Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Term Structure Estimation with Survey Data on Interest Rate Forecasts

Don H. Kim and Athanasios Orphanides

2005-48

Term Structure Estimation with Survey Data on Interest Rate Forecasts*

Don H. Kim[†] and Athanasios Orphanides[‡]

Board of Governors of the Federal Reserve System

October 2005

Abstract

The estimation of dynamic no-arbitrage term structure models with a flexible specification of the market price of risk is beset by a severe small-sample problem arising from the highly persistent nature of interest rates. We propose using survey forecasts of a short-term interest rate as an additional input to the estimation to overcome the problem. The three-factor pure-Gaussian model thus estimated with the U.S. Treasury term structure for the 1990-2003 period generates a stable estimate of the expected path of the short rate, reproduces the well-known stylized patterns in the expectations hypothesis tests, and captures some of the short-run variations in the survey forecast of the changes in longer-term interest rates.

KEYWORDS: Dynamic term structure models, survey data, interest rate forecasts, term premia, expectations hypothesis.

JEL CLASSIFICATION: E43, E47, G12

*We thank Jim Clouse, Brian Sack, and Jonathan Wright for helpful inputs and encouragements, and Greg Duffee, Antti Ilmanen, Peter Tinsley and participants at the 11th International Conference on Computing in Economics and Finance, June 23-25, 2005, Washington, D.C., for useful comments. The opinions expressed in this paper do not necessarily reflect those of the Board of Governors of the Federal Reserve System.

[†] Corresponding author. Requests for data used in the estimations in the paper can be sent to the corresponding author. Tel.: (202) 452-3303, e-mail: Don.H.Kim@frb.gov.

[‡] Tel.: (202) 452-2654, e-mail: Athanasios.Orphanides@frb.gov.

1 Introduction

The term structure of interest rates at any moment contains information regarding interest rates that markets expect to prevail later on. This information is of tremendous interest to financial practitioners and policymakers alike. Policymakers carefully monitor expectations of future monetary policy to gauge the effectiveness of their communications strategy. For practitioners, the availability of accurate interest rate forecasts can be the key to a successful trading strategy.

Extracting interest rate expectations from the yield curve, though, rests on the validity of additional necessary assumptions. The expectations hypothesis, for instance, takes the current forward rate curve as the expected path of the short rate by abstracting from the presence of term premia. It is by now widely acknowledged that the yield curve contains time-varying term premia, however, so the expectations hypothesis may give a misleading picture of interest rate expectations.

In recent years, a new generation of dynamic term structure models has focused on flexible modeling of risk premia that allows for interesting departures from the expectations hypothesis. As highlighted by Duffee (2002) and Dai and Singleton (2002), these models have shown considerable promise for capturing the dynamic behavior of the term structure while retaining analytical tractability. In many practical situations, however, the empirical implementation of these models runs into the problem that typical data samples may be too short to provide a reliable characterization of the physical dynamics of the interest rate process. The fundamental difficulty is the highly persistent nature of interest rates. In a term structure sample spanning 5 to 15 years, one may not observe a sufficient number of "mean-reversions" and it may thus be impossible to estimate with precision model parameters related to the drift of the underlying state variables.

One way to overcome this problem is to provide additional relevant information that can facilitate the estimation. To this end, we investigate the usefulness of supplementing actual term structure data with information from surveys of financial market participants' forecasts of a short-term interest rate. The basic idea is that this additional information on the expected short-rate path can help pin down the model parameters pertaining to the physical drift of the state variables underlying the model; together with the information on the risk-neutral drift of the state variables provided by the cross-sectional dimension of the term structure data, this should improve the precision of parameters related to the risk premium.

To the extent that surveys of financial market participants are useful proxies for the market expectations implicit in the term structure at any point in time, they should be a rich source of information that can supplement the available interest rate data for estimation of dynamic term structure model. A critical issue is that surveys may not be informative in this regard.

Indeed, some earlier studies, including Friedman (1980) and Froot (1989), have questioned some aspects of "rationality" of interest-rate survey forecasts, such as their unbiasedness and efficiency characteristics. Cognizant of these concerns regarding the reliability of the surveys, in our approach we posit that the survey forecasts only serve as a noisy source of information on expectations and allow the data to determine the extent of this noise. In doing so, our estimation nests the traditional approach that ignores the information in the surveys as a limiting case. As we explain, this can be easily accommodated within the Kalman-filter-based maximum likelihood estimation approach.

We apply this methodology to estimation of the three-factor pure-Gaussian model of the U.S. Treasury term structure over the 1990-2003 period, supplementing the term structure data used in conventional estimation with survey forecasts of the 3-month Treasury bill yield presented in the *Blue Chip Financial Forecasts* publication.

Our main results can be summarized as follows. We find that the use of survey forecasts of the short-term interest rate helps produce stable and sensible results in the estimation of the three-factor pure-Gaussian model. Problems encountered in "conventional" estimation of the model are greatly alleviated. The model estimated with survey data is also successful in reproducing the well-known pattern of deviation from the expectations hypothesis regressions and generates an implied forecast of long-term interest rates that captures some of the deviations of survey forecasts of long-term interest rates from the expectations hypothesis. We also report Monte Carlo evidence that documents the presence of a substantial bias and imprecision in the parameter estimates in the conventional estimation and the improvement brought by estimation with survey data.

The rest of the paper is organized as follows. Section 2 provides some contexts in which the small-sample issue may arise. Section 3 describes symptoms of the problem that arises in the conventional estimation of the three-factor pure-Gaussian model and discusses the source of the difficulty in uncovering the physical dynamics and potential solutions. Section 4 reviews basic features of the survey forecasts we propose to use as a supplement to the term structure data. We present the Kalman-filter approach for estimating the term structure model with survey data in section 5 and describe the estimation results in section 6, focusing on the differences from the conventional estimation. In section 7 we examine the expectations of long-term interest rates implied by the model and how they compare to the survey forecast of long-term interest rates and the expectations hypothesis. We compare the performance of the conventional estimation and the estimation with (artificial) survey data in a Monte Carlo experiment in section 8 and conclude in section 9.

2 Limited data, structural change, and small samples

The size of the sample that may be usefully employed in estimation of a dynamic term structure model may be limited for a variety of reasons. In some applications, short-sample estimation is unavoidable because the underlying data may simply not exist. For example, the literature on the term structure estimation with LIBOR-swap data has often relied on samples shorter than 10 years due to the short history of that market.[1] On other occasions, a major break in regime or institutional structure may necessitate discarding much of the available sample. Consider, for instance, the creation of the European Central Bank in 1999, and the associated harmonization of short-term interest rates for all nations that adopted the euro as their common currency. While interest rates set by each national central bank earlier in the sample reflected importantly macroeconomic conditions specific to that nation, euro area conditions have become the common determinant of interest rates since then. Arguably, this has rendered most national-currency term-structure data from the 1990s and earlier periods of questionable use for current estimation of dynamic term structure models in the euro area.

Even when a long record of historical term structure data is available without major institutional breaks (as in the U.S. Treasury market), one might wish to focus on a relatively recent subsample out of concern about a structural break in the data-generating mechanism. In the United States, a burgeoning literature has pointed to possible changes in monetary policy, economic structure and/or macroeconomic volatility over the past two decades or so.[2] Business cycles may have become less frequent and milder. The relative importance of the *ex ante* real rate and inflation expectation in the variation of nominal rates may have changed, and the volatility of short-term interest rates appears to have gotten significantly smaller. These changes may have potentially induced a change in the dynamic behavior of the term structure and numerous authors have presented estimates pointing to large differences in the behavior of the term structure over subsamples.[3] However, there is considerable disagreement about the nature and extent of structural change in the U.S. economy.[4] It is not entirely clear that what appears as structural change in simpler models cannot be accommodated within an unchanged structure in a rich, flexibly specified term structure model. But this issue is hard to settle if estimation of these models runs into problems in small samples.

[1]Examples include Dai and Singleton (2000) and Jagannathan, Kaplin, and Sun (2003). Gupta and Subrahmanyam (2000) argue that swaps may have been (incorrectly) priced without convexity premium adjustments up to early 1990s; this may mean that the length of the "correctly" priced swap data is even shorter.

[2]See e.g., Kim and Nelson (1999), Clarida, Gali and Gertler (2000) and Cogley and Sargent (2001).

[3]These studies include Campbell and Viceira (2001), Goto and Torous (2003), and Rudebusch and Wu (2004).

[4]With regard to monetary policy, for example, the evidence of dramatic changes reported by Clarida, Gali and Gertler (2000) and Cogley and Sargent (2001) stands in sharp contrast to the subtle changes presented by Orphanides (2004) and Sims and Zha (2004).

Determining the evolution of long-horizon expectations of interest rates beyond business cycle frequencies is particularly vexing. The term-structure literature has not studied the long-horizon expectations of the short rate in detail,[5] possibly because it is difficult to assess the evolution of these expectations by purely econometric or theoretical means. Though long-horizon expectations may be hard to pin down, one can still point to information suggesting changes in the *range* of reasonable values over the past few decades. For example, when the Federal Reserve began to tighten policy in the summer of 2004, much of the market discussion and press commentary centered on the "neutral level" of the federal funds rate, which can be viewed as the rate expected to prevail at the medium-to-long term. Since then, some statements from policymakers suggested the range of 3.5 to 4.5 percent and market commentary has seldom been outside the range of 3 to 5 percent.[6] We can contrast this situation with that prevailing during the 1980s when interest rates were generally much higher. The federal funds rate was always above this range during the 1980s, and by a considerable margin. Indeed, with the exception of two months in the fall of 1986, the fed funds rate exceeded 6 percent throughout the 1980s and market participants certainly would have regarded recent estimates of the neutral rate as too low. It thus appears that perceptions of the "neutral level" have changed over time, and not insignificantly.

The reduction in perceptions regarding medium-to-long-term expectations of interest rates since the late 1980s is also consistent with the apparent evolution of the Federal Reserve's medium-term inflation objective during this period. Following the Volcker disinflation, it appeared that Federal Reserve policy during the 1980s allowed inflation to stabilize at a level close to 4 percent, resisting incipient increases, but waiting for favorable supply shocks and unforeseen recessions to deliver the additional desired reduction in inflation.[7] With the 1990-91 recession delivering the desired convergence to price stability, inflation has since remained at a level closer to 2 percent—a full 2 percentage points lower than what might have been expected to prevail in the late 1980s.

The evolution of long-term inflation expectations provides additional clues about how comparable interest-rate expectations may have evolved over the past two decades. Figure 1 shows two survey measures of long-term inflation expectations, one from the Federal Reserve Bank of Philadelphia Survey of Professional Forecasters (SPF) and the other from the Michigan survey

[5] A notable exception is Kozicki and Tinsley (2001).

[6] See, for example, Derby (2004). He reviews comments by several policymakers and cites Yellen (President of the FR Bank of San Francisco) as suggesting the 3.5 to 4.5 percent range. More recently, Hoenig (2005) (President of the FR Bank of Kansas City) reiterated that "most estimates of the neutral fed funds rate fall within a range of 3.5 to 4.5 percent." (p. 8).

[7] This strategy came to be known as the "opportunistic approach" to disinflation. See Orphanides and Wilcox (2002) for additional details and related policymaker citations describing this approach.

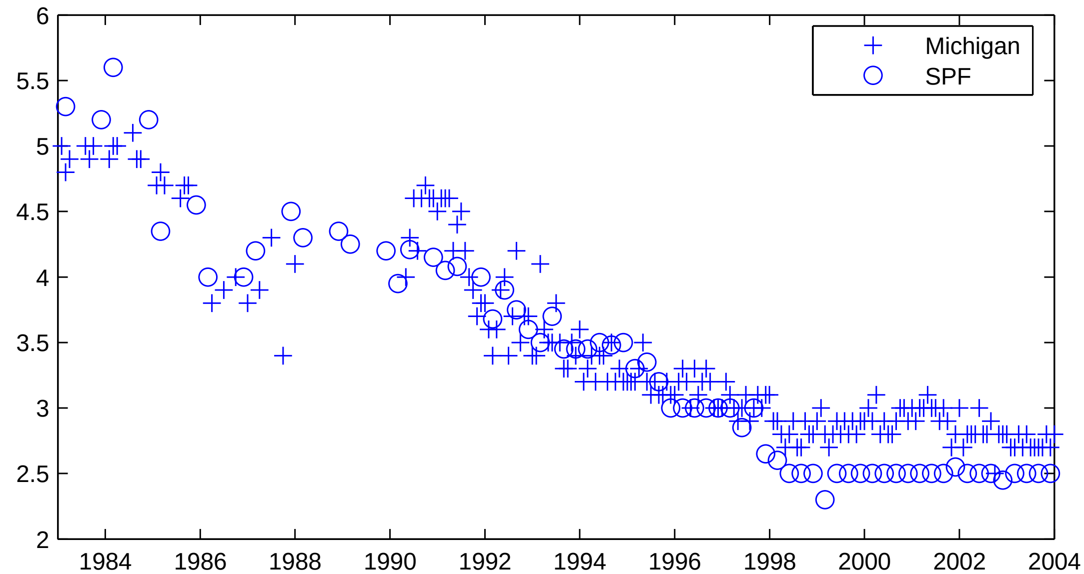


Figure 1: Long-term inflation forecasts from the Michigan survey of households (inflation forecast over the next 5 to 10 years) and the Survey of Professional Forecasters (10-year inflation forecast).

of households. Despite the differences in the forecasted object and the nature of participants, both measures give a clear indication that the long-term inflation expectation has gradually declined over time.[8] In light of this substantial change in long-term inflation expectations, one might reasonably suspect that the long-term expectation of the short-term nominal interest embedded in the term structure of interest rates has also changed significantly over the sample.

To produce the kind of variation suspected for the "long-term" expectation of the short rate, a dynamic term structure model would need a persistent factor with a fairly long half-life, say, longer than 5 years. Unfortunately, this very persistence makes the conventional estimation of the model especially difficult and unreliable. Next, we review how this problem is manifested in conventional estimation of a flexible dynamic term structure model with data on U.S. Treasury securities.

3 Conventional estimation of a dynamic term structure model

We focus our attention on the three-factor pure-Gaussian model with a flexible specification of market price of risk (the "essentially affine" $EA_0(3)$ model, in Duffee's (2002) terminology). Duffee (2002), Dai and Singleton (2002), and Duarte (2004) have already estimated this model with relatively long samples of US Treasury term structure data, and noted that this model

[8]The SPF is a forecast of the CPI, while the Michigan survey is a forecast of the prices of "things that you buy"; the SPF survey participants are business forecasters, while the Michigan survey participants are consumers (households).

can reproduce the stylized facts from the expectations hypothesis (Campbell-Shiller) regression, and forecasts future interest rates (in-sample) better than other affine models (e.g., $EA_1(3)$ and $EA_2(3)$). Though this model has certain shortcomings, in particular, it implies time-invariant volatility of yields, it is a natural model to consider here, as our paramount concern is extracting interest rate expectations from the yield curve.

A general n-factor pure-Gaussian model specifies the short rate as

$$r_t = \rho_o + \rho' x_t, \tag{1}$$

where ρ_o is a constant, ρ is an n-dimensional constant vector, and x_t is an n-dimensional vector of state variables which follow the multivariate Ornstein-Uhlenbeck process:

$$dx_t = \mathcal{K}(\mu - x_t)dt + \Sigma dB_t \tag{2}$$

in which $\mathcal{K}$ and Σ are $n \times n$ constant matrices, μ is an n-dimensional vector, and B_t is an n-dimensional standard Brownian motion. The n-dimensional vector of the market price of risk of the state variables is given by

$$\lambda_t = \lambda_a + \Lambda_b x_t, \tag{3}$$

where λ_a is an n-dimensional vector, and Λ_b is an $n \times n$ constant matrix. In this model, the yield on a zero-coupon bond with time-to-maturity τ is given by

$$y_{t,\tau} = a_\tau + b'_\tau x_t, \tag{4}$$

in which the factor loadings a_τ, b_τ (given in Appendix A) are functions of $\mathcal{K}^*, \mu^*, \Sigma, \rho, \rho_o$, where

$$\mathcal{K}^* = \mathcal{K} - \Sigma\Lambda_b \tag{5}$$

$$\mu^* = \mathcal{K}^{*-1}(\mathcal{K}\mu - \Sigma\lambda_a). \tag{6}$$

Implementation of the model requires a normalization to rule out "invariant transformations" (see, Dai and Singleton (2000)). To that end we choose the following normalization: We restrict $\mathcal{K}$ to be lower-triangular, Σ to be diagonal, $\mu_1 = \mu_2 = \mu_3 = 0$, and $\rho_1 = \rho_2 = \rho_3 = 1$.

