using a specific survey (the value of i) in this expression or some combination of surveys to compute the private sector expectations. Furthermore, by focusing on expectations $E_t^i(\pi_{t+1})$ in (3.24), we can

⁷Appendix B provides the details of the procedure.

evaluate how the various data sources (macro variables, yields, or survey forecasts) contribute to the determination of these expectations.

Having established the appropriate way to determine inflation expectations, we can use our model to determine the effects of monetary policy. Our framework offers three main ways of addressing this topic. First, the macroeconomic literature on optimal monetary policy suggests that if the policy responds correctly to the expectations of the private sector, expectations about long-term inflation will be anchored. Therefore, we can investigate the properties of long-term forecasts and whether these properties have changed over time. Second, another aspect of the anchoring of expectations is how fast long-term inflation expectations adjust to transient inflation shocks. In other words, expectations are anchored better when inflation is less persistent and reverts back to its long-run mean soon after a shock. Third, we can use the finance side of our model to extract inflation risk premia. If long-term inflation expectations are anchored, the long-term inflation premium should be small.

The inflation premium (ignoring convexity terms) can be extracted from a generalized version of the Fisher equation:

$$y_t(\tau) = y_t^R(\tau) + \bar{p}_{t,0}(\tau) + IP_t(\tau), \quad (3.25)$$

where y^R refers to real yield, expected inflation \bar{p} under the \mathbb{P} measure is defined in (3.16), and IP denotes the inflation premium. Because inflation is one of the observed factors, we can use our model to identify the real yields and inflation premium by observing inflation and nominal yields (see, e.g., Ang, Bekaert, and Wei, 2004). Specifically, the real discount factor is

$$\log \xi_t^R = -r_{t-1} + \pi_t - \frac{1}{2} \Lambda'_{t-1} \Lambda_{t-1} - \Lambda_{t-1} \epsilon_t, \quad (3.26)$$

This allows us to compute real yields, $y_t^R(\tau)$, by analogy with the nominal yields (3.9). Before that, we establish the appropriate measure of \mathbb{P} -expectations. As a result, we will be able to obtain an accurate estimate of the inflation premium.

Finally, if inflation expectations are stable, we should be able to forecast inflation more accurately. We check this by conducting an out-of-sample forecasting analysis.

4 Empirical Approach

4.1 Data

We use three types of data in this paper. Direct measures of macro variables (inflation and GDP) represent two observable state variables in our model. Treasury yields and survey forecasts are the observable data that help us to learn about the model parameters and latent state variables. Using

these data, we construct a quarterly panel from 1970 to 2004. We describe the data details in appendix C.

4.2 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.⁸ We place estimation errors on all yields and survey forecasts so that the latent factors are not associated with pre-specified observables. We assume that the macro variables are observed without error.

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

$$y_t(\tau) = a^{\mathbb{Q}}(\tau) + b^{\mathbb{Q}}(\tau)'z_t + \omega_t, \quad (4.1)$$

$$\bar{p}_{t,s}^i(\tau) = a^i(s, \tau) + b^i(s, \tau)'z_t + \chi_{t,s}^i(\tau). \quad (4.2)$$

where y represents the yields of maturities from one month to ten years, and \bar{p}^i represents all the forecasts made in all the surveys (as described above).

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 forecast 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 for addressing our questions. Therefore, for each survey, we restrict the errors in the measurement of long-term forecasts to be no less than those of short-term forecasts and those of yields.⁹ While such a flexible specification runs the risk of overparameterization, we decided to use it because of diverging properties of forecasts. As we can see from the description above, the forecasts within one survey but at different horizons might be available for different data spans, different quarters, and different frequencies.

⁸Other 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 Ait-Sahalia and Kimmel (2002), the simulated maximum likelihood of Brandt and He (2002), and the Bayesian MCMC of Collin-Dufresne, Goldstein, and Jones (2003).

⁹Long-term forecasts are available for LS and SPF only.

4.3 Additional Considerations

In this section, we review briefly additional matters that we took into account when estimating.

Missing observations

We determine the main span of our dataset by the availability of the yield data. Thus, we end up with 140 quarters from 1970 to 2004. However, as noted above, the result of this choice is that many forecast observations are missing. There are two reasons for this: (i) some surveys were only available from a date later than 1970, and some are available less frequently than quarterly. Regardless, missing observations are not problematic. They can be handled easily in the Kalman filter framework. We simply do not update, or partially update, the state vector when the observations are missing. This procedure is automatic when one uses a new forecast vector $\bar{\varphi}$ that is an old forecast vector \bar{p} scaled by a matrix that has zeros in place of missing observations, and ones in place of the available ones (see Harvey (1989)).¹⁰

Number of factors and identification

We largely follow the setup of Bikbov and Chernov (2005) in selecting the number of latent factors and the scheme for identifying parameters. We decided to use a four-factor model (two latent factors) because it is the smallest model that can reasonably fit the yield curve. We were hesitant to increase the number of states because we wanted to make sure that the same latent variables drive both yields and forecasts, so that we could learn about these latent states from a rich set of observations.

Risk premia and survey-specific biases

The elaborate risk-premia specification combined with forecast-specific biases leaves one with a concern about overfitting. We follow Bikbov and Chernov (2005) and augment the standard log-likelihood function, \mathcal{L} , with a penalization term that is proportional to the variation of the term premium in (3.9):

$$\begin{aligned} \mathcal{L}_p &= \mathcal{L} - \frac{1}{2\sigma_p^2} \sum_{\tau} (a^{TP}(\tau))^2 + b^{TP}(\tau)' \cdot \text{Diag}(\text{var}(z_t)) \cdot b^{TP}(\tau) \\ &\quad - \frac{1}{2\sigma_p^2} \sum_{i,s,\tau} (a^{TBi}(s,\tau))^2 + b^{TBi}(s,\tau)' \cdot \text{Diag}(\text{var}(z_t)) \cdot b^{TBi}(s,\tau), \end{aligned} \quad (4.3)$$

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 biases are equal to zero, the term premium

¹⁰This approach was also used by Kim and Orphanides (2005), Kozicki and Tinsley (2006), Lu and Wu (2005), and Pennacchi (1991).

and term biases will be equal to zero as well. Therefore, \mathcal{L}_p imposes an extra burden on the model to use the risk premia and survey biases 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.

Optimization

We need to estimate 77 parameters in our model. We have a large cross-section of observations, which should help in pinning these parameters down. However, with a time-series of 140 observations there remains a concern 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 multi-dimensional 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 twenty thousand points as starting values. We optimize alternating between simplex and SQP algorithms and eliminating half of the likelihoods at each stage.

5 Findings

We estimate six versions of our model. The full model that uses the survey and yield data is labelled AS (All data, Subjective expectations). A model that uses all the data, but restricts subjective measures to coincide with objective measure, is labelled AO (All data, Objective expectations). A model that uses only yields for estimation is referred to as 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 forecasts and label it OF (Only Forecasts). In this implementation, the risk-neutral measure cannot be estimated. In order to implement various robustness checks we have estimated two more models. ASR (All data, Subjective expectations, Restricted) restricts the values of \mathbb{P} -measure parameters to the values estimated in AO and allows other parameters (\mathbb{P}^i and \mathbb{Q}) to be free. OFO (Only Forecasts, Objective expectations) is a version of OF that restricts subjective measures to coincide with objective measure.¹¹

We do not report the technical details of the estimation results, such as parameter values, tests of their statistical significance, and measures of fit. There are too many parameters to discuss and most of them are difficult to interpret. We will comment on the relevant aspects of the fit as we discuss the model implications.

¹¹In spirit, OFO is similar to the time-series model of Kozicki and Tinsley (2006).

5.1 Measuring inflation expectations

First, we wish to establish a measure of private sector expectations about inflation. Specifically, in the context of our model, we want to understand whether a subjective forecast from a specific survey i , or objective forecast, or some combination thereof, could represent such expectations. We compare forecasts from the different surveys to the \mathbb{P} -measure expectations. Then we evaluate the importance of the heterogeneous agent framework for forecasting purposes.

5.1.1 The term structure of forecast biases

In Figure 2 we plot the term structure of term biases – the deviations of the survey forecasts from the optimal forecasts by maturity (see formula (3.17)). We consider term biases of survey forecasts made today over the horizon matching the maturity of the bonds in our sample, that is:

$$TB_t^i(\tau) = E_t^i(\pi_{t+\tau}) - E_t(\pi_{t+\tau}), \quad \tau = 1, 2, 4, 8, 12, 20, 28, 40 \text{ qtr} \quad (5.1)$$

We use the AS model for the computations. We report the unconditional term structure, which measures the average bias for each horizon, and the time-series of conditional biases for the horizons of one ($\tau = 4$) and 10 ($\tau = 40$) years.

The first panel shows the unconditional term structure of the survey biases. All the surveys exhibit a bias that increases as the horizon lengthens. This is to be expected, because most of the forecasts are concentrated at the short horizons and longer-term horizons are available for a shorter time span. BCEI is uniformly worse and varies from -0.7% at the three-month horizon to -1.9% at the 10-year horizon.

Turning to the conditional biases in Figure 2, we see that there is virtually no difference between the different surveys. If anything, MCS seems to be more in line with the optimal forecast at longer horizons. The shaded regions represent the NBER dated recessions and help to identify the reasons for the biases. The peak of inflation overestimation generally coincides with the end of a recession. This property of biases is consistent with the view that monetary policy tends to be accommodative during recessions, thereby increasing fears of inflation.

The extant analysis (e.g., Thomas, 1999) finds, on the basis of the split-sample regressions, that surveys overestimate inflation when it is low and vice versa. Our conditional framework allows us to perform a more refined analysis period by period instead of comparing the pre- and post-Volcker samples. We check the conclusion from the regression-based work by plotting the one-year term bias against the demeaned inflation. Indeed, we observe a general pattern that suggests that inflation is underpredicted during periods of high inflation. This impression is particularly strong because the two subperiods previously considered in the literature (pre- and post-Volcker) could be characterized as the periods of “great inflation” and “disinflation,” respectively. Nonetheless, this relationship between underprediction and inflation can be traced even during the disinflation period.

Business cycles and inflation are, of course, related. Thus, it is not surprising that we find links from the overestimation of inflation to both. Which link is more informative about the behaviour of forecasters? It appears that the business cycle indicator is more robust: the end of a recession always coincides with a local maximum of the bias, while the end of an expansion always coincides with a local minimum of the bias. In contrast, inflation could peak after or before the bias' minimum was reached. Sometimes, inflation does not move much, as in the mid-1990s, but biases have very strong patterns.

We characterize the disagreement between the various surveys by computing the difference between the largest and the smallest bias:

$$D_t(\tau) = \max_i (TB_t^i(\tau)) - \min_i (TB_t^i(\tau)) \tag{5.2}$$

The time series of the one- and 10-year disagreement are presented in Figure 2 as well.

There is greater disagreement between the surveys at the one-year horizon than at the 10-year horizon. The ranges are from 0 to 2 percent and from 0.2 to 3.5 percent, respectively. While overprediction is countercyclical, the disagreement is procyclical. Consistent with our reasoning about the overprediction of inflation, it appears that, at the end of a recession, all forecasters tend to agree that future inflation will be high.

5.1.2 Subjective and objective expectations

We have demonstrated that our model is sufficiently rich to describe the evolution of subjective and objective inflation expectations. However, it is not clear which type should be used as a measure of the private sector expectations. To address this question, we use information from the other three models that we have estimated (AO, NF, and OF). These versions of our main model differ in the restrictions that they place on parameters and the data used for estimation. These differences should help us to understand the primary sources of expectations.

We start by evaluating the model-implied mean absolute errors (MAE), which are reported in Table 1. Overall, the models match the data well; depending on the model, the average MAE ranges from 10 to 20 basis points for surveys and from 10 to 30 basis points for yields. As one would expect, the model clearly trades off the fit of surveys and yields. OF and NF provide a lower bound on the errors. The survey MAE in AS are comparable to those in OF. The biggest difference between the two models is for the BCEI longer maturity forecasts. Clearly, AS sacrifices the yield fit in favour of the survey fit, because the differences between the NF and AS yield errors are larger than those between the OF and AS survey errors.

AO should have more difficulty fitting the surveys than AS because we restrict 20 parameters that control the subjective probability measures of the four surveys to the values of their objective measure counterparts. However, in practice, we see small differences.

The largest discrepancy between AO and AS is in how they fit the MCS forecast. In addition, regardless of the model, the MCS error is noticeably large. The type of bias implicit in MCS is likely to be different from other surveys, either because consumers do not predict CPI explicitly or because an average consumer forms expectations differently from an average professional.¹² It appears that MCS does contain useful information. However, some features of the survey either have to be ignored via the error term, while some could be accommodated via the subjective probability measure. Our model can be extended to fit MCS better, but this improvement is not likely to affect any other surveys or yields.

Because AO is so similar to the larger AS in terms of MAE, it is natural to wonder whether the richness provided by the subjective probability measures is required, either (i) for a good model fit or (ii) inflation forecasting. The first issue can be addressed via the likelihood ratio (LR) test because AO (the null hypothesis) is nested in AS (the alternative hypothesis). The likelihood-ratio test statistic is equal to 1.5, which is insignificant.¹³ Clearly, there is no statistical difference between the two models.

Our analysis of the term biases suggests that, despite the lack of statistical significance, there could be economic differences between the two models. We evaluate economic differences using a metric that is important for our study: objective inflation expectations. Figure 3 plots the time series of inflation expectations that are computed from AO, OF, and NF. The AS-based expectations are similar to those from NF and are omitted to avoid clutter.¹⁴ We also plot, as a benchmark, the time series of realized inflation and the 10-year inflation expectations that are implied by TIPS.¹⁵

Despite the similarity in the survey fit, the model-implied expectations diverge quite dramatically, especially at the 10-year horizon. The NF expectations float around 5% in the 1970s, during the great inflation and in the 1990s, during the disinflation. The AO expectations increase to 9% in 1980 and decline to 2% in 2004. In addition, the 10-year expectations implied by AO fall right on top of the TIPS-based expectations. This provides informal evidence that AO-based forecasts have a reasonable level. In addition, using the visual inspection of the realized inflation as a basis, AO fares better than NF and AS after 1985.

We provide more formal evidence regarding the differences between AO and NF by displaying the in-sample inflation forecast RMSE in Table 2. In the full sample, the RMSE ratios are very close to one, which suggests that there is not much difference between the two models. We attempt to reconcile these numbers with Figure 3, where AO appears to be dominant, by splitting the sample

¹²Ang, Bekaert, and Wei (2007) provide a detailed numerical analysis of the biases in LS, MCS, and SPF.

¹³The five percent critical value of the chi-squared distribution with 20 degrees of freedom is 31.4.

¹⁴Our interpretation is that the extra flexibility allowed for by the AS specification leads to objective expectations that are unaffected by the information from the survey data.

