A Forward-looking Model of the Term Structure of Interest Rates

Albert Lee Chun*

July 2, 2016

Abstract

We build dynamic term structure models using a generalized structure of observable, forward-looking factors, where the dynamics of multi-horizon survey forecasts of inflation, output growth and monetary policy are modelled jointly with the physical process driving their realisations. When multiple-horizon forecasts drive the short rate, it takes on the novel interpretation of a forward-looking multiple-horizon monetary policy rule, which facilitates a decomposition of monetary policy and the yield curve into short and longer horizon expectations. Although short horizon expectations of real output growth are obscured in the cross section of yields, longer horizon growth expectations are strongly manifest in the yield curve's slope. We conclude by exploring the models' implications by linking expectations with bond risk premia. Our models provide central bankers and market participants with a tool for linking the dynamic properties of the yield curve to the multiple horizon structure of market expectations, including those possibly imputed to forward guidance.

KEYWORDS: macro-term structure model, forward-looking multi-horizon policy rule, survey expectations, bond risk premia (JEL G12, E37, E43, E44, E52, C13)

*UQ Business School, University of Queensland, Brisbane, Australia, email: albertleechun@gmail.com. I would like to thank Caio Almedia, Bogdan Dima, Oleysa Greishenko, Jesper Lund, and participants at the Annual SoFiE Conference, ESSFM Gerzensee Asset Pricing Week (Evening Sessions), Econometric Society European Meeting, D-CAF Members Meeting, 2nd Humbolt-Copenhagen Financial Econometrics Conference, World Congress of the Bachelier Finance Society, SAFE Conference on New Directions in Term Structure Modelling, Northern Finance Association, French Finance Association (AFFI), FIRN Art of Finance Conference, Bank of Canada, UC Riverside, KAIST Graduate School of Finance, Seoul National University, Yonsei School of Business, SKK GSB, Vrije Universiteit Amsterdam, Erasmus School of Economics Rotterdam, University of Queensland, Copenhagen Business School, CEFIMO at ASE Bucharest, West University of Timisoara and the Stony Brook University Center of Finance for comments on earlier drafts of this article.

1 Introduction

Both theory and empirical evidence lend support to a dynamic link between bond yields and market expectations. Even casual observations reveal that interest rates tend to rise during episodes of higher inflationary expectations, along with tighter expectations of central bank monetary policy. Such linkages can partly be motivated by a forward-looking monetary authority and expectations about the future stance of monetary policy, and partly by the impact of investor expectations on bond risk premiums. Since market expectations play a central role in the determination of the term structure of interest rates, they ought to also play a prominent role in the development of academic models.

By including macro variables in the state vector, the recent slate of yield curve models has made large strides in helping to uncover the links between macroeconomics, monetary policy and the yield curve.[1] However, the macro-economic factors used in most of these studies, as for example Ang and Piazzesi (2004), are essentially backward-looking, as they measure realized, historically observed economic quantities. Moreover, these models are heavily dependent on the presence of latent factors, linking them with traditional models of interest rate behavior, such as Duffie and Kan (1996) and Dai and Singleton (2000). Latent factors, by nature of their construction, result in a near perfect fit of the yield curve. Due to the fact that latent factors are generally void of economic meaning, we eschew their use throughout this study. We employ only observable, forward-looking factors and allow for the presence of errors on all model-implied yields, making no claims that our factors span the yield curve perfectly. Naturally, we also do not impose the converse condition that yields span our factors, as is the premise in Joslin, Singleton, and Priebsch (2014). Therefore, we stress that the factors in our model comprise spanned as well as unspanned components vis-à-vis both yields and risk premia, the identification of which is not the central focus of our research.

In interpreting short rate equations driven by macroeconomic expectations as a forward-looking monetary policy rule, Chun (2011) studies term structure models using survey-based

[1] In pre-dating the now extensive macro-term structure literature by several decades, Langetieg (1980) argues that the term structure of interest rates is embodied in a large macroeconomic system and that yield curve models should be rich enough to accommodate a large number of economic relationships. Piazzesi (2005) first introduced the role of Federal Reserve policy into interest rate models by incorporating policy moves when modelling the short rate.

factors comprising expert analysts' views from the Blue Chip Financial Forecasts.[2] He decomposes and highlights the contribution of macroeconomic and interest rate expectations on the shape of the yield curve and identifies real GDP growth as a dominant driver of time variation in risk premia. However, his models only examine the role of the 3-month ahead forecasts, thus failing to incorporate information across the complete set of horizons available in the surveys.

In this research, we develop a new framework that facilitates the examination of the yield curve's relation with the multi-horizon structure of market expectations. We leverage this additional information in defining and describing the dynamics of the model's factors, as well as in interpreting the risk premiums implied by these dynamics. We study three different variants of our model, each progressively increasing in the amount of survey information used in the estimation. First, we select a forecast horizon h^* to define the state vector and then estimate the model using only information corresponding to this particular horizon in the surveys. Second, we select a forecast horizon h^* to define the state vector, and use information across the multi-horizon structure of the forecasts in the estimation. Our third specification extends the model to allow multiple horizon forecasts to synchronously define the short rate equation. This conceptual innovation gives the short rate the interpretation of a *forward-looking multiple-horizon monetary policy rule*, which facilitates a decomposition of monetary policy into short and longer horizon expectations. Indeed, a forward-looking central bank would consider concurrent expectations across various horizons when setting the target rate, and our model is novel in being the first to model this aspect of monetary policy.

We highlight a fundamental trade-off in a model's ability to simultaneously fit the short and long ends of the yield curve, as the ability of survey forecasts to describe short maturity

[2]Several studies have relied on forward-looking information from surveys to better define various model properties. Pennacchi (1991) uses expectations from surveys to help to identify the real rate and expected inflation factors in a latent variable model. He also suggests that surveys could help resolve the small sample problem and hence provide more powerful estimates of the dynamics of bond prices. More recently, Kim and Orphanides (2012) use 6-month and 12-month ahead forecasts of the 3-month maturity yield from surveys in estimating the drift parameters under the physical measure. Their model is also based on a latent factor approach, with the surveys playing only an ancillary role in the estimation. In a similar vein, Orphanides and Wei (2012) incorporate survey forecasts of inflation and the 10-year yield from the SPF survey in estimating a model with time-varying coefficients. Chernov and Mueller (2012) who fit their model to the term structures of inflation forecasts across several different surveys and allow for the subjective expectations driving inflation surveys to potentially differ from expectations generated by the physical measure.

yields monotonically decreases as the forecast horizon increases. A model defined using a short forecast horizon best describes the dynamics of short maturity yields, while longer maturity yields are best captured using longer horizon forecasts. This demonstrates that there is differing information across the different forecast horizons, all of which may be useful for describing different segments of the yield curve. Thus admitting only a single forecast horizon into the short rate equation creates a tension between jointly fitting the short and long segments of the yield curve, a tension that persists even when additional horizons are used in the estimation.

We break this tension by allowing multiple horizon forecasts to determine the short rate equation so as to allow yields to be concurrently driven by information available across various forecast horizons. In a single-horizon model, expected inflation is a level factor, that explains essentially all the dynamics of long maturity yields. However, in a model driven by two sets of horizons, expected inflation can be decomposed into mirrored slope factors where long yield movements are coupled with longer horizon inflation expectations. In addition, we find that short-horizons forecasts of real GDP growth appear to be nearly 'hidden', in the sense of Duffee (2011), in that they have little impact on the contemporaneous cross-section of yields. In stark contrast, we find that longer horizon forecasts of real GDP growth are clearly not 'hidden,' and their role as a slope factor is distinctly manifest in the yield curve.

We find that the expectations implied by our models can capture up to 20% of the variation in realised excess returns. Interestingly, including additional information from multiple horizons in the estimation does not appear to help the models better describe the level or the variation in excess returns. In addition, we find evidence that in addition to capturing a good part of the component of excess returns explained by information in the yield curve, our model generated expectations also capture part of the component that is unexplained by a linear combination of information in forward rates. Finally, we note that forward-looking models that incorporate both macroeconomic and short rate expectations provide a natural framework for studying the impact of the central bank's forward guidance on the yield curve.

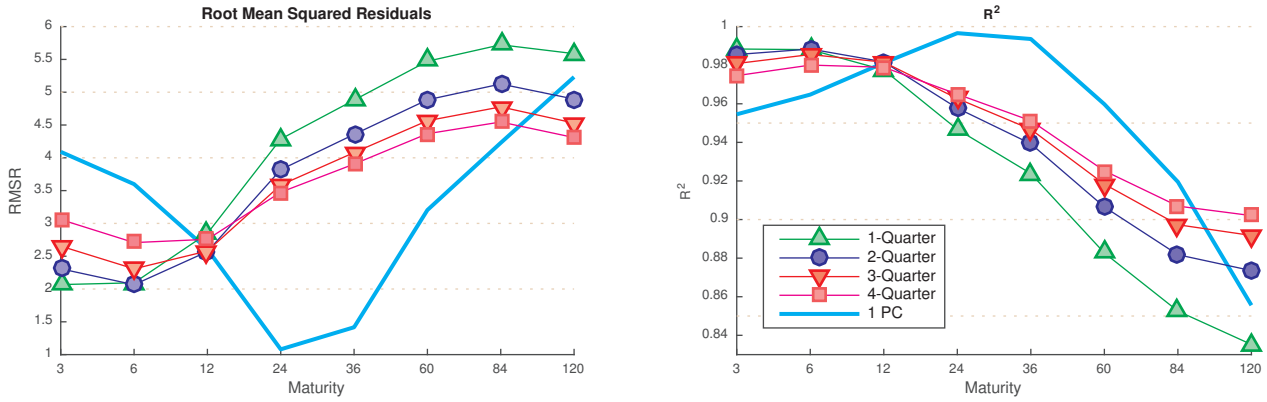

Figure 1: **Yield Curve Dynamics and Survey Expectations** This figure plots the Root Mean Squared Residuals (RMSRs) in basis points (left panel), and R^2s (right panel) obtained from a linear regression of yields of maturities 3, 6, 12, 24, 36, 60, 84 and 120 months onto three forward-looking factors taken from the Blue Chip Financial Forecasts - real GDP growth, CPI inflation and the federal funds rate for forecast horizons ranging from one to four quarters ahead. The result of regressing yields on the first principal component is also shown as a benchmark.

2 Motivating Regressions

To motivate our study, Figure 1 plots the square root of the mean squared residuals (RMSRs) and R^2s from a linear regression of yields of different maturities on survey forecasts of inflation, GDP growth and the fed funds rate one to four quarters ahead. Suppose we are constrained in selecting a single forecast horizon for use in defining the factors in a yield curve model. Such a constraint would naturally emphasize matching one segment of the yield curve over another. We see that the 1-quarter ahead forecast, as used in Chun (2011), best captures the behavior of only the shortest maturity yields, but is less than ideal for explaining the long end of the yield curve, for which the 4-quarter ahead forecast would be a better modelling choice. This suggests that incorporating the full structure of available forecast horizons could lead to a more informative model of the yield curve.

Figure 2 looks deeper into the variations in the yields that are explainable by using various subsets of the survey forecasts. As would be expected, including the foretasted funds rate is sufficient to capture variation in short yields. However, note that for longer maturity yields, foretasted inflation alone captures nearly all of the variation. In all three of the left hand side panels, the R^2 results using contemporaneously observable, historical counterparts are plotted, and clearly demonstrate the advantage of using the forecasts as

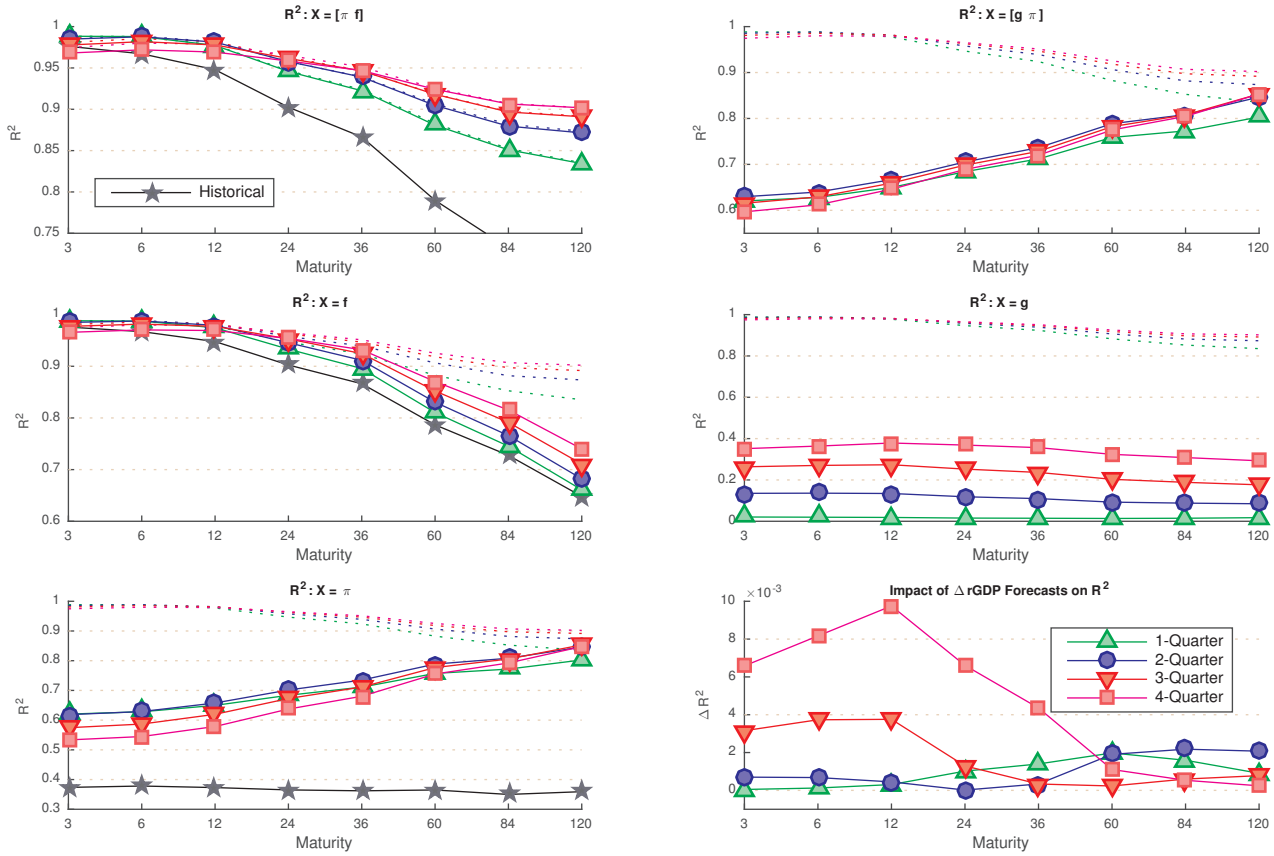

Figure 2: **Decomposing Yield Dynamics into Survey Expectations** This figure plots the R^2s (right panel) obtained from a linear regression of yields of maturities 3, 6, 12, 24, 36, 60, 84 and 120 months onto various subsets of the three forward-looking factors taken from the Blue Chip Financial Forecasts - real GDP growth(g) , CPI inflation(π) and the federal funds rate(f) for forecast horizons ranging from one to four quarters ahead. The upper right panel, shows the incremental difference in the adjusted R^2 from adding real GDP growth conditioned on the other two forecasts. The left panels also report the R^2 based on contemporaneously available, historical inflation and funds rate data. As a benchmark, the R^2 results from Figure 1 are presented using dotted lines.

opposed to contemporaneous variables.[3] Conditioned on forecasted inflation and the funds rate, real GDP forecasts have little explanatory power, increasing the R^2 by less than 1% as seen in the bottom right panel. However, note that 4-quarter ahead real GDP forecasts, when used as the sole explanatory variable, is an important driver of the yield curve, explaining nearly 40% of the variation in yields.

Returning to Figure 1, the R^2 results based on the first principal component suggests that the fit from using the three survey forecasts as factors would, on average, be comparable to

[3]The comparison using contemporaneous variables is only done for the funds rate and inflation, which can measured at the same monthly frequency as the survey forecasts. We omit these results in the right panel, as real GDP growth is only reported quarterly.

a one-factor latent variable model.[4] Note that the survey forecasts explain a greater amount of the variation in short maturities yields and also in the 10-year yield, relative to the first principal component of yields. It must be stressed, however, that the motivation of this research is not one of competing with latent factor models in fitting yields, as these latent factors (or equivalently, observable combinations of contemporaneous yields such as principal components), will, by construction, maximize the total variation explained. Latent factor models explain yields only tautologically, that is by "using yields to explain yields." Our motivation is to study the economic drivers that shape the yield curve, by linking yields to a set of observable factors that are external, rather than internal to yields themselves. The only factors we employ are quantifiable measures of macroeconomic expectations, and we focus on investigating the extent to which these forward-looking measures, across different horizons, can explain key features of the yield curve.

3 The Model

In providing an overview of our framework, let X_t be a vector of fundamental state variables such as inflation, output growth and monetary policy, which are important for driving bond prices. Suppose that the physical $\mathbb{P}$-dynamics of X_t is governed by $dX_t = \mathcal{K}(\theta - X_t)\,dt + \Sigma_h(X_t)\,dB_t$ and we represent the h-month ahead forecast of the k-month average of X_t by $\bar{X}_{t,h} \equiv E^{\mathbb{P}}[(1/k)\int_{t+h}^{t+h+k} X_s ds|X_t]$.[5] We show that $\bar{X}_{t,h} = (1 - \Psi(h))\,\theta + \Psi(h)\,X_t$, which is a weighted average of the long run mean θ and the random walk forecast X_t. We further show that the weighting matrix on the random walk forecast, $\Psi(h)$, also drives the volatility of $\bar{X}_{t,h}$, which evolves according to $d\bar{X}_{t,h} = \mathcal{K}\left(\theta - \bar{X}_{t,h}\right)dt + \Psi(h)\,\Sigma(X_t)\,dB_t$. Thus, the parameters that govern the drift of the forecasts are identical to the parameters that govern the drift of the fundamental state vector; the instantaneous volatility however is dependent on the forecast horizon via $\Psi(h)$. Accordingly, our model implies that information about the parameters driving X_t should be perfectly revealed either through its own time-series

[4] All yields and hence RMSRs are given as monthly rates. To annualize for comparison with other studies, one needs to multiply by 12. So a residual of 2 basis points, translates to 24 basis points in annual terms.

[5] Without loss of generality, although we assume that the average forecasts are generated under the physical or data generating measure $\mathbb{P}$, the same modelling framework supports the interpretation that the forecasts are generated under a subjective probability measure $\mathbb{P}^s$, under the assumption that the surveys deviate from rational expectations and that bonds are priced based on this belief.

dynamics, or, for any choice of h, through the time-series dynamics of $\bar{X}_{t,h}$. If this is the case, then given any arbitrary forecast horizon, inference about $\mathcal{K}$, Σ and θ could be made from the time-series dynamics of the forecasts, which we proxy for using survey data. If the survey expectations are well described by our model, then including data from across multiple forecast horizons should lead to a more accurate description of the underlying system.

Furthermore, due to the survey participants being asked to predict an average over a particular calendar quarter, a nontrivial issue that arises when estimating models at a monthly frequency is the resulting time-varying nature of the forecast horizon. Hence all studies that incorporate monthly information from this survey, either ignore this fact, or rely on simple approximations. We derive a consistency relation between forecasts across different horizons by which an h^*-horizon and an h-horizon forecast are related by $\bar{X}_{t,h^*} = \theta + e^{-\mathcal{K}(h^*-h)}\left(\bar{X}_{t,h} - \theta\right)$. This key result enables the construction of a set of constant horizon forecasts which then define the model factors at a monthly frequency, in effect, resolving the time-varying forecast horizon issue. To illustrate our framework, we focus on special cases of forward-looking macro-term structure model where survey expectations of real GDP growth, inflation and the fed funds rate follow multivariate Gaussian dynamics, which then drive bond yields within an affine term structure framework.

3.1 A Model of the Forward-looking Factors

Suppose there is an N-dimensional real-valued vector of state variables X_t important for driving bond prices. Suppose the dynamics of X_t under the physical measure $\mathbb{P}$ evolve according to the following stochastic differential equation

$$dX_t = \mathcal{K}\left(\theta - X_t\right) dt + \Sigma(X_t)\, d\mathcal{B}_t \tag{1}$$

where $\mathcal{K}$ is an $N \times N$ matrix, $\Sigma(X_t)$ is a function defining an $N \times N$ (potentially stochastic) volatility matrix, θ is an N-dimensional central tendency vector and $\mathcal{B}$ is an N-dimensional vector of independent Brownian motions under $\mathbb{P}$. A unit of time is defined to be one month.

At time t, we denote the h-period ahead conditional expectation of the average realisation

of X_t over the k-period interval $[t+h, t+h+k]$ by

$$\bar{X}_{t,h} \equiv E^{\mathbb{P}}\left[\frac{1}{k}\int_{t+h}^{t+h+k} X_s ds | X_t\right]. \tag{2}$$

It follows from Proposition 1 of Appendix A that $\bar{X}_{t,h}$ is a linear combination of the random walk forecast X_t and the long run mean θ

$$\bar{X}_{t,h} = (1 - \Psi(h))\,\theta + \Psi(h)\,X_t, \tag{3}$$

where the weight on the random walk forecast is given by $\Psi(h) \equiv \frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}h}\left[I - e^{\mathcal{K}k}\right]$.

