JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 47, No. 1, Feb. 2012, pp. 241–272
COPYRIGHT 2012, MICHAEL G. FOSTER SCHOOL OF BUSINESS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195
doi:10.1017/S0022109011000627

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

Term Structure Estimation with Survey Data on Interest Rate Forecasts

Don H. Kim and Athanasios Orphanides*

Abstract

The estimation of dynamic no-arbitrage term structure models with a flexible specification of the market price of risk is beset by severe small-sample problems 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. To illustrate the methodology, we estimate the 3-factor affine-Gaussian model with U.S. Treasury yields data and demonstrate that incorporating information from survey forecasts mitigates the small-sample problem. The model thus estimated for the 1990–2003 sample generates a stable and sensible 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.

I. 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 this information to infer market-based expectations of future monetary policy and 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 (EH), for instance, takes the current forward rate curve as the expected path of

*Kim, donhkim@yonsei.ac.kr, School of Business, Yonsei University, 262 Seongsanno, Seodaemun-gu, Seoul 120-749, Korea; Orphanides, athanasios.orphanides@centralbank.gov.cy, Central Bank of Cyprus, 80 Kennedy Ave., Nicosia 1076, Cyprus. We thank Jim Clouse, Brian Sack, and Jonathan Wright for helpful input and encouragement, and Hendrik Bessembinder (the editor), Mikhail Chernov, Jefferson Duarte (associate editor and referee), Greg Duffee, Antti Ilmanen, Monika Piazzesi, Peter Tinsley, and participants of presentations at the Bank for International Settlements (BIS), Bank of England, University of Frankfurt, Universite Libre de Bruxelles, University of Basel, the 11th International Conference on Computing in Economics and Finance (Washington, DC, 2005), the Annual Conference of the German Economic Association (Bayreuth, 2006), and the 2007 American Finance Association (AFA) Chicago meetings, for useful comments.

the short rate by abstracting from the presence of term premia. It is widely acknowledged that the yield curve contains time-varying term premia, however, so the EH 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 EH. 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 situations, however, the implementation of these models is subject to considerable practical difficulty. Estimation with maximum likelihood methods may not yield suitable estimates due to the "shallowness" of the likelihood function, and the resulting estimated model may display features that most practitioners would dismiss as implausible. The purpose of the present paper is to draw attention to this problem and to explore the use of survey forecast data as a potential remedy.

We argue that the basic difficulty is the "small-sample problem": Typical data samples used in a dynamic term structure estimation may be too short to provide a reliable characterization of the physical dynamics of the interest rate process. The fundamental issue 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. Another aspect of the small-sample problem (besides imprecision) is bias: In short samples the estimated persistence of the time series can be often substantially lower than the actual persistence.

Given that the problem is insufficient information in the data, it would help to provide additional relevant information. If a longer data sample is available, provided that the sample is sufficiently long and homogeneous (i.e., free of structural change), then its use would provide the most straightforward way to incorporate new information and solve the problem. In many practical situations, however, a longer sample may not be available. The most prominent case is the estimations involving London Interbank Offered Rate (LIBOR)-swap data. Due to data availability problems, most studies with interest rate swap data use data from the late 1980s or early 1990s. For example, Dai and Singleton's (2000) study uses 10 years' data starting from 1987, and Jagannathan, Kaplin, and Sun's (2003) study uses 5 years' data starting from 1995.

Furthermore, often a researcher may intentionally choose to use a relatively short sample out of concern about the structural change in a longer sample. For example, for the euro area, one might wish to explore estimations with only a post-Economic and Monetary Union (EMU) sample. Even in the case of the U.S. Treasury term structures where data are available, one might be worried that a long sample may not be structurally homogeneous (see, e.g., Kim and Nelson (1999), Cogley and Sargent (2001)). A proposed model may give a reasonably good description of the properties of a short sample but may need additional elements to describe a longer sample. Estimating over a longer sample may require complicating the model (e.g., allowing for time variation in volatility). For example, U.S. interest rate volatility currently is substantially lower and more stable

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

than in the late 1970s and early 1980s, therefore an affine-Gaussian model (which assumes constant volatility of yields) could be a good approximation of the data from the 1990s and 2000s, but less so for the long sample going back to the 1970s; an affine-Gaussian model estimated with a long U.S. sample would imply a substantially higher volatility of yields than those implied by interest rate option market prices.

We propose incorporating the information in the 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 help improve the overall precision of the estimated parameters.

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 the estimation of a dynamic term structure model. However, one may be concerned about the reliability of the survey forecast information; the relative scarcity of studies employing surveys in the finance literature may reflect such concerns. Therefore, 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. Furthermore, we allow for a flexible structure in the noise (in particular, the possibility of serial correlation) to accommodate a realistic departure of market forecasts from survey forecasts, though we also examine the simpler case of serially uncorrelated noise. As we explain, the Kalman-filter-based maximum likelihood estimation approach provides a convenient framework for incorporating survey forecast information.

We apply this methodology to the estimation of the 3-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 T-bill yield presented in the *Blue Chip Financial Forecasts* (BCFF) publication. 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 3-factor pure-Gaussian model. The model estimated with survey data is also successful in reproducing the well-known pattern of deviation from EH 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 EH. 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 the estimation with survey data. We also compare the use of survey interest rate forecast data and the use of a longer yield sample and find that the use of survey data can be even more effective than the use of a longer sample.

The rest of this paper is organized as follows. In Section II, we illustrate the small-sample problem in a specific practical setting and discuss strategies to address the problem. In Section III, we introduce the survey forecasts and develop a Kalman-filter-based approach to incorporate the information from the surveys

in the estimation. Section IV describes our empirical results, and Section V concludes.

II. The "Small-Sample" Problem

A. Problems with the Conventional Estimation

We focus on the affine-Gaussian model, which has been used in a number of studies on dynamic no-arbitrage models. Duffee (2002) and others have noted that this model is promising for capturing time-varying term premia; the affine-Gaussian framework has also been a workhorse for efforts to model term structure jointly with macro variables such as inflation and growth, as in Ang and Piazzesi (2003).

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 that 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. 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 affine-Gaussian model is a flexible and tractable model with many nice features. 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

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

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

Interest rate forecasts in affine-Gaussian models are also easy to compute. For example, the expected short rate w-periods ahead is given by[1]

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

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

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

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, $\mathrm{E}_t(r_{t+\infty})$, is trivially a constant. Although this contrasts with 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 over such long horizons.[2]

Let us now consider the following practical application: Suppose today is December 31, 2003, and we would like to estimate the 3-factor Gaussian model based on the term structure data from 1990 to today and use the estimated model to extract the market's expectation of the future short rate. We would also like to analyze the time variation of nominal yields and forward rates, in particular, decomposing them into an expectations component and a term premium component.

While what the model says about the short-rate expectation within the next few years is of obvious practical interest, what the model says about expectations over longer horizons is also of considerable importance to policymakers and market participants. Policymakers have an interest in long-term trends beyond the business-cycle time frame and the information contained in long-term interest rates. Furthermore, if the "long end" of the expectations term structure was incorrect, this would cast doubt on the shorter-horizon expectations, as the 2 sides would have to connect smoothly.

The output from the affine-Gaussian model, based on the standard maximum likelihood estimation, is shown in Figure 1.[3] Graph A of Figure 1 shows the expected path of the short rate, together with the forward rate curve ($f_{t,\tau}$ as a function of horizon τ), on Dec. 31, 2003. Graph B of Figure 1 shows the time series of the model-implied expectation of the short rate 10 years ahead, together with the 10-year instantaneous forward rate and the forward term premium $\varphi_{t,10y}$ (instantaneous forward rate minus expected short rate), that is,

$$\varphi_{t,\tau} = f_{t,\tau} - \mathrm{E}_t(r_{t+\tau}), \qquad \text{for} \quad \tau = 10y. \tag{8}$$

Expectations are inherently unobservable objects, and hence assessing and evaluating expectations is a difficult task. Often, the presence of problems in estimates of expectations may not be easy to discern. Unreasonable estimates of expectations can still appear to provide a good in-sample forecast performance, while out-of-sample forecast performance can be difficult to interpret due to weak power. In small samples, unforecastable interest rate shocks may lead to substantial forecast errors even for the true model.

These difficulties notwithstanding, the results shown in Figure 1 are sufficiently implausible that most policymakers and market participants would dismiss them immediately. Specifically, the short rate is projected to rise too slowly

[2] Kozicki and Tinsley (2001) present a model with a time-varying infinite-horizon expectation of the short-term interest rate, $\mathrm{E}_t(r_{t+\infty})$. They mention that their work is in spirit similar to the so-called "time-varying tendency model," which is nested in our specification.

[3] These results are based on the "$\mathcal{NS}90$ estimate" presented in Section IV.A.

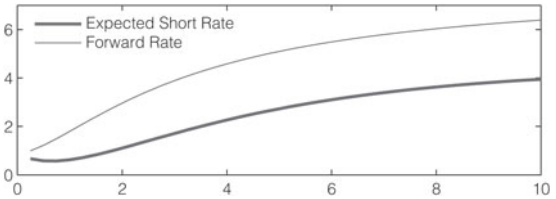
https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

FIGURE 1

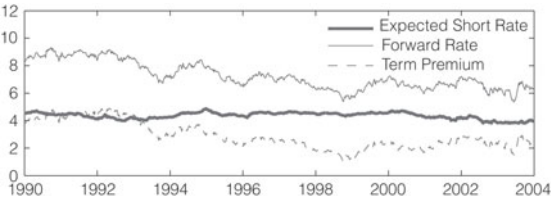
Cross Section and Time Series of Expected Short Rate

Graph A of Figure 1 shows the expected path of the short rate on Dec. 31, 2003, and the forward rate curve (horizon in years is shown on the *x*-axis). Graph B shows the time series of 10-year-ahead expected short rate, forward rate, and the forward term premium.

Graph A. Expected Short-Rate Path on Dec. 31, 2003



Graph B. Expected Short Rate 10 Years Ahead



according to the estimated model. Our reading of market participants' views based on anecdotal reports and market commentary is that rates were expected to rise considerably faster than the thick solid line in Graph A of Figure 1, as indeed turned out to be the case in 2004 and 2005.[4] Also, the time series of the 10-year-ahead expectation implied by the model is far too stable: During most of the 1990–2003 period, the 10-year-ahead expectation stayed in a fairly narrow range between 4.0% and 4.6%. This, in turn, implies that most of the longer-horizon forward rate variation is attributed to the variation in term premium, rather than the variation in the short-rate expectation.

