Monetary policy surprises and their transmission through term premia and expected interest rates

Iryna Kaminska, Haroon Mumtaz and Roman Sustek

ISSN 1473-0278

Working Paper No. 917 November 2020

School of Economics and Finance

Queen Mary University of London

Monetary policy surprises and their transmission through term premia and expected interest rates*

Iryna Kaminska†, Haroon Mumtaz‡§ and Roman Šustek¶

November 13, 2020

Abstract

Monetary policy moves the yield curve. How much is due to expected interest rates vs. term premia? And does it matter for macroeconomic outcomes? Using an affine term structure model, we shed new light on these questions. Estimation is subject to restrictions addressing an estimation bias in expected interest rates obtained by previous studies. High-frequency yield curve decomposition around FOMC announcements into term premia and expected interest rates then provides instruments for a local projection model. The effects of interest rate expectations and term premia are found equally important for the transmission mechanism and broadly consistent with macroeconomic theory.

JEL Classification Codes: E43, E52, E58, G12, C58.

Keywords: High-frequency data, monetary policy transmission mechanism, restricted affine term structure models, yield curve decomposition, local projection method, Bayesian estimation.

*We thank Refet Gürkaynak and Silvia Miranda-Agrippino for valuable comments and suggestions. The views expressed are those of the authors and not necessarily of the Bank of England or its Monetary Policy Committee.

†Bank of England; Iryna.Kaminska@bankofengland.co.uk.

‡Queen Mary University of London; h.mumtaz@qmul.ac.uk.

§Corresponding author: Haroon Mumtaz, address: School of Economics and Finance, Queen Mary University of London, Mile End Road, London, E1 4NS, UK, tel: +44 20 7882 8839, e-mail: h.mumtaz@qmul.ac.uk.

¶Queen Mary University of London and the Centre for Macroeconomics; r.sustek@qmul.ac.uk.

1 Introduction

A classic question in macroeconomics concerns the transmission of monetary policy surprises into the economy. The interest in this question stems from the notion that empirical impulse-responses can guide the development of theory (eg, Christiano, Eichenbaum, and Evans, 1999). This research strategy, however, rests on the assumption that one can identify the relevant impulses (shocks) in the data. The traditional approach to this identification problem relies on monthly or quarterly vector auto-regressions (VAR) combining macroeconomic data with a short-term nominal interest rate, taken as a proxy for a policy instrument. Various identification schemes have been proposed within this approach (see Ramey, 2016, for a review). What they have in common, however, is that the identified shocks at best reflect monetary policy surprises relative to the mathematical expectations of the regression model.[1] Furthermore, VAR-based identification is limiting once financial data are included. How does one invert the VAR residuals to identify monetary policy shocks when, at monthly or quarterly frequency, financial markets react to monetary policy and policy makers partially base their decisions on information contained in asset prices? At the same time, ignoring financial data is inefficient, as asset prices reflect expectations of future monetary policy and some sectors, for instance the housing market, are sensitive to asset prices (long-term interest rates).[2]

High-frequency (HF) data can ameliorate the identification problem (see Kuttner, 2001; Cochrane and Piazzesi, 2002; Gürkaynak, Sack, and Swanson, 2005b, for early contributions). The idea is that the announcement of the outcome of a policy meeting is the only event impacting on asset prices in a tight enough window around the news release and that, in that window, the content of the news release is exogenous to asset prices. The HF movements in asset prices thus provide instruments that can identify policy shocks.[3] Once the shock

[1] An alternative identification strategy, proposed by Romer and Romer (2004), is based on central bank narrative.

[2] Evans and Marshall (1998) is an example of an attempt to include long-term interest rates in a macro VAR model with monetary policy shocks identified in one of the traditional ways.

[3] There is a debate about the interpretation of the shocks identified this way and we briefly discuss this debate below.

is identified, it can be used in an empirical model to trace out its effect on macro variables in a standard way. Gertler and Karadi (2015) carry out such an exercise and arrive at a stark conclusion: monetary policy surprises transmit into the economy almost exclusively through changes in term premia, with movements in expected future interest rates playing a minuscule role.[4] This finding presents a challenge to standard macro models used for monetary policy analysis. In standard models, monetary policy transmits through changes in the conditional mean of the nominal pricing kernel, not its variance, the relevant part for movements in term premia (eg, Atkeson and Kehoe, 2009).

In this paper, we first revisit the question to what extent monetary policy surprises affect the yield curve through term premia vs. expected future nominal interest rates. We then proceed to investigate the implications of this decomposition for the macro-economy, including housing market variables. Specifically, in the first step, we employ an estimated affine term structure model (ATSM) to decompose the HF movements in yields around Federal Open Market Committee (FOMC) announcements into expected interest rates and term premia.[5] The key feature of the estimated ATSM is that it is subject to recently proposed restrictions leading to more precise estimates of expected interest rates, and thus term premia, than either VAR models or unrestricted ATSMs would suggest. In the second step, we take the HF changes in the two yield curve components to derive instruments that we use in local projections (Jordà, 2005) to trace out the dynamic effects of policy shocks on macro variables. The instruments generalise proxies for monetary policy shocks employed in recent studies and have intuitive economic interpretation.

Why use an ATSM for the decomposition and why is it important to estimate the term structure model subject to restrictions? As the term premium is often estimated as a difference between the observed yield of a given maturity and the expected future path of the

[4]Term premia reflect risk compensation for holding a long-term bond and can be estimated as a difference between the observed long-term interest rate of a given maturity and a forecast of the path of the short rate over that time horizon (ignoring technical details such as measurement errors and Jensen's inequality).

[5]ATSMs are the go-to models in empirical finance to study the term structure of interest rates. See Diebold, Piazzesi, and Rudebusch (2005), Piazzesi (2006), Duffee (2012), or Gürkaynak and Wright (2012) for an introduction.

short rate, one could in principle estimate a VAR on yields (and possibly macro variables) and then iterate it forward J times to obtain forecasts of the short rate between now and the Jth period ahead, thus obtaining the term premium for the Jth maturity. There are two problems with this approach, which is the approach taken by Gertler and Karadi (2015). First, the VAR-based forecasts of future yields of different maturities may imply arbitrage opportunities, which are unlikely to be present in such a deep market, like the government bond market. Second, nominal interest rates are highly persistent, which, in samples of the length typically observed, leads to both a downward bias in the persistence of the VAR process and high standard errors of its estimates. This problem arises because we do not observe frequent enough mean reversions of interest rates in the data to estimate the parameters of the driving process well.[6] The implication of the downward bias is that VAR-based forecasts of future interest rates at the long end of the yield curve are insensitive to current shocks and thus any observed movements in long yields are prescribed to term premia.