Furthermore, we show in Corollary 1 of Appendix A that the time series dynamics of the forecasts $\bar{X}_{t,h}$ evolve according to

$$d\bar{X}_{t,h} = \mathcal{K}\left(\theta - \bar{X}_{t,h}\right)dt + \Sigma_h(X_t)\,d\mathcal{B}_t. \tag{4}$$

where $\Sigma_h(X_t) = \left[\frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}h}\left(I - e^{\mathcal{K}k}\right)\right]\Sigma(X_t) = \Psi(h)\,\Sigma(X_t)$.[6] The parameters that govern the drift of the conditional expectation are identical to the parameters that govern the drift of the underlying state vector as both processes X_t and $\bar{X}_{t,h}$ are characterized by the same speed of mean reversion $\mathcal{K}$ and the same long run mean θ. However, the instantaneous volatility is now dependent on the forecast horizon h via $\Psi(h)$. Intuitively, $\Sigma_h(X_t)$ is 'proportional' to $\Sigma(X_t)$, in the sense that it scales according to the weighting matrix on the random walk forecast, $\Psi(h)$. The insight this provides is that the closer the conditional forecast is to a random walk, the closer the volatility of the forecast is to the volatility of X_t, while the greater the distance between the conditional forecast and the random walk (in the direction of the long run mean θ), the more dampened is the volatility of $\bar{X}_{t,h}$. In this setting, information about the dynamics of the underlying state vector is reflected not only in the time-series dynamics of the X_t but in the time-series dynamics of $\bar{X}_{t,h}$ as well, as both are driven by the common parameter set of $\mathcal{K}$, θ and Σ.

To study the theoretical properties implied by the model, for values of $\mathcal{K}$ that imply stationarity (which we do not strictly impose in the estimation), we can see that as h in-

[6] In contrast to our Markovian setting, Feunou and Fontaine (2014) develop a model with expected inflation where the underlying risk-factors are non-Markovian under $\mathbb{P}$. However, as in our model, conditional expectations in their setting are also Markovian.

creases, the magnitude of the elements of the $\Psi(h)$ matrix decreases, going to zero in the limit.[7] This implication of the model is consistent with our intuition that in the long run, the expectation of a stationary process should converge to its long run mean θ, and in doing so have zero dynamics and hence also zero volatility. A practical implication of this is that volatility would tend to fall as the forecast horizon increases.

3.2 A Forward-looking Yield Curve

Given a single-horizon framework with a fixed horizon h^*, we take $\bar{X}_{t,h^*}$ as factors in an affine term structure model. In defining the instantaneous short rate, $r_t = \rho_0 + \rho_1 \bar{X}_{t,h^*}$, the forecast horizon h^* determines the degree to which the central bank is forward looking. The short rate equation is given economic life as a descriptive interpretation of a forward-looking Taylor (1993) rule, as in Clarida, Gali, and Gertler (2000). Such a rule links the behavior of the central bank to forecasts of economic conditions as opposed to contemporaneous or backward looking observations. The market prices of risk, $\lambda_t = \lambda_0 + \lambda_1 \bar{X}_{t,h^*}$, and hence all risk premiums implied by the model are also forward looking. The model specification follows the expectations-based variant of the arbitrage-free affine term structure model as described in Section 2 of Chun (2011). We will later generalize this to a multi-horizon setting.

The physical process driving $\bar{X}_{t,h^*}$ is assumed to follow a Gaussian version of the dynamics in (4) and under the equivalent martingale measure $\mathbb{Q}$ we have

$$d\bar{X}_{t,h} = \mathcal{K}^{\mathbb{Q}} \left(\theta^{\mathbb{Q}} - \bar{X}_{t,h} \right) dt + \Sigma_h d\mathcal{B}_t^{\mathbb{Q}}. \tag{5}$$

This falls within the setting of Duffie and Kan (1996) so that for an n-period bond the yield and price are given by $Y_t^{(n)} = A(n) + B(n)' \bar{X}_{t,h}$ and $P_t^{(n)} = e^{-nA(n) - nB(n)' \bar{X}_{t,h}}$, respectively. We impose no restrictions on $K^{\mathbb{Q}}$ and $\theta^{\mathbb{Q}}$ that follow from bounds on the market prices of risk, so that the estimation of the drift under $\mathbb{P}$ is unencumbered by the estimation of the drift under $\mathbb{Q}$ and vice versa. As such the $\mathbb{P}$ parameters matter for constructing $\bar{X}_{t,h^*}$ but do not affect the $\mathbb{Q}$-dynamics of $\bar{X}_{t,h^*}$. This motivates our two-step estimation procedure where

[7] The required condition for stationarity is for the smallest eigenvalue of the matrix $\mathcal{K}$ to be strictly greater than zero. From the decomposition, $e^{-\mathcal{K}h} = U e^{-\Lambda h} U^{-1}$, where Λ is a diagonal matrix with the eigenvalues of $\mathcal{K}$, it follows that stationarity implies that $e^{-\mathcal{K}h}$, and thus $\Psi(h) \equiv \frac{1}{k} \mathcal{K}^{-1} e^{-\mathcal{K}h} \left[I - e^{\mathcal{K}k} \right]$ goes to zero as h goes to infinity.

the physical drift and volatility parameters are estimated in the first step. The volatility parameters are taken to be the same in the second step when estimating the $\mathbb{Q}$ parameters due to the diffusion invariance principal where the volatility matrix is invariant to a change of measure. Once all the information in the surveys are accounted for, information in the shape of the yield curve is assumed to be incrementally uninformative about the evolving nature of the forecasts, so that conditionally, yields are not useful for describing the $\mathbb{P}$ dynamics of the surveys. This can be further justified by our use of survey forecasts available at the beginning of the month to describe average yields over that month. Once the parameters are estimated, it is straight forward to back out estimates of the market price of risk parameters as $\lambda_0 = \Sigma_h^{-1}[K\theta - K^{\mathbb{Q}}\theta^{\mathbb{Q}}]$ and $\lambda_1 = \Sigma_h^{-1}[K^{\mathbb{Q}} - K]$.

3.3 Definition of the Fundamental State Vector

The definition of inflation and real output growth used in this article is somewhat non-standard and hence requires some explanation. Recall, that the unit time interval is defined to be one month. Let P_t denote the average level of the consumer price index (CPI) over the 3-month interval $[t-3, t]$. Let π_t^A denote the continuously compounded, annual growth rate in the price level P_t over the 3-month period from t to $t+3$ and note that π_t^A is defined by the following expression[8]

$$e^{\frac{1}{4}\pi_t^A} = \frac{P_{t+3}}{P_t} \tag{6}$$

Likewise, define π_t as the continuously compounded, annualized *instantaneous* growth rate in the price level P_t, so that

$$e^{\frac{1}{12}\int_t^{t+3}\pi_s ds} = \frac{P_{t+3}}{P_t}. \tag{7}$$

Thus we immediately see that π_t^A, the annualized inflation rate over the period from t to $t+3$, is equal to the average of the instantaneous inflation rate π_t over the same 3 month

[8] With π_t^A being an annual inflation rate implies a monthly inflation rate of $\frac{1}{12}\pi_t^A$ and a quarterly inflation rate of $\frac{1}{4}\pi_t^A$.

period

$$\pi_t^A = \frac{1}{3}\int_t^{t+3} \pi_s ds. \tag{8}$$

The h-month ahead conditional expectation of inflation can then be defined as

$$\bar{\pi}_{t,h} \equiv E_t^{\mathbb{P}}\left[\pi_{t+h}^A\right] = E_t^{\mathbb{P}}\left[\frac{1}{3}\int_{t+h}^{t+h+3} \pi_s ds\right]. \tag{9}$$

Denoting g_t as the instantaneous growth rate of quarterly averages of real GDP, we construct analogous definitions of g_t^A and $\bar{g}_{t,h}$ in exactly the same manner as above

$$\bar{g}_{t,h} \equiv E_t^{\mathbb{P}}\left[g_{t+h}^A\right] = E_t^{\mathbb{P}}\left[\frac{1}{3}\int_{t+h}^{t+h+3} g_s ds\right]. \tag{10}$$

Denoting the fed funds rate by f_t, we define $\bar{f}_{t,h}$ as the h-month ahead conditional expectation of the average federal funds rate over a 3 month interval

$$\bar{f}_{t,h} \equiv E_t^{\mathbb{P}}\left[\frac{1}{3}\int_{t+h}^{t+h+3} f_s ds\right]. \tag{11}$$

Finally, we define our fundamental state vector by

$$X_t = [g_t \ \pi_t \ f_t]' \tag{12}$$

and the conditional expectation of the 3-month average of X_t by

$$\bar{X}_{t,h} \equiv [\bar{g}_{t,h} \ \bar{\pi}_{t,h} \ \bar{f}_{t,h}]' = E_t^{\mathbb{P}}\left[\frac{1}{3}\int_{t+h}^{t+h+3} X_s ds\right]. \tag{13}$$

In the following section, we show how to map observable expectations from surveys into $\bar{X}_{t,h}$.

3.4 Surveys as Observable Proxies of Expectations

A literal interpretation of the short rate equation as a forward-looking monetary policy rule requires that the short rate, and hence all yields, be driven by policy maker's expectations. We take a slightly more liberal interpretation of the short rate as an economically-motivated modelling device though which expectations enter the yield curve model. Naturally, both

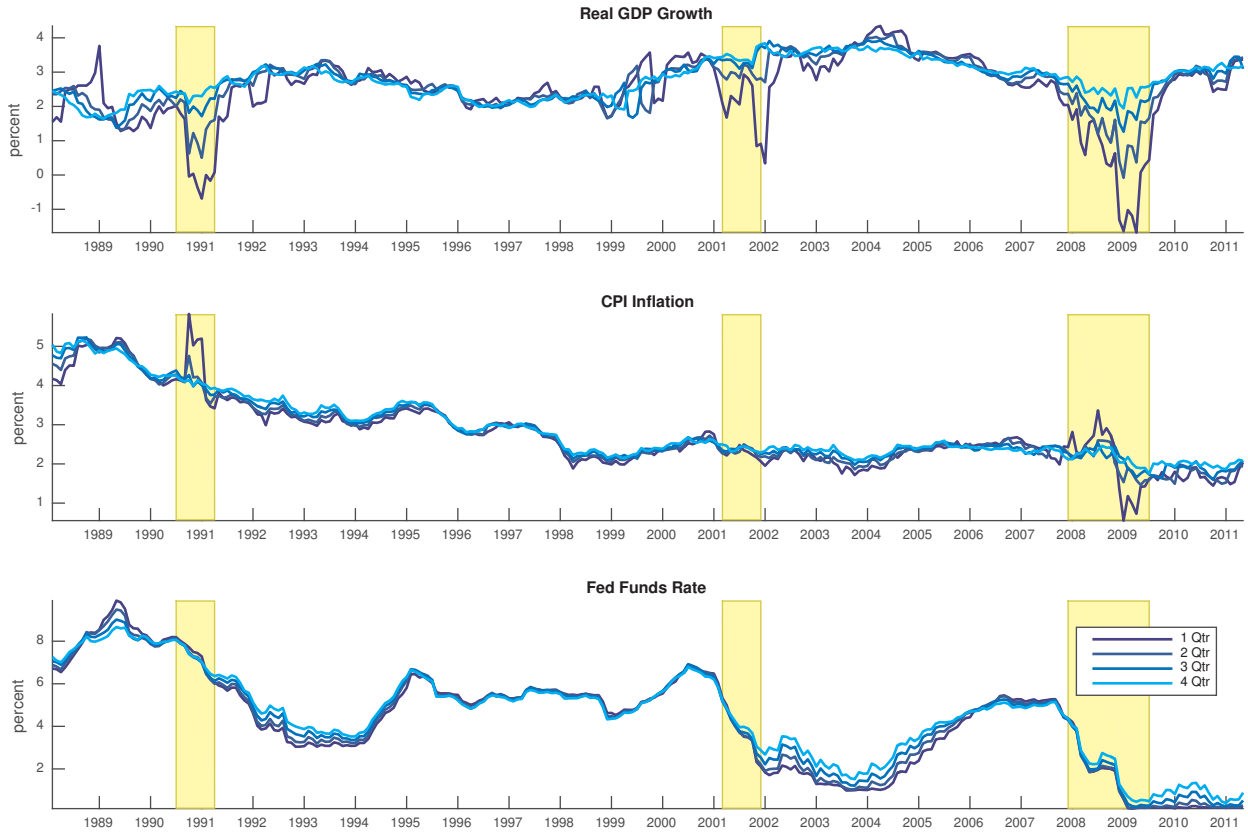

Figure 3: **Forward Looking Factors from Surveys** The figure plots the time series of the 3 forward-looking factors taken from the Blue Chip Financial Forecasts - real GDP growth (top panel), CPI inflation (middle panel) and the federal funds rate (bottom panel) for forecast horizons ranging from 1 to 4 quarters ahead. The official NBER recessions are given by the shaded regions.

the expectations of the Fed and those of the market matter for determining the shape of the yield curve, as these may certainly not coincide with one another, nor with survey data, at any point in time. That said, we justify the use of survey expectations in our models via the following argument.

Consider the following information sets $\mathcal{F}_t \subset \mathcal{F}_t^S$ and $\mathcal{F}_t \subset \mathcal{F}_t^F$, where $\mathcal{F}_t^S$ denotes the information set of the survey participants, $\mathcal{F}_t^F$ the information set of the Fed and $\mathcal{F}_t$ the information set of a naive econometrician. Using the law of iterated expectations, note that $\bar{X}_{t,h} \equiv E[X_{t+h}|\mathcal{F}_t] = E[E[X_{t+h}|\mathcal{F}_t^F]|\mathcal{F}_t] = E[E[X_{t+h}|\mathcal{F}_t^S]|\mathcal{F}_t]$. According to the naive econometrician wishing to estimate the models, conditioned on $\mathcal{F}_t$, the expectation of the Fed's forecast is equal to the expectation of the survey forecast, which is then taken to proxy for the market expectation. From this perspective, we justify the interpretation the short

rate equation as a forward-looking monetary policy rule.[9]

We will thus take the mean expectations of real GDP growth, CPI inflation and the federal funds rate from the Blue Chip Financial Forecasts as our empirical proxy for $\bar{X}_{t,h}$.[10]. The Blue Chip respondents are asked to forecast of the quarterly average of the fed funds rate, which matches the definition of $\bar{f}_{t,h}$. However, the respondents are asked to provide forecasts of the annualized quarter-over-quarter percentage change in the consumer price index and real output (seasonally adjusted). It now becomes clear as to why we chose the particular definitions of inflation and real GDP growth used to construct $\bar{X}_{t,h}$ as this matches the definition of variables from the survey.[11]

By taking the mean survey forecasts as an empirical proxy of $\bar{X}_{t,h}$, we are implicitly making the assumption that the survey forecasts are generated in a way that is consistent with our underlying econometric model. Figure 3 plots the mean values of the surveys over our sample period from January of 1988 to April 2011. Each plot shows four different forecast horizons ranging from one to four quarters ahead. Consistent with the theoretical properties of our model, it appears that volatility varies with the forecast horizon — on average appearing to decrease as the forecast horizon increases.

3.5 Extracting Constant Horizon Forecasts from Surveys

The Blue Chip surveys are released monthly but the forecasts are of the average over a specific calender quarter. This introduces the nuisance of a time-varying forecast horizon

[9]In fact, it would not be unreasonable to assume that $\mathcal{F}_t^S$ is encompassed by $\mathcal{F}_t^F$, given the fact that the Fed can observe the survey forecasts at the beginning of the month. Furthermore studies such as Romer and Romer (2000) find that the Fed forecasts typically outperform those of survey participants. Interestingly in this case $E[E[X_{t+h}|\mathcal{F}_t^F]|\mathcal{F}_t^S] = E[X_{t+h}|\mathcal{F}_t^S]$, so that the survey expectations of the Fed's expecations equal the survey expectations.

[10]A detailed description of the Blue Chip surveys along with an analysis of the forecasting performance of the inflation and funds rate forecasts relative to a set of econometric benchmarks are presented in Chun (2016). The Federal Reserve is primarily interested in core inflation as it excludes volatile energy and food prices, yet we do not have forecasts of core inflation. An interpretation of the forecasting evidence in Ajello, Benzoni, and Chyruk (2014) suggests that survey forecast of inflation, which are more persistent and less volatile than total inflation, may indeed be capturing a significant component of the core.

[11]To be precise, one might argue that the survey participants are providing forecasts of $\%\Delta P_t$, the annualized quarter-over-quarter percentage change, where $1 + \%\Delta P_t = [P_{t+3}/P_t]^4 = e^{\pi_t^A}$, following from our definition of π_t^A. Thus for small changes, $\%\Delta P_t$ is approximately equal to π_t^A. In fact, when computing the mean square error between the survey forecasts and the time series of both $\%\Delta P_t$ and π_t^A, we find the latter series results in slightly lower RMSEs across all forecast horizons. Thus we conclude that the survey forecasts are closer in nature to π_t^A than $\%\Delta P_t$.

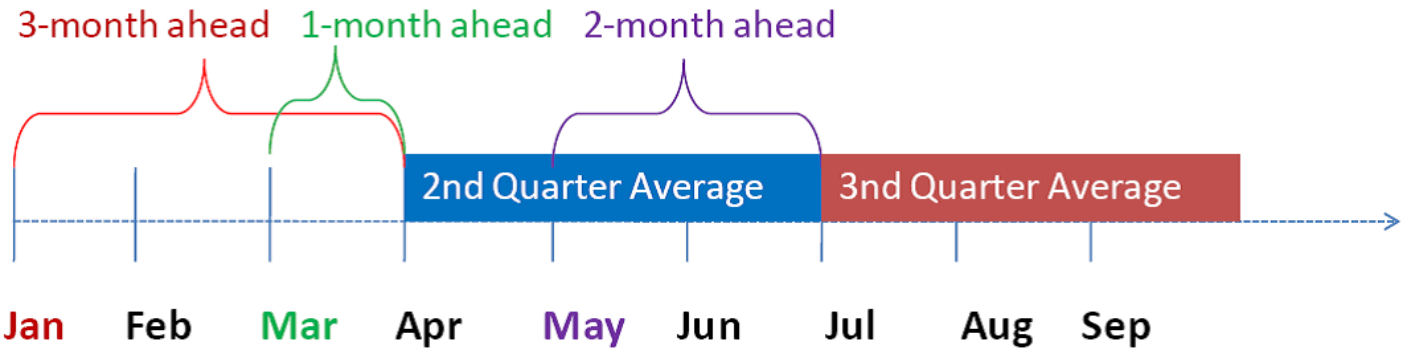

Figure 4: **Time-variation in the Forecast Horizon** The figure illustrates the nature of the time varying forecast horizon problem for the 1-quarter ahead forecast.

into the data. Figure 4 illustrates the nature of this time-varying forecast horizon problem. In January the one quarter ahead forecast is of the expected average realization over April, May and June of that year, thus the forecast horizon is three months. In March, the one quarter ahead forecast horizon is now 1 month, as it pertains to the average over the same period - April, May and June. The observed time series of the survey forecasts, which are of the average realization of X_t over a calendar quarter, is represented as follows

$$\{\bar{X}_{1,l}, \bar{X}_{2,l-1}, \bar{X}_{3,l-2}, \bar{X}_{4,l}, \bar{X}_{5,l-1}, \bar{X}_{6,l-2}, \bar{X}_{7,l}, \ldots\} \tag{14}$$

where, for example, if the forecast horizon is given as 1 quarter ahead in the survey, then $l = 3$, and if the horizon is given as 2 quarters ahead, then $l = 6$. For many applications, constant horizon h^*-period ahead forecasts would be useful, including our immediate application of extracting factors for use in a term structure model. In dealing with this issue, Chun (2011) proposes a simple, non-rigorous approximation of using a weighted average over 2 consecutive calender quarters to construct fixed horizon forecasts.[12]

Orphanides and Wei (2012) and Chernov and Mueller (2012) sample the forecasts at a quarterly frequency whereby obviating this issue completely. Using end of quarter yields may be justified if the research question is specific to observed real output growth, as higher frequency macro data are simply not available. In general, however, using quarterly data only paints a very coarse picture of the yield curve. The frequency of Fed policy moves is often best described at a monthly frequency, and yield curve fluctuations that are sampled

[12]This approximation method is also replicated by others, including Grishchenko and Huang (2013) when extracting constant horizon forecasts of inflation, as well as by Le and Singleton (2013).

quarterly miss much of this activity. For example, a quarterly model would miss that the monthly average of the 1-year maturity yield went from 1.07 in November 2008, to 0.50 in December, then to 0.44 in January 2009 and before moving up to 0.62 in February. From a modelling perspective, the attractiveness of the availability of monthly macro-forecasts in the Blue Chip surveys is, at least partially, offset by the presence of the time varying horizon issue.

Since we are interested in obtaining and using survey expectations at a monthly frequency, we begin by proposing a method to extract constant horizon forecasts in a theoretically consistent and rigorous manner. To do so, we show that the data generating process governing the underlying dynamics of the fundamental vector X_t imposes a consistency condition between a constant horizon h^*-period ahead forecast and a time-varying $h(t)$-period ahead forecast. The following relation is derived in Proposition 2 of Appendix A

$$\bar{X}_{t,h^*} = \theta + e^{-\mathcal{K}(h^* - h(t))} \left(\bar{X}_{t,h(t)} - \theta \right). \tag{15}$$

Hence, we can extract an estimate of an unobserved forecast $\bar{X}_{t,h^*}$ with fixed horizon h^*, conditioned on having observed a forecast with horizon $h(t)$, all the while maintaining consistency with the forecast dynamics under $\mathbb{P}$. Replacing the parameters above with their estimated quantities, $\hat{\mathcal{K}}$ and $\hat{\theta}$, we have an econometrician's estimate of the unobserved h^*-period ahead forecast. Potential applications of this technique include any modelling situation requiring constant horizon survey forecasts. Its usefulness certainly extends beyond the domain of yield curve modeling to other economics and finance applications where constant horizon forecasts are needed at a monthly frequency.[13]

[13] Although we focus on quarterly averages, a slight modification allows us to back out constant horizon forecasts for the state vector itself at a particular point in the future. Kim and Orphanides (2012), for example, extract 3 and 6 month ahead forecasts for the 3-month yield using an interpolation that assumes the forecasts are for the midpoint in the quarter. Our technique suggests a method that could generate missing forecasts in a way that is consistent with the underlying dynamics of the forecasts themselves.