The stability of "long-horizon" short-rate expectations (Graph B of Figure 1) is particularly troublesome in view of the consensus among market participants that the underlying trend of inflation and of interest rates had shifted down significantly from the beginning of the sample to its end. In the decade following the Volcker disinflation, expectations adjusted only slowly to the Federal Reserve's march toward price stability. This is confirmed when examining surveys of expectations of both inflation and short-term interest rates over long horizons. Figure 2 presents a long-horizon survey forecast of the short-term nominal interest rate, which shows a downward trend, amid other variations.[5] Some might question how quantitatively reliable the long-horizon interest rate forecast survey

[4]At that time (end of 2003), the federal funds rate had been at the historical low 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 tightening.

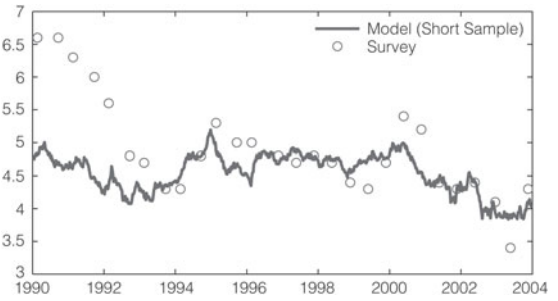
[5]The survey data are from the Blue Chip survey and are explained in greater detail later. For the comparable evolution of inflation expectations from surveys, see Figure 1 of the working paper version of this study.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

FIGURE 2

Long-Horizon Short-Rate Expectation

Figure 2 shows the long-horizon expectation of short-term interest rate from the Blue Chip survey (expected 3-month yield in the next 6–11 years) and the model-implied expectation.



is, but surely most would agree on the qualitative statement that there has been a significant variation in long-horizon interest rate expectation (and inflation expectation) over time. The model-implied value, on the other hand, shows a more stable behavior, and shows little "downward trend."[6]

Findings such as these are in fact quite common with flexibly specified no-arbitrage models; one often obtains expectations that are at odds with reasonable priors and term premia that seem too volatile. In addition, the estimation of these models often faces practical difficulties: The relevant likelihood function may have multiple local maxima, and the parameter estimates as well as estimates of variables of interest (such as term premia) can be very imprecise and lack robustness.

B. Difficulty of Identifying the Physical Dynamics

The basic source of the 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: 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 equations (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

$$\hat{x}_t = A(\mathcal{K})\hat{x}_{t-h} + b(\mathcal{K}, \mu) + \epsilon_t, \tag{9}$$

[6]The model-implied value is $\mathrm{E}_t(\bar{y}_{3m,T_1,T_2})$, described in Section III.C. This quantity is more variable than the 10-year-ahead expectation of the short rate (shown in Graph B of Figure 1), as it is an average over the next 6–11 years (and thus includes parts (e.g., 6-year horizon) that are more variable than the 10-year horizon).

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

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 equations (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 equation (9) cannot be measured very accurately when $\mathcal{K}$ is unrestricted (other than the stationarity and normalization conditions) and the sample period is relatively short (no matter how often $\hat{x}_t$ is sampled).[7] As noted earlier, long-term short-rate expectations (and long-term inflation expectations) show a visible downward trend, which means that there exists in nominal interest rates a very persistent factor with long half-life (more than 5 years). The dynamics of such a factor would be hard to pin down with just a couple of decades worth of data. Furthermore, in our sample period (1990–2003) there were only 2 recessions (1990–1991, 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 factors would then feed into other aspects of the estimation.

It is also well known that estimation of systems like equation (9) may result in an upward bias in the estimated mean reversion of the system.[8] 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, resulting 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.

Although this discussion was based on the 3-factor affine-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.[9]

In discussing the difficulties caused by the small-sample problem, it is useful to distinguish among 2 related problems. One issue is the potential presence of bias due to the size of the sample. Once identified, the presence of such a bias may be addressed with a treatment designed to provide an appropriate bias adjustment.[10]

A thornier issue is the potential *imprecision* and *lack of robustness* in parameter estimates, associated with a nearly flat likelihood function or the presence of multiple inequivalent local maxima. Obviously, this issue cannot be addressed

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

[8]Note that the 1-factor version of equation (9) 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 3-factor affine-Gaussian model make the bias assessment more complicated.

[9]Duffee and Stanton (2008) provide an example of small-sample problems in a model with a stochastic volatility factor.

[10]For example, the study by Bekaert, Hodrick, and Marshall (1997) 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.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

by "2-step" procedures adjusting for bias. The "1st-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 1st step of a 2-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*. As an example of "model constraint," consider restricting Λ_b to be a 0 matrix. In this case $\mathcal{K}$ in equation (9) would be constrained and more effectively pinned down (as we have $\mathcal{K} = \mathcal{K}^*$), thus we would not expect as much small-sample difficulty.[11] However, the $\Lambda_b = 0$ case corresponds to the EH; in view of strong evidence for the failure of the EH, such a constraint may be problematic. Overall, many of the parameters of the latent factor model like ours do not have specific economic meaning, hence it is difficult to come up with a well-motivated constraint on them.

In what follows, we propose using survey data on interest rate forecasts as such an additional source of information that would help alleviate the aforementioned problems.

III. Survey Forecasts of Interest Rates

Our basic idea is that surveys are a direct (though likely imperfect) measure of expectations in the physical measure (rather than risk-neutral measure), hence they can help pin down the parameters like $\mathcal{K}$ (and risk premium parameters Λ_b). Note also that surveys provide *real-time* information; thus they may help alleviate concerns about look-ahead bias that can arise from the use of future information in the computation of forecasts (e.g., model-based in-sample forecast).

A. Simple Analysis of Survey Forecasts

Studies using survey forecasts have been rather scarce in the finance literature;[12] this is likely due to the fact that many of the early investigations of surveys during the 1980s did not yield encouraging conclusions regarding the usefulness and reliability of surveys.[13] However, the interest rate environment has changed a lot since then, and one might expect that along with the continued development and innovations in the financial market, the sophistication of the survey participants has also increased. Thus, it is useful to take a fresh look at the survey interest rate forecasts, as we do here.

In our analysis, we use monthly data on the 6- and 12-month-ahead forecasts of the 3-month T-bill yield based on the BCFF. We denote these 2 forecasts, for a survey conducted at t, $\mathrm{E}_t^{\mathrm{svy}}(y_{t+6m,3m})$ and $\mathrm{E}_t^{\mathrm{svy}}(y_{t+12m,3m})$, respectively. The BCFF

[11] Indeed, this example would be similar to that in Ball and Torous (1996), where the use of cross-sectional yield curve information helps to better identify the time-series parameters of the short-rate process. Duffee and Stanton (2008) emphasize the specialness of the Ball and Torous case.

[12] Pennacchi (1991) is one exception, but the focus of his study is quite different from ours. He uses survey inflation forecasts to identify the dynamics of inflation expectations.

[13] See, for example, Friedman (1980) and Froot (1989). Thomas (1999) notes that although surveys have performed poorly in forecasting inflation in the 1970s, they have done much better in the 1990s.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

is published on the 1st 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 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 over shorter horizons. We denote the long-term forecast by $\mathrm{E}_t^{\mathrm{svy}}(\bar{y}_{3m,T_1,T_2})$, where $T_1 \approx 6$ and $T_2 \approx 11$. Additional details on the data are provided in Appendix B.[14]

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 EH and the random walk hypothesis (RWH). In the case of the 3-month yield $y_{t,3m}$, forecasts corresponding to these 2 hypotheses can be stated as

$$\mathrm{E}_t^{\mathrm{RWH}}(y_{t+w,3m}) = y_{t,3m}, \tag{10}$$

$$\mathrm{E}_t^{\mathrm{EH}}(y_{t+w,3m}) = f_{t,w,3m}, \tag{11}$$

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 EH and the survey forecasts; the expected change based on the RWH is trivial (simply 0). 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 forecast error based on the EH (root mean square error (RMSE) of 0.99%) compared to the RWH (RMSE of 0.85%). By comparison, the RMSE based on the survey forecast is 0.79%, somewhat smaller than that of the random walk forecast.

The very poor forecast performance of the EH suggests that, at minimum, some "term premium correction" is required to understand the data. The survey forecast can be viewed as containing such a correction: 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), while the sample mean of the expected change based on the survey is smaller than that based on the EH (0.08% vs. 0.48%).

The regression of the realized interest rate change onto the survey-forecasted change gives[15]

$$y_{t+6m,3m} - y_{t,3m} = -0.35_{(0.12)} + 1.24_{(0.31)}(\mathrm{E}_t^{\mathrm{svy}}(y_{t+6m,3m}) - y_{t,3m}) + \xi_{t+6m}, \tag{12}$$

[14]The BCFF survey also provides information on the individual respondents' forecast. This is a potentially interesting source of additional information, but in this study we shall confine our attention to the mean forecast information, as even this simple measure (interest rate forecast) has not been utilized in a term structure model estimation setting heretofore.

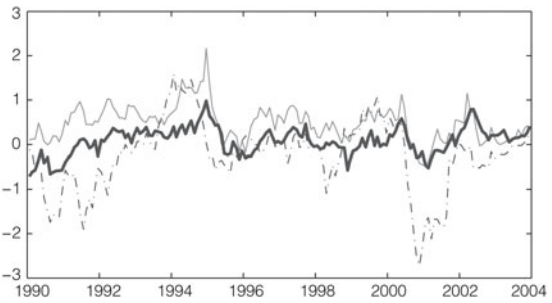
[15]In all regressions, the coefficients are reported in units of percentage points (e.g., 0.35 means 35 basis points (bp)).