By construction, ATSMs resolve the first issue by requiring that the evolution of yields, determined by a VAR for a small number of risk factors, respects no-arbitrage conditions. ATSMs can also resolve the second issue, but only if they are estimated subject to restrictions. As ATSMs are estimated on both time series and cross-sectional (across maturities) data, they use more information than a VAR. In particular, the cross section of yields at a point in time can *potentially* provide very precise information for the model dynamics.[7] However, a well-documented time-variation in risk premia (Fama and Bliss, 1987, and a long list of studies that followed), effectively eliminates this connection between the time-series and cross-sectional dimensions of ATSMs. Loosely speaking, there always exist some (unobserved) prices of risk that make the time series data consistent with the cross-sectional data. In terms of the ATSM terminology, they make the physical '$\mathbb{P}$-dynamics' of the state

[6] See the classic results of Kendall (1954), Nicholls and Pope (1988), and Shaman and Stine (1988) and, for a discussion in the context of ATSMs, Bauer, Rudebusch, and Wu (2012). As demonstrated by Pierse and Snell (1995), increasing the sampling frequency does not resolve the problem.

[7] To illustrate this, suppose investors were risk neutral (ie, prices of risk were equal to zero) and so observed yields were equal to expected future interest rates. Then one could simply read off expected future interest rates from the cross-section, thus avoiding the problematic time series data altogether.

space consistent with the risk-adjusted '$\mathbb{Q}$-dynamics' implied by the cross-section. In fact, as demonstrated by Joslin, Singleton, and Zhu (2011), in a canonical ATSM—the maximally flexible model that is subject only to normalizing restrictions required for identification—the cross-sectional data convey no information for the time-series properties of the model (see also Hamilton and Wu, 2012). As a result, the parameter estimates of the underlying state space process are equivalent to those obtained from a simple VAR estimated on time series of yields, with the problems noted above.

Following the common practice of using as risk factors the principal components (PCs) of interest rates, we estimate an ATSM subject to two types of restrictions, broadly defined. First, we employ a statistical search procedure to set some parameters of the affine mapping from the factors to risk prices equal to zero. By restricting how individual risk prices respond to the factors, this procedure constrains the wedge between the $\mathbb{P}$- and $\mathbb{Q}$-dynamics, thus allowing to exploit the cross-sectional information for the estimation of the VAR parameters of the state space. This strategy has been employed, in various forms, by Cochrane and Piazzesi (2008), Duffee (2011), Joslin et al. (2011), Joslin, Priebsch, and Singleton (2014), and Bauer (2018). For our main exercise, we use the Bayesian procedure proposed by Bauer (2018). Second, we estimate a bias-corrected ATSM as in Bauer et al. (2012). Specifically, the model is estimated subject to the restriction that, if taken as the data-generating process, it produces the same small-sample bias as in the data. Both restrictions lead to the same result, namely higher persistence of the factors, relative to estimates from a simple VAR, and thus more volatile expected future interest rates. To ensure robustness, the estimation is carried out on monthly as well as daily data. Monthly data conform in terms of frequency to macro variables, and as such are commonly used in the literature, while daily data are closer to the HF decomposition.

The estimated ATSMs are then used to decompose movements in the yield curve in a tight window around FOMC meetings into changes in term premia and expected future interest rates. The HF changes in the two components of the yield curve serve as instruments to

identify monetary policy shocks in a local projection macro model. We are interested in how the different shocks affect the macro-economy. As the shocks are identified from unexpected movements of the entire yield curve, not just the short rate, they are conceptually different from the traditional shocks identified in structural VARs. For instance, a shock can occur even if the Fed maintains the fed funds rate unchanged.

Our instruments generalise the proxies for monetary policy shocks employed in recent studies. Gertler and Karadi (2015) identify monetary policy shocks by using the change in a single interest rate around FOMC meetings. In contrast, we exploit information in the entire term structure. Our approach is thus closer to the work of Gürkaynak, Sack, and Swanson (2005a) who show that FOMC announcements are adequately described by two factors obtained from the cross-section of high frequency yields. However, we go beyond their approach and use ATSMs to decompose the HF changes in the yield curve into term premia and interest rate expectations and study their dynamic effects on selected macro variables. Such decomposition provides a sharper characterization of the transmission mechanism of monetary policy than instruments that bundle the two components together.

Our main findings, based on the period 1996-2007, characterized by conventional monetary policy, can be summarized by two points. First, the reaction of expected future interest rates in response to FOMC announcements in this period was significant. Specifically, in all ATSM specifications that are estimated subject to restrictions, expected interest rates at the ten-year horizon are twice as volatile as what a VAR analysis would suggest. In the monthly models, in particular, the HF reaction of expected interest rates is as large as the reaction of term premia. Monetary policy surprises contained in FOMC announcements thus had an equally important effect on expected interest rates as on term premia.

Second, a PC decomposition of the HF change in expected interest rates and term premia identifies three types of policy shocks. Two of them account for most of the variation in the HF changes in the yield curve, with the third one playing quantitatively a less important role. One accounts for a bulk of the HF movements in expectations (the first PC of expectations)

and the other in term premia (the first PC of term premia). The first PC of expectations identifies what appears to be a persistent shock associated with an increase in expected future nominal interest rates, inflation, and the 30-year mortgage rate, and a decline in housing market activity and aggregate output. The propagation of this shock into the economy fits the narrative of models with a role for housing finance. The first PC of term premia, on the other hand, identifies a policy shock associated with a temporary increase in corporate credit spreads, and a decline in inflation and output. Its effects seem consistent with financial accelerator models. The third shock has similar aggregate implications as the shock moving term premia, but is identified from a HF change in expected interest rates (the second PC of expectations). It appears to fit a narrative of New-Keynesian models. In sum, the economy's responses to the three instruments seem broadly consistent with existing theories.

HF, intra-day, data have been increasingly used to study yield curve responses to various macroeconomic events and to identify monetary policy shocks. Broadly speaking, the literature can be divided into two mutually non-exclusive categories: the decomposition of the yield curve into its various components (including real and nominal) and the structural interpretation of monetary policy surprises. The first category includes, for instance, Beechey (2007), Beechey and Wright (2009), Bauer (2015), Gertler and Karadi (2015), Hanson and Stein (2015), and Hördahl, Remolona, and Valente (2015). Daily data are used by Abrahams, Adrian, Crump, Moench, and Yu (2016). Some of these studies employ ATSMs, while others use simple regressions. The second category includes, among others, Nakamura and Steinsson (2018), Cieslak and Schrimpf (2019), and Jarocinski and Karadi (2020). These authors investigate the information content of policy surprises, as opposed to genuine policy shocks. Gürkaynak et al. (2005a), in contrast, look at surprises in current policy actions vs. communication about future policy. A few studies (Gertler and Karadi, 2015; Nakamura and Steinsson, 2018; Jarocinski and Karadi, 2020) proceed, like we do, to study the macroeconomic consequences of monetary policy surprises identified in HF data. We focus on the most fundamental decomposition of the nominal yield curve, into term

premia and expected future interest rates, and one specific event, FOMC announcements. In terms of the housing market, a subset of our findings is consistent with those reported by Hamilton (2008), who follows a different methodology.[8]

The paper proceeds as follows. Section 2 discusses the HF data, Section 3 introduces the ATSM framework and the necessary notation, Section 4 describes estimation of the ATSMs, Section 5 reviews the estimates, combines the models with HF data for the purposes of the HF decomposition, and carries out the local projection analysis. Finally, Section 6 concludes. Additional robustness checks and other details are included in an online Appendix.