The expected short rate w-periods ahead implied by the model is given by[9]

$$E_t(r_{t+w}) = \rho_o + \rho'(e^{-\mathcal{K}w} x_t + (I - e^{-\mathcal{K}w})\mu). \tag{7}$$

Following most implementations, in estimation we assume stationarity so that the eigenvalues of $\mathcal{K}$ are all positive. In the limit $w \to \infty$, we have $e^{-\mathcal{K}w} \to 0$, so the behavior of the infinite-horizon expectation, $E_t(r_{t+\infty})$, is trivially a constant. Although this contrasts sharply with

[9] Here and elsewhere in the paper, the notation e^X, where X is a square matrix, denotes the matrix exponential $e^X = I + X + X^2/2 + X^3/6 + \cdots$.

models that require the infinite-horizon expectation of the short-term interest rate to vary over time, we note that, as a practical matter, the stationary model we consider is sufficiently flexible that it can accommodate considerable time variation in "long-horizon" forecasts (say 5 to 10 years) and it may be hard to distinguish from nonstationary models even at such long horizons.[10]

This model can be estimated with the Kalman-filter-based maximum-likelihood method (e.g., de Jong (2000), Duffee and Stanton (2004), Kim (2005)) or with the maximum likelihood method used by Duffee (2002) and Dai and Singleton (2002). We refer to estimation using only Treasury yield data with either of these ML methods as "conventional" estimation, but in what follows we concentrate on the Kalman-filter-based approach which facilitates direct comparison of estimation with and without the use of survey forecasts in our setting.

For our conventional estimation, let the observation vector be

$$o_t = [y_{t,\tau_1}, ..., y_{t,\tau_M}]' \tag{8}$$

where $\tau_1, ..., \tau_M$ denote a set of M yields used in the estimation. Then,

$$o_t = \begin{bmatrix} a_{\tau_1} \\ \vdots \\ a_{\tau_M} \end{bmatrix} + \begin{bmatrix} b'_{\tau_1} \\ \vdots \\ b'_{\tau_M} \end{bmatrix} x_t + \begin{bmatrix} \eta_{t,\tau_1} \\ \vdots \\ \eta_{t,\tau_M} \end{bmatrix}, \tag{9}$$

in which $\eta_{t,\tau_1}, .., \eta_{t,\tau_M}$ denote yield measurement errors. Eq. (9), together with the equation

$$x_t = e^{-\mathcal{K}h} x_{t-h} + (I - e^{-\mathcal{K}h})\mu + \epsilon_t, \tag{10}$$

with $\epsilon_t \sim N(0, \Omega_h)$, where $\Omega_h = \int_0^h e^{-\mathcal{K}s} \Sigma\Sigma' e^{-\mathcal{K}'s} ds$,[11] constitute the basic Kalman filter observation equation and state equation. The construction of the likelihood function and the estimation is then straightforward; see, for example, de Jong (1999), Duffee and Stanton (2004), and Kim (2005).

3.1 Estimation over two alternative samples

Figure 2 compares estimation results of the $EA_0(3)$ model presented above over two samples: A long sample using nearly forty years of data, 1965-2003, and a shorter sample with 1990-2003 data. We refer to these two estimates of the model as NS1965 and NS1990, respectively. The figure suggests that the two estimates provide quite different pictures of the evolution of interest

[10] Kozicki and Tinsley (2001) present a model with a time-varying infinite-horizon expectation of the short-term interest rate, $E_t(r_{t+\infty})$. They mention that their work is in a spirit similar to the "time-varying tendency model," in which the short rate r_t mean-reverts to a slower-moving target θ_t: $dr_t = \kappa_r(\theta_t - r_t)dt + \sigma_r dB_{rt}$, $d\theta_t = \kappa_\theta(\mu - \theta_t)dt + \sigma_\theta dB_{\theta t}$. This time-varying tendency model is nested in our specification.

[11] This integral can evaluated to give $vec(\Omega_h) = -((\mathcal{K} \otimes I) + (I \otimes \mathcal{K}))^{-1} vec(e^{-\mathcal{K}h} \Sigma\Sigma' e^{-\mathcal{K}'h} - \Sigma\Sigma)$.

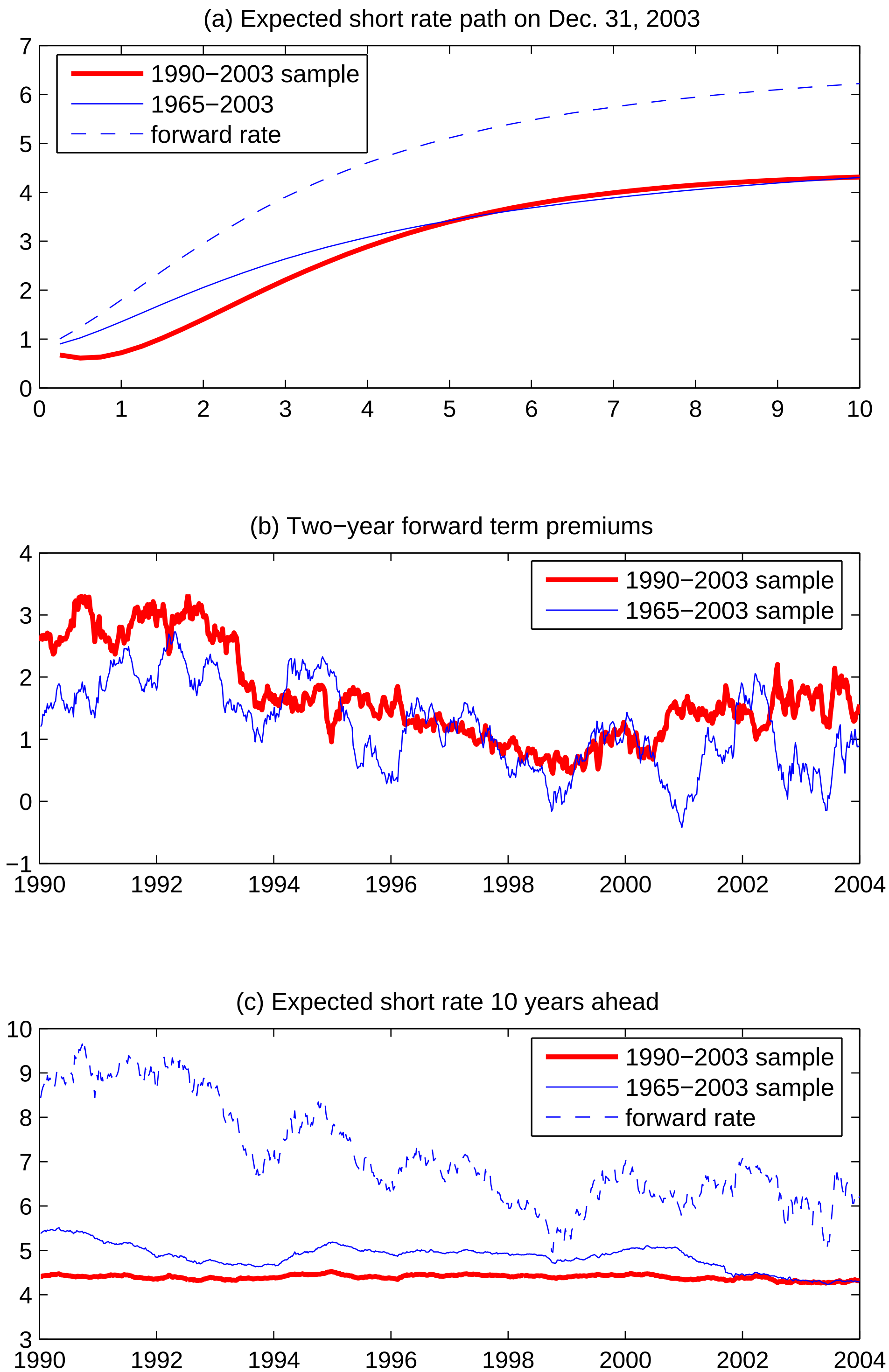


Figure 2: Conventional model estimation: (a) The expected paths of the short rate on Dec. 30, 2003, based on the conventional (maximum-likelihood) estimation with the 1990-2003 and 1965-2003 samples. (Horizon in years shown on axis.) The corresponding forward rate curve is also shown for comparison (dashed line). (b) Two-year forward term premium. (c) Ten-year-ahead short-rate expectations and the forward rate time series.

rate expectations and risk premia since 1990. In principle, these differences could be the result of structural change. However, the estimates of the shorter sample appear unreasonable in some respects, raising suspicion of an estimation problem likely arising from the nature of the shorter sample. We discuss some symptoms of the small-sample problem in these estimates below.

The top panel of the figure shows the expected path of the short rate at the end of 2003. Since this corresponds to the last data point in our estimation sample, this path represents a real-time out-of-sample forecast of the short rate. The instantaneous forward rate, which we denote $f_{t,w}$ (for time t and horizon w), is also shown for comparison. According to the naive version of the expectations hypothesis, the forward rate curve would be the expected short rate path, but few people (policy makers and market participants) at the time would have believed the implications of the expectations hypothesis for long horizons: short rate of more than 6% in 10 years would have seemed too high. On the other hand, few would have taken seriously the expected short rate from the model estimated with the short sample (NS1990) either: the short rate is projected to rise *too slowly*. Our reading of market participants' views based on anecdotal reports and market commentary is that rates were expected to rise faster than the thick solid line in Figure 2a, as indeed turned out to be the case in 2004 and 2005.[12]

The "unreasonable" short rate forecast is not a feature of the specification of the model but a feature of the specific estimates obtained with conventional estimation over the shorter sample. As can be seen, the estimates shown in the figure that are obtained with the longer sample (NS1965) suggest a more reasonable forecast, notwithstanding the concern that the presence of structural change over the longer sample may have introduced some bias.

Signs of the small-sample problem can be found not only in the real-time forecast but also in the model-implied times series of the expected short rate. Figure 2b shows the model-implied forward term premium $\varphi_{t,w}$ (the difference between the forward rate and the expected short rate),

$$\varphi_{t,w} = f_{t,w} - E_t(r_{t+w}), \tag{11}$$

for $w = 2$ years. Note that the NS1990-based $\varphi_{t,2y}$ is quite large in the early 1990s and between 2001-2003, while the NS1965 estimate shows substantially smaller values in these periods. Because the federal funds rate indeed declined or stayed low during these periods, the term premium estimate based on the short sample implies a smaller in-sample forecasting error. However, this "better" forecasting performance is likely an artifact of the statistical procedure (look-ahead bias) that becomes severe in small samples. Again, anecdotal evidence (reports from market participants, surveys, press commentaries about interest rates) would

[12]At that time (end of 2003), the federal funds rate had been at the historically lowest level of 1%, macroeconomic data were predominantly positive and inflationary concerns were receiving increasing attention in the financial press, pointing to a sustained period of monetary policy tightenings.

not support the case of such high term premia. In addition, information on the inflation forecasts of FOMC members and business forecasters suggests that both policymakers and market participants tended to overpredict inflation in the 1990s.[13] In other words, a significant part of the disinflation experienced during the 1990s may have been *unanticipated* which, in turn, suggests that a significant part of the decline in the short rate may have been unanticipated as well.[14]

Perhaps an even more striking sign of the small sample problem is the stability of the long-horizon short rate expectation based on the NS1990 estimation (thick solid line in Figure 2c). The estimated model attributes *almost all* of the variation in the forward rate for long horizons to the variation in term premia. The level of the long-horizon expected short rate in 1990 is not much different from that in 2003. Given the evolution of long-horizon inflation expectations and beliefs regarding reasonable estimates of the neutral interest rate (as discussed earlier), this seems implausible.

Estimation with a sample that is "too short" presents technical problems as well. In particular, the likelihood function seems to have multiple inequivalent local maxima which have similar likelihood values but substantially different implications for economic quantities of interest.[15] Furthermore, the likelihood function around these maxima can be quite flat along many directions in the parameter space; the standard errors for the parameters and quantities of economic interest are often too large to be useful. A common practice is to restrict several parameters with large standard errors to zero and re-estimate the model. (See e.g. Duffee (2002), Dai and Singleton (2002), and Duarte (2004).) This produces somewhat smaller confidence intervals for objects like model-implied interest rate forecasts. The point estimates in the original case and in the modified (re-estimated) case often produce similar economic implications.[16] Should the implications differ, it is unclear which is "better." The issue is that setting some parameters to zero and reestimating is a rather arbitrary procedure, since the individual parameters in a flexibly specified model often do not have a simple economic meaning. Indeed, the procedure is sensitive to the normalization employed in estimation: A model with a smaller number of free parameters (obtained by setting some parameters to zero) in one normalization would not necessarily have the same number of free parameters in a different normalization. More generally, it is unclear how one should determine how small the associated standard errors have to be in order to set a parameter to zero. Setting parameters to zero simply because they are estimated

[13]See, for example, Kohn (1999) and Croushore (1998).

[14]The accounts of Blinder and Yellen (2001) and Meyer (2004), who served on the Federal Reserve Board during the 1990s, are also consistent with this view.

[15]By inequivalent local maxima, we mean the maxima which are not connected by an invariant transformation (such as rotations, permutations, etc.) of the state variables.

[16]For example, Duffee (2002) presents results for both cases, which lead to similar conclusions.

imprecisely also risks introducing significant biases in the resulting estimated model.[17]

The plots in Figure 2 for the 1990-2003 sample estimation are based on the parameter estimate with the highest likelihood that we could find, re-estimated with some of the parameters with large standard errors to zero. We have also tried different choices of parameters to be set to zero and other local maxima. Most of these other estimates had similar qualitative problems as reported here (too slow expected short rate path at the end of 2003, too large forward term premia, too stable long term expectations), but some exhibited economically significant quantitative differences. The risk that different researchers might arrive at different conclusions with the same specification (due to minor differences in the data, differences in initial parameter guesses leading to different local maxima, and different choices in reducing the dimension of the model) is certainly an unsatisfactory situation.

3.2 The difficulty of uncovering the physical dynamics

The source of these difficulties is that the sample (1990-2003) is too short to provide accurate information about the *physical* dynamics of the state variables underlying the model. From a cross-section of the term structure (yield curve) the risk-neutral parameters $\mathcal{K}^*, \mu^*$ (and Σ) can be determined, but not the physical parameters $\mathcal{K}, \mu$, which enter the bond pricing equation only in combination with the market-price-of-risk parameters λ_a, Λ_b.

The physical parameters can be determined (heuristically) as follows.[18] Given the values of $\mathcal{K}^*$, μ^* and Σ, one can "back out" the implied state-variables $\hat{x}_t$ by solving a system of eq. (4) consisting of n yields (or by minimizing the least squares deviation when there are more than n yields). With the time series $\hat{x}_t$ given, one can obtain the physical parameters $\mathcal{K}, \mu$ by estimating a VAR

$$\hat{x}_t = A(\mathcal{K})\hat{x}_{t-h} + b(\mathcal{K}, \mu) + \epsilon_t, \tag{12}$$

in which $A(\mathcal{K}) \approx I - \mathcal{K}h$, $b(\mathcal{K}, \mu) \approx \mathcal{K}\mu h$, and $\epsilon_t \sim N(0, h\Sigma\Sigma')$. The market price of risk parameters can be obtained by computing Λ_b, λ_a from eq. (5) and (6), i.e., $\Lambda_b = \Sigma^{-1}(\mathcal{K} - \mathcal{K}^*)$, $\lambda_a = \Sigma^{-1}(\mathcal{K}\mu - \mathcal{K}^*\mu^*)$.