4 Three Versions of the Model: Estimation and Results

The aforementioned framework is general in nature and thus allows for models with stochastic volatility, however, in illustrating our forward-looking model of the term structure of interest rates, we focus on a few special cases where the factors $\bar{X}_{t,h^*}$ follow a multivariate Gaussian process. We estimate three different versions of our model, with each version progressively incorporating more and more information from the survey forecasts. First, in Section 3.1 we describe model SS_{h*}, where a single forecast horizon defines our state vector, and only information from that horizon is used to estimate the model parameters. Second, in Section 3.2 we introduce model SM_{h*}, where a single forecast horizon defines our state vector, but information across the multi-horizon structure of the forecasts is used to estimate the model parameters. Finally, in Section 3.3 we introduce our most significant innovation, the MM-class of models, which permit forecasts across multiple horizons to simultaneously drive the state vector. Information across all horizons entering the model are used to estimate the model parameters. We then interpret the short rate as a *multiple-horizon forward-looking monetary policy rule*.

4.1 Estimating Model SS_{h^*}

We denote as SS_{h^*} models where the state equation is defined by an single h^*-month ahead constant maturity forecast vector that is estimated using a single forecast horizon from the surveys. Given an observed time series of survey forecasts, we estimate the physical parameters $\mathcal{K}$, θ and Σ using maximum likelihood. In this setting where only a single horizon is used for the estimation, we demonstrate how to write down the likelihood function directly. In subsequent sections, where information across multiple horizons enters the estimation, it is natural to estimate the model via a Kalman Filter.

Due to the presence of a time-varying forecast horizon in the data, the standard likelihood function needs to be modified slightly. We derive in Corollary 3 of Appendix A the appropriate from of $\tilde{f}\left(\bar{X}_{t+1,h(t+1)}|\bar{X}_{t,h(t)}\right)$, the conditional density function of $\bar{X}_{t+1,h(t+1)}$

given $\bar{X}_{t,h(t)}$, where $h(t)$ denotes the forecast horizon observed at time t.[14] The parameters $\mathcal{K}$, Σ and θ are estimated by maximizing the sample log-likelihood function $\mathcal{L}_X = 1/T \sum_{t=1}^{T} log\tilde{f}\left(\bar{X}_{t+1,h(t+1)}|\bar{X}_{t,h(t)}\right)$. Table S1 of the Supplementary Appendix shows four different sets of maximum likelihood parameter estimates, for estimates based on forecast horizons ranging from 1 to 4 quarters ahead.

This framework permits the parameters to be estimated despite the absence of a time series of constant horizon forecast data. Once the parameters are estimated then the constant horizon forecasts can be backed out as described in (15). Thus the constant horizon forecasts are extracted, not for estimating the dynamics of the survey forecasts, but in a manner consistent with the dynamics of the forecasts. We construct fixed-horizon forecasts for a set of 12 different monthly horizons with h^* ranging from 1 to 12. When h^* is 1, 2 or 3, estimates of the constant horizon forecasts, $\bar{X}_{t,h^*}$, are derived from the estimates of the parameters $\phi_1 = \{\mathcal{K}, \Sigma, \theta\}$, obtained from 1-quarter ahead forecasts, and when h^* is 4, 5 or 6, the constant horizon forecasts are derived from parameters obtained from the 2-quarter ahead forecasts, etc. In the second step, we estimate the term structure parameters $\phi_2 = \{\rho_0, \rho_1, \lambda_0, \lambda_1\}$. All yields are assumed to be measured with error. The results of this step of the estimation are given in Table S2 of the Supplementary Appendix with two-step adjusted standard errors in parenthesis.[15]

From Equation 15 we see that forecasts with horizon h^* are modelled as an affine rotation of forecasts with horizon h, which, as seen from Equation 3, are in turn an affine rotation of the fundamental state vector X_t. An implication of this is that the estimation results should be invariant to the choice of forecast horizon: a yield curve model driven by a forward-looking state vector with horizon h^* is *observationally equivalent* to any other model given a different choice of h.[16] Thus, a strict interpretation of the model implies that $\bar{X}_{t,h}$, for all values of h, contains the exact same information in X_t, which would render a forward-looking model to be redundant. Naturally, the data tells us that this is not the case. To visualize the

[14]For example, given the 1-quarter ahead forecasts, if t corresponds to the beginning of January then $h(t) = 3$ and $h(t+1) = 2$, if t corresponds to February then $h(t) = 2$ and $h(t+1) = 1$ and if t corresponds to March then $h(t) = 1$ and $h(t+1) = 3$.

[15]The two-step estimation strategy and the two-step asymptotic standard error reported here closely follows Appendix D of Chun (2011).

[16]Any two such models will generate the same distributions for the short rate and bond prices as one model can be written as an *invariant affine transformation* of the other in the sense of Appendix A of Dai and Singleton (2000).

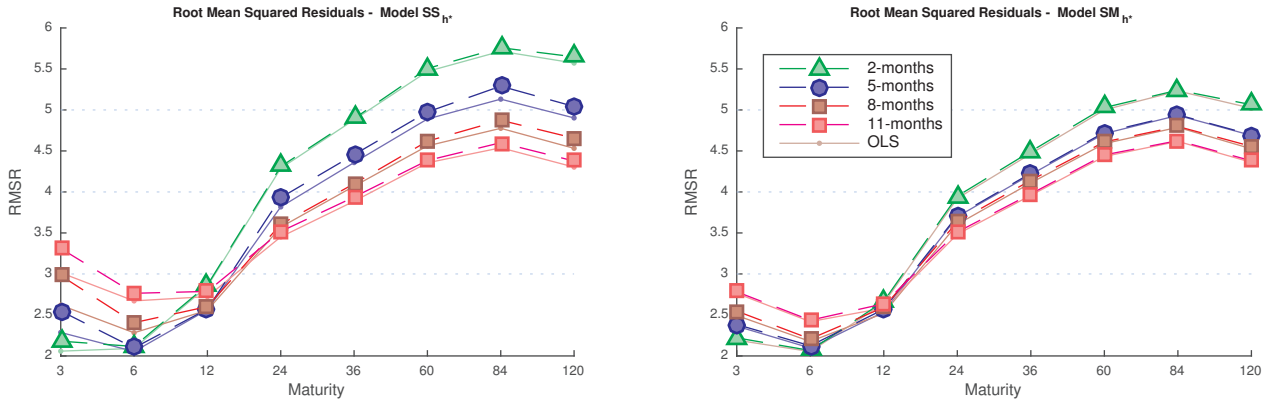

Figure 5: **Root Mean Squared Residuals.** The left panel shows the RMSRs (in basis points) for Models SS_{h^*}, with $h^* = 2$, 5, 8 and 11, and for yields with maturities of 3, 6, 12, 24, 36, 60 and 120 months. The right panel shows the RMSRs for Model SM_{h^*}, $h^* = 2$, 5, 8 and 11. As a benchmark, the corresponding OLS regression RMSRs are also shown as solid lines.

trade-off in a model's ability to fit long versus short maturity yields, the left panel in Figure 5 depicts RMSRs for four models with $h^* = 2$, 5, 8 and 11. To provide a connection with the OLS results in Panel A of Figure 1, the OLS RMSRs from regressing yields on the constant horizon forecasts are also plotted. These unrestricted OLS results serve as a benchmark reflecting the lower limit of what our affine term structure models, with their cross-equation restrictions, are able to explain.

When we look across the models estimated with the 1, 2, 3 or 4 quarter ahead forecast horizons, we see clearly that one set of estimates is not an invariant transformation of the other, as each generates different yield curve patterns. The best fit of the 3-month yield is obtained using the short-horizon forecast. Increasing h^* leads to a better fit at the long end of the yield curve, at the expense of fitting the short end of the yield curve. This is consistent with our intuition that information in short horizon forecasts should be reflected in short maturity yields, while the information in longer horizon forecasts should be reflected in long maturity yields. Yet note that a small incremental improvement in fitting the short end comes at a much larger cost in terms of fitting the long end. This evidence points to the presence of idiosyncratic, horizon specific information across the term structure of the survey forecasts.

4.2 Estimating Model SM_{h^*}

Our framework specifies a parsimonious, dynamic representation of a multiple-horizon system of expectations, based on a common set of parameters and a common set of stochastic shocks. If so, a natural question is whether the additional information available across the entire set of available horizons, help us better pin down the estimation of the volatility and the physical drift parameters?[17] In our setting, as the physical parameters determine the construction of the constant horizon forecasts, additional information used to estimate the parameters might also refine the construction of the factors in our yield curve model. Additional information across the surveys would also impact the model's performance by altering the estimates of the volatility parameters. Yet, and perhaps more importantly, as the $\mathbb{Q}$-parameters do not explicitly depend on the $\mathbb{P}$-parameters due to the lack of restrictions imposed on the market prices of risk, the ability to accurately estimate the physical parameters has implications for model implied measures of risk premia, such as the excess returns generated by the model. We will return to discuss this point later on in the article.

In this section, we introduce additional information from the surveys into the estimation. As in the previous section, we only let the h^*-horizon forecasts drive bond yields directly, however we leverage the information across the multi-horizon structure of the surveys, 1 to 4 quarters ahead, when inferring the physical drift and volatility parameters. We denote these models as SM_{h^*} models as the state equation is defined by an single h^*-month ahead constant maturity forecast, but estimated using information across multiple forecast horizons.

Recall that due to the time-varying pattern of horizons in the surveys, the vector $\bar{X}_{t,h^*}$ for any given h^* is only precisely observed every third month. During the remaining two months, it must be estimated. The estimation uses all four sets of observed forecasts, ranging from one to four quarters ahead. Each of these forecasts are related to the partially observed vector $\bar{X}_{t,h^*}$ by a relation of the form given in 15. Unlike the direct likelihood approach employed in the previous section, we map our model into a state space framework and estimate the parameters with the aid of a Kalman Filter derived likelihood function, extracting the smoothed estimates of the state vector as our estimate of $\bar{X}_{t,h^*}$. The maximum

[17] This line of thinking is consistent with the spirit of Kim and Orphanides (2012), who, within a latent variable framework, use Blue Chip survey forecasts of the 3-month yield to aid in pinning down the drift of the physical parameters.

likelihood estimates for our example with $h^* = 2$, 5, 8 and 11 are reported in Table S3 of the Supplementary Appendix. To streamline our discussion, we only summarize the estimation results here and provide the details of our estimation procedure in Appendix B.

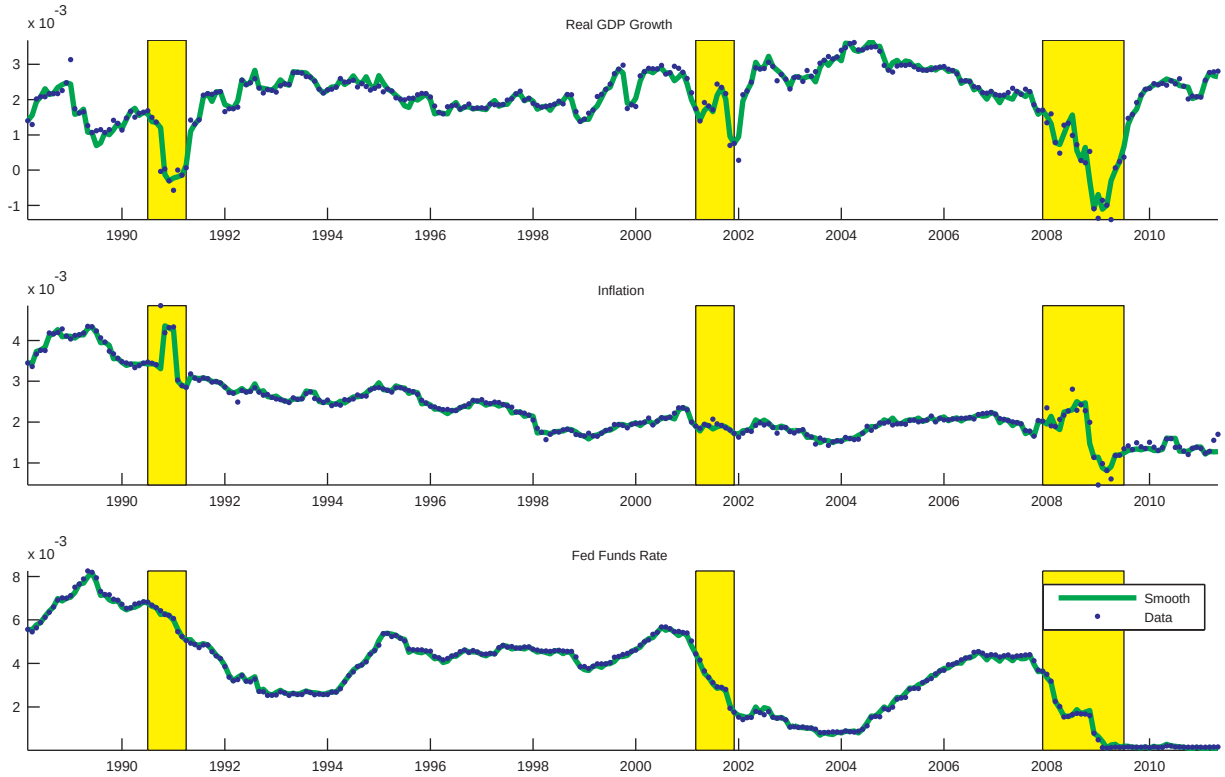

Figure 6: **Constant Horizon Forecasts - Model SM$_2$** As an example, the three panels in the figure plot the estimated constant 2-month horizon forecasts using information across multiple forecast horizons for each of the three model factors. The raw 1-quarter ahead forecasts from the survey are also shown as individual blue dots.

For each h^*, the smoothed estimate of the h^*-month ahead forecast is extracted conditioned on the entire sample. For example, Figure 6 shows the smoothed constant horizon forecasts for the case of $h^* = 2$. The 1-quarter ahead forecast from the surveys are given by the blue dots. Note that by construction, the smoothed 2-month ahead forecasts pass through every third point where it exactly matches the 1-quarter ahead forecast. During the remaining two months, the smoothed forecasts will deviate from the observed 1-quarter ahead forecast, reflecting the filtering and the additional information in the 2, 3 and 4 quarter ahead surveys. Letting $\bar{X}_{t,h^*}$ be the smoothed estimate of the h^*-month ahead forecast and following a two-step estimation procedure, the remaining term structure parameters are estimated conditioned on the first stage Kalman filter estimates of $\mathcal{K}$, θ and Σ. For $h^* =$ 2, 5, 8 and 11, the estimation results of the term structure parameters for Model SM$_{h^*}$ are

reported in Table S4 of the Supplementary Appendix.

In contrast with the RMSRs of Model SS_{h^*} in the left panel, the right panel of Figure 5 displays the RMSRs of Model SM_{h^*} along with the RMSRs of the corresponding OLS model that serves as a benchmark. Since the constant-horizon state vector is exactly observed every three months, the differences between the estimated state vectors across the two models are manifest every two out of three months. Overall we see that including additional horizons in the estimation, leads to some, but not dramatic, improvements relative to Model SS_{h*}. Regardless if we use a single horizon or multiple-horizons to perform the estimation, there is still clearly a tension in the model when simultaneously attempting to fit both the short and long ends of the yield curve.

The presence of latent factors infiltrates most of the macro-finance literature. Naturally, the inclusion of latent factors, which we purposefully avoid, could explain much of the remaining variation. For example, Orphanides and Wei (2012) build on a forward-looking 3-factor specification (forecasts are included in the estimation, but not in the state vector) and study the effect of adding an additional latent factor. By not including factors that are extracted from yields themselves (latent factors), our model allows us to gauge the extent to which one can explain the yield curve without resorting to using the yield curve itself.[18] In the next section we continue with this vision and in lieu of throwing in latent factors we build on the insights we have learned so far and propose extending the state vector across multiple horizons.

4.3 Multi-horizon Policy Rule and Yield Curve Model

In theory, the factors that enter the short rate equation, if to be interpreted a monetary policy rule, should well describe the short end of the curve. If the goal is to maximize our ability to explain the short rate, and its interpretation as a forward-looking monetary policy rule, then our results lend support to the choice of a short forecast horizon for the policy variables in Chun (2011). However, if we choose h^* large to better pin down the long end of the curve, we

[18] In Chun (2011), where the forecast horizon is (approximately) fixed at 3 months, a 4th factor is required to explain the dynamics of long yields. That 4th factor is the forecasted slope of the yield curve defined using survey forecasts of the 10-year maturity yield. Although this factor is driven by long run expectations of monetary policy, it is stretch to interpret this as a macro-factor, as it is by its very nature close to that of a latent slope factor.

need to be more cautious in interpreting the short rate equation as a policy rule. To resolve this fundamental tension in fitting both short and long yields simultaneously, we propose the novel idea of expanding the short rate equation to incorporate multiple forecast horizons. This is the main conceptual innovation of the paper and allows the model to directly link up information across distinct forecast horizons with movements in bond yields. In addition, we further embellish the modelling framework to allow for different sets of innovations across different forecast horizons.

4.3.1 Forward-looking Multiple-Horizon Monetary Policy Rule

A forward-looking central bank would naturally look at expectations across various horizons when setting the target rate. We capture this idea by generalizing the factors driving the instantaneous short rate equation to include multiple horizon forecasts:

$$r_t = \rho_0 + \rho_{h_1}\bar{X}_{t,h_1} + \cdots + \rho_{h_M}\bar{X}_{t,h_M}. \tag{16}$$

The short rate then takes on the interpretation of a *forward-looking multiple-horizon monetary policy rule*. Interpreting the instantaneous short rate as a proxy for the policy instrument allows for a more realistic approximation of the overnight target rate than, say, the 3-month short rate used in many macro-term structure studies. To the best of our knowledge, the idea of a forward-looking multi-horizon policy rule is novel to the literature on monetary policy rules, and certainly novel within the term structure literature.

In the application below, we simplifiy the exposition by examaining a special case selecting only the 2- and the 11-month ahead forecasts as policy variables:

$$r_t = \rho_0 + \rho_{11}\bar{X}_{t,11} + \rho_2\bar{X}_{t,2}. \tag{17}$$

4.3.2 A Dynamic Model of the Multiple-Horizon Forecasts

Simply extending the state-vector by stacking several forecast horizons together, all of which are driven by a common stochastic component, leads to a type of stochastic singularity, as movements in the forecasts of any horizon can be explained by another. To break this

singularity, we need to introduce a new set of stochastic shocks so as to allow for idiosyncratic variation across different forecast horizons. To illustrate, let's stack the 11- and 2-month horizon forecasts into the state vector $\bar{X}_t \equiv [\bar{X}'_{t,11}\ \bar{X}'_{t,2}]'$, and assume that its dynamics are driven by the following Gaussian stochastic differential under $\mathbb{P}$

$$d\bar{X}_t = d\begin{pmatrix}\bar{X}_{t,11}\\ \bar{X}_{t,2}\end{pmatrix} = \begin{pmatrix}\mathcal{K} & 0\\ 0 & \mathcal{K}\end{pmatrix}\left[\begin{pmatrix}\theta\\ \theta\end{pmatrix} - \begin{pmatrix}\bar{X}_{t,11}\\ \bar{X}_{t,2}\end{pmatrix}\right]dt + \begin{pmatrix}\Sigma_{11} & 0\\ \Sigma_2 & D_\sigma\end{pmatrix}\begin{pmatrix}dB_{F,t}\\ dB_{S,t}\end{pmatrix} \tag{18}$$

where the diagonal matrix D_σ in the bottom right of the volatility matrix is defined as

$$D_\sigma = \begin{pmatrix}\sigma_1 & 0 & 0\\ 0 & \sigma_2 & 0\\ 0 & 0 & \sigma_3\end{pmatrix}. \tag{19}$$

To be consistent with our existing framework, we would set D_σ equal to the zero matrix, and note that for all horizons, $\bar{X}_{t,h}$ is driven by the same set of innovations given by the Brownian motions, $dB_{F,t}$, with the horizon specific volatility matrices Σ_h. In resolving this stochastic singularity, we introduce D_σ into the volatility matrix, allowing an additional set of innovations $dB_{S,t}$ to enter the model, and specifying the correlation structure across the two sets of shocks. Note that the product of each diagonal element of the matrix D_σ with the corresponding element of $dB_{S,t}$, captures deviations of the actual forecast from our aforementioned model. Next, we assume that under the equivalent martingale measure, the dynamics of $\bar{X}_t = [\bar{X}'_{t,11}\ \bar{X}'_{t,2}]'$ follows

$$d\begin{pmatrix}\bar{X}_{t,11}\\ \bar{X}_{t,2}\end{pmatrix} = \begin{pmatrix}\mathcal{K}^{\mathbb{Q}}_{11} & 0\\ \mathcal{K}^{\mathbb{Q}}_{2,11} & \mathcal{K}^{\mathbb{Q}}_{2}\end{pmatrix}\left[\begin{pmatrix}\theta^{\mathbb{Q}}_{11}\\ \theta^{\mathbb{Q}}_{2}\end{pmatrix} - \begin{pmatrix}\bar{X}_{t,11}\\ \bar{X}_{t,2}\end{pmatrix}\right]dt + \begin{pmatrix}\Sigma_{11} & 0\\ \Sigma_2 & D_\sigma\end{pmatrix}\begin{pmatrix}dB^{\mathbb{Q}}_{F,t}\\ dB^{\mathbb{Q}}_{S,t}\end{pmatrix} \tag{20}$$

and these parameters along with the short rate coefficients are inferred in the second step of our two-step estimation procedure.