2 High-frequency movements around FOMCs

In this section we describe the yield curve data used to estimate surprises in FOMC announcements. We also briefly analyse some of their properties, which will then serve as a basis for the advanced term structure analysis presented in the later sections of the paper.

In order to study high-frequency yield curve reactions to monetary policy decisions, we measure yields at various maturities in a narrow window around FOMC announcements. In doing so, we build on the available literature studying monetary policy shocks within the high-frequency approach (see Gertler and Karadi, 2015; Miranda-Agrippino, 2016; Jarocinski and Karadi, 2020, among others). This literature focuses on the short end of the curve and uses the change in the three-month fed funds future (FF4) as a baseline measure of the interest rate surprise. Importantly, we extend these studies and measure monetary policy surprises along the spectrum of maturities, capturing the reaction of the whole yield curve.

[8] A part of the literature, Kim and Orphanides (2012) being an early example, complements yield curve data with surveys of professional economists as a source of data for expected future interest rates. To keep the paper focused on the improvement of the estimation relative to VARs, we confine ourselves only to yield curve data.

2.1 Intra-day data

Our main high-frequency data source is coming from *Refinitiv Tick History*. As in the earlier literature, the surprises are measured in a 30-minute window starting 10 minutes before and ending 20 minutes after the announcement.

We focus on the period from January 1996 to August 2007, characterized by conventional monetary policy. Unfortunately, the data could be scarce, especially in the 1990s, with only a dozen of intra-day observations available in some cases. Therefore, for a few announcement dates our window has to be wider than 30 minutes. Despite this, the estimated changes in rates are very similar to those reported by other studies. For example, our constructed FF4 reactions are similar to the raw FF4-based shocks reported by Miranda-Agrippino (2016), with the correlation between the two being 0.9.

At the beginning of the sample, Treasury bonds with maturities longer than 10 years were traded relatively infrequently. Therefore, our longest maturity is based on the 10-year Treasury yield series. At medium-term maturities, Treasuries were not as frequently traded as LIBOR-based swaps. Hence we faced a trade-off between having the same instrument but captured at different times due to relative illiquidity (which is especially the case for the 1990s), or having all rates captured at the same time but taken from similar assets rather than the same instruments. We chose the latter and estimated the high-frequency changes at 2-, 3-, and 5-year maturities from LIBOR-based swaps, which enabled us to create consistent narrow windows around the announcements. The observed changes across the various maturities around the announcements are shown in Figure 1, which displays a consistent response pattern across all maturities. Although the strongest reactions seem to be in the short rates, the response is evident at all maturities, and longer-term rates are rarely responding by much less than medium-term rates. The next subsection investigates the simple statistic properties of these reactions in more detail.

Finally, to analyse the behaviour of the yield curve around the announcements in a systematic way, we constructed a consistent yield curve across all maturities, adjusting for

observed daily LIBOR spreads. We do this by estimating the spreads between LIBOR swap rates and the corresponding maturity yields observed at the close of business on the pre-announcement dates and then apply them to LIBOR rates around the announcements.

2.2 Basic properties of yield curve reactions

Table 1 presents basic statistics for the responses across maturities. Several observations follow. First, during 1996-2007, monetary policy surprises were slightly negative on average, with the shortest maturities affected the most and the impact declining with maturity. Also, judging by the volatilities, the shortest maturities were the most sensitive to the announcements and the longest maturities were the least sensitive. Second, despite the declining volatility pattern, all maturities display a strong reaction to the announcements, with yield volatility at medium and long maturities being, respectively, more than 80 and 60 percent of the volatility of FF4. Third, the yield curve tends to respond to the announcement in a consistent way, as indicated by the positive correlations between reactions across maturities, although the correlations are declining with maturity.

Interestingly, the responses are highly correlated across medium and long maturities, with all the correlations between them being around 0.9.[9] The decompositions carried out with the ATSMs into term premia and interest rate expectations, and the subsequent decomposition of these yield curve components into their respective PCs will provide a clear interpretation of the observed correlations.

3 The ATSM framework

The aim of this section is to provide a brief overview of the ATSM and introduce concepts and notation used in the rest of the paper. An underlying assumption behind an ATSM is the fundamental principal of finance, applied to default-free zero-coupon bonds of different

[9] According to Hanson and Stein (2015), for instance, this strong co-movement in maturities beyond the policy horizons is due to monetary policy affecting predominantly term premia at longer horizons.

maturities. Specifically,

$$E_t\left[M_{t+1}R_{t+1}^{(j)}\right] = 1, \tag{1}$$

where the expectation operator is with respect to information in period t, the scalar $M_t > 0$ is a kernel that prices all bonds and $R_{t+1}^{(j)}$ is a one-period gross return on a bond of any maturity j. That is, $R_{t+1}^{(j)} = P_{t+1}^{(j-1)}/P_t^{(j)}$, where $P_t^{(j)}$ is the price in period t of a bond of maturity j, which becomes a bond of maturity $j-1$ one period later. Of course, $P_t^{(0)} = 1$, as one dollar today has a value of one dollar.

ATSMs assume a specific functional form for the pricing kernel

$$-\log M_{t+1} = r_t + \frac{1}{2}\lambda_t'\lambda_t + \lambda_t'\varepsilon_{t+1}. \tag{2}$$

The popularity of this functional form lies in its practicality: when combined with the state space described below, it leads to a convenient affine solution for yields satisfying the no-arbitrage condition (1). Here, r_t is the continuously compounded short-term nominal interest rate, λ_t is a $N \times 1$ vector of risk prices for N underlying risk factors, and ε_{t+1} is a $N \times 1$ vector of innovations specified below. The N factors are assumed to follow a first-order Gaussian VAR

$$X_t = \mu + \Phi X_{t-1} + \Sigma\varepsilon_t, \tag{3}$$

with $\varepsilon_t \sim N(0, I_N)$. This VAR process, describing the stochastic evolution of the risk factors, is referred to as the '$\mathbb{P}$-measure' and the implied dynamics as the '$\mathbb{P}$-dynamics'. That is, this is the physical probability measure describing the physical dynamics of the state space.

Both the short rate and the risk prices are assumed to be related to the N factors through affine mappings

$$r_t = \delta_0 + \delta_1'X_t, \tag{4}$$

$$\lambda_t = \Sigma^{-1}(\lambda_0 + \lambda_1 X_t), \tag{5}$$

where δ_0 δ_1, Σ^{-1}, λ_0, and λ_1 are commensurate to the variables. In particular, λ_1 is $N \times N$.