¹⁵We use TIPS-based expectations that have been adjusted for risk and liquidity premia. The series are provided by the Federal Reserve Bank of Cleveland. The TIPS are available beginning from 1997. D'Amico, Kim, and Wei (2007) find that TIPS provide useful proxies for inflation expectations.

at the first quarter of 1983. This is when most surveys started to report their forecasts on a regular basis. Therefore, we should be able to generate useful results by using the surveys in the post-1983 subsample. Indeed, NF dominates AO by 5% to 20%, depending on the forecasting horizon, prior to 1983, while AO dominates NF by 15% to 45% afterwards.¹⁶

Given the importance of the inflation expectations that we obtained from AO, on the one hand, and the distinct patterns in term biases that we obtained from AS, on the other, we perform a robustness check by estimating ASR. By construction, this model produces inflation expectations that are identical to AO, but it has the flexibility to capture the term biases as well. We find that this model has MAE that are similar to the results for AO and AS that were presented in Table 1. As highlighted earlier, MAE in AO and AS differ substantively only in the case of MCS. The MCS' MAE in ASR is very similar to that in the more flexible AS. The term biases are close to zero (conditionally and unconditionally) for all surveys except for MCS.¹⁷ These two results are consistent with our earlier observations that MCS is the only survey that could benefit from the extra flexibility that is offered by the subjective measures.

Given the importance of the $\mathbb{P}^i = \mathbb{P}$ restriction for identifying accurate forecasts, we perform another robustness check by estimating OFO. This model generates MAE that are larger than those of OF and a likelihood-ratio test rejects OFO in favour of the more flexible OF. Nonetheless, the model produces inflation expectations that are very similar to the ones from AO.¹⁸ These findings mean that if a researcher is confronted with a survey-only dataset, OF is a better statistical description of the data and OFO is a better description of the data from the economic viewpoint. Thus, if one is interested in accurate inflation forecasts at all dates and for all horizons, OFO will provide a set of accurate forecasts.¹⁹

We conclude that using surveys is extremely important, because they help us to identify the objective dynamics of the state variables. However, the surveys have to be used in conjunction with yields, because the latter shed light on the issues addressed in our paper, that is, the interaction between monetary policy and inflation expectations and the behaviour of the inflation risk premium. Subjective probability measures could be important for interpreting the information contained in the survey forecasts, but from a statistical and forecasting perspective, a simpler model appears to be more attractive. We provide out-of-sample analysis in later sections.

¹⁶The sample values of RMSE ratios must be interpreted with caution, because the bootstrapped confidence bounds reported in Table 2 are quite wide. In particular, the ratios are not significantly different from one even in the post-1983 subsample for forecasting horizons of five years or more. The wide confidence bounds for RMSE ratios that we compute are consistent with similar findings by Stock and Watson (2007).

¹⁷The details of ASR-based results are available upon request.

¹⁸The details of OFO-based results are available upon request.

¹⁹In this regard, our conclusion supports the approach of Kozicki and Tinsley (2006).

5.1.3 The term structure of inflation expectations

Figure 4 shows the time-series of the inflation expectations at multiple horizons. These expectations are computed from AO. In contrast to the survey forecasts in Figure 1, these objective, or marginal, expectations can be computed each period at any horizon.

The term structure effects are pronounced. The inflation curve becomes inverted in 1973, right before the recession, and continues to be inverted until early 1982. This period coincides with the unstable period of monetary policy during the Burns and Miller chairmanship of the US Federal Bank and the monetary policy experiment under Volcker’s chairmanship. The curve became inverted again briefly in the early part of Greenspan’s tenure from 1987 to 1991. Afterwards, it had a normal, nearly flat, shape.

5.2 The determinants of the inflation expectations

5.2.1 Factor loadings

Given that the different versions of the model produce different inflation expectations, we would like to understand better how various factors influence these expectations. We begin our characterization of the forecasts’ determinants by plotting the factor loadings $b^{\mathbb{P}}(0, \tau)$ in (3.17) for the AO and NF models, where subjective and objective measures coincide. Figure 5 plots the loadings by factor.

AO model uses inflation as a level factor, because the loadings are almost identical throughout the forecasting horizon spectrum. In contrast, inflation works as a slope in the NF model. The current inflation value features prominently at the short horizons, and there is almost no effect at long horizons. GDP is used as a slope by AO and as a curvature by NF.

It is more difficult to interpret the loadings on the latent factors, because there are multiple combinations, or rotations, of the factors that produce identical yields and expectations (Dai and Singleton, 2000). To aid the interpretation of the latent factors, we follow Mueller (2008) and rotate them so that the two latent factors are orthogonal to each other and x_1 is interpreted as the factor that affects surveys. The last property is achieved by maximizing the one-year forecast’s loading on x_1 (see Appendix D for details).

The bottom panels of Figure 5 show the effect of such a rotation for the two models. We see that the largest loading on x_1 occurs at the one-year horizon. Moreover, in AO, the loadings on x_1 are similar across all the forecast horizons, which suggests that it is appropriate to interpret it as a “survey” factor. The orthogonal factor x_2 has a curvature effect in both models, but the effect is very small; it is an order of magnitude less than for other factors. This property of x_2 justifies our interpretation of x_1 further, because this is the only one of the two latent factors that has a material impact on expectations.

5.2.2 The nature of latent factors

To gain further insight into the latent factors, we study their determinants in our models. The latent variables are filtered at the estimation stage by constructing a linear function of the observables (real activity, inflation, yields, and forecasts). We will make an attempt to establish which ones feature most prominently in the filters.

We use the following *ad hoc* procedure. Suppose a filter with a steady-state forecast error covariance matrix can be represented as

$$\hat{x}_{it} = A_{i0} + A_{iy}y_t + A_{ip}\bar{p}_t + A_{im}m_t + A_{il}\hat{x}_{it-1}, \quad (5.3)$$

where y , \bar{p} and m represent vectors of observable yields, forecasts, and macro variables, respectively. We visually inspect the vectors of loadings A_y , A_p and A_m and select the ones with the largest absolute values. Then we rerun the Kalman filter, using as a basis the parameter values estimated from the full setup, but using only the observables that correspond to the selected weights. To be conservative, we implement this procedure in several steps and stop when a new, simplified, filter has a correlation of less than 0.98 with the original full filter.

Table 3 lists the variables that we use in each filter. We can make the following observations. First, our procedure allowed us to eliminate most of the observables from the filter. We need three or four observable series, depending on a factor and a model. Second, the two models use the macro variables in a different way: the first latent factor loads on both and the second factor does not use them at all in AO, while they load only on GDP or only on inflation in NF. Third, the factor x_1 in AO can be estimated using output, inflation, and one-year forecasts from the Livingston and SPF surveys.²⁰ This property is consistent with our interpretation of x_1 as a “survey” factor.

5.2.3 The incremental impact of surveys

The two previous sections provide evidence that survey forecasts are important for the determination of inflation expectations in our model. They channel through the latent variables. Still unaddressed is the quantitative impact on the inflation expectations of the surveys versus the macro variables. The difficulty in addressing this issue lies in the conditional and unconditional dependence of the macro and latent factors. One needs to disentangle the information in surveys that is not related to the macro variables. In the discussion that follows, we focus on AO, because, as the previous section shows, the latent factor x_1 is not related to yields in this model.

We distinguish the surveys and macro variables by applying the projection procedure in equation (3.21). If surveys have any incremental information over the macro variables, the residual component f will be correlated with some components of the inflation forecasts that are unrelated to the macro variables. We prewhiten the forecasts by regressing them on ten lags of real activity and inflation.

²⁰We explain the need for both surveys by the longer available span of LS and higher available frequency for SPF.

This will enable us to relate the factors f to the innovations in forecasts. We correlate f_1 and f_2 with the innovations in all 20 available forecasts. In the case of f_1 , only 20% of all correlations are below 0.4 and the highest value is 0.72. In the case of f_2 , 85% of all correlations are below 0.4 and the highest value is 0.45. Thus, the residual factor f_1 has inherited the “survey” interpretation from the original factor x_1 .

The top panels of Figure 6 show the breakdown of expected inflation into the contribution of M (inflation, real activity, and their history), f_1 and f_2 . It is obvious that inflation and its lags play a major role in explaining expected inflation. In contrast, factor f_2 makes almost no contribution. Real activity and the “survey” factor f_1 make a modest, but non-negligible contribution. These observations are supported by the contents of Table 4A, which presents the variance of expectations explained by M and f . Regardless of the forecast horizon, M explains approximately 90% of the variation in expectations and f_1 explains the remaining 10%, according to AO.²¹ The forecasts affect expectations not only through factor f_1 , but also through parameter values. As we argued in detail in section 5.1.2, an identical model with different parameter values yields less accurate expectations.

These findings lead us to the following three conclusions. First, the private sector expectations in the forward-looking interest rate rule representation (3.24) reflect information outside of the history of macro variables and yields. This information is captured by the residual f_1 . Second, this information is contained in the survey forecasts. Third, the residual f_1 may reflect observable variables that are not incorporated in our model or genuinely unobservable components that affect survey forecasts, such as the professional judgements of the participants in the surveys.²² Regardless of its origins, the information lies outside the standard variables that are contemplated in the theoretical literature on monetary policy. As highlighted earlier, this information is required if the objective dynamics of the state variables are to be identified. Taken on their own, yields are not very informative about this aspect of the data-generating process, because pricing takes place under the risk-neutral probability measure.

5.2.4 The interest rate response

The main finding of the previous section is that private sector expectations about inflation incorporate information that is orthogonal to the history of macro variables and, in the case of AO, to the history of yields. This bit of information is captured by the factor f_1 . The remaining issue is whether the monetary policy responds to f_1 .

We can evaluate how surveys affect nominal yields by breaking the latter down into the contribution of macro variables M and residual factors f using the projection (3.21). The residual factor f_1 is associated with information unique to the surveys and, therefore, represents the link between yields and forecasts. The bottom panels of Figure 6 display the projection-based breakdown. Ta-

²¹NF implies slightly more variation in these numbers, but they are qualitatively similar.

²²Kozicki and Tinsley (2006) review the literature on the judgmental aspects of forecasting.

ble 4B quantifies the contribution of M vis-a-vis f . As is the case with inflation expectations, inflation and its lags play a big explanatory role. Overall, M explain 40% to 50% of the yields' variation. Therefore, factors f contribute much more to the variation of yields than they do to the variation of expectations. In contrast to expectations, much of the contribution comes from factor f_2 .

This analysis does not provide the full picture of the policy response to f_1 . So far, we have seen only the total impact of f_1 as a part of the latent factor x_1 that follows from equations (3.6) and (3.21). Equation (3.24) shows that f_1 affects the spot interest rate r via two channels: directly as a shock in the interest rate rule and indirectly via the inflation expectation. Because we want to gauge whether monetary policy responds to the determinants of expectations, we are interested primarily in the latter channel.

Equations (3.19) and (3.20) allow us to compute the specific impact of f_1 via the expectations channel. The total contribution of f_1 , $\delta_{x,1}$, can be broken down as

$$\delta_{x,1} = \frac{\delta_\pi}{\Phi_{22}}\Phi_{23} + \gamma_{x,1}, \quad (5.4)$$

or, in relative terms,

$$1 = \frac{\delta_\pi\Phi_{23}}{\delta_{x,1}\Phi_{22}} + \frac{\gamma_{x,1}}{\delta_{x,1}}, \quad (5.5)$$

where Φ_{ij} denotes an individual element of the matrix Φ . The loading on the contribution of f_1 via the expectations channel is the first member of the sum. Our parameter estimates imply a value of 0.73.²³ Thus, the contribution of f_1 to the volatility of r via the expectation component represents 73% of the overall contribution of f_1 , which is 19% in AO.

5.3 Monetary policy effectiveness

We have established that it is important to use information from surveys and yields to construct an accurate measure of inflation expectations. Moreover, short-term inflation expectations are driven by output, inflation, and only one of the two latent factors. This factor affects inflation expectations at all horizons and is, therefore, termed a “survey” factor. Various breakdowns of this factor show that it is driven by the macro variables, their history, and a residual factor that reflects information that is orthogonal to that contained in macro variables and the yield curve. The residual factor affects the short-term interest rate mostly via the inflation expectation channel.

Thus, we have established that the interest rate rule responds to the determinants of the expectations of the private sector, in addition to the standard macro variables. This conclusion is consistent with the prescription of the literature on the stability of monetary policy. The next natural issue to address is whether such an interest rate rule leads to monetary policy stability, or expectations

²³This value is statistically significant, because the 95% bootstrapped confidence interval is [0.11, 1.94].

anchoring, as predicted by the literature. We explore anchoring from three angles: the magnitude of inflation premia, long-run inflation expectations, and inflation persistence.

5.3.1 Inflation premia

If inflation expectations are anchored, the inflation premium should be low. We extract inflation premia from nominal yields via the Fisher equation (3.25). We report summary statistics associated with the determinants of nominal yields in Table 5.

There are quite substantive differences between the nominal yield components across the two models. All of them, with the exception of the long-run expected inflation, are more volatile in NF. As we showed before, the NF-based inflation expectations are slightly higher, about 20 to 30 basis points on average. These modest differences translate into large divergences in real rates and inflation premia. The short-term real rate averages at 2% for both models.²⁴ However, the long-term rate declines to 1.2% in AO and increases to 3.3% in NF.²⁵ These differences feed back into the magnitudes of inflation premia. In AO, the average premium increases from 0.2% to 2.2%, while in the NF the average premium is negative and declines from -0.1% to -0.3%.

Thus, using two versions of one and the same model, we were able to replicate the varying signs of inflation premium that have been reported in the existing research. Hordahl and Tristani (2007) provide a thorough discussion of this issue. In particular, the existing calibrations of various general equilibrium models imply a positive inflation premium. On the empirical side, the sign of the inflation premium depends on how inflation expectations are measured in a particular implementation. The AO model produces more accurate expectations and positive inflation premia. Thus, it is natural to conclude that the estimates of the real rates are more reasonable in AO as well.

Figure 7 displays the time series of the breakdown of nominal yields via the Fisher equation. We see that the 10-year inflation premium varies between -5% and 10% in NF. This appears to be excessive, particularly because expected inflation is stable around 5% in this model. Further, during the so-called disinflation period of the 1980s and 1990s, one would expect the absolute value of the inflation premium to decline, regardless of its sign. This conjecture is consistent with AO, but not with NF.

The last observation is precisely what we were looking for when gauging the anchoring of expectations. The more reasonable AO model implies that the 10-year inflation premium was gradually falling after the monetary experiment and was hovering between zero and 2.5 per cent after 1990. This is precisely what we would expect to see if inflation were stabilized.

²⁴This number is consistent with Buraschi and Jiltsov (2005) and is slightly higher than that reported by Ang, Bekaert, and Wei (2004).

²⁵Buraschi and Jiltsov (2005) report a mild increase to 2.5%. Ang, Bekaert, and Wei (2004) do not investigate long maturities.