We define $dB_{F,t}$ as *fundamental shocks* to the model, as they play a central roll in driving all sets of forecasts (along with the underlying fundamental state vector X_t). On the other hand, $dB_{S,t}$ captures horizon specific innovations to market expectations, and hence we call them *sentiment shocks* as they capture deviations of market expectations from that

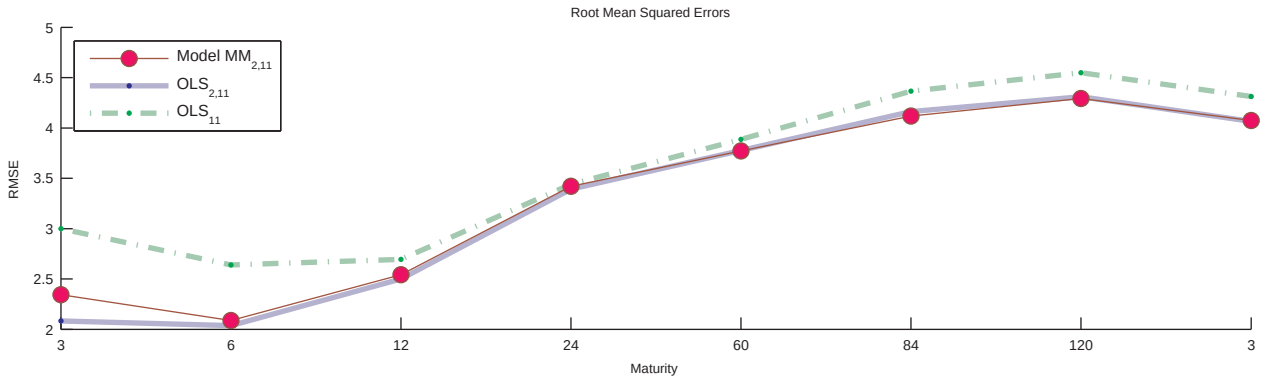

Figure 7: **Multiple-horizon Forward-looking Model $\mathbf{MM}_{2,11}$** The red circular makers show the RMSRs of the multiple-horizon model estimated with the 2- and 11-month ahead forecast horizons. As a benchmark, the OLS RMSR fit of the regression using the same factors are given by the shaded blue line. The RMSR from using only the 11 month forecast horizon (dashed green line) is also given for comparison.

generated by fundamental shocks. These shocks are a modelling device that effectively breaks the stochastic singularity in the system of equations. In the above volatility matrix, note that these shocks only affect the second set of forecasts through D_σ. So fundamental shocks affect forecasts across all horizons, while sentiment shocks are used to drive a wedge between forecasts across different horizons. We denote this multiple-horizon forward-looking model as Model $\text{MM}_{11,2}$.

4.3.3 Estimating Model $MM_{11,2}$

To facilitate the estimation, we proceed in several steps. In the first step, the state parameters $\phi_1 = \{\mathcal{K}, \theta, \Sigma, D_\sigma\}$ are estimated by constructing the sample log likelihood function using the Kalman Filter as in the previous section. This pins down the dynamics under the $\mathbb{P}$ measure and generates the *smoothed* constant horizon estimates for use as factors in the term structure model.

Figure 7 displays the RMSRs of this model for yields of different maturities. The RMSRs of the OLS regressions using both the 11 and 2 month ahead horizons ($\text{OLS}_{2,11}$) as well as for the model using only the 11 month ahead forecasts (OLS_{11}) are also provided as benchmarks. Our multi-horizon model shows considerable improvement in fitting the yield curve vis-à-vis what is obtainable using only the 11-month ahead forecasts. The model effectively pins down the short end of the yield curve, while slightly improving the explanatory power at

the long end. Without using a latent factor, we have effectively relaxed the tension between simultaneously fitting short and long yields by introducing multiple horizons directly into a forward-looking short rate equation. This essentially illustrates the descriptive limits of a term structure model using the full multi-horizon structure of our selected survey forecasts.

5 Interpreting the Results

Our goal is to link bond yields to expected inflation and real GDP growth. However, as we will see, the forecasted funds rate picks up nearly all of the variation in bond yields, in effect driving out the macro-forecasts. Thus we prioritize the role of the macro-forecasts in the model by rotating the state vector as follows. For Model $\text{MM}_{2,11}$, let $\bar{X}_t = [\bar{X}_t^{M\prime} \ \bar{X}_t^{F\prime}]'$ be the reordered original factors where $\bar{X}_t^M = [\bar{g}_{t,11} \ \bar{\pi}_{t,11} \ \bar{g}_{t,2} \ \bar{\pi}_{t,2}]'$ and $\bar{X}_t^F = [\bar{f}_{t,11} \ \bar{f}_{t,2}]'$. Using $Z_t = [1 \ \bar{X}_t^{M\prime}]$ to define the t^{th} row of Z in the projection matrix $P = Z(Z'Z)^{-1}Z'$ and letting $\bar{X}_t^{F\phi\prime} \equiv [\bar{f}_{t,11}^{\phi} \ \bar{f}_{t,2}^{\phi}]$ denote the t^{th} row of $\bar{X}^{F\phi} = (I - P)\bar{X}^F$, the rotated vector is given by $\bar{X}_t^{\phi} = [\bar{X}_t^{M\prime} \ \bar{X}_t^{F\phi\prime}]'$. The vector $\bar{X}_t^{F\phi}$ is orthogonal to the macro-forecasts in $\bar{X}_t^M$. We refer to this as an *orthogonal projection rotation*, further details of which are provided in Appendix C. Following Chun (2011), we can refer to the rotated funds rate factors as the anticipated monetary policy factors as they are the component of the fed funds rate forecasts that are orthogonal to the macro forecasts, effectively capturing a part of what would traditionally be considered a monetary policy shock. For the single-horizon models, $\bar{X}_t = [\bar{g}_{t,h^*} \ \bar{\pi}_{t,h^*} \ \bar{f}_{h^*}]'$ and $Z_t \equiv [1 \ \bar{g}_{t,h^*} \ \bar{\pi}_{t,h^*}]$ so that the rotated funds rate factor is given by $\bar{f}_{h^*}^{o} = (I - P)\bar{f}_{h^*}$.

First, we examine the factor loading structure across our model specifications, highlighting the increasingly prominent role of real GDP growth over longer forecast horizons. Second, we focus on the policy interpretation behind the multi-horizon forward-looking monetary policy rule. Finally, we conclude our analysis by looking at the models' implications for predictability of bond risk premia.

5.1 Interpreting the Yield Coefficients

We now examine the sensitivity of bond yields to each of the forward-looking factors as given by the coefficients $B(n)$ in the yield equation, $Y_t^{(n)} = A(n) + B(n)'\bar{X}_{t,h^*}$.

5.1.1 Factor Loadings: Single Horizon Models

Figure 8 plots the factor loadings for Model SS_{h^*} estimated for constant horizon forecasts with h^* equal to 2, 5, 8 and 11 months.[19] The left panel plots the loadings on the original specification. Note that short yields are primarily driven by the expected funds rate, while long yields are almost purely driven by expected inflation. However, it is natural to suspect that much of the explanatory power in the macro-forecasts are being driven out by the funds rate forecast.

The right panel plots the loadings on the orthogonal projection rotations, where the loadings have been rotated to give the two macro forecasts first priority in explaining yield dynamics. First, across all horizons h^*, a shock to forecasted inflation results in a parallel shift in the yield curve and thus clearly takes on the interpretation of a level factor. Second, exactly as in the left panel, the anticipated monetary policy factor has a slope effect in actuating only short maturity yields. Expected inflation is clearly the key driver of long bond yields.

Finally, note that when $h^* = 2$ the effect of forecasted real GDP growth is essentially obscured, loading negligibly across all maturities. However, as h^* increases, shocks to forecasted real activity becomes increasingly important.[20] Clearly, the impact of a shock to longer-horizon ($h^* = 11$) GDP growth is inversely related to the level of short maturity yields, yet this effect gradually dissipates with increasing maturity. The difference in pattern between the loadings on longer and short horizon forecasts is striking and has important implications for how we interpret the role of real output growth expectations in macro-finance models.

5.1.2 Is Expected GDP Growth a Hidden Factor?

Since our model reveals that short-horizons forecasts of real GDP growth have no immediate impact on the shape of the yield curve, they appear to be nearly 'hidden', in the spirit of Duffee (2011). However, longer horizon forecasts are clearly manifest in the cross section of

[19] Figure S1 of the Supplementary Appendix plots nearly similar factor loadings for Model SM_{h^*}.

[20] In slight contrast, based on essentially the same orthogonal rotation, with (approximately) $h^* = 3$ and over a shorter sample period ending in 2006 and hence excluding the most recent recession, Chun (2011) finds evidence of negative short rate coefficients on forecasts of real GDP growth.

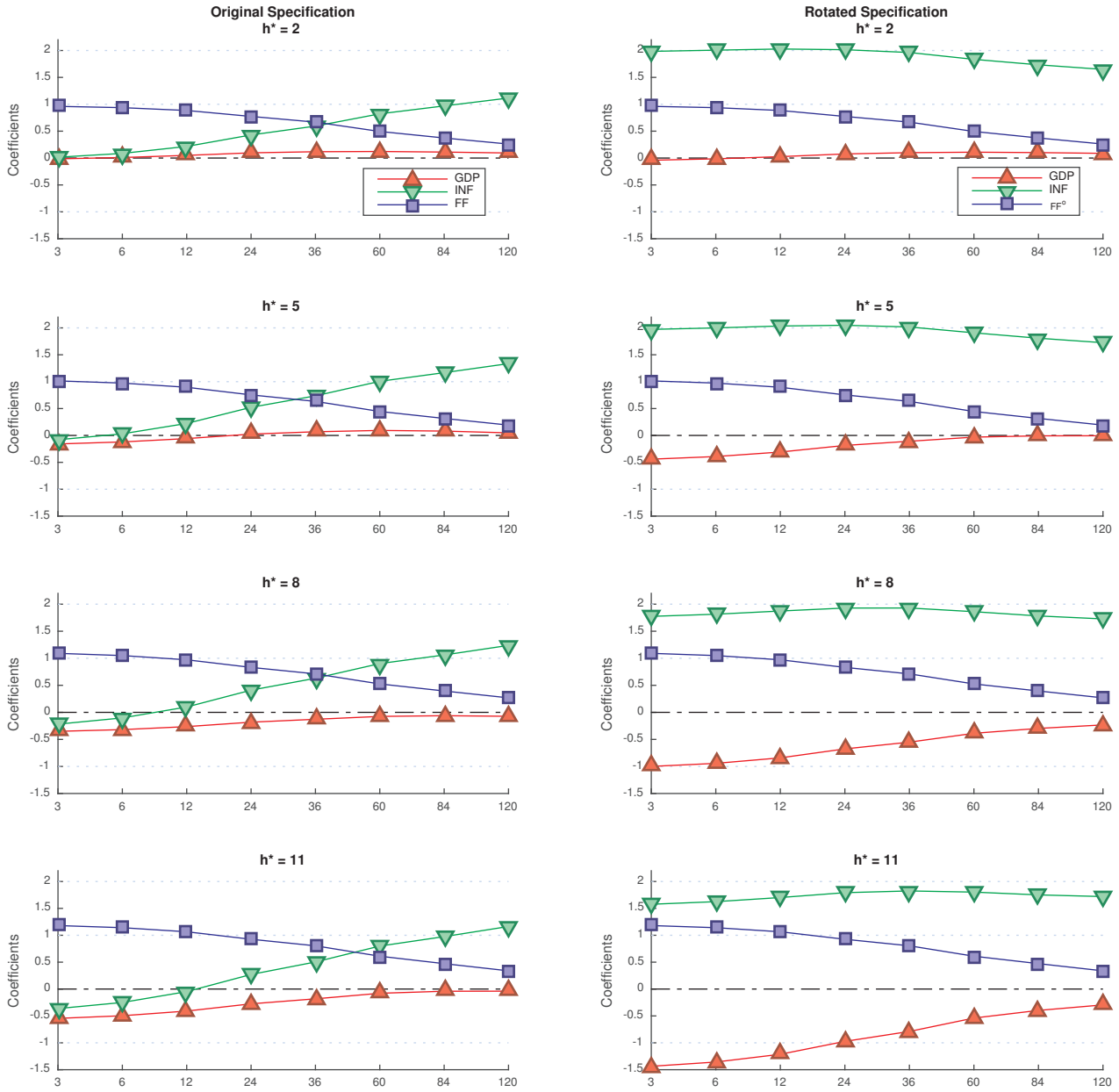

Figure 8: **Factor Loadings - SS_{h^*}** The figures plot the factor loadings for Model SS_{h^*} for $h^* = 2$, 5, 8 and 11 months. The left panel plots the loadings on the original specification, while the right panel plots the corresponding loadings on the orthogonal projection rotation specification which highlights the role of the two macro forecasts in explaining the yield curve. The x-axis gives the time to maturity.

yields. Thus an important finding is that, although short horizon output forecasts remain obscured in the cross section, longer horizon forecasts of real GDP growth are not hidden, as they are strongly linked to the slope of the yield curve. To further explore the hidden properties of the short horizon forecasts of real output growth we focus on the coefficient

matrix $B(n)$ in the bond yield equation $y(n) = A(n) + B(n)\bar{X}_{t,h^*}$, which takes on the form

$$B(n) = M_n\rho_1 \equiv \frac{1}{n}(\mathcal{K}^{Q\prime})^{-1}(I - e^{-\mathcal{K}^{Q\prime}n})\rho_1 \tag{21}$$

where $\mathcal{K}^Q$ denotes the mean reversion matrix under the equivalent martingale measure $\mathbb{Q}$. For forecasted real GDP growth to have no impact on the cross section of yields implies that the first element of $B(n)$, which corresponds to the loading on expected real output growth, be equal to zero. In the model of Duffee (2011), a particular normalization requires that changes in risk premia exactly offset expectations of future bond yields, such that the short rate process under $\mathbb{Q}$ and hence bond yields do not depend on the hidden factor.[21] In our setting, as seen from the above equation, the cross product of the first row of the matrix M_n with ρ_1 needs to be small in magnitude. In fact, the hidden nature of the first factor in our model when $h^* = 2$ is due to the insignificant and small estimate of ρ_{11} along with very small values in the last two elements of the first row of the matrix M_n.

Joslin, Singleton, and Priebsch (2014) investigate the unspanned component of realized macro factors, that by construction, only impact the $\mathbb{P}$ dynamics in a model using yield principal components as state variables. In contrast, our observable macro factors are not apriori designated as 'hidden', and can become manifest depending on the data. That said, the meaning of a 'hidden' factor in our context is somewhat different than in models with latent factors, as the first couple of yield principal components would essentially drive out most of the descriptive power of the macro-forecasts, rending all observable variables to be 'hidden.' In contrast, we go so far as to prioritize the role of the macro-forecasts in explaining the cross-section via our orthogonal projection rotations. Short horizon forecasts of real output growth remain 'hidden' despite this rotation to bring them out into the open.[22]

[21] Unlike the general framework of Duffee (2011) keeping a factor hidden in the cross-section is independent of the corresponding element of ρ_1 being zero, as $\mathcal{K}^Q$ is not normalized to be diagonal in our case.

[22] Interestingly, Chernov and Mueller (2012) find evidence in support of expected inflation being a partly hidden factor, which at first may appear incongruent with our evidence. In fact, this is probably due to the original, non-rotated inflation forecasts having a slope effect on the yield curve. The presence of a latent slope factor would drive out expected inflation, in essence, obscuring it in the cross-section.

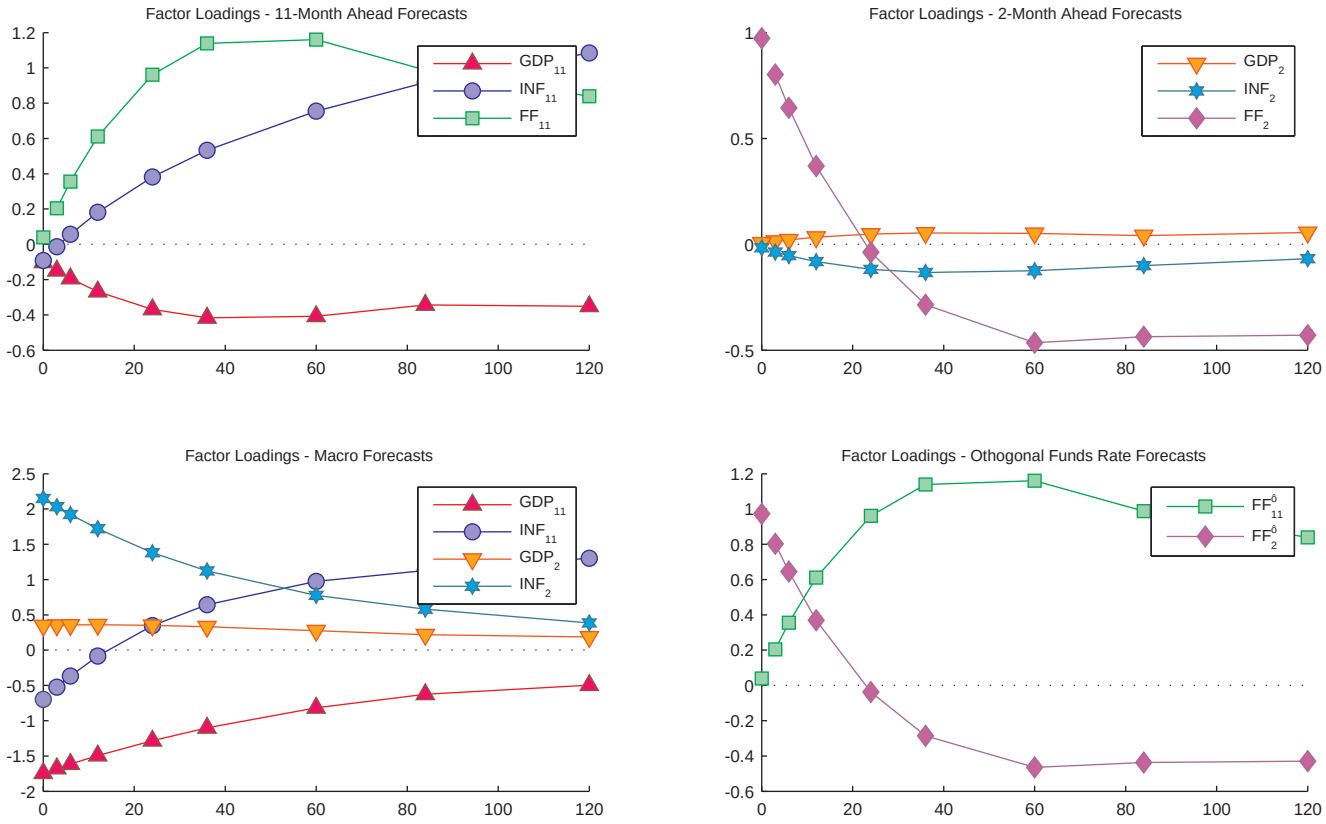

Figure 9: **Factor Loadings - MM$_{11,2}$** The upper panel displays the loadings from the original specification. The bottom panel plots the factor loadings on the rotated model with the four macro forecasts on the left and the two orthogonalized funds rate forecasts on the right. The short rate coefficients also are plotted along the y-axis. Time to maturity (in months) is on the x-axis.

5.1.3 Factor Loadings: Model MM$_{11,2}$

The yield factor loadings for the original specification of Model MM$_{11,2}$ are plotted in the top two panels of Figure 9, while the bottom two panels plots their corresponding orthogonal projection rotations. The multiple horizon policy rule (short rate) coefficients are plotted along the y-axis. As previously noted the original loadings in the top panel are difficult to interpret due to the dominance of the two funds rate factors. Hence we focus the remainder of our attention on the bottom panels, which plot the factor loadings on our preferred rotation. The loadings on the four macro forecasts reflect the fact that they are given first priority in explaining the yield curve, as do the loadings on the two orthogonalized funds rate forecasts in bottom right panel.

Although on average a level factor, short ($h^* = 2$) and longer horizon ($h^* = 11$) forecasts

of inflation act like slope factors, albeit each tilting the yield curve in different directions. Short horizon forecasts of inflation decrease the slope by pulling up short rates. Intuitively, this is what one would expect when thinking about the response of contractionary monetary policy to high levels of inflation. In contrast, longer horizon forecasts of inflation increase the slope by shifting long rates upward and short yields downward. Following recessions, near term inflationary expectations tend to be lower than longer term expectations, and such an episode is typically characterized by lower rates accompanied by higher short rate expectations about the future.

Similar to the pattern exhibited by the single horizon model, the 11-month ahead real output forecast is clearly not 'hidden' as it plays a prominent role in explaining the slope of the term structure. This is consistent with the interpretation that the central bank accommodates longer horizon growth expectations by lowering rates. Short horizon output growth forecasts, albeit exhibiting a slight positive effect, are the only factor that could plausibly be interpreted as being 'hidden' in the cross selection of yields.

Finally, we shift our attention to the two orthogonalized funds rate forecasts that have been purged, in a linear sense, of the information in the four macro forecasts. In contrast with the single horizon models, the inclusion of two funds rate factors allows the model to disentangle the effect of longer-horizon expectations from that of short-horizon expectations of monetary policy. The 2-month ahead forecast sharply shifts short rates positively, while having a negative impact on longer maturity rates. The positive impact on short maturity yields is consistent with its role as an anticipated shock in the monetary policy rule. The 11-month ahead loadings on the funds rate forecast appear as the mirror image of the 2-month ahead loadings, which shows that the longer horizon forecasts actuate long yields while having essentially no effect on short maturity yields. This makes intuitive sense, as conditioned on the 2-month ahead forecasts, longer horizon forecasts of monetary policy should be linked to information embedded in the long end of the yield curve, while providing little information about short term interest rates.

5.2 Interpreting the Policy Rule Coefficients

Any changes to the short rate are propagated across the entire yield curve, therefore the entire yield curve should reflect the policy actions of the Federal Reserve. Thus information across the entire spectrum of yields may be informative for estimating the policy rule parameters.[23] We focus on the policy rule coefficients implied by the multiple horizon model as given below:

$$r_t = 0.0002 - .10\bar{g}_{t,11} - 0.01\bar{\pi}_{t,11} + 0.04\bar{f}_{t,11} + 0.01\bar{g}_{t,2} - 0.02\bar{\pi}_{t,2} + 0.97\bar{f}_{t,2}. \tag{22}$$

However the ability to interpret this as a policy rule is impaired by the fact that the forecasted funds rate twice appears on the right hand side. As to be expected, the 2-month ahead forecast of the federal funds rate picks up nearly all of the variation in the short rate, in effect driving out the macro-forecasts. In contrast, a proper interpretation of the central bank's reaction function should focus on forecasted inflation and real activity. Thus we prioritize the role of the macro-forecasts in the model and rotate the state vector to highlights the role of expected inflation and real output growth in the policy rule, while allowing the funds rate forecasts to explain the remaining variation. The policy rule coefficients on our transformed model take on the form:

$$r_t = 0.0034 - 1.74\bar{g}_{t,11} - 0.69\bar{\pi}_{t,11} + 0.35\bar{g}_{t,2} + 2.15\bar{\pi}_{t,2} + 0.04\bar{f}^{\phi}_{t,11} + 0.97\bar{f}^{\phi}_{t,2} \tag{23}$$

The magnitude and sign of the different coefficients on the longer and short horizon forecasts of the macro-variables shed light on the degree to which the central bank is forward looking. Although we characterize this equation as a policy rule, note the absence of an explicit monetary policy shock term — the rotated funds rate forecasts in $\bar{X}_t^{F\phi'} \equiv [\bar{f}^{\phi}_{t,11} \;\; \bar{f}^{\phi}_{t,2}]$ are, in fact, picking up the anticipated components of monetary policy, much of what previous studies would call a monetary policy shock (Christiano, Eichenbaum, and Evans (1999)). Furthermore, these anticipate policy factors have the added interpretation of capturing the possible effect of other, omitted policy variables (besides inflation and output growth) on the target rate.