That is, the risk price of a particular factor can be affected by all factors. Observe that under risk neutrality (zero risk prices), the pricing kernel is simply $M_{t+1} = \exp(-r_t)$. That is, future cash flows are discounted with the short rate. Equations (1)-(5) summarize the ATSM.

Starting with $P_t^{(0)} = 1$, the model can be solved recursively for equilibrium bond prices. Given the functional assumptions on the pricing kernel and the state space, the solution is an affine mapping from factors to the logarithm of bond prices (see, eg, Gürkaynak and Wright, 2012). Continuously compounded yields can then be inferred from the bond prices through standard discounting, $P_t^{(j)} = \exp(-jy_t^{(j)})$, which can be inverted to obtain yields as $y_t^{(j)} = (-1/j)\log P_t^{(j)}$. Yields are thus also affine in factors. For $j = 1$, we get the short rate: $y_t^1 = r_t$.

The vector of any J yields, $\hat{Y}_t$, can be written as

$$\hat{Y}_t = A + BX_t, \tag{6}$$

where $\hat{Y}_t$ is a $J \times 1$ vector. Equation (6) describes the model-implied yield curve—the cross-section of yields at a point in time that is consistent with no-arbitrage. In an empirical implementation of the model, model-implied yields can potentially differ from observed yields due to measurement error and the lack of fit.

The arbitrage-free loadings A and B are non-linear, recursive, functions of the model parameters δ_0, δ_1, λ_0, λ_1, μ, Φ, and Σ (see, eg, Gürkaynak and Wright, 2012). It can then be shown that the A and B in equation (6) would be the same if pricing was risk neutral, $M_{t+1} = \exp(-r_t)$, but the law of motion for the risk factors was given by a risk-adjusted VAR

$$X_t = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_{t-1} + \Sigma\varepsilon_t, \tag{7}$$

where

$$\mu^{\mathbb{Q}} = \mu - \lambda_0 \quad \text{and} \quad \Phi^{\mathbb{Q}} = \Phi - \lambda_1. \tag{8}$$

The VAR process (7) is referred to as the '$\mathbb{Q}$-measure', describing the '$\mathbb{Q}$-dynamics'. That is, dynamics under risk neutral pricing. Observe that under risk neutral pricing, the model is parameterised in terms of δ_0, δ_1, $\mu^{\mathbb{Q}}$, $\Phi^{\mathbb{Q}}$, and Σ. Thus, to derive the cross-sectional implications of the model summarized by equation (6), all that is required is the $\mathbb{Q}$-measure. The knowledge of the $\mathbb{P}$-measure and the risk prices λ_t is not required. To put it differently, the cross-section identifies the parameters of the $\mathbb{Q}$-measure, not the $\mathbb{P}$-measure.

Under the $\mathbb{Q}$-measure, the expected value of the short rate j periods ahead can be obtained from the short rate equation (4) and the VAR process (7). The effect of X_t on the expected value is given by $(\Phi^{\mathbb{Q}})^j$. The effect of X_t on the average expected short rate over the forecast horizon under the $\mathbb{Q}$-measure is thus

$$B_j = \frac{1}{j}\delta_1' \left[I + \Phi^{\mathbb{Q}} + \cdots + (\Phi^{\mathbb{Q}})^{j-1} \right], \tag{9}$$

which is the jth row in the loading B in equation (6). Under the $\mathbb{P}$-measure, the expected value of the short rate j periods ahead can be obtained from the short rate equation (4) and the VAR process (3). In this case, the effect of X_t on the expected value is given by Φ^j and the average expected short rate over the forecast horizon is given by

$$B_j^{\mathbb{P}} = \frac{1}{j}\delta_1' \left[I + \Phi + \cdots + \Phi^{j-1} \right]. \tag{10}$$

The difference $(B_j - B_j^{\mathbb{P}})$ accounts for the effect of X_t on the term premium in yield $y_t^{(j)}$ and, as follows from the relationship (8), depends on λ_1. Thus, while the knowledge of the $\mathbb{P}$-measure is not required for the cross-sectional implications of the model, it is necessary to derive a decomposition between term premia and expected interest rates. Observe that the $\mathbb{P}$-measure can be identified either from the time-series of X_t or the cross-section of yields and the knowledge of λ_0 and λ_1 through the relationship (8).

4 Estimation of the ATSM

To ensure identification, normalising restrictions need to be imposed on the no-arbitrage framework introduced in the previous section. We employ the normalising restrictions of Joslin et al. (2011), leading to their canonical representation—the maximally flexible form that is identified. Under this representation, the N risk factors are defined as linear combinations of yields, $X_t = W\hat{Y}_t$, where W is a weighting matrix, and the model parameters are mapped into the set of unknowns $k^{\mathbb{Q}}$, $\phi^{\mathbb{Q}}$, μ, Φ, and Σ, which fully characterize the $\mathbb{P}$- and $\mathbb{Q}$-dynamics, (μ, Φ) and $(\mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}})$ respectively. Here, $k^{\mathbb{Q}}$ determines the mean of the short rate under the $\mathbb{Q}$-measure and $\phi^{\mathbb{Q}}$ is a $N \times 1$ vector that contains the eigenvalues of $\Phi^{\mathbb{Q}}$. Following Joslin et al. (2011), the risk factors X_t are calculated as the first N PCs of the yields and W is the associated $N \times J$ loading matrix. Finally, the observed yields Y_t are assumed to be measured with error: $Y_t = \hat{Y}_t + e_t$. Under the assumption that X_t are observed in the estimation (ie, N linear combinations of yields using the weights W are estimated exactly by the model), the $J - N$ independent measurement errors are normal with variance σ_e^2.

We estimate three versions of the model, with details of the estimation provided in the next sections. Model $\mathcal{M}_0$ is the maximally flexible benchmark that is only subject to the Joslin et al. (2011) normalising restrictions. As discussed in the Introduction, such a specification does not fully exploit the information in the cross-section of yields and estimates of the VAR parameters under the $\mathbb{P}$-measure (μ, Φ) are based on time-series information only. Due to the persistence of interest rates, the estimates of the VAR parameters can be biased in small samples. The maximally flexible model is subject to this problem. By placing no restrictions on risk prices, $\lambda_0 = \mu - \mu^{\mathbb{Q}}$ and $\lambda_1 = \Phi - \Phi^{\mathbb{Q}}$ are simply obtained from the separate estimates of the $\mathbb{P}$- and $\mathbb{Q}$-measure parameters, (μ, Φ) and $(\mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}})$ respectively. Therefore, one approach to deal with this issue involves placing zero restrictions on λ_0 and λ_1.