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

FIGURE 3

Expected and Realized Change in the 3-Month Yield over 6 Months

In Figure 3, the thin solid line is the expectations-hypothesis-based expected change, the thick solid line is the survey-based expected change, and the realized change is shown in a dot-dashed line. All are plotted at the time of the forecast.



where the Newey-West (1987) standard errors are given in subscripts. The coefficient on the $(E_t^{svy}(y_{t+6m,3m}) - y_{t,3m})$ term is insignificantly different from 1; in other words, there is a 1-for-1 correspondence between survey-expected and realized changes. Note also that the coefficient on the constant term is indeed significantly negative. This implies a "bias," but this is a desirable feature: There is evidence that neither policymakers nor market participants fully anticipated the (relatively) low inflation in the 1990s,[16] and, consequently, one can reason 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 give an insignificant coefficient. For example,

$$(13) \qquad y_{t+6m,3m} - E_t^{svy}(y_{t+6m,3m}) = -0.39_{(0.23)} + 0.03_{(0.09)}(y_{t,10y} - y_{t,3m}) + \xi_{t+6m}.$$

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

Overall, these results are encouraging, and suggest at least that it would be worth exploring what kind of effect it would have on dynamic term structure model estimation.

B. Caveats

Some caveats do apply in the use of survey forecasts. First of all, 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. Incentive issues, that is, lack of incentive (careless

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

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

reporting) or wrong incentive (misrepresentation of true belief), may also cause complications. Lastly, no matter how accurately surveys are conducted, they are prone to some measurement error (e.g., 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 errors when using information from surveys. Therefore, 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,

$$\mathrm{E}_t^{\mathrm{svy}}(y_{t+w,3m}) \quad = \quad \mathrm{E}_t^{\mathrm{mkt}}(y_{t+w,3m}) + e_{t,w}, \tag{14}$$

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 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.

C. Term Structure 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 convenient 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 observation vector be[17]

$$o_t \quad = \quad [y_{t,\tau_1}, \ldots, y_{t,\tau_M}, \mathrm{E}_t^{\mathrm{svy}}(y_{t+6m,3m}), \mathrm{E}_t^{\mathrm{svy}}(y_{t+12m,3m}), \mathrm{E}_t^{\mathrm{svy}}(\bar{y}_{3m,T_1,T_2})]'. \tag{15}$$

It is straightforward to show

$$o_t \quad = \quad \begin{bmatrix} a_{\tau_1} \\ \vdots \\ a_{\tau_M} \\ a_{3m} + b_{3m}'\left(I - e^{-\mathcal{K}/2}\right)\mu \\ a_{3m} + b_{3m}'\left(I - e^{-\mathcal{K}}\right)\mu \\ a_{3m} + b_{3m}'(I - W_{T_1,T_2})\mu \end{bmatrix} + \begin{bmatrix} b_{\tau_1}' \\ \vdots \\ b_{\tau_M}' \\ b_{3m}' e^{-\mathcal{K}/2} \\ b_{3m}' e^{-\mathcal{K}} \\ b_{3m}' W_{T_1,T_2} \end{bmatrix} x_t + \begin{bmatrix} \eta_{t,\tau_1} \\ \vdots \\ \eta_{t,\tau_M} \\ e_{t,6m} \\ e_{t,12m} \\ e_{t,\mathcal{L}} \end{bmatrix}, \tag{16}$$

in which

$$W_{T_1,T_2} \quad = \quad \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} e^{-\mathcal{K}s} ds \quad = \quad \frac{1}{T_2 - T_1}\mathcal{K}^{-1}\left(e^{-\mathcal{K}T_1} - e^{-\mathcal{K}T_2}\right), \tag{17}$$

[17] That the survey data are not available at all dates can be handled by treating them as missing data. We use the modified observation vector $\tilde{o}_t = S_t o_t$, where S_t is a matrix of varying dimension $d(t) \times (M + 2 + 1)$ that selects only the nonmissing variables at time t (see, e.g., Harvey (1989)).

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

and $\eta_{t,\tau_1},\ldots,\eta_{t,\tau_M}$ denote yield measurement errors, and $e_{t,6m}$, $e_{t,12m}$, and $e_{t,\mathcal{L}}$ denote 6-month, 12-month, and long-term (approximately 6–11 years) survey expectation measurement errors, respectively. In deriving equation (16), we have used

$$\mathrm{E}_t(x_{t+w}) = e^{-\mathcal{K}w}x_t + (I - e^{-\mathcal{K}w})\mu, \tag{18}$$

$$\mathrm{E}_t\left(\frac{1}{T_2 - T_1}\int_{T_1}^{T_2} ds\; x_{t+s}\right) = W_{T_1,T_2}x_t + (I - W_{T_1,T_2})\mu. \tag{19}$$

The discretely measured state vector x_t follows the process

$$x_t = e^{-\mathcal{K}h}x_{t-h} + (I - e^{-\mathcal{K}h})\mu + \epsilon_t, \tag{20}$$

with $\epsilon_t \sim N(0, \Omega_h)$, where $\Omega_h = \int_0^h e^{-\mathcal{K}s}\Sigma\Sigma' e^{-\mathcal{K}'s}ds$.[18]

Equation (20), together with equation (16), with independent and identically distributed (i.i.d.) (serially and cross-sectionally uncorrelated) measurement errors

$$\eta_t = [\eta_{t,\tau_1},\ldots,\eta_{t,\tau_M}, e_{t,6m}, e_{t,12m}, e_{t,\mathcal{L}}]' \sim N(0, \Omega^\eta), \tag{21}$$

$$\Omega^\eta = \mathrm{diag}([\omega^2_{\tau_1},\ldots,\omega^2_{\tau_M}, \tilde{\omega}^2_{6m}, \tilde{\omega}^2_{12m}, \tilde{\omega}^2_{\mathcal{L}}]), \tag{22}$$

constitute perhaps the simplest Kalman-filter observation and state equations, from which the likelihood function can be constructed for estimating our model with survey forecasts.[19] Note that the parameters of the model are still estimable without the last 3 rows in equation (16); indeed, that case corresponds to the conventional estimation. The Kalman-filter setup (including survey forecasts in the observation vector) is thus a natural and convenient way to incorporate the information contained in the survey data, without any change in the underlying term structure model.[20]

The Kalman-filter setup can also accommodate 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-0 AR(1) processes with correlated innovations:

$$e_{t,6m} = \phi_{6m}e_{t-h,6m} + \varepsilon_{t,6m}, \tag{23}$$

$$e_{t,12m} = \phi_{12m}e_{t-h,12m} + \varepsilon_{t,12m}, \tag{24}$$

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

[19]The construction of the likelihood function and the estimation is then straightforward (see, e.g., de Jong (2000), Duffee and Stanton (2008), and Babbs and Nowman (1999)).

[20]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, $\mathrm{E}_t(y_{t+6m,3m})$, as a state variable (i.e., write the state vector as $\tilde{x}_t = [\mathrm{E}_t(y_{t+6m,3m}), x_{2t}, x_{3t}]'$). Besides the potential complication of survey measurement errors, this may also face a "consistency" problem: by definition, $\mathrm{E}_t(y_{t+6m,3m}) = a_{3m} + b'_{3m}(e^{-\mathcal{K}/2}[\mathrm{E}_t(y_{t+6m,3m}), x_{2t}, x_{3t}]' + (I - e^{-\mathcal{K}/2})\mu)$, which is an inconvenient condition to impose. A recent study by Chun (2011) uses survey forecasts of inflation and gross domestic product (GDP) as part of the state vector.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

in which

$$[\varepsilon_{t,6m}, \varepsilon_{t,12m}]' \sim N(0, \tilde{\Omega}^{\varepsilon}), \tag{25}$$

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

$$\tilde{\Omega}^{\varepsilon} = \tilde{\Omega}^{e} - \text{diag}([\phi_{6m}, \phi_{12m}]')\tilde{\Omega}^{e}\text{diag}([\phi_{6m}, \phi_{12m}]'). \tag{26}$$

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 nonzero 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

$$o_t = \begin{bmatrix} a_{\tau_1} \\ \vdots \\ a_{\tau_M} \\ a_{3m} + b_{3m}'\left(I - e^{-\mathcal{K}/2}\right)\mu \\ a_{3m} + b_{3m}'\left(I - e^{-\mathcal{K}}\right)\mu \\ a_{3m} + b_{3m}'(I - W_{T_1,T_2})\mu \end{bmatrix} + \begin{bmatrix} b_{\tau_1}' & 0 & 0 \\ \vdots & 0 & 0 \\ b_{\tau_M}' & 0 & 0 \\ b_{3m}'e^{-\mathcal{K}/2} & 1 & 0 \\ b_{3m}'e^{-\mathcal{K}} & 0 & 1 \\ b_{3m}'W_{T_1,T_2} & 0 & 0 \end{bmatrix} z_t + \begin{bmatrix} \eta_{t,\tau_1} \\ \vdots \\ \eta_{t,\tau_M} \\ 0 \\ 0 \\ e_{t,\mathcal{L}} \end{bmatrix}, \tag{27}$$

and

$$z_t = \begin{bmatrix} (I - e^{-\mathcal{K}h})\mu \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} e^{-\mathcal{K}h} & 0 & 0 \\ 0 & \phi_{6m} & 0 \\ 0 & 0 & \phi_{12m} \end{bmatrix} z_{t-h} + \begin{bmatrix} \epsilon_t \\ \varepsilon_{t,6m} \\ \varepsilon_{t,12m} \end{bmatrix}, \tag{28}$$

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

In estimating the model, survey measurement error parameters such as $\tilde{\omega}_{6m}$, $\tilde{\omega}_{12m}$, and $\tilde{\omega}_{\mathcal{L}}$ can be either fixed according to some a priori judgments or estimated as free parameters. Treating the measurement error variances as free parameters is attractive, as 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, to rely heavily on them if they closely match the expectations of interest rates implied by the model. Note that the conventional estimation (without survey data) corresponds to fixing the measurement error variances at a very large number $(\tilde{\omega}_{6m}, \tilde{\omega}_{12m}, \tilde{\omega}_{\mathcal{L}} = \infty)$. One could also fix the variance of the measurement errors at values higher than suggested by estimating them as free parameters, if one wishes to be more conservative in incorporating the information from the surveys in the estimation than suggested by an unconstrained fitting of the model.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

For the "baseline" survey-based estimation results presented in the next section, we shall use the serially and cross-sectionally correlated survey measurement error version (observation and state equations given by equations (27) and (28)). This version is conceptually better justified and is a "gentler" way of introducing survey data than the i.i.d. survey measurement error version ($\phi_{6m} = \phi_{12m} = \tilde{\omega}_c = 0$; equivalently, using observation and state equations given by equations (16) and (20)).[21] We let the parameters of the 6- and 12-month survey forecast measurement errors be freely estimated, while we fix the $\tilde{\omega}_{\mathcal{L}}$ (size of long-horizon survey measurement error) at a sizeable number, 0.75%, to downplay the role of the long-horizon survey forecast.[22]

We have also experimented with alternative choices regarding the incorporation of survey data, including the use of the i.i.d. measurement error version (equations (16) and (20)), the practice of not fixing $\tilde{\omega}_{\mathcal{L}}$ but letting it be estimated, and so on. The results from these implementations will also be briefly mentioned in the next section.