Longer horizon ($h^* = 11$) forecasts of real output growth figure prominently in the short

[23] Ang, Dong, and Piazzesi (2007) and Chun (2011) estimate policy rules using information in the entire yield curve.

rate equation as the coefficient is large and strikingly negative. In contrast, short horizon forecasts of real output growth are only weakly related to the short rate. During economic downturns, when the economy is below trend and thus expected to grow again in the future, longer-run growth forecasts tend to be much higher than short-run forecasts. Meanwhile, the central bank continues to stimulate the economy by lowering rates, leading to a starkly negative coefficient on longer-horizon forecasts. In fact, when market participants expect higher growth rates in the future, the monetary authority may be accommodating these longer run expectations with lower interest rates. Over short horizons, however, expectations about economic growth are correlated with the current state of the economy. Thus, in bad times, market participants expect low growth rates to continue in the near future, and this is coincident with an expansionary monetary policy regime, yielding a positive coefficient on short-run expectations.

Moreover, short and longer horizon forecasts of inflation have opposing effects on the short rate. Consistent with a pre-emptive central bank policy, short run forecasts of inflation play a positive role on the short rate with a relatively large coefficient of 2.15. In contrast, longer horizon forecasts of inflation have a negative impact on the short rate with a coefficient of -0.69. A forward-looking central bank has to offset the need for stimulating the economy in the short run with concerns of potentially generating higher inflation in the long run.

Finally, note that the 2-month ahead forecast of the funds rate, and by construction the corresponding orthogonal factor $\bar{f}^{\phi}_{t,2}$, captures what is essentially a one-to-one response in the short rate. In contrast, the 11-month ahead funds rate forecast is economically insignificant. This is not at all surprising as longer horizon forecasts of the funds rate should not embody significant information about the current short rate once information in the short horizon forecast is accounted for. On its own, the short horizon funds rate forecast essentially drives out all of the other variables, and for this reason performing the orthogonal projection rotation is important for providing an economically meaningful interpretation of the coefficients.

The advantage of our multiple horizon model is in allowing for a greater clarity of differentiation between longer and short horizons effects of macro forecasts on the policy rule, and consequently on the yield curve itself. Central bankers should be interested in these patterns as they describe the potential impact of forward guidance on the shape of the yield curve. In

addition to guiding expectations about the future path of the macroeconomy via statements and announcements, by influencing long rate expectations, policy makers may effectively convey information about the path of monetary policy shocks. We find that expectations of longer horizon shocks are linked to the slope of the yield curve, thus credibly managing these expectations could provide a key to influencing long yields in a manner that is consistent with the policymaker's objectives.

5.3 An Analysis of Predictable Excess Returns

The anticipated excess holding period return should, to a significant degree, reflect a risk premium for holding an n-period risky bond for m-periods. Given the estimated parameters, there are a couple of ways to generate excess returns, depending largely on the choice of the conditioning information set. First, model-implied expected excess returns can be computed by conditioning only on $\bar{X}_{t,h}$, that is by calculating both model-implied yields, and the model-implied expected yield, $E_t^{\mathbb{P}}[Y_{t+m}^{(n-m)}]$. We will refer to this method as the $\mathcal{F}_t^{\bar{X}}-$*conditional expected excess return* where all yield quantities are model-implied. In addition, by decreasing the holding period length, in the limit, we can obtain a model-implied *expected instantaneous excess return*. The specific expressions used to compute the various measures of excess returns are provided in Appendix D.

5.3.1 Factor Loadings on Model-implied Risk-Premia

Model-implied expected excess returns are an affine function of $\bar{X}_{t,h}$ and Figure 10 plots the factor loadings on the $\mathcal{F}_t^{\bar{X}}-$conditional expected excess return equations for Models SS_{11} and SM_{11} in the left panel and Model $\text{MM}_{11,2}$ in the right panels. The dashed lines with solid markers depict loadings for the instantaneous excess returns implied by the model, while the corresponding solid lines with open markers show the model-implied 12-month ahead expected excess returns. The larger the absolute magnitude of the coefficient, the more important the corresponding factor is in explaining excess returns. Note that forecasted GDP growth is the dominant driver of model-implied risk premia in longer horizon models (Models SS_{11} and SM_{11}). In the multi-horizon Model $\text{MM}_{11,2}$ real GDP growth appears to be the most important source of time variation in bond risk premia.

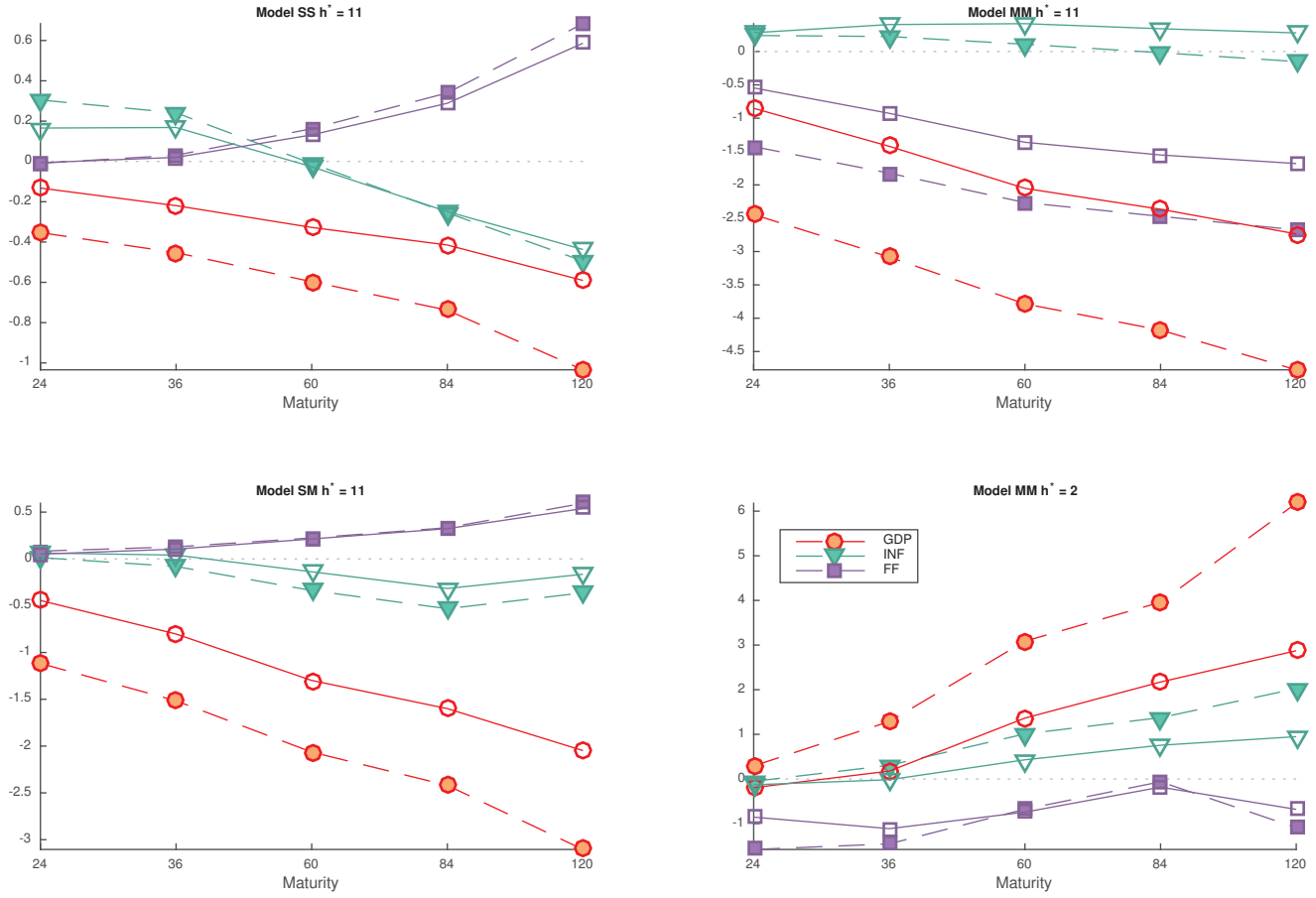

Figure 10: **Factor Loadings on $\mathcal{F}_t^{\bar{X}}$−conditional Model-Implied Expected Excess Returns** The plots depict the factor loadings for both instantaneous excess returns (dashed line, solid markers) as well as model-implied $m = 12$ month excess holding period returns (solid lines, open markers). Each plot shows expected excess return coefficients for yields with maturities from 24 to 120 months. Coefficients on model-implied expected excess returns for the single horizon Models SS_{11} and SM_{11} are depicted in the left panel and for the two sets of horizons in Model $\text{MM}_{2,11}$ in the right panel.

Our findings complement the existing evidence of a relation between macro factors and bond risk premia: Ludvigson and Ng (2009) find that the cyclical fluctuations in bond risk premia are linked to macroeconomic activity that is orthogonal to information in the yield curve. Chun (2011) finds evidence that market prices of risk are linked to the expected state of the economy – noting, in particular, that forecasted GDP growth appears to be an important driver of model-implied bond risk premia. Complementary evidence is provided by Dick, Schmeling, and Schrimpf (2013) who find a positive link between SPF survey expectations about real GDP growth and term premia expectations.[24]

[24]In a parallel finding in the equity markets, Cooper and Priestly (2009) find evidence suggesting excess stock return predictability using the output gap.

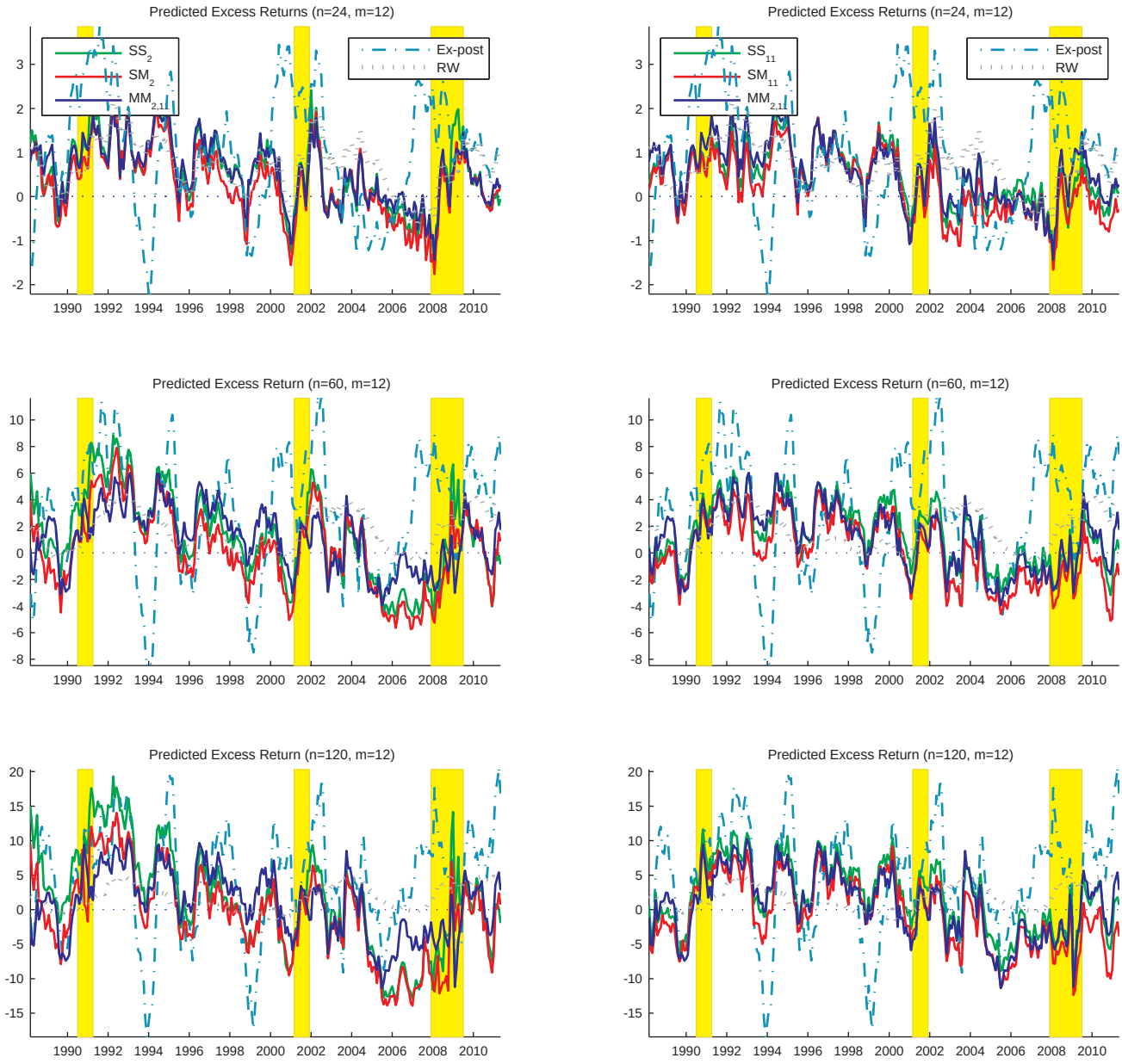

Figure 11: **Ex-ante Expected Excess Returns** The plots display 12-month holding period $\mathcal{F}_t^{Y,\bar{X}}$−conditional expected excess returns, calculated using the model-generated 12-month ahead expectation. All contemporaneous yields quantities are observed and not derived from the model. In each panel these conditional expected excess returns are plotted for Model SS_{h^*} (green), Model SM_{h^*} (red) and for Model $\text{MM}_{2,11}$ (blue). The top panel plots the excess returns for the 24-month bond, the middle panel for the 60-month bond and the bottom panel for the 120-month bond. As a benchmark the ex-post realized excess log returns are plotted as dashed lines. The lightly dotted line plots the predicted excess returns using the random walk forecast in lieu of the 12-month ahead expectation when calculating the expected return.

5.3.2 Predictability of Excess Returns

Due to the presence of fitting errors in the model-implied yield curve, the $\mathcal{F}_t^{\bar{X}}-$conditional excess returns are not necessarily what investors would predict if given access to our models. Rather, an investor could use yields that are contemporaneously observed, using the models to generate only the expectation, $E_t^{\mathbb{P}}[Y_{t+m}^{(n-m)}]$. We will refer to this method as the $\mathcal{F}_t^{Y,\bar{X}}-$*conditional expected excess return*, as these excess returns are computed by conditioning on both $\bar{X}_{t,h}$ and the observed yield curve.

Figure 11 plots the $\mathcal{F}_t^{Y,\bar{X}}-$conditional expected excess return series, which uses observed, rather than model-implied, bond yields. The left panel shows the expected excess returns implied by Model SS_2 (green), Model SM_2 (red) and Model $\text{MM}_{2,11}$ (blue), while the right panel shows the plots for Model SS_{11}, Model SM_{11} and Model $\text{MM}_{2,11}$ (repeated). The ex-post realized excess returns are plotted as dashed lines and serve as a benchmark. The predicted excess returns using the random walk forecast, which takes contemporaneously observed yields to define the expectation in the excess return expression, is denoted RW and plotted as dotted lines. Note that all the excess return series implied by the three models track each other for the most part. The magnitudes of the excess returns are increasing in the maturity of the bond. In general the predicted series are smaller in amplitude relative to the realized excess return series., which is intuitive given that expected excess returns are not predictions of extreme fluctuations.

Panel A of Table 1 reports RMSRs as measures of how well the $\mathcal{F}_t^{Y,\bar{X}}-$conditional expected excess returns forecast realized excess returns for bonds with maturities of 24, 48, 60 and 120 months. This measure gives an idea as to which model does the best at predicting the level of excess returns. A few patterns are discernible. First note that the measure derived from the random walk model described above (final column) does nearly as well as the most informative models, especially for shorter maturities. Second, note that the SS class of models generates lowers RMSRs than their SM counterparts, with Model SS_{11} best describing the level of excess returns. One might have expected that models using multiple horizon forecasts to estimate the $\mathbb{P}$ dynamics of the survey forecasts would do better at predicting excess returns. However the use of multiple forecast horizons in Models SM and MM does not appear to help the models better describe the level of excess returns.

Next, to examine the ability of the models to predict the total variation in excess returns, we linearly regress excess returns on their predictions and look at the R^2s. Panel B reports the R^2s from regressing the realized ex-post excess returns on the $\mathcal{F}_t^{Y,\bar{X}}-$conditional expected excess returns. In contrast with its ability to predict the level of excess returns the measure derived from the random walk model does poorly at predicting the variation in excess returns. Interestingly, as with the level of excess returns, Model SS_{11} appears to best capture the variance of excess returns, obtaining the highest adjusted R^2s on average. The final row suggests that up to 20 percent of the variation in excess returns of the 10-year bond can be predicted by the $\mathcal{F}_t^{Y,\bar{X}}-$conditional expectations generated by our models.[25] Once again, including additional information from multiple horizons in Models SM and MM does not appear to help the models better describe the variation in excess returns.

Panel C reports the R^2s from a linear regression of the realized ex-post excess returns on the $\mathcal{F}_t^{Y,\bar{X}}-$conditional excess returns and a set of forward rates up to 10 years. By including forward rates in the regressions, we can ascertain the extent to which our model-implied excess returns contain information not in the yield curve. The final column presents the adjusted R^2 from regressing excess returns on only the forward rates. As we can see, for all of our models, the adjusted R^2s are in fact higher when augmenting the forward rate regressions with our conditional expected excess returns. Based on information from the Blue Chip surveys, Le and Singleton (2013) suggest that a significant amount of variation in predicted excess returns can be attributed to unspanned risks, that is risks unrelated to information in the yield curve. Our use of forwards up to 10 years is consistent with their evidence that long maturity forwards beyond 5 years embed important information for predicting excess returns, recalling that Cochrane and Piazzesi (2005) only employ forwards up to 5 years in their regressions. We conclude that a part of our model generated expectations are also capturing a component of risk premia unrelated to information in forward rates and hence the yield curve. The idea that that macro factors explain excess returns that are not explained by information in the yield curve is discussed in Duffee (2011), Joslin, Singleton, and Priebsch (2014) and Ludvigson and Ng (2009). Once again, conditional on information in forward rates, including additional information from multiple horizons in the estimation

[25]The goal of our estimations is not to maximize the model's ability to explain excess returns, rather the excess returns explained are a by product of our estimated model and serves as a check on the reasonableness of the model-generated expectations.

Table 1: Measures of Predictability of Excess Returns

Panel A: Root Mean Squared Errors

	Model				
Maturity	SS_{11}	SM_{11}	SS_2	SM_2	$MM_{2,11}$	RW
24	0.015	0.016	0.014	0.016	0.014	0.014
60	0.046	0.053	0.050	0.054	0.049	0.045
84	0.059	0.069	0.068	0.073	0.064	0.060
120	0.076	0.089	0.092	0.097	0.082	0.080

Panel B: $AdjR^2$ of Excess Returns Regressions

	Model				
Maturity	SS_{11}	SM_{11}	SS_2	SM_2	$MM_{2,11}$	RW
24	0.028	0.032	0.053	0.027	0.032	0.010
60	0.108	0.092	0.090	0.092	0.069	0.053
84	0.146	0.126	0.096	0.114	0.113	0.042
120	0.201	0.170	0.119	0.142	0.185	0.063

Panel C: $AdjR^2$ of Excess Returns Regressions with Forwards Rates

	Model				
Maturity	SS_{11}	SM_{11}	SS_2	SM_2	$MM_{2,11}$	Forwards
24	0.239	0.239	0.246	0.239	0.239	0.216
60	0.248	0.246	0.237	0.236	0.238	0.212
84	0.245	0.241	0.219	0.221	0.234	0.195
120	0.288	0.281	0.235	0.242	0.278	0.212

Panel A reports the RMSE of the difference between the predicted excess returns for our five selected models and the ex-post realized excess returns. The RMSE using the random walk forecast of the yield in the excess return expression is also given in the final column. Panel B reports the R^2s from projecting realized ex-post excess returns on predicted expected excess returns. The final column of Panel C reports the Adjusted R^2s from projecting realized excess returns on ten forward rates with horizons up to 10 years; the first three columns of Panel C report the Adjusted R^2s from projecting realized excess returns on the predicted (fitted) values from the forward rate regression and the predicted excess returns based on each of our selected models.

does not appear to help the models better describe the level or variation in excess returns.

6 Conclusion

We build forward-looking models linking the shape of the yield curve to the multi-horizon structure of macroeconomic expectations, attributing the dynamics behind interest rate movements to the forward-looking nature of central bank policy. Our modeling framework and subsequent empirical analysis offer several significant innovations. Our set-up is designed to accommodate survey data plagued by a time-varying forecast horizon, and allows for the state parameters to be estimated in a manner that is consistent across the multi-horizon structure of the survey forecasts.