To impose such restrictions on risk prices, our first alternative specification of the ATSM (model $\mathcal{M}_1$) uses a stochastic search variable selection (SSVS) algorithm employed by Bauer

(2018). It is clear from the definition of λ_0 and λ_1 that setting some risk prices to zero has the effect of 'pulling up' the VAR parameters μ and Φ towards $\mu^{\mathbb{Q}}$ and $\Phi^{\mathbb{Q}}$, thus ameliorating a small sample bias.

Our second alternative specification (model $\mathcal{M}_2$) is based on the analysis of Bauer et al. (2012), who propose statistical methods to estimate and correct the small sample bias in μ and Φ. In this case, the model is estimated subject to the restriction that, assuming it is the data-generating process, it produces the same small sample bias in the persistence of yields as in the data. As a result, this procedure increases the persistence of the VAR under the $\mathbb{P}$-measure, relative to that under $\mathcal{M}_0$.

4.1 MCMC algorithm

As shown by Joslin et al. (2011), an important implication of their canonical representation is that the likelihood of the model factors into components that simplify estimation greatly. The likelihood function is defined as:

$$f\left(Y_t|Y_{t-1},\Theta\right) = f\left(Y_t|X_t,\phi^{\mathbb{Q}},k^{\mathbb{Q}},\Sigma,\sigma_e^2\right) \times f(X_t|X_{t-1},\mu,\Phi,\Sigma), \tag{11}$$

where $\Theta = \left(\phi^{\mathbb{Q}},k^{\mathbb{Q}},\Sigma,\sigma_e^2,\mu,\Phi\right)$ denotes the parameters to be estimated. Note that the first term in this factorisation is the '$\mathbb{Q}$-likelihood', as it incorporates information from the cross-section of yields. In contrast, the second term is the '$\mathbb{P}$-likelihood', based on information derived from the time-series of the risk factors.[10]

We employ a Bayesian approach to estimate the different versions of the model, using the Gibbs sampling algorithm proposed by Bauer (2018). The Bayesian approach is particularly useful as it provides a systematic and efficient method to impose restrictions on μ and Φ (or equivalently on λ_0 and λ_1). This means that there is no need to carry out an explicit model comparison exercise that can involve estimation of a large number of restricted specifications.

[10] As Joslin et al. (2011) show, the fact that the two likelihoods share Σ does not affect the estimates of the other parameters.

Moreover, maximisation of the likelihood of the ATSM is a non-trivial task that is made even more challenging by the small sample of the typical data set. Bayesian estimation does not rely on maximisation of the likelihood function and, instead, aims to approximate the joint posterior distribution of the model parameters. MCMC methods make this task easier by working with the conditional distributions associated with the joint posterior, thus breaking a complex problem into smaller portions. Finally, as the Bayesian approach approximates the posterior distribution, error bands for parameter estimates are obtained directly. In contrast, frequentist approaches rely on asymptotic standard errors that may be inaccurate in small samples.[11]

Details of the MCMC algorithm and the conditional posterior distributions are provided in the Appendix. Here we present a summary of the algorithm for each version of the model. The Gibbs algorithms use 1,000,000 iterations. Every 20th draw after a burn-in period of 500,000 is used for inference. The Appendix presents inefficiency factors that are fairly low, suggesting convergence.

4.1.1 Model $\mathcal{M}_0$

As described in the Appendix, we employ un-informative priors for the parameters. The Gibbs algorithm samples from the following conditional posterior distributions:

1. $g(\mu, \Phi|\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma, \sigma_e^2)$. Given the factorisation of the likelihood in equation (11), this conditional posterior is standard. In other words, conditional on $\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma$ and σ_e^2, the model collapses to a standard Bayesian VAR with flat priors where the posterior distribution of the $\mathbb{P}$-coefficients is known to be normal with mean and variance given by OLS estimates.

2. $g(\phi^{\mathbb{Q}}, k^{\mathbb{Q}}|\mu, \Phi, \Sigma, \sigma_e^2)$. The conditional posterior of the $\mathbb{Q}$-parameters is non-standard. Therefore, a random walk Metropolis-Hastings step is used to sample these parameters.

[11] Note that bootstrap methods require repeated re-estimation of the model by maximising the likelihood. This approach is, therefore, subject to problems associated with the task of finding the mode of the likelihood.

Note that the acceptance probability is easily computed using the $\mathbb{Q}$-likelihood in the factorisation shown in equation (11).

3. $g(\Sigma|\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \mu, \Phi, \sigma_e^2)$. Σ appears in both terms of the likelihood factorisation. As a consequence the conditional posterior is not known in closed form. We therefore sample from this conditional posterior using a random walk Metropolis-Hastings step.

4. $g(\sigma_e^2|\Sigma, \phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \mu, \Phi)$. Conditional on the remaining parameters, the model collapses to the multi-variate regression $Y_t = A + BX_t + e_t$ where $var(e_t) = \sigma_e^2$. Given an Inverse Gamma (IG) prior for σ_e^2, the conditional posterior is also IG with scale parameter $\tilde{e}_t'\tilde{e}_t + \sigma_0^2$ and degrees of freedom $T(J-N) + T_0$ where $\tilde{e}_t = vec(e_t)$ and σ_0^2 and T_0 are the prior scale parameter and degrees of freedom, respectively.

4.1.2 Model $\mathcal{M}_1$

The estimation of model $\mathcal{M}_1$ differs from the estimation of model $\mathcal{M}_0$ only in step 1. The model is re-parameterised in terms of the risk prices $\lambda' = (\lambda_0', vec(\lambda_1)')$, where λ has dimension $N(N+1) \times 1$. The $\mathbb{P}$-dynamics parameters are consequently obtained from equations (8). Bauer (2018) shows that this results in a restricted VAR under the $\mathbb{P}$-measure. Restrictions on λ are implemented using the SSVS algorithm of George and McCulloch (1993). This algorithm involves a hierarchical prior on the elements of λ. In particular, the prior for λ is assumed to be normal: $(1-\gamma)\,N(0, \tau_0^2) + \gamma N(0, \tau_1^2)$ where γ is a $N(N+1) \times 1$ vector of ones and zeros, τ_0^2 is a small number while τ_1^2 is chosen to be large number.[12] If the ith element of γ equals zero, the prior for the ith risk price is tightly centered around zero. On the other hand if the ith element of γ equals 1, the prior for the corresponding risk price is loose and the posterior can be different from zero. Here, γ is treated as an additional set of unknown parameters. We employ an uninformative prior for the elements of γ, implying a-priori a 50% chance for risk price to be included or excluded (ie, set to zero).

[12] τ_0 is set equal to $\frac{1}{c}\hat{\sigma}_\lambda$ and $\tau_1 = c\hat{\sigma}_\lambda$ where $\hat{\sigma}_\lambda^2$ is an estimate of the variance of λ based on maximum likelihood estimates of the model parameters. c is set to 100.