IV. Empirical Results

A. Effects of Survey Data on the Estimation

Our baseline parameter estimate from the estimation with 1990–2003 term structure data and survey forecast data is presented in Table 1 (henceforth we refer to this case as the $\mathcal{S}90$ estimation). The parameter estimate from the estimation with the same term structure data but without survey data ($\mathcal{NS}90$ estimation) is also provided for comparison.

Because our model is a flexibly specified latent factor model, the individual parameters from the 2 estimations cannot be easily compared.[23] However, the comparison of certain parameters is still meaningful, such as the smallest eigenvalue of the $\mathcal{K}$ matrix, which can be denoted min(eig($\mathcal{K}$)). Note that min(eig($\mathcal{K}$)) = 0.052 for the $\mathcal{S}90$ estimation is substantially smaller than that of the $\mathcal{NS}90$ estimation (min(eig($\mathcal{K}$)) = 0.264).[24] This proves important for the long-horizon forecasts implied by the model, which we discuss later on. Another important effect of the use of survey data in the estimation is that it tends to make the parameter estimates more precise (i.e., make standard errors smaller). For example, the $\mathcal{NS}90$ estimation has 3 of the $\Sigma\Lambda_b$ matrix elements, which have standard errors 2 times (or more) larger than the estimates, while the $\mathcal{S}90$ estimation has

[21] Note from equation (28) we are in effect introducing additional latent factors in the case of serially correlated errors.

[22] Our rationale is that the long-horizon survey forecast may be less reliable than the short-horizon forecasts (due to a greater uncertainty); some might even question whether it can be viewed as "pure data" on the same footing as the yield data. Choosing a sizeable value beforehand represents a compromise between not using the data at all and treating the data as a faithful counterpart to the theory.

[23] Even though our normalization restrictions discussed in Section II.A give an identified model with a nonsingular Hessian, there are discrete symmetries (e.g., exchanging x_{1t} and x_{2t}) that create equivalent (but different looking) parameter sets. Thus, there is little point in comparing the values of a given parameter, such as $\mathcal{K}_{11}$, across models.

[24] Note that min(eig($\mathcal{K}$)) controls the degree of persistence of the process; in discrete time models, this concept is equivalent to the largest unit root.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

TABLE 1

Parameter Estimates and Standard Errors for Short-Sample Estimation

Table 1 reports the parameter estimates for short-sample estimation with survey data ($\mathcal{S}90$) and short-sample estimation without survey data ($\mathcal{NS}90$). Standard errors are given in parentheses.

	With Survey	Without Survey
$\mathcal{K}_{11}$	0.0517 (0.0944)	1.1097 (0.5983)
$\mathcal{K}_{21}$	−0.1919 (0.1300)	4.7981 (1.5640)
$\mathcal{K}_{31}$	0.3768 (1.7098)	3.2445 (1.4457)
$\mathcal{K}_{22}$	0.4639 (0.0857)	0.2642 (0.3098)
$\mathcal{K}_{32}$	−5.6679 (1.6666)	−0.1965 (0.3070)
$\mathcal{K}_{33}$	2.1807 (0.3906)	0.4477 (0.2497)
Σ_{11}	0.0022 (0.0000)	0.0028 (0.0007)
Σ_{22}	0.0023 (0.0005)	0.0047 (0.0006)
Σ_{33}	0.0061 (0.0003)	0.0043 (0.0002)
ρ_o	0.0483 (0.0126)	0.0442 (0.0197)
λ_{a1}	−0.7273 (0.3264)	−0.7191 (0.5963)
λ_{a2}	−0.7194 (1.1324)	−1.8202 (0.4374)
λ_{a3}	−1.7799 (1.5729)	−1.5834 (0.3688)
$[\Sigma\Lambda_b]_{11}$	−0.0921 (0.2879)	0.5916 (0.4703)
$[\Sigma\Lambda_b]_{21}$	−0.2511 (0.2562)	−0.4095 (0.8150)
$[\Sigma\Lambda_b]_{31}$	−1.3363 (0.7059)	0.1080 (0.7666)
$[\Sigma\Lambda_b]_{12}$	−0.6977 (0.2642)	−0.1172 (0.1262)
$[\Sigma\Lambda_b]_{22}$	−0.9627 (0.1525)	0.0787 (0.1636)
$[\Sigma\Lambda_b]_{32}$	−0.9062 (0.9085)	0.3025 (0.1980)
$[\Sigma\Lambda_b]_{13}$	0.2650 (0.0492)	0.0903 (0.0764)
$[\Sigma\Lambda_b]_{23}$	0.3365 (0.1323)	−0.1965 (0.1861)
$[\Sigma\Lambda_b]_{33}$	0.3897 (0.3084)	−0.4593 (0.1991)
ω_{3m}	0.0010 (0.0000)	0.0010 (0.0000)
ω_{6m}	0.0003 (0.0000)	0.0002 (0.0000)
ω_{1y}	0.0005 (0.0000)	0.0005 (0.0000)
ω_{2y}	0.0000 (0.0001)	0.0000 (0.0002)
ω_{4y}	0.0003 (0.0000)	0.0003 (0.0000)
ω_{7y}	0.0000 (0.0001)	0.0000 (0.0003)
ω_{10y}	0.0005 (0.0000)	0.0005 (0.0000)
$\tilde{\omega}_{6m}$	0.0023 (0.0003)	
$\tilde{\omega}_{12m}$	0.0016 (0.0003)	
$\tilde{\omega}_c$	0.0020 (0.0003)	
ϕ_{6m}	0.9575 (0.0113)	
ϕ_{12m}	0.9721 (0.0075)	

only 1. As discussed previously, however, the individual parameters of latent factor models like ours are hard to interpret, therefore it is more revealing to compare the confidence intervals for economically meaningful quantities, such as interest rate expectations and term premia. As an example, in Figure 4 we show the width of the 95% confidence interval for the 2-year-ahead short-rate expectation (i.e., $\mathrm{E}_t(r_{t+2y})$) implied by both $\mathcal{NS}90$ and $\mathcal{S}90$ estimations, computed using the delta method. It can be seen that the confidence interval for $\mathcal{S}90$ estimation is fairly narrow, while that for $\mathcal{NS}90$ estimation is quite wide.

The increased precision suggests that the likelihood function for the estimation with survey data is less "shallow" than the standard likelihood function, and hence it is easier to find its maximum. Furthermore, should multiple inequivalent local maxima (of comparable likelihood value) occur in the standard estimation, the additional information in survey data can help select the "correct" maximum. 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.

Let us now look more closely at the ability of the $\mathcal{S}90$ estimation to capture the information in survey forecast data and yields data. As shown in Graph A of Figure 5, the $\mathcal{S}90$ estimation fits the survey forecast of the 3-month yield

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

reasonably well. The regression of the survey-based 6-month-horizon expected change in the 3-month yield, $\mathrm{E}_t^{\mathrm{svy}}(y_{t+6m,3m}) - y_{t,3m}$, on the model-implied (fitted) 6-month-horizon expected change in the 3-month yield, $\hat{\mathrm{E}}_t(y_{t+6m,3m}) - y_{t,3m}$, gives

$$\begin{aligned} \mathrm{E}_t^{\mathrm{svy}}(y_{t+6m,3m}) - y_{t,3m} &= 0.00_{(0.01)} + 1.06_{(0.06)}(\hat{\mathrm{E}}_t(y_{t+6m,3m}) - y_{t,3m}) + \epsilon_t, \\ R^2 &= 0.67. \end{aligned}$$

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

FIGURE 4

Width of the Confidence Interval

Figure 4 shows the width of the 95% confidence interval for the 2-year-ahead short-rate expectation implied by $\mathcal{S}90$ and $\mathcal{NS}90$ estimations.

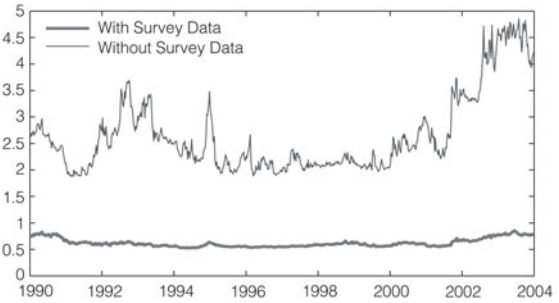
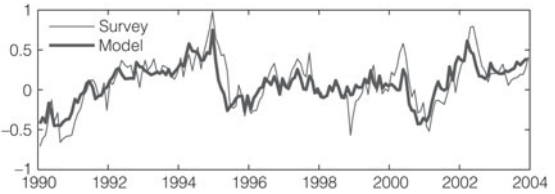


FIGURE 5

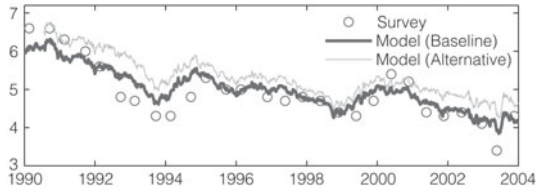
Comparison of the Model-Implied Expectation and Survey Forecast

Graph A of Figure 5 shows the survey forecast of the 6-month-horizon expected change in the 3-month yield and the forecast implied by the model. Graph B shows the survey forecast of the expected 3-month yield in the next 6–11 years and the model forecasts.

Graph A. 6-Month-Horizon Expected Change in the 3-Month Yield



Graph B. Expected 3-Month Yield over the 6- to 11-Year Horizon



https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

Graph B of Figure 5 plots the model-implied expectation of the 3-month yield over the 6- to 11-year horizon (long-horizon expectation) 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 1990–1994 period. We have also performed an estimation in which $e_{t,\mathcal{L}}$ is treated as a free variable to be estimated; in this case the fit of the 6- to 11-year survey forecasts is further improved, and the estimated standard deviation of $e_{t,\mathcal{L}}$ is less than 30 bp. Graph B of Figure 5 also shows the model-implied expectation of a 6- to 11-year horizon expectation based on an estimation that uses only the $6m$- and $12m$-ahead survey forecast (thin solid line). As can be seen, even without using long-horizon survey forecasts in the estimation, the model can produce a long-horizon expectation that closely tracks the variation of long-horizon survey forecasts.