We rotate the models to highlight the role of the macro-forecasts and find that short horizon expectations best capture short maturity yields, while longer horizon expectations best explain long maturity yields. We find that short-horizons forecasts of real output growth are nearly obscured in the cross-section of yields, suggestive of the hidden properties of real activity in Duffee (2011); in stark contrast, however, we find that longer horizon forecasts of real output growth are not hidden, but clearly manifest in the slope of the yield curve. Our main conceptual innovation is the introduction of a *forward-looking multiple-horizon monetary policy rule*. This novel idea facilitates the decomposition of both monetary policy and the yield curve into short and longer horizon expectations of the macroeconomy. Although forecasted inflation stands out as a level factor responsible for driving long yields in the single horizon models, when introducing multiple horizons, it admits a decomposition into two slope components: short horizon expectations are tied to short yield movements while longer horizon expectations primarily drive long yields movements. Our analysis of model-implied bond risk premia suggest that up to 20 percent of the variation in excess returns can be capture by our models. In addition, expectations generated by our models are able to explain a greater amount of variation in ex-post excess returns than comparable regressions based on forward rates.

In theory, our framework is general enough so that the factors can be taken from survey data, as well as from forecasts derived from financial market data, as well as including

historical data. Although we illustrate our framework in a Gaussian setting, the estimation of a model admitting state-dependent volatility may potentially provide a richer framework for the study of risk premia. Admittedly, our models do not capture the nuisance of every factor affecting yields, such as the forces driving supply and demand, including those created by government policies such as quantitative easing; nor have we modeled liquidity or the market's perceptions about the credit worthiness of the US government. We abstract, in some sense, from most of these realities, and focus our attention on forecasts of a few key variables central to the monetary policy channel. In this sense our models are not completely specified, yet we believe such an abstraction is justified in light of the abundant insights that the models provide in the absence of latent factors.

Finally, our models provide a framework for central bankers to study the impact of forward guidance on the yield curve. To the extent that policy makers can credibly influence not only macroeconomic expectations, but longer horizon short rate expectations via statements and announcements, they can alter the shape of the yield curve. Our framework ties movements in bond yields to not only the multiple horizon structure of market expectations, but also to the risk premia as by generated these expectations. Financial markets participants would naturally also be interested in our models, as they provide insights on how the yield curve reacts to changes in market expectations due to both endogenous as well as monetary policy induced shocks. A more detailed exploration of these linkages is reserved for future research.

A Proofs of the Model Propositions

We assume that the dynamics of the vector X_t are described by the stochastic differential equation in (1), except in Corollaries 2 and 3 where we assume a Gaussian version, $dX_t = \mathcal{K}(\theta - X_t)\,dt + \Sigma d\mathcal{B}_t$, where $\mathcal{K}$ is assumed to be a diagonalizable NxN matrix, θ is an N-vector, $\Sigma\Sigma'$ is an NxN variance-covariance matrix and $\mathcal{B}_t$ is an N-dimensional Brownian motion under the physical measure $\mathbb{P}$.[26]

[26]Note the assumption that the square matrix $\mathcal{K}$ is *diagonalizable*, that is there exists a matrix U with linearly independent eigenvectors of $\mathcal{K}$ (so that U is invertible) in its columns and a diagonal matrix Λ containing the eigenvalues of $\mathcal{K}$ such that $\mathcal{K} = U\Lambda U^{-1}$ holds. This assumption is innocuous as a small perturbation in the matrix elements will render the matrix diagonalizable without significantly affecting the

Lemma 1. *Define* $\Psi(h) \equiv \frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}h}\left[I - e^{\mathcal{K}k}\right]$ *then*

$$(i) \quad \frac{1}{k}\int_h^{h+k} e^{-\mathcal{K}\tau}d\tau = \Psi(h) \tag{24}$$

$$(ii) \quad \Psi(h)\,\mathcal{K}\Psi(\ell)^{-1} = e^{-\mathcal{K}(h-\ell)}\mathcal{K} \tag{25}$$

Proof. (i) Since $\mathcal{K}$ is assumed to be diagonalizable, then $\mathcal{K} = U\Lambda U^{-1}$ and it follows that $e^{-\mathcal{K}\tau} = Ue^{-\Lambda\tau}U^{-1}$ for any scalar constant τ.[27] For a diagonal matrix $\int_h^{h+k} e^{-\Lambda(\tau)}d\tau = \Lambda^{-1}\left[e^{-\Lambda(h)} - e^{-\Lambda(h+k)}\right]$, hence

$$\begin{aligned}
\int_h^{h+k} e^{-\mathcal{K}(\tau)}d\tau &= U\int_h^{h+k} e^{-\Lambda\tau}d\tau\; U^{-1} \\
&= U\Lambda^{-1}\left(e^{-\Lambda(h)} - e^{-\Lambda(h+k)}\right)U^{-1} \\
&= U\Lambda^{-1}U^{-1}\left[Ue^{-\Lambda(h)}U^{-1} - Ue^{-\Lambda(h+k)}U^{-1}\right] \\
&= \mathcal{K}^{-1}\left[e^{-\mathcal{K}(h)} - e^{-\mathcal{K}(h+k)}\right] = \Psi(h)
\end{aligned}$$

where the 3rd line inserts the identity matrix, $U^{-1}U$, and the final line uses the identity $\mathcal{K}^{-1} = U\Lambda^{-1}U^{-1}$ which follows from the diagonalizable assumption on $\mathcal{K}$.

(ii) Note that $\Psi(h)^{-1} = \left[I - e^{-\mathcal{K}(k)}\right]^{-1}e^{\mathcal{K}(h)}\mathcal{K}k$, for any h. It follows that

$$\begin{aligned}
\Psi(h)\,\mathcal{K}\Psi(\ell)^{-1} &= \frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}(h)}\left[I - e^{-\mathcal{K}(k)}\right]\mathcal{K}\left[I - e^{-\mathcal{K}(k)}\right]^{-1}e^{\mathcal{K}(\ell)}\mathcal{K}k \\
&= \frac{1}{k}\mathcal{K}^{-1}\mathcal{K}e^{-\mathcal{K}(h)}\left[I - e^{-\mathcal{K}(k)}\right]\left[I - e^{-\mathcal{K}(k)}\right]^{-1}e^{\mathcal{K}(\ell)}\mathcal{K}k \\
&= e^{-\mathcal{K}(h-\ell)}\mathcal{K}
\end{aligned}$$

where the 2nd and 3rd lines simply follow from repeatedly making use of the property, $\mathcal{K}e^{-\mathcal{K}(h)} = e^{-\mathcal{K}(h)}\mathcal{K}$.[28]

□

mean reversion path implied by $\mathcal{K}$.

[27] We will use the identity $[U\Lambda U^{-1}]^2 = U\Lambda U^{-1}U\Lambda U^{-1} = U\Lambda^2U^{-1}$ that generalizes to $[U\Lambda U^{-1}]^j = U\Lambda^jU^{-1}$, and the definition of a matrix exponential $e^{\mathcal{K}} = \sum_{j=0}^{\infty}\frac{1}{j!}\mathcal{K}^j$. Taking the exponential of the matrix $\mathcal{K}\tau$ we have $e^{\mathcal{K}\tau} = e^{U\Lambda\tau U^{-1}} = \sum_{j=0}^{\infty}\frac{1}{j!}[U\Lambda\tau U^{-1}]^j = \sum_{j=0}^{\infty}\frac{1}{j!}U(\Lambda\tau)^jU^{-1} = U[\sum_{j=0}^{\infty}\frac{1}{j!}(\Lambda\tau)^j]U^{-1} = Ue^{\Lambda\tau}U^{-1}$.

[28] From the definition of the exponential of a matrix $\mathcal{K}$, we have $\mathcal{K}e^{-\mathcal{K}} = \mathcal{K}\sum_{j=0}^{\infty}\frac{1}{j!}\left(-\mathcal{K}^j\right) = \sum_{j=0}^{\infty}\frac{1}{j!}\mathcal{K}(-\mathcal{K})^j = \sum_{j=0}^{\infty}\frac{1}{j!}(-\mathcal{K})^j\mathcal{K} = e^{-\mathcal{K}}\mathcal{K}$.

Proposition 1. *Let* $\bar{X}_{t,h} \equiv E^{\mathbb{P}}\left[\frac{1}{k}\int_{t+h}^{t+h+k} X_s ds|X_t\right]$ *denote the time* t *conditional* h*-period ahead forecast under* $\mathbb{P}$ *of the average of* X_s *over a fixed time interval of length* k*, then* $\bar{X}_{t,h}$ *is a linear combination of the random walk forecast* X_t *and the long run mean* θ

$$\bar{X}_{t,h} = (1 - \Psi(h))\,\theta + \Psi(h)\,X_t, \tag{26}$$

where the weight on the random walk forecast is given by $\Psi(h) \equiv \frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}h}\left[I - e^{\mathcal{K}k}\right]$.

Proof. As shown in Appendix B of Duan and Simonato (1999), for a stochastic process following (1), the conditional expectation of X_s given X_t is given by

$$E^{\mathbb{P}}[X_s|X_t] = \theta + e^{-\mathcal{K}(s-t)}\,(X_t - \theta) \tag{27}$$

Using this expression we obtain[29]

$$\begin{aligned}
E^{\mathbb{P}}\left[\frac{1}{k}\int_{t+h}^{t+h+k} X_s ds|X_t\right] &= \frac{1}{k}\int_{t+h}^{t+h+k} E^{\mathbb{P}}[X_s|X_t]ds \\
&= \frac{1}{k}\int_{t+h}^{t+h+k} \theta + e^{-\mathcal{K}(s-t)}\,(X_t - \theta)\,ds \\
&= \theta + \frac{1}{k}\left[\int_{h}^{h+k} e^{-K\tau}d\tau\right](X_t - \theta) \\
&= \theta + \Psi(h)\,(X_t - \theta)
\end{aligned}$$

where the last line follows from Lemma 1(i). □

Proposition 2. *Let* $\bar{X}_{t,h} \equiv E^{\mathbb{P}}\left[\frac{1}{k}\int_{t+h}^{t+h+k} X_s ds|X_t\right]$ *and* $\bar{X}_{t,\ell} \equiv E^{\mathbb{P}}\left[\frac{1}{k}\int_{t+\ell}^{t+\ell+k} X_s ds|X_t\right]$ *denote, respectively, the time* t *conditional* h*-period and* ℓ*-period ahead forecasts of the average of* X_s *over a fixed time interval of length* k*, then there exists a* $\mathbb{P}$*-consistent relation between these two forecasts such that*

$$\bar{X}_{t,h} = \theta + e^{-\mathcal{K}(h-\ell)}\left(\bar{X}_{t,\ell} - \theta\right). \tag{28}$$

Proof. Let $\Psi(h) = \frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}h}\left[I - e^{-\mathcal{K}k}\right]$ then from Proposition 1, we have $\bar{X}_{t,h} = (I - \Psi(h))\,\theta +$

[29] Note that when $h = 0$, this proposition reduces to the derivation seen in Appendix C of Chun (2011).

$\Psi(h) X_t$ and $X_t = \Psi(\ell)^{-1} [\bar{X}_{t,\ell} - (I - \Psi(\ell))\theta]$. Therefore,

$$\begin{aligned}
\bar{X}_{t,h} &= (I - \Psi(h))\theta + \Psi(h)\Psi(\ell)^{-1}[\bar{X}_{t,\ell} - (I - \Psi(\ell))\theta] \\
&= \theta - \Psi(h)\theta + \Psi(h)\Psi(\ell)^{-1}\bar{X}_{t,\ell} - \Psi(h)\Psi(\ell)^{-1}\theta + \Psi(h)\theta \\
&= \theta - e^{-\mathcal{K}(h-\ell)}\theta + e^{-\mathcal{K}(h-\ell)}\bar{X}_{t,\ell}
\end{aligned}$$

where the final equality follows from the identity $\Psi(h)\Psi(\ell)^{-1} = e^{-\mathcal{K}(h-\ell)}$, which is a special case of Lemma 1 (ii). □

Corollary 1. *Let* $\bar{X}_{t,h} \equiv E^{\mathbb{P}}\left[\frac{1}{k}\int_{t+h}^{t+h+k} X_s ds | X_t\right]$, *then the* $\mathbb{P}$*-dynamics of* $\bar{X}_{t,h}$ *follows the following stochastic differential equation*

$$d\bar{X}_{t,h} = \mathcal{K}[\theta - \bar{X}_{t,h}]dt + \Sigma_h d\mathcal{B}_t \tag{29}$$

where $\Sigma_h = \Psi(h)\Sigma(X_t) = \frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}h}\left[I - e^{-\mathcal{K}k}\right]\Sigma(X_t)$.

Furthermore, the conditional expectation of $\bar{X}_{s,h}$ *given* $\bar{X}_{t,h}$ *is*

$$E^{\mathbb{P}}[\bar{X}_{s,h}|\bar{X}_{t,h}] = \theta + e^{-\mathcal{K}(s-t)}\left(\bar{X}_{t,h} - \theta\right). \tag{30}$$

Proof. Let $\Psi(h) = \frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}(h)}\left[I - e^{-\mathcal{K}(k)}\right]$ then from Proposition 1, we have $\bar{X}_{t,h} = (I - \Psi(h))\theta + \Psi(h)X_t$, which is a twice continuously differential function of the Ito process given by (1). Noting that $\frac{\partial \bar{X}_{t,h}}{\partial t} = 0$, $\frac{\partial^2 \bar{X}_{t,h}}{\partial X_t^2} = 0$, $\frac{\partial \bar{X}_{t,h}}{\partial X_t} = \Psi(h)$, and invoking Ito's lemma we have

$$d\bar{X}_{t,h} = \Psi(h)\mathcal{K}(\theta - X_t)dt + \Psi(h)\Sigma(X_t)d\mathcal{B}_t.$$

Since $X_t = \Psi(h)^{-1}[\bar{X}_{t,h} - (I - \Psi(h))\theta]$, we can write the drift of the above process as

$$\begin{aligned}
\Psi(h)\mathcal{K}(\theta - X_t) &= \Psi(h)\mathcal{K}\left(\theta - \Psi(h)^{-1}[\bar{X}_{t,h} - (I - \Psi(h))\theta]\right) \\
&= \Psi(h)\mathcal{K}\left(\theta - \Psi(h)^{-1}\bar{X}_{t,h} + \Psi(h)^{-1}\theta - \theta\right) \\
&= \Psi(h)\mathcal{K}\Psi(h)^{-1}\left(\theta - \bar{X}_{t,h}\right) \\
&= \mathcal{K}\left(\theta - \bar{X}_{t,h}\right)
\end{aligned}$$

where the final equality follows from the property $\Psi(h)\mathcal{K}\Psi(h)^{-1} = \mathcal{K}$ obtained by setting

$\ell = h$ in Lemma 1 (ii).

Having established that the dynamics of $\bar{X}_{t,h}$ are governed by (29), by direct analogy, substitute $\bar{X}_{t,h}$ for X_t into (27) to obtain the expression for the conditional expectation.

□

Corollary 2. *If X_t follows a Gaussian version of (1), then the conditional variance-covariance matrix $V_t(\bar{X}_{s,h})$ can be expressed using matrix vectorization as*[30]

$$vec\left(V_t(\bar{X}_{s,h})\right) = (K \otimes I + I \otimes K)^{-1}\left(I - e^{-(K\otimes I + I\otimes K)(s-t)}\right) vec(\Sigma_h \Sigma_h') \tag{31}$$

where $\Sigma_h = \Psi(h)\,\Sigma = \frac{1}{k}\mathcal{K}^{-1}e^{-\mathcal{K}h}\left[I - e^{-\mathcal{K}k}\right]\Sigma$.

When $s = t+1$, it immediately follows that the process in (29) can be discretized as

$$\bar{X}_{t+1,h} = \theta + e^{-\mathcal{K}}\left(\bar{X}_{t,h} - \theta\right) + V_t(\bar{X}_{t+1,h})^{1/2}\epsilon_{t+1}. \tag{32}$$

where ϵ_t is a standard Gaussian white noise vector.

Proof. Assuming a Gaussian version of the stochastic process for X_t, we first derive an expression for the variance-covariance matrix $V_t(X_s)$ of the state vector X_s conditioned on X_t. The expression for $E[X_s|X_t]$ is given in (27). We follow Fisher and Gilles (1996) by fixing s, and using Ito's lemma we obtain[31]

$$X_s = E[X_s|X_s] = E[X_s|X_t] + \int_t^s dE[X_s|X_v] = E[X_s|X_t] + \int_t^s e^{-\mathcal{K}(s-v)}\Sigma dB_v$$

[30] It is worth mentioning that this customized formulation of the conditional variance matrix requires evaluating only a single matrix exponent, which by far is the most computationally intensive component of all the required matrix operations.

[31] Making use of the partial derivatives of the conditional expectation in (27)

$$\frac{\partial E[X_s|X_t]}{\partial t} = e^{-\mathcal{K}(s-t)}\mathcal{K}[X(t) - \theta],\ \frac{\partial E[X_s|X_t]}{\partial X_t} = e^{-\mathcal{K}(s-t)}\mathcal{K},\ \frac{\partial^2 E[X_s|X_t]}{\partial X_t^2} = 0$$

and applying Ito's Lemma we note that the drift disappears $[e^{-\mathcal{K}(s-t)}\mathcal{K}(\theta - X_t) + e^{-\mathcal{K}(s-t)}\mathcal{K}(X_t - \theta)]dt = 0$ so that $dE[X_s|X_t] = e^{-\mathcal{K}(s-t)}\Sigma dB_t$. Expressing this in integration notation yields the result.

and taking the conditional variance of both sides, we obtain[32].

$$\begin{aligned} V_t[X_s] &= V_t\left[\int_t^s e^{-\mathcal{K}(s-v)}\Sigma d\mathcal{B}_v\right] \\ &= E_t\left[\int_t^s e^{-\mathcal{K}(s-v)}\Sigma\Sigma' e^{-\mathcal{K}(s-v)'}dv|X_t\right] \\ &= \int_t^s e^{-\mathcal{K}(s-v)}\Sigma\Sigma' e^{-\mathcal{K}(s-v)'}dv \end{aligned}$$

Left as an integral the above expression is not computationally all that tractable, thus we roughly follow the outline of Fackler (2000) in obtaining an explicit expression. First we vectorize the expression as[33]

$$\begin{aligned} vec(V_t[X_s]) &= \int_t^s vec(e^{-\mathcal{K}(s-v)}\Sigma\Sigma' e^{-\mathcal{K}(s-v)'})dv \\ &= \int_t^s \left[e^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)}\right] vec(\Sigma\Sigma')dv \end{aligned}$$

Next we differentiate with respect to s to obtain the differential equation[34]

$$\begin{aligned} \frac{d[vec(V_t[X_s])]}{ds} &= (e^{-\mathcal{K}(0)}\otimes e^{-\mathcal{K}(0)})vec(\Sigma\Sigma') + \int_t^s \frac{d[e^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)}]vec(\Sigma\Sigma')}{ds}dv \\ &= (I\otimes I)vec(\Sigma\Sigma') + \int_t^s \left[\frac{d[e^{-\mathcal{K}(s-v)}]}{ds}\otimes e^{-\mathcal{K}(s-v)} + e^{-\mathcal{K}(s-v)}\otimes\frac{d[e^{-\mathcal{K}(s-v)}]}{ds}\right] vec(\Sigma\Sigma')dv \\ &= vec(\Sigma\Sigma') + \int_t^s \left[-Ke^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)} + e^{-\mathcal{K}(s-v)}\otimes -Ke^{-\mathcal{K}(s-v)}\right] vec(\Sigma\Sigma')dv \\ &= vec(\Sigma\Sigma') + \int_t^s \left[-\left(K\otimes I + I\otimes K\right)\left(e^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)}\right]\right. vec(\Sigma\Sigma')dv \\ &= vec(\Sigma\Sigma') - \left(K\otimes I + I\otimes K\right)\int_t^s \left[\left(e^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)}\right]\right. vec(\Sigma\Sigma')dv \\ &= vec(\Sigma\Sigma') - \left(K\otimes I + I\otimes K\right)vec(V_t[X_s]). \end{aligned}$$

[32]The middle equality follows from the fact the integrated process has mean zero. The final equality follows from an extension of the Ito isometry to multi-dimensions, a proof of which can be found in the Appendix to Fackler (2000)

[33]We use the relation $vec(ABC) = (A'\otimes C)vec(B)$.

[34]For the second line, we use the relation $\frac{dA(s)\otimes B(s)}{ds} = \frac{dA(s)}{ds}\otimes B(s) + A(s)\otimes\frac{dB(s)}{ds}$. The trick to obtaining the fourth line is the relation $AC\otimes BC = (A\otimes C)(B\otimes C)$. It follows that $-Ke^{-\mathcal{K}(s-v)}\otimes Ie^{-\mathcal{K}(s-v)} = -(K\otimes I)(e^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)})$ and I $e^{-\mathcal{K}(s-v)}\otimes -Ke^{-\mathcal{K}(s-v)} = -(I\otimes K)(e^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)})$ thus we can write $-Ke^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)} + e^{-\mathcal{K}(s-v)}\otimes -Ke^{-\mathcal{K}(s-v)} = -\left(K\otimes I + I\otimes K\right)\left(e^{-\mathcal{K}(s-v)}\otimes e^{-\mathcal{K}(s-v)}\right)$.

Let $A = (K \otimes I + I \otimes K)$, and making use of the integrating factor e^{As} we obtain

$$e^{As}\frac{d[vec(V_t[X_s])]}{ds} + Ae^{As}vec(V_t[X_s]) = e^{As}vec(\Sigma\Sigma')$$

the integral of which is

$$\begin{aligned} e^{As}vec(V_t[X_s]) &= \left[\int e^{As}ds\right] vec(\Sigma\Sigma') + c \\ &= A^{-1}e^{As}vec(\Sigma\Sigma') + c \\ &= A^{-1}e^{As}vec(\Sigma\Sigma') - A^{-1}e^{At}vec(\Sigma\Sigma') \end{aligned}$$

where we have used the initial condition $vec(V_t[X_t]) = 0$ (as the variance is equal to zero when $s = t$) to solve for $c = -A^{-1}e^{At}vec(\Sigma\Sigma')$. Thus we obtain

$$\begin{aligned} vec(V_t[X_s]) &= A^{-1}vec(\Sigma\Sigma') - A^{-1}e^{A(t-s)}vec(\Sigma\Sigma') \\ &= A^{-1}[I - e^{A(t-s)}]vec(\Sigma\Sigma') \\ &= (K \otimes I + I \otimes K)^{-1}\left(I - e^{-(K\otimes I + I\otimes K)(s-t)}\right)vec(\Sigma\Sigma') \end{aligned}$$

The conditional variance-covariance matrix $V_t(\bar{X}_{s,h})$ follows by direct analogy by substituting Σ_h for Σ into the above expression. □

Corollary 3. *The density function of $\bar{X}_{s,\ell}$ conditional on $\bar{X}_{t,h}$ is given by*

$$\tilde{f}\left(\bar{X}_{s,\ell}|\bar{X}_{t,h}\right) = f\left(\theta + e^{-\mathcal{K}(h-\ell)}\left(\bar{X}_{s,\ell} - \theta\right) \mid \bar{X}_{t,h}\right)\|e^{-\mathcal{K}(h-\ell)}\| \tag{33}$$

where

$$f\left(\theta + e^{-\mathcal{K}(h-\ell)}\left(\bar{X}_{s,\ell} - \theta\right)|\bar{X}_{t,h}\right) = (2\pi)^{-N/2}|\Omega_{\bar{X}}|^{-\frac{1}{2}}e^{-\frac{1}{2}\left(\epsilon_t^X\right)'\Omega_X^{-1}\left(\epsilon_t^X\right)} \tag{34}$$

$$\epsilon_t^X = e^{-\mathcal{K}(h-\ell)}\left(\bar{X}_{s,\ell} - \theta\right) - e^{-\mathcal{K}(s-t)}\left(\bar{X}_{t,h} - \theta\right) \tag{35}$$

and where the conditional variance matrix $\Omega_{\bar{X}} \equiv V_t(\bar{X}_{s,h})$ can be vectorized as in (31).

Proof. Fix a forecast horizon h. From the multivariate Gaussian density function and Corol-

lary 1, the conditional density function of $\bar{X}_{s,h}$ given $\bar{X}_{t,h}$ is

$$f\left(\bar{X}_{s,h}|\bar{X}_{t,h}\right) = (2\pi)^{-N/2}\left|\Omega_{\bar{X}}\right|^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\epsilon_t^X\right)'\Omega_{\bar{X}}^{-1}\left(\epsilon_t^X\right)}. \tag{36}$$

Using a standard change-of-variables method, we obtain $\tilde{f}\left(\bar{X}_{s,\ell}|\bar{X}_{t,h}\right) = f\left(\bar{X}_{s,h}|\bar{X}_{t,h}\right)\|J\|$ where $\|J\|$ denotes the absolute value of the Jacobian determinant of the transformation. From Proposition 2 we have, $\bar{X}_{s,h} = \theta + e^{-\mathcal{K}(h-\ell)}\left(\bar{X}_{s,\ell} - \theta\right)$ and the Jacobian in this case is given by $J = \partial\bar{X}_{s,h}/\partial\bar{X}'_{s,\ell} = e^{-\mathcal{K}(h-\ell)}$. Thus the density function is given by

$$\tilde{f}\left(\bar{X}_{s,\ell}|\bar{X}_{t,h}\right) = f\left(\theta + e^{-\mathcal{K}(h-\ell)}\left(\bar{X}_{s,\ell} - \theta\right) \mid \bar{X}_{t,h}\right)|det\left(e^{-\mathcal{K}(h-\ell)}\right)|. \tag{37}$$

Since $E^{\mathbb{P}}[\bar{X}_{s,h}|\bar{X}_{t,h}] = [\theta + e^{-\mathcal{K}(s-t)}\left(\bar{X}_{t,h} - \theta\right)]$, we can write $\epsilon_t^X = \bar{X}_{s,h} - E^{\mathbb{P}}[\bar{X}_{s,h}|\bar{X}_{t,h}]$ as

$$\epsilon_t^X = \theta + e^{-\mathcal{K}(h-\ell)}\left(\bar{X}_{s,\ell} - \theta\right) - [\theta + e^{-\mathcal{K}(s-t)}\left(\bar{X}_{t,h} - \theta\right)] \tag{38}$$

$$= e^{-\mathcal{K}(h-\ell)}\left(\bar{X}_{s,\ell} - \theta\right) - e^{-\mathcal{K}(s-t)}\left(\bar{X}_{t,h} - \theta\right) \tag{39}$$

and the conditional variance takes on the expression given in (31). □

B Kalman Filter Estimation of the $\mathbb{P}$ Parameters

We define $\xi_t \equiv [\bar{X}_{t,h^*} - \theta]$ where h^* is fixed to be the forecast horizon driving bond yields in the model. From Corollary 2 of Appendix A, the discretized dynamics of $\bar{X}_{t,h^*}$ can be expressed as our state equation

$$\xi_{t+1} = \bar{X}_{t+1,h^*} - \theta = e^{-\mathcal{K}}[\bar{X}_{t,h^*} - \theta] + v_{t+1} \tag{40}$$

where v_t is Gaussian, white noise innovation term with a variance-covariance matrix $Q \equiv V_t(\bar{X}_{t+1,h^*})$ that is expressed as a column vector in (31).

Given the time-varying pattern of horizons in the Blue Chip Financial surveys, note that the vector $\bar{X}_{t,h^*}$ for values of h^* between 1 and 12 is only observed every third month. During the remaining 2 months, it is unobserved, however we do observe $\bar{X}_{t,h(t,q)}$, for $q = 1$ to 4, where $h(t,q)$ defines a mapping that gives the horizon, in months, of the q-quarter ahead

forecast horizon. Thus $h(t,q)$ translates a q-quarter ahead horizon to an h-month ahead horizon at each discrete time step t. From Proposition 2, the relation between the observed q-quarter ahead forecast, $\bar{X}_{t,h(t,q)}$, and the h^*-horizon forecast, $\bar{X}_{t,h^*}$, can be expressed to include a measurement error

$$\bar{X}_{t,h(t,q)} = \theta + e^{-\mathcal{K}(h(t,q)-h^*)}\left[\bar{X}_{t,h^*} - \theta\right] + w_t^q, \tag{41}$$

where w_t^q is a forecast horizon-dependent, Gaussian white noise error vector. Stacking these observed forecasts, the observation or measurement equation is given as

$$\begin{bmatrix} \bar{X}_{t,h(t,1)} \\ \vdots \\ \bar{X}_{t,h(t,4)} \end{bmatrix} = \begin{bmatrix} \theta \\ \vdots \\ \theta \end{bmatrix} + \begin{bmatrix} e^{-\mathcal{K}(h(t,1)-h^*)} \\ \vdots \\ e^{-\mathcal{K}(h(t,4)-h^*)} \end{bmatrix} \xi_t + w_t. \tag{42}$$

where $w_t = [w_t^{1\prime} \ldots w_t^{4\prime}\ ]'$ is by assumption a Gaussian white noise vector with a diagonal variance matrix $E[w_t w_t'] = R_t$.[35] This diagonal assumption on R_t implies that the disturbances, or the deviation of the actual forecast from the relation implied by the model, are uncorrelated across the different forecast horizons. To capture horizon-dependent volatility, we allow for a different volatility parameter for each variable and forecast horizon. Thus, R_t is time-varying, where the same matrix reappears every 3rd month. During the month when the state vector is observed, the measurement errors, and hence the variance of the errors, on the corresponding equations will be taken to be zero.

To estimate the parameters $\mathcal{K}$, θ, Σ and the diagonals of the three R_t matricies, the log of the likelihood function obtained via the Kalman filter recursions is maximized with respect to the unknown parameters.[36] Obtaining the global maximum of the likelihood function can

[35] Although we choose not to do this, by including the current quarter ($q = 0$) forecast or observed lags in the observation equation, such as $\bar{X}_{t,h(t,-L)}$, would allow the estimation to also consider historical information.

[36] We map the set up of the state space representation described here to the textbook treatment in Hamilton (1994), with the state equation is $\xi_{t+1} = F\xi_t + v_{t+1}$ where $F = e^{-K}$ and the measurement equation given by $y_t = (1_{M\times 1} \otimes \theta) + H_t'\xi_t + w_t$. Forecasts of ξ_{t+1} and its associated MSE, $P_{t+1|t}$, are based on iterating $\hat{\xi}_{t+1|t} = F\hat{\xi}_{t|t-1} + K_t\left(y_t - (1_{M\times 1}\otimes\theta) - H_t'\hat{\xi}_{t|t-1}\right)$ and $P_{t+1|t} = (F - K_tH_t')\,P_{t|t-1}\,(F' - H_tK_t') + K_tR_tK_t' + Q$ where the Kalman gain matrix K_t is defined as $K_t \equiv FP_{t|t-1}H_t\left(H_t'P_{t|t-1}H_t + R_t\right)^{-1}$. The Kalman filter is initialized by setting $\hat{\xi}_{1|0} = 0$ and $vec\left(P_{1|0}\right) = [I - (F\otimes F)]^{-1}\,vec\,(Q)$, with the constraint that the eigenvalues of F lie inside the unit circle. These recursions gives us the forecast $\hat{y}_{t+1|t} = \theta + H_t'\hat{\xi}_{t+1|t}$ and the MSE of the forecast of $E[\left(y_{t+1} - \hat{y}_{t+1|t}\right)\left(y_{t+1} - \hat{y}_{t+1|t}\right)'] = H_t'P_{t+1|t}H_t + R_t$ and given these conditional

be rather tricky. To reduce the likelihood that our final parameter estimates are not the result of an optimization algorithm having converged to a local optimum, we polish off each estimation step by employing a simulated annealing algorithm as described in Goffe, Ferrier, and Rogers (1994). Our experience is that the simulated annealing step always increases final likelihood values over that obtained via a Nelded-Mead simplex algorithm employed with grid of multiple starting values. Finally, we follow Hamilton (1994) to extract the smooth estimates of $\xi_t \equiv [\bar{X}_{t,h^*} - \theta]$, to which we add back the estimate of θ to back out the smoothed state vector, $\bar{X}_{t,h^*}$.

C Orthogonal Projection Rotations

We refer to the procedure of constructing a new state vector $X^o = [X_1 \; X_2^o]$ from the original vector $X = [X_1 \; X_2]$ an *orthogonal projection rotation* if $X_2^o = (I - P)X_2$ is the residual vector from an orthogonal projection of X_2 on the space spanned by $Z = [1 \; X_1]$, where $P = Z(Z'Z)^{-1}Z'$ is the usual orthogonal projection matrix. In other words, X_2^o is simply the residual vector from a linear regression, representing X_2 purged of the linear influence of X_1. It is the case that X^o is an affine rotation of X, such that $X_t^o = v + LX_t$. Specific to our single-horizon setting where $X = [g \; \pi \; f]$, we let $X_1 = [g \; \pi]$ and $X_2 = f$, and expressing $f_t = B_0 + B_g g_t + B_\pi \pi_t + \epsilon_t$ we have

$$X_t^o = \begin{bmatrix} g_t \\ \pi_t \\ f_t^o \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -B_0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -B_g & -B_\pi & 1 \end{bmatrix} \begin{bmatrix} g_t \\ \pi_t \\ f_t \end{bmatrix}. \tag{43}$$

This can be seen from $f_2^o = (I - P)f = f - Z(Z'Z)^{-1}Z'f = f - [1 \; g \; \pi]B$ where the vector $B = [B_0 \; B_g \; B_\pi]' = (Z'Z)^{-1}Z'f$. More generally, as in the case of the multi-horizon version

moments, the sample likelihood function is

$$f\left(y_{t+1}|y_t\right) = (2\pi)^{-N/2} \left|H_t' P_{t+1|t} H_t + R_t\right|^{-\frac{1}{2}} e^{-\frac{1}{2}\left(y_t - \theta - H_t'\hat{\xi}_{t+1|t}\right)'\left(H_t' P_{t+1|t} H_t + R_t\right)^{-1}\left(y_t - \theta - H_t'\hat{\xi}_{t+1|t}\right)}.$$

of the model where we let $X_1 = [g_{11}\ \pi_2\ g_2\ \pi_2]'$ and $X_2 = [f_{11}\ f_2]'$, the rotation becomes

$$X_t^o = \begin{bmatrix} X_1 \\ X_2^o \end{bmatrix} = \begin{bmatrix} 0_{k\times 1} \\ -B_0 \end{bmatrix} + \begin{bmatrix} I_{k\times k} & 0_{k\times 1} \\ -B_{1:k} & 1_{2\times 1} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}. \tag{44}$$

where B_0' now denotes the first row (the intercepts) and the matrix $B_{1:k}'$ the remaining rows of the 3×2 coefficient matrix $B = (Z'Z)^{-1}Z'X_2$.

It follows that a yield curve model based on the state vector X is an invariant affine transformation, as described in Appendix A of Dai and Singleton (2000), of a yield curve model based on the state vector X^o as both models generate observationally equivalent distributions for bond yields. Thus we can estimate either model and rotate one set of parameters to obtain the other set.[37]

D Model-implied Expected Excess Returns

Conditional on knowing the model parameters and on $\bar{X}_{t,h}$, we can define the model-implied, $\mathcal{F}_t^{\bar{X}}-$conditional expected excess log return from holding an n-period bond over m-months as

$$\begin{aligned} xr_{t,t+m}^{(n)} \equiv & E^{\mathbb{P}}\left[log\left(\frac{P_{t+m}^{(n-m)}}{P_t^{(n)}}\right)\right] - mY_t^{(m)} && (45) \\ = & -(n-m)E^{\mathbb{P}}\left[Y_{t+m}^{(n-m)}\right] + nY_t^{(n)} - mY_t^{(m)} && (46) \\ = & -(n-m)A(n-m) - (n-m)B(n-m)'E_t^{\mathbb{P}}\left[\bar{X}_{t+m,h}\right] \\ & + nA(n) + nB(n)'\bar{X}_{t,h} - mA(m) - mB(m)'\bar{X}_{t,h} && (47) \end{aligned}$$

[37] A application of Ito's lemma shows that given the physical drift and volatility parameters of the original model $\Phi_X = \{\mathcal{K}_X, \theta_X, \Sigma_X\}$, the transformed model parameters are $\Phi_{X^o} = \{L\mathcal{K}_X L^{-1}, L\theta_X + v, L\Sigma_X\}$. The risk neutral drifts have the exact same analogous transformation, that is, $\Phi_X^{\mathbb{Q}} = \{\mathcal{K}^{\mathbb{Q}}{}_X, \theta_X^{\mathbb{Q}}\}$ is transformed as $\Phi_{X^o}^{\mathbb{Q}} = \{L\mathcal{K}_X^{\mathbb{Q}}L^{-1}, L\theta_X^{\mathbb{Q}} + v\}$, which when parameterized using market prices of risk results in $\lambda_X = \{\lambda_0, \lambda_1\}$ being transformed as $\lambda_{X^o} = \{\lambda_0 - \lambda_1 L^{-1}v, \lambda_1 L^{-1}\}$. The short rate is similarly transformed from $\rho_X = \{\rho_0, \rho_1\}$ to $\rho_X^o = \{\rho_0 - \rho_1 L^{-1}v, \rho_1 L^{-1}\}$.

Evaluating the conditional expectation as $E_t^{\mathbb{P}}\left[\bar{X}_{t+m,h}\right] = \left[I - e^{-Km}\right]\theta + e^{-Km}\bar{X}_{t,h}$, we can rewrite the above expression as an affine function of $\bar{X}_{t,h}$

$$xr_{t,t+m}^{(n)} \equiv \Xi_0 + \Xi_1 \bar{X}_{t,h} \tag{48}$$

$$\Xi_0 = nA(n) - mA(m) - (n-m)A(n-m) - (n-m)B(n-m)'\left[I - e^{-Km}\right]\theta \tag{49}$$

$$\Xi_1 = nB(n)' - mB(m)' - (n-m)B(n-m)'e^{-Km} \tag{50}$$

The model-implied instantaneous expected excess return is the volatility multiplied by the price of risk. It follows from an application of Ito's lemma that $-nB(n)'\Sigma_h$ is the volatility matrix on the instantaneous return from holding an n-period bond. Given the market price of risk $\lambda_t = \lambda_0 + \lambda_1 \bar{X}_{t,h} = \Sigma_h^{-1}[K\theta - K^{\mathbb{Q}}\theta^{\mathbb{Q}}] + \Sigma_h^{-1}[K^{\mathbb{Q}} - K]\bar{X}_{t,h}$ we obtain

$$ixr_t^{(n)} \equiv -nB(n)'\Sigma_h\lambda_t = -nB(n)'\left[\left(\mathcal{K}\theta - \mathcal{K}^{\mathbb{Q}}\theta^{\mathbb{Q}}\right) + \left(\mathcal{K}^{\mathbb{Q}} - \mathcal{K}\right)\bar{X}_{t,h}\right]. \tag{51}$$

Conditioning on more information, that is on Y_t and $\bar{X}_{t,h}$, we can define the model-implied, $\mathcal{F}_t^{Y,\bar{X}}-$conditional expected excess log return from holding an n-period bond over m-months as

$$\begin{aligned} xr_{t,t+m}^{(n)} &= -(n-m)E^{\mathbb{P}}\left[Y_{t+m}^{(n-m)}\right] + nY_t^{(n)} - mY_t^{(m)} && (52)\\ &= -(n-m)\left[I - e^{-Km}\right]\theta - (n-m)e^{-Km}\bar{X}_{t,h} + nY_t^{(n)} - mY_t^{(m)} && (53)\end{aligned}$$

Finally, the excess return predicted using the random walk forecast of the $(n-m)$-period yield is denoted RW and given by

$$rwxr_{t,t+m}^{(n)} = -(n-m)Y_t^{(n-m)} + nY_t^{(n)} - mY_t^{(m)} \tag{54}$$

References

Ajello, A., L. Benzoni, and O. Chyruk (2014): "Core and 'Crust': Consumer Prices and the Term Structure of Interst Rates," *Federal Reserve Bank of Chicago Working paper.*

Ang, A., S. Dong, and M. Piazzesi (2007): "No-Arbitrage Taylor Rules," *NBER Working Paper No. 13448*.

Ang, A., and M. Piazzesi (2004): "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 50(4), 745–787.

Chernov, M., and P. Mueller (2012): "The Term Structure of Inflation Expectations," *Journal of Financial Economics*, 106, 367 – 394.

Christiano, L. J., M. Eichenbaum, and C. Evans (1999): "Monetary Policy Shocks: What Have We Learned and to What End?," in *Handbook of Macroeconomics*, ed. by J. Taylor, and M. Woodford, pp. 65–148. Elsevier Science, North Holland.

Chun, A. L. (2011): "Expectations, Bond Yields and Monetary Policy," *Review of Financial Studies*, 24(2), 208–247.

——— (2016): "Forecasting Interest Rates and Inflation: Blue Chip Clairvoyants or Econometrics?," *Working paper: http://ssrn.com/abstract=946667*.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115(1), 147–180.

Cochrane, J. H., and M. Piazzesi (2005): "Bond Risk Premia," *American Economic Review*, 95(1), 138–160.

Cooper, I., and R. Priestly (2009): "Time-Varying Risk Premiums and the Output Gap," *The Review of Financial Studies*, 22(7), 2801–2833.

Dai, Q., and K. Singleton (2000): "Specification Analysis of Affine Term Structure-Models," *Journal of Finance*, 55, 1943–1978.

Dick, C., M. Schmeling, and A. Schrimpf (2013): "Macro Expectations, Aggreate Uncertainty and Expected Term Premia," *European Economic Review*, 58, 58–80.

Duan, J., and J. Simonato (1999): "Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter," *Review of Quantitative Finance and Accounting*, 13, 111–135.

Duffee, G. (2011): "Information in (and not in) the Term Structure," *Review of Financial Studies*, 9(24), 2895–2934.

Duffie, D., and R. Kan (1996): "A Yield Factor Model of Interest Rates," *Mathematical Finance*, 6(4), 379–406.

Fackler, P. L. (2000): "Moments of Affine Diffusions," *Working Paper, North Carolina State University*.

Feunou, B., and J.-S. Fontaine (2014): "Non-Markov Gaussian Term Structure Models: The case of Inflation," *Review of Finance*, 18(5), 1953–2001.

Fisher, M., and C. Gilles (1996): "Term Premia in Exponential-Affine Models of the Term Structure," Working Paper, Board of Governers of the Federal Reserve System.

Goffe, W. L., G. D. Ferrier, and J. Rogers (1994): "Global Optimization of Statistical Functions with Simulated Annealing," *Journal of Ecnometrics*, 60, 65 – 99.

Grishchenko, O. V., and J. Huang (2013): "Inflation Risk Premium: Evidence from the TIPS Market," *Journal of Fixed Income*, 22, 5–30.

Hamilton, J. (1994): *Time Series Analysis*. Princeton University Press.

Joslin, S., K. J. Singleton, and M. Priebsch (2014): "Risk Premiums in Dynamic Term Structure Models and Unspanned Macro Risks," *Journal of Finance*, 69(3), 1197–1233.

Kim, D. H., and A. Orphanides (2012): "Term Structure Estimation with Survey Data on Interest Rate Forecasts," *Journal of Financial and Quantitative Analysis*, 47(1), 241 – 272.

Langetieg, T. C. (1980): "A Multivariate Model of the Term Structure," *Journal of Finance*, 25(1), 71–97.

Le, A., and K. J. Singleton (2013): "The Structure of Risks in Equilibrium Affine Term Structure of Bond Yields," *Working Paper*.

Ludvigson, S. C., and S. Ng (2009): "Macro Factors in Bond Risk Premia," *Review of Financial Studies*, 22(12), 5027–5067.

Orphanides, A., and M. Wei (2012): "Evolving Macroeconomic Perceptions and the Term Structure of Interest Rates," *Journal of Economic Dynamics and Control*, 36(2), 239–254.

Pennacchi, G. (1991): "Identifying the Dynamics of Real Interst Rates and Inflation: Evidence Using Survey Data.," *Review of Financial Studies*, 4, 53–86.

Piazzesi, M. (2005): "Bond Yields and the Federal Reserve," *Journal of Political Economy*, 113(2), 311–344.

Romer, D., and C. Romer (2000): "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 90(3), 429–457.

Taylor, J. B. (1993): "Discretion verses Policy Rules in Practice," *Carnegie-Rohester Conference Series on Public Policy*, 39, 195–214.

E Supplementary Appendix: Tables and Figures

Table S1: Maximium Likelihood Estimates of State Parameters

	Forecast Horizon: Quarters Ahead			
	1	2	3	4
$\mathcal{K}_{11}$	0.09597	0.09576	0.05146	0.04454
	(0.01224)	(0.01265)	(0.00995)	(0.00756)
$\mathcal{K}_{21}$	-0.00913	-0.02282	-0.01097	-0.00173
	(0.00572)	(0.00482)	(0.00619)	(0.00653)
$\mathcal{K}_{31}$	-0.05048	-0.04794	-0.03950	-0.02681
	(0.01468)	(0.01344)	(0.01583)	(0.01965)
$\mathcal{K}_{12}$	-0.01697	-0.00363	0.00883	0.01662
	(0.01395)	(0.00974)	(0.00623)	(0.00444)
$\mathcal{K}_{22}$	0.05161	0.00549	-0.00587	-0.00825
	(0.00587)	(0.00368)	(0.00281)	(0.00328)
$\mathcal{K}_{32}$	-0.02294	-0.03291	-0.03845	-0.03737
	(0.01480)	(0.01148)	(0.01242)	(0.01150)
$\mathcal{K}_{13}$	0.03241	0.02212	0.00970	0.00392
	(0.00867)	(0.00534)	(0.00353)	(0.00317)
$\mathcal{K}_{23}$	-0.00922	0.00010	0.00445	0.00686
	(0.00648)	(0.00250)	(0.00214)	(0.00215)
$\mathcal{K}_{33}$	0.00630	0.01267	0.01909	0.02377
	(0.01157)	(0.00887)	(0.00706)	(0.00772)
Σ_{11}	0.00322	0.00267	0.00179	0.00147
	(0.00008)	(0.00014)	(0.00015)	(0.00012)
Σ_{21}	0.00017	-0.00022	-0.00011	0.00000
	(0.00007)	(0.00005)	(0.00006)	(0.00009)
Σ_{31}	0.00059	0.00025	-0.00004	-0.00024
	(0.00016)	(0.00014)	(0.00023)	(0.00030)
Σ_{22}	0.00186	0.00104	0.00073	0.00070
	(0.00004)	(0.00004)	(0.00004)	(0.00004)
Σ_{32}	0.00039	0.00058	0.00090	0.00099
	(0.00015)	(0.00011)	(0.00014)	(0.00015)
Σ_{33}	0.00162	0.00176	0.00183	0.00192
	(0.00005)	(0.00007)	(0.00009)	(0.00014)
θ_1	0.00207	0.00213	0.00208	-0.00184
	(0.00014)	(0.00010)	(0.00029)	(0.03116)
θ_2	0.00252	0.00291	0.00311	0.01146
	(0.00024)	(0.00031)	(0.00076)	(0.06702)
θ_3	0.00464	0.00527	0.00533	0.01388
	(0.00042)	(0.00059)	(0.00099)	(0.06180)

This table resports the maximum likelihood estimates of the state parameters for the single horizon estimation with corresponding asymptotic stardard errors in parenthesis. All Σ estimates and their standard errors are multiplied by 10,000.

Table S2: Maximum Likelihood Estimates of Term Structure Parameters

	Forecast Horizon in Months - h^*											
	1	2	3	4	5	6	7	8	9	10	11	12
ρ_0	-0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\rho_{1,1}$	-0.008	-0.045	-0.103	-0.139	-0.201	-0.270	-0.324	-0.386	-0.445	-0.539	-0.599	-0.648
	(0.015)	(0.016)	(0.018)	(0.025)	(0.307)	(0.213)	(0.042)	(0.043)	(0.046)	(0.061)	(0.072)	(0.076)
$\rho_{1,2}$	-0.039	-0.061	-0.086	-0.165	-0.200	-0.235	-0.283	-0.343	-0.381	-0.436	-0.486	-0.539
	(0.018)	(0.033)	(0.035)	(0.023)	(0.459)	(0.199)	(0.032)	(0.033)	(0.033)	(0.040)	(0.041)	(0.042)
$\rho_{1,3}$	0.975	0.983	0.983	1.046	1.056	1.064	1.108	1.133	1.147	1.198	1.221	1.245
	(0.015)	(0.014)	(0.014)	(0.017)	(0.078)	(0.054)	(0.019)	(0.019)	(0.019)	(0.023)	(0.023)	(0.023)
$\lambda_{0,1}$	0.6	0.7	0.9	0.3	0.2	0.1	1.5	1.6	1.5	1.7	1.3	1.4
	(0.3)	(0.3)	(0.4)	(0.3)	(0.5)	(8.5)	(0.5)	(0.5)	(0.5)	(0.7)	(0.6)	(0.7)
$\lambda_{0,2}$	-0.6	-0.5	-0.6	0.3	0.3	0.4	0.5	0.5	0.4	0.3	-0.0	0.0
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.8)	(0.2)	(0.2)	(0.2)	(0.2)	(0.1)	(0.1)
$\lambda_{0,3}$	0.2	0.4	0.3	0.2	0.2	0.3	0.2	0.2	0.2	0.4	0.4	0.4
	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)	(4.4)	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)
$\lambda_{1,11}$	33.9	83.4	45.8	249.6	281.0	315.4	162.6	208.0	218.6	271.1	246.3	234.2
	(34.2)	(37.4)	(44.1)	(48.3)	(130.6)	(3.4)	(93.1)	(92.3)	(106.7)	(129.6)	(129.7)	(133.3)
$\lambda_{1,21}$	-25.3	-45.4	-40.9	-199.5	-220.7	-243.9	-182.9	-189.9	-191.7	-85.4	-76.0	-78.2
	(11.6)	(10.5)	(12.4)	(11.3)	(50.4)	(21.5)	(11.9)	(10.5)	(10.7)	(14.3)	(14.7)	(15.4)
$\lambda_{1,31}$	-140.3	-221.3	-208.8	-156.8	-174.3	-193.5	-79.3	-92.7	-89.5	-34.5	-12.3	-15.8
	(35.9)	(37.4)	(41.6)	(41.6)	(3.6)	(5.8)	(27.4)	(24.7)	(29.3)	(21.7)	(29.1)	(29.8)
$\lambda_{1,12}$	-247.9	-236.0	-244.8	-388.5	-393.6	-399.0	-626.2	-746.0	-672.2	-634.5	-536.7	-566.0
	(119.5)	(82.0)	(80.5)	(97.8)	(1628.6)	(13.6)	(165.5)	(176.9)	(173.5)	(204.1)	(200.0)	(209.7)
$\lambda_{1,22}$	326.0	332.5	352.0	95.5	97.6	99.9	-68.1	-78.8	-41.6	-112.0	-45.5	-46.7
	(19.1)	(15.4)	(15.9)	(27.5)	(589.3)	(31.0)	(49.4)	(48.9)	(49.4)	(54.1)	(41.9)	(44.3)
$\lambda_{1,32}$	202.3	187.3	200.8	294.3	300.0	306.0	185.6	212.4	191.5	70.1	61.1	68.6
	(85.9)	(71.3)	(70.3)	(54.2)	(570.2)	(165.3)	(50.7)	(51.5)	(50.1)	(52.1)	(49.9)	(50.6)
$\lambda_{1,13}$	-45.6	-81.8	-113.4	56.4	62.9	70.3	-33.6	6.4	-6.6	-81.2	-63.6	-61.7
	(50.5)	(55.8)	(67.2)	(37.4)	(1629.1)	(39.0)	(65.7)	(68.0)	(66.1)	(79.4)	(67.9)	(69.3)
$\lambda_{1,23}$	-32.9	-35.5	-34.6	-20.5	-25.5	-30.9	28.9	24.2	25.4	59.4	75.0	73.3
	(10.5)	(8.3)	(8.5)	(6.4)	(364.1)	(40.7)	(6.8)	(5.4)	(6.7)	(8.3)	(10.7)	(10.9)
$\lambda_{1,33}$	-21.1	-21.0	-1.9	-104.5	-109.5	-114.9	-60.2	-77.0	-76.7	-72.1	-73.0	-73.7
	(48.7)	(36.5)	(40.0)	(27.9)	(1321.8)	(0.7)	(18.8)	(17.9)	(17.1)	(11.5)	(12.2)	(12.0)

This table resports maximum likelihood estimates of the yield curve parameters for Models SS_{h^*}, with corresponding asymptotic stardard errors in parenthesis.

Table S3: Kalman Filter Estimates of State Parameters

	Forecast horizon in months - h^*			
	2	5	8	11
$\mathcal{K}_{11}$	0.1301	0.1024	0.0663	0.0323
	(0.03023)	(0.02467)	(0.02114)	(0.01615)
$\mathcal{K}_{21}$	-0.0053	-0.0035	-0.0050	-0.0095
	(0.01732)	(0.00884)	(0.00826)	(0.01153)
$\mathcal{K}_{31}$	-0.0337	-0.0422	-0.0455	-0.0496
	(0.02144)	(0.02246)	(0.03582)	(0.03020)
$\mathcal{K}_{12}$	-0.0089	0.0002	0.0008	0.0143
	(0.07091)	(0.03067)	(0.01659)	(0.01259)
$\mathcal{K}_{22}$	0.0066	-0.0095	-0.0107	-0.0097
	(0.01619)	(0.00997)	(0.00768)	(0.00821)
$\mathcal{K}_{32}$	0.0173	-0.0407	-0.0422	-0.0429
	(0.03475)	(0.02343)	(0.02918)	(0.03122)
$\mathcal{K}_{13}$	0.0229	0.0215	0.0153	0.0066
	(0.02665)	(0.01239)	(0.00750)	(0.00648)
$\mathcal{K}_{23}$	0.0026	0.0082	0.0085	0.0078
	(0.00785)	(0.00391)	(0.00296)	(0.00377)
$\mathcal{K}_{33}$	0.0027	0.0244	0.0251	0.0262
	(0.01485)	(0.00983)	(0.01103)	(0.01188)
Σ_{11}	0.0037	0.0024	0.0017	0.0012
	(0.00023)	(0.00027)	(0.00024)	(0.00020)
Σ_{21}	0.0005	0.0002	0.0000	-0.0000
	(0.00017)	(0.00011)	(0.00008)	(0.00013)
Σ_{31}	0.0010	0.0005	0.0001	-0.0002
	(0.00019)	(0.00026)	(0.00043)	(0.00047)
Σ_{22}	0.0016	0.0008	0.0007	0.0006
	(0.00006)	(0.00003)	(0.00003)	(0.00005)
Σ_{32}	0.0010	0.0009	0.0012	0.0014
	(0.00016)	(0.00011)	(0.00016)	(0.00024)
Σ_{33}	0.0019	-0.0017	0.0018	0.0018
	(0.00010)	(0.00007)	(0.00012)	(0.00021)
θ_1	0.0093	0.0017	0.0017	-0.0010
	(0.05722)	(0.00234)	(0.00528)	(0.09049)
θ_2	0.0223	0.0046	0.0045	0.0087
	(0.16385)	(0.00887)	(0.01911)	(0.16112)
θ_3	-0.0273	0.0075	0.0074	0.0089
	(0.30351)	(0.01044)	(0.02110)	(0.09196)

This table resports multiple horizon Kalman filter maximum likelihood estimates of the state parameters for Models SM_2, SM_5, SM_8 and SM_{11}, with corresponding asymptotic stardard errors in parenthesis. All Σ estimates and their standard errors are multiplied by 10,000.

Table S4: Maximum Likelihood Estimates of Term Structure Parameters

	Forecast Horizon in Months - h^*			
	2	5	8	11
ρ_0	-0.000	0.000	0.001	0.001
	(0.000)	(0.000)	(0.007)	(0.000)
$\rho_{1,1}$	0.002	-0.169	-0.394	-0.563
	(0.018)	(0.033)	(1.700)	(0.056)
$\rho_{1,2}$	-0.023	-0.115	-0.192	-0.404
	(0.042)	(0.046)	(7.346)	(0.057)
$\rho_{1,3}$	0.947	1.001	1.053	1.163
	(0.021)	(0.019)	(3.841)	(0.020)
$\lambda_{0,1}$	1.796	0.726	-1.233	1.132
	(1.559)	(0.661)	(12.168)	(0.786)
$\lambda_{0,2}$	-0.984	-0.763	0.336	0.185
	(0.676)	(0.273)	(2.597)	(0.177)
$\lambda_{0,3}$	-0.624	-0.602	0.474	0.700
	(0.654)	(0.138)	(0.851)	(0.158)
$\lambda_{1,11}$	21.599	343.104	394.789	579.489
	(78.600)	(71.443)	(16590.776)	(149.046)
$\lambda_{1,21}$	6.935	-96.293	-168.249	-268.705
	(33.840)	(43.948)	(6027.999)	(5.457)
$\lambda_{1,31}$	-153.154	291.360	-174.570	-64.182
	(37.374)	(32.134)	(4620.674)	(22.667)
$\lambda_{1,12}$	3212.933	-256.778	158.522	-976.962
	(685.178)	(149.524)	(963.363)	(474.642)
$\lambda_{1,22}$	-1332.945	64.472	-111.886	71.502
	(286.229)	(72.743)	(5184.575)	(65.919)
$\lambda_{1,32}$	-1036.196	-134.921	78.987	-92.179
	(348.567)	(34.868)	(4875.455)	(61.892)
$\lambda_{1,13}$	-2933.723	-332.569	-96.071	-141.738
	(271.075)	(175.222)	(1872.504)	(152.233)
$\lambda_{1,23}$	1267.032	239.406	77.869	65.203
	(110.131)	(62.784)	(4150.381)	(25.173)
$\lambda_{1,33}$	1128.040	-2.351	-7.913	-45.448
	(176.585)	(34.876)	(23.837)	(17.530)

This table reports the maximum likelihood estimates of the yield curve paramters for Models SM_2, SM_5, SM_8 and SM_{11}, with corresponding asymptotic stardard errors in parenthesis.

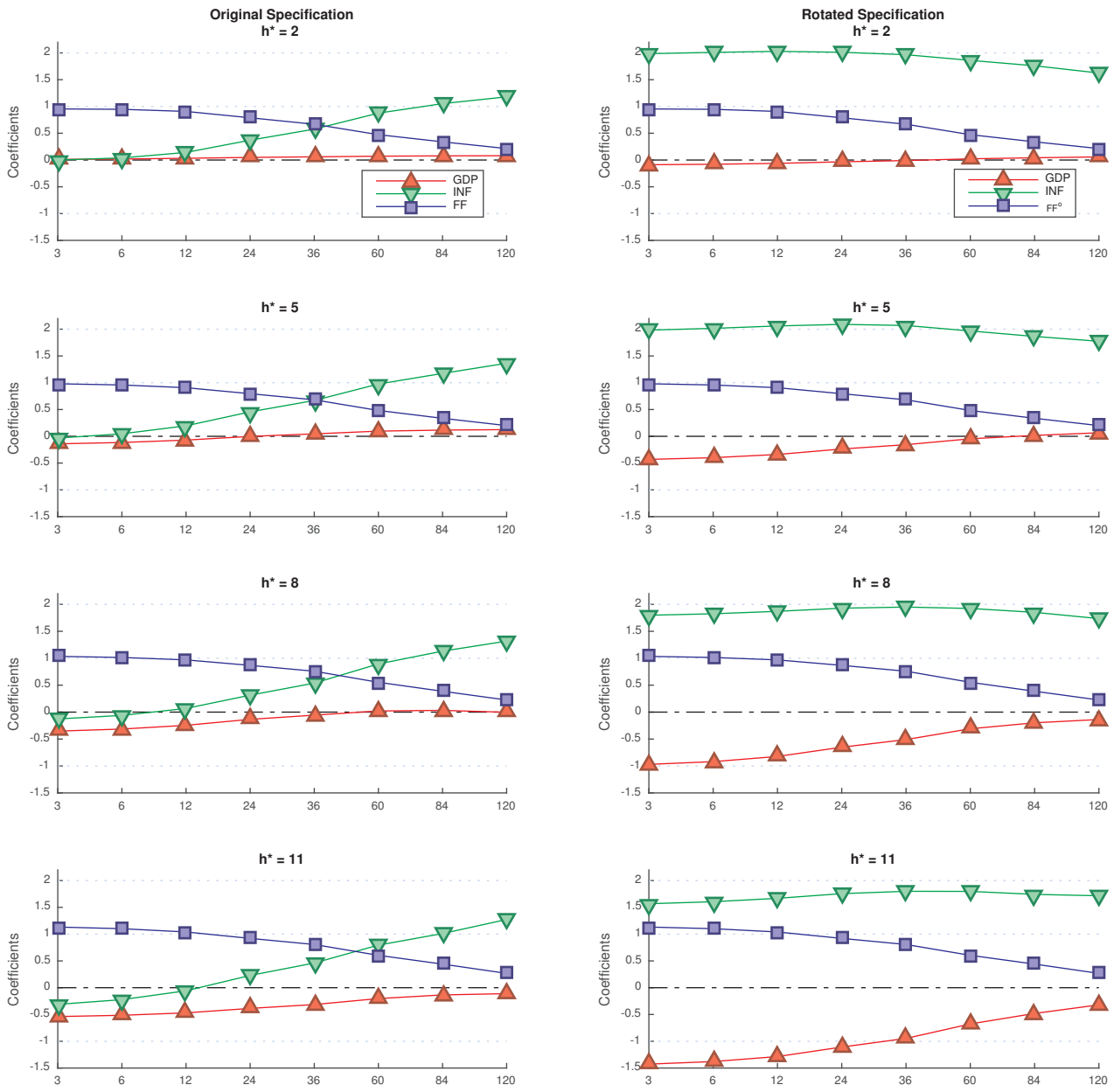

Figure S1: **Factor Loadings - Model SM** The figures plot the factor loadings, that is the coefficient on the yield equations, for Model SM_{h^*} for $h^* = 2$, 5, 8 and 11 months. The left panel plot the loading for the original specification and the right panel the corresponding loadings on the orthogonal projection rotation specification which give the two macro forecasts priority in explaining the yield curve.

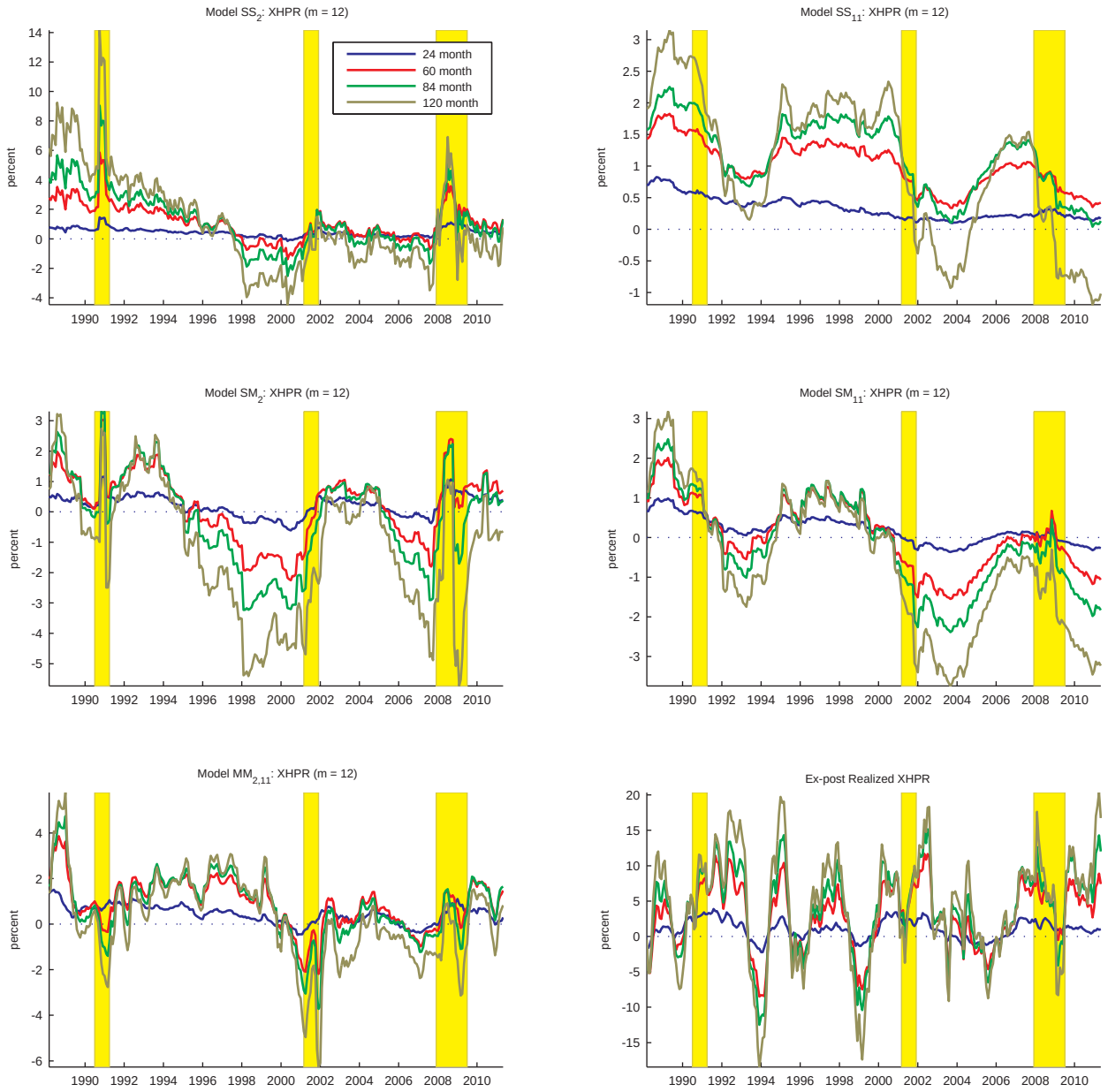

Figure S2: **Model-Implied Risk Premiums** The figures plots model-implied $\mathcal{F}_t^{\bar{X}}-$conditional expected excess returns for the various models over $m = 12$ months. Each plot depicts expected excess returns for yields of four different maturities - 24, 60, 84 and 120 months. The realized ex-post excess returns are plotted in the bottom right figure.