The vector of parameters to be estimated for model $\mathcal{M}_1$ is given by $\Theta = \left(\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma, \sigma_e^2, \lambda, \gamma\right)$. The conditional posteriors for $\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma, \sigma_e^2$ remain unchanged from the ones described in steps 2-4 of the algorithm for model $\mathcal{M}_0$. As shown in George and McCulloch (1993), the conditional posterior for γ is Bernoulli. Given a draw of γ, the prior for λ is specified as described above. Bauer (2018) shows that the conditional posterior for λ is normal and derives the mean and the variance.

4.1.3 Model $\mathcal{M}_2$

The estimation of model $\mathcal{M}_2$ differs from that of the maximally flexible specification again only in step 1. Specifically, it differs in terms of the prior used for the VAR parameters $\beta' = vec\left(\mu, \Phi\right)'$. We assume the following prior for β: $N\left(\beta_b, V_b\right)$, where β_b denotes the bootstrap bias corrected estimator of the VAR coefficients used by Bauer et al. (2012). The covariance V_b is a diagonal matrix with elements on the main diagonal set to $1e-04$. Therefore, this prior reflects a very strong belief that the VAR coefficients are close to the bias corrected estimates. The rest of the steps of the MCMC algorithm are identical to the ones used for model $\mathcal{M}_0$.

4.2 Data for estimation of the ATSMs

We estimate the models using both monthly and daily data for yields at maturities of 1, 3 and 6 months and 1 through 10 years. That is, thirteen maturities in total. We use both frequencies for robustness. Monthly data conform in terms of frequency to the macro model, and are commonly used in the ATSM literature, while daily data are closer to the intra-day decompositions. The data at maturities of one year and above are obtained from the Federal Reserve Board database on the nominal yield curve (the Gürkaynak-Sack-Wright data set), with rates at shorter maturities taken from the FRED database. The sample runs from January 1990 to December 2008.

5 Results

This section consists of three parts. First, we present key results for the three models estimated using monthly and daily data, respectively. The purpose of inspecting these results is to gauge the effects of imposing the restrictions on the ATSM and to determine whether the estimated models display reasonable properties. Second, we use the estimated models to decompose movements in the yield curve around FOMC announcements into term premia and interest rate expectations. And third, we use the intra-day yield curve components as instruments in a local projection model to study the dynamic effects of monetary policy shocks thus identified on selected macro variables.

5.1 Estimated ATSMs

All models display a good fit to the data with root mean squared errors that are below 5 basis points. The estimates of the $\mathbb{P}$ and $\mathbb{Q}$ parameters, and the implied λ's, are shown in Tables 1-3 of the Appendix.

5.1.1 Persistence and volatility

Figure 2 shows the estimated posterior distributions of the largest eigenvalues of $\Phi^{\mathbb{Q}}$ and Φ obtained from the three monthly models. The top panel of the figure shows that under the $\mathbb{Q}$-measure, the three models have a very similar profile in terms of persistence. This, of course, is expected as the estimates are based on the same cross-sectional information and the partial likelihoods for the $\mathbb{Q}$-measure differ across the models only in terms of Σ. The results, however, are very different for the eigenvalues under the $\mathbb{P}$-measure. The bottom panel of the figure shows that the maximally flexible specification has the lowest persistence under the $\mathbb{P}$-measure. The median estimate of the highest eigenvalue for this model is 0.9827 with a 68% error band of 0.9685 and 0.9937. In contrast, the posterior estimates for the restricted models $\mathcal{M}_1$ and $\mathcal{M}_2$ are shifted to the right. The median estimate of the maximum eigenvalue for model $\mathcal{M}_1$ is 0.9925 with an error band of $(0.9784, 0.9970)$. The estimates

for model $\mathcal{M}_2$ are more precise, with the median and error bands given by 0.9961 and $(0.9923, 0.9987)$, respectively. Thus, restrictions on risk prices or statistical bias correction lead to a substantial increase in persistence. To illustrate this, take the median values to the power of 120 to derive their effect on excepted interest rates ten years ahead. This exercise results in 0.12, 0.41, and 0.63 percentage point increase in the nominal short rate in ten-years time for the three models respectively, for one percentage point increase in the current short rate.

Figure 3 tries to assess more formally the consequences of the differences in persistence under the $\mathbb{P}$-measure on the volatility of the two main components of yields. The term premium is calculated as the difference between the fitted yield $\hat{Y}_t$ and the forecast of the average short rate over the relevant horizon, based on the VAR parameters under the $\mathbb{P}$-measure. Note that the latter represents the expectations component of yields. We focus on the 10-year yield. The top panel of Figure 3 plots the posterior distribution of the unconditional standard deviation of the term premium for the three models, while the bottom of the figure shows the posterior distribution for the expectations component (as the two components may be correlated, the respective volatilities do not necessarily add up to the volatility of yields). The figure shows that the term premium from the maximally flexible model is the most volatile when compared with the alternative specifications. In contrast, models $\mathcal{M}_1$ and $\mathcal{M}_2$ produce a more volatile expectations component than the unrestricted model.[13]

Figure 4 presents the measure of persistence for the estimated daily models.[14] There are two features of these results that distinguish them from the monthly estimates. First, the largest eigenvalue under the $\mathbb{P}$-measure is higher than in the monthly estimates, with the posterior median estimated to be 0.9992 for all three models. This, of course, is not

[13]The charts seem to suggest that the standard deviation may be negative. This, however, is just an artefact of the kernel density used to approximate the distribution.

[14]The headline specification of the daily model contains only one lag of the risk factors in the state space. This is theoretically consistent with the framework outlined in Section 3 under spanning (see Duffee, 2012). Nevertheless, to allow for the possibility that all relevant information is not reflected in bond prices on the same day, we also estimated the model with up to 21 lags (one month), with AIC criterion picking up 12 lags. This turned out not to affect the persistence of the models.

surprising, as the daily models are estimated on data sampled at a higher frequency than the monthly counterparts. When converted to monthly frequency, the persistence is similar to that of the unrestricted monthly model $\mathcal{M}_0$.[15] Second, restrictions on risk prices and bias correction do not affect the results substantially. One consequence of this result is that the properties of the components of yields do not vary much across the daily models, as demonstrated in Figure 5.[16] Nevertheless, comparing Figures 3 and 5, we see that the expectations component in all three daily models is more volatile than in the unrestricted monthly model. This occurs despite the aforementioned similarity of the $\mathbb{P}$-measure persistence between the daily models and the unrestricted monthly model. It is due to a change in the estimated correlation between the term premium and the expectations component, when compared to the unrestricted monthly model. Estimating the ATSM, with or without restrictions, on daily data thus brings the volatility of the expectations component closer to the restricted versions of the ATSM estimated on monthly data. It appears that daily time series data contain information that makes the expectations component of yields more volatile than what monthly time series data would imply.[17]

In sum, imposing restrictions on the monthly ATSM to correct for the small sample bias, or estimating the ATSM on daily data, increases the volatility of the expectations component above what a simple monthly VAR model would suggest.