Note that the "good fit" of survey forecasts is not achieved at the expense of the fit of yields data. As can be seen in Table 1, the yield fitting errors ($\omega_{3m}, \omega_{6m}, \omega_{1y}, \ldots$) are about the same for the estimations with and without survey forecasts (e.g., the 10-year yield fitting error size (ω_{10y}) is 5 bp for both $\mathcal{S}90$ and $\mathcal{NS}90$ estimations). That the survey forecasts and the term structure model are largely consistent is an encouraging and nontrivial result. Our estimation procedure does not guarantee that the model can match the survey data well; a misspecified 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.

Figure 6 presents some comparison of the implications of the estimated models. It is of interest to compare the effects of estimation with long sample (without survey data) and the estimation with short sample and survey data. Thus, we also include results based on estimation with yields data over 1965–2003, but without survey data. We refer to this as the $\mathcal{NS}65$ estimation.

The expected short-rate paths for Dec. 31, 2003, based on the $\mathcal{S}90$, $\mathcal{NS}90$, and $\mathcal{NS}65$ estimations are shown in Graph A of Figure 6. As can be seen, the $\mathcal{S}90$ estimation gives a more reasonable, faster-rising expected path of the short rate for that date (compared to the $\mathcal{NS}90$). Interestingly, the longer-sample conventional estimation ($\mathcal{NS}65$) gives an expected path that is closer to the $\mathcal{S}90$ estimation.

As shown in Graph B of Figure 6, the time series of the 10-year-ahead short-rate expectation based on the $\mathcal{S}90$ estimation shows a visible downward trend and has a greater variability, compared to the $\mathcal{NS}90$ estimation. This qualitative difference is due to the fact that the smallest eigenvalue of the $\mathcal{K}$ matrix estimated without survey forecast data ($\mathcal{NS}90$ estimation) is fairly large, hence it implies that expectations converge fast to a long-run mean (fast mean reversion), hence the 10-year-ahead expectations do not move a lot. The long-sample conventional estimation ($\mathcal{NS}65$) also displays downward-trending 10-year-ahead expectations, though not as strong as that of the $\mathcal{S}90$ estimation.

The 2-year forward term premia $\varphi_{t,2y}$ based on the 3 estimations are shown in Graph C of Figure 6. The $\mathcal{S}90$ result is on average smaller than the $\mathcal{NS}90$ result, and comes close to 0 (i.e., gets close to the EH) quite often. Recall that

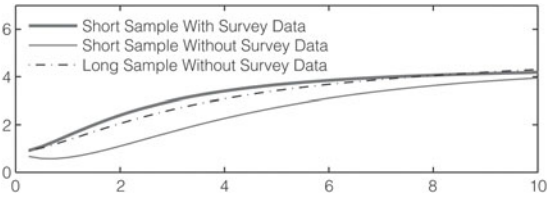
https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

FIGURE 6

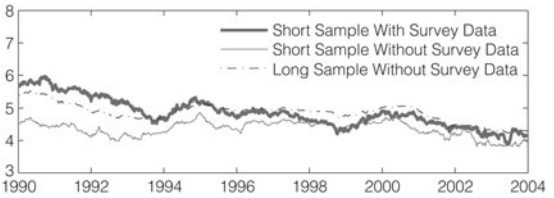
Model-Implied Short-Rate Expectation and Term Premia

Graph A of Figure 6 shows the expected paths of the short rate on Dec. 31, 2003, based on estimations with and without survey data ($\mathcal{S}$ and $\mathcal{NS}$) (horizon in years is shown on the *x*-axis). Graph B shows the 10-year-ahead expectation of the short rate. Graph C shows the 2-year forward term premium for the $\mathcal{S}90$ estimation (thick solid line), $\mathcal{NS}90$ estimation (thin solid line), and $\mathcal{NS}65$ estimation (thin dot-dashed line).

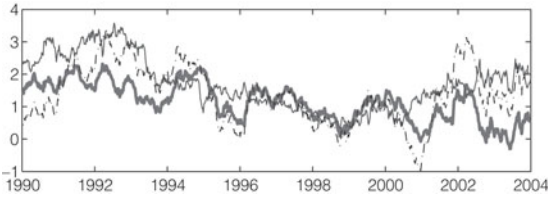
Graph A. Expected Short-Rate Path on Dec. 31, 2003



Graph B. 10-Year-Horizon Short-Rate Expectations



Graph C. 2-Year Forward Term Premia



the forward term premium is the amount to be subtracted from the forward rate (or futures rate) to obtain the expected short rate. Because the federal funds rate indeed declined or stayed low during these periods, the term premium estimate based on the $\mathcal{NS}90$ estimation implies a smaller in-sample forecast error than the $\mathcal{S}90$ estimation. However, this "better" forecast performance is likely an artifact of the statistical procedure (look-ahead bias) that becomes severe in small samples.[25] Note also that the $\varphi_{t,2y}$ estimate based on the $\mathcal{NS}90$ estimation is quite large for 2001–2003; recent research, on the other hand, has noted that the futures curve (without a term premium correction) is a good indicator of the near-term short-rate expectations (e.g., Gurkaynak, Sack, and Swanson (2007)), thus most market participants would be skeptical about the large 2-year-horizon term premium correction for 2001–2003 suggested by the $\mathcal{NS}90$ estimation. Interestingly, the 2-year forward premium implied by the long-sample conventional estimation

[25] As discussed in Section III.A, the historical evidence suggests that a significant part of the decline in the short rate has been unanticipated by the market and policymakers. Thus, most practitioners would be skeptical about in-sample forecasts that imply the contrary.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

($\mathcal{NS}65$) shows many features that are similar to those of $\mathcal{S}90$. For example, both $\mathcal{NS}65$ and $\mathcal{S}90$ estimations indicate that the 2-year term premium fell in 1995 and 2000 and rose in 2001.

The similarity between the implications of the $\mathcal{S}90$ and $\mathcal{NS}65$ estimations is an encouraging sign: The use of additional data in the form of longer yield sample or other data-like survey forecasts can help alleviate the problems in the conventional estimation with a short sample. It bears noting, however, that the use of survey forecast data and the use of a longer sample do not always lead to largely similar answers. For example, the $\mathcal{NS}65$ model-implied values of the standard deviation of a 1-week change in the 1-year yield is 26 bp, which is much larger than the sample standard deviation of a 1-week change in the 1-year yield in the 1990–2003 sample, which is 11 bp. On the other hand, both the $\mathcal{NS}90$ and $\mathcal{S}90$ estimations produce model-implied values of 11 bp, in agreement with sample standard deviation. Thus it can be seen that while the $\mathcal{NS}65$ estimation produces more sensible term premium implications than the $\mathcal{NS}90$ estimation, it comes at the cost of the volatility fit.

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 Jan. 2004 to Feb. 2005 (realized between July 2004 and Aug. 2005) were 0.69% and 0.46% for the $\mathcal{NS}90$ and $\mathcal{S}90$ models, respectively. The conventional short-sample estimation performed worse. However, our out-of-sample window is quite short, so care is required to avoid reading too much into this. In fact, 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 forecast error. For 6-month-ahead forecasts of the 3-month yield in a 4.5-year period based on the $\mathcal{S}$ parameter estimate, we find that there is a substantial chance (about 20%) of the ratio being larger than 1 (i.e., of the true RMSE 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 $\mathcal{NS}90$ and $\mathcal{S}90$ estimates is of economic significance. For example, in 2002, the 2-year forward term premium estimates based on $\mathcal{S}$ and $\mathcal{NS}$ 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 some remarks about the output from some alternative implementations. We find that if the i.i.d. survey measurement error version is used, the results are largely similar to our baseline results from the serially and cross-sectionally correlated error version ($\mathcal{S}90$ estimation); this should be of interest to practitioners, as the i.i.d. version is simpler to implement. Our results also do not depend on the fine-tuning of the long-horizon survey measurement

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

error size $\tilde{\omega}_{\mathcal{L}}$. Estimation with an even shorter sample, 1995–2003, produced similar results. In addition, we have undertaken estimation with the Survey of Professional Forecasters (SPF) survey forecast of the 3-month T-bill rate (instead of the BCFF survey), and obtained similar results. These findings speak well for the robustness of the survey-based estimation results presented in this paper.

B. Monte Carlo Evidence

In this section, we perform a small Monte Carlo experiment to examine the severity of the small-sample problem in our application, and to assess the extent to which the use of survey forecast information alleviates the problem. The Monte Carlo setting is particularly useful because it allows the examination of the small-sample problem in isolation (i.e., not confounded by potential misspecification issues).

We generate 50 artificial samples of term structure data and artificial survey forecasts of length equal to our 1990–2003 sample from our estimated $\mathcal{S}90$ 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. In order to compare the use of a longer-yield sample with the use of survey data, we also perform conventional estimations with samples that are 3 times as long. Specifically, for the long-sample estimations, we generate simulated data with length of 2,100 weeks (40.4 years), which is about the length of the sample in the $\mathcal{NS}65$ estimations. For the short-sample estimations, we use only the first 700 weeks (13.5 years) of the simulated long samples. We note that in setting the experiment in this manner, we enforce homogeneity in the data and rule out the possibility of structural change. In this way, we present the best possible scenario for using a long sample without survey data.

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. In addition, we add measurement errors with characteristics similar to the fitting errors in Section IV.A. 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 fitted errors $[\hat{\eta}_{t,3m}, \ldots, \hat{\eta}_{t,10y}, \hat{e}_{t,6m}, \hat{e}_{t,12m}]$ based on the $\mathcal{S}$ estimate and the data in Section IV.A. The measurement error $e_{t,\mathcal{L}}$ is drawn from an i.i.d. 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 Graph B of Figure 5).

Table 2 displays the median (50th), 25th, and 75th percentile values of the model parameters from the short-sample estimation with (artificial) survey data and short-sample conventional estimation and long-sample conventional estimation. There is a notable upward bias in the mean-version parameters ($\mathcal{K}$) in the short estimation without the survey data: True value of the smallest eigenvalue of $\mathcal{K}$ is 0.0517, but the median of the conventional short-sample estimation is 0.1539. The median of the conventional long-sample estimation (0.0948) is closer to the true value, but the short-sample estimation with survey data does even better (median of 0.0715). Besides the reduction of the bias, there is also a decrease in

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

TABLE 2

Monte Carlo Results (Parameters)

Table 2 reports the results of Monte Carlo experiments with artificially generated yields and survey data. The median (50th percentile), 25th percentile, and 75th percentile values of the estimated parameters are shown for short-sample estimations with survey data, short-sample estimations without survey data, and long-sample estimations without survey data.

		With Survey (short)			Without Survey (short)			Without Survey (long)		
	True	Med.	25%	75%	Med.	25%	75%	Med.	25%	75%
$\mathcal{K}_{11}$	0.0517	0.0715	0.0560	0.1029	0.1539	0.0552	0.4137	0.0948	0.0450	0.1858
$\mathcal{K}_{21}$	−0.1919	−0.2243	−0.2786	−0.1766	−0.2336	−0.4733	−0.0825	−0.2465	−0.3394	−0.1435
$\mathcal{K}_{31}$	0.3768	0.5324	0.3521	0.7724	0.4406	−0.0921	0.9305	0.3729	0.2086	0.7704
$\mathcal{K}_{22}$	0.4639	0.4698	0.4071	0.5366	0.7356	0.5189	1.0509	0.5148	0.4357	0.6808
$\mathcal{K}_{32}$	−5.6679	−5.7365	−6.1758	−5.5800	−5.7320	−6.2875	−5.4721	−5.7470	−6.0246	−5.5021
$\mathcal{K}_{33}$	2.1807	2.2131	2.1059	2.5703	2.4715	2.0952	3.0316	2.3093	2.1687	2.5318
Σ_{11}	0.0022	0.0021	0.0020	0.0023	0.0021	0.0019	0.0023	0.0020	0.0019	0.0022
Σ_{22}	0.0023	0.0023	0.0022	0.0024	0.0024	0.0022	0.0025	0.0024	0.0023	0.0024
Σ_{33}	0.0061	0.0062	0.0061	0.0064	0.0062	0.0060	0.0063	0.0062	0.0060	0.0063
ρ_o	0.0483	0.0476	0.0443	0.0501	0.0465	0.0423	0.0522	0.0448	0.0406	0.0486
λ_{a1}	−0.7273	−0.7240	−0.8035	−0.6609	−0.7345	−0.8611	−0.5992	−0.6885	−0.7644	−0.5927
λ_{a2}	−0.7194	−0.6779	−0.9709	−0.4209	−0.7300	−1.2682	−0.4317	−0.6320	−0.8545	−0.4460
λ_{a3}	−1.7799	−1.7990	−2.0020	−1.6184	−1.8666	−2.3262	−1.3752	−1.6322	−1.9724	−1.2788
$[\Sigma\Lambda_b]_{11}$	−0.0921	−0.1050	−0.1824	−0.0817	−0.2402	−0.4681	−0.1037	−0.1426	−0.2575	−0.0672
$[\Sigma\Lambda_b]_{21}$	−0.2511	−0.2258	−0.3079	−0.1372	−0.2466	−0.4922	0.0176	−0.2436	−0.3375	−0.1670
$[\Sigma\Lambda_b]_{31}$	−1.3363	−1.4600	−1.8677	−1.1766	−1.6103	−1.9331	−0.9726	−1.4536	−1.6686	−1.2740
$[\Sigma\Lambda_b]_{12}$	−0.6977	−0.6934	−0.7170	−0.5654	−0.6586	−0.7557	−0.4904	−0.6213	−0.7042	−0.5029
$[\Sigma\Lambda_b]_{22}$	−0.9627	−0.9637	−1.0557	−0.9154	−1.2769	−1.5635	−1.0242	−1.0746	−1.2600	−0.9919
$[\Sigma\Lambda_b]_{32}$	−0.9062	−0.9226	−1.1360	−0.4054	−0.8194	−1.0922	0.1606	−0.8365	−1.1577	−0.2405
$[\Sigma\Lambda_b]_{13}$	0.2650	0.2610	0.2186	0.2810	0.2644	0.2249	0.3121	0.2450	0.2060	0.2735
$[\Sigma\Lambda_b]_{23}$	0.3365	0.3405	0.3046	0.3895	0.3781	0.3118	0.4485	0.3655	0.3317	0.4045
$[\Sigma\Lambda_b]_{33}$	0.3897	0.3389	0.1306	0.4158	0.1747	−0.1740	0.4392	0.3267	0.0984	0.4916
ω_{3m}	0.0009	0.0009	0.0009	0.0010	0.0009	0.0009	0.0010	0.0009	0.0009	0.0010
ω_{6m}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ω_{1y}	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
ω_{2y}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ω_{4y}	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
ω_{7y}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ω_{10y}	0.0005	0.0005	0.0004	0.0005	0.0005	0.0004	0.0005	0.0005	0.0005	0.0005
$\tilde{\omega}_{6m}$	0.0018	0.0017	0.0016	0.0018						
$\tilde{\omega}_{12m}$	0.0019	0.0018	0.0017	0.0020						
$\tilde{\omega}_c$	0.0016	0.0015	0.0014	0.0017						
ϕ_{6m}	0.9205	0.9157	0.9045	0.9225						
ϕ_{12m}	0.9554	0.9531	0.9452	0.9603						

the dispersion (standard deviations) of the estimated parameters in the estimation with survey data (i.e., the survey data improve the precision of the estimates). For example, the interval between the 25th and 75th percentile values for $\mathcal{K}_{11}$ is 0.0469 $(= 0.1029 - 0.0560)$ for $\mathcal{S}90$, while the corresponding value for $\mathcal{NS}90$ is 0.3585 $(= 0.4137 - 0.0552)$, a much larger number.

As noted earlier, in a flexible 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. To this end, we examine the following quantities:

$$\Phi_w^a = \left(\frac{1}{T}\right)\int_0^T dt|\hat{\varphi}_{t,w} - \varphi_{t,w}|, \tag{29}$$

$$\Phi_w^b = \frac{\int_0^T dt(\hat{\varphi}_{t,w} - \overline{\hat{\varphi}_{t,w}})^2}{\int_0^T dt(\varphi_{t,w} - \overline{\varphi_{t,w}})^2}, \tag{30}$$

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

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

$$\Gamma_w = \frac{\int_0^T dt\left(\hat{\mathrm{E}}_t(r_{t+w}) - \overline{\hat{\mathrm{E}}_t(r_{t+w})}\right)^2}{\int_0^T dt\left(\mathrm{E}_t(r_{t+w}) - \overline{\mathrm{E}_t(r_{t+w})}\right)^2}, \tag{31}$$

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

$$\mathcal{T}_w = \frac{\int_0^T dt(y_{t+w,3m} - \hat{\mathrm{E}}_t(y_{t+w,3m}))^2}{\int_0^T dt(y_{t+w,3m} - \mathrm{E}_t(y_{t+w,3m}))^2}, \tag{32}$$

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 3, the deviation of the estimated term premium from the true values is quite substantial (Φ_w^as are large) in the case of the short-sample 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_w^a values and pulling the Φ_w^b, Γ_w numbers toward 1.

TABLE 3

Monte Carlo Results (Term Premia and Expected Short Rates)

Table 3 reports the median, 25th percentile, and 75th percentile values of Φ_w^a, Φ_w^b, Γ_w (for $w = 2y$, $10y$), and $\mathcal{T}_w$ (for $w = 6m$, $1y$) from short-sample estimations with/without survey data and long-sample estimations without survey data. Here, Φ_w^a measures how close the estimated term premium is to the true term premium and is 0 if they are identical; Φ_w^b measures the relative variability of the estimated term premium (with respect to the variability of the true term premium); Γ_w similarly measures the relative variability of the estimated short-rate expectation; and $\mathcal{T}_w$ compares the estimated RMSE forecast error with actual RMSE. When there is no bias, Φ_w^b, Γ_w, and $\mathcal{T}_w$ are 1.

	With Survey (short)			Without Survey (short)			Without Survey (long)		
	Med.	25%	75%	Med.	25%	75%	Med.	25%	75%
Φ_{2y}^a	0.0734	0.0439	0.0944	0.2317	0.1839	0.3311	0.1886	0.1211	0.2856
Φ_{10y}^a	0.1129	0.0560	0.1600	0.4141	0.3075	0.6243	0.3903	0.1813	0.5877
Φ_{2y}^b	1.0143	0.9719	1.0824	1.1583	1.0098	1.3801	1.0778	0.9917	1.1917
Φ_{10y}^b	1.0961	0.9950	1.1867	1.4561	1.0349	1.5526	1.2074	1.0803	1.4109
Γ_{2y}	0.9965	0.9196	1.0369	0.8570	0.5417	1.0380	0.9439	0.8208	1.0252
Γ_{10y}	0.8510	0.7304	0.9618	0.2325	0.0436	0.9365	0.6558	0.2401	0.9190
$\mathcal{T}_{6m}$	1.0019	0.9928	1.0105	0.9715	0.9347	0.9992	0.9857	0.9746	1.0048
$\mathcal{T}_{12m}$	1.0011	0.9888	1.0211	0.9551	0.8849	0.9975	0.9769	0.9455	1.0044

The use of a longer sample (without survey) is also seen to help, though not as much as the use of survey forecast data. For example, the median value of the Φ_{10y}^b based on the conventional short-sample estimation is 1.46 (i.e., the conventional short sample leads to 10-year term premium variance that typically

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

is 1.46 times larger than the true value). This bias is reduced to 1.21 in the case of the long-sample conventional estimation, while the short-sample estimation with survey data gives an even smaller bias (1.10).

It is also interesting that in the estimations 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 forecast errors). 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.

Overall, our results suggest that in our application, incorporating the information from a survey forecast is more effective than using a sample that is 3 times as long, even if we assume that the long sample is free of structural change. 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 parameters for Monte Carlo simulation (the $\mathcal{S}90$ estimate) 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 a key feature of the data, and 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 point to a more general cautionary tale: If a model exhibits a substantial degree of persistence, the conventional estimation can be suspect.

C. 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 EH reflected in Campbell-Shiller (1991) regressions, and whether expectations of long-term yields implied by our model compare with survey-forecasts of comparable-maturity yields.

Recall that in the Campbell-Shiller (1991) regression

$$y_{t+w,\tau-w} - y_{t,\tau} = \alpha + \beta \frac{w}{\tau - w}(y_{t,\tau} - y_{t,w}) + \xi_{t+w}, \tag{33}$$

the EH corresponds to $\beta = 1$. Campbell and Shiller 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 EH is weaker. We therefore reexamine the Campbell-Shiller regression for our sample and look at how well our model matches the resulting pattern of β.

Panel A of Table 4 presents estimates of the Campbell-Shiller (1991) regression for several τ with $w = 3$ months, using the T-bill yield ($y^{\text{Tbill}}_{t,3m}$), and with $w = 6$ months, using both the T-bill yield ($y^{\text{Tbill}}_{t,6m}$) and coupon-securities-based

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

TABLE 4

Beta Coefficients from the Expectations Hypothesis Regression

Panel A of Table 4 reports the β coefficients from the 1990–2003 sample, with Newey-West (1987) standard errors in parentheses. Panel B reports the implied β coefficients from the $\mathcal{S}90$ estimate.

		$\tau = 2y$	$\tau = 3y$	$\tau = 5y$	$\tau = 10y$
Panel A. 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)
Panel B. Model					
$w = 3m$		−1.06 (0.76)	−1.32 (1.01)	−1.69 (1.45)	−2.58 (2.56)
$w = 6m$		−0.89 (0.73)	−1.09 (1.06)	−1.41 (1.40)	−2.23 (2.51)

yield ($y^{\text{coupon}}_{t,6m}$).[26] The recent sample still displays the classic pattern of the departure from the EH, in that β is negative and more so for longer maturities, but the point estimates are less negative than Campbell and Shiller, and the standard errors are large so that $\beta = 1$ (EH) 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 (1991) regression, we simulated 500 samples with the length of 750 weeks (which approximately corresponds to the length of the true sample) and ran regression (33) with the simulated data. Panel B of Table 4 gives the median and standard deviation of the βs. The β estimates do deviate from 1 and increasingly so with larger τ; the median βs imply somewhat greater departure from the EH than the true sample point estimate. However, the dispersions (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.

In addition to the survey forecasts of the short-term interest rate (3-month T-bill) that we use in estimating our model, BCFF 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 and 10 years. It is of interest to compare these survey forecasts of long-term yields with those based on the EH 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 EH. More specifically, in the regression

$$\text{(34)} \qquad \mathrm{E}^{\text{svy}}_t(y_{t+w,\tau}) - y_{t,w} \quad = \quad \alpha + \beta\,(f_{t,w,\tau} - y_{t,w}) + \epsilon_{t+w}$$

(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.

[26] 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 EH.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

We have also looked at comparable specifications of regression (34) with the survey forecasts of the TCM yields in our sample and obtained similar results when τ is large and w is small. However, equation (34) has relatively little power for tests relating to the EH. Even if the survey expectations were inconsistent with the EH, equation (34) with large τ and small w would still give β close to 1 and high R^2 due to the persistence of term spreads.[27]

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 equation (34)). In performing this regression, we also 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 is 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

$$\mathcal{O}_{t,\tau} = Y_{t,\tau} - Y_{t,\tau}^{\text{TCM}}, \tag{35}$$

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

$$\delta_w^{\text{svy}} Y_{t,\tau} = \text{E}_t^{\text{svy}}(Y_{t+w,\tau}^{\text{TCM}}) - Y_{t,\tau}^{\text{TCM}} + \delta_w \mathcal{O}_{t,\tau}, \tag{36}$$

where $\delta_w \mathcal{O}_{t,\tau}$ is the expected w-horizon change in $\mathcal{O}_{t,\tau}$.[28] Substituting the survey forecast of the TCM yield in equation (36), we obtain the "survey-based" forecast of the changes in the (off-the-run) par yield, $\delta_w^{\text{svy}} Y_{t,\tau}$. The corresponding EH-based expected change in the (off-the-run) par yield is

$$\delta_w^{\text{EH}} Y_{t,\tau} = F_{t,w,\tau} - Y_{t,\tau}, \tag{37}$$

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

The results for the 5-year par yield forecasts regression

$$\delta_w^{\text{svy}} Y_{t,5y} = \alpha + \beta\, \delta_w^{\text{EH}} Y_{t,5y} + \xi_t \tag{38}$$

for horizons of 6 and 12 months ($w = 6m, 12m$) are shown in Table 5. The low value of R^2 in the case of the 6-month horizon ($R^2 = 0.05$) suggests that the EH 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.

[27] To elaborate, consider the short-horizon forecast of $y_{t,\tau}$. The forecast $\text{E}_t(y_{t+w,\tau})$ would differ from $y_{t,\tau}$ only by a small amount, that is, $\text{E}_t(y_{t+w,\tau}) = y_{t,\tau} + \delta_t$, where δ_t is a small number that goes to 0 as w goes to 0. 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 $y_{t,\tau} - y_{t,w}$.

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

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

TABLE 5

Coefficients from the Regression of Survey Forecasts on Model-Implied Forecasts

In Table 5, ND (no differencing) refers to equations (38) and (40), and D (3-month differencing) refers to equations (39) and (41). Numbers in parentheses represent standard errors.

		α	β	R^2
Panel A. EH				
ND	6*m*	−0.01 (0.03)	0.36 (0.13)	0.05
	12*m*	−0.13 (0.04)	0.83 (0.08)	0.38
D	6*m*	0.03 (0.02)	−1.53 (0.29)	0.15
	12*m*	0.03 (0.02)	−0.51 (0.19)	0.04
Panel B. Model				
ND	6*m*	0.04 (0.02)	0.31 (0.06)	0.11
	12*m*	0.09 (0.02)	0.71 (0.06)	0.49
D	6*m*	0.02 (0.02)	0.30 (0.09)	0.06
	12*m*	0.02 (0.02)	0.40 (0.07)	0.18

It can be seen from Graph A of Figure 7 that the survey-based 6-month-horizon expected change in the 5-year par yield ($\delta_{6m}^{\text{svy}} Y_{t,5y}$) has substantial short-run movements, while the EH-based expected change $\delta_{6m}^{\text{EH}} Y_{t,5y}$ 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 $\delta_{12m}^{\text{svy}} Y_{t,5y}$ tracks $\delta_{12m}^{\text{EH}} Y_{t,5y}$ fairly well.

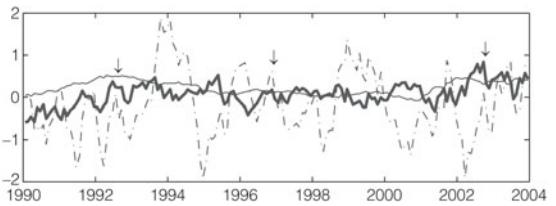
The inability of the EH to explain the short-run variations in the survey forecast can be highlighted by the regression of the 3-month difference in $\delta_{w}^{\text{svy}} Y_{t,5y}$

FIGURE 7

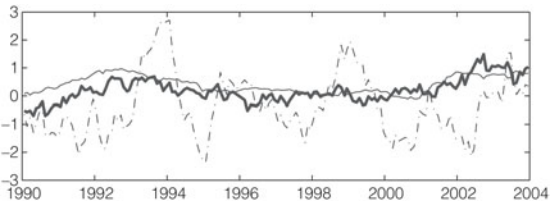
Comparison of Survey and Expectations Hypothesis

Graph A of Figure 7 shows 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. The dot-dashed line is the realized change. Graph B is for the 12-month horizon.

Graph A. 6-Month Horizon



Graph B. 12-Month Horizon



https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

($\Delta^{\text{svy}}_{3m,w} Y_{t,5y} \equiv \delta^{\text{svy}}_{w} Y_{t,5y} - \delta^{\text{svy}}_{w} Y_{t-3m,5y}$) onto the 3-month difference in $\delta^{\text{EH}}_{w} Y_{t,5y}$ ($\Delta^{\text{EH}}_{3m,w} Y_{t,5y} \equiv \delta^{\text{EH}}_{w} Y_{t,5y} - \delta^{\text{EH}}_{w} Y_{t-3m,5y}$) for $w = 6$ and 12 months:

$$\Delta^{\text{svy}}_{3m,w} Y_{t,5y} \quad = \quad \alpha + \beta \, \Delta^{\text{EH}}_{3m,w} Y_{t,5y} + \xi_t. \tag{39}$$

If the survey expectation were consistent with the EH, 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 $\delta^{\text{svy}}_{w} Y_{t,5y}$ 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 $\delta^{\text{svy}}_{w} Y_{t,5y}$ do correspond to the changes that were realized later. For example, in Graph A of Figure 7, $\delta^{\text{svy}}_{6m} Y_{t,5y}$ peaks around June 1992, Oct. 1996, and Aug. 2002; the realized values $Y_{t+3m,5y} - Y_{t,5y}$ also peak at those times.

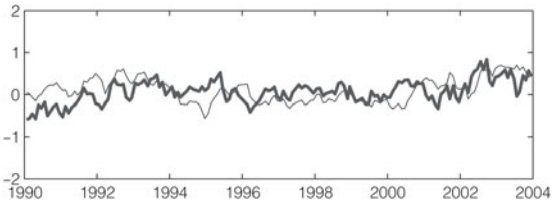
We also compared the survey-based expected change in long-term yield with the expected change implied by the model estimated with survey data on the forecast of short-term interest rates ($\mathcal{S}90$ estimation). Graph A of Figure 8 displays the model-implied 6-month-ahead expected change in the 5-year par yield, together with the survey-based expected change; the corresponding plot for 12-month-ahead expected change is shown in Graph B of Figure 8.[29] As can be seen, $\delta^{\text{model}}_{w} Y_{t,5y}$ captures many of the short-run movements in $\delta^{\text{svy}}_{w} Y_{t,5y}$. The visual agreement between the model and the survey-based forecast is particularly notable for the 12-month-ahead forecasted change, $\delta^{\text{model}}_{12m} Y_{t,5y}$.

FIGURE 8

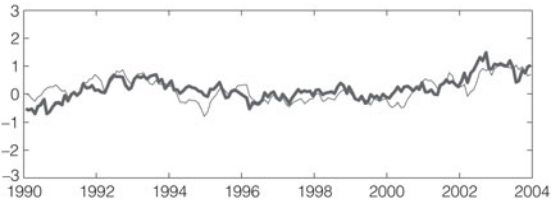
Comparison of Survey and Model

Graph A of Figure 8 shows the survey-based expected change in the 5-year par yield (thick solid line) and the model-based expected change (thin solid line) for the 6-month horizon. Graph B is for the 12-month horizon.

Graph A. 6-Month Horizon



Graph B. 12-Month Horizon



[29]Because the formula for the expectation of par yields is not available in a closed form, we have evaluated the expectation of par yields using simulation.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

The results from the regressions corresponding to equations (38) and (39), that is,

$$\delta_w^{\text{svy}} Y_{t,5y} = \alpha + \beta\, \delta_w^{\text{model}} Y_{t,5y} + \xi_t, \tag{40}$$

$$\Delta_{3m,w}^{\text{svy}} Y_{t,5y} = \alpha + \beta\, \Delta_{3m,w}^{\text{model}} Y_{t,5y} + \xi_t, \tag{41}$$

are given in Panel B of Table 5. The R^2s in regression (40) are larger than those in equation (38), and the sign of β in equation (41) is now significantly 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. 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). 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.

V. 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, despite considerable attention policymakers and market participants pay to them. In this paper we have shown 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 potential solution to the problem, we have explored incorporating 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. Monte Carlo evidence indicates that using survey forecast information can be more effective than using a substantially longer sample. In addition, our analysis highlights features of the survey forecasts that are of interest on their own.

Although in this paper we focus on the estimation of a Gaussian model of the nominal term structure, the technique of using the information in survey forecasts to help better identify the parameters pertaining to expectations and risk premia can be expected to be useful in other settings as well, such as macro-finance models that examine jointly the dynamics of interest rates and macroeconomic variables, models with stochastic volatility, and so on. This points to the exciting possibility for further fruitful research using surveys of market participants' forecasts.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

Appendix 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 are given by equation (4), with

$$a_\tau = -\frac{1}{\tau}\Big[(\mathcal{K}^*\mu^*)'(M_{1,\tau}-\tau I)\mathcal{K}^{*-1'}\rho + \frac{1}{2}\rho'\mathcal{K}^{*-1}(M_{2,\tau}-\Sigma\Sigma'M_{1,\tau}-M_{1,\tau}'\Sigma\Sigma'+\tau\Sigma\Sigma') \times \mathcal{K}^{*-1'}\rho-\tau\rho_o\Big], \tag{A-1}$$

$$b_\tau = \frac{1}{\tau}M_{1,\tau}\rho, \tag{A-2}$$

where

$$M_{1,\tau} = -\mathcal{K}^{*-1'}(e^{-\mathcal{K}^{*'}\tau}-I), \tag{A-3}$$

$$M_{2,\tau} = -\mathrm{vec}^{-1}((\mathcal{K}^*\otimes I)+(I\otimes\mathcal{K}^*))^{-1} \times \mathrm{vec}\,(e^{-\mathcal{K}^*\tau}\Sigma\Sigma'e^{-\mathcal{K}^{*'}\tau}-\Sigma\Sigma')). \tag{A-4}$$

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

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

$$Y_{t,\tau} = 2\frac{1-P_{t,\tau}}{\sum_{j=1}^{2\tau}P_{t,j/2}}, \tag{A-5}$$

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 equation (37)) is

$$F_{t,w,\tau} = 2\frac{P_{t,w}-P_{t,w+\tau}}{\sum_{j=1}^{2\tau}P_{t,w+j/2}}. \tag{A-6}$$

The (zero-coupon) forward rate $f_{t,w,\tau}$ (in equation (11)) is simply

$$f_{t,w,\tau} = (\log(P_{t,w})-\log(P_{t,\tau+w}))/\tau. \tag{A-7}$$

The instantaneous forward rate (in equation (8)) is

$$f_{t,w} = -\frac{d}{dw}\log(P_{t,w}). \tag{A-8}$$

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.[30] 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), for example, $\hat{\mathrm{E}}_t(y_{t+w,\tau})$ and $\hat{\varphi}_{t,w}$ in Section IV.

[30] In practice, the difference between the smoothed state variable $x_{t|T}$ and the filtered state variable $x_{t|t}$ is negligible.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

Appendix B. Notes on the Data

We use weekly (Wednesday) zero-coupon term structure data, with maturities of 3*m*, 6*m*, 1*y*, 2*y*, 4*y*, 7*y*, and 10*y* 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 data set 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 (3*m* and 6*m*). For the long-sample estimation (1965–2003), we use monthly (end of the month) data with maturities of 6*m*, 1*y*, 2*y*, 5*y*, and 10*y*; this data set is an update of the data used in Duffee (2002).

The monthly forecasts of 3-month T-bill yields are based on the BCFF. The BCFF is published on the 1st day of each month and presents forecasts from a survey conducted during 2 consecutive business days 1 to 2 weeks earlier.[31] 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 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 yields 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- 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.[32] 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- and 12-month-ahead forecasts from Dec. 17, 2003 as 1.31 and 1.97, respectively.

Twice a year, generally on the June 1 and Dec. 1 issues, the BCFF also provides a long-horizon forecast of the 3-month T-bill approximately 6–11 years out. The approximation is due to the fact that survey participants are asked to provide a forecast for a 5-calendar-year average. For example, in the June 1, 2002 issue, the survey participants were asked to forecast the 5-year average T-bill yield for 2008–2012, thus it is, more accurately, a 5½- to 10½-year horizon forecast, that is, $T_1 = 5½, T_2 = 10½$ in equation (16).

The TCM yields are available from the Federal Reserve Board's H.15 Web site (http://federalreserve.gov/releases/h15).

References

Ang, A., and M. Piazzesi. "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables." *Journal of Monetary Economics*, 50 (2003), 745–787.

Babbs, S. H., and K. B. Nowman. "Kalman Filtering of Generalized Vasicek Term Structure Models." *Journal of Financial and Quantitative Analysis*, 34 (1999), 115–130.

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

[32] 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).

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press

Ball, C. A., and W. N. Torous. "Unit Roots and Estimation of Interest Rate Dynamics." *Journal of Empirical Finance*, 3 (1996), 215–238.
Bekaert, G.; R. J. Hodrick; and D. A. Marshall. "On Biases in Tests of Expectations Hypothesis of the Term Structure of Interest Rates." *Journal of Financial Economics*, 44 (1997), 309–348.
Campbell, J. Y. "Some Lessons from the Yield Curve." *Journal of Economic Perspectives*, 9 (1995), 129–152.
Campbell, J. Y., and R. J. Shiller. "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *Review of Economic Studies*, 58 (1991), 495–514.
Chun, A. L. "Expectations, Bond Yields, and Monetary Policy." *Review of Financial Studies*, 24 (2011), 208–247.
Cogley, T., and T. J. Sargent. "Evolving Post-World War II U.S. Inflation Dynamics." In *NBER Macroeconomics Annual*, 16 (2001), 331–373.
Croushore, D. "Low Inflation: The Surprise of 1990s." *Federal Reserve Bank of Philadelphia Business Review*, July/Aug. (1998), 3–12.
Dai, Q., and K. J. Singleton. "Specification Analysis of Affine Term Structure Models." *Journal of Finance*, 55 (2000), 1943–1978.
Dai, Q., and K. J. Singleton. "Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure." *Journal of Financial Economics*, 63 (2002), 415–441.
de Jong, F. "Time Series and Cross-Section Information in Affine Term-Structure Models." *Journal of Business Economics and Statistics*, 18 (2000), 300–314.
Duarte, J. "Evaluating an Alternative Risk Preference in Affine Term Structure Models." *Review of Financial Studies*, 17 (2004), 379–404.
Duffee, G. R. "Idiosyncratic Variation of Treasury Bill Yields." *Journal of Finance*, 51 (1996), 527–551.
Duffee, G. R. "Term Premia and Interest Rate Forecasts in Affine Models." *Journal of Finance*, 57 (2002), 405–443.
Duffee, G. R., and R. H. Stanton. "Evidence on Simulation Inference for Near Unit-Root Processes with Implications for Term Structure Estimation." *Journal of Financial Econometrics*, 6 (2008), 108–142.
Friedman, B. M. "Survey Evidence on the 'Rationality' of Interest Rate Expectations." *Journal of Monetary Economics*, 6 (1980), 453–465.
Froot, K. A. "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates." *Journal of Finance*, 44 (1989), 283–305.
Gurkaynak, R. S.; B. P. Sack; and E. T. Swanson. "Market-Based Measures of Monetary Policy Expectations." *Journal of Business and Economic Statistics,* 25 (2007), 201–212.
Harvey, A. C. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge, UK: Cambridge University Press (1989).
Jagannathan, R.; A. Kaplin; and S. Sun. "An Evaluation of Multi-Factor CIR Models Using LIBOR, Swap Rates, Cap and Swaption Prices." *Journal of Econometrics*, 116 (2003), 113–146.
Kim, C.-J., and C. R. Nelson. "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 (1999), 608–616.
Kohn, D. "Forward-Looking Rules for Monetary Policy: Comment." In *Monetary Policy Rules*, J. B. Taylor, ed. Chicago, IL: University of Chicago Press (1999), 192–199.
Kozicki, S., and P. A. Tinsley. "Shifting Endpoints in the Term Structure of Interest Rates." *Journal of Monetary Economics*, 47 (2001), 613–652.
Merton, R. C. "On Estimating the Expected Return on the Market: An Exploratory Investigation." *Journal of Financial Economics*, 8 (1980), 323–361.
Newey, W. K., and K. D. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* 55 (1987), 703–708.
Pennacchi, G. G. "Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data." *Review of Financial Studies*, 4 (1991), 53–86.
Rudebusch, G., and T. Wu. "The Recent Shift in Term Structure Behavior from a No-Arbitrage Macro-Finance Perspective." Federal Reserve Bank of San Francisco Working Paper No. 2004-25 (2004).
Svensson, L. E. O. "Estimating Forward Interest Rates with the Extended Nelson and Siegel Method." *Quarterly Review, Sveriges Riksbank* (1995), 13–26.
Thomas, L. B. Jr. "Survey Measures of Expected U.S. Inflation." *Journal of Economic Perspectives*, 13 (1999), 125–144.

https://doi.org/10.1017/S0022109011000627 Published online by Cambridge University Press