Information in (and not in) interest rates surveys[*]

Laura Coroneo
University of York

Adam Golinski
University of York

26 January 2023

Abstract

We show that standard term structure models for observed interest rates fail to capture interest rate survey expectations. We therefore propose a joint term structure model for observed interest rates and interest rate surveys that allows for separate objective and subjective probability measures. Our results contradict the previous term structure literature and provide evidence that interest rate surveys do not help identify observed interest rate dynamics. Yet, despite this evidence against the rational expectation hypothesis, we find that surveys provide valuable information as a priced risk factor that is not spanned by observed interest rates.

Keywords: bond premia, asset pricing, risk premia, surveys

JEL Classification: D84, E4, G12

[*]We thank seminar participants at the Federal Reserve Board, the Banque de France, the University of Liverpool and the University of York for useful comments.

1 Introduction

Central banks, government and market participants closely monitor long term interest rates and routinely use term structure models to separate the information about investors' expectations of the economy from the one about the risks they perceive. A well-known drawback of term structure models is that they suffer from identification issues (see Ang and Piazzesi 2003, Hamilton and Wu 2012), and minor changes in the estimated coefficients can lead to a misestimation of the risk premium and a misperception of the market's outlook. One way to address these identification problems is to use survey expectations on short-term rates to aid in the identification of the physical parameters (Kim and Wright 2005, Kim and Orphanides 2012, d'Amico, Kim and Wei 2018), to proxy for expectations (Crump, Eusepi and Moench 2018) or to proxy for state variables (Chun 2011). These approaches rely on the rational expectation hypothesis, i.e. the assumption that the probability measure used by economic agents is the same as the statistical probability measure.

The empirical evidence on the rational expectation hypothesis, however, is rather weak, as there is increasing evidence of large and persistent errors in investors' expectations about the short-term interest rate (Cieslak 2018, Farmer, Nakamura and Steinsson 2021). This indicates a departure from rational expectations, for example because agents overestimate the persistence of the pricing factors (Piazzesi, Salomao and Schneider 2015) or are learning about the real-world parameters (Farmer et al. 2021).

Still, there can be a discrepancy between survey expectations and forecasts implied by observed long-term interest rates even under the rational expectation hypothesis. This can happen when observed interest rates do not contain all the information necessary to identify the drivers of interest rate expectations due to the presence of hidden factors. A hidden (or unspanned) factor arises when a state variable has offsetting effects on the expectation and the risk premium components. As a consequence, it is not possible to extrapolate it from observed interest rates, but any variable that depends separately on these two components (or weights them differently) will reveal such unspanned factors (see Duffee 2011).

In this paper we propose using survey expectations to directly extrapolate the information about any unspanned factors while accounting for possible deviations from rational expectations. This is motivated by three stylised facts about interest rate surveys that we document. First, yield forecasts implied by the term structure of observed interest rates are at odds with survey expectations. Second, interest rate surveys, apart from yield curve factors, are spanned by an additional, survey-specific factor. Third, the survey-specific factor Granger-causes the yield curve factors and drives the term premium.

To account for these stylised facts, we develop a joint term structure model for observed zero-coupon yields and survey expectations that allows for separate real-world and survey dynamics, and that also includes an additional state variable that is unspanned by observed interest rates but that drives surveys. In addition, we explicitly enforce a zero-lower bound on observed interest rates and interest rate surveys. We estimate the joint shadow-rate model for observed zero-coupon yields and survey expectations by maximum likelihood on quarterly US interest rate data from 1983:Q1 to 2020:Q3 and Blue Chip Financial Forecasts. We then use our maximally-flexible model to test the rational expectation hypothesis and for the presence of a priced survey factor unspanned by observed rates.

Contrary to earlier approaches postulated in the term structure literature, we find that surveys do not help to identify the physical dynamics parameters. Under the full information rational expectations (FIRE) hypothesis, agents' subjective probability measure coincides with the real-world measure, so that rational agents' forecast errors should not be biased or persistent.[1] However in a formal test we overwhelmingly reject the FIRE hypothesis.

[1]This definition of rationality is different from the one used by Adam and Marcet (2011) and Singleton

In particular, we find that imposing FIRE implies that the factor dynamics collapse to subjective expectations, but then they fail to forecast interest rates. On the other hand, when we allow for different objective and subjective dynamics, the estimated objective dynamics are very similar to those obtained from a yields-only model, but in this case they cannot be used to determine the risk premium demanded by investors.

Our rejection of the rational expectation hypothesis using survey expectations is in line with previous results in macroeconomic (see, among others, Mankiw, Reis and Wolfers 2003, Coibion and Gorodnichenko 2015) and finance (see Greenwood and Shleifer 2014, Adam, Marcet and Beutel 2017, De La O and Myers 2021). An important implication is that the perceived risk premium of market agents is not measured by the objective risk premium (determined by the real-world dynamics of observed interest rates) but instead by the subjective risk premium (determined by the dynamics of survey expectations, see Nagel and Xu 2022).

A second important result from our analysis is that surveys contain important information as they are driven by a priced risk factor that is not spanned by observed interest rates. We show that accounting for this unspanned survey-specific factor is important for a more reliable measurement of interest rate forecasts and of the risk premium. It is usually difficult to determine a priori which variables are good candidates for unspanned risk factors. The literature has focused on macroeconomic variables (see Joslin, Priebsch and Singleton 2014, Coroneo, Giannone and Modugno 2016), but as noted in Duffee (2011), macroeconomic variables explain only a small fraction of the variation in the hidden factor. In this paper, instead, we follow the intuition in Duffee (2011) that "the most obvious choice is survey data on interest rate forecasts." Indeed, our results indicate that while surveys cannot be used to aid the identification of the physical parameters, they contain priced information that helps

(2021), in which the agent does not know the data generating process but forms rational beliefs.

predict observed interest rates.[2]

The paper is organized as follows. In Section 2 we describe the data and present the three stylized facts that show the deficiency of standard term structure models. In Section 3 we present the general modelling framework and the shadow rate term structure model. In Section 4 we describe the estimation approach. In Section 5 we report the empirical results, and, finally, in Section 6 we conclude.

2 Data and preliminary evidence

2.1 Data

We use quarterly observations for zero-coupon yields and surveys for the period 1983:Q1 to 2020:Q3. For interest rates, we use end-of-quarter rates for maturities 3 and 6 months, 1, 2, 5, 7 and 10 years. Data for 3 and 6-month maturities are from the FRED dataset, while data for 1, 2, 5, 7 and 10-year maturities are from the Federal Reserve Board website.

For surveys we use Blue Chips Financial Forecasts (BCFF) consensus (mean) forecasts at 1 through 5 quarters ahead, which is the longest forecast horizon available throughout the whole data sample. Although the BCFF forecasts are published at the beginning of a month, they are collected in the last few days of the preceding month. Thus, to align the timing of the conditioning information set of the survey forecasts with that in spot interest rates, we assume that the surveys are observed as soon as they are collected (i.e. at the end of the month), so that we use BCFF surveys published in the first few days of April, July, October and January as observed at the end of preceding months: March, June, September and December, respectively.

[2]In a closely related paper Piazzesi et al. (2015) also use survey data on interest rate forecasts to construct subjective bond risk premia, but they do not consider unspanned information in surveys, and they do not formally test for the equality between subjective and objective dynamics.

We use forecasts for all available rates, which are generally different at each point in time. The 3-month rate is available in surveys throughout the entire sample. With the exception of the first three quarters of 1983, when surveys also include the 10-year rate, up to 1987:Q4, only three survey rates are available (3 months, 3 years and 30 years). From 1988:Q1 the surveys also always include the 6-month, 1-year and 10-year rates. For almost the whole sample period the longest available rate is the 30-year bond, with the exception of the period 2002:Q2 to 2006:Q1, during which the longest rates are for 10-year (2002:Q2-2004:Q2) and 20-year bonds (2004:Q3-2006:Q1).

Since the surveys refer to T-bill discount rates and par yields, we extract zero-coupon rates using the Nelson and Siegel (1987) and Svensson (1994) model.[3] In particular, we parameterize the zero-coupon yield curve and use it to compute the implied T-bill discount rates and par yields, which we then fit to the survey data. To prevent overfitting, in the period 1983:Q1 to 1987:Q4, we fit the simple 4-parameter Nelson and Siegel (1987) model, where the parameters are the three shape parameters and the decay parameter. From 1988:Q1 the BCFF surveys include forecasts for between 6 and 8 different maturities, which allows us to fit the more flexible Svensson (1994) model that includes four shape parameters and two decay parameters. From this procedure, throughout this paper, we use 3-month, 1, 5 and 10-year zero-coupon survey yields.

Finally, we note that the BCFF forecasts are for average rates over calendar quarters. To align the forecasts with spot interest rates, following the standard approach in the literature (e.g. Kim and Orphanides 2012, Buraschi, Piatti and Whelan 2022, Nagel and Xu 2022), we assume that the survey forecasts are for mid-quarter rates, and interpolate linearly between quarters. We use forecasts for up to 5-quarters ahead, which after interpolation gives us

[3] Although T-bills are quoted on the actual-to-360 days basis, for simplicity we assume that each calendar quarter has 90 days. For a textbook treatment of quoting conventions and the Nelson and Siegel (1987) model, see Veronesi (2010).

quarterly forecasts up to 1 year ahead. Thus, in the empirical analysis we use 16 interest rate survey forecasts: 4 maturities (0.25, 1, 5 and 10 years) with 4 forecast horizons (from one to four quarters ahead) each.

2.2 Preliminary evidence

In this section, we document three empirical findings that show the deficiency of standard yields-only term structure models. First, as explained in the introduction, the standard approach in term structure modelling is to assume that the probability measure used by economic agents is the same as the statistical probability measure.[4] It follows that the model-implied term premium is interpreted as the risk premium that investors require for holding long-maturity bonds. In this conceptual framework, investors' expectations are filtered out from the physical dynamics of interest rates. Then, a valid question is how the expectations implied by the model are related to market expectations.

To address this question in Figure 1 we plot the median forecasts of the Blue Chip Financial Forecasts for the 1−year yield over the period 1983:Q1 to 2020:Q3 for horizons 1 through 4 quarters ahead.[5] In the same figure we plot the forecasts implied by a term structure model estimated using only yields (in the rest of the paper we refer to this specification as Case 0). Since a large part of the sample size is the period of the zero lower bound, we use a shadow rate model that prohibits negative interest rates (for more details please see Section 3.3). Visual inspection reveals that forecasts from the yields-only model seem to have little relation to market-based expectations. Even more striking is the pattern that indicates a different direction of predicted interest rates. Thus, our first fact is that subjective dynamics

[4] In the literature, the statistical measure is also called the real-world measure, the physical measure or the objective measure. In this paper, we use these terms interchangeably.

[5] Figures for other maturities look fundamentally the same and they are available from the authors upon request.

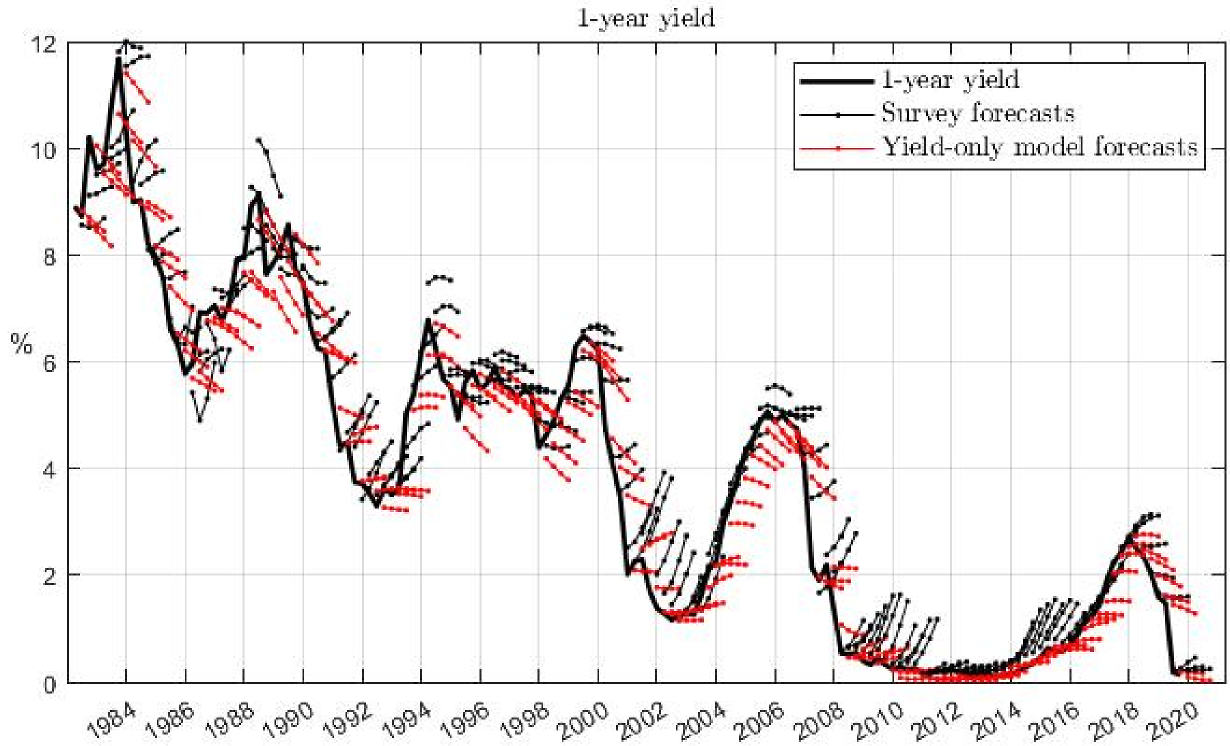

Figure 1. Market expectations measured by surveys (black-dotted lines) and expectations extracted from a yields-only term structure model of interest rates (red-dotted lines), along with the observed 1-year yield (bold black line).

are different from real-world dynamics.

Fact 1 *The real-world probability distribution that drives observed interest rates is different from the subjective probability distribution that drives interest rate survey expectations.*

Our second fact is a finding that the space spanned by interest surveys is different from the space spanned by spot interest rates. In Table 1 we report the principal component analysis for yields and surveys. We report both the percentage of the total variance explained by the principal components and, to get some economic intuition of the magnitudes, also the (time-series and cross-sectional) average $RMSE$ of the remaining fitting errors (in basis points).

Yields: fit with first k YPC

k	1	2	3	4	5
Variance explained	97.30	99.88	99.98	100.00	100.00
Average RMSE	45.45	9.33	3.49	1.12	0.24

Surveys: fit with first k SPC

k	1	2	3	4	5
Variance explained	97.15	99.81	99.90	99.95	99.98
Average RMSE	44.95	11.68	8.40	5.90	3.94

Surveys: fit with 3 YPC and k $SPC_{\perp}$

k	0	1	2	3	4
Variance explained	97.97	99.63	99.85	99.93	99.96
Average RMSE	39.35	16.48	10.15	7.26	5.08

Table 1. Principal component analysis
Cumulative proportions of variance (in percentage points) and average RMSEs (in basis points) of observed yields (top panel) and surveys (mid and bottom panels) explained by the principal components extracted from yields (YCY), surveys (SPC), or the residuals of the projection of surveys in the first three yields principal components ($\mathrm{SPC}_{\perp}$).

The results indicate that three principal components extracted from observed yields (YPC) explain virtually all variation in yields, and also that three principal components extracted from surveys (SPC) explain 99.90% of the total variation in surveys. The resulting fit is very tight, slightly more than 3 basis points for yields and 8 basis points for surveys. This seems to indicate that three state variables extracted from interest rates drive the dynamics of interest rates surveys as well. More precisely, if there are no arbitrage opportunities in the bond market and there are no unspanned risk factors that drive the interest rates dynamics under the subjective probability measures, then the first three principal components of yields should be sufficient to span the whole survey space.[6]

The bottom panel of Table 1, however, contradicts this conjecture, since the space

[6]If the pricing model is nonlinear, as in our application, this statement is valid only up to a first-order approximation.

spanned by these two sets of variables is different. Here we report the percentage of the total variance of surveys explained by the first 3 principal components of yields plus up to 4 first principal components of the residuals from the regression of surveys on the first three YPCs, denoted $SPC_{\perp}$. We can see that the 3 yield principal components explain only 97.97% of the surveys variance and leave measurement errors with an average $RMSE$ of about 39 basis points. A survey-specifc factor that is (by construction) orthogonal to the principal components of observed yields increases this percentage to 99.63%, an increase by almost 1.7 percentage points, which reduces the average $RMSE$ of the fitting errors to 16 basis points. Adding more principal components explains little additional survey variation, suggesting that the remaining variation in surveys is effectively just idiosyncratic noise.

This provides evidence that there is systematic comovement in surveys that can be captured by a factor unspanned by bond prices. As such, in the following analysis we focus on the first principal component that captures the common variation specific to interest rate surveys, which we dub the $s-$factor.[7]

Fact 2 *Interest rates surveys, apart from yield curve factors, are spanned by an additional, survey-specific factor. In other words, there is a factor unspanned by yields that drives the subjective dynamics of interest rates.*

Finally, we examine the role the survey-specific factor plays in factor dynamics under the real-world probability measure. To this end, we estimate a $VAR(1)$ with the first three principal components extracted from yields (YPC) and the $s-$factor ($SPC_{\perp}$). In the top panel of Table 2, we report the coefficient estimates, which indicate that the survey factor Granger-causes the first YPC at the 1% significance level and the third YPC at the 5%

[7]We also check the possibility that surveys are spanned by higher principal components of yields. The analysis of survey residuals from the regression on the 5 first YPCs is virtually identical to that reported in the bottom panel of Table 1 - results available upon request.

	$VAR(1)$			
	$YPC_{1,t-1}$	$YPC_{2,t-1}$	$YPC_{3,t-1}$	$SPC_{\perp,t-1}$
$YPC_{1,t}$	0.9835^{***}	0.0062	-0.0216^{**}	0.1204^{***}
$YPC_{2,t}$	-0.0026	0.9238^{***}	0.1250^{***}	-0.0396
$YPC_{3,t}$	0.0699^{*}	-0.0218	0.6552^{***}	-0.1515^{**}
$SPC_{\perp,t}$	-0.0086	0.0864	0.0078	0.2319^{***}

	Excess returns			
	$YPC_{1,t-1}$	$YPC_{2,t-1}$	$YPC_{3,t-1}$	$SPC_{\perp,t-1}$
$xret_{1,t}$	-1.9777^{***}	0.4302^{***}	-0.0594^{**}	-0.2456^{***}
$xret_{2,t}$	-1.9106^{***}	0.5118^{***}	-0.1041	-0.4261^{***}
$xret_{5,t}$	-1.8587^{***}	0.7612^{***}	-0.2573	-0.7565^{***}
$xret_{7,t}$	-1.8501^{***}	0.9646^{***}	-0.3107	-0.8931^{***}
$xret_{10,t}$	-1.8332^{***}	1.1786^{***}	-0.3558	-1.0314^{***}

Table 2. Estimates of VAR(1) coefficients (top panel) and coefficients of predictive regression of excess returns on yield curve factors and the survey factor (bottom panel). ***, ** and * denote significance at the 1%, 5% and 10% level.

level. In the bottom panel of Table 2 we report forecasting regressions of quarterly excess returns for 1, 2, 5, 7 and 10−year zero coupon bonds on the yield curve factors (YPC) and the $s-$factor ($SPC_{\perp}$). In our sample bond excess returns are strongly predictable by the first and the second YPC. Importantly, however, the $s-$factor is statistically significant in forecasts of all excess returns at the 1% significance level.

Although this preliminary analysis provides plenty of evidence that the survey-specific factor is important for the real-world dynamics of interest rates, we should note that the statistical significance presented in this section is likely underestimated due to the possible presence of nonlinearities in the factor dynamics induced by the zero lower bound of interest rates. All in all, this evidence indicates that, despite not being spanned by the observed rates by construction, the survey factor contains useful information to predict future interest rates.

Fact 3 *The survey-specific factor (s−factor) drives the real-world dynamics of (Granger-causes) the yield curve factors.*

Taken together, Facts 1-3 indicate that in order to correctly estimate the term premium and interest rate expectations any term structure model must allow for separate real-world and survey dynamics and also include an additional survey-specific state variable that is unspanned by observed interest rates.

3 Framework

In this section we detail our term structure modeling framework which involves both the standard specification of the risk-neutral and the physical dynamics, and also the subjective dynamics. We embed the underlying Gaussian dynamics of the factors in a shadow rate model, which prohibits negative interest rates.

3.1 Bond prices and real-world factor dynamics

We assume that bond prices are driven by a $K-$dimensional state vector $\mathbf{x}_t$, and that there exists a risk-neutral probability measure $\mathbb{Q}$ that prices all financial assets, under which the state variable follows a first-order Gaussian vector autoregression ($VAR(1)$):

$$\mathbf{x}_t = \mu^{\mathbb{Q}} + \mathbf{\Phi}^{\mathbb{Q}}\mathbf{x}_{t-1} + \mathbf{u}_t^{\mathbb{Q}}, \tag{1}$$

where $\mathbf{u}_t^{\mathbb{Q}} \overset{\mathbb{Q}}{\sim} i.i.d.\mathcal{N}(\mathbf{0}, \mathbf{\Sigma_x})$.

We denote the logarithm of the price of a zero-coupon bond at time t with remaining maturity n by $p_{n,t}$ and the corresponding yield by $y_{n,t} = -p_{n,t}/n$. In the absence of arbitrage opportunities, the price of a zero-coupon bond can be expressed as

$$p_{n,t} = \log E_t^{\mathbb{Q}} \left[\exp \left(- \sum_{j=0}^{n-1} r_{t+j} \right) \right], \tag{2}$$

where r_t is the short rate driven solely by the state vector $\mathbf{x}_t$, that is $r_t = r(\mathbf{x}_t)$.[8] It follows that bond yields are spanned by the state vector, which we denote as

$$y_{n,t} \equiv y(\mathbf{x}_t, n; \mathbf{\Psi}), \tag{3}$$

where $\mathbf{\Psi}$ is a vector of relevant risk-neutral parameters and $y(.)$ is the (yield) pricing function.

The risk-neutral dynamics of the risk factors are not observable and can only be inferred from asset prices. Instead, what we measure are the physical dynamics of the term structure variables $\mathbf{x}_t$, which can be also influenced by other unspanned factors $\mathbf{s}_t$. We assume that the system of physical (or real-world) dynamics follows a Gaussian $VAR(1)$:

$$\begin{bmatrix} \mathbf{x}_t \\ \mathbf{s}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_x^{\mathbb{P}} \\ \boldsymbol{\mu}_s^{\mathbb{P}} \end{bmatrix} + \begin{bmatrix} \mathbf{\Phi}_{xx}^{\mathbb{P}} & \mathbf{\Phi}_{xs}^{\mathbb{P}} \\ \mathbf{\Phi}_{sx}^{\mathbb{P}} & \mathbf{\Phi}_{ss}^{\mathbb{P}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{t-1} \\ \mathbf{s}_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{x,t}^{\mathbb{P}} \\ \mathbf{u}_{s,t}^{\mathbb{P}} \end{bmatrix}, \tag{4}$$

where $[\mathbf{u}_{x,t}^{\mathbb{P}}{}', \mathbf{u}_{s,t}^{\mathbb{P}}{}']' \overset{\mathbb{P}}{\sim} i.i.d.\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, with the $K \times K$ top left matrix block of $\mathbf{\Sigma}$ equal to $\mathbf{\Sigma}_{\mathbf{x}}$ defined in Eq.(1). As long as the matrix block $\mathbf{\Phi}_{xs}^{\mathbb{P}}$ contains non-zero elements, we say that the variable $\mathbf{s}_t$ is unspanned by yields because it cannot be recovered from bond prices (see (3)) but, nonetheless, it drives the expectation (under the $\mathbb{P}$ measure) of future interest rates and thus the objective risk premium.

Under FIRE the real-world probability measure defines the risk premium demanded by risk-averse investors for holding risky assets. In particular, if market participants were

[8] If the pricing function includes lags of $\mathbf{x}_t$, we can always redefine the state vector as $\tilde{\mathbf{x}}_t = [\mathbf{x}_t', \mathbf{x}_{t-1}', \ldots]'$.

rational and risk indifferent, the log bond price would be given by the real-world expectation of interest rates in the future (adjusted for convexity), a theory commonly known as the expectations hypothesis of interest rates. Building on this concept, Bauer, Rudebusch and Wu (2012) define a 'risk-neutral' yield that reflects expectations of the short term rate over the lifetime of the bond under the real-world probability measure

$$y_{n,t}^{\mathbb{P}} = -\frac{1}{n} \log E_t^{\mathbb{P}} \left[\exp \left(-\sum_{j=0}^{n-1} r_{t+j} \right) \right]. \tag{5}$$

Hence, the yield risk premium on an $n-$period zero-coupon bond with respect to the real-world probability measure is defined as

$$\begin{aligned} rp_{n,t}^{\mathbb{P}} &= y_{n,t} - y_{n,t}^{\mathbb{P}} \\ &= \frac{1}{n} \log E_t^{\mathbb{P}} \left[\exp \left(-\sum_{j=0}^{n-1} r_{t+j} \right) \right] - \frac{1}{n} \log E_t^{\mathbb{Q}} \left[\exp \left(-\sum_{j=0}^{n-1} r_{t+j} \right) \right]. \end{aligned} \tag{6}$$

3.2 Surveys and subjective risk premia

Assume now that, in addition to the zero-coupon bond prices, we also observe market agents' expectations about future zero-coupon yields. Typically in the literature the rational expectations (or FIRE) hypothesis is assumed, so that the subjective dynamics are the same as the dynamics under the real-world probability measure (4). This is a useful assumption, since it allows us to 'measure' market expectations from the history of bond prices, which in turn allows us to determine the objective risk premium. However, in general, the subjective probability measure $\mathbb{S}$ used by market agents can differ from the real-world probability measure $\mathbb{P}$. Indeed, if market expectations are observable, the equality of the $\mathbb{S}$ and $\mathbb{P}$ measures is a testable hypothesis.

Denote the $h-$period ahead expectation at time t of a yield on a zero-coupon bond with an $n-$period tenor by $y^s_{n,t,h}$, such that $y^s_{n,t,h} \equiv E^{\mathbb{S}}_t[y(\mathbf{x}_{t+h}, n; \mathbf{\Psi})]$. In general, it might not be possible to find an analytic solution to this expectation as, e.g., in the case of the shadow rate model that we are going to adopt, see Section 3.3. Therefore, we are going to work with a first-order approximation

$$y^s_{n,t,h} \approx y(E^{\mathbb{S}}_t[\mathbf{x}_{t+h}], n; \mathbf{\Psi}). \tag{7}$$

Eq.(7) is exact for a fully Gaussian (Vasicek) model and almost exact for a shadow rate model when interest rates are away from the zero lower bound, while at the zero lower bound there are omitted higher-order terms.

In this paper, we assume that the factor dynamics under $\mathbb{S}$ also follow a $VAR(1)$

$$\begin{bmatrix} \mathbf{x}_t \\ \mathbf{s}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}^{\mathbb{S}}_x \\ \boldsymbol{\mu}^{\mathbb{S}}_s \end{bmatrix} + \begin{bmatrix} \mathbf{\Phi}^{\mathbb{S}}_{xx} & \mathbf{\Phi}^{\mathbb{S}}_{xs} \\ \mathbf{\Phi}^{\mathbb{S}}_{sx} & \mathbf{\Phi}^{\mathbb{S}}_{ss} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{t-1} \\ \mathbf{s}_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{u}^{\mathbb{S}}_{x,t} \\ \mathbf{u}^{\mathbb{S}}_{s,t} \end{bmatrix}, \tag{8}$$

where $[\mathbf{u}^{\mathbb{S}}_{x,t}{}', \mathbf{u}^{\mathbb{S}}_{s,t}{}']' \overset{\mathbb{S}}{\sim} i.i.d.\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$. Note that, by the Girsanov theorem, the conditional variance matrix of the factor innovations is the same under different probability measures. If $\mathbf{s}_t$ is a scalar variable (i.e. $\mathbf{s}_t = s_t$), as we assume following the preliminary evidence set out in Section 2, $\mathbf{\Phi}^{\mathbb{S}}_{xs}$ is a $K \times 1$ block of parameters.

An advantage of using survey data is that it allows us to readily recover the unspanned risk factor under the $\mathbb{S}$ probability measure. To see this, note that unless the block $\mathbf{\Phi}^{\mathbb{S}}_{xs}$ includes only zeros, the price expectation becomes also a function of $\mathbf{s}_t$, since

$$E^{\mathbb{S}}_t[\mathbf{x}_{t+h}] = \boldsymbol{\mu}^{\mathbb{S}}_{x,h} + \mathbf{\Phi}^{\mathbb{S}}_{x,h}\mathbf{x}_t + \mathbf{\Phi}^{\mathbb{S}}_{s,h}\mathbf{s}_t, \tag{9}$$

where $\boldsymbol{\mu}^{\mathbb{S}}_{x,h}$, $\mathbf{\Phi}^{\mathbb{S}}_{x,h}$ and $\mathbf{\Phi}^{\mathbb{S}}_{s,h}$ are functions of terms defined in (8). As such, the existence of a survey-specific factor documented in Section 2 is equivalent to the statement that $\mathbf{\Phi}^{\mathbb{S}}_{s,h}$ is nonzero, which holds if $\mathbf{\Phi}^{\mathbb{S}}_{xs}$ is nonzero.

We can define the subjective risk premium analogously to the definition of the objective risk premium in Eq.(6), that is:

$$rp^{\mathbb{S}}_{n,t} = \frac{1}{n}\log E^{\mathbb{S}}_t\left[\exp\left(-\sum_{j=0}^{n-1} r_{t+j}\right)\right] - \frac{1}{n}\log E^{\mathbb{Q}}_t\left[\exp\left(-\sum_{j=0}^{n-1} r_{t+j}\right)\right]. \tag{10}$$

The subjective risk premium has a clear economic interpretation as the compensation required by risky investors for holding risky assets, since it corresponds to the investors' beliefs. This is corroborated by the finding in Giglio, Maggiori, Stroebel and Utkus (2021) that survey data is highly informative about individuals' portfolio decisions. On the other hand, if the subjective probability measure is different from the real-world measure, the objective risk premium in (6) is difficult to interpret, as it merely represents the measured ex-post risk premium based on the history of bond prices.

Since the $s-$factor exists if and only if it drives the interest rates dynamics under the $\mathbb{S}$ probability measure, it drives the subjective risk premium by construction. Moreover, if we assume that the subjective probability measure coincides with the real-world measure, the survey-specific factor is also a predictor of interest rates and, as such, it is a risk factor under the $\mathbb{P}$ probability measure. This observation could be useful if the modeller is uncertain about which variables should be included as risk factors in a term structure model and she is willing to assume that the subjective and real-world measures are the same.[9] In this case, she can avoid the inclusion of arbitrary macro variables and substitute them with the factor extracted from surveys. However, in the more general case in which the subjective and the objective dynamics are different, it is an empirical question whether the s-factor also drives

[9] In the empirical section, we reject the hypothesis about the equality of the subjective and real-world probability measures.

Is s_t a risk factor under $\mathbb{P}$? $\mathbf{\Phi}^{\mathbb{P}}_{xs} \neq \mathbf{0}$	Rational expectations ($\mathbb{S} \sim \mathbb{P}$) $\boldsymbol{\mu}^{\mathbb{P}}_x = \boldsymbol{\mu}^{\mathbb{S}}_x$ and $\mathbf{\Phi}^{\mathbb{P}}_{xx} = \mathbf{\Phi}^{\mathbb{S}}_{xx}$	No rational expectations ($\mathbb{S} \nsim \mathbb{P}$) $\boldsymbol{\mu}^{\mathbb{P}}_x \neq \boldsymbol{\mu}^{\mathbb{S}}_x$ and $\mathbf{\Phi}^{\mathbb{P}}_{xx} \neq \mathbf{\Phi}^{\mathbb{S}}_{xx}$
No	**Case 1** (S$\sim \mathbb{P}$, no s_t) surveys informative for econometric identification of $\mathbb{P}-$parameters	**Case 3** (S$\nsim P$, s_t is not $\mathbb{P}-$risk factor) surveys not informative
Yes	**Case 2** (S$\sim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor) surveys informative for both the physical risk factors and $\mathbb{P}-$parameters	**Case 4** (S$\nsim P$, s_t is $\mathbb{P}-$risk factor) surveys informative for physical risk factors

Table 3. Model specifications.

the objective dynamics.

In our framework, both the assumptions of rational expectations and about the existence of a priced unspanned survey factor can be tested. To facilitate the interpretation of these assumptions and their implications, in Table 3 we tabulate the four different cases that arise.

In Case 1 the subjective and the real-world measures coincide and all the risk factors can be extracted from observed rates (no hidden factors). In this case surveys can be used to aid the econometric identification of the physical parameters of the term structure model (Kim and Orphanides 2012, d'Amico et al. 2018), to proxy for expectations (Crump et al. 2018), or to proxy for state variables (Chun 2011), but they do not convey any additional information about the risk factors. In terms of the model parameters, in Case 1 the factor s_t and the corresponding parameters do not appear in the factor dynamics equations (4) and (8), so that $\mathbf{\Phi}^{\mathbb{P}}_{xs} = \mathbf{\Phi}^{\mathbb{S}}_{xs} = \mathbf{0}$ and the remaining parameters coincide, i.e. $\boldsymbol{\mu}^{\mathbb{P}}_x = \boldsymbol{\mu}^{\mathbb{S}}_x$ and $\mathbf{\Phi}^{\mathbb{P}}_{xx} = \mathbf{\Phi}^{\mathbb{S}}_{xx}$.

In Case 2 we maintain the rational expectations hypothesis, and we have a hidden factor under the physical measure (which, as a consequence, cannot be extracted from observed yields). Consequently, surveys have a dual role. First, they convey information about the hidden factor that drives real-world expectations and risk premium. Second, they help identify econometrically the physical parameters of the term structure model. This amounts to the restriction that the real-world dynamics in Eq.(4) and the subjective dynamics in Eq.(8) are the same, including $\mathbf{\Phi}^{\mathbb{P}}_{xs} = \mathbf{\Phi}^{\mathbb{S}}_{xs} \neq \mathbf{0}$.

In Case 3 we abandon the rational expectations hypothesis, and we assume that all the risk factors under the objective measure can be extracted from observed rates (no hidden factors under the objective measure). In this case surveys do not convey additional information about the risk factors extracted from observed rates nor help to identify econometrically the model parameters. In terms of the model specification, the parameters in Eq.(4) and Eq.(8) are allowed to be different but we have the restriction $\mathbf{\Phi}^{\mathbb{P}}_{xs} = \mathbf{0}$.

In Case 4 we have hidden factors under the physical measure and no rational expectations. In this case, surveys convey information about the hidden factor that drives observed risk premium and real-world expectations. This is the general model presented in the previous section, in which all parameters in Eq.(4) and Eq.(8) are allowed to be different and freely estimated. As such, it nests all the other cases and allows us to conduct statistical tests of the corresponding restrictions.

3.3 The shadow rate term structure model

To enforce a lower bound on interest rates, we assume that there exists a shadow short rate that is linear in the Gaussian state variables

$$ssr_t = \delta_0 + \boldsymbol{\delta}_1' \mathbf{x}_t \tag{11}$$

and that the observed short-term interest rate r_t is equal to the shadow short rate only when the latter is above the lower bound and equal to the lower bound otherwise,

$$r_t = \max\{ssr_t, \underline{r}\}. \tag{12}$$

The short rate equation (12) together with the state dynamics under the risk-neutral measure (1) and under the subjective measure (8) represent the complete term structure model.

The advantage of the shadow rate model is that it allows us to operate within the Gaussian framework described in Section 3 but at the same time it imposes the lower bound on interest rates, which is essential for the consistency of the model estimates if the sample period includes such episodes. The idea of a shadow rate was first introduced by Black (1995) and is based on the observation that negative interest rates are hard to enforce if investors have the option to convert to currency. As a result of this option, all term rates and forward rates are bounded and do not have a reflecting boundary condition, as opposed to affine models with factors that follow square-root processes or quadratic Gaussian models. Another feature of the shadow rate term structure model is that, when short-term interest rates are far from the lower bound, interest rates behave approximately as in a Gaussian affine term structure model.

Although in the Gaussian setting the pricing functional for the shadow rate model is available in closed form (see Priebsch 2013), it is computationally demanding, which makes it inconvenient in practical applications. Wu and Xia (2016), however, provide a convenient approximation in discrete time. In particular, denote by $f_{n,t}$ the time t one period forward rate for a loan starting at $t + n$. Wu and Xia (2016) show that, under (1) and (12), the forward rate $f_{n,t}$ is approximately equal to

$$f_{n,t} \approx \underline{r} + \sigma_n^{\mathbb{Q}} g\left(\frac{a_n + \mathbf{b}_n'\mathbf{x}_t - \underline{r}}{\sigma_n^{\mathbb{Q}}}\right) \tag{13}$$

where the function $g(z) = z\mathtt{N}(z) + \mathtt{n}(z)$ with $\mathtt{N}(z)$ and $\mathtt{n}(z)$ the CDF and the PDF of z, respectively, and $\sigma_n^{\mathbb{Q}}$, a_n and $\mathbf{b}_n$ are known coefficients, given the risk-neutral parameters.[10]

4 Estimation and identification

We estimate the joint model of zero-coupon yields and interest rate survey expectations by maximum likelihood. We implement it using the factor extraction method proposed by Golinski and Spencer (2022), which generalizes the estimation technique by factor rotation with observable factors introduced by Joslin, Singleton and Zhu (2011) to non-linear models. This method does not require any approximation (such as e.g. that required by the extended Kalman filter) and is computationally efficient since it allows for concentrating the likelihood function with respect to the $\mathbb{P}$ parameters.

We denote the vector of model-implied zero-coupon yields at time t by $\mathbf{y}_t$ and the associated vector of pricing errors by $\mathbf{v}_t \sim N(\mathbf{0}, \mathbf{\Sigma_y})$, such that the observed yields are $\widetilde{\mathbf{y}}_t = \mathbf{y}_t + \mathbf{v}_t$. Similarly, we denote the vector of observable surveys with different maturities and forecasting horizons as $\widetilde{\mathbf{y}}_t^s = \mathbf{y}_t^s + \boldsymbol{\eta}_t$, where entries of the model-implied surveys $\mathbf{y}_t^s$ are given in Eq.(7) and $\boldsymbol{\eta}_t$ are the associated zero mean measurement errors with variance $\mathbf{\Sigma_s}$.

Following Joslin et al. (2011), we assume that the first three principal components of yields are measured without error. If the eigenvectors associated with the three largest eigenvalues of the covariance matrix of observed yields are collected in a $N \times 3$ matrix $\mathbf{W}$ (where N

[10]For more details and definition of the terms in (13), please refer to Wu and Xia (2016), Appendix A.

is the number of yields), the first three principal components of observed yields are given by $\mathbf{q}_{y,t} \equiv \mathbf{W}'\widetilde{\mathbf{y}}_t = \mathbf{W}'\mathbf{y}_t$, where the last equality follows from the observability assumption ($\mathbf{W}'\mathbf{v}_t = \mathbf{0}$). As shown in Golinski and Spencer (2022), this observability assumption allows us to extract the underlying state vector $\mathbf{x}_t$ from the nonlinear system, conditionally on the risk-neutral parameters.

In a similar fashion, we extract the factor unspanned by interest rates but spanned by interest rates surveys. We assume that the first principal component of surveys (across maturities and forecast horizons) $q_{s,t}$ is observable without error. Then, conditional on the $\mathbb{Q}$ parameters (which is equivalent to observable $\mathbf{x}_t$) and the $\mathbb{S}$ parameters, we can solve the system of interest rates surveys with respect to s_t.

We adopt the parameterization of the model of Joslin et al. (2011) in which the $\mathbb{Q}$ dynamics depend on $\{\mu_1^{\mathbb{Q}}, \boldsymbol{\lambda}^{\mathbb{Q}}, \boldsymbol{\Sigma}\}$, where $\mu_1^{\mathbb{Q}}$ is the first element of $\boldsymbol{\mu}^{\mathbb{Q}}$, other elements being zeros, and $\boldsymbol{\lambda}^{\mathbb{Q}}$ is the vector of eigenvalues of $\boldsymbol{\Phi}^{\mathbb{Q}}$. The short shadow-rate equation (11) is identified with $\delta_0 = 0$ and $\boldsymbol{\delta}_1 = \mathbf{1}$, where $\mathbf{1}$ denotes a vector of ones. Since the $s-$factor is essentially a (nonlinear) rotation of the system of surveys, it is not identified for unrestricted $\mathbb{S}$ parameters since s_t can be linearly rotated to an observationally equivalent representation, s_t^*, by changing its mean and scale: $s_t^* = c_0 + c_1 s_t$, where c_0 and c_1 are arbitrary constants. Thus, to identify the $s-$factor we impose $\mu_s^{\mathbb{S}} = 0$; and $\sum_{j=1}^{K} \phi_{sx,j}^{\mathbb{S}} = 1$, where $\phi_{sx,j}^{\mathbb{S}}$ are elements of the $1 \times K$ vector $\boldsymbol{\Phi}_{sx}^{\mathbb{S}}$ defined in (8). In addition, we impose $\underline{r} = 0$, which guards against model overfitting. Denote $\boldsymbol{\Theta} \equiv \{\mu_1^{\mathbb{Q}}, \boldsymbol{\lambda}^{\mathbb{Q}}, \boldsymbol{\Sigma}_{\mathbf{y}}, \boldsymbol{\mu}^{\mathbf{S}}, \boldsymbol{\Phi}^{\mathbb{S}}, \boldsymbol{\Sigma}_{\mathbf{s}}, \boldsymbol{\mu}^{\mathbb{P}}, \boldsymbol{\Phi}^{\mathbb{P}}, \boldsymbol{\Sigma}\}$ and the conditional likelihood function as

$$\log \mathcal{L}(\boldsymbol{\Theta}) = \sum_{t=2}^{T} \log \ell(\widetilde{\mathbf{y}}_t, \widetilde{\mathbf{y}}_t^s | \widetilde{\mathbf{y}}_{t-1}, \widetilde{\mathbf{y}}_{t-1}^s; \boldsymbol{\Theta}). \tag{14}$$

Then, the time$-t$ conditional likelihood can be decomposed as

$$
\begin{aligned}
\ell(\widetilde{\mathbf{y}}_t, \widetilde{\mathbf{y}}^s_t | \widetilde{\mathbf{y}}_{t-1}, \widetilde{\mathbf{y}}^s_{t-1}; \mathbf{\Theta}) \quad = \quad & \ell^{\mathbb{Q}}(\widetilde{\mathbf{y}}_t | \mathbf{q}_{y,t}; \mu^{\mathbb{Q}}_1, \boldsymbol{\lambda}^{\mathbb{Q}}, \mathbf{\Sigma}_{\mathbf{y}}, \mathbf{\Sigma}) \\
& \times \ell^{\mathbb{S}}(\widetilde{\mathbf{y}}^s_t | \mathbf{q}_{y,t}, q_{s,t}; \mu^{\mathbb{Q}}_1, \boldsymbol{\lambda}^{\mathbb{Q}}, \boldsymbol{\mu}^{\mathbb{S}}, \mathbf{\Phi}^{\mathbb{S}}, \mathbf{\Sigma}_{\mathbf{s}}, \mathbf{\Sigma}) \\
& \times \ell^{\mathbb{P}}(\mathbf{q}_{y,t}, q_{s,t} | \mathbf{q}_{y,t-1}, q_{s,t-1}; \boldsymbol{\mu}^{\mathbb{P}}, \mathbf{\Phi}^{\mathbb{P}}, \mathbf{\Sigma}).
\end{aligned}
$$

Note that the last part of the likelihood, $\ell^{\mathbb{P}}$, involves only the $\mathbb{P}$ parameters, which, upon rotation of the principal components $\mathbf{q}_{y,t}$ and $q_{s,t}$ to $\mathbf{x}_t$ and s_t, allows us to estimate the conditional mean parameters $\boldsymbol{\mu}^{\mathbb{P}}$ and $\mathbf{\Phi}^{\mathbb{P}}$ by *OLS* due to the classic result in Zellner (1962).[11]

Since the factor extraction of $\mathbf{x}_t$ is conditional on the $\mathbb{Q}$ parameters and the extraction of the $s-$factor is conditional on both the $\mathbb{Q}$ and $\mathbb{S}$ parameters, these have to be jointly estimated numerically. However, assuming that $\mathbf{x}_t$ and s_t are observable allows us to estimate the $\mathbb{P}$ dynamics in Eq.(4) by *OLS*, which greatly facilitates the estimation since it considerably reduces the number of parameters in numerical optimization. For more details, please see Appendix A.

5 Results

In this section we report the empirical results focussing on the comparison between the four different cases in Table 3, and also Case 0 (the yields-only model). After reporting the model fit for each case, we perform formal model specification testing. For each case, we then analyse the factor dynamics, the in-sample and out-of-sample forecasting performance, and the risk premium estimates. Lastly, we analyze the determinants of the unspanned survey factor.

[11]The factor rotation introduces in the likelihood a Jacobian term - see Golinski and Spencer (2022) for details.

Model	3m	6m	1y	2y	5y	7y	10y	Av. RMSE
Yields								
Case 0 (yield-only model, no s_t)	4.16	1.87	4.33	2.21	4.81	2.44	4.35	3.45
Case 1 ($\mathbb{S} \sim \mathbb{P}$, no s_t)	4.15	1.87	4.34	2.26	4.81	2.47	4.36	3.47
Case 2 ($\mathbb{S} \sim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor)	4.11	1.86	4.24	2.32	4.84	2.57	4.44	3.48
Case 3 ($\mathbb{S} \nsim \mathbb{P}$, s_t is not $\mathbb{P}-$risk factor)	4.13	1.87	4.26	2.30	4.84	2.55	4.44	3.48
Case 4 ($\mathbb{S} \nsim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor)	4.10	1.86	4.24	2.33	4.84	2.57	4.45	3.48

Model	Horizon	3m	1y	5y	10y	Av. RMSE
Surveys						
Case 0 (yield-only model, no s_t)	1q	37.97	44.58	41.38	41.85	41.45
	2q	40.67	53.87	48.57	44.46	46.89
	3q	47.63	64.97	59.42	51.78	55.95
	4q	60.20	79.78	72.96	62.91	68.96
Case 1 ($\mathbb{S} \sim \mathbb{P}$, no s_t)	1q	37.79	41.90	39.87	42.21	40.44
	2q	38.06	45.10	39.71	40.79	40.92
	3q	37.42	44.49	38.84	39.67	40.10
	4q	37.91	42.45	38.49	40.16	39.75
Case 2 ($\mathbb{S} \sim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor)	1q	18.69	20.21	12.63	17.85	17.34
	2q	14.64	18.23	10.89	15.71	14.87
	3q	12.84	17.26	12.83	16.18	14.78
	4q	15.67	16.43	15.72	18.49	16.58
Case 3 ($\mathbb{S} \nsim \mathbb{P}$, s_t is not $\mathbb{P}-$risk factor)	1q	18.47	20.18	12.60	17.40	17.17
	2q	14.49	18.10	11.08	15.35	14.75
	3q	12.64	17.08	13.19	15.96	14.72
	4q	15.36	16.09	16.03	18.47	16.49
Case 4 ($\mathbb{S} \nsim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor)	1q	18.45	20.12	12.59	17.56	17.18
	2q	14.43	18.14	11.05	15.56	14.80
	3q	12.70	17.22	13.07	16.12	14.78
	4q	15.55	16.33	15.86	18.52	16.56

Table 4. Model fit measured by the root mean-square error ($RMSE$) for yields (top panel) and surveys (bottom panel). The numbers are reported in basis points.

5.1 Fit to data

In Table 4 we present root mean-square errors for fitting errors for yields and surveys for different cases tabulated in Table 3, and the model estimated only with yields (Case 0). The first observation we can make is that the yield fit for all models (about 3.5 basis points) is very close to the unrestricted fit of first three principal components reported in Table 1. The yields-only model (Case 0) has the greatest flexibility to fit yields and so it also has the smallest average $RMSE$, although generally the differences among different models are not economically significant.

Similarly, due to the presence of a survey-specific factor in Cases 2, 3 and 4, the model fit of subjective expectations reported in Table 1 is close to the unrestricted fit of three YPC and the first $SPC_{\perp}$, 16.5 basis points on average, while the model without the survey factor (Case 1) is unable to fit the surveys. The $RMSE$ for surveys in Case 1 is about 40 basis points for all forecast horizons, more than double those reported for Cases 2-4, about $15-17$ basis points with a $U-$shaped fit pattern. It is also remarkable that restrictions placed on the objective probability measures have little effect on the fit of the model to surveys expectations.

5.2 Likelihood function and likelihood ratio tests

From the cross-sectional fit statistics presented in Table 4, it is clear that as long as we have a survey-specific factor in the model, the fitting errors do not provide us with sufficient information to discriminate among different hypotheses. Thus, for statistical comparison, we use the likelihood ratio test, where the most flexible model of Case 4 is the unrestricted model. The relevant models for comparison are Cases 2 and 3, i.e., the model specifications with the $s-$factor. The number of restrictions amounts to 20 for Case 2, and 3 for Case 3 (see the discussion in Section 3.2). In Table 5 we report the value of the log-likelihood function at the estimated maximum as well as its components as in Eq.(15). In the last column of the table we report the $p-$value of the likelihood ratio test for Cases 2 and 3.

First of all, the evidence in favour of the rejection of the rational expectations hypothesis (Case 2) is overwhelming, with the $p-$value equal to zero. We also reject the hypothesis that the $s-$factor is not priced under the $\mathbb{P}$ probability measure even when we allow for deviation from rational expectations (Case 3), with the $p-$value equal to zero. As such, our results

Model	Restrictions (# of restr.)	Total L	$L^{\mathbb{Q}}$	$L^{\mathbb{S}}$	$L^{\mathbb{P}}$	$p-$value
Case 4	$\mathbb{S} \nsim \mathbb{P}, s_t$ is P-risk factor None (0)	17,120.52	3,729.22	10,866.16	2,525.14	-
Case 2	$\mathbb{S} \sim \mathbb{P}, s_t$ is $\mathbb{P}-$risk factor $\boldsymbol{\mu}_x^{\mathbb{P}} = \boldsymbol{\mu}_x^{\mathbb{S}}$, $\boldsymbol{\Phi}^{\mathbb{P}} = \boldsymbol{\Phi}^{\mathbb{S}}$ (20)	17,080.31	3,729.54	10,855.72	2,495.04	0.0000
Case 3	$\mathbb{S} \nsim \mathbb{P}, s_t$ is not $\mathbb{P}-$risk factor $\boldsymbol{\Phi}_{xs}^{\mathbb{P}} = \mathbf{0}$ (3)	17,082.57	3,728.84	10,865.14	2,488.59	0.0000
Case 5	$\mathbb{S} \nsim \mathbb{P}, s_t$ is $\mathbb{P}-$risk factor $\mathbf{m}^{\mathbb{P}} = \mathbf{m}^{\mathbb{S}}$ (4)	17,116.19	3,729.56	10,884.45	2,502.19	0.0700

Table 5. Likelihood ratio test. Total likelihood values (first column), and its components as in Equation (15), along with likelihood ratio $p-$values (last column) for different cases.

provide us with strong evidence in favour of a $\mathbb{P}-$priced survey factor and an unequivocal rejection of the rational expectations hypothesis.

The inspection of the likelihood decomposition allows us to explore the reasons for these rejections. Confirming our former informal analysis of the fit of the model in Table 4, we can see that the parts of the likelihood corresponding to the cross-sectional fit of yields ($L^{\mathbb{Q}}$) and surveys ($L^{\mathbb{S}}$) for all cases are similar and, therefore, they are not the source of these rejections. However, the value of the likelihood corresponding to the physical dynamics of the factors ($L^{\mathbb{P}}$) deteriorates substantially for the restricted models.

In a closely related application, Piazzesi et al. (2015) allow for different objective and subjective factor dynamics but they assume that agents are rational in the long run and, thus, they impose a restriction on the long-run mean of the factors. In our Gaussian framework, the long-run mean can be defined as $\mathbf{m}^{\mathbb{P}} \equiv (\mathbf{I} - \boldsymbol{\Phi}^{\mathbb{P}})^{-1}\boldsymbol{\mu}^{\mathbb{P}}$ and $\mathbf{m}^{\mathbb{S}} \equiv (\mathbf{I} - \boldsymbol{\Phi}^{\mathbb{S}})^{-1}\boldsymbol{\mu}^{\mathbb{S}}$ under the objective and subjective probability measure, respectively. Since we can rewrite Eq.(4) as

$$\mathbf{x}_t - \mathbf{m}^{\mathbb{P}} = \boldsymbol{\Phi}^{\mathbb{P}}(\mathbf{x}_{t-1} - \mathbf{m}^{\mathbb{P}}) + \mathbf{u}_t^{\mathbb{P}}, \tag{15}$$

under the long-run mean restriction, $\mathbf{m}^{\mathbf{P}} = \mathbf{m}^{\mathbb{S}}$, we can estimate $\boldsymbol{\Phi}^{\mathbb{P}}$ by OLS conditionally

Model	$\mu_1^{\mathbb{Q}} \times 100$	$eig(\Phi^{\mathbb{Q}})$	
Case 0 (yields-only model)	0.0567	0.9981	0.8491±0.0118
Case 1 ($\mathbb{S} \sim \mathbb{P}$, no s_t)	0.0565	0.9982	0.8542±0.0372
Case 2 ($\mathbb{S} \sim \mathbb{P}$, s_t is $\mathbb{P}-$ risk factor)	0.0661	0.9966	0.8372±0.0196
Case 3 ($\mathbb{S} \nsim \mathbb{P}$, s_t is not $\mathbb{P}-$risk factor)	0.0724	0.9966	0.8355±0.0169
Case 4 ($\mathbb{S} \nsim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor)	0.0657	0.9965	0.8357±0.0181

Table 6. Estimates of risk-neutral parameters.

on the $\mathbb{S}$ parameters:

$$\mathbf{x}_t - \mathbf{m}^{\mathbb{S}} = \mathbf{\Phi}^{\mathbb{P}}(\mathbf{x}_{t-1} - \mathbf{m}^{\mathbb{S}}) + \mathbf{u}_t^{\mathbb{P}}. \tag{16}$$

The $p-$value for the likelihood-ratio test with 4 degrees of freedom is 7%, which indicates that the rationality of agents in the long-run cannot be rejected at a 5% significance level. A closer inspection at the likelihood components, however, reveals that the high overall likelihood value for this case is achieved by a closer fit to surveys forecasts (higher $L^{\mathbb{S}}$) at a cost of worse likelihood of the $\mathbb{P}$ dynamics. Indeed, in unreported results (available upon request), the improvement in the survey fit is about 0.01 basis point as measured by cross-sectional $RMSE$. In Appendix B we confirm that this model specification has inferior predictive performance relative to Case 4, which renders it unreliable for the purposes of the objective risk premium decomposition, and therefore we are not going to consider it further.

5.3 Factor dynamics

We now inspect the factor dynamics in each of the cases considered. In Table 6 we report the estimates of the risk-neutral parameters. As typically found in the literature, our results show that the factors are highly persistent under the risk-neutral measure. For all cases, the $\mathbb{Q}$ dynamics are close to a unit root process with the largest eigenvalue exceeding 0.996.

In Table 7 we report the estimates of the factor dynamics under the real-world measure.

Model	$\boldsymbol{\mu}^{\mathbb{P}}$	$\boldsymbol{\Phi}^{\mathbb{P}}$				Eigenvalues	$\mu^{\mathbb{P}}_{\infty}$ (%)
Case 0	0.0017	0.9816	0.0534	-0.4125		0.9866	-0.75
yields-only model	-0.0024	0.0055	0.9278	0.4859		0.9199	
	0.0014	0.0003	-0.0007	0.7535		0.7564	
Case 1	0.0061	0.9759	0.1231	-0.7506		0.9544	5.97
$\mathbb{S} \sim \mathbb{P}$	0.0001	-0.0073	0.9268	0.3793		0.9322	
no s_t	0.0006	0.0025	-0.0036	0.7125		0.7285	
Case 2	0.0036	0.9999	-0.1641	-2.0787	1.7099	0.9692	3.87
$\mathbb{S} \sim \mathbb{P}$	-0.0002	-0.0050	0.9123	0.3043	0.0837	0.9429	
s_t is $\mathbb{P}-$risk factor	-0.0000	0.0035	0.0163	0.7438	-0.0306	0.7010	
	0	-0.0114	0.1902	0.8211	0.0368	0.0796	
Case 3	0.0013	0.9804	0.0581	-0.3223	0	0.9885	-1.66
$\mathbb{S} \nsim \mathbb{P}$	-0.0025	0.0087	0.9297	0.3999	0	0.9215	
s_t is not $\mathbb{P}-$risk factor	0.0014	-0.0011	-0.0010	0.7973	0	0.7974	
	0.0020	-0.0124	0.1262	0.5715	0.2733	0.2733	
Case 4	-0.0043	1.0188	-0.1557	-1.5087	1.6478	0.9891	-1.83
$\mathbb{S} \nsim \mathbb{P}$	-0.0023	0.0064	0.9343	0.4632	-0.4632	0.9007	
s_t is $\mathbb{P}-$risk factor	0.0017	-0.0023	0.0061	0.8200	-0.0608	0.7794	
	0.0025	-0.0191	0.1235	0.5757	0.2043	0.3082	

Table 7. Estimates of factor (shadow PC) dynamics under the real-world probability measure. The last column reports the long-run mean of the shadow short rate in percentage points.

The penultimate column reports the eigenvalues of the feedback matrix $\boldsymbol{\Phi}^{\mathbb{P}}$. In all cases, the eigenvalues are within the unit circle, which indicates stationarity. The highest eigenvalue for the model estimated only with yields (Case 0) amounts to 0.987, a value typical in the term structure literature. If in the estimation process we include interest rates surveys and impose the rational expectation assumption, the persistence of interest rates decreases considerably, to 0.954 and 0.969 for Case 1 and Case 2, respectively. If we allow for different objective and subjective probability measures in Case 3 and Case 4, the physical dynamics are freed and become similar to those obtained from the yields-only model, with the value of the highest eigenvalue of about 0.989.

In the last column of Table 7 we report the long-run mean of the shadow short rate under the real-world probability measure, which, for cases with the survey factor, can be found as

$$\mu_{\infty}^{\mathbb{P}} = [\mathbf{1}',\ 0] \left(\mathbf{I} - \mathbf{\Phi}^{\mathbb{P}}\right)^{-1} \boldsymbol{\mu}^{\mathbb{P}}. \tag{17}$$

When there is no survey factor (Case 0), the formula is the same without a zero in the square bracket.

When surveys are not included in the estimation, the physical long-run mean is estimated at −0.75%, due to the protracted downward trend of interest rates in our sample. Including surveys in the estimation process and imposing rational expectations turns the estimate of the long-run mean to a positive value, 5.97% and 3.87% for Case 1 and Case 2, respectively. These estimates are negative again when the physical measure and subjective measure are allowed to be different and amount to −1.66% and −1.83% for Case 3 and Case 4, respectively.

Results in Table 7 suggest that survey expectations exhibit lower persistence than measured under the real-world measure,[12] but also that they dominate the physical dynamics if the rational expectation hypothesis is maintained. Estimates for the subjective dynamics, reported in Table 8, confirm this conjecture. The highest eigenvalue for subjective dynamics for Case 3 and Case 4 amounts to 0.97, about the same as for Case 2. Generally, the point estimates of dynamics for Case 3 and Case 4 under the subjective measure are very similar to each other. Also the estimated long-run mean under the subjective measure is similar in the two cases, about 3.7%.

[12] This is in line with the finding in Gourinchas and Tornell (2004) that agents misperceive interest rate shocks to be more transitory than what they actually are.

Model	$\boldsymbol{\mu}^{\mathbb{S}}$	$\boldsymbol{\Phi}^{\mathbb{S}}$				Eigenvalues	$\mu_{\infty}^{\mathbb{S}}$ (%)
Case 3	0.0036	0.9880	-0.2489	-2.2179	1.8585	0.9718	3.70
$\mathbb{S} \nsim \mathbb{P}$	-0.0002	-0.0050	0.9019	0.2901	0.1062	0.9429	
s_t is not $\mathbb{P}-$risk factor	-0.0001	0.0045	0.0211	0.6939	-0.0239	0.6515	
	0	-0.0043	0.2217	0.7826	0.0448	0.0624	
Case 4	0.0034	1.0006	-0.1929	-2.1262	1.7365	0.9689	3.65
$\mathbb{S} \nsim \mathbb{P}$	-0.0002	-0.0051	0.9098	0.3179	0.0858	0.9475	
s_t is $\mathbb{P}-$risk factor	-0.0000	0.0042	-0.0179	0.6896	-0.0190	0.6549	
	0	-0.0108	0.2093	0.8015	0.0289	0.0576	

Table 8. Estimates of factor (shadow PC) dynamics under the subjective probability measure. The last column reports the long run mean of the shadow short rate in percentage points.

5.4 Yield forecasts

5.4.1 In-sample

In this section we evaluate the predictive performance for different model specifications. First, we focus on in-sample forecasts. For all periods from 1983:Q1 to 2019:Q3 we forecast 0.25, 1, 5 and 10−year yields for 1 through 4 quarters ahead, so that for each forecast horizon we have a time series of 147 predictions.

In Table 9 we report the root mean-square forecasting error ($RMSFE$) for forecasts made with the physical dynamics of different model specification. For completeness, we also report the $RMSFE$ statistic for the random walk forecasts, which is a common benchmark for predictive exercises in the term structure literature. We highlight in bold the smallest $RMSFE$ for each yield and at each forecasting horizon. The most flexible model with different physical and subjective dynamics, and with the $s-$factor (Case 4) clearly dominates other specifications by providing the smallest $RMSFE$ for all yields and forecast horizons with the exception only of the 10−year yield at 3 quarters ahead (best $RMSFE$ obtained from Case 0) and the 10−year yield at 4 quarters ahead (best $RMSFE$ obtained from Case 3), but in both cases the difference in the $RMSFE$ is not economically meaningful. With the

Model	Horizon\Yield	3m	1y	5y	10y
Case 0	1q	45.93	51.76	51.65	48.04
(yields-only model)	2q	74.04	80.40	73.49	67.06
	3q	96.80	101.20	86.42	**78.33**
	4q	119.48	123.49	101.88	99.49
Case 1	1q	45.93	53.30	53.48	48.92
($\mathbb{S} \sim \mathbb{P}$, no s_t)	2q	76.21	84.75	78.84	70.08
	3q	101.97	109.95	96.67	84.69
	4q	128.41	137.17	117.23	100.43
Case 2	1q	31.73	35.66	40.82	41.06
($\mathbb{S} \sim \mathbb{P}$, s_t is $\mathbb{P}-$ risk factor)	2q	63.69	73.74	73.46	67.46
	3q	92.37	103.01	94.54	84.79
	4q	119.57	131.80	115.01	99.91
Case 3	1q	45.95	51.78	51.70	48.05
($\mathbb{S} \nsim \mathbb{P}$, s_t is not $\mathbb{P}-$risk factor)	2q	74.62	80.48	73.57	67.12
	3q	97.47	101.34	86.53	78.47
	4q	119.94	123.60	102.05	**90.70**
Case 4	1q	**31.58**	**34.43**	**39.40**	**39.49**
($\mathbb{S} \nsim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor)	2q	**61.62**	**69.18**	**69.58**	**65.08**
	3q	**86.23**	**92.28**	**85.37**	79.66
	4q	**107.52**	**113.23**	**98.96**	90.71
Random walk	1q	49.12	52.83	52.39	48.55
	2q	81.68	83.59	75.90	68.70
	3q	109.87	108.28	91.18	81.96
	4q	136.58	134.28	109.15	96.89

Table 9. Root mean-square forecasting error for in-sample model-implied objective expectations and the random walk.

aforementioned exceptions, the forecast precision delivered by other specifications is inferior to Case 4, with $RMSFE$ often higher by more than 10 basis points at both short and long forecast horizons.

The models with the rational expectations assumption, Case 1 and Case 2, generally produce the worst predictions, especially at long horizons. At the 1 quarter ahead horizon, the forecasts produced by Case 2 are worse but not far from Case 4, which can indicate the importance of the unspanned $s-$factor for short-run dynamics. Nonetheless, the results confirm our conclusions from previous sections that forcing the model to fit survey expectations with real-world dynamics distorts the latter. As such, this is another manifestation of the rejection of the rational expectations hypothesis by our model. In fact, it would be preferable to estimate the model without surveys (Case 0) rather than try to learn about the objective dynamics of interest rates from surveys expectations assuming rational expectations (Case 1 and Case 2).

A separate question is the role of the survey-specific factor. Its importance for yield predictions is particularly pronounced for short-horizon forecasts: the specifications without the $s-$factor driving the $\mathbb{P}$ dynamics (Case 1 and Case 3) attain the $RMSFE$ about $10-15$ basis points worse than their counterparts that allow for this interaction (Case 3 and Case 4, respectively) for $1-$quarter ahead forecasts. At longer forecast horizons this effect remains strong for the $3-$month and $1-$year yields but it dissipates for longer maturity yields.

These results indicate that to obtain reliable yield forecasts using the information in interest rates surveys, we should a) allow for different objective and subjective factor dynamics, and b) allow the survey-specific factor to drive the real-world dynamics. Both of these elements seem to play important role in the determination of the real-world dynamics.

Moving to the subjective dynamics, in Figure 2 we report survey expectations for the 1-year yield, the subjective expectations from Case 4 and the realised 10-year yield. Contrary to Figure 1, we can see that our maximally flexible model is able to replicate the survey expectations, further validating the evidence that subjective dynamics are different from objective ones. However, given the strong rejection of the rational expectations hypothesis by our model, we should not expect subjective dynamics to prove useful in the predictive exercise.

To corroborate this conjecture, in Table 10 we report the $RMSFE$ for observed survey forecasts and for the forecasts made using subjective expectations from Case 3 and Case 4.

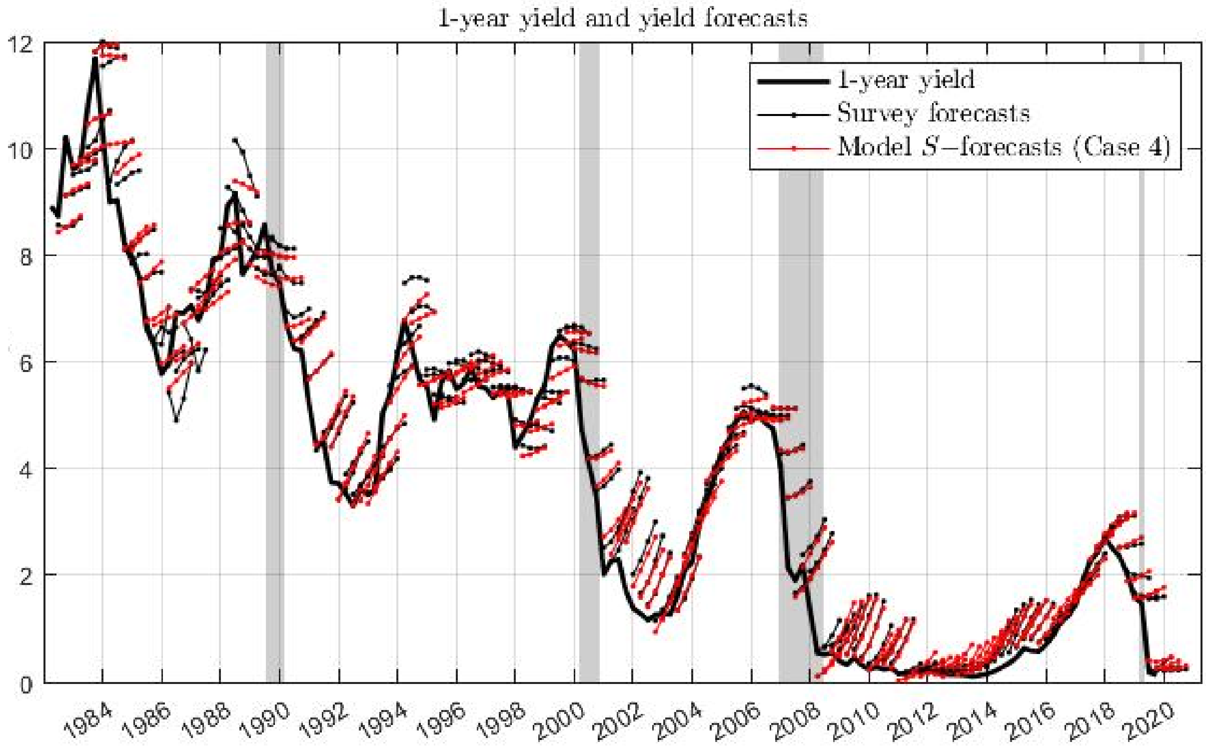

Figure 2. Market expectations measured by surveys (black-dotted lines) and in-sample subjective expectations extracted from the Case 4 model with different subjective and objective probability measures and with a priced survey-specific factor (red-dotted lines), along with the observed 1-year yield (bold black line)

First, confirming our conjecture regarding the subjective probability measure, we observe that the $RMSFE$ for models specified in Case 3 and Case 4 are much larger than the forecasts made with the physical dynamics. However, for three yields the survey forecasts at the 1−quarter horizon produce $RMSFE$ slightly lower than the model-generated predictions from Table 9. For longer horizons, however, surveys perform significantly worse than $\mathbb{P}$ forecasts produced by our preferred specification in Case 4.

Finally, we note that the forecasts obtained with real-world dynamics in models specified in Case 0 and Case 4 reported in Table 9 generally comfortably outperform the random walk

Model	Horizon\Yield	3m	1y	5y	10y
Surveys	1q	**26.56**	36.89	**37.56**	**38.91**
	2q	61.81	76.59	73.04	65.53
	3q	92.83	104.66	93.95	83.65
	4q	122.67	134.02	113.82	100.10
Case 3 (subjective)	1q	31.85	35.65	40.95	41.47
($\mathbb{S} \nsim \mathbb{P}$, s_t is not $\mathbb{P}-$risk factor)	2q	63.60	73.89	73.60	67.67
	3q	92.26	103.38	94.75	84.94
	4q	119.58	132.33	115.19	99.99
Case 4 (subjective)	1q	31.73	35.61	40.90	41.34
($\mathbb{S} \nsim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor)	2q	63.55	73.88	73.56	67.56
	3q	92.25	103.36	94.73	84.87
	4q	119.61	132.31	115.21	99.97

Table 10. Root mean-square forecasting error for surveys and in-sample model-implied subjective expectations.

benchmark, with particularly large margins at longer forecast horizons.

5.4.2 Out-of-sample

Next, we turn to the results of the out-of-sample forecasting exercise. We use the expanding-window scheme on the last 15 years of our data sample. In particular, using the initial sample from 1983:Q1 to 2004:Q4 (88 observations) we estimate the model and make forecasts (under the $\mathbb{P}$ measure) for the next 4 quarters. Next, we add the data on 2005:Q1 (89 observations) and re-estimate the model to make predictions for the following 4 quarters. We repeat this procedure until 2019:Q3, so that we obtain 60 predictions for each forecast horizon.

In Table 11 we report the $RMSFE$ for the out-of-sample forecasts. As a benchmark we also report the $RMFSE$ produced by the random walk model and with bold font we highlight the smallest $RMFSE$ across all cases. Similarly to the in-sample results, predictions generated by Case 4 have generally the smallest $RMFSE$ among other model specifications. The two exceptions are forecasts of the 3−month and 5−year yields at the 1−quarter hori-

Model	Horizon\Yield	3m	1y	5y	10y
Case 1	1q	33.31	34.28	38.50	40.74
($\mathbb{S} \sim \mathbb{P}$, no s_t)	2q	58.96	61.55	61.16	58.86
	3q	84.92	88.43	80.27	74.74
	4q	110.62	116.68	101.61	90.70
Case 2	1q	**25.56**	28.46	**31.97**	37.54
($\mathbb{S} \sim \mathbb{P}$, s_t is $\mathbb{P}-$ risk factor)	2q	47.32	55.85	58.16	55.15
	3q	73.75	85.73	81.91	73.44
	4q	98.76	114.48	101.63	86.90
Case 3	1q	33.95	34.04	40.35	42.45
($\mathbb{S} \nsim \mathbb{P}$, s_t is not $\mathbb{P}-$risk factor)	2q	59.14	59.16	60.55	59.57
	3q	83.55	82.03	74.97	72.14
	4q	105.87	105.09	91.55	84.98
Case 4	1q	26.34	**26.53**	32.41	**36.57**
($\mathbb{S} \nsim \mathbb{P}$, s_t is $\mathbb{P}-$risk factor)	2q	**46.31**	**48.61**	**54.28**	**53.80**
	3q	**67.50**	**69.08**	71.21	68.90
	4q	**84.91**	**86.76**	83.22	79.08
Random walk	1q	39.54	36.75	37.33	39.60
	2q	68.99	62.46	55.69	54.38
	3q	95.06	85.13	**67.77**	**64.95**
	4q	118.04	107.37	**81.49**	**75.54**

Table 11. Out-of-sample $RMSFE$ for model-implied objective expectations for the period from 2004:Q4 to 2019:Q3. Entries in bold are the lowest for the specific horizon and yield.

zon, for which the smallest $RMSFE$ was attained by Case 2, but the difference is marginal, smaller than 1 basis point. However, in all other cases, Case 4 produces more accurate predictions than other specifications, especially at long forecast horizons. In particular, the $RMSFE$ for the 4$-$quarters ahead forecasts produced by Case 4 are about $10 - 30$ basis points smaller than obtained by other specifications. Case 4 also produces mostly smaller $RMSFE$ than the random walk benchmark except for the 5 and 10$-$year yields for the forecasts 3 and 4 quarters ahead.

Having found that Case 4 delivers predictions superior in economic terms, we turn to the statistical significance of differences in predictability. In Table 12 we report the ratio of the

Benchmark	Horizon\Yield	3m	1y	5y	10y
Case 1	1q	0.791*	0.774*	0.842*	0.898*
($\mathbb{S} \sim \mathbb{P}$, no s_t)	2q	0.785**	0.790**	0.888	0.914
	3q	0.795**	0.781**	0.887	0.922
	4q	0.768**	0.744**	0.819*	0.872
Case 2	1q	1.031	0.932	1.014	0.974
($\mathbb{S} \sim \mathbb{P}$, s_t is $\mathbb{P}-$ risk factor)	2q	0.979	0.870**	0.933	0.976
	3q	0.915**	0.806**	0.869**	0.938
	4q	0.860**	0.758**	0.819**	0.910*
Case 3	1q	0.776*	0.780*	0.803*	0.862*
($\mathbb{S} \nsim \mathbb{P}$, s_t is not $\mathbb{P}-$risk factor)	2q	0.783**	0.822*	0.897	0.903
	3q	0.808*	0.842	0.950	0.955
	4q	0.802*	0.826*	0.909	0.931
Random walk	1q	0.666*	0.722*	0.868*	0.924
	2q	0.671**	0.778*	0.975	0.989
	3q	0.710**	0.811**	1.051	1.061
	4q	0.719**	0.808**	1.021	1.047

Table 12. Out-of-sample forecasting performance of Case 4. The table reports out-of-sample $RMSFE$s for the model-implied objective expectations of Case 4 relative to the ones of the other cases and the random walk. The evaluation period is 2004:Q4 to 2019:Q3. ***, ** and * denote one-side significance at 1%, 5% and 10% level of the Diebold and Mariano (1995) test of equal predictive ability using fixed$-b$ asymptotics as in Coroneo and Iacone (2020).

$RMSFE$ of the forecasts generated by Case 4 to the $RMSFE$ of other specifications. A value smaller than 1 means that the forecasts generated by Case 4 are more accurate, as is indeed the case for most maturities and forecasting horizons, as also noted in Table 11.

To assess whether the predictive performance of Case 4 is statistically different than other specifications, we apply the Diebold and Mariano (1995) test of equal predictive ability using fixed-b asymptotics as in Coroneo and Iacone (2020). This approach allows us to obtain correctly sized tests with our small sample (60 out-of-sample observations) and in the presence of autocorrelated forecast errors.[13]

[13] As in Coroneo and Iacone (2020), we use a quadratic loss and a weighted covariance estimate for the long-run variance with Bartlett kernel and truncation $\lfloor T^{1/2} \rfloor$ (which equals 7 for our out of sample size of

The null hypothesis is that two models produce forecasts that are not statistically different, and the alternative is that the benchmark (Case 4) produces more accurate forecasts. As shown in Table 12, the null hypothesis is frequently rejected, especially for 3-months and $1-$year yields. The rejection for yields with maturities of 5 and 10 years is less frequent. Regarding comparison with the random walk model, although, as indicated earlier, random walk predictions produce the smallest $RMSFE$ for 5 and $10-$year yields at the 3 and $4-$quarters horizon, Case 4 provides significantly more accurate forecasts for the 3-months and $1-$year yields at all forecasting horizons.

5.5 Risk premium

In this section we examine the model's predictions in terms of the decomposition of interest rates into expected future rates and risk premium.

Focusing on the real-world probability measure first, in Figure 3, we plot the objective $10-$year risk premium calculated as in Eq.(6) for different model specifications. For reference, we also plot the $10-$year yield. To get a more nuanced view, we split the plots of the term premium into two subperiods: 1983-2000 (top panel) and 2001-2020 (bottom panel). The most visible feature in this figure, which also reflects our discussion in previous sections, is that imposing rational expectations distorts objective probabilities which, in turn, results in anomalous risk premia for Case 1 and Case 2. Indeed, the risk premia for these rational expectation specifications are very different from the premia implied by other models: they are much lower, starting at about 3% in 1983, they gradually decline, reach zero around 2005, and after 2010 stay negative for the rest of the sample. This is due to the relatively low persistence of interest rates implied by surveys expectations, so the model predicts that interest rates would quickly return to their long-run mean of about $4-6\%$, see Table 7.

60 observations).

Since the 10−year yield drops below 4% after 2010, the risk-neutral yield $y_{10}^{\mathbb{P}}$ becomes higher than the observed yield and the risk premium, being the residual in this relation, becomes negative.

Although relative to the models in Case 1 and Case 2 that impose rational expectations the risk premia implied by the three other cases in Figure 3 seem to be close to each other, a careful inspection reveals that there are important differences in the risk premium dynamics implied by these models, which underlines the importance of accounting for the survey-specific factor as a $\mathbb{P}-$risk factor. As expected, since the real-world dynamics of the model estimated only with yields (Case 0) and the model with different probability measures but with the risk premium restriction on the $s-$factor (Case 3) are very similar, the implied risk premia for these two specifications are nearly identical throughout the whole sample period. The risk premium implied by our preferred, fully-fledged model (Case 4), however, follows the same secular trend but exhibits locally very different dynamics. This is particularly visible in the second part of the sample presented in the bottom panel of Figure 3. First, it exhibits clear countercyclical behaviour, increasing substantially in all three recession episodes in the 21st century. Although the term premia implied by Case 0 and Case 3 also increase in the dot-com recession in 2001 and the financial crisis 2008-2009, they decline throughout the coronavirus recession in 2020. Second, following the recessionary episodes, during the expansionary periods the term premium of Case 4 declines faster than those implied by other specifications and gradually becomes larger at the later stages of the business cycle. For example, although during the peak of the great financial crisis at the end of 2008 the term premia implied by all cases with flexible $\mathbb{P}$ dynamics were about the same (about 3.3%), at the end of 2010 the Case 4 term premium amounted to 1.2%, about 80 basis points lower than those of Case 0 and Case 3. On the other hand, when the coronavirus pandemic shook the global economy in the first quarter of 2020, the Case 4 term premium increased from

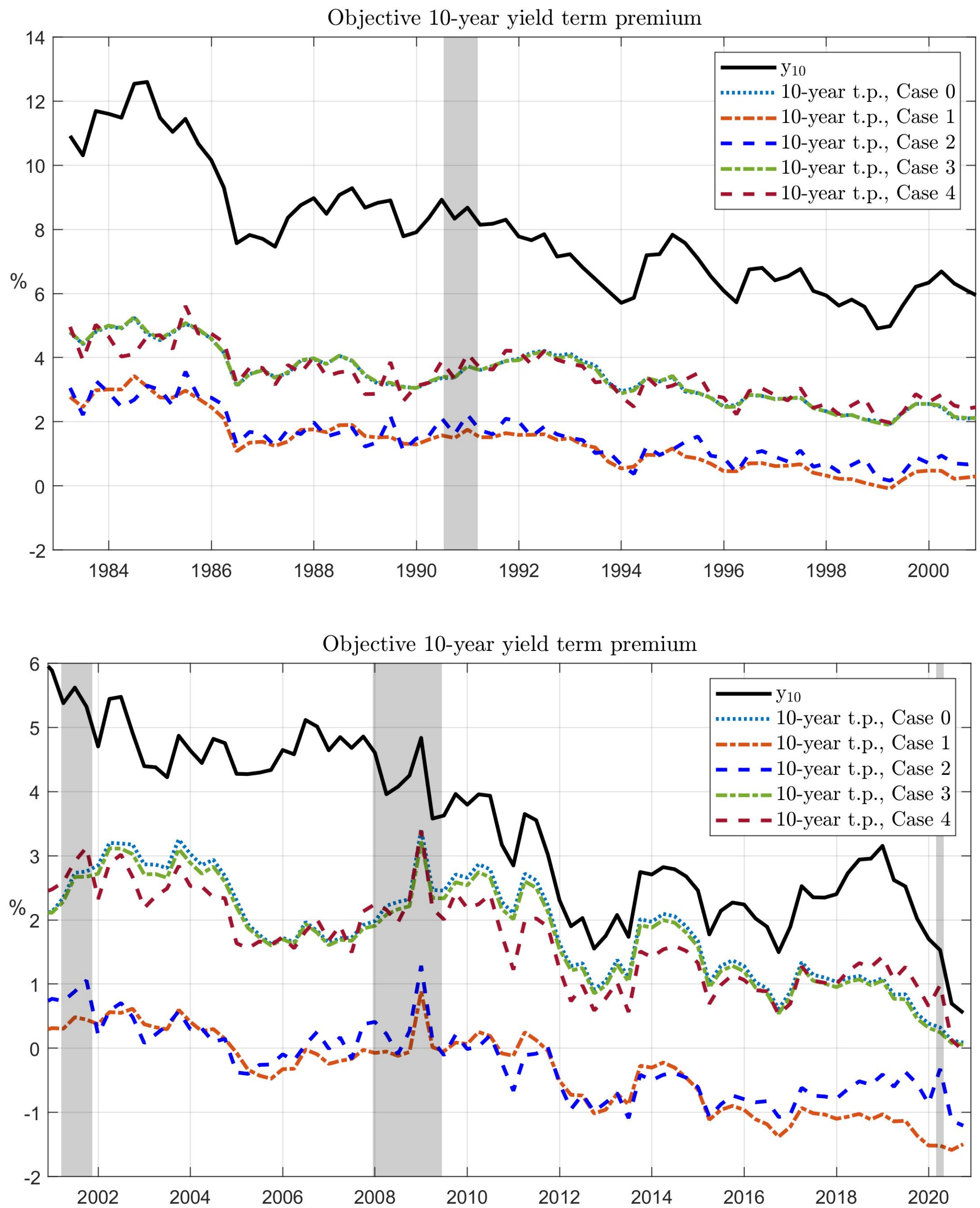

Figure 3. Objective 10−year risk premia implied by different model specifications for the period from 1983:Q1 to 2000:Q4 (top panel) and from 2001:Q1 to 2020:Q3 (bottom panel).

0.65% to 1%, while those of Case 0 and Case 3 declined from 0.30% to 0.25%.

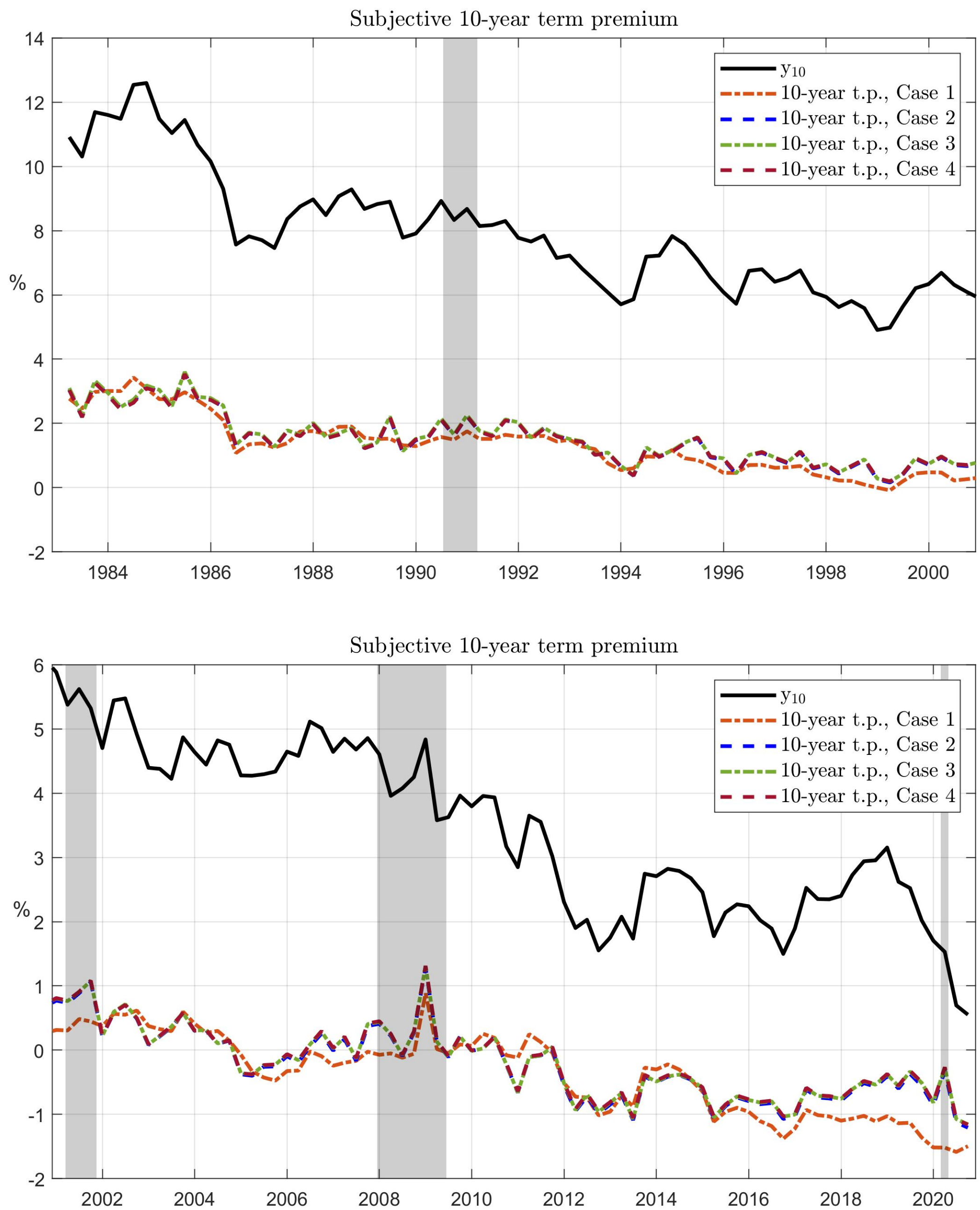

Figure 4. Subjective 10−year risk premia implied by different model specifications for the period from 1983:Q1 to 2000:Q4 (top panel) and from 2001:Q1 to 2020:Q3 (bottom panel).

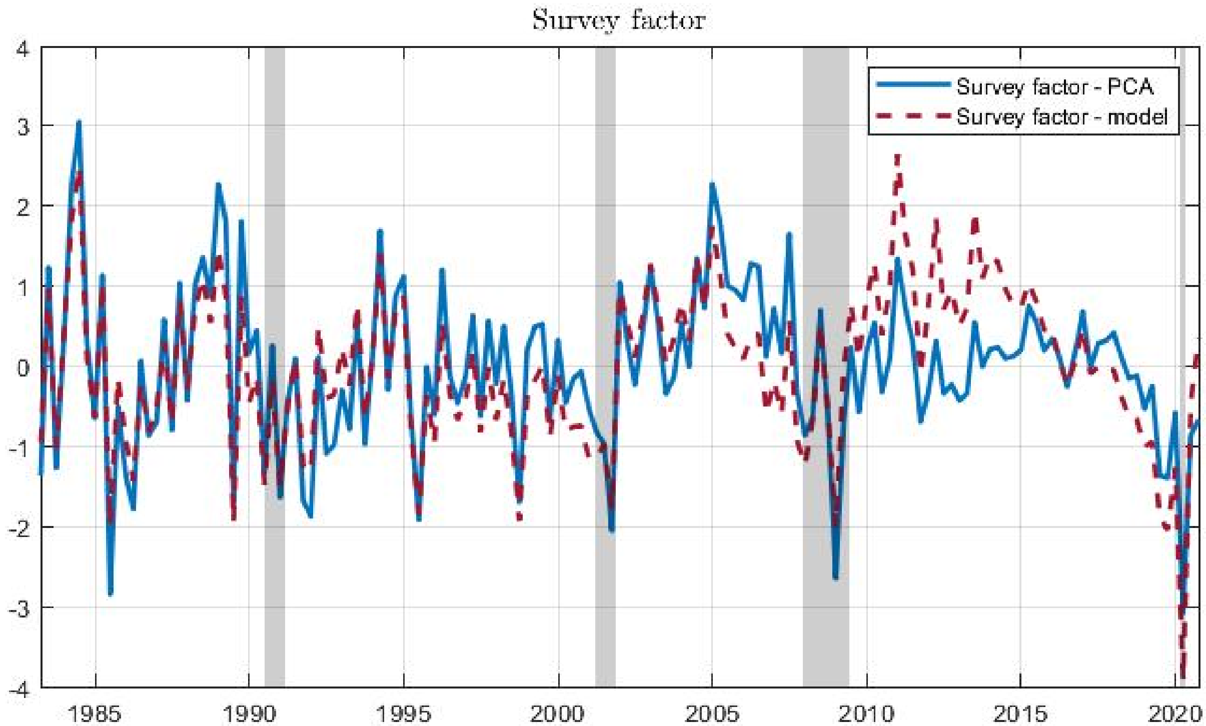

Figure 5. $PC-$based $s-$factor estimated in Section 2 (continuous blue line) and the $s-$factor extracted from the term structure model in Section 3 specified as Case 4 (red dashed line). Shaded areas denote the NBER recessions.

In Figure 4 we plot the subjective risk premia for different model specifications throughout the sample period. Naturally, the subjective risk premia for the models with rational expectations constraint, Case 1 and Case 2, are the same as presented in Figure 3. However, due to the extremely strong identification of subjective dynamics parameters, subjective expectations for all models are virtually the same, which is manifested by overlapping lines for Cases 2, 3 and 4. The estimates of the risk premium for Case 1 are slightly different since this model omits the survey-specific factor in the dynamics equation (8).

5.6 A closer look into the unspanned survey factor

In Figure 5 we plot the time series of the $PC-$based $s-$factor estimated in Section 2 (continuous blue line) and the $s-$factor extracted from the term structure model in Section 3 specified as Case 4 (red dashed line) with shaded areas denoting the NBER recessions.[14] The two time series comove together throughout the whole sample (the correlation amounts to 0.83) with a notable exception of the zero lower bound period 2009-2015. If the $s-$factor picks up just random correlation in noise in surveys, we should not expect it to have any business cycle variation. However, the cyclical nature of the survey factor is striking: the $s-$factor increases during economic expansions and strongly declines in recessions, such as in 1990, 2001, 2007-2008 and 2020. Other large negative spikes, although not related to official recession periods, can also be identified with important economic and political events, such as the 1985:Q2 collapse of Home State Savings Bank that set off a series of savings-and-loan closures across the US; the fall of communism in Central and Eastern Europe in 1989:Q2; dissolution of the Soviet Union in 1991:Q4; Clinton's bailout to Mexico in response to the peso crisis in 1995:Q2; Russian financial crisis in 1998:Q3. This indicates that there is a systematic factor in surveys, unspanned by interest rates, that contains information about the business cycle.

To investigate this possibility, we analyze the determinants of the unspanned survey factor by regressing estimates of the s-factor on a set of potential proxies. In particular, in Table 13 we report results for the estimated s-factor from the joint term structure model in Section 3 specified as Case 4, and in Table 14 we report results for the $PC-$based s-factor estimated in Section 2. We use as regressors a set of variables that include: industrial production growth and CPI inflation (which are commonly used in macro-finance models of

[14] It should be kept in mind that the factor sign is unidentified, so our choice of the pro-cyclical character of the $s-$factor is arbitrary.

Table 13. Determinats of the model-based $s-$factor

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IP	13.084**						8.830**
	(5.246)						(4.451)
CPI		15.884					11.919
		(15.731)					(16.059)
VIX			-0.034**				-0.008
			(0.014)				(0.015)
EPU				-0.002*			0.002
				(0.001)			(0.001)
EPU-MP					-0.006***		-0.004**
					(0.002)		(0.002)
EM-EUI						-0.823**	-0.577*
						(0.330)	(0.300)
Constant	-0.062	-0.103	0.703**	0.280*	0.571***	0.320*	0.519
	(0.137)	(0.189)	(0.313)	(0.165)	(0.198)	(0.175)	(0.389)
Observations	151	151	120	140	140	140	120
R^2	0.034	0.002	0.067	0.034	0.135	0.084	0.166

The dependent variable is the estimated s-factor from the model in Section 3 specified as Case 4. The regressors are: the Industrial Production growth rate (IP), the Consumer Price Index growth rate (CPI), the CBOE Volatility Index (VIX), the Economic Policy Uncertainty index (EPU), the Economic Policy Uncertainty Index: Monetary policy (EPU-MP), and the Equity Market-related Economic Uncertainty Index (EM-EUI). All variables are quarterly from 1983:Q1 (when available) to 2020:Q3. Newey-West standard errors with 4 lags in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the term structure), the VIX index, the Economic Policy Uncertainty index of Baker, Bloom and Davis (2016) (both the full index and the categorical index for monetary policy) and also the Equity Market-related Economic Uncertainty Index (also from Baker et al. 2016).

Results indicate that, when considering each regressor individually, the Economic Policy Uncertainty categorical index for monetary policy explains the largest proportion of the variance of both estimates of the s-factor. In particular, it explains 13.5% of the variance

Table 14. Determinants of the $PC-$based $s-$factor

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IP	17.063***						4.664
	(4.447)						(4.032)
CPI		41.314***					23.659*
		(12.765)					(13.451)
VIX			-0.047***				-0.018*
			(0.011)				(0.010)
EPU				-0.004***			0.000
				(0.001)			(0.001)
EPU-MP					-0.007***		-0.005***
					(0.002)		(0.002)
EM-EUI						-0.624**	-0.118
						(0.257)	(0.213)
Constant	-0.081	-0.268*	0.880***	0.469***	0.646***	0.219	0.657**
	(0.107)	(0.141)	(0.253)	(0.154)	(0.156)	(0.139)	(0.296)
Observations	151	151	120	140	140	140	120
R^2	0.062	0.051	0.179	0.140	0.214	0.052	0.286

The dependent variable is the $PC-$based s-factor estimated in Section 2. The regressors are: the Industrial Production growth rate (IP), the Consumer Price Index growth rate (CPI), the CBOE Volatility Index (VIX), the Economic Policy Uncertainty index (EPU), the Economic Policy Uncertainty Index: Monetary policy (EPU-MP), and the Equity Market-related Economic Uncertainty Index (EM-EUI). All variables are quarterly from 1983:Q1 (when available) to 2020:Q3. Newey-West standard errors with 4 lags in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

of the model-based estimate and 21.4% of the variance of the non-parametric estimate. The adjusted$-R^2$ does not improve much when other regressors are added in the last column of the two tables, as it reaches 16.6% and 28.6%.

Other variables that seem related to the s-factor are the VIX and the other two uncertainty indices, the EPU and the EM-EUI. The VIX explains, respectively, the third and the second largest share of the model-based and the non-parametric estimate of the s-factor. The equity market-related uncertainty index explains the second largest share of the variance of the model-based s-factor and the economic policy uncertainty index explains the third largest share of the non-parametric estimate of the s-factor. The macro indicators, while significant in some cases, do not seem to have an important role, as they explain at most 6.2% of the variance of the s-factor estimates.

These results indicate that the s-factor is more related to uncertainty than to macroeconomic fundamentals. This is in line with the findings in Tillmann (2020) that the effects of monetary policies designed to reduce long-term bond yields can become less effective if monetary policy uncertainty is high. In particular, his finding that term premia would increase and partly offset the stimulating policy impulse seems in line with our results about the presence of a priced unspanned survey factor related to monetary policy uncertainty.

6 Conclusion

One of the popular approaches to deal with the poor identification of the parameters that determine the dynamics of interest rates and thus the risk premium is to augment the term structure model of interest rates with expectations obtained from interest rates surveys (see Kim and Wright 2005, Kim and Orphanides 2012, d'Amico et al. 2018). It has been assumed offhand that the resulting decomposition can be taken as *both* the risk premium demanded by investors for holding long-range Treasury securities *and* the expectation of the short interest rate with respect to the real-world probability measure. Unfortunately, the results presented in this paper show that such conclusions are unwarranted.

To examine the properties of interest rates surveys, we develop a joint model of interest rates and surveys that allows for separate objective and subjective probability measures. Contrary to earlier approaches postulated in the literature, our results indicate that surveys do not help to identify the parameters of physical dynamics. On the contrary, the model overwhelmingly rejects the equality of subjective and objective dynamics, which can be interpreted as evidence against the rational expectations hypothesis. Yet, we find that under both objective and subjective probability measures there exists a common risk factor that is not spanned by observed interest rates. Modelling jointly interest rates and interest rates surveys allows us to extract this common unspanned factor within the model. Our results suggest that taking into account the survey-specific factor is important for a more reliable measurement of the risk premium and interest rate forecasts.

Undoubtedly, our results should be taken as the first step towards deeper understanding of the role of surveys as a gauge of market expectations. It should be noted that we rely solely on the consensus forecasts from Blue Chip Financial Forecasts. It is plausible that a deeper examination of the granular structure of individual forecasts could reveal more information. In particular, an important development for future research should explain the underlying causes of the divergence between the objective and subjective probability measure of a representative (consensus) agent. The possible explanation might involve microfounded aggregation results with conditioning information or some market microstructure effects, along the lines of Giacoletti, Laursen and Singleton (2021), Singleton (2021) and Buraschi et al. (2022), but focussed on understanding the superior out-of-sample predictability under the objective, not subjective, probability measure.

Also, it is plausible that the survey-specific factor documented in this paper, being a systematic risk factor in the bond market, is also a pricing factor in other asset classes, such as equities or derivatives. This conjecture can be supported by our finding that the survey factor is related to economic uncertainty which is generally market-wide. Also, since the survey factor contributes to the subjective risk premium, it might be used for monitoring market's perceptions of the monetary policy. These are all interesting directions for future research.

Finally, a natural extension of the model presented in this paper could involve a joint real and nominal term structure model with survey expectations of macro variables. A related application involves a joint term structure model of inflation swaps with survey inflation expectations. The latter is particularly relevant in the current times of elevated inflation. Further research in this direction is currently underway.

References

Adam, Klaus, Albert Marcet, and Johannes Beutel (2017) 'Stock price booms and expected capital gains.' *American Economic Review* 107(8), 2352–2408

Adam, Klaus, and Albert Marcet (2011) 'Internal rationality, imperfect market knowledge and asset prices.' *Journal of Economic Theory* 146(3), 1224–1252

Ang, Andrew, and Monika Piazzesi (2003) 'A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables.' *Journal of Monetary Economics* 50(4), 745–787

Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2016) 'Measuring economic policy uncertainty.' *The Quarterly Journal of Economics* 131(4), 1593–1636

Bauer, Michael D, Glenn D Rudebusch, and Jing Cynthia Wu (2012) 'Correcting estimation bias in dynamic term structure models.' *Journal of Business & Economic Statistics* 30(3), 454–467

Black, Fischer (1995) 'Interest rates as options.' *Journal of Finance* 50(5), 1371–1376

Buraschi, Andrea, Ilaria Piatti, and Paul Whelan (2022) 'Subjective bond returns and belief aggregation.' *The Review of Financial Studies* 35(8), 3710–3741

Chun, Albert Lee (2011) 'Expectations, bond yields, and monetary policy.' *The Review of Financial Studies* 24(1), 208–247

Cieslak, Anna (2018) 'Short-rate expectations and unexpected returns in Treasury bonds.' *The Review of Financial Studies* 31(9), 3265–3306

Coibion, Olivier, and Yuriy Gorodnichenko (2015) 'Information rigidity and the expectations formation process: A simple framework and new facts.' *American Economic Review* 105(8), 2644–78

Coroneo, Laura, and Fabrizio Iacone (2020) 'Comparing predictive accuracy in small samples using fixed-smoothing asymptotics.' *Journal of Applied Econometrics* 35(4), 391–409

Coroneo, Laura, Domenico Giannone, and Michele Modugno (2016) 'Unspanned macroeconomic factors in the yield curve.' *Journal of Business & Economic Statistics* 34(3), 472–485

Crump, Richard K., Stefano Eusepi, and Emanuel Moench (2018) 'The term structure of expectations and bond yields.' Staff Reports 775, Federal Reserve Bank of New York

De La O, Ricardo, and Sean Myers (2021) 'Subjective cash flow and discount rate expectations.' *The Journal of Finance* 76(3), 1339–1387

Diebold, Francis X, and Roberto S Mariano (1995) 'Comparing predictive accuracy.' *Journal of Business & Economic Statistics* 13(3), 253–263

Duffee, Gregory R (2011) 'Information in (and not in) the term structure.' *The Review of Financial Studies* 24(9), 2895–2934

d'Amico, Stefania, Don H Kim, and Min Wei (2018) 'Tips from TIPS: the informational content of Treasury Inflation-Protected Security prices.' *Journal of Financial and Quantitative Analysis* 53(1), 395–436

Farmer, Leland, Emi Nakamura, and Jón Steinsson (2021) 'Learning about the long run.' NBER Working Papers 29495

Giacoletti, Marco, Kristoffer T. Laursen, and Kenneth J. Singleton (2021) 'Learning from disagreement in the U.S. Treasury bond market.' *Journal of Finance* 76(1), 395–441

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus (2021) 'Five facts about beliefs and portfolios.' *American Economic Review* 111(5), 1481–1522

Golinski, Adam, and Peter Spencer (2022) 'Unconventional monetary policies and the yield curve: Estimating non-affine term structure models with unspanned macro risk by factor extraction.' Working paper

Gourinchas, Pierre-Olivier, and Aaron Tornell (2004) 'Exchange rate puzzles and distorted beliefs.' *Journal of International Economics* 64(2), 303–333

Greene, W. H. (2011) *Econometric Analysis* (7th Edition, Pearson Education)

Greenwood, Robin, and Andrei Shleifer (2014) 'Expectations of returns and expected returns.' *The Review of Financial Studies* 27(3), 714–746

Hamilton, James D, and Jing Cynthia Wu (2012) 'Identification and estimation of Gaussian affine term structure models.' *Journal of Econometrics* 168(2), 315–331

Joslin, Scott, Kenneth J Singleton, and Haoxiang Zhu (2011) 'A new perspective on Gaussian dynamic term structure models.' *The Review of Financial Studies* 24(3), 926–970

Joslin, Scott, Marcel Priebsch, and Kenneth J Singleton (2014) 'Risk premiums in dynamic term structure models with unspanned macro risks.' *Journal of Finance* 69(3), 1197–1233

Kim, Don H, and Athanasios Orphanides (2012) 'Term structure estimation with survey data on interest rate forecasts.' *Journal of Financial and Quantitative Analysis* 47(1), 241–272

Kim, Don H., and Jonathan H. Wright (2005) 'An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates.' Finance and Economics Discussion Series 2005-33, Board of Governors of the Federal Reserve System (U.S.)

Mankiw, N Gregory, Ricardo Reis, and Justin Wolfers (2003) 'Disagreement about inflation expectations.' *NBER macroeconomics annual* 18, 209–248

Nagel, Stefan, and Zhengyang Xu (2022) 'Dynamics of subjective risk premia.' NBER Working Papers 29803

Nelson, Charles R, and Andrew F Siegel (1987) 'Parsimonious modeling of yield curves.' *The Journal of Business* 60(4), 473–489

Piazzesi, Monika, Juliana Salomao, and Martin Schneider (2015) 'Trend and cycle in bond premia.' Working paper

Priebsch, Marcel A. (2013) 'Computing arbitrage-free yields in multi-factor gaussian shadow-rate term structure models.' Finance and Economics Discussion Series 2013-63, Board of Governors of the Federal Reserve System

Singleton, Kenneth J. (2021) 'Presidential address: How much "rationality" is there in bond-market risk premiums?' *Journal of Finance* 76(4), 1611–1654

Svensson, Lars E. O. (1994) 'Estimating and interpreting forward interest rates: Sweden 1992-1994.' IMF Working Papers 1994/114, International Monetary Fund

Tillmann, Peter (2020) 'Monetary policy uncertainty and the response of the yield curve to policy shocks.' *Journal of Money, Credit, and Banking* 52(4), 803–833

Veronesi, Pietro (2010) *Fixed income securities: Valuation, risk, and risk management* (John Wiley & Sons)

Wu, Jing Cynthia, and Fan Dora Xia (2016) 'Measuring the macroeconomic impact of monetary policy at the zero lower bound.' *Journal of Money, Credit and Banking* 48(2-3), 253–291

Zellner, A. (1962) 'An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias.' *Journal of the American Statistical Association, 57, no 298, 348-368* .

Appendix A Factor extraction

This appendix is based on Golinski and Spencer (2022), please see there for more details.

First, for convenience, we restate the notation. We denote the model-implied zero coupon yields at time t by $\mathbf{y}_t$ and the associated vector of pricing errors by $\mathbf{v}_t \sim N(\mathbf{0}, \mathbf{\Sigma}_y)$, such that the observed yields are

$$\widetilde{\mathbf{y}}_t = \mathbf{y}_t + \mathbf{v}_t. \tag{A-1}$$

We assume that there are $K < N-1$ yields or fixed combinations (or 'portfolios') of yields, given by a fixed $N \times K$ weighting matrix $\mathbf{W}$, that are nevertheless fitted without error:

$$\mathbf{q}_{y,t} \equiv \mathbf{W}'\widetilde{\mathbf{y}}_t = \mathbf{W}'\mathbf{y}_t \tag{A-2}$$

for all t, which is equivalent to assuming $\mathbf{W}'\mathbf{v}_t = \mathbf{0}$. Following Golinski and Spencer (2022), we refer to Eq.(A-2) as the *observability restriction*. As in Joslin et al. (2011), we assume that the first three principal components of yields are observed without error, so $\mathbf{W}$ is a $N \times 3$ matrix that consists of eigenvectors corresponding to the three largest eigenvalues of $cov(\mathbf{y}_t)$.

Denote by $\mathbf{q}_y(\mathbf{x}_t; \mathbf{\Psi})$ the vector-valued function in $\mathbb{R}^K$ that maps the latent state vector $\mathbf{x}_t$ to the observable principal components $\mathbf{q}_{y,t}$, where $\mathbf{\Psi}$ denotes a generic vector of relevant parameters, possibly different in every equation. Also, denote the inverse of $\mathbf{q}_y(\mathbf{x}_t; \mathbf{\Psi})$ by $\mathbf{q}_y^{-1}(\mathbf{q}_{y,t}; \mathbf{\Psi})$, so that:

$$\mathbf{q}_{y,t} = \mathbf{q}_y(\mathbf{x}_t; \mathbf{\Psi}) = \mathbf{W}'\mathbf{y}(\mathbf{x}_t; \mathbf{\Psi}) \Longleftrightarrow \mathbf{x}_t = \mathbf{q}_y^{-1}(\mathbf{q}_{y,t}; \mathbf{\Psi}) = \mathbf{x}(\mathbf{y}_t^o; \mathbf{\Psi}). \tag{A-3}$$

Substituting $\mathbf{x}_t$ back into (A-1) gives a non-linear econometric model of the cross section of N observed yields:

$$\widetilde{\mathbf{y}}_t = \mathbf{y}(\mathbf{x}(\widetilde{\mathbf{y}}_t; \mathbf{\Psi}); \mathbf{\Psi}) + \mathbf{v}_t. \tag{A-4}$$

Thus, given a set of parameters $\mathbf{\Psi}$, we can find $\mathbf{x}_t$ that solves the equation

$$\mathbf{q}_{y,t} = \mathbf{W}'\mathbf{y}(\mathbf{x}(\widetilde{\mathbf{y}}_t; \mathbf{\Psi}); \mathbf{\Psi}) \tag{A-5}$$

for any t. Since there is a direct mapping between $\mathbf{q}_{y,t}$ and $\mathbf{x}_t$, the latter can be treated as observable, conditional on $\mathbf{\Psi}$. The parameters that are necessary for solving Eq.(A-5) are those necessary for cross-sectional spanning of asset prices, i.e. the risk-neutral parameters: $\boldsymbol{\mu}^{\mathbb{Q}}$, $\mathbf{\Phi}^{\mathbb{Q}}$ and $\mathbf{\Sigma}$.

Similarly, assuming that the first principal component of surveys is observed without measurement errors

$$q_{s,t} \equiv \mathbf{W}'_s\widetilde{\mathbf{y}}^s_t = \mathbf{W}'_s\mathbf{y}^s_t, \tag{A-6}$$

we can then extract the $s-$factor:

$$q_s(\{\mathbf{x}_t, s_t\}; \mathbf{\Psi}) = \mathbf{W}'_s\mathbf{y}^s(\{\mathbf{x}_t, s_t\}; \mathbf{\Psi}) \Longleftrightarrow s_t = q_s^{-1}(\{q_{s,t}, \mathbf{x}_t\}; \mathbf{\Psi}),$$

where $\mathbf{W}_s$ is the weighting matrix (vector) for surveys. Note that extraction of s_t is conditional on $\mathbf{x}_t$, since those are required for forming model-based expectations that are fitted to surveys. Given that, we can solve

$$q_{s,t} = \mathbf{W}'_s\mathbf{y}^s(\{\mathbf{x}_t, s_t\}; \mathbf{\Psi}) \tag{A-7}$$

for s_t for all t. Since surveys span both cross-sectional and temporal dimensions, the vector of parameters required to solve Eq.(A-7) must include both the $\mathbb{Q}$ parameters $\boldsymbol{\mu}^{\mathbb{Q}}$, $\boldsymbol{\Phi}^{\mathbb{Q}}$ and $\boldsymbol{\Sigma}$ and the $\mathbb{S}$ parameters $\boldsymbol{\mu}^{\mathbb{S}}$ and $\boldsymbol{\Phi}^{\mathbb{S}}$.

Due to the observability restriction, the parts of the likelihood function that correspond to cross sections of yields and surveys are identical regardless of whether we condition on the observed principal components or on the underlying factors, i.e.

$$\begin{aligned} \ell^{\mathbb{Q}}(\widetilde{\mathbf{y}}_t|\mathbf{q}_{y,t};\mu_1^{\mathbb{Q}},\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\Sigma}_{\mathbf{y}},\boldsymbol{\Sigma}) &= \ell^{\mathbb{Q}}(\widetilde{\mathbf{y}}_t|\mathbf{x}_t;\mu_1^{\mathbb{Q}},\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\Sigma}_{\mathbf{y}},\boldsymbol{\Sigma}) \\ \ell^{\mathbb{S}}(\widetilde{\mathbf{y}}_t^s|\mathbf{q}_{y,t},q_{s,t};\mu_1^{\mathbb{Q}},\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\mu}^{\mathbb{S}},\boldsymbol{\Phi}^{\mathbb{S}},\boldsymbol{\Sigma}_{\mathbf{s}},\boldsymbol{\Sigma}) &= \ell^{\mathbb{S}}(\widetilde{\mathbf{y}}_t^s|\mathbf{x}_t,s_t;\mu_1^{\mathbb{Q}},\boldsymbol{\lambda}^{\mathbb{Q}},\boldsymbol{\mu}^{\mathbb{S}},\boldsymbol{\Phi}^{\mathbb{S}},\boldsymbol{\Sigma}_{\mathbf{s}},\boldsymbol{\Sigma}) \end{aligned}$$

The rotation of the factors, however, requires an adjustment to the part of the likelihood that corresponds to the physical dynamics of the factors $\ell^{\mathbb{P}}$. This can be done through the application of the change-of-variable technique:[15]

$$\ell^{\mathbb{P}}(\mathbf{q}_{y,t},q_{s,t}|\mathbf{q}_{y,t-1},q_{s,t-1};\boldsymbol{\mu}^{\mathbb{P}},\boldsymbol{\Phi}^{\mathbb{P}},\boldsymbol{\Sigma}) = \ell^{\mathbb{P}}(\mathbf{x}_t,s_t|\mathbf{x}_{t-1},s_{t-1}) \times |\det(\mathbf{J}_{x,t})|^{-1} \times |(J_{s,t})|^{-1}, \quad \text{(A-8)}$$

where $\mathbf{J}_{x,t}$ and $J_{s,t}$ are the Jacobian terms resulting from the change of variables:

$$\mathbf{J}_{x,t} \equiv \left[\frac{\partial \mathbf{q}_{y,t}}{\partial x_{1,t}},\cdots,\frac{\partial \mathbf{q}_{y,t}}{\partial x_{K,t}}\right] \quad \text{and} \quad J_{s,t} \equiv \left[\frac{\partial q_{s,t}}{\partial s_t}\right].$$

The rotation of factors from $\mathbf{q}_{y,t}$ and $q_{s,t}$ to $\mathbf{x}_t$ and s_t, respectively, allows us to find the conditional mean parameter under the $\mathbb{P}$ measure in Eq.(4), $\boldsymbol{\mu}^{\mathbb{P}}$ and $\boldsymbol{\Phi}^{\mathbb{P}}$, by *OLS* due to the result in Zellner (1962), conditional on the risk-neutral and subjective parameters that are necessary for factor rotation.

In the shadow rate model the first derivative for a $j-th$ yield with the remaining maturity

[15] See Greene (2011), Appendix B.

m_j is

$$\frac{\partial y_{j,t}}{\partial \mathbf{x}_t} = \frac{1}{m_j}\frac{\partial}{\partial \mathbf{x}_t}\sum_{m=0}^{m_j-1} f(\mathbf{x}_t, m; \Psi) \approx \frac{1}{m_j}\sum_{m=0}^{m_j-1} \Phi\left(\frac{a_{f,m} + \mathbf{b}'_{f,m}\mathbf{x}_t - \underline{r}}{\sigma_m^{\mathcal{Q}}}\right)\mathbf{b}_{f,m}. \tag{A-9}$$

Denote by $\mathbf{H}_{x,t}$ stacked vectors $\partial y_{j,t}/\partial \mathbf{x}_t$:

$$\mathbf{B}_t = \frac{\partial \mathbf{y}_t}{\partial \mathbf{x}'_t} = \left[\frac{\partial y_{1,t}}{\partial \mathbf{x}'_t}, \cdots, \frac{\partial y_{J,t}}{\partial \mathbf{x}'_t}\right]'. \tag{A-10}$$

Thus, the Jacobian $\mathbf{J}_{y,t}$ is given by

$$\mathbf{J}_{y,t} = \mathbf{W}'\mathbf{H}_{x,t}. \tag{A-11}$$

In a similar fashion we find the Jacobian $J_{s,t}$. First, note that given our assumption of the subjective yield expectations in Eq.(7), the shadow rate yield expectation is

$$\begin{aligned} y^s_{j,t,h} &= \frac{1}{m_j}E_t\left[\sum_{m=0}^{m_j-1} f_{m,t+h}\right] \qquad \text{(A-12)}\\ &\approx \underline{r} + \frac{1}{m_j}\sum_{m=0}^{m_j-1}\sigma_m^{\mathbb{Q}} g\left(\frac{a_{f,m} + \mathbf{b}'_{f,m}E_t[\mathbf{x}_{t+h}] - \underline{r}}{\sigma_m^{\mathbb{Q}}}\right)\\ &= \underline{r} + \frac{1}{m_j}\sum_{m=0}^{m_j-1}\sigma_m^{\mathbb{Q}} g\left(\frac{a_{f,m} + \mathbf{b}'_{f,m}\boldsymbol{\mu}^{\mathbb{S}}_{x,h} + \mathbf{b}'_{f,m}\boldsymbol{\Phi}^{\mathbb{S}}_{x,h}\mathbf{x}_t + \mathbf{b}'_{f,m}\boldsymbol{\Phi}^{\mathbb{S}}_{s,h}s_t - \underline{r}}{\sigma_m^{\mathbb{Q}}}\right), \end{aligned}$$

where the last line uses the conditional expectation of $\mathbf{x}_{t+h}$ at time t is defined in Eq.(9). The parameters $\boldsymbol{\mu}^{\mathbb{S}}_{x,h}$ and $\boldsymbol{\Phi}^{\mathbb{S}}_{x,h}$ are found by forward propagating the subjective dynamics equation (8). Since the first derivative of $y_{m,t}$ with respect to s_t is

$$\frac{\partial y^s_{j,t,h}}{\partial s_t} = \frac{1}{m_j}\sum_{m=0}^{m_j-1}\Phi\left(\frac{a_{f,m} + \mathbf{b}'_{f,m}\boldsymbol{\mu}^{\mathbb{S}}_{x,h} + \mathbf{b}'_{f,m}\boldsymbol{\Phi}^{\mathbb{S}}_{x,h}\mathbf{x}_t + \mathbf{b}'_{f,m}\boldsymbol{\Phi}^{\mathbb{S}}_{s,h}s_t - \underline{r}}{\sigma_m^{\mathbb{Q}}}\right)\boldsymbol{\Phi}^{\mathbb{S}\prime}_{s,h}\mathbf{b}_{f,m}, \tag{A-13}$$

the Jacobian $J_{s,t}$ is

$$\mathbf{J}_{s,t} = \mathbf{W}_s' \mathbf{H}_{s,t}, \tag{A-14}$$

where

$$\mathbf{H}_{s,t} = \frac{\partial \mathbf{y}_t^s}{\partial s_t} = \left[\frac{\partial y_{1,t}^s}{\partial s_t}, \cdots, \frac{\partial y_{J_s,t}^s}{\partial s_t} \right]'.$$

Appendix B Forecasting errors for Case 5

In this appendix, we report in-sample and out-of-sample $RMSFE$ for Case 5.

Model	Horizon\Yield	3m	1y	5y	10y
Case 5 (in-sample)	1q	31.74	35.57	40.87	41.33
($\mathbb{S} \nsim \mathbb{P},\ \mathbf{m}^{\mathbb{S}} = \mathbf{m}^{\mathbb{S}}$)	2q	63.56	73.86	73.55	67.56
	3q	92.27	103.35	94.72	84.87
	4q	119.67	132.31	115.21	99.99
Case 5 (out-of-sample)	1q	26.54	28.31	33.12	37.24
($\mathbb{S} \nsim \mathbb{P},\ \mathbf{m}^{\mathbb{S}} = \mathbf{m}^{\mathbb{S}}$)	2q	46.24	49.92	55.00	54.25
	3q	68.01	71.27	72.70	69.41
	4q	85.54	89.56	85.13	79.52

Table 15. Root mean-square forecasting error for Case 5, in-sample and out-of-sample.