[15]The implied eigenvalue at the monthly frequency can be obtained by raising the daily estimate to the power of 21. The posterior median and error band of the implied monthly eigenvalue are 0.983 and $(0.97, 0.993)$, respectively, and are essentially the same as those of the unrestricted monthly model. However, relative to the unrestricted monthly model, the distribution of the estimates for the daily model is skewed to the right (ie, towards higher persistence).

[16]One possible reason why the restrictions do not substantially affect the estimates of the $\mathbb{P}$-measure is that daily persistence is already quite high and any attempt to increase it, by subjecting the model to restrictions, operates in a region of significant digits that make any numerical refinements difficult to achieve.

[17]The similarity of the estimates across the three versions of the daily model is not a general property of daily models and depends on the sample period. For instance, when the daily model is estimated using data starting in 1982 (that is, including the Volker disinflation period), the persistence under the $\mathbb{P}$-measure is estimated to be higher in the two restricted versions of the daily model than in the maximally flexible one. Estimating an ATSM subject to restrictions thus seems, a-priori, a robust strategy even when the estimation is based on daily data and ex-post the restrictions turn out not to bind in a given sample.

5.1.2 The implied time series

The time series of the term premia and the expectations components for the 10-year yield (based on the median posterior estimates of the parameters) are presented in Figures 6 and 7, for the monthly and daily models, respectively. The left panel of Figure 6 shows that the estimated term premium from the monthly maximally flexible model has a clear declining trend, falling from about 5 percent in early 1990s to around 1 percent at the end of the sample. The term premia from models $\mathcal{M}_1$ and $\mathcal{M}_2$ are more stable across the sample period. It is also apparent from the right panel of the figure that the expectations component from models $\mathcal{M}_1$ and $\mathcal{M}_2$ shows variation over the sample that is larger than in the unrestricted model. In all specifications, the correlation between the term premium and the expectations component is positive, slightly above 0.5.

As noted above, whereas the persistence under the $\mathbb{P}$-measure (when converted to monthly frequency) in all three daily models is comparable to the persistence of the unrestricted monthly model, the volatility of the expectations component in the three daily models is closer to the volatility in the restricted monthly models $\mathcal{M}_1$ and $\mathcal{M}_2$. This finding shows up also in Figure 7. The major swings in the expectations component—during the early 1990s, the early 2000s and in 2008—are twice as large in the daily models as in the unrestricted monthly model.

Figures 6 and 7 suggest that in the restricted monthly models $\mathcal{M}_1$ and $\mathcal{M}_2$, as well as in the daily models, term premia are elevated during the NBER recessions (shaded areas), whereas expected interest rates decline. To provide a better sense of this comovement, we compute correlations of term premia and expected interest rates with the index of industrial production, which we take as a monthly proxy for the business cycle.[18]

Focusing on the 10-year bond and the monthly models, Table 2 shows that the term premium in both restricted monthly models exhibits a negative correlation with industrial production, whereas in model $\mathcal{M}_0$ the correlation is essentially zero. The correlation coeffi-

[18] A number of theoretical models point towards countercyclical prices of risk; ie, high prices of risk in downturns (eg, Campbell and Cochrane, 1999; Wachter, 2006; Atkeson and Kehoe, 2009).

cient is more negative for model $\mathcal{M}_1$ than model $\mathcal{M}_2$, -0.33 vs. -0.17. In all three models, the expectations component, in contrast, is positively correlated with industrial production. For completeness, the table also includes correlations for the 3-month and the 10-year yields and their spread. The 3-month yield is more strongly positively correlated with the business cycle than the 10-year yield, with the spread exhibiting a strong negative correlation. These statistics are thus consistent with the narrative suggesting a decline in nominal interest rates and interest rate expectations in downturns, which are partially offset for long yields by an increase in term premia. The restricted monthly models fit better this narrative, especially model $\mathcal{M}_1$, whereas the unrestricted model fails to generate countercyclical term premia.

5.2 High-frequency decomposition

With the estimated ATSMs in hand, we turn to the first key question of the paper: Are yield movements around FOMC meetings driven by changes in expectations or term premia? Note that as we estimate both restricted and unrestricted ATSMs and use monthly and daily data, we are also able to investigate how the small sample bias affects the decomposition. This is an advantage over relying on simple VARs (eg, Gertler and Karadi, 2015) to decompose yields into term premia and expectations.

5.2.1 Implementation

While the decomposition differs slightly across the monthly and daily models, in nutshell, our procedure is based on feeding into the estimated models changes in the risk factors observed in a narrow window around FOMC meetings and using the structure of the models to decompose the changes in fitted yields into changes in term premia and the expectations component.[19] In what follows, the results are based on the median posterior estimates of the parameters.

For the monthly ATSMs, the decomposition follows a simple procedure. Let $\Delta\tilde{X}_t$ denote

[19] As our shortest maturity in the HF dataset is three months, all maturities used in the HF analysis are determined in the market and thus reflect the market reaction to a policy announcement.

the first four PCs obtained from the time series on the *changes* in yields in a tight window around FOMC meetings (recall that the observed yields are recorded 10 minutes before and 20 minutes after the meeting and the data consist of maturities of 3 months, and 2, 3, 5, and 10 years).[20] The changes in term premia and expectations are then computed using $\Delta\tilde{X}_t$ and the estimated models in the usual way. In terms of the notation of Section 3, the vector of changes in the expectations component, for the five maturities, is given by $B^{\mathbb{P}}\Delta\tilde{X}_t$ and the vector of changes in term premia is given by $(B - B^{\mathbb{P}})\Delta\tilde{X}_t$, where $B^{\mathbb{P}}$ and B are the estimated $\mathbb{P}$- and $\mathbb{Q}$-measure loadings given by (10) and (9), respectively. We also derive $\Delta\tilde{Y}_t = B\Delta\tilde{X}_t$, where $\Delta\tilde{Y}_t$ is the change in the fitted HF yields.[21]

The decomposition using daily models proceeds in a similar manner, but is fine-tuned to align the models more closely with the high frequency data. In particular, before carrying out the decomposition, the daily models are re-estimated by replacing the observations on the day before and on the day after a FOMC meeting with the high frequency data recorded 10 minutes before and 20 minutes after the meeting. The parameter estimates are then used to calculate the loadings and decompose the changes in yields around FOMC meetings as for the monthly models (the re-estimation has no effect on the findings reported in Section 5.1).

5.2.2 Findings

Figures 8 and 9 show the contributions of term premia and expectations components to the changes in yields (for 3-month, 5- and 10-year bonds) around FOMC meetings in the three monthly and daily models, respectively, together with the observed changes in yields. Starting with the monthly models, Figure 8 shows that changes at the short end of the

[20] Given that the set of maturities in the HF dataset is only a subset of the maturities used to estimate the models, one may wonder how different the estimated parameters of the ATSMs would be if only the maturities of the HF dataset were used in the estimation. It turned out that the estimates are almost identical. The maturities in the HF dataset thus seem to capture all of the main movements in the yield curve over time.