5.3.2 Stability

To address the issue of the stability of inflation, we can study the evolution of the long-run inflation expectations (monetary policy is stable if inflation expectations are low and do not change) and the persistence of inflation (the long-run expectations about inflation are influenced by current transient shocks if inflation is persistent). We conduct an out-of-sample analysis to measure these quantities.

The out-of-sample framework offers two advantages. First, the Gaussian model that we use to describe the evolution of state variables implies a constant long-run mean and a constant persistence of these variables. This is an undesirable feature if one wants to evaluate how long-run inflation expectations or the responses of these expectations to shocks have changed over time.²⁶ Second, the out-of-sample approach brings us closer to the adaptive expectations framework that has been used to evaluate optimal monetary policy (e.g., Evans and Honkapohja, 2003, or Preston, 2006). Indeed, if our model is reestimated for each period, it can be interpreted as a reduced-form representation of the optimal monetary policy with adaptive expectations. This view can be rationalized within the anticipated utility framework of Kreps (1998). This approach essentially allows decision makers to update parameter values as new data arrive and to simultaneously make decisions as if current parameter values remain unchanged.²⁷

We implement the out-of-sample analysis by reestimating our model every quarter as new information arrives. This allows us to construct “online” expectations of inflation. We take a pre-disinflation period (1970 to 1982) as a burn-in sample to obtain the initial parameter estimates. There are at least two reasons to do so. First, it would be difficult, if not impossible, to estimate our relatively large model using too few data points. Taking roughly a third of the observations to obtain the initial values seems reasonable. Second, most of our surveys do not have observations during this period, so it would be difficult to evaluate the role of the forecasts.

Changing parameters

We provide the time-series of some of the estimated parameters in Figure 8. We focus on the parameters that are related to inflation and real activity, because the parameters that pertain to the latent factors are more difficult to interpret. Indeed, the latent factors will change every period, not only because of the time variation in the state of the economy, but also because of changes in the values of parameters.²⁸ On each plot, the last point corresponding to the fourth quarter of 2004 shows the estimated parameter values for the full sample.

Panel (a) reports the loadings on inflation (dashed lines) and real activity (solid lines), that is,

²⁶Kozicki and Tinsley (2006) offer an alternative approach for incorporating the time-varying long-run mean of inflation.

²⁷Cogley and Sargent (2006) discuss the connection between anticipated utility, rational expectations, and Bayesian decision-making.

²⁸Collin-Dufresne, Goldstein, and Jones (2006) provide a detailed analysis of this phenomenon.

parameters δ_m in the spot interest rate equation (3.6). Inflation loadings are similar across the two models. They gradually decline from 0.4 to 0.2. The real activity loadings are similar for the two models until 1990, when the loading in AO jumps from -0.2 to 0.2. The 10-year forecasts were introduced to the surveys after this date, which, we believe, affects the estimated values of the real activity loading.

The situation is the opposite for the long-run mean reported in panel (b). Now, the value of parameters controlling the real activity do not change much from one model to another. The long-run inflation declines from 6.6% to 2.3% in AO, while it does not change as dramatically in NF and declines only to 4.6%. This plot suggests that inflation has stabilized over time because, according to AO, its long-run expectations fell dramatically in the post-monetary-experiment period and were moving around 2-2.5% from the mid-1990s.

It is difficult to attribute persistence to a particular factor because the matrix Φ in (3.2) has off-diagonal elements. Instead, we report in panel (c) the most and the least persistent factors as measured by the eigenvalues of Φ . Here again, the source of instability is associated with 1990. The least persistent component (dashed lines) was similar across the two models before the introduction of long-term forecasts, and became different after that. While in AO, it continued to be close to zero, in NF it has gradually increased in persistence. The most persistent component (solid lines) was similar across the two models throughout the full sample. Thus, relative to the information in yields only, the surveys suggest the presence of a transient factor in the inflation dynamics, which helps to pull long-run expectations back on track after a shock.

The standard deviation of real activity (solid lines) in panel (d) is similar across the models and stable over time. The standard deviation of inflation (dashed lines) is higher in AO. It increases in both models until the early 1990s and then declines to approximately four per cent by the end of the sample. This evidence is again consistent with more stable inflation in recent history.

Changing impulse responses

We can also gauge the impact of the changing persistence of inflation on the anchoring of expectations via impulse responses. Our model allows the construction of impulse responses of expected inflation to shocks in inflation. As we add more data, estimated persistence of inflation! will change and so will the response of the long-run expected inflation. Monitoring these changes over time allows us to acquire a different perspective on the perceived stability of inflation.

We use the recursive identification scheme to compute the impulse responses. The order of the state variables is the same as in the original model specification. Figure 9 reports responses of the 10-year expected inflation to a one-basis-point inflation shock for four representative dates in our sample: 1987:1, 1989:2, 2000:1, and 2004:4.

The common characteristic of all responses is that expected inflation declines after the initial

spike. The critical difference between the early and late periods is whether the effect of the shock is eliminated completely. In the early years, the initial response of 0.20 - 0.35 basis points quickly drops to the level of 0.05 - 0.10 basis points, but then it lingers at this level for at least three years. In the later part of the sample, the initial response dissipates almost completely in two quarters. We interpret these impulse responses as evidence that inflation became less persistent and better anchored.

Conditional inflation expectation

We complement the inflation that are presented in Figure 8(b) by revisiting the inflation expectations that are presented in Figure 3. However, because we want to see how long-run expectations are affected by new data, it is more appropriate to study out-of-sample expectations. Figure 10 presents such out-of-sample expectations for AO and NF and benchmarks them against the realized inflation and TIPS-implied expectations. Because of our focus on the stability of long-run expectations, we replace one-year forecasts in Figure 3 by five-year forecasts.

In contrast to in-sample estimation, we see that the NF 10-year forecasts decline over time from 6-7% to 3.5-4% at the end of the sample. However, the NF expectations are quite variable. Thus, yields-based estimation would lead one to conclude that inflation expectations are not anchored.

The AO-based expectations are much more stable, even at the five-year horizon. These expectations decline to a more intuitive level of 2%. The example of the five-year forecasts also shows that, beginning from 1999, the level of realized inflation matches up with expectations quite well. The 10-year realized inflation is not yet available for the most recent period. The out-of-sample expectations continue to agree well with TIPS-based expectations.

Changing term structure of expectations

Figure 11 provides additional evidence regarding monetary policy by displaying the average term structure of expectations. The out-of-sample period is split into four five-year subsamples. We compute the time series of the expectations using out-of-sample parameter values, as in Figure 10, and then average them across the subsamples. This exercise provides a summary of how the expectations curve evolves over time.

The shape of the NF curve is normal first, then flattens out, and then becomes normal again in the last subsample. The normal shape of the curve suggests that the monetary policy is less effective in handling inflation as it bounces around its long-run mean. AO paints a different picture of the monetary policy. The average curve becomes flatter over time, with declines in both the short and long ends of the curve. This phenomenon could be interpreted as evidence that monetary policy is effective.

5.4 The implication of stable expectations

5.4.1 Inflation Forecasting

Table 6A 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 forecasts about inflation that are made one year in advance. Our results should complement their findings along 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.

The reported RMSE ratios send a clear message. First, AO significantly dominates NF in different subsamples and across the entire range of forecasting horizons. The improvement ranges from 25% to 65%, depending on the sample and horizon. Second, AO performs on par with raw surveys. It does much better than MCS, but this is to be expected, given our earlier discussions. For almost all other surveys, the ratio of AO to surveys is greater than one, but the difference is statistically insignificant and it is so mild that it is difficult to imagine that it would matter economically. The 10-year forecasts from LS and SPF stand out, because both ratios are equal to 1.3. However, we have only 9 and 13 observations of these forecasts, respectively, so these ratios are hardly reliable.

The fact that there is essentially no difference between raw surveys and a complicated AO model may be disappointing to some readers. 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 have presented and defended a model that can produce survey-quality forecasts, but at any time and at any horizon. An example of this is provided in Table 6B. Given that Ang, Bekaert, and Wei (2007), after searching through a list of 30 various models, could not find any such model even for one forecasting horizon, our achievement appears to be important.

5.4.2 Yield Forecasting

Table 7 displays out-of-sample results for the forecasting of yields. We use the random walk model (RW) as a reference and report all the results in the form of RMSE ratios of a model (AO or NF) to RW. To the best of our knowledge, this is the first analysis 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 sub-sample, or reestimated each period over the short subsample. The range of forecasting horizons is typically narrow, from the next period to one year ahead. Finally, statistical inference is not provided, that is, numerical values of RMSE are simply compared to each other.

Using point values of RMSE ratios as a basis, we see that in our first sample, which ranges from 1983 to 2004, RW dominates both models and AO, overall, fares better than NF. The exceptions are

most forecasts that are made one-quarter ahead.²⁹ However, the statistical uncertainty is so great that it is difficult to distinguish AO from NF. Moreover, AO cannot be distinguished from RW at longer forecasting horizons or for yields that have maturities of three years or more.

We noted earlier that AO can exploit the full benefit of the surveys only after 1990, when long-term forecasts began to appear in surveys. For this reason, our second sample ranges from 1990 to 2004. AO still underperforms NF for most short-term forecasts. However, it frequently beats RW. We still cannot distinguish AO from RW once statistical uncertainty is taken into account.

We conclude that incorporating forecasts about inflation helps when forecasting yields. However there remains room for improvement, because the results for forecasts of yields are not as clear as for those for forecasts of inflation. The analysis in this paper suggests that survey forecasts are very useful, so incorporating the forecasts of real activity or yields should help when forecasting yields and can reduce the statistical uncertainty.

6 Conclusion

We have built a dynamic macro-finance model that incorporates jointly the behaviour 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 the observed yields and survey forecasts are internally consistent with each other, at least in the context of our model. Moreover, both yields and forecasts are important for producing realistic expectations about future inflation and yields.

Using our model, we find evidence of systematic biases in survey forecasts. They overpredict inflation the most, and disagree the least, at the end of recessions. However, this does not imply that surveys are useless for forecasting. A model that restricts survey expectations so that they coincide with the expectations of the private sector and still uses information from surveys produces the most realistic inflation forecasts as judged by in- and out-of-sample RMSE, and by implied real yields and inflation premia.

The out-of-sample analysis of the model suggests that monetary policy became more effective over time. The long-run expectations are anchored at about 2%. The term structure of inflation expectations has flattened out over time. This suggests that the arrival of new data does not affect long-run expectations much, perhaps because the monetary policy is expected to address all short-term fluctuations successfully.

²⁹Our conjecture for the AO underperformance at short forecasting horizons is that NF fits yields better than AO (Table 1). The near-term forecasts should be similar to the current yields, so worse fit translates into more poor forecasting performance.

A A Model of Learning with Heterogeneous Beliefs

We start out by describing the evolution of inflation π_t in the most basic form:

$$\pi_t - \pi_{t-1} = p_{t-1} + \sigma e_t \quad (\text{A.1})$$

where p_t represents expected inflation rate. We fill this equation with more content by assuming that p_t is determined by the inflation π_t and some state variable s_t that is unobservable to the agents. We further assume that the vector $u_t = (\pi_t, s_t)'$ follows a VAR(1) process

$$u_t = \mu^u + \Phi^u u_{t-1} + \Sigma^u \epsilon_t^u \quad (\text{A.2})$$

$$= \begin{bmatrix} \mu^\pi \\ \mu^s \end{bmatrix} + \begin{bmatrix} \phi^{\pi\pi} & \phi^{\pi s} \\ \phi^{s\pi} & \phi^{ss} \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ s_{t-1} \end{bmatrix} + \begin{bmatrix} \sigma^{\pi\pi} & 0 \\ \sigma^{s\pi} & \sigma^{ss} \end{bmatrix} \begin{bmatrix} \epsilon_t^\pi \\ \epsilon_t^s \end{bmatrix}. \quad (\text{A.3})$$

In particular, this specification implies that the spot expectation of inflation is:

$$p_t = E_t(\pi_{t+1}) = \mu^\pi + \phi^{\pi\pi} \pi_t + \phi^{\pi s} s_t, \quad (\text{A.4})$$

Since s_t is not observable the agents must filter it. We use a setup that is similar to the one in Scheinkman and Xiong (2003), who develop stock pricing in the context of investors with heterogenous beliefs. For transparency, we assume that there are only two forecast surveys being conducted: A and B . Participants of the surveys signals θ_t^A and θ_t^B about s_t . Members of survey A think of the signal θ^A as their own, but can observe both. Specifically, forecaster A believes that only her signal is correlated with innovations in s_t , i.e., the vector $w_t = (\pi_t, \theta_t^A, \theta_t^B, s_t)'$ follows a restricted VAR(1) process:³⁰

$$w_t = \mu^w + \Phi^w w_{t-1} + \Sigma^w \epsilon_t^w \quad (\text{A.5})$$

$$= \begin{bmatrix} \mu^\pi \\ 0 \\ 0 \\ \mu^s \end{bmatrix} + \begin{bmatrix} \phi^{\pi\pi} & 0 & 0 & \phi^{\pi s} \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ \phi^{s\pi} & 0 & 0 & \phi^{ss} \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ \theta_{t-1}^A \\ \theta_{t-1}^B \\ s_{t-1} \end{bmatrix} + \begin{bmatrix} \sigma^{\pi\pi} & 0 & 0 & 0 \\ -\frac{\sigma^{s\pi} \sigma^{s\theta^A}}{\sigma^{\theta^A \theta^A}} & \sigma^{\theta^A \theta^A} & 0 & 0 \\ 0 & 0 & \sigma^{\theta^B \theta^B} & 0 \\ \sigma^{s\pi} & \sigma^{s\theta^A} & 0 & \sigma^{ss} \end{bmatrix} \begin{bmatrix} \epsilon_t^\pi \\ \epsilon_t^{\theta^A} \\ \epsilon_t^{\theta^B} \\ \epsilon_t^s \end{bmatrix},$$

where $\epsilon_t \sim N(0, I)$. The restrictions ensure that the private signal is not correlated with inflation rate π_t , and that the expected change in the private signal is equal to the unobserved state variable. The element $\sigma^{s\theta^A}$ controls the degree of informativeness of the signal regarding the state variable s .

We solve the forecaster's filtering problem using the results from appendix B, expression (B.7), in particular. In our Gaussian setup the filtered value of the state variable s is going to be a linear function of inflation π , the private signals and their lags:

$$\hat{s}_t^i = c_0^i + \sum_{j=0}^t (c_{\pi, t-j}^i \pi_{t-j} + c_{\theta, t-j}^i \theta_{t-j}^i + c_{\theta, t-j}^{-i} \theta_{t-j}^{-i}), \quad (\text{A.6})$$

³⁰A symmetric argument applies to the members of survey B .

where i and $-i$ generically refer to one of the surveys, and all others, respectively. Therefore, the survey i expected inflation value of inflation is equal to:

$$p_t^i = E_t^i(\pi_{t+1}) = \mu^\pi + \phi^{\pi\pi}\pi_t + \phi^{\pi s}\hat{s}_t^i. \quad (\text{A.7})$$

As a next step, we adopt this result to our empirical setting. An econometrician does not observe the private signals θ^i , therefore she has to estimate them, which would require adding a second layer of filtering equations. We want to avoid this complication and introduce additional notations and some approximations.

First, we assume that first p lags of the variables are sufficient to accurately approximate \hat{s}^i in (A.6). We stack up the contemporaneous and lagged values of the variables in (A.6) into a vector: $\zeta_t^p = (\pi_{t-1}, \dots, \pi_{t-p}, \theta_t^i, \dots, \theta_{t-p}^i, \theta_t^{-i}, \dots, \theta_{t-p}^{-i})'$. Now we can rewrite \hat{s}^i as

$$\hat{s}_t^i \approx c_0^i + c_{\pi,t}^i \pi_t + c^{i'} \zeta_t^p \quad (\text{A.8})$$

Second, we assume that a vector x_t of first k principal components of ζ_t^p explains most of the variation in this variable. Therefore,

$$\hat{s}_t^i \approx c_0^i + c_{\pi,t}^i \pi_t + \alpha^{i'} x_t \quad (\text{A.9})$$

where α^i is a convolution of c^i and the principal components loadings on ζ_t^p . In particular, these assumptions combined with the inflation forecast equation (A.4) implies that, regardless of the survey, inflation forecasts are going to be linear functions of π_t and x_t , but the weights in these function will be survey specific:

$$p_t^i = E_t^i(\pi_{t+1}) = \nu_0^i + \nu_\pi^i \pi_t + \nu_x^i x_t, \quad (\text{A.10})$$

Note that, because the vector $(\pi_t, \zeta_t^{p'})'$ follows a VAR(1), the vector $(\pi_t, x_t)'$ will do so as well. This observation yields our setup in section 3.³¹

B Projection

The model controlling the evolution of state z in (3.2) does not represent a state-space system. Nonetheless, Liptser (1997) and Liptser and Shiryaev (2001) derive the projection of one element of the VAR(1) on the other using the same ideas as in the Kalman filtering. In particular, these authors provide the following expression for the conditional mean, often referred to as “forecast,”

³¹In practice the accuracy of the approximation can be achieved by using a sufficient number of the latent factors x_t . The sufficiency can be established via model specification analysis.

and variance of the forecast error:

$$\begin{aligned}\hat{x}(M_t) &= \mu^x + \Phi^{xx}\hat{x}(M_{t-1}) + \Phi^{xm}m_{t-1} \\ &+ (\Sigma^{xx}\Sigma^{mx'} + \Sigma^{xm}\Sigma^{mm'} + \Phi^{xx}P_{t-1}\Phi^{mx'}) (\Sigma^{mx}\Sigma^{mx'} + \Sigma^{mm}\Sigma^{mm'} + \Phi^{mx}P_{t-1}\Phi^{mx'})^{-1} \\ &\times (m_t - \mu_m - \Phi^{mx}\hat{x}(M_{t-1}) - \Phi^{mm}m_{t-1})\end{aligned}\quad (\text{B.1})$$

$$\begin{aligned}P_t &= \Phi^{xx}P_{t-1}\Phi^{xx'} + (\Sigma^{xx}\Sigma^{xx'} + \Sigma^{xm}\Sigma^{xm'}) \\ &- (\Sigma^{xx}\Sigma^{mx'} + \Sigma^{xm}\Sigma^{mm'} + \Phi^{xx}P_{t-1}\Phi^{mx'}) (\Sigma^{mx}\Sigma^{mx'} + \Sigma^{mm}\Sigma^{mm'} + \Phi^{mx}P_{t-1}\Phi^{mx'})^{-1} \\ &\times (\Sigma^{xx}\Sigma^{mx'} + \Sigma^{xm}\Sigma^{mm'} + \Phi^{xx}P_{t-1}\Phi^{mx'})',\end{aligned}\quad (\text{B.2})$$

where we generically refer to m and x as vectors of observable and latent variables, respectively.

We introduce additional notations to describe the projection initialization. The long run mean z is:

$$(I - \Phi)^{-1} \mu = \begin{bmatrix} \Theta^m \\ \Theta^x \end{bmatrix} \quad (\text{B.3})$$

$$\quad (\text{B.4})$$

The steady-state matrix P satisfies:

$$\begin{aligned}P &= \Phi^{xx}P\Phi^{xx'} + (\Sigma^{xx}\Sigma^{xx'} + \Sigma^{xm}\Sigma^{xm'}) \\ &- (\Sigma^{xx}\Sigma^{mx'} + \Sigma^{xm}\Sigma^{mm'} + \Phi^{xx}P\Phi^{mx'}) (\Sigma^{mx}\Sigma^{mx'} + \Sigma^{mm}\Sigma^{mm'} + \Phi^{mx}P\Phi^{mx'})^{-1} \\ &\times (\Sigma^{xx}\Sigma^{mx'} + \Sigma^{xm}\Sigma^{mm'} + \Phi^{xx}P\Phi^{mx'})'\end{aligned}\quad (\text{B.5})$$

Then the projection is initialized as follows:

$$\hat{x}(m_0) = \Theta^x + V^{xm}(V^{mm})^{-1}(m_0 - \Theta^m), \quad P_0 = P \quad (\text{B.6})$$

In this case $P_t = P$ and the projection is time-stationary. An alternative strategy is to initialize P_0 at the unconditional variance of z . In this case, the sequence P_t will converge to P . In our model it happens in twelve steps.

The lags of the projected x in the expression (B.1) could be recursively substituted out so that the current projection is expressed as a distributed-lag function of macro variables:

$$\hat{x}(M_t) = c(\Theta) + \sum_{j=0}^t c_{t-j}(\Theta)m_{t-j}, \quad (\text{B.7})$$

where the matrices c are functions of parameters Θ that control the dynamics of the state variables z in (3.2).

C Details of the Dataset

C.1 Yields

We use quarterly time series of bond data from 1970 to 2004. We use an unsmoothed Fama-Bliss approximation of the zero coupon bond prices of maturities at three and six months, and one, two, three, five, seven, and ten years.³² It is important to measure the full yield curve because its slope is correlated with the macro environment (Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998). Given this constraint, we do not consider earlier years, because the longest maturity available was five years. In addition, using rich yield data helps to identify the risk premia.

C.2 Macro Variables

We use quarterly time series of linearly detrended real per capita GDP and log changes in seasonally adjusted CPI to proxy for g_t , and π_t , respectively. GDP and CPI numbers are available from FRED. 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 2000 USD (seasonally adjusted annual rate).

All the inflation forecasts that we are using are released some time in the third month of a quarter. Therefore, in order 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.

To construct the GDP growth series we follow two steps. First, we divide the quarterly GDP number by the mid-month population in the third month of the quarter. Second, we linearly detrend real per capita GDP. The detrending is done on a period-by-period basis to avoid a look-ahead bias. Because GDP numbers for quarter t are released in quarter $t + 1$ (first estimate in the first month of quarter $t + 1$ and then revisions in the two following months), we shift the time series by one period to account for the fact that in quarter t we only know information about quarter $t - 1$.

C.3 Survey Forecasts

We use the following survey data:

1. Michigan Consumer Survey (MCS). This is a monthly survey conducted by the University of Michigan from 1978. The quarterly forecasts that we use are available from 1960. The forecasts of the annualized percentage price change $P_{t+\tau}/P_t - 1$, are released early in the third month of a quarter. The respondents answer the question: “By about what percent do you expect prices

³²We are grateful to Robert Bliss for providing us with the data.

to go (up/down) on the average, during the next 12 months?” Therefore, the forecast is not tied to a particular CPI statistic. We have observations for $\bar{p}_{t,0}^1(4)$.

2. Livingston Survey (LS). This is a semi-annual survey of economists from industry, government, banking, and academia. The forecasts of the price level $P_{t+\tau}$, specifically of non-seasonally 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 $\bar{p}_{t,0}^2(\tau)$, $\tau = 2, 4$ (available from 1947). We also have annual forward forecasts available from 1974: $\bar{p}_{t,2}^2(4)$ (released in the second quarter), and $\bar{p}_{t,4}^2(4)$ (released in the fourth quarter). Finally, ten-year forecasts, $\bar{p}_{t,0}^2(40)$, are available semi-annually from 1991, and two-year forecasts, $\bar{p}_{t,0}^2(8)$, 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, $\bar{p}_{t,0}^2(2)$. We perform a simple seasonal adjustment of the forecast. We compute the average annual inflation, and the average inflation over the first and second half-a-year in our sample. Then we adjust the six-month forecasts by the respective differences in annual and semi-annual annualized average inflations. Ghysels and Osborn (2001) provide the details of, and 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.

3. Survey of Professional Forecasters (SPF). This 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 $\bar{p}_{t,s}^3(1)$, $s = 0, 1, 2, 3$; $\bar{p}_{t,0}^3(\tau)$, $\tau = 4, 40$. The ten-year forecast is available from the fourth quarter of 1991.
4. Blue Chip Economic Indicators (BCEI).³³ This is a monthly survey of economic forecasters at approximately 50 banks, corporations, and consulting firms. 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 $\bar{p}_{t,s}^4(1)$, $s = 0 - 6$. The forecasts with $s = 4, 5, 6$ are available in the first three, two, and one quarters of a year, respectively.

³³We are grateful to Randell Moore for providing us with the data.

D Latent Factor Indeterminacy

Dai and Singleton (2000) point out that identifying 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 will use the invariant affine transformation, which scales factors by a matrix. Appendix A of Dai and Singleton (2000) describes how such a transformation affects model parameters.

We examine two types of rotations. The first rotation, \mathcal{O} , ensures that the two factors are orthogonal to each other. We define a rotation $\mathcal{O} = Rx_t$, so that the variance-covariance matrix of x becomes diagonal. The matrix R is not unique; i.e., the rotation of type \mathcal{O} can generate many pairs of orthogonal factors x . Our second proposed rotation, \mathcal{M} , can be applied after any of the rotations from the class \mathcal{O} , resolves this type of indeterminacy. Define $\mathcal{M} = Ux_t$, where the matrix U is the orthogonal matrix; i.e., $UU' = I$, that preserves the correlation structure between the factors. In our two-dimensional case, the matrix U is determined by a single parameter, which is established by maximizing the loading of one-year inflation expectation on x_1 .

References

- Ait-Sahalia, Yacine, and Robert Kimmel, 2002, Estimating Affine multifactor term structure models using closed-form likelihood expansions, Working paper, Princeton University.
- Ang, Andrew, Geert Bekaert, and Min Wei, 2004, The term structure of real rates and expected inflation, *Journal of Finance*, forthcoming.
- , 2007, Do macro variables, asset markets or surveys forecast inflation better?, *Journal of Monetary Economics* 54, 1163–1212.
- Ang, Andrew, Sen Dong, and Monika Piazzesi, 2004, No-arbitrage Taylor rules, Working paper, Columbia University.
- Ang, Andrew, and Monika Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics* 50, 745–787.
- Backus, David, Silverio Foresi, and Chris Telmer, 1999, Discrete time models of bond pricing, in Narasimhan Jegadeesh, and Bruce Tuckman, ed.: *Advanced Fixed-Income Valuation Tools* (Wiley and Sons: New York, Ny).
- Basak, Suleyman, 2005, Asset pricing with heterogeneous beliefs, *Journal of Banking and Finance* 29, 2849–2881.
- Bikbov, Ruslan, and Mikhail Chernov, 2005, No-arbitrage macroeconomic determinants of the yield curve, Working paper, Columbia University.
- Brandt, Michael W., and Ping He, 2002, Simulated likelihood estimation of affine term structure models from panel data, Working paper, Wharton.
- Buraschi, Andrea, and Alexei Jiltsov, 2005, Inflation risk premia and the expectation hypothesis, *Journal of Financial Economics* 75, 429–490.
- Carlson, John A., 1977, A study of price forecasts, *Annals of Economic and Social Measurement* 6, 27–56.
- Chen, Ren-raw, and Louis O. Scott, 1993, Maximum likelihood estimation of a multi-factor equilibrium model of the term structure of interest rates, *Journal of Fixed Income* pp. 14–31.
- Chun, Albert Lee, 2005, Expectations, bond yields and monetary policy, Working paper, Stanford University.
- Clarida, Richard, Jordi Gali, and Mark Gertler, 1998, The science of monetary policy: A New-keynesian perspective, *Journal of Economic Literature* 37, 1661–1707.

- Cogley, Timothy, and Thomas Sargent, 2006, Anticipated utility and rational expectations as approximations of bayesian decision making, Working paper, UC Davis and NYU.
- Collin-Dufresne, Pierre, Robert Goldstein, and Christopher Jones, 2003, Identification and estimation of ‘maximal’ affine term structure models: An application to stochastic volatility, Working Paper, Washington University.
- , 2006, Identification of maximal affine term structure models, Working Paper, USC.
- Dai, Qiang, and Kenneth Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.
- D’Amico, Stefania, Don H. Kim, and Min Wei, 2007, Tips from TIPS: The informational content of Treasury inflation-protected security prices, Working paper, Federal Reserve Board.
- de Jong, Frank, 2000, Time series and cross-section information in affine term-structure models, *Journal of Business and Economic Statistics* 18, 300–314.
- Detemple, Jerome, and S. Murthy, 1994, Intertemporal asset pricing with heterogeneous beliefs, *Journal of Economic Theory* 62, 294–320.
- Duffee, Gregory R., 2005, Term structure estimation without using latent factors, forthcoming, *Journal of Financial Economics*.
- , 2006, Are variations in term premia related to the macroeconomy?, Working paper, UC Berkeley.
- , and Richard Stanton, 2004, Estimation of dynamic term structure models, Working paper, University of California at Berkeley.
- Dumas, Bernard, Alexander Kurshev, and Raman Uppal, 2005, What can rational investors do about excessive volatility?, Working Paper, London Business School.
- Estrella, Arturo, and Gikas A. Hardouvelis, 1991, The term structure as a predictor of real economic activity, *Journal of Finance* 46, 555–576.
- Estrella, Arturo, and Frederic Mishkin, 1998, Predicting U.S. recessions: Financial variables as leading indicators, *Review of Economics and Statistics* 80, 45–61.
- Evans, George, and Seppo Honkapohja, 2003, Expectations and the stability problem for optimal monetary policies, *Review of Economic Studies* 70, 807–824.
- Evans, Martin D. D., 1998, Real rates, expected inflation, and inflation risk premia, *Journal of Finance* 53, 187–218.

- Friedman, Milton, 1968, The role of monetary policy, *American Economic Review* 58, 1–17.
- Ghysels, Eric, and Denise Osborn, 2001, *The Econometric Analysis of Seasonal Time Series* (Cambridge University Press: Cambridge, UK).
- Harrison, J., and David Kreps, 1978, Speculative investor behaviour in a stock market with heterogeneous expectations, *Quarterly Journal of Economics* 92, 323–336.
- Harvey, Andrew C., 1989, *Forecasting, structural time series models and the Kalman filter* (Cambridge University Press: Cambridge, UK).
- Hordahl, Peter, and Oreste Tristani, 2007, Inflation risk premia in the term structure of interest rates, Working paper, European Central Bank.
- Howitt, Peter, 1992, Interest rate control and nonconvergence to rational expectations, *Journal of Political Economy* 100, 776–800.
- Joyce, Mike, Peter Lildholdt, and Steffen Sorensen, 2007, Extracting inflation expectations and inflation risk premia from the term structure: a joint model of the UK nominal and real yield curves, Working paper, Bank of England.
- Kim, Don H., and Athanasios Orphanides, 2005, Term structure estimation with survey data on interest rate forecasts, Working paper, Board of Governors of the Federal Reserve System.
- Kozicki, Sharon, and P.A. Tinsley, 2006, Survey-based estimates of the term structure of expected U.S. inflation, Working paper, Bank of Canada.
- Kreps, David, 1998, Anticipated utility and dynamic choice, in D. Jacobs, E. Kalai, and M. Kamien, ed.: *Frontiers of Research in Economic Theory* (Cambridge University press: Cambridge, UK).
- Liptser, Robert S., 1997, Stochastic control, Lecture Notes, Tel Aviv University.
- , and Albert N. Shiryaev, 2001, *Statistics of Random Processes: I. General Theory* (Springer: Berlin Heidelberg) 2nd edn.
- Lu, Biao, and Liuren Wu, 2005, Systematic movements in macroeconomic releases and the term structure of interest rates, Working Paper, Baruch College.
- Mueller, Philippe, 2008, Credit spreads and real activity, Working paper, Columbia Business School.
- Patton, Andrew, and Allan Timmermann, 2007, Generalized forecast errors, a change of measure, and forecast optimality conditions, Working Paper, Oxford.

- Pennacchi, George G., 1991, Identifying the dynamics of real interest rates and inflation: Evidence using survey data, *Review of Financial Studies* 4, 53–86.
- Piazzesi, Monika, and Martin Schneider, 2008, Bond positions, expectations, and the yield curve, Working paper, Chicago GSB.
- Press, William H., Saul A. Teukovsky, William T. Vetterling, and Brian P. Flannery, 1992, *Numerical Recipes in C* (Cambridge University Press: Cambridge) second edn.
- Preston, Bruce, 2006, Adaptive learning, forecast-based instrument rules and monetary policy, *Journal of Monetary Economics* 53, 507–535.
- Primiceri, Giorgio, 2006, Why inflation rose and fell: Policy-makers' beliefs and U.S. postwar stabilization policy, *Quarterly Journal of Economics* 121, 867–901.
- Scheinkman, Jose, and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183–1219.
- Stock, James, and Mark Watson, 2007, Why has U.S. inflation become harder to forecast?, *Journal of Money, Banking and Finance* 39.
- Stock, James H., and Mark W. Watson, 2003, Forecasting output and inflation: The role of asset prices, *Journal of Economic Literature* 41, 788–829.
- Thomas, Lloyd B., 1999, Survey measures of expected U.S. inflation, *Journal of Economic Perspectives* 13, 125–144.

Table 1 : Mean absolute errors

We report mean absolute fitting errors for four versions of our model. The full model that uses the survey and yield data is labeled as AS (All data, Subjective expectations). A model that uses all the data, but restricts subjective measures to coincide with objective measure is labeled as AO (All data, Objective expectations). A model that uses only yields for estimation is referred to as NF (No Forecasts). In this implementation the subjective measures cannot be estimated because they are not identified. Finally, we estimate a model that uses only forecasts and label it as OF (Only Forecasts). The notation for surveys is as follows: MCS - Michigan consumer survey; LS - Livingston survey; SPF - survey of professional forecasters; BCEI - Blue Chip economic indicators.

Survey	Obs	Horizon, qtr		Models			
		s	τ	AS	AO	OF	NF
MCS	108	0	4	0.45	0.80	0.45	
LS	70	0	2	0.18	0.22	0.19	
	70	0	4	0.12	0.13	0.12	
	13	0	8	0.12	0.13	0.07	
	29	0	40	0.17	0.15	0.11	
	32	2	4	0.16	0.17	0.14	
	31	4	4	0.27	0.28	0.14	
SPF	94	0	1	0.14	0.14	0.11	
	94	0	4	0.06	0.06	0.07	
	53	0	40	0.16	0.15	0.07	
	94	1	1	0.07	0.07	0.03	
	94	2	1	0.07	0.07	0.07	
	94	3	1	0.11	0.12	0.10	
BCEI	100	0	1	0.23	0.21	0.21	
	100	1	1	0.13	0.15	0.12	
	100	2	1	0.20	0.21	0.15	
	99	3	1	0.15	0.16	0.07	
	73	4	1	0.19	0.19	0.04	
	48	5	1	0.24	0.23	0.01	
	24	6	1	0.28	0.28	0.05	
	Yields	140		1	0.47	0.47	
140			2	0.39	0.38		0.07
140			4	0.29	0.28		0.13
140			8	0.15	0.14		0.13
140			12	0.15	0.15		0.11
140			20	0.26	0.26		0.07
140			28	0.34	0.35		0.07
140			40	0.45	0.45		0.14

Table 2 : In-sample Inflation Forecast RMSE ratios

We compare the in-sample inflation forecasting performance of the AO and NF models. We arbitrarily 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 RMSE ratio of AO to NF. In addition to the full sample, The RMSE ratios are reported for two sub-samples. The first sub-sample (from 1970 to 1982) represents the period when survey forecasts were rarely available. The second sub-sample (from 1983 to 2004) represents the full collection of the surveys. The bootstrapped 95% confidence bounds are provided in brackets.

Horizon, qtr	AO/NF		
	Full Sample	1970-1982	1983-2004
1	1.07 [1.04, 1.09]	1.26 [1.21, 1.32]	0.82 [0.79, 0.86]
2	1.05 [1.01, 1.09]	1.22 [1.14, 1.30]	0.75 [0.70, 0.80]
4	1.05 [0.98, 1.12]	1.22 [1.10, 1.35]	0.69 [0.61, 0.78]
8	0.98 [0.86, 1.12]	1.10 [0.89, 1.36]	0.64 [0.51, 0.81]
12	0.94 [0.78, 1.15]	1.05 [0.78, 1.44]	0.63 [0.44, 0.87]
20	0.93 [0.65, 1.30]	1.05 [0.62, 1.70]	0.57 [0.24, 1.07]
28	0.98 [0.57, 1.55]	1.14 [0.54, 2.14]	0.56 [0.07, 1.39]
40	1.07 [0.40, 2.11]	1.26 [0.35, 2.75]	0.63 [0.01, 1.96]

Table 3 : Simplified Filters

We summarize which data can nearly replicate the full filter of latent variables x . The results are reported for two versions of our model. In the case of AO, the estimation was conducted using all the data, that is, macro variables, m_t , yields, $y_t(\tau)$, and survey forecasts $\bar{p}_t(s, t)$. Therefore, the full filter was computed using all these data. In the case of NF, only m_t and $y_t(\tau)$ were used. The table reports, which data are instrumental in constructing the filter as evidenced by the correlation of the simplified filter with the full filter (last column, labeled “corr”). The numbers in the $y_t(\tau)$ column refer to yields’ maturity τ in quarters. The notation in the $\bar{p}_t(s, t)$ column refers to a particular survey (LS for Livingston, SPF for Survey of Professional Forecasters) and starting quarter and forecasting horizon (in quarters).

Model	Factor	m_t	$y_t(\tau)$	$\bar{p}_t(s, t)$	corr
AO	x_1	g, π	–	LS(0,4), SPF(0,4)	0.99
	x_2	–	1, 40	LS(0,4), SPF(0,4)	0.98
NF	x_1	g	1, 40	–	0.99
	x_2	π	1, 40	–	0.98

Table 4 : Theoretical R^2 for expected inflation and yields

We establish which fraction of the expectations, or yield curve variation is explained by various factors. Specifically, we distinguish M , the entire history of inflation and real activity, and f , the projection residual factors. Because the residual factors are orthogonal with respect to each other, we can further distinguish their individual contributions. The bootstrapped 95% confidence bounds are provided in brackets.

Panel A. Expectations

Horizon (qtr)	M		f_1		f_2	
	AO	NF	AO	NF	AO	NF
1	0.93	0.95	0.07	0.05	0.00	0.00
	[0.63, 0.95]	[0.89, 0.98]	[0.05, 0.36]	[0.02, 0.11]	[0.00, 0.01]	[0.00, 0.00]
8	0.89	0.85	0.11	0.15	0.00	0.00
	[0.44, 0.93]	[0.71, 0.95]	[0.06, 0.56]	[0.05, 0.29]	[0.00, 0.01]	[0.00, 0.01]
40	0.88	0.83	0.11	0.12	0.01	0.05
	[0.38, 0.93]	[0.61, 0.98]	[0.06, 0.58]	[0.01, 0.32]	[0.00, 0.05]	[0.00, 0.12]

Panel B. Yields

Horizon (qtr)	M		f_1		f_2	
	AO	NF	AO	NF	AO	NF
1	0.50	0.54	0.19	0.30	0.31	0.16
	[0.14, 0.86]	[0.26, 0.70]	[0.06, 0.43]	[0.11, 0.57]	[0.10, 0.63]	[0.02, 0.61]
8	0.41	0.52	0.21	0.20	0.38	0.29
	[0.06, 0.84]	[0.23, 0.70]	[0.06, 0.43]	[0.03, 0.50]	[0.12, 0.72]	[0.06, 0.74]
40	0.39	0.50	0.23	0.07	0.39	0.43
	[0.02, 0.84]	[0.21, 0.71]	[0.07, 0.47]	[0.01, 0.34]	[0.12, 0.74]	[0.18, 0.79]

Table 5 : Sample statistics for nominal yields decomposition

We report the annualized mean and standard deviation of the nominal yield components in the decomposition

$$y_t(\tau) = y_t^R(\tau) + \bar{p}_{t,0}(\tau) + IP_t(\tau),$$

where $y_t^R(\tau)$ is the real yield, $\bar{p}_{t,0}(\tau)$ is the inflation expectation under the \mathbb{P} measure, and $IP_t(\tau)$ is the inflation premium.

Model	τ		$y_t^R(\tau)$	$\bar{p}_{t,0}(\tau)$	$IP_t(\tau)$
AO	4	Mean	2.00	4.39	0.19
		Std. dev.	1.44	0.86	1.14
	40	Mean	1.19	4.27	2.16
		Std. dev.	0.20	0.60	1.28
NF	4	Mean	2.05	4.64	-0.07
		Std. dev.	3.60	2.06	2.12
	40	Mean	3.29	4.63	-0.31
		Std. dev.	1.20	0.44	2.06

Table 6 : Out-of-sample Inflation Forecast RMSE Ratios

We compare the out-of-sample inflation forecasting performance of the AO and NF models and the observed survey forecasts. Panel A compares the models to the actual surveys by reporting RMSE ratio of a model to a survey. The notation for surveys is as follows: MCS - Michigan consumer survey; LS - Livingston survey; SPF - survey of professional forecasters; BCEI - Blue Chip economic indicators. In Panel B we select alternative forecasting horizons. We report RMSE ratio of AO to NF as the surveys are not available for most of the selected horizons. Actual survey forecasts have missing observations. Therefore, the RMSE ratios will be 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 2004. 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, qtr		1983-2004		1990-2004	
	s	τ	AO/Survey	NF/Survey	AO/Survey	NF/Survey
MCS	0	4	0.78 [0.60, 0.96]	1.52 [1.27, 1.76]	0.61 [0.40, 0.84]	1.35 [1.07, 1.65]
LS	0	2	1.08 [0.84, 1.33]	1.65 [1.37, 1.94]	0.99 [0.68, 1.30]	1.63 [1.29, 2.00]
	0	4	1.06 [0.80, 1.32]	1.97 [1.62, 2.29]	1.01 [0.67, 1.40]	2.15 [1.67, 2.62]
	0	8	1.07 [0.51, 1.74]	3.15 [2.34, 4.05]	1.07 [0.51, 1.74]	3.15 [2.34, 4.05]
	0	40	1.32 [0.68, 2.38]	2.50 [1.92, 3.34]	1.32 [0.68, 2.38]	2.50 [1.92, 3.34]
	2	4	1.05 [0.78, 1.34]	2.08 [1.68, 2.50]	1.06 [0.64, 1.55]	2.41 [1.78, 3.08]
	4	4	1.13 [0.83, 1.46]	2.35 [1.91, 2.83]	1.17 [0.68, 1.73]	3.05 [2.34, 3.86]
SPF	0	1	1.05 [0.87, 1.25]	1.42 [1.26, 1.61]	0.99 [0.76, 1.25]	1.36 [1.14, 1.59]
	0	4	1.10 [0.86, 1.38]	2.17 [1.82, 2.51]	0.99 [0.66, 1.37]	2.20 [1.74, 2.68]
	0	40	1.34 [0.59, 2.59]	2.82 [2.12, 3.86]	1.34 [0.59, 2.59]	2.82 [2.12, 3.86]
	1	1	1.05 [0.87, 1.25]	1.59 [1.37, 1.83]	0.99 [0.75, 1.26]	1.58 [1.28, 1.91]
	2	1	1.05 [0.87, 1.24]	1.71 [1.46, 1.98]	1.00 [0.76, 1.28]	1.68 [1.34, 2.07]
	3	1	1.06 [0.89, 1.26]	1.84 [1.57, 2.12]	1.00 [0.74, 1.30]	1.91 [1.52, 2.34]
BCEI	0	1	1.07 [0.89, 1.28]	1.45 [1.29, 1.64]	1.00 [0.76, 1.26]	1.37 [1.15, 1.61]
	1	1	1.04 [0.86, 1.23]	1.57 [1.36, 1.81]	1.00 [0.79, 1.28]	1.60 [1.29, 1.93]
	2	1	0.96 [0.80, 1.14]	1.57 [1.34, 1.82]	1.04 [0.79, 1.33]	1.75 [1.39, 2.15]
	3	1	1.06 [0.88, 1.25]	1.82 [1.55, 2.10]	1.03 [0.76, 1.35]	1.97 [1.57, 2.41]
	4	1	1.08 [0.89, 1.29]	1.79 [1.50, 2.10]	1.07 [0.77, 1.43]	2.04 [1.59, 2.54]
	5	1	1.06 [0.79, 1.33]	1.99 [1.62, 2.40]	1.10 [0.70, 1.53]	2.24 [1.71, 2.85]
	6	1	1.10 [0.68, 1.55]	2.28 [1.77, 2.85]	1.09 [0.54, 1.69]	2.35 [1.69, 3.09]

Panel B. Arbitrary horizons.

Survey	Horizon, qtr		1983-2004	1990-2004
	s	τ	AO/NF	AO/NF
	0	1	0.73 [0.71, 0.77]	0.73 [0.69, 0.77]
	0	2	0.61 [0.57, 0.67]	0.57 [0.51, 0.63]
	0	4	0.51 [0.43, 0.60]	0.45 [0.35, 0.55]
	0	8	0.46 [0.33, 0.63]	0.38 [0.21, 0.58]
	0	12	0.47 [0.29, 0.72]	0.37 [0.14, 0.65]
	0	20	0.47 [0.15, 0.97]	0.38 [0.01, 0.93]
	0	28	0.51 [0.02, 1.34]	0.44 [0.01, 1.50]
	0	40	0.62 [0.01, 1.95]	0.55 [0.01, 2.27]

Table 7 : Out-of-sample Yield Forecast RMSE Ratios

We compare the out-of-sample yield forecasting performance of the AO, NF and RW (random walk) models. We report RMSE ratio 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 2004. 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 Horizons	1983-2004		1990-2004	
		AO/RW	NF/RW	AO/RW	NF/RW
1	1	1.45 [1.38, 1.52]	1.18 [1.13, 1.22]	1.73 [1.64, 1.81]	1.02 [0.95, 1.06]
	4	1.29 [1.25, 1.48]	1.75 [1.29, 3.38]	1.09 [0.92, 1.32]	1.49 [1.01, 3.03]
	8	1.17 [0.95, 1.45]	1.53 [0.99, 3.29]	1.01 [0.75, 1.37]	1.44 [0.87, 3.14]
	20	1.35 [1.04, 1.88]	1.30 [0.66, 2.78]	0.72 [0.31, 1.38]	1.29 [0.55, 3.06]
2	1	1.34 [1.27, 1.41]	1.27 [1.22, 1.31]	1.57 [1.49, 1.65]	1.18 [1.12, 1.23]
	4	1.27 [1.13, 1.45]	1.73 [1.30, 3.01]	1.11 [0.94, 1.33]	1.51 [1.05, 2.84]
	8	1.15 [0.93, 1.44]	1.52 [1.00, 2.90]	1.02 [0.76, 1.37]	1.47 [0.92, 2.86]
	20	1.30 [0.98, 1.90]	1.30 [0.66, 2.65]	0.69 [0.29, 1.34]	1.28 [0.54, 2.94]
4	1	1.26 [1.19, 1.33]	1.21 [1.15, 1.24]	1.48 [1.39, 1.56]	1.24 [1.18, 1.29]
	4	1.24 [1.09, 1.44]	1.66 [1.22, 2.76]	1.09 [0.91, 1.33]	1.50 [1.04, 2.49]
	8	1.12 [0.90, 1.42]	1.49 [0.95, 2.51]	0.99 [0.73, 1.38]	1.46 [0.88, 2.63]
	20	1.24 [0.91, 1.87]	1.26 [0.59, 2.61]	0.67 [0.26, 1.38]	1.20 [0.42, 2.91]
8	1	1.10 [1.03, 1.16]	1.12 [1.06, 1.16]	1.20 [1.12, 1.28]	1.17 [1.10, 1.22]
	4	1.17 [1.01, 1.39]	1.60 [1.07, 2.52]	1.03 [0.85, 1.29]	1.50 [0.95, 2.46]
	8	1.09 [0.85, 1.45]	1.51 [0.88, 2.53]	0.97 [0.69, 1.40]	1.55 [0.87, 2.77]
	20	1.23 [0.88, 1.91]	1.26 [0.48, 2.59]	0.68 [0.25, 1.47]	1.16 [0.26, 2.89]
12	1	1.02 [0.95, 1.08]	1.04 [0.99, 1.08]	1.06 [0.98, 1.14]	1.07 [1.01, 1.11]
	4	1.12 [0.96, 1.37]	1.56 [0.92, 2.47]	0.96 [0.77, 1.25]	1.49 [0.83, 2.46]
	8	1.08 [0.83, 1.45]	1.54 [0.79, 2.63]	0.95 [0.66, 1.41]	1.63 [0.83, 2.91]
	20	1.26 [0.90, 2.02]	1.30 [0.43, 2.71]	0.71 [0.26, 1.54]	1.17 [0.16, 2.97]
20	1	1.05 [0.98, 1.11]	1.04 [0.99, 1.08]	1.03 [0.95, 1.11]	1.03 [0.97, 1.08]
	4	1.10 [0.93, 1.37]	1.55 [0.71, 2.64]	0.88 [0.69, 1.20]	1.49 [0.62, 2.67]
	8	1.07 [0.82, 1.48]	1.59 [0.66, 2.88]	0.92 [0.62, 1.39]	1.79 [0.80, 3.18]
	20	1.31 [0.93, 2.10]	1.35 [0.36, 2.83]	0.77 [0.31, 1.73]	1.17 [0.04, 3.16]
28	1	1.11 [1.05, 1.18]	1.05 [1.00, 1.09]	1.12 [1.04, 1.20]	1.02 [0.96, 1.08]
	4	1.09 [0.92, 1.35]	1.51 [0.55, 2.76]	0.86 [0.66, 1.17]	1.50 [0.51, 2.81]
	8	1.06 [0.81, 1.50]	1.58 [0.53, 3.04]	0.91 [0.61, 1.41]	1.93 [0.84, 3.43]
	20	1.36 [0.99, 2.19]	1.42 [0.37, 2.99]	0.79 [0.33, 1.76]	1.20 [0.03, 3.23]
40	1	1.19 [1.12, 1.26]	1.10 [1.04, 1.14]	1.30 [1.22, 1.38]	1.13 [1.06, 1.18]
	4	1.10 [0.93, 1.36]	1.51 [0.47, 2.94]	0.86 [0.66, 1.16]	1.51 [0.42, 3.02]
	8	1.09 [0.84, 1.55]	1.63 [0.52, 3.12]	0.90 [0.60, 1.39]	2.07 [0.91, 3.67]
	20	1.40 [1.02, 2.21]	1.42 [0.34, 3.08]	0.80 [0.34, 1.73]	1.10 [0.01, 3.25]

Figure 1. The Term Structure of Survey Forecasts

We plot the survey inflation forecasts that we use as inputs in our model. Our sample period is 1970 to 2004 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. A single number, like 4q, indicates a four-quarter forecast horizons from today. Two numbers, like 1q (5q fw), indicates a one-quarter long forecasting horizon starting in five quarters from today.

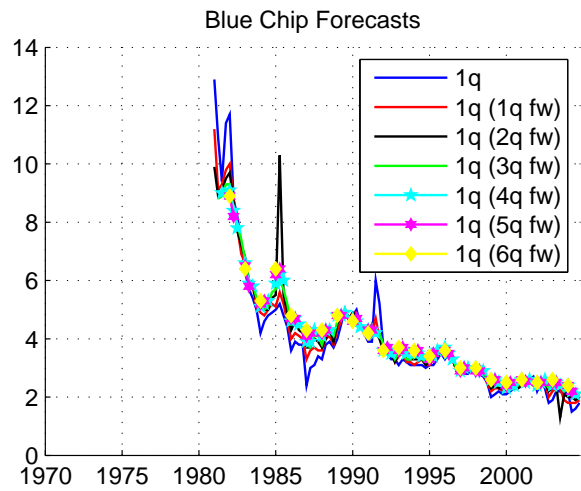
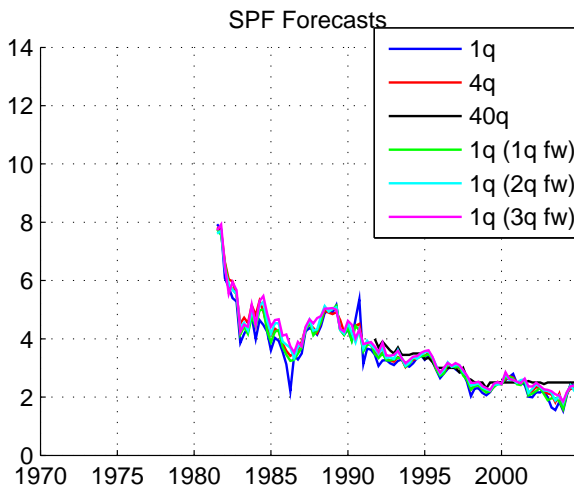
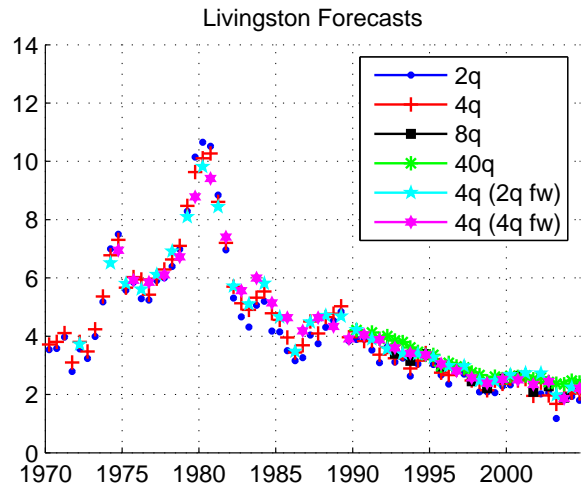
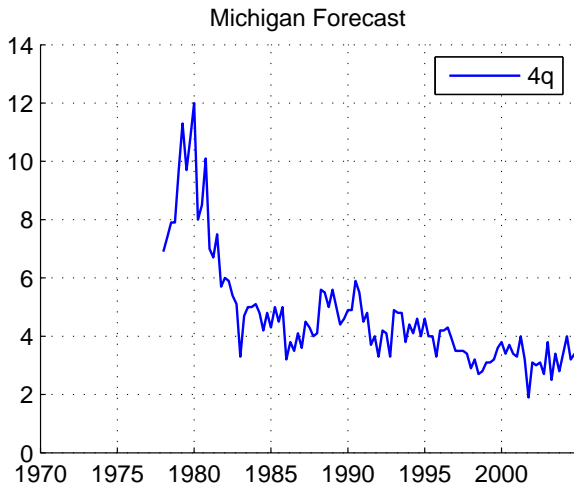


Figure 2. Term Bias

The figure shows unconditional and conditional deviations of survey forecasts from the optimal \mathbb{P} -forecasts. The conditional bias is compared with business cycles (shaded areas) and demeaned inflation (thick black line). The solid dark line in the right column of panels shows a measure of the survey disagreement. The shaded regions show the NBER recessions.

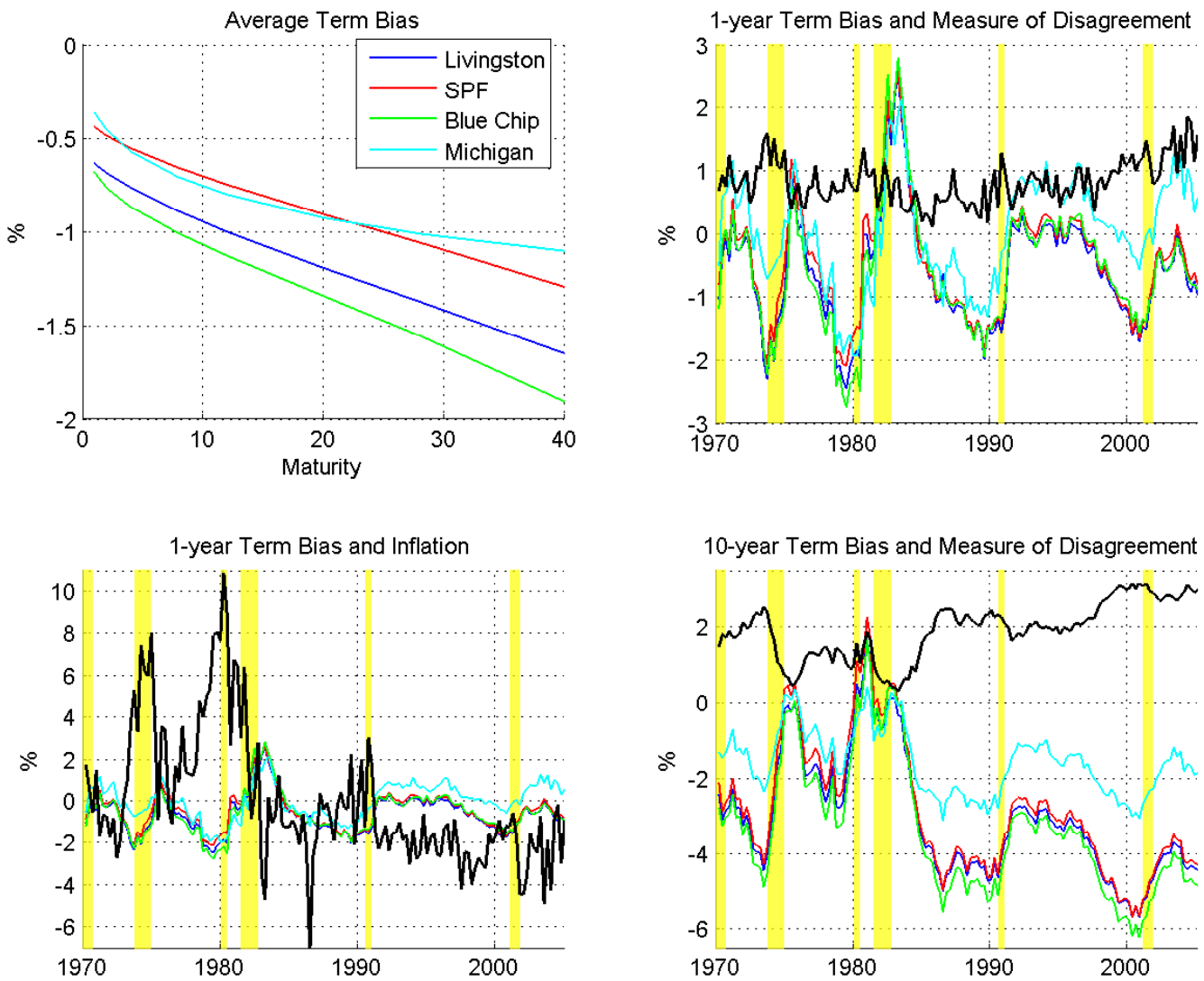


Figure 3. Inflation expectations

The figure shows realized inflation and optimal inflation forecasts computed from three different versions of our model. The shaded regions show the NBER recessions.

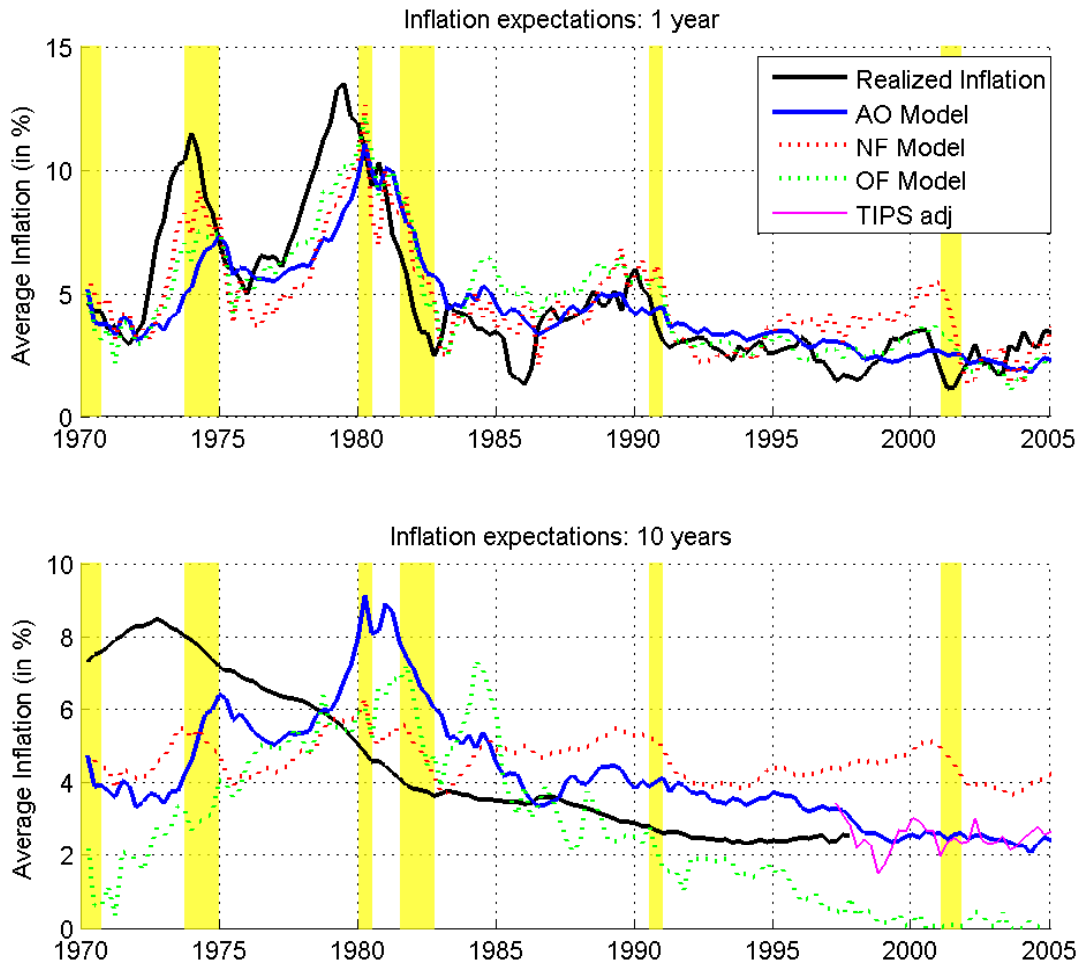


Figure 4. The Term Structure of Inflation Expectations

The figure displays the marginal inflation expectations computed from the AO model. The shaded regions show the NBER recessions.

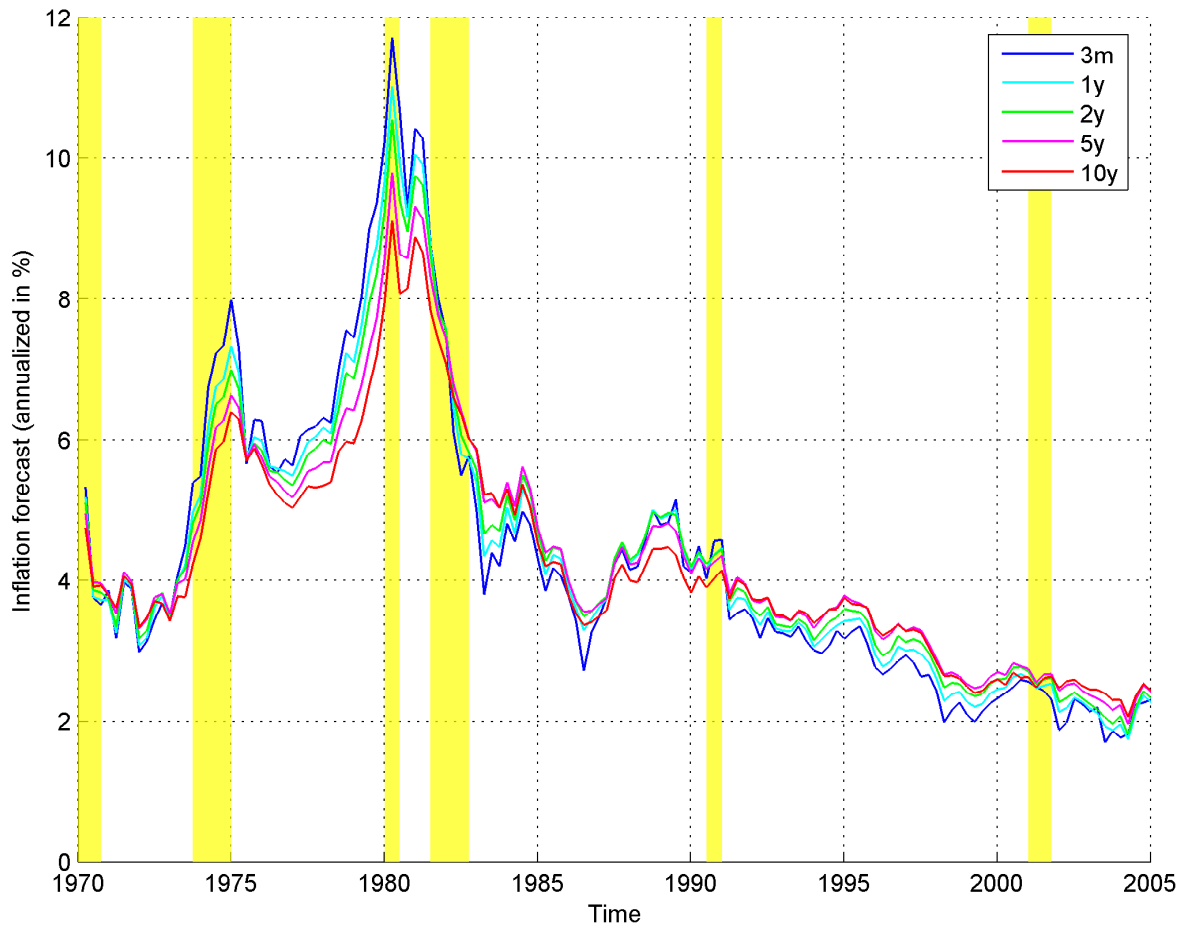


Figure 5. Factor loadings for inflation expectations

We plot factor loadings that are used to compute inflation expectations and multiple horizons (0 to 40 quarters).

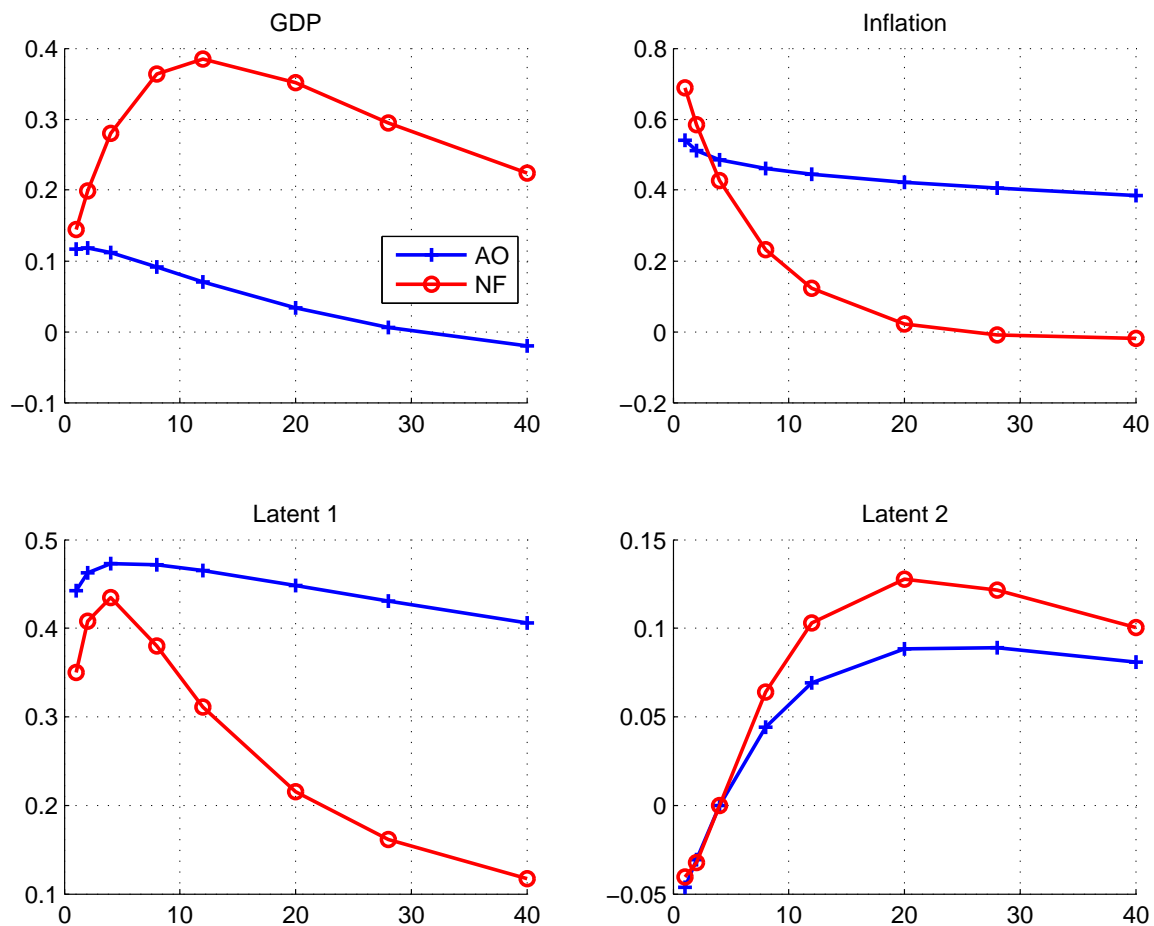


Figure 6. Expectations and Yields decomposition

The figure shows the time series of the one-quarter and ten-year inflation expectations (top panels) and yields (bottom panels) decomposed into the contributions of output, g , and its history, inflation π , and its history, the “survey” residual factor f_1 and the residual factor f_2 . The shaded regions show the NBER recessions. The decompositions are reported for AO only.

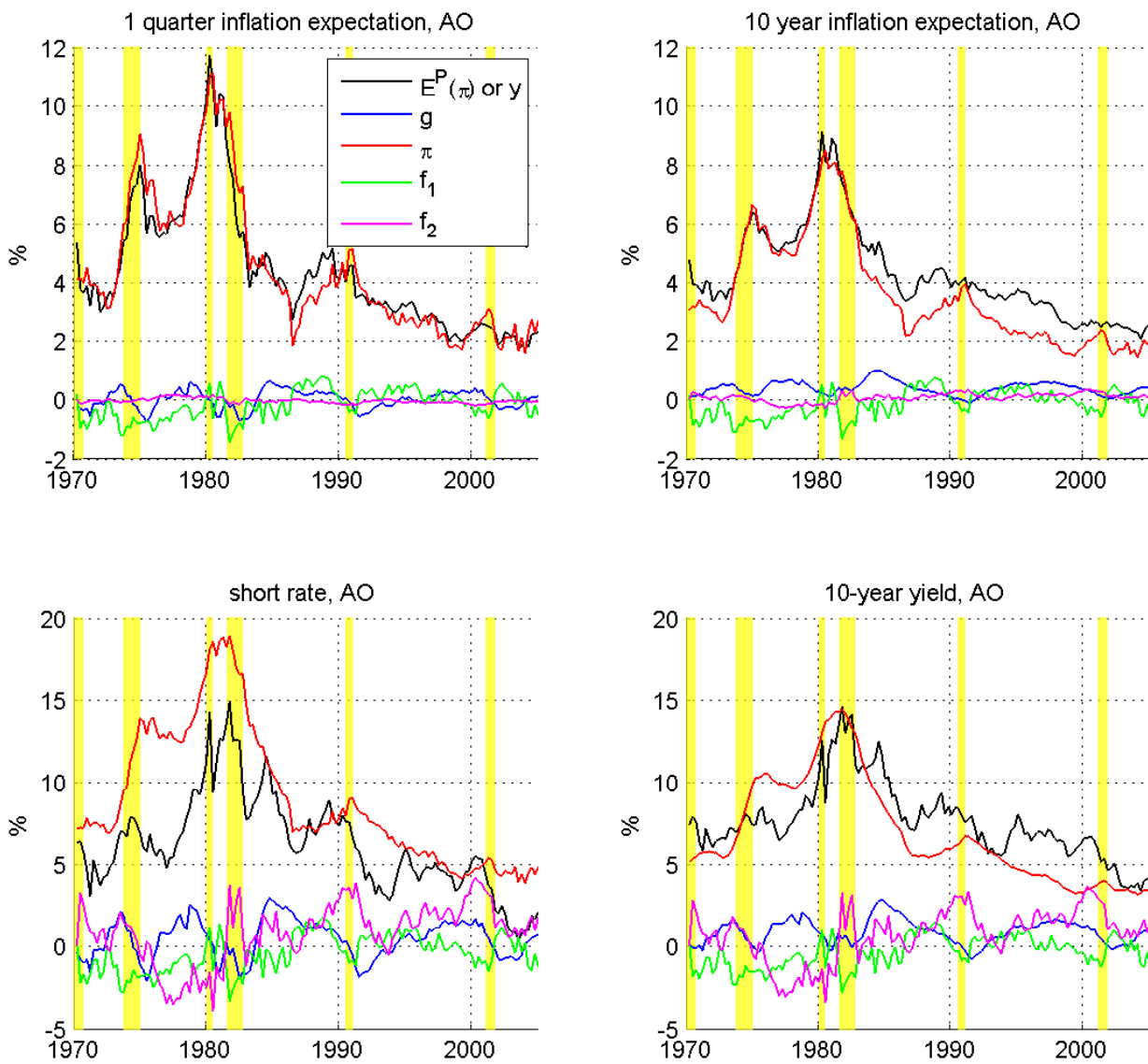


Figure 7. The Fisher equation

The figure plots the time-series of the nominal yield components in the decomposition

$$y_t(\tau) = y_t^R(\tau) + \bar{p}_{t,0}(\tau) + IP_t(\tau),$$

where $y_t^R(\tau)$ is the real yield, $\bar{p}_{t,0}(\tau)$ is the inflation expectation under the \mathbb{P} measure, and $IP_t(\tau)$ is the inflation premium. The shaded regions show the NBER recessions.

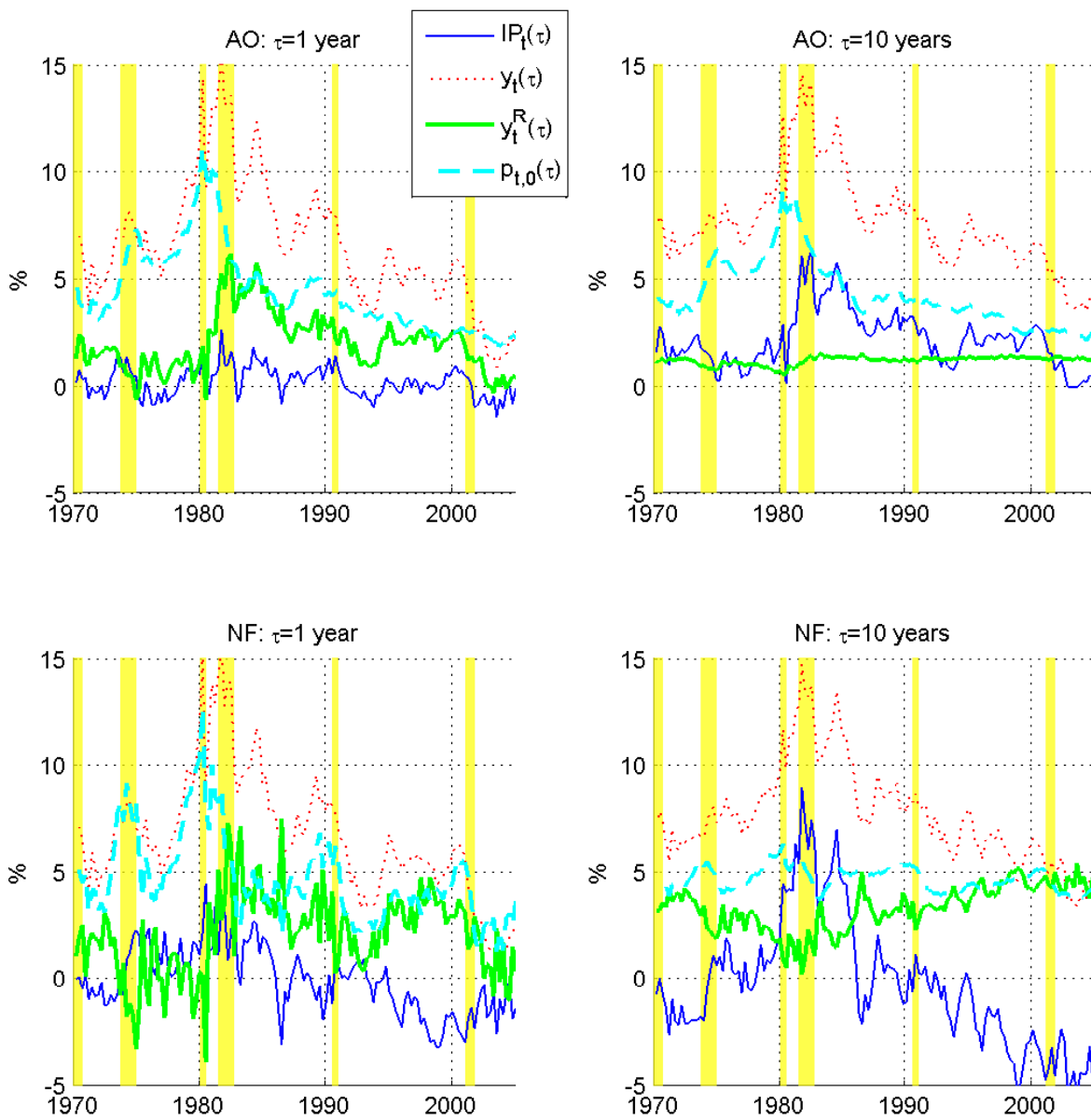


Figure 8. Out-of-sample parameters

The figure plots the time-series of the parameters of the AO and NF models that were reestimated every quarter beginning from 1983. We mostly display parameters pertaining to inflation (dashed lines in panels (a), (b) and (d)) and real activity (solid lines in panels (a), (b) and (d)) because the latent ones are difficult to interpret. Panel (c) provides the highest (solid line) and the lowest (dashed line) eigenvalues of the matrix Φ (see (3.2)) to gauge the persistence of the factors.

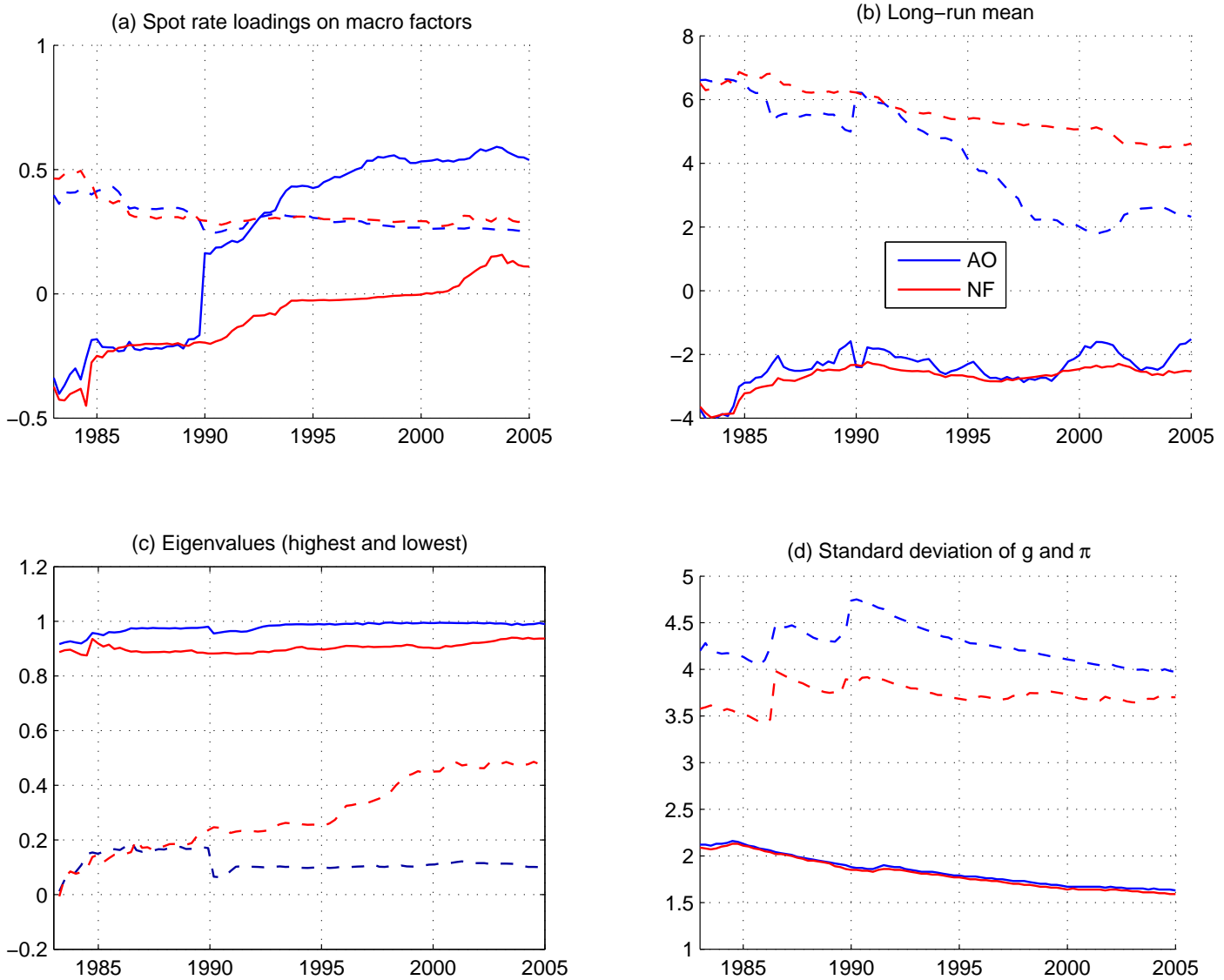


Figure 9. Out-of-sample impulse responses

The figure plots responses of 10-year expected inflation to one basis point shocks in inflation. The responses are computed from AO using parameters estimated using data available at four different dates.

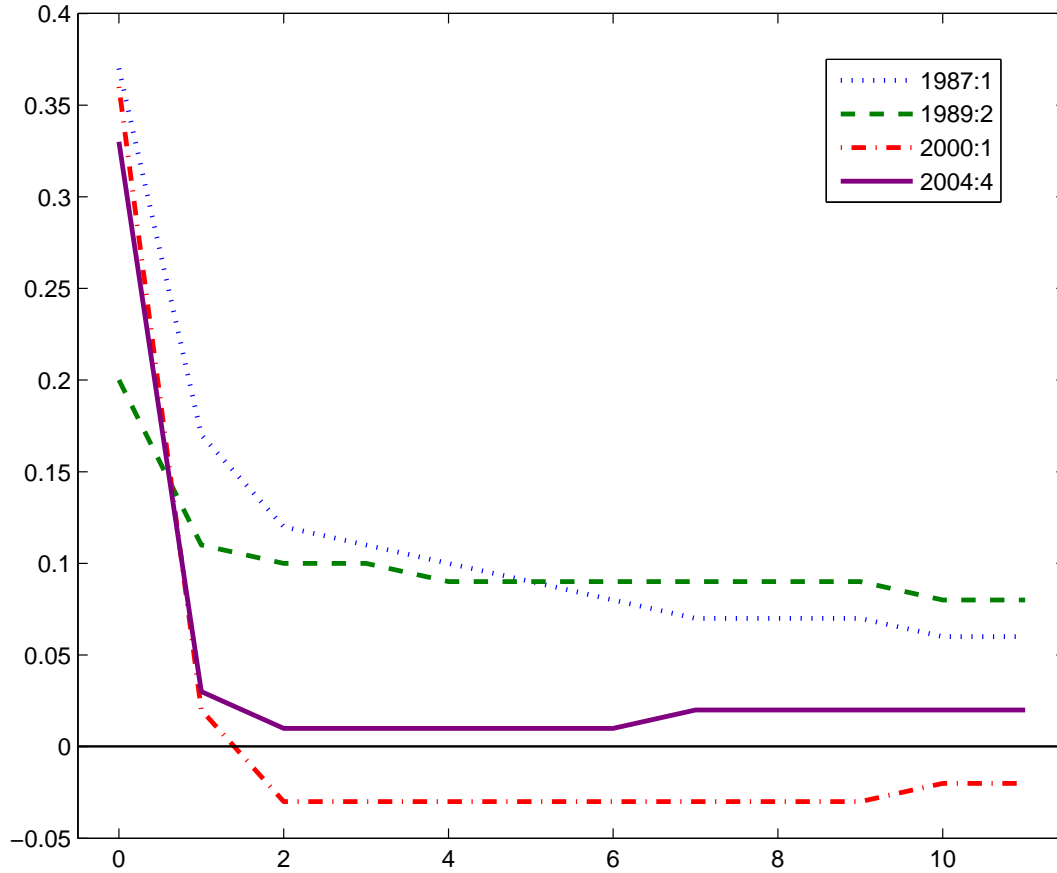


Figure 10. Out-of-sample inflation expectations

The figure shows realized inflation and out-of-sample inflation forecasts computed from AO and NF models. The shaded regions show the NBER recessions.

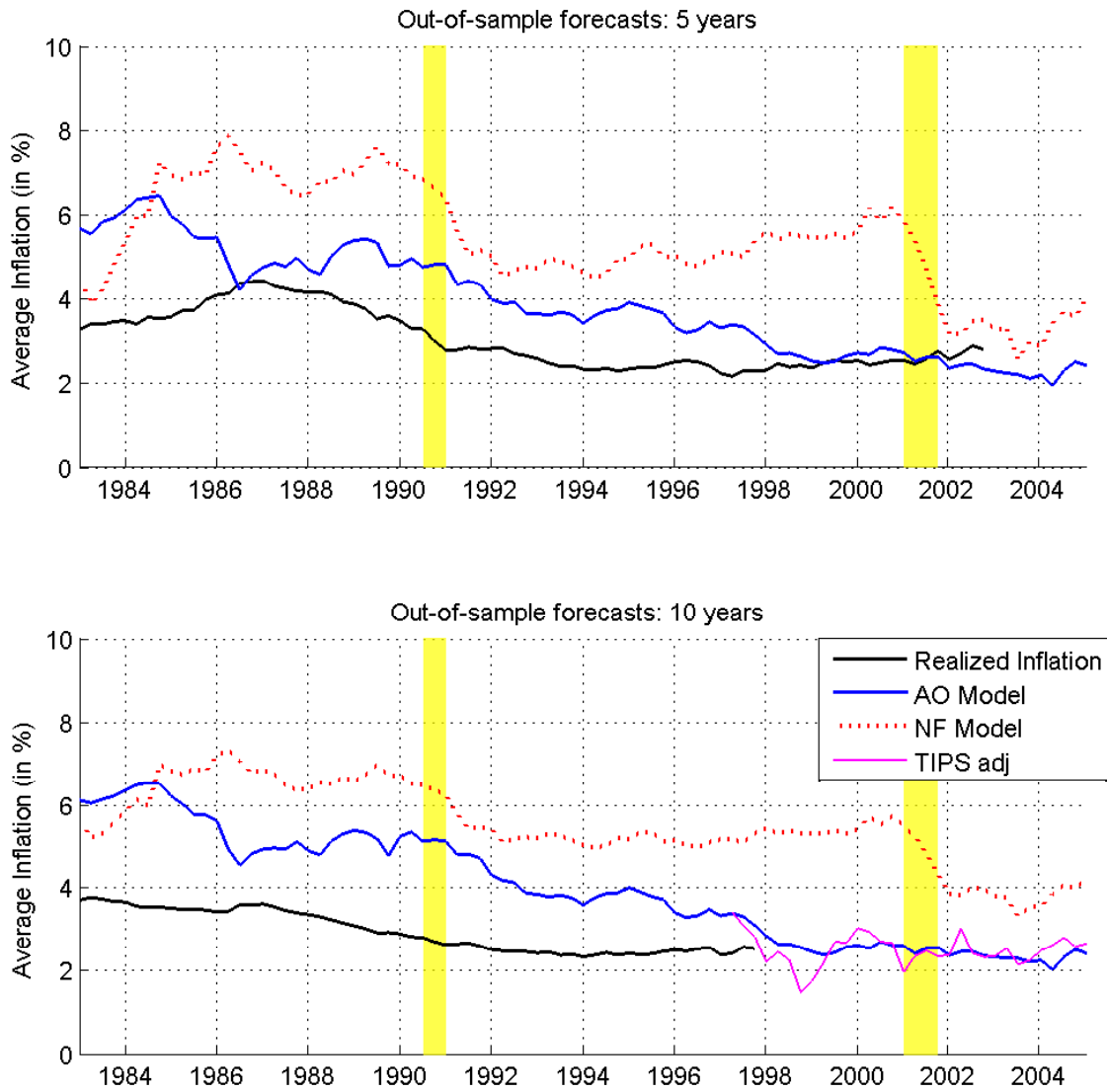


Figure 11. Out-of-sample average term structure of inflation expectations

The figure displays average term structure of inflation expectations, as implied by AO and NF based on the out-of-sample parameter values.