The problem is that $\mathcal{K}$ in eq. (12) cannot be measured accurately when $\mathcal{K}$ is unrestricted (other than the stationarity and normalization conditions) and the sample period is relatively

[17] For example, consider estimating the standard Vasicek model with data on long-term yields:

$$dr = \kappa(\mu - r_t)dt + \sigma dB_t, \quad \lambda_t = \lambda.$$

The market price of risk λ is usually estimated imprecisely. Setting $\lambda = 0$ and re-estimating the model in this case, however, could bias the long-run mean estimate (μ), because the average level of yields depend on λ.

[18] These steps heuristically describe the maximum likelihood estimation in Duffee (2002) and Dai and Singleton (2002).

short (no matter how often $\hat{x}_t$ is sampled).[19] As noted earlier, to capture the characteristics of the interest rate process, the system should contain a slow-moving factor (with half-life of 5 years or more), which might show very few mean reversions in a sample of 10 to 15 years. Furthermore, over the past two decades, there were only two recessions (1990-91, 2001), which may not be sufficient to determine accurately variation in the term structure over the business cycle. The resulting difficulties in estimating the drift of the persistent factors would then feed into other aspects of the estimation. It is also well known that estimation of equations like (12) may result in an upward bias in the estimated mean-reversion of the system.[20] In other words, the estimation tends to understate the persistence of the process, hence implying a quicker convergence of the short-rate path to the asymptotic value. This results in an overstatement of the variation of term premia and understatement of the variation in long-horizon expectations of the short-term interest rate implied by the model, both symptoms evident in our short sample estimates in Figure 2.

Note that if Λ_b were known or could be sufficiently restricted, then the problem would not arise because $\mathcal{K}$ in eq. (12) would be constrained and more effectively pinned down. For example, if $\Lambda_b = 0$, we have $\mathcal{K} = \mathcal{K}^*$; since $\mathcal{K}^*$ is determined reasonably well without a long time series, we would not expect difficulties in this case.[21] These are special cases, however, and do not extend to the more empirically relevant case with a flexibly specified market price of risk.

Although this discussion was based on the three-factor pure-Gaussian model, we note that the underlying arguments are more general, and do not depend on yields being a linear function of risk factors, nor on the risk factors following a multivariate Gaussian process. In general, similar problems would be expected in any model with a persistent factor and a flexible specification of the factor dynamics and of the market price of risk.[22]

3.3 Potential remedies

It is useful to distinguish among two related problems associated with small samples in our context. One issue is the potential presence of a bias due to the size of the sample. Once

[19]Basically, accuracy depends more critically on the length of the observation period rather than the number of observations, analogous to Merton's (1980) point about estimating the expected return on stocks.

[20]Note that the one-factor-version of eq. (12) is just an AR(1) model, $x_t = \phi x_{t-1} + c + \epsilon_t$, in which case the bias in the ϕ estimate is given by $-(1 + 3\phi)/T$ (if the value of the constant term c is known). We are not aware of analytical results applicable to the present case, as the closeness to the unit-root situation and the multifactor nature of the three-factor pure-Gaussian model make the bias assessment more complicated.

[21]Indeed, this example would be similar to that in Ball and Torous (1996), where the use of cross-sectional yield curve information helps to identify the time series parameters of the short rate process better. Duffee and Stanton (2004) emphasize the specialness of the Ball and Torous (1996) case.

[22]Duffee and Stanton (2004) provide an example of small-sample problems in a model with a stochastic volatility factor.

identified, the presence of such a bias may be addressed with a Monte Carlo treatment designed to provide an appropriate bias adjustment.[23]

A thornier issue is the potential *imprecision* and *lack of robustness* in parameter estimates that arises from multiple inequivalent local maxima or from a flat likelihood function around those maxima. Obviously, this issue cannot be addressed by "two-step" procedures adjusting for bias. The "first step" parameter estimates to which a bias adjustment would be applied may be poor; indeed, a greater concern could be that different researchers may arrive at different estimates in the first step of a two-step procedure rather than what bias correction should be applied.

Alleviating the imprecision and lack of robustness associated with a small sample requires additional information. This "information" could take the form of a *model constraint* or additional *data*. In the absence of concerns of structural change, if a longer data sample is available, its use would provide the most straightforward way to incorporate new information. If a longer sample is not available or suspicion of structural change risks introducing additional problems with it, additional information corresponding to the original shorter sample must be sought. In what follows, we advocate using survey data on interest rate forecasts as such an additional source of information.[24]

4 Survey forecasts of interest rates

Surveys of experts in the financial community can potentially serve as a rich source of additional information for estimating a dynamic term structure model, provided the survey forecasts are useful proxies for the interest rate expectations embedded in the actual term structure when the survey is conducted. At the same time, the use of survey forecasts could raise concerns for several reasons. The rationality and efficiency of forecasts reported in any particular survey might be questioned. Furthermore, in our context specifically, it is not entirely clear that the reported measures of survey expectations *should* correspond to the interest rate expectations implicit in the market prices of financial instruments. Survey forecasts provide a measure of *average* expectations, which might not necessarily correspond to the expectation of the *marginal investor* relevant for observed market prices.[25] There are stylized models where expectations

[23]For example, Bekaert et al (1997)'s study includes a VAR estimation (with a short-term interest rate and a term spread) that corrects for the small-sample bias using a Monte-Carlo methodology.

[24]There are potentially other kinds of data that may be useful. For example, options data might help, as suggested by Phillips and Yu (2005). However, their example is rather special (they consider a CIR model with *zero* market price of risk); the case of richer models has not been investigated in this regard.

[25]Mishkin (1981) expresses this view as follows: "Not all market participants have to be rational in order for a market to display rational expectations. The behavior of a market is not necessarily the same as the behavior of the average individual. As long as unexploited profit opportunities are eliminated by some participants in a market (this is analogous to an arbitrage condition), then the market will behave as though expectations are

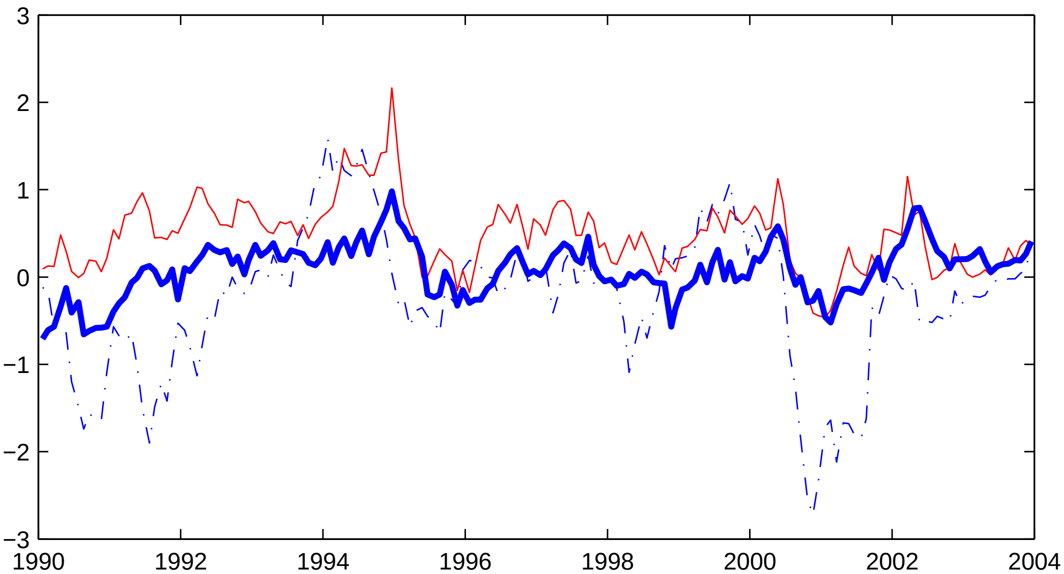


Figure 3: Expected and realized change in the 3-month yield over 6 months. Expectations-hypothesis based (thin solid line); survey based (thick solid line); realized (dot-dashed line). All are plotted at the time of the forecast.

of the average investor and the marginal investor differ.[26] However, we are not aware of a realistic model of differential information suitable for an empirical implementation, so in our context we need to rely on the "average" as a proxy for the "marginal."[27] As well, no matter how accurately surveys are conducted, they are prone to some measurement error, for example due to differences in the information sets and precise timing when various participants may prepare their responses. For these reasons, it is important to allow for the possible presence of substantial errors when using information from surveys and we do so in the term-structure estimation we propose. Nonetheless, we find that the survey forecasts we employ in our sample are surprisingly informative.

In our analysis, we use monthly data on the 6-month- and 12-month-ahead forecasts of the 3-month T-bill yield based on the *Blue Chip Financial Forecasts* (BCFF). We denote these two forecasts, for a survey conducted at t, $E_t^{svy}(y_{t+6m,3m})$ and $E_t^{svy}(y_{t+12m,3m})$, respectively. The BCFF is published on the first day of each month and presents forecasts of financial industry participants and professional forecasters, including banking institutions, conducted a few days earlier. Twice a year, the survey also collects information on long-range forecasts (average

rational despite irrational participants in the market."

[26] See, for example, Harrison and Kreps (1978) and Mayshar (1983).

[27] This issue is not uncommon in macroeconomics. Recall, for instance, Tobin's Q problem where the distinction between average Q and marginal Q may greatly influence econometric inference. Likewise, estimation of so-called micro-founded models of aggregate inflation dynamics invariably relies on average costs where marginal cost estimates would be required by the theory.

expected 3-month T-bill yield between (approximately) the next 6 and 11 years) and we use this information to supplement the monthly readings of expectations at shorter horizons. We denote the long-term forecast by $E_t^{svy}(\bar{y}_{3m,T_1,T_2})$, where $T_1 \approx 6$ and $T_2 \approx 11$. Additional details on the data are provided in Appendix B.

As an overview of the behavior of the survey forecasts, it is informative to compare the expected change in the short-term yield according to the surveys with those based on the expectations hypothesis (EH) and the random walk hypothesis (RWH). In the case of the 3-month yield $y_{t,3m}$, forecasts corresponding to these two hypotheses can be stated as:

$$E_t^{RWH}(y_{t+w,3m}) = y_{t,3m} \tag{13}$$

$$E_t^{EH}(y_{t+w,3m}) = f_{t,w,3m}, \tag{14}$$

where $f_{t,w,\tau}$ is the forward rate for contracting to borrow for τ period starting at w period from time t.

Figure 3 shows the expected change over the 6-month horizon in the 3-month yield based on the expectations hypothesis and the survey forecasts; the expected change based on the random walk hypothesis is trivial (simply zero). A notable characteristic of the data in this sample, 1990-2003, is that the realized change tends to be negative. In contrast, the EH-based expected change is mostly positive. The sample means of the realized change and the expected change according to the EH are -0.25% and 0.48%, respectively. This results in a larger forecasting error based on the expectations hypothesis (RMSE of 0.99%) compared to the random walk hypothesis (RMSE of 0.85%). By comparison, the root-mean-square error based on the survey forecast is 0.79%, somewhat smaller than that of random walk forecast.

The very poor forecasting performance of the EH suggests that, at minimum, some "term premium correction" is required to understand the data, and more accurate forecasts would be obtained once a term premium is introduced. The survey forecast has precisely these features: the expected interest rate change based on the survey forecast is *positively* correlated with that based on the EH (6-month-horizon expected changes have correlation of 0.64), and the sample mean of the expected change based on the survey is smaller than that based on the EH (0.08% vs. 0.48%). This suggests that the survey forecast incorporates a term premium correction. The figure also suggests that the term premium embedded in the survey forecasts is not constant but rather it varies over time.

A key question is how large the "proper" term premium correction should be. Note that the average expected change based on the survey is still larger than the average realized change, suggesting a bias. This can be also seen in the unbiasedness regression:[28]

$$y_{t+6m,3m} - y_{t,3m} = -0.35_{(0.12)} + 1.24_{(0.31)}(E_t^{svy}(y_{t+6m,3m}) - y_{t,3m}) + \xi_{t+6m}, \qquad (15)$$

where the Newey-West standard errors are given in subscripts. The coefficient on the constant term is indeed significantly negative. We argue, however, that this is a *desirable* feature: Recall that our concern about the 2-year forward term premium estimate based on the 1990-2003 data (Figure 2b) was that it seemed too large and seemed to imply spuriously small forecasting errors. At issue here is how much the market anticipated (in real time) the future changes (or lack of changes) in the interest rates; we have argued earlier that the extent of the decline in the short-rate over this period is unlikely to have been fully anticipated.

It is also interesting to note that the *efficiency* of the survey forecast is not rejected; survey forecast errors regressed on the term spread at the time of the forecast gives an insignificant coefficient. For example,

$$y_{t+6m,3m} - E_t^{svy}(y_{t+6m,3m}) = -0.39_{(0.23)} + 0.03_{(0.09)}(y_{t,10y} - y_{t,3m}) + \xi_{t+6m}. \qquad (16)$$

Coefficients on other information variables we have examined were also not significant.

These results are encouraging enough that we take the survey forecast of the 3-month T-bill yield as a proxy for the market expectation of the future 3-month T-bill yield, that is,

$$E_t^{svy}(y_{t+w,3m}) = E_t^{mkt}(y_{t+w,3m}) + e_{t,w}, \qquad (17)$$

in which $e_{t,w}$ is an error term that accounts for a possible discrepancy between the survey expectation and market expectation. Part of this error, like those stemming from the uncertainty about the exact time of the forecast, would have a white-noise-type characteristic. Other parts of this error can be expected to possess more structure; for example, if there is a divergence between the average investor expectation and the marginal investor expectation due to some fundamental economic mechanism, such errors are likely to be serially correlated. If the survey is drawn from the same (or slowly changing) small panel of participants every period, that would also create persistent measurement errors. In this paper we assume that the market expectation equals the model expectation, that is, $E_t^{mkt}(y_{t+w,3m}) = E_t(y_{t+w,3m})$, but the model may have some specification error, which may also introduce a serial correlation in $e_{t,w}$.

5 Estimation with survey data

It is straightforward to implement term structure estimation with survey data within a Kalman-filter-based maximum-likelihood estimation. The Kalman-filter approach is especially conve-

[28] In all regressions, the coefficients are reported in units of percentage points. (For example, -0.35 means 35 basis points.)

nient in this case because it can handle easily the feature that the yields data are available at a finer interval (weekly sampling) than the survey data (monthly and semiannual sampling).

Let the expanded observation vector be

$$o_t = [y_{t,\tau_1}, ..., y_{t,\tau_M}, E_t^{svy}(y_{t+6m,3m}), E_t^{svy}(y_{t+12m,3m}), E_t^{svy}(\bar{y}_{3m,T_1,T_2})]' \tag{18}$$

It is straightforward to show

$$o_t = \begin{bmatrix} a_{\tau_1} \\ \vdots \\ a_{\tau_M} \\ a_{3m} + b'_{3m}(I - e^{-\mathcal{K}/2})\mu \\ a_{3m} + b'_{3m}(I - e^{-\mathcal{K}})\mu \\ a_{3m} + b'_{3m}(I - W_{T_1,T_2})\mu \end{bmatrix} + \begin{bmatrix} b'_{\tau_1} \\ \vdots \\ b'_{\tau_M} \\ b'_{3m}e^{-\mathcal{K}/2} \\ b'_{3m}e^{-\mathcal{K}} \\ b'_{3m}W_{T_1,T_2} \end{bmatrix} x_t + \begin{bmatrix} \eta_{t,\tau_1} \\ \vdots \\ \eta_{t,\tau_M} \\ e_{t,6m} \\ e_{t,12m} \\ e_{t,\mathcal{L}} \end{bmatrix}, \tag{19}$$

in which

$$W_{T_1,T_2} = \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} e^{-\mathcal{K}s} ds = \frac{1}{T_2 - T_1}\mathcal{K}^{-1}(e^{-\mathcal{K}T_1} - e^{-\mathcal{K}T_2}), \tag{20}$$

and $\eta_{t,\tau_1}, .., \eta_{t,\tau_M}$ denote yield measurement errors and $e_{t,\mathcal{L}}$ denotes long-term (approximately 6-to-11-years) survey expectation measurement errors. In deriving eq. (19), we have used

$$E_t(x_{t+w}) = e^{-\mathcal{K}w}x_t + (I - e^{-\mathcal{K}w})\mu, \tag{21}$$

$$E_t\left(\frac{1}{T_2 - T_1}\int_{T_1}^{T_2} ds\ x_{t+s}\right) = W_{T_1,T_2}x_t + (I - W_{T_1,T_2})\mu. \tag{22}$$

Equation (19), together with equation (10) constitute the Kalman filter observation and state equations from which the likelihood function can be constructed for estimating our model with survey forecasts.

Note that the parameters of the model are still estimable without the last three rows in eq. (19); indeed, that case corresponds to the conventional estimation in our setting (eqs. (9) and (10)). The Kalman-filter setup (simply expanding the observation vector) is a natural and convenient way to incorporate the information contained in the survey data, without any change in the underlying term structure model.[29]

Survey forecasts could also be introduced in a term structure model in other ways. For example, one may treat the 6-month-ahead forecast of the 3-month yield, $E_t(y_{t+6m,3m})$, as a state variable, i.e., write the state vector as:

$$x_t = [E_t(y_{t+6m,3m}), x_{2t}, x_{3t}]'. \tag{23}$$

[29]That the survey data are not available at all dates can be handled by treating them as missing data, that is, we take the new observation vector as $\tilde{o}_t = S_t o_t$, where S_t is a matrix of varying dimension $d(t) \times (M + 2 + 1)$ that selects only the non-missing variables at time t. (See, for example, Harvey (1989).)

There is, however, nothing special about the 6-month horizon; one may also use the 12-month-horizon forecast or both, as in

$$x_t = [E_t(y_{t+6m,3m}), E_t(y_{t+12m,3m}), x_{3t}]'. \tag{24}$$

Besides the potential complication of survey measurement errors in these cases, eqs. (23) and (24) may face a "consistency" problem. For example, eq. (23) implies

$$E_t(y_{t+6m,3m}) = a_{3m} + b'_{3m}(e^{-\mathcal{K}/2}[E_t(y_{t+6m,3m}), x_{2t}, x_{3t}]' + (I - e^{-\mathcal{K}/2})\mu), \tag{25}$$

an inconvenient condition to impose. In the case of eq. (24), because the expectations $E_t(y_{t+6m,3m})$ and $E_t(y_{t+12m,3m})$ are correlated in a certain way, treating the two expectations as state variables with dynamics described by a triangular $\mathcal{K}$ matrix without restrictions may result in a misspecification or inefficiency.[30]

To implement the Kalman filter, we need to specify the measurement error structure for the observation equation. A standard choice would be the serially- and cross-sectionally uncorrelated normal errors:

$$\eta_t = [\eta_{t,\tau_1}, ..., \eta_{t,\tau_M}, e_{t,6m}, e_{t,12m}, e_{t,\mathcal{L}}]' \sim N(0, \Omega^\eta), \tag{26}$$

$$\Omega_\eta = \text{diag}([\omega^2_{\tau_1}, ..., \omega^2_{\tau_M}, \tilde{\omega}^2_{6m}, \tilde{\omega}^2_{12m}, \tilde{\omega}^2_{\mathcal{L}}]). \tag{27}$$

Note that in the limit $\tilde{\omega}_{6m}, \tilde{\omega}_{12m}, \tilde{\omega}_{\mathcal{L}} \to \infty$, the estimation collapses to that of the standard estimation (i.e., estimation without survey data). In estimating the model, objects such as $\tilde{\omega}_{6m}, \tilde{\omega}_{12m}, \tilde{\omega}_{\mathcal{L}}$ can be either fixed according to some *a priori* judgments or estimated as free parameters. Nesting the traditional estimation in our model in this manner and allowing treating the measurement variances as free parameters is particularly useful in this setting as it it allows the model to ignore the survey forecasts if these prove inconsistent with the dynamics in the actual term structure data (and are thus uninformative for the model) or, alternatively, rely heavily on them if they closely match the expectations of interest rates implied by the model. The two extreme cases simply correspond to infinite or zero estimates of the variance of the measurement errors. This setting also allows fixing the variance of the measurement errors at values higher than suggested by estimating them as free parameters, if we wish to be more conservative in incorporating the information from the surveys in the estimation than suggested by an unconstrained fitting of the model.

[30] We note incidentally that survey forecasts of other variables may also be used in estimating term structure models. For example, Pennacchi (1991) employs survey forecasts of inflation to study the dynamics of real interest rates. Chun (2005) employs 3-month-ahead survey forecasts of macroeconomic variables as the state variables in a dynamic term structure model. (However, his setup does not utilize the forecasts at other available horizons, such as 6-month-ahead forecasts, and thus does not enforce the dynamic consistency of expectations at different horizons.)

The estimation may also account for potential serial correlation and cross-sectional correlation in the measurement error of the short-horizon survey forecasts by modeling $e_{t,6m}$ and $e_{t,12m}$ as mean-zero AR(1) processes with correlated innovations:

$$
\begin{aligned}
e_{t,6m} &= \phi_{6m} e_{t-h,6m} + \varepsilon_{t,6m} \qquad (28)\\
e_{t,12m} &= \phi_{12m} e_{t-h,12m} + \varepsilon_{t,12m} \qquad (29)
\end{aligned}
$$

in which

$$[\varepsilon_{t,6m}, \varepsilon_{t,12m}]' \sim N(0, \tilde{\Omega}^{\varepsilon}), \qquad (30)$$

where $\tilde{\Omega}^{\varepsilon}$ can be expressed in terms of the unconditional covariance matrix $\tilde{\Omega}^{e}$ $(= \mathrm{cov}([e_{t,6m}, e_{t,12m}]')$ as

$$\tilde{\Omega}^{\varepsilon} = \tilde{\Omega}^{e} - \mathrm{diag}([\phi_{6m}, \phi_{12m}]')\tilde{\Omega}^{e}\mathrm{diag}([\phi_{6m}, \phi_{12m}]'). \qquad (31)$$

The matrix $\tilde{\Omega}^{e}$ can be parameterized in terms of the Cholesky decomposition,

$$\tilde{\Omega}^{e} = \begin{bmatrix} \tilde{\omega}_{6m} & 0 \\ \tilde{\omega}_{c} & \tilde{\omega}_{12m} \end{bmatrix} \begin{bmatrix} \tilde{\omega}_{6m} & 0 \\ \tilde{\omega}_{c} & \tilde{\omega}_{12m} \end{bmatrix}',$$

which can accommodate cross-sectional correlation of $e_{t,6m}$ and $e_{t,12m}$ with a non-zero value of $\tilde{\omega}_{c}$.

This specification of the measurement error structure can be incorporated into the Kalman filter by augmenting the state space. Let $z_t = [x_t', e_{t,6m}, e_{t,12m}]'$. We can now write the measurement equation and the state equation as

$$o_t = \begin{bmatrix} a_{\tau_1} \\ \vdots \\ a_{\tau_M} \\ a_{3m} + b_{3m}'(I - e^{-\mathcal{K}/2}) \\ a_{3m} + b_{3m}'(I - e^{-\mathcal{K}}) \\ a_{3m} + b_{3m}'(I - W_{T_1,T_2}) \end{bmatrix} + \begin{bmatrix} b_{\tau_1}' & 0 & 0 \\ \vdots & 0 & 0 \\ b_{\tau_M}' & 0 & 0 \\ b_{3m}' e^{-\mathcal{K}/2} & 1 & 0 \\ b_{3m}' e^{-\mathcal{K}} & 0 & 1 \\ b_{3m}' W_{T_1,T_2} & 0 & 0 \end{bmatrix} z_t + \begin{bmatrix} \eta_{t,\tau_1} \\ \vdots \\ \eta_{t,\tau_M} \\ 0 \\ 0 \\ e_{t,\mathcal{L}} \end{bmatrix}, \qquad (32)$$

and

$$z_t = \begin{bmatrix} (I - e^{-\mathcal{K}h})\mu \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} e^{-\mathcal{K}h} & 0 & 0 \\ 0 & \phi_{6m} & 0 \\ 0 & 0 & \phi_{12m} \end{bmatrix} z_{t-h} + \begin{bmatrix} \epsilon_t \\ \varepsilon_{t,6m} \\ \varepsilon_{t,12m} \end{bmatrix}, \qquad (33)$$

which is again in a form that can be easily implemented.

To reduce the number of estimated parameters, we posit that the measurement errors of the yields we use are serially and cross-sectionally uncorrelated that is, $\eta_t = [\eta_{t,\tau_1}, \ldots, \eta_{t,\tau_M}] \sim N(0, \mathrm{diag}([\omega_{\tau_1}^2, \ldots, \omega_{\tau_M}^2]))$, and also assume $E_{t-h}(\eta_t[e_{t,6m}, e_{t,12m}, e_{t,\mathcal{L}}]) = 0$ and $E_{t-h}([e_{t,6m}, e_{t,12m}]e_{t,\mathcal{L}})$ $= 0$.

6 Basic results from estimation with survey data

Our baseline parameter estimates from the estimation with 1990-2003 term structure data and survey forecast data are presented in Table 1. (Henceforth we refer to this case as the S1990 estimates.) The parameter estimate from the estimation with the same term structure data but without survey data (NS1990) is also provided for comparison. To obtain the estimates shown, we followed the common practice of first performing a preliminary unrestricted estimation and then reestimating the model setting some of the $\mathcal{K}, \Lambda_b$ parameters with large standard errors in the preliminary estimation to zero. Four parameters are set to zero in the estimation without the survey data while only one in the estimation with the survey data. As noted earlier, in the estimation with the survey data several options are presented with respect to the treatment of the measurement error of the survey forecasts. In the baseline case presented in the table, we allow for serial correlation in the measurement error associated with the short-term survey forecasts and free estimation of the variances of these errors (eq. (32) and (33)). To downplay the role of the long-horizon survey forecast, however, we fixed the value of the corresponding measurement error, $\tilde{\omega}_{\mathcal{L}}$, to 75 basis points, more than twice as large as what free estimation of that parameter would suggest.

	S1990	NS1990		S1990	NS1990
$\mathcal{K}_{11}$	0.0539(0.0901)	0.6360(0.3646)	$[\mathcal{K}\Lambda_b]_{31}$	-1.2934(0.6479)	0
$\mathcal{K}_{21}$	-0.1486(0.0781)	5.8219(1.6881)	$[\mathcal{K}\Lambda_b]_{12}$	-0.8328(0.2877)	-0.0459(0.0419)
$\mathcal{K}_{31}$	0	4.5024(1.1315)	$[\mathcal{K}\Lambda_b]_{22}$	-0.9081(0.0956)	0
$\mathcal{K}_{22}$	0.4841(0.0846)	0.5822(0.2021)	$[\mathcal{K}\Lambda_b]_{32}$	-1.0071(0.9260)	0.3100(0.1566)
$\mathcal{K}_{32}$	-6.1308(1.7814)	0	$[\mathcal{K}\Lambda_b]_{13}$	0.2974(0.0370)	0.0912(0.0345)
$\mathcal{K}_{33}$	2.1521(0.3175)	0.4985(0.1776)	$[\mathcal{K}\Lambda_b]_{23}$	0.2889(0.1157)	-0.1052(0.1208)
Σ_{11}	0.0023(0.0000)	0.0022(0.0005)	$[\mathcal{K}\Lambda_b]_{33}$	0.4122(0.2794)	-0.4572(0.1849)
Σ_{22}	0.0021(0.0004)	0.0049(0.0004)	ω_{3m}	0.0010(0.0001)	0.0010(0.0001)
Σ_{33}	0.0062(0.0002)	0.0045(0.0002)	ω_{6m}	0.0003(0.0000)	0.0002(0.0000)
ρ_o	0.0486(0.0112)	0.0441(0.0143)	ω_{1y}	0.0005(0.0000)	0.0005(0.0000)
λ_{a1}	-0.7828(0.3214)	-0.5065(0.4613)	ω_{2y}	0.0000(0.0001)	0.0000(0.0001)
λ_{a2}	-0.6960(0.9861)	-1.9103(0.3590)	ω_{4y}	0.0003(0.0000)	0.0003(0.0000)
λ_{a3}	-1.8420(1.3677)	-1.6761(0.3509)	ω_{7y}	0.0000(0.0001)	0.0000(0.0001)
$[\mathcal{K}\Lambda_b]_{11}$	-0.1615(0.1467)	0.7744(0.3055)	ω_{10y}	0.0005(0.0000)	0.0005(0.0000)
$[\mathcal{K}\Lambda_b]_{21}$	-0.2564(0.1075)	0	$\tilde{\omega}_{6m}$	0.0023(0.0003)	
			$\tilde{\omega}_{12m}$	0.0017(0.0002)	
			$\tilde{\omega}_{c}$	0.0020(0.0002)	
			ϕ_{6m}	0.9577(0.0108)	
			ϕ_{12m}	0.9730(0.0078)	

Table 1: Parameter estimates with and without survey data. (S1990 and NS1990).

Incorporating the information from the survey forecasts in the estimation results in notable differences. One such difference is that the smallest eigenvalue of the $\mathcal{K}$ matrix in the estimation with survey data is indeed small (0.054), much smaller than its counterpart (0.490) from the estimation without survey data.[31] This proves important for the long-horizons forecasts implied by the model which we detail later on.[32] Note also that the estimation with survey data tends to result in smaller standard errors. (Recall that more of the NS1990 parameters have been set to zero). This improved precision has implications for confidence intervals associated with implied quantities of economic interest such as interest rate expectations and term premia.[33] Use of the survey data also makes a significant difference in the behavior of the likelihood function and alleviates estimation problems encountered with the conventional estimation of the model. Importantly, the presence of multiple inequivalent local maxima is much less of an issue. In sum, using the survey data helps overcome the imprecision and lack of robustness issues associated with the estimation without them in a small sample.

To confirm the robustness of the estimation we also examined several alternative implementations of the estimation with survey forecasts. For example, estimation without imposing the sole parameter restriction $\mathcal{K}_{31} = 0$ yields similar results. In addition, we examined estimation with survey data under the assumption of cross-sectionally and serially uncorrelated (i.i.d.) measurement error. (That is, $\phi_{6m} = \phi_{12m} = \tilde{\omega}_c = 0$.) Note that in this case we can simply use eqs. (19) and (10) as the observation equation and state equation, instead of eqs. (32) and (33). We obtained largely similar results as the baseline S1990 estimate. This fact is of interest, as the i.i.d. version is somewhat simpler to implement in practice. Also, estimation with an even shorter sample, 1995-2003, produced similar results. Furthermore, we have tried imposing the condition that ρ_o be determined by the condition that the unconditional mean $E(r_t) = \rho_o + \rho'\mu$ equals a predetermined value, e.g., $E(r_t) = 0.045$. We find that this helps stabilize the model's implications about the long-horizon expectations, and can serve as a substitute for using the long-horizon survey data in the estimation, to some extent. These findings speak well for the robustness of the survey-based estimation results presented in this paper.

The model estimated with survey data (S1990) fits the survey forecast of the 3-month yield reasonably well, as can be seen in Figure 4a. The regression of the survey-based 6-month-horizon expected change in the 3-month yield, $E_t^{svy}(y_{t+6m,3m}) - y_{t,3m}$, on the model-implied

[31] In a single factor model, a value of 0.054 would correspond to a half-life of 13 years.

[32] We note, however, that having a persistent factor is not a *sufficient* condition. For example, suppose that an independent-factors model ($\mathcal{K}, \Lambda_b, \Sigma$ diagonal in the notation detailed below) has a factor x_{1t} that is very persistent (small $\mathcal{K}_{11}$). If the diffusion of the factor (Σ_{11}) is very small, that factor wouldn't have much "weight" in the model, and the model might not produce a substantially variable long-horizon expectation.

[33] For example, in the case of the confidence interval for the 2-year forward term premium, the part of the confidence interval due to sampling uncertainty (as opposed to the forecast uncertainty) for S1990 is about half as small as that for NS1990.

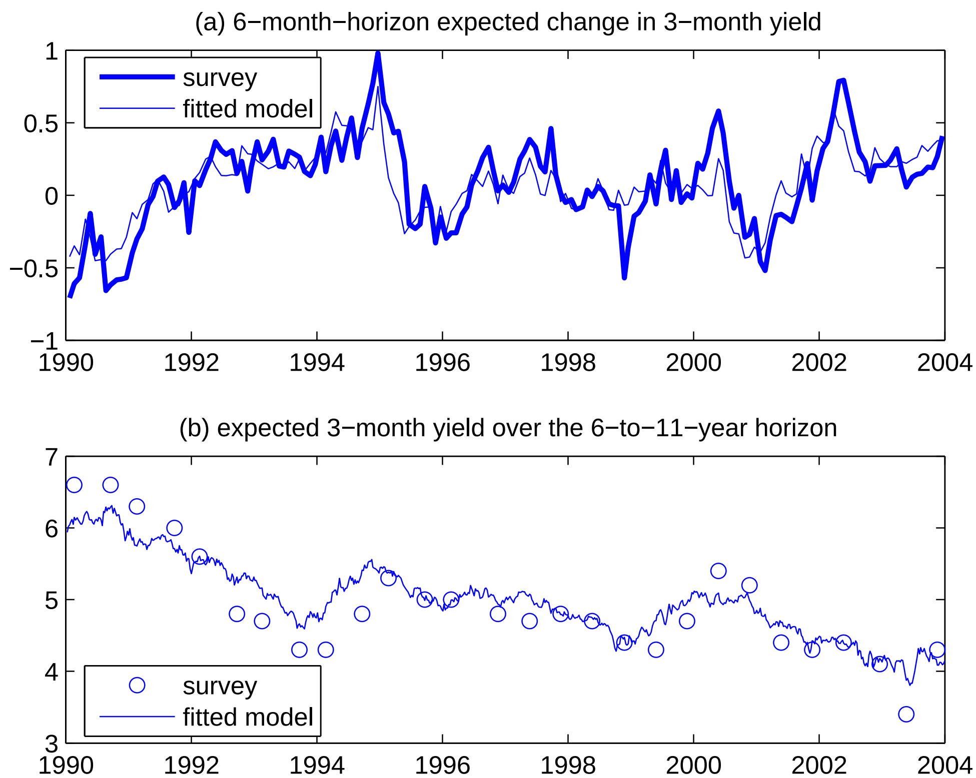


Figure 4: (a) Survey forecast of the 6-month-horizon expected change in the 3-month yield and the forecast implied by the model. (b) Survey forecast of the expected 3-month yield in the next 6 to 11 years and the model forecast.

(fitted) 6-month-horizon expected change in the 3-month yield, $\hat{E}_t(y_{t+6m,3m}) - y_{t,3m}$, gives:

$$E_t^{svy}(y_{t+6m,3m}) - y_{t,3m} = 0.00_{(0.01)} + 1.07_{(0.06)}(\hat{E}_t(y_{t+6m,3m}) - y_{t,3m}) + \epsilon_t, \quad R^2 = 0.67.$$

The sample standard deviation of $\hat{E}_t(y_{t+6m,3m}) - y_{t,3m}$ is 24 basis points, while the sample standard deviation of $E_t^{svy}(y_{t+6m,3m}) - \hat{E}_t(y_{t+6m,3m})$ is 18 basis points, thus the "signal-to-noise ratio" is larger than 1, but the "noise" amount is substantial.[34]

Figure 4b plots the model-implied expectation of the 3-month yield over the 6-to-11 year horizon and the corresponding semiannual long-horizon survey forecast. Interestingly, even though we imposed a fairly large measurement error for $e_{t,\mathcal{L}}$, the model-implied long-term expectation tracks the survey-forecasts remarkably well, though a persistent deviation is visible in the 1990-94 period.

Figure 5a shows the expected short-rate path for Dec. 10, 2003, based on the S1990 estimation. The result from the NS1990 estimation is also reproduced for comparison. As can be seen, estimation with the survey data gives a more reasonable (faster rising) expected path of

[34] Part of this "noise" may be due to misspecification.

the short rate for that date. The two-year forward term premium $\varphi_{t,2y}$ based on the S1990 estimation, shown in Figure 5b, is on average smaller than the NS1990 result, and becomes near-zero (i.e., gets close to the expectations hypothesis) quite often. Finally, the ten-year-ahead expected short rate, $E_t(r_{t+10y})$, based on the S1990 estimation (thin solid line in Figure 5c) exhibits considerably more variation than its NS1990 counterpart, and, in agreement with our priors, reflects the downward trend one might have expected in light of the evolution of the long-term inflation expectations presented in Figure 1.

These results indicate that the use of survey forecasts not only helps stabilize the estimation of our dynamic term structure model but also produces "sensible" results. To be sure, it may not be surprising that the short-term interest-rate expectations suggested by the model are more or less in line with the survey forecasts, since survey forecasts of the 3-month yield were an input in the estimation of the model. Still, these results are nontrivial for several reasons. Our estimation procedure does not guarantee that the model can match the survey data well; a misspecified model or a model that is too restrictive model might not do so. Further, if the survey forecasts were in fact inconsistent with the expectations implied by the yield curve data, the estimated model would not agree with the survey data and the estimated parameters would instead interpret the surveys as noise.

Perhaps the most intriguing reason for appreciating the estimates of the model with the survey data, however, is suggested by comparing Figures 2 and 5. This comparison presents evidence of considerable qualitative agreement between the estimated model with survey data over the 1990-2003 sample (S1990) and that without survey data but with a longer sample (NS1965), which raises the credence of both results. In particular, note that the two estimates suggest remarkably similar two-year forward term premia (in the middle panels of the two figures) and both show a significant downward drift in the 10-year-ahead expected short-term rate (bottom panels). That said, the two estimates are not the same. The S1990 estimate exhibits somewhat greater high-frequency variation in long-horizon expectations and picks up some interesting features that are missing from the NS1965 estimate. One such feature is the the brief dip of the 10-year-ahead short-rate expectation in 2003 to a level below 4 percent. At the time, the Federal Reserve expressed concern of an "unwelcome" fall in inflation, and indicated that monetary policy would maintain an easy stance for a considerable period—as long as a substantial fall in inflation remained a risk. (See Bernanke (2003).)

The similarity of the NS1965 and S1990 estimates is also notable for additional reasons. First, it suggests that from the perspective of our model, the dynamic behavior of the term structure has not changed as much as is suggested when the model is estimated over small samples without survey data. In turn, this suggests that the model is sufficiently rich and flexible

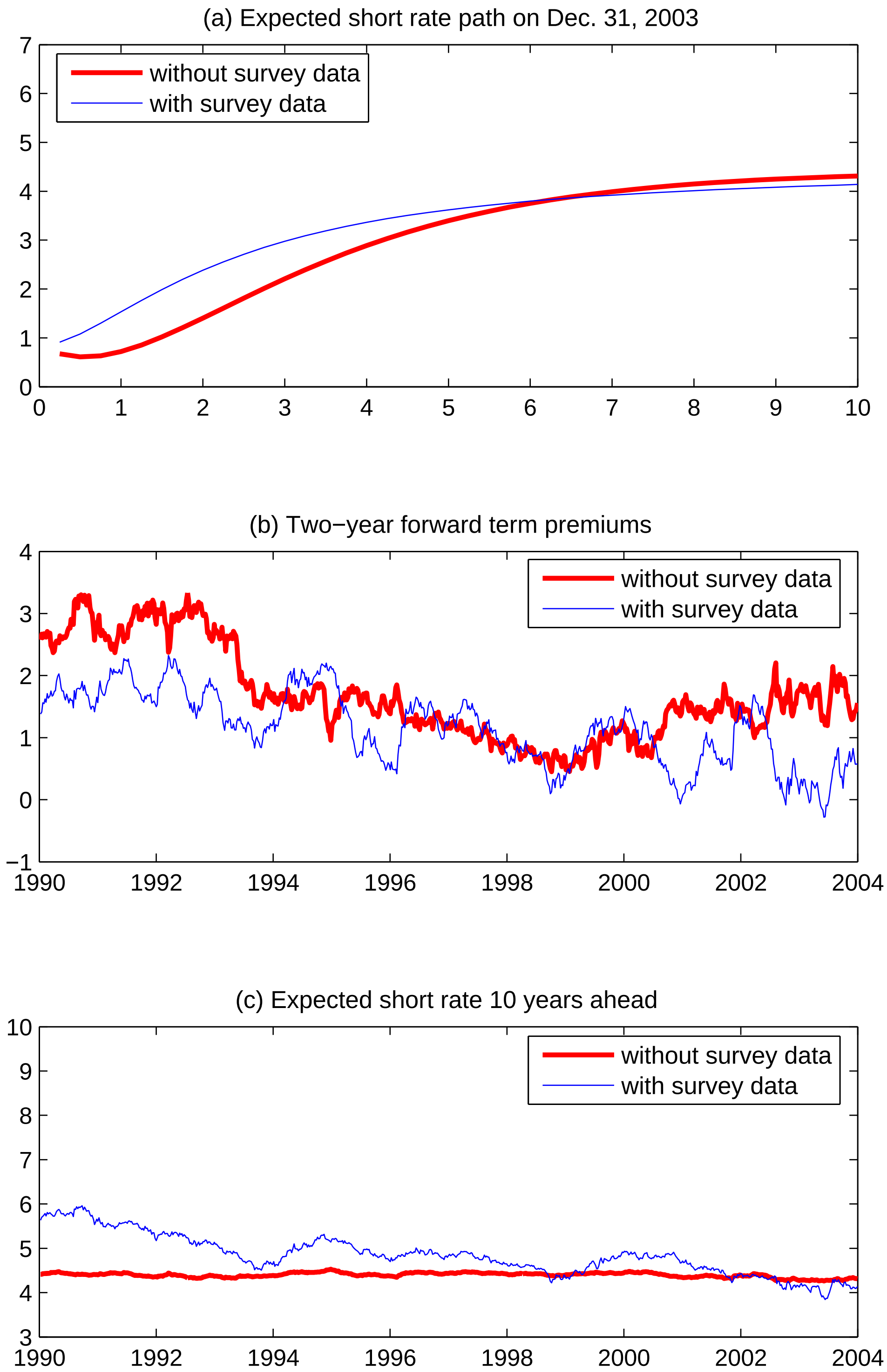


Figure 5: (a) The expected paths of the short rate on Dec. 30, 2003, based on estimations with and without survey data (S1990 and NS1990). (Horizon in years shown on axis.) (b) Two-year forward term premium. (c) Expected 10-year-ahead short-rate expectations.

	RMSE	α_e	β_e	α_u	β_u
NS1990	0.65	-0.28(0.16)	0.08(0.07)	-0.11(0.10)	1.05(0.22)
NS1965	0.77	-0.49(0.22)	0.09(0.09)	-0.46(0.11)	2.95(0.54)
S1990	0.76	-0.36(0.22)	0.02(0.08)	-0.39(0.11)	2.02(0.43)

Table 2: Root-mean-square errors, coefficients from the "efficiency regression," and coefficients from the "unbiasedness regression," with the model-implied 6-month-horizon forecasting errors (1990-2003).

to encompass much of what would appear to reflect "structural change" and expected to result in substantially different estimates over alternative samples in simpler, more restrictive models.[35] Furthermore, it confirms that the 1990-2003 is "too short" to produce reliable estimates with the conventional estimation that does not incorporate the survey data. If the small-sample problem were not severe, the NS1990 characterization of the term structure would not have differed so much from that implied by the NS1965 estimation. Finally, this result strengthens the case that the survey forecasts of interest rates we employ can serve as quite useful proxies of the rational expectations implicit in our model.

We have examined some criteria which Duffee (2002) suggested are useful for evaluating term-structure models. Table 2 lists the root-mean-square values, the coefficients α_e, β_e from the efficiency regression, and the coefficients α_u, β_u from the unbiasedness regression, based on the 6-month-horizon *in-sample* forecasts (in the 1990-2003 period) from the S1990, NS1990, and NS1965 estimations. Not surprisingly, the RMSE value is the smallest in the NS1990 case, but this may be due to a substantial look-ahead bias in the short sample, rather than a genuine forecastability; indeed, in the Monte Carlo example discussed later (Section 7), the RMSE based on in-sample estimated forecast errors tends to be smaller than the RMSE of the "true" forecast errors. The unbiasedness ($\alpha_u = 0$, $\beta_u = 1$) is not rejected in the NS1990, but this might not be a desirable feature, recalling our discussion about disinflation in Section 2. The β_e coefficient in the efficiency regression is not significant in all three cases, hence does not distinguish the models.

We have also looked at the out-of-sample forecasts for the brief period since the end of our estimation sample. The RMSE of forecast errors for the 6-month-ahead forecasts of the short rate made in the period from January 2004 to February 2005 (realized between July 2004 and August 2005) were 0.69%, 0.51%, and 0.46% for the NS1990, NS1965, and S1990 models, respectively. The conventional short-sample estimation performed worst. However, our out-of-sample window is quite short so care is required to avoid reading too much in this. In fact,

[35] An illustration of this point is that the "central tendency model" (discussed in footnote 10), which is nested in our model, can be viewed as a "structural change" version of the univariate model in which the short rate mean-reverts to a constant.

even the somewhat longer out-of-sample test periods used in Duffee (2002) and Duarte (2004) are not likely to be decisive. Since much of the variation in interest rates is *unanticipated*, in a relatively short out-of-sample period—say 5 years—the "correct model" may underperform the martingale benchmark quite often simply due to randomness (sampling variability). With a Monte Carlo simulation, we have examined the distribution of the ratio of the RMSE based on the true forecast error to the RMSE based on the random walk forecasting error. For 6-month-ahead forecasts of the 3-month yield in a 4.5-year period based on the S1990 parameter estimate, we find that there is a substantial chance (about 20 percent) of the ratio being larger than 1 (i.e., of the true RSME being larger than the random walk RMSE).

These results point to a relative lack of power in asymptotic statistics and conventional model evaluation criteria for distinguishing among alternative models in a setting like this. Unfortunately, the difference between the NS1990 and S1990 estimates is of economic significance. For example, in 2002, the two-year forward term premium estimates based on S1990 and NS1990 differ by more than 1%. To the extent that reliance on the incorrect model might have resulted in such a substantial misreading of market expectation of the path of interest rates it could also have led to a significant policy error.

We conclude this section with remarks on a few practical implications of the results. First, the results suggest some kind of "middle ground" in interpreting movements of long-horizon forward rates. As Figures 4b and 5c show, short-rate expectations can have a significant variation even at long horizons, but variation in term premia is also substantial. Thus, interpreting long-horizon forward rate movements entirely as movements in long-horizon expectations (e.g., in Gurkaynak, Sack, and Swanson (2005)) may be an extreme view, but interpreting the long-horizon forward rate movements purely in terms of the movements in term premium may be also extreme.[36] Second, the finding that the survey forecasts are generally informative makes it especially significant that they line up better with the forward rates in recent years than in the past. (See the post-2001 period in Figure 3.) This means that reading the futures curve or forward rate curve as the expected short-rate path (without a term premium correction) *for relatively short horizons* may be less problematic in this recent period than has been in the past. This point may help explain the continued reference to the expectations hypothesis in the market place despite its poor historical performance.[37]

[36]Best, Byrne, and Ilmanen (1998) draw similar conclusions. Kim and Wright (2005) provide a detailed discussion of the model's implications as regards the forward rate movements in the recent period.

[37]Market commentators and traders often refer to the eurodollar futures curve (with a simple adjustment for the difference between the federal funds rate and the 3-month LIBOR rate) as the "expected policy path." See Blinder (1997, p. 16) for an interesting remark on this phenomenon. Gurkaynak, Sack, and Swanson (2002) also emphasize the informativeness of the federal funds futures and eurodollar futures curves in the post-1994 period.

7 Expectations of longer-term interest rates

In estimating the model we have used survey forecasts of only the short-term interest rates and not those of long-term yields. Hence, examination of what the model has to say about expectations of long-term yields can serve as a useful check of the model and the survey. In particular, we examine whether the model is consistent with the stylized facts regarding the failure of the expectations hypothesis reflected in Campbell-Shiller regressions, and whether expectations of long-term yields implied by our model compare with survey-forecasts of comparable-maturity yields.

7.1 Campbell-Shiller regressions

Recall that in the Campbell-Shiller regression

$$y_{t+w,\tau-w} - y_{t,\tau} = \alpha + \beta \frac{w}{\tau - w}(y_{t,\tau} - y_{t,w}) + \xi_{t+w}, \tag{34}$$

the expectations hypothesis corresponds to $\beta = 1$. Campbell and Shiller (1991) found that empirically $\beta = 1$ is rejected, often being negative and increasingly so with larger τ. Rudebusch and Wu (2004) have noted that in the more recent sample, the evidence against the expectations hypothesis is weaker. We therefore re-examine the Campbell-Shiller regression for our sample and look at how well our model matches the resulting pattern of β.

The upper panel of Table 3 presents estimates of the Campbell-Shiller regression for several τ with $w = 3$ months, using the T-bill yield ($y^{Tbill}_{t,3m}$), and with $w = 6$ months, using both the T-bill yield ($y^{Tbill}_{t,6m}$) and coupon-securities-based yield ($y^{coupon}_{t,6m}$).[38] The recent sample still displays the classic pattern of the departure from the expectations hypothesis, in that β is negative and more so for longer maturities, but the point estimates are less negative than Campbell and Shiller (1991),[39] and the standard errors are large so that $\beta = 1$ (expectations hypothesis) is not rejected at the 5% significance level.

To investigate how well the estimated model matches this pattern of slope estimates of the Campbell-Shiller regression, we simulated 500 samples of length of 750 weeks (which approximately corresponds to the length of the true sample) and run regression (34) with the simulated data. The bottom panel of Table 3 shows the median and standard deviation of the β's. The β estimates do deviate from 1 and increasingly so with larger τ; the median β's imply somewhat greater departure from the EH than the true sample point estimate. However, the dispersions

[38] Rudebusch and Wu (2004) have focused on the case $w = 1$ month, but this requires data on the 1-month yield, which is not measured very well; Duffee (1996), for example, has noted that the 1-month T-bill yield has a substantial amount of idiosyncratic variation and cautioned against the use of the 1-month T-bill yield in tests of the expectations hypothesis.

[39] For example, for $w = 3m, \tau = 10y$, Campbell and Shiller obtain a β estimate of -2.52 with their (older and longer) sample, while the corresponding number in the present paper is -0.74

			$\tau = 2y$	$\tau = 3y$	$\tau = 5y$	$\tau = 10y$
SAMPLE	$w = 3m$	T-bill	-0.06 (0.95)	-0.57 (1.09)	-0.86 (1.22)	-0.74 (1.32)
	$w = 6m$	T-bill	-0.17 (0.99)	-0.67 (1.06)	-0.96 (1.10)	-0.96 (1.25)
	$w = 6m$	coupon	-0.43 (0.94)	-0.82 (1.02)	-1.05 (1.06)	-1.01 (1.22)
MODEL	$w = 3m$		-1.00 (0.78)	-1.21 (1.08)	-1.53 (1.60)	-2.38 (2.75)
	$w = 6m$		-0.83 (0.78)	-0.98 (1.06)	-1.23 (1.54)	-2.03 (2.63)

Table 3: β coefficients from the expectations hypothesis regression. Top panel ("SAMPLE") shows the results from the 1990-2003 sample, with Newey-West standard errors in parenthesis. Bottom panel shows the implied coefficients from the S1990 estimate.

(the standard errors) for $\tau = 5y$ and $\tau = 10y$ are large enough that the EH is not rejected in the simulated data either, suggesting that the EH regression tests are not powerful enough to distinguish between the EH and interesting alternative hypotheses in a sample of the length considered here.

7.2 Survey forecasts of longer-term interest rates

In addition to the survey forecasts of the short-term interest rate (3-month T-bill) that we use in estimating our model, *Blue Chip Financial Forecasts* provides survey forecasts of longer-term interest rates. In particular, monthly forecasts of Treasury constant maturity (TCM) yields are available every month in our sample for maturities such as 5 years and 10 years. It is of interest to compare these survey forecasts of long-term yields with those based on the expectations hypothesis and those implied by our model.

Since our model points to systematic deviations from the EH, for the survey forecasts to be consistent with our model they should also exhibit similar deviations. In this regard, comparison with results reported by Froot (1989) is also of interest. In particular, Froot noted that survey forecasts of long-term yields appeared to be in line with the expectations hypothesis. More specifically, in the regression

$$E_t^{svy}(y_{t+w,\tau}) - y_{t,w} = \alpha + \beta\,(f_{t,w,\tau} - y_{t,w}) + \epsilon_{t+w} \tag{35}$$

(eq. (14) in Froot (1989)), he reported estimates of β that are close to 1 and very high R^2 (often close to 1) when τ is large and w is small.

We have also looked at comparable specifications of the regression (35) with the survey forecasts of the TCM yields in our sample and obtained similar results when τ is large and w is small. However, eq. (35) has relatively little power for tests relating to the expectations hypothesis. Even if the survey expectation were inconsistent with the expectations hypothesis, eq. (35) with large τ and small w would still give β close to 1 and high R^2 due to the persistence of term spreads.[40]

A more informative test of whether the survey forecasts are consistent with the EH is provided by the regression of *expected changes* in yields (i.e., subtracting $y_{t,\tau}$ instead of $y_{t,w}$ in eq. (35)). In performing this regression, we also need to account for the fact that the TCM yields correspond to the yields on on-the-run par Treasury securities. In order to obtain the survey forecast that can be compared with the model (which are estimated with the off-the-run yield data), we convert the TCM yield forecast to the off-the-run par yield forecast. Note that we can define the "on-the-run premium" as

$$\mathcal{O}_{t,\tau} = Y_{t,\tau} - Y_{t,\tau}^{TCM}, \tag{36}$$

where $Y_{t,\tau}$ is the par yield based on the off-the-run Treasury securities and $Y_{t,\tau}^{TCM}$ is the TCM yield. The w-horizon expected change in par yield $\delta_w^{svy} Y_{t,\tau} = E_t^{svy}(Y_{t+w,\tau}) - Y_{t,\tau}$ can be thus written

$$\delta_w^{svy} Y_{t,\tau} = E_t^{svy}(Y_{t+w,\tau}^{TCM}) - Y_{t,\tau}^{TCM} + \delta_w \mathcal{O}_{t,\tau}, \tag{37}$$

where $\delta_w \mathcal{O}_{t,\tau}$ is the expected w-horizon change in $\mathcal{O}_{t,\tau}$.[41] Substituting the survey forecast of the TCM yield in eq. (37), we obtain the "survey-based" forecast of the changes in the (off-the-run) par yield, $\delta_w^{svy} Y_{t,\tau}$. The corresponding EH-based expected change in the (off-the-run) par yield is

$$\delta_w^{EH} Y_{t,\tau} = F_{t,w,\tau} - Y_{t,\tau}, \tag{38}$$

where $F_{t,w,\tau}$ is the par forward rate for borrowing for τ period at time $t+w$ (See Appendix A).

The results for the 5-year par yield forecasts regression

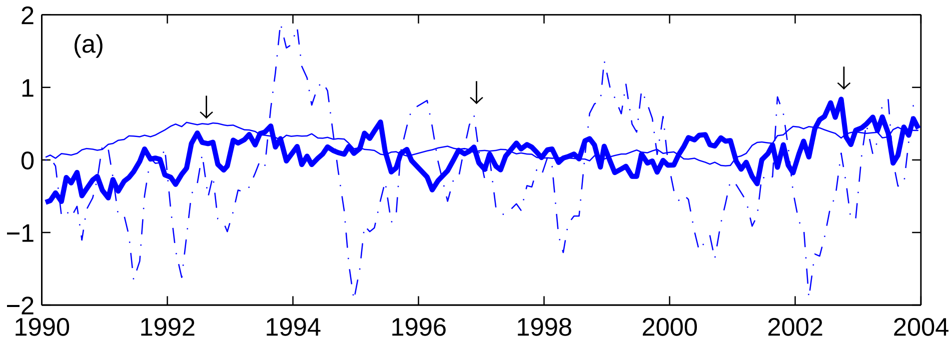
$$\delta_w^{svy} Y_{t,5y} = \alpha + \beta\, \delta_w^{EH} Y_{t,5y} + \xi_t \tag{39}$$

for horizons of 6 months and 12 months ($w = 6m, 12m$) are shown in Table 4. The low value of R^2 in the case of the 6-month horizon ($R^2 = 0.05$) suggests that the expectations hypothesis explains little of the short-horizon survey forecast of the changes in long-term interest rates. For forecasts at a longer horizon (12 months), the results still point to deviations from the EH but are closer to it, with larger values of the β coefficient and R^2.

It can be seen from Figure 6a that the survey-based 6-month-horizon expected change in the 5-year par yield ($\delta_{6m}^{svy} Y_{t,5y}$) has substantial short-run movements, while the EH-based expected

[40]To elaborate, consider the short-horizon forecast of $y_{t,\tau}$. The forecast $E_t(y_{t+w,\tau})$ would differ from $y_{t,\tau}$ only by a small amount, i.e., $E_t(y_{t+w,\tau}) = y_{t,\tau} + \delta_t$, where δ_t is a small number that goes to zero as w goes to zero. Because the regression of $y_{t,\tau} - y_{t,w}$ on $f_{t,\tau,w} - y_{t,w}$ gives the values of β and R^2 close to 1 for large τ, so would the regressions with $y_{t,\tau} + \delta_t - y_{t,w}$ (with the correct forecast) and $y_{t,\tau} - \delta_t - y_{t,w}$ (with a wrong forecast) in place of $y_{t,\tau} - y_{t,w}$.

[41]We use rolling AR(1) estimation to calculate the real-time estimation of $\delta_w \mathcal{O}_{t,\tau}$.



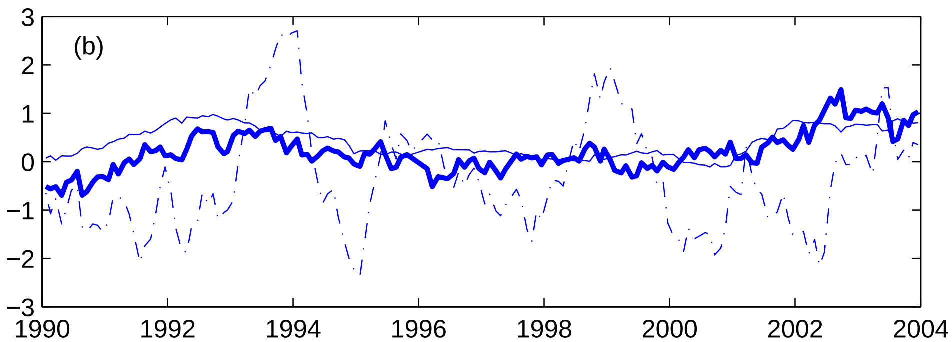


Figure 6: Comparison of the survey-based expected change in the 5-year par yield (thick solid line) and the expectations-hypothesis-based expected change (thin solid line) for the 6-month horizon (a) and the 12-month horizon (b).

change $\delta^{EH}_{6m}Y_{t,5y}$ is quite smooth. As the horizon grows, the agreement between the survey and the EH gets better, as the *business cycle* variation aspects play a greater role; note that the envelop of $\delta^{svy}_{12m}Y_{t,5y}$ tracks $\delta^{EH}_{12m}Y_{t,5y}$ fairly well.

The inability of the expectations hypothesis to explain the short-run variations in the survey forecast can be highlighted by the regression of the 3-month difference in $\delta^{svy}_{w}Y_{t,5y}$ ($\Delta^{svy}_{3m,w}Y_{t,5y} \equiv \delta^{svy}_{w}Y_{t,5y} - \delta^{svy}_{w}Y_{t-3m,5y}$) onto the 3-month difference in $\delta^{EH}_{w}Y_{t,5y}$ ($\Delta^{EH}_{3m,w}Y_{t,5y} \equiv \delta^{EH}_{w}Y_{t,5y} - \delta^{EH}_{w}Y_{t-3m,5y}$) for w = 6-month and 12-month:

$$\Delta^{svy}_{3m,w}Y_{t,5y} = \alpha + \beta\ \Delta^{EH}_{3m,w}Y_{t,5y} + \xi_t. \tag{40}$$

If the survey expectation were consistent with the expectations hypothesis, α should be 0 and β should be 1. Instead, we find that the β coefficients have the wrong sign. Some of the high-frequency variation in $\delta^{svy}_{w}Y_{t,5y}$ is likely to be due to "measurement errors," but a significant fraction can be expected to reflect the variation in expectations; indeed, the comparison of the survey-based expected change with the actual change shows that many of the sharp movements in $\delta^{svy}_{w}Y_{t,5y}$ do correspond to the changes that were realized later. For example, as shown with

			α	β	R^2
EH	ND	$6m$	-0.01(0.03)	0.36(0.13)	0.05
		$12m$	-0.13(0.04)	0.83(0.08)	0.38
EH	D	$6m$	0.03(0.02)	-1.53(0.29)	0.15
		$12m$	0.03(0.02)	-0.51(0.19)	0.04
Model	ND	$6m$	0.04(0.02)	0.32(0.07)	0.11
		$12m$	0.09(0.03)	0.73(0.06)	0.48
Model	D	$6m$	0.02(0.02)	0.31(0.10)	0.06
		$12m$	0.02(0.02)	0.41(0.07)	0.18

Table 4: Coefficients from the regression of survey forecasts on model-implied forecasts. ND (no-differencing) refers to eqs. (39) and (41), and D (3-month differencing) refers to eqs. (40) and (42).

arrows in Figure 6a, $\delta_{6m}^{svy} Y_{t,5y}$ peaks around Jun. 1992, Oct. 1996, and Aug. 2002; the realized values $Y_{t+3m,5y} - Y_{t,5y}$ also peak at those times.

We also compared the survey-based expected change in long-term yield with the expected change implied by the model estimated with survey data on the forecast of short-term interest rates (S1990). Figures 7a displays the model-implied 6-month-ahead expected change in the 5-year par yield, together with the survey-based expected change; the corresponding plot for 12-month-ahead is shown in Figures 7b.[42] As can be seen, $\delta_w^{model} Y_{t,5y}$ captures many of the short-run movements in $\delta_w^{svy} Y_{t,5y}$. The visual agreement between the model and the survey-based forecast is particularly notable for the 12-month-ahead forecasted change, $\delta_{12m}^{model} Y_{t,5y}$.

The results from the regressions corresponding to eqs. (39) and (40), i.e.,

$$\delta_w^{svy} Y_{t,5y} = \alpha + \beta\, \delta_w^{model} Y_{t,5y} + \xi_t \tag{41}$$

$$\Delta_{3m,w}^{svy} Y_{t,5y} = \alpha + \beta\, \Delta_{3m,w}^{model} Y_{t,5y} + \xi_t \tag{42}$$

are shown in the bottom panel of Table 4. The R^2s in regression (41) are larger than those in eq. (39), and the sign of β in eq. (42) is now positive.

These findings suggest that the survey forecasts of long-term yields are broadly in line with the configuration of expectations and time-varying risk premia in our model. The ability of the model to capture many of the short-run movements in the survey expected change in long-term yields can be linked to the time-variation of market price of risk: according to rational asset pricing (in the three-factor pure-Gaussian model), the expected change in bond yield with maturity τ for short horizon w is approximately given by

$$E_t(y_{t+w,\tau} - y_{t,\tau}) \approx (y_{t,\tau} - r_t)w/\tau + w\tau b_\tau \Sigma \lambda_t. \tag{43}$$

[42] Because the formula for the expectation of par yields is available in a closed form, we have evaluated the expectation of par yields using simulation. We have also tried an analytical approximation (useful if the long-yield forecast data were also used in the Kalman filter), and obtained very similar results.

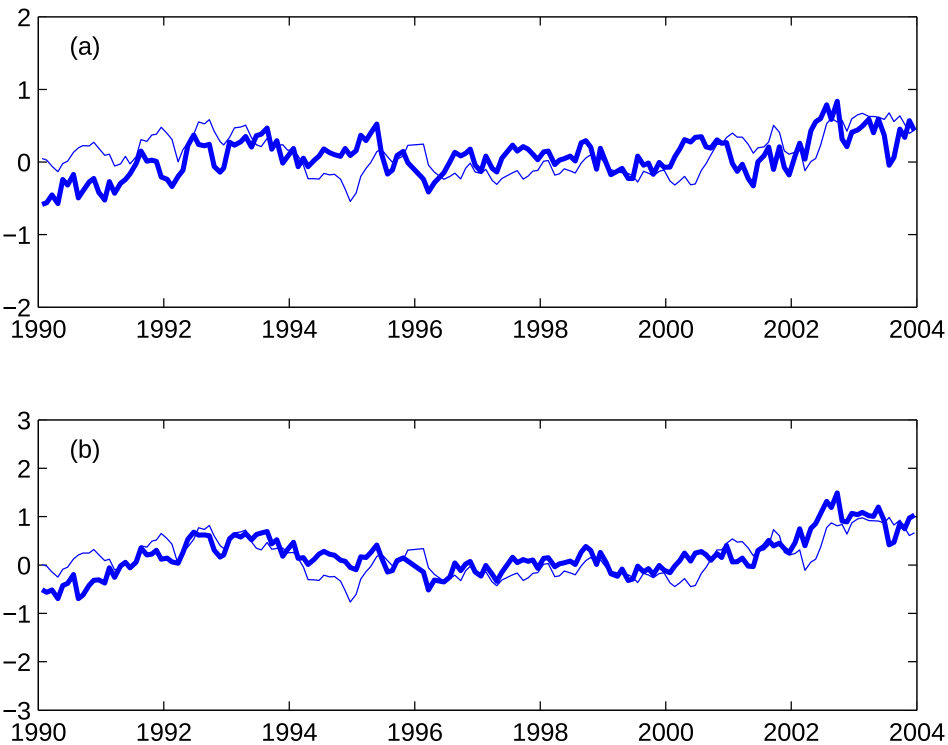


Figure 7: Comparison of the survey-based expected change in the 5-year par yield (thick solid line) and the model-based expected change (thin solid line). (a) 6-month horizon, (b) 12-month horizon.

Because the "term spread" $(y_{t,\tau} - r_t)$ varies mainly over the business cycle and is typically fairly smooth over monthly and quarterly frequencies, a large number of relatively sharp movements in $\delta_w Y_{t,5y}$ point to substantial time-variation in the market price of risk λ_t.

However, at certain periods in our sample period, for example, in 1994 and 2000, the model-based expected change and the survey-based expected change do display a notable difference. The 1994 episode concerns the (still rather poorly understood) large increase in long-term bond yields following the beginning of the Federal Reserve's monetary tightening in February of that year. (Campbell (1995) presents a detailed review of this episode.) A more recent example is the monetary tightening episode that began in June 2004 when the survey forecasts of long-term yields performed poorly: The survey participants forecasted a fairly rapid rise in the long yields in the latter half of 2004 and the first half of 2005 which did not materialize; on the other hand, the term structure model did not predict a rapid rise in long yields. It is unclear whether such instances of deviations reflect a fundamental economic mechanism that is absent in our model (e.g., differences between the average investor and the marginal investor), model specification error, institutional effects, or occasional temporary departures from rationality—as defined in the context of our model.

8 Monte Carlo evidence

In this section, we examine the severity of the small-sample problem via a Monte Carlo experiment (i.e., in a setting where the problem is not confounded with a potential misspecification). We generate 50 artificial samples of term structure data and artificial survey forecasts of length equal to our 1990-2003 sample from the S1990 model, and perform both the conventional estimation and the estimation with (artificial) survey data to study the dispersion and bias of the estimates of the parameters and the quantities of economic interest.[43] This exercise can also serve as a consistency check of the approach that we proposed in this paper.

The artificial yield and survey forecast data were generated by computing the model-implied yields and expectations from the state variables simulated according to the model. On top of these, we add measurement errors with characteristics similar to the fitting errors in Section 4. More specifically, we simulate yield measurement errors and 6-month-ahead and 12-month-ahead 3-month-yield forecast measurement errors such that their autocorrelation and cross-sectional correlation are similar to those of the the fitted errors $[\hat{\eta}_{t,3m}, ..., \hat{\eta}_{t,10y}, \hat{e}_{t,6m}, \hat{e}_{t,12m}]$ based on the S1990 estimate and the data in Section 4. The measurement error $e_{t,\mathcal{L}}$ is drawn from an iid normal distribution with variance 0.0020^2, which is about the size of the fitting errors $\hat{e}_{t,\mathcal{L}}$ in the 1995-2004 period (see Figure 4b).

Table 5 displays the true parameter values, the mean, median and standard deviation of the model parameters estimated without (artificial) survey data, and the mean, median and standard deviation of the model parameters estimated with (artificial) survey data. There is a notable upward bias in the mean-version parameters ($\mathcal{K}$) in the estimation without the survey data, which is substantially reduced when the survey data is used. Incidentally, we note that in the estimation with survey data, there are occasionally outlier-like cases; excluding these cases would make the improvements even more spectacular, as can be seen from the comparison of the mean and median of the $\mathcal{K}_{11}$ estimates. Besides the reduction of the bias, there is also a decrease in the dispersion (standard deviations) of the estimated parameters in the estimation with survey data.

In a flexibly specified model like this, many of the individual parameters of the models are not easily interpretable; more interesting are the estimates of the quantities of economic interest, such as forward term premia and their variability. To this end, we examine the mean,

[43] We are limited to a small number of samples because the estimation of multifactor models is fairly time-consuming, but even with only 50 samples we seem to get fairly indicative results.

	True	Without survey	With survey
$\mathcal{K}_{11}$	0.0539	0.4300 [0.3149] (0.3724)	0.2446 [0.0984] (0.3257)
$\mathcal{K}_{21}$	-0.1486	-0.5101 [-0.2857] (0.7925)	-0.3913 [-0.2557] (0.3569)
$\mathcal{K}_{31}$	0	0	0
$\mathcal{K}_{22}$	0.4841	0.8649 [0.7697] (0.4406)	0.4937 [0.4857] (0.1483)
$\mathcal{K}_{32}$	-6.1308	-6.5044 [-6.5417] (1.1764)	-5.8669 [-5.9161] (1.2970)
$\mathcal{K}_{33}$	2.1521	2.6334 [2.5841] (0.5382)	2.3640 [2.3662] (0.5489)
Σ_{11}	0.0023	0.0019 [0.0020] (0.0007)	0.0020 [0.0021] (0.0005)
Σ_{22}	0.0021	0.0023 [0.0023] (0.0003)	0.0024 [0.0023] (0.0004)
Σ_{33}	0.0062	0.0062 [0.0062] (0.0003)	0.0061 [0.0061] (0.0004)
ρ_o	0.0486	0.0486 [0.0479] (0.0120)	0.0489 [0.0488] (0.0060)
λ_{a1}	-0.7828	-0.6477 [-0.6025] (0.3162)	-0.6750 [-0.6539] (0.2164)
λ_{a2}	-0.6960	-0.7649 [-0.7413] (0.7581)	-0.7627 [-0.6911] (0.6288)
λ_{a3}	-1.8420	-1.7935 [-1.8011] (1.0762)	-1.7655 [-1.7582] (0.7062)
$[\mathcal{K}\Lambda_b]_{11}$	-0.1615	-0.4925 [-0.4152] (0.3739)	-0.3302 [-0.2487] (0.3225)
$[\mathcal{K}\Lambda_b]_{21}$	-0.2564	-0.2134 [-0.2608] (0.6913)	-0.2638 [-0.2544] (0.2864)
$[\mathcal{K}\Lambda_b]_{31}$	-1.2934	-1.6499 [-1.5142] (1.2332)	-1.5343 [-1.4187] (1.0310)
$[\mathcal{K}\Lambda_b]_{12}$	-0.8328	-0.5705 [-0.5633] (0.4117)	-0.5942 [-0.6134] (0.3001)
$[\mathcal{K}\Lambda_b]_{22}$	-0.9081	-1.3733 [-1.3131] (0.4398)	-0.9945 [-0.9969] (0.1260)
$[\mathcal{K}\Lambda_b]_{32}$	-1.0071	-0.0754 [-0.3415] (1.4984)	-0.5022 [-0.8176] (1.2381)
$[\mathcal{K}\Lambda_b]_{13}$	0.2974	0.2240 [0.2067] (0.1538)	0.2428 [0.2342] (0.1128)
$[\mathcal{K}\Lambda_b]_{23}$	0.2889	0.3505 [0.3433] (0.1072)	0.3822 [0.3383] (0.1567)
$[\mathcal{K}\Lambda_b]_{33}$	0.4122	-0.0426 [-0.0274] (0.6030)	0.2433 [0.2203] (0.5035)
ω_{3m}	0.0008	0.0009 [0.0009] (0.0001)	0.0009 [0.0009] (0.0001)
ω_{6m}	0	0	0
ω_{1y}	0.0003	0.0005 [0.0005] (0.0000)	0.0005 [0.0005] (0.0000)
ω_{2y}	0	0	
ω_{4y}	0.0002	0.0003 [0.0003] (0.0000)	0.0003 [0.0003] (0.0000)
ω_{7y}	0	0	0
ω_{10y}	0.0004	0.0005 [0.0005] (0.0000)	0.0005 [0.0005] (0.0000)
$\tilde{\omega}_{6m}$	0.0017		0.0017 [0.0017] (0.0001)
$\tilde{\omega}_{12m}$	0.0019		0.0018 [0.0018] (0.0002)
$\tilde{\omega}_c$	0.0016		0.0015 [0.0015] (0.0002)
ϕ_{6m}	0.9203		0.9141 [0.9129] (0.0135)
ϕ_{12m}	0.9555		0.9486 [0.9470] (0.0116)

Table 5: Mean, median (in [.]), and standard deviation (in (.)) of the estimated parameters according to estimations with and without "survey" data.

	Without survey	With survey
Φ^a_{2y}	0.27 [0.27] (0.09)	0.08 [0.08] (0.04)
Φ^a_{10y}	0.50 [0.40] (0.27)	0.12 [0.10] (0.06)
Φ^b_{2y}	1.27 [1.28] (0.24)	1.00 [1.01] (0.09)
Φ^b_{10y}	1.42 [1.45] (0.17)	1.16 [1.13] (0.16)
Γ_{2y}	0.74 [0.71] (0.29)	1.12 [1.01] (0.10)
Γ_{10y}	0.23 [0.08] (0.29)	0.67 [0.78] (0.29)
$\mathcal{T}_{6m}$	0.95 [0.96] (0.05)	1.00 [1.00] (0.02)
$\mathcal{T}_{12m}$	0.92 [0.94] (0.08)	1.00 [1.00] (0.03)

Table 6: Mean, median (in [.]), and standard deviation (in (.)) of Φ^a_w, Φ^b_w, Γ_w (for $w = 2y, 10y$) and $\mathcal{T}_w$ (for $w = 6m, 1y$) from estimations with and without "survey" data.

median and standard deviation of the following quantities:

$$\Phi^a_w = (1/T)\int_0^T dt|\hat{\varphi}_{t,w} - \varphi_{t,w}|, \tag{44}$$

$$\Phi^b_w = \frac{\int_0^T dt(\hat{\varphi}_{t,w} - \overline{\hat{\varphi}_{t,w}})^2}{\int_0^T dt(\varphi_{t,w} - \overline{\varphi_{t,w}})^2}, \tag{45}$$

where $\hat{\varphi}$ is the estimated forward term premium, φ is the true term premium, and $\overline{\hat{\varphi}}$, $\overline{\varphi}$ are sample averages ($\overline{\hat{\varphi}_{t,w}} = 1/T\int_0^T dt\hat{\varphi}_{t,w}$, $\overline{\varphi_{t,w}} = 1/T\int_0^T dt\varphi_{t,w}$). The quantity Φ^a_w simply measures how far the estimated forward term premium is from the true forward term premium, and the quantity Φ^b_w measures whether the variability of the estimated term premium is overstated (> 1) or understated (< 1). We also examine the median and standard deviation of

$$\Gamma_w = \frac{\int_0^T dt(\hat{E}_t(r_{t+w}) - \overline{\hat{E}_t(r_{t+w})})^2}{\int_0^T dt(E_t(r_{t+w}) - \overline{E_t(r_{t+w})})^2}, \tag{46}$$

which measures the variability of the estimated expected short rate (compared to the variability of the true expected short rate), and

$$\mathcal{T}_w = \frac{\int_0^T dt(y_{t+w,3m} - \hat{E}_t(y_{t+w,3m}))^2}{\int_0^T dt(y_{t+w,3m} - E_t(y_{t+w,3m}))^2}, \tag{47}$$

which is the ratio of the estimated RMSE of the short-term yield ($3m$ yield) forecast and the true RMSE.

As can be seen in Table 6, the deviation of the estimated term premium from the true values are quite substantial (Φ^a_ws are large) in the case of the estimation without survey data. Furthermore, the tendency to overstate the variation of the term premia and to understate the variation of the expected short rate is clearly visible ($\Phi^a > 1, \Gamma < 1$). The use of survey data is shown to correct these problems, reducing the Φ^a_w values and pulling the Φ^b_w, Γ_w numbers toward 1.

It is also interesting that in the estimation without survey data, the estimated RMSE of the short-term yield forecast error is on average smaller than the true RMSE ($\mathcal{T} < 1$), i.e., the estimated forecast errors are more likely to be smaller than actual forecasting error. This may appear somewhat surprising, since one could have imagined that the sampling error (uncertainty in the parameter estimate) would increase the forecast error; look-ahead-bias apparently can have more than a countervailing effect.

To be sure, the extent of the small sample problem in a Monte Carlo experiment depends on the parameter set from which the artificial data are generated. The dramatic small-sample effects reported here owe much to the fact that the chosen parameter estimates imply a lot of persistence in the interest rates. Had this persistence been lower, the small-sample problems would have been less severe. However, as we noted earlier, the persistence reflected in the estimates of our baseline model is necessary for producing a realistic variation in long-horizon expectations and is thus a desirable characteristic of the model. Viewed in this light, these Monte Carlo results provide a cautionary example that if a model exhibits a substantial degree of persistence, the conventional estimation is likely to miss that.

9 Concluding remarks

Extant studies of flexibly specified no-arbitrage models of the term structure of interest rates have not examined long-horizon expectations of the short-term interest rate in much detail. In this paper we show that conventional estimation of dynamic term structure models that contain a plausible behavior of the long-horizon expectations may face a severe small sample problem leading to imprecision, bias, and lack of robustness in estimates of objects of economic interest.

As a simple and effective solution to the problem, we propose to incorporate the information embedded in survey forecasts of short-term interest rates. With this approach, we are able to obtain more reasonable and stable estimation results of the underlying term structure model. In addition, our analysis highlights features of the survey forecasts which are of interest on their own.

By focusing on the U.S. Treasury market, where data are available for a much longer sample than the 1990-2003 period we concentrate on, we are also able to compare our short-sample estimates with those from a much longer sample (1965-2003) where we would not expect a similar small-sample problem. In the longer sample, potential misspecification due to structural change would be a greater concern. Conventional estimation does indicate notable differences in the 1965-2003 and 1990-2003 estimates, which could be interpreted as supporting the concern of structural change. Surprisingly, the estimated term structure model for the shorter sample once we include the survey data is similar to that obtained with the conventional estimation over the 1965-2003 sample. This suggests that the evidence for structural change in the U.S. Treasury term structure over the past few decades may be considerably weaker than suggested by comparisons of conventional estimates.[44] In addition, it speaks well for the flexibility of the underlying model, as it appears sufficiently flexible to accommodate variation that might be interpreted as structural change in simpler or more rigid models.

Regarding the surveys, we find that the survey forecasts of *longer-term* interest rates show significant departures from the expectations hypothesis but also find that these departures are broadly in line with the departures implied by the expectations embedded in the dynamic term structure model with a flexible risk premium specification that we estimate. This suggests that *risk* is an important element in understanding the evolution of survey expectations. We also find some episodes when the survey forecasts are not well explained by the model. A better understanding of these cases would enrich our understanding of market expectations and improve our characterization of the rich behavior of the term structure of interest rates. More generally, however, our results suggest that surveys of financial market participants' forecasts are quite informative, which points to their potential usefulness in other related term-structure applications.

A Formulae and notations

In this appendix, we provide formulae for several quantities that appear in the paper. Zero-coupon bond yields in the Gaussian model is given by eq. (4), with

$$\begin{aligned} a_\tau &= -\frac{1}{\tau}[(\mathcal{K}^*\mu^*)'(M_{1,\tau} - \tau I)\mathcal{K}^{*-1'}\rho \\ &+ \frac{1}{2}\rho'\mathcal{K}^{*-1}(M_{2,\tau} - \Sigma\Sigma' M_{1,\tau} - M_{1,\tau}\Sigma\Sigma' + \tau\Sigma\Sigma')\mathcal{K}^{*-1'}\rho - \tau\rho_o] \qquad (48) \\ b_\tau &= \frac{1}{\tau}M_{1,\tau}\rho, \qquad (49) \end{aligned}$$

where

$$\begin{aligned} M_{1,\tau} &= -\mathcal{K}^{*-1'}(e^{-\mathcal{K}^{*'}\tau} - I) \qquad (50) \\ M_{2,\tau} &= -vec^{-1}((\mathcal{K}^* \otimes I) + (I \otimes \mathcal{K}^*))^{-1} vec(e^{-\mathcal{K}^*\tau}\Sigma\Sigma' e^{-\mathcal{K}^{*'}\tau} - \Sigma\Sigma')). \qquad (51) \end{aligned}$$

Alternative closed-form expressions for a_τ and b_τ can be found in Langetieg (1980) and Dai and Singleton (2002).

[44] We note, however, that the question of structural change depends on the issues considered and here we have focused on the expectation of nominal interest rates. In a more detailed model that decomposes the nominal term structure into inflation expectations, *ex ante* real rates, and other factors, evidence for a structural break is likely to be stronger. In addition, the volatility of the short-term yield based on the S1990 estimate is substantially lower than that based on the NS1965 estimate, but a fully satisfactory treatment of the volatility issue likely requires a more sophisticated model with stochastic volatility.

The par yield (for semiannual coupon payments, as in the US Treasury notes) is given by

$$Y_{t,\tau} = 2\frac{1 - P_{t,\tau}}{\sum_{j=1}^{2\tau} P_{t,j/2}}, \tag{52}$$

where $P_{t,\tau} = \exp(-\tau y_{t,\tau})$ is the price of a zero-coupon bond with time-to-maturity τ. The par forward rate $F_{t,w,\tau}$ (appearing in eq. (38)) is

$$F_{t,w,\tau} = 2\frac{P_{t,w} - P_{t,w+\tau}}{\sum_{j=1}^{2\tau} P_{t,w+j/2}}. \tag{53}$$

The (zero-coupon) forward rate $f_{t,w,\tau}$ (in eq. (14)) is simply

$$f_{t,w,\tau} = (\log(P_{t,w}) - \log(P_{t,\tau+w}))/\tau. \tag{54}$$

The instantaneous forward rate (in eq. (11)) is

$$f_{t,w} = -\frac{d}{dw}\log(P_{t,w}). \tag{55}$$

The "model-implied" quantities in the paper are calculated by substituting Kalman-smoothed estimate of the state vector $x_{t|T}$ ($\equiv \hat{x}_t$) into relevant expressions.[45] For example, the model-implied w-period-ahead expectation of the 3-month yield at time t is evaluated as $a_{3m}(\hat{\theta}) + b_{3m}(\hat{\theta})'(I - e^{-\hat{\mathcal{K}}w})\hat{\mu} + b_{3m}(\hat{\theta})'e^{-\hat{\mathcal{K}}w}\hat{x}_t$, where $\hat{\theta}$ collectively denotes estimated parameters. For clarity, we use the "hat" symbol at several places in the main text to emphasize the estimated character of the quantities (as opposed to the data values or true values), e.g., $\hat{E}_t(y_{t+w,\tau})$ in Section 4 and $\hat{\varphi}_{t,w}$ in Section 6.

B Notes on the data used in this paper

We use weekly (Wednesday) zero-coupon term structure data, with maturities of $3m, 6m, 1y, 2y, 4y, 7y, 10y$ for estimation over the 1990-2003 sample. The zero-coupon yields for maturities greater than or equal to a year used in the estimation are taken from a zero-coupon yield dataset constructed at the Federal Reserve Board based on the prices of off-the-run coupon securities using the flexible form of smoothing suggested by Svensson (1995). (This "Svensson data" is currently available daily, from 1988 and on, up to maturities of 20 years, at intervals of 3 months.) Besides their use in the estimation, the smoothed term structure data are also used to construct the par yields and forward rates used in the paper. Because the estimation uses survey data on the 3-month T-bill yield forecasts, we use the T-bill yields data for short-maturity yields in the estimation ($3m$ and $6m$). The data used for the longer-sample estimation (1965-2003) are an update of the data used in Duffee (2002).

The monthly forecasts of 3-month T-bill yield are based on the *Blue Chip Financial Forecasts* (BCFF). The BCFF is published on the first day of each month and presents forecasts from a survey conducted during two consecutive business days one to two weeks earlier.[46] The precise dates of the survey vary and are not generally noted in the publication. To match the yields used in the estimation, which are sampled on Wednesdays, we treat the surveys as if they are

[45]See, for smoothing formulae, Kim (2005). In practice, the difference between the smoothed state variable $x_{t|T}$ and the filtered state variable $x_{t|t}$ is negligible.

[46]We thank Randy Moore for providing information on the surveys and when they are conducted.

conducted on the Wednesday closest to the likely actual dates of the survey. We used the following assumption: if the last day of the month (in which the survey is done) is Monday or Tuesday, we take the forecast date to be the last Wednesday of the month. If the last day of the month is between Wednesday and Sunday, we take the forecast date to be the next to last Wednesday of the month. For December, the survey is prepared before Christmas so the selected date may be somewhat earlier. If Christmas falls on a weekend we select the Wednesday before Christmas. Otherwise, we select the Wednesday of the week prior to the week where Christmas falls. For example, in the BCFF survey with date marked Jan 1, 2004, we take Dec 17, 2003 (Wednesday) as the actual date of the forecast.

The *Blue Chip* forecasts of the 3-month T-Bill yield are forecasts of the *quarterly average* of the 3-month T-Bill yield for the current quarter, next quarter, and so on, out to 5 or 6 quarters ahead. From these quarter-average forecasts, we construct 6-month-ahead and 12-month-ahead forecasts by treating the quarter-averaged forecast as approximately equal to the forecast for mid-quarter (Feb. 14, May 15, Aug. 15, Nov. 15) and interpolating between the mid-quarter points.[47] For example, in the Jan 1, 2004 *Blue Chip* survey, the forecast of 3-month T-bill yields for 2004Q2, 2004Q3, and 2004Q4 are 1.2, 1.5, and 1.8, respectively. From these, we obtain the 6-month-ahead and 12-month-ahead forecast from Dec 17, 2003 as 1.31 and 1.97.

Twice a year, generally on the June 1 and December 1 issues, the BCFF also provides a long-horizon forecast of the 3-month T-bill approximately 6-to-11 year out. The approximation is due to the fact that survey participants are asked to provide a forecast for a five-calendar-year average. For example, in the June 1, 2002 issue, the survey participants were asked to forecast the five-year average T-bill yield between 2008-2012, thus it is, more accurately, a $5\frac{7}{12}$-to-$10\frac{7}{12}$ year horizon forecast, i.e., $T_1 = 5\frac{7}{12}, T_2 = 10\frac{7}{12}$ in eq. (19).

The Treasury Constant Maturity yields are available from the Federal Reserve Board's H.15 website.

References

Ball, C. A. and W. N. Torous (1996), "Unit roots and estimation of interest rate dynamics," *Journal of Empirical Finance*, 3, 215-238.

Bekaert, G. , R. Hodrick and D. Marshall (1997), "On biases in tests of expectations hypothesis of the term structure of interest rates," *Journal of Financial Economics*, 44, 309-348.

Bernanke, B. S. (2003), "An Unwelcome Fall in Inflation?" Remarks before the Economics Roundtable, University of California, San Diego, La Jolla, California, July 23.

Best, P., A. Byrne, and A. Ilmanen (1998), "What really happened to U.S. bond yields," *Financial Analyst Journal*, May/June, 41-49.

Blinder, A. S. (1997), "Distinguished lecture on economics and government: What central bankers could learn from academics – and vice versa," *Journal of Economic Perspectives*, 11, 3-19.

Blinder, A. S. and J. L. Yellen (2001), *The Fabulous Decade: Macroeconomic Lessons from the 1990s*, The Century Foundation Press, New York.

Campbell, J. Y. (1995), "Lessons from the yield curve," *Journal of Economic Perspectives*, 9, 129-152.

[47] Because there is not a lot of structure in the forecast as a function of the forecasted quarters (especially for several quarters ahead), this approximation is quite accurate. Nearly identical results to this piecewise-linear interpolation are obtained with other interpolation procedures (e.g, cubic splines).

Campbell, J. Y. and R. J. Shiller (1991) "Yield spreads and interest rate movements: a bird's eye view," *Review of Economic Studies*, 58, 495-514.

Campbell, J. Y. and L. M. Viceira (2001), "Who should buy long-term bonds?" *American Economic Review*, 91(1), 99-127.

Clarida, R., J. Gali and M. Gertler (2000), "Monetary policy rules and macroeconomic stability," *Quarterly Journal of Economics*, 65, 147-180.

Chun, A. L. (2005), "Expectations, bond yields, and monetary policy," working paper.

Cogley, T. and T. Sargent (2001), "Evolving post-world war II U.S. inflation dynamics," in *NBER Macroeconomics Annual*, 16, 331-373.

Croushore, D. (1998), "Low inflation: The surprise of 1990s", *Federal Reserve Bank of Philadelphia Business Review*, July/August, 3-12.

Dai, Q. and K. J. Singleton (2000), "Specification analysis of affine term structure models," *Journal of Finance*, 55, 1943-1978.

Dai, Q. and K. J. Singleton (2002), "Expectations puzzles, time-varying risk premia, and dynamic models of the term structure," *Journal of Financial Economics*, 63, 415-441.

de Jong, F. (2000), "Time series and cross-sectional information in affine term structure models," *Journal of Business Economics and Statistics*, 18, 300-314.

Derby, M. S. (2004), "Fed watch: Fed struggles to define neutral rate policy," *Dow Jones Capital Markets Report*.

Duarte, J. (2004), "Evaluating an alternative risk preference in affine term structure models," *Review of Financial Studies*, 17(2), p. 379-404.

Duffee, G. R. (1996), "Idiosyncratic variation of Treasury bill yields," *Journal of Finance*, 51, 527-551.

Duffee, G. R. (2002), "Term premia and interest rate forecasts in affine models," *Journal of Finance*, 57, 405-443.

Duffee, G. R. and R. Stanton (2004), "Estimation of dynamic term structure models," working paper.

Friedman, B. M. (1980), "Survey evidence on the 'rationality' of interest rate expectations," *Journal of Monetary Economics*, 6, 453-465.

Froot, K. A. (1989), "New hope for the expectations hypothesis of the term structure of interest rates," *Journal of Finance*, 44, 283-305.

Goto, S. and W. N. Torous (2003), "The Conquest of U.S. Inflation: Its Implications for the Fisher Hypothesis and the Term Structure of Interest Rates," working paper, November.

Gupta, A. and Subrahmanyam, M. (2000), "An empirical examination of the convexity bias in the pricing of interest rate swaps," *Journal of Financial Economics*, 55, 239-279.

Gurkaynak, R., B. Sack, and E. Swanson (2002), "Market-based measures of monetary policy expectations" Finance and Economics Discussion Series 2002-40, Federal Reserve Board, Washington, D.C.

Gurkaynak, R., B. Sack, and E. Swanson (2005), "The sensitivity of long-term interest rates to economic news: evidence and implications for macroeconomic models," *American Economic Review*, 95, 425-436.

Harrison and Kreps (1978), "Speculative investor behavior in a stock market with heterogeneous expectations," *Quarterly Journal of Economics*, 92, 323-336.

Harvey, A. C. (1989), *Forecasting, structural time series models and the Kalman filter*, Cambridge; New York: Cambridge University Press.

Hoenig, T. M. (2005), “The U.S. economic outlook and monetary policy under an inflation watch,” remarks at the Wichita Bankers’ Forum, Wichita, Kansas, June 16.

Jagannathan, R., A. Kaplin, and S. Sun (2003), “An evaluation of multi-factor CIR models using LIBOR, swap rates, cap and swaption prices,” *Journal of Econometrics*, 116, 113-146.

Kim, C. J. and C. R. Nelson (1999), “Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle,” *Review of Economics and Statistics*, 81(4), p. 608-616, November.

Kim, D. H. (2005), “Time-varying risk and return in the quadratic-Gaussian model of interest rates,” working paper.

Kim, D. H. and J. Wright (2005), “An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates,” Finance and Economics Discussion Series 2005-33, Federal Reserve Board, Washington D.C.

Kohn, D. (1999), “Forward-Looking Rules for Monetary Policy: Comment,” in *Monetary Policy Rules*, ed. by J. B. Taylor, 192-199.

Kozicki, S. and P. A. Tinsley (2001), “Shifting endpoints in the term structure of interest rate,” *Journal of Monetary Economics*, 47(3), 613-652.

Langetieg, T. C. (1980), “A multivariate model of the term structure,” *Journal of Finance*, 71-97.

Mayshar, J. (1983), “On divergence of opinion and imperfections in capital markets,” *American Economic Review*, 74, 114-128.

Merton, R. C. (1980), “On estimating the expected return on the market,” *Journal of Financial Economics*, 9, 323-361.

Meyer, L. H. (2004). *A Term at the Fed: An Insider’s View*, Harper Collins, New York.

Mishkin, F. S. (1981), “Are market forecasts rational?” *American Economic Review*, 71, 295-306.

Orphanides, A. (2004), “Monetary policy rules, macroeconomic stability and inflation: A view from the trenches,” *Journal of Money, Credit and Banking*, 36(2), p. 151-175, April.

Orphanides, A. and D. Wilcox (2002), “The opportunistic approach to disinflation,” *International Finance*, 5(1), 47-71, Spring.

Pennacchi, G. (1991), “Identifying the dynamics of real interest rates and inflation: Evidence using survey data,” *Review of Financial Studies*, 4(1), p. 53-86.

Phillips, P. C. B. and J. Yu (2005), “Jackknifing bond option prices,” *Review of Financial Studies*, 18(2), p. 707-742.

Rudebusch, G. and Wu, T. (2004), “The recent shift in term structure behavior from a no-arbitrage macro-finance perspective,” working paper, November.

Sims, C. and T. Zha (2004), “Were there regime switches in US monetary policy?” working paper

Svensson, L. E. O. (1995), “Estimating forward interest rates with the Extended Nelson and Siegel Method,” *Quarterly Review, Sveriges Riksbank*, 1995:3, 13-26