[21] The root mean squared error of the fit of the models at the high frequency is about 3 to 4 basis points across all models, comparable to their fit at the monthly and daily frequencies. The models, however, have a hard time fitting well the largest HF changes at the shortest maturity, as is apparent from Figures 8 and 9 below.

yield curve are largely explained in all three specifications by changes in the expectations component. The figure further reveals that the maximally flexible model attributes the bulk of the movement in yields at the 5- and 10-year horizon to the term premium component. This matches the conclusions reached by Gertler and Karadi (2015), who show that the responses of long-term rates to monetary policy surprises are almost completely due to a reaction of term premia. However, once the small sample bias in the VAR parameters under the $\mathbb{P}$-measure is corrected (models $\mathcal{M}_1$ and $\mathcal{M}_2$), the resulting decomposition at the 5- and 10-year horizons offers conclusions that are different. As is apparent from Figure 8, the chunk of the light bars denoting the contribution of term premia that dominate the charts for model $\mathcal{M}_0$ turn into a dark color, denoting the contribution of expectations, as we move from model $\mathcal{M}_0$ to models $\mathcal{M}_1$ and $\mathcal{M}_2$. Imposing restrictions on risk prices or correcting the statistical bias thus prescribes a more important role to the expectations component in accounting for the movement of interest rates at medium and long horizons. Using daily data has a similar, though not as strong, effect, as emerges from Figure 9. All daily models have higher volatility of the expectations component at the 5- and 10-year horizon, relative to the unrestricted monthly model.

Figure 10 provides a summary of the properties of the time series plotted in Figures 8 and 9. It plots the volatility curve of the HF changes in the two components. Figure 10 clearly demonstrates that imposing restrictions on estimation, as well as using daily data, substantially increases the reaction of expected future interest rates to FOMC announcements. In the restricted monthly models $\mathcal{M}_1$ and $\mathcal{M}_2$, the volatility of the change in the expectations component at the 10-year horizon is as large as the volatility of term premia, and about twice as large as in the unrestricted model $\mathcal{M}_0$. In the daily models, while term premia are more volatile than expectations, the volatility of expected interest rates at the 10-year horizon is again about twice as large as in the unrestricted monthly model. These findings demonstrate that a simple VAR can severely underestimate the effects of monetary policy surprises on expected interest rates.

Note that the plotted volatilities in the two components of the yield curve in Figure 10 do not represent a variance decomposition. Any of the risk factors contained in $\Delta\tilde{X}_t$ can potentially move both term premia and expectations. Term premia and expectations can thus be correlated. The solid line at the bottom of the charts in Figure 10 plots the correlation at a given maturity. We note here that, in the HF movements, only the monthly model $\mathcal{M}_1$ exhibits a positive correlation between the two components at the 10-year horizon, of around 0.5, which is about the same as the correlation in all the models at their respective either monthly or daily frequencies.

5.2.3 Potential interpretations

While it is beyond the scope of the paper to provide an economic model to justify the findings from a theoretical perspective, it is worth asking if the results can be, in any way, informative about possible mechanisms at play. To this end, we first carry out a PC decomposition of the change in *yields* around FOMCs. As in Gürkaynak et al. (2005a), we find that HF monetary policy surprises are two dimensional. Two PCs explain the bulk of the variation in the change in yields across maturities in the FOMC window, with contributions of 85 and 11 percent, respectively. With the ATSMs in hand, however, we can go further and ask about the dimensionality of the HF changes in the two *components* of yields, term premia and interest rate expectations. The broad picture that emerges across the various specifications of the ATSM is that the HF change in interest rate expectations is two-dimensional, with its first and second PCs explaining 83-90 percent and 10-17 percent, respectively. The HF change in term premia, in contrast, has a first PC accounting for up to 99 percent of its variation, depending on the model.

The loadings of the PCs of expectations and term premia for the different maturities are presented in Figure 11. They are broadly consistent across the different models. An interesting feature of the loadings for interest rate expectations is that the patterns are consistent with theoretical responses of nominal interest rates to certain monetary policy shocks.[22] The

[22]Under rational expectations, the dynamic impulse-responses should correspond to the expectations em-

pattern for the first PC (the upper left chart) resembles responses to a persistent shock inducing a shift in inflation expectations (along the lines of the papers referred to by Ireland, 2007), or an information shock about the long-term real rate, as in Nakamura and Steinsson (2018). The pattern for the second PC (the upper right chart) resembles responses to the standard temporary monetary policy shock in New-Keynesian models (eg, Galí, 2015), albeit with not so well anchored long-term inflation expectations along the lines of Gürkaynak et al. (2005b), accounting for the negative loadings at the medium and long end. The loadings of the first PC of term premia, unsurprisingly, increase with maturity.

The patterns in Figure 11 can help us interpret the basic correlations in the HF change in yields reported in Table 1. Based on the patterns in Figure 11, the high correlations across medium and long rates in Table 1 can occur due to both the first PC of the expectations component and the first PC of term premia. The pattern for the second PC of expectations, producing a negative correlation between the short and medium to long end, is then responsible for the weaker correlation in Table 1 between the short rate on the one hand and the medium and long rates on the other.

5.3 Local projections

To estimate the impact of policy shocks on macroeconomic variables of interest, we use Bayesian local projections (BLP) introduced by Miranda-Agrippino and Ricco (2015). As in Jordà (2005), the model is

$$Z_{t+h} = c^{(h)} + B_1^{(h)} Z_t + \sum_{j=1}^{P} b_j^{(h)} Z_{t-j} + v_{t+h}, \tag{12}$$

where Z_t denotes the M variables of interest, h is the impulse-response horizon and v_{t+h} denotes residuals that represent combination of forecast errors and, therefore, are serially correlated and heteroscedastic. The impulse-response to the shock of interest at horizon h can be calculated as $B_1^{(h)} A_0$, where A_0 denotes the contemporaneous impact matrix. As

bedded in the yield curve.

discussed below, the column of the A_0 matrix corresponding to the shock of interest is estimated using instruments based on the high-frequency change in the components of yields identified above.

Relative to VARs, the main advantage of local projections (LP) is that they are not subject to biases that may arise if the VAR model is misspecified or subject to a severe small sample bias, as in the case of interest rates. By using LPs, we mitigate these concerns. However, LPs are likely to be less efficient than VARs. This is due to the non-spherical nature of v_{t+h} and because of the fact that LPs involve an estimation of a large number of coefficients. We address this bias vs. variance trade-off by following the Bayesian approach of Miranda-Agrippino and Ricco (2015) and use prior information to reduce the variance of the estimates. To elicit priors for the coefficients of the LP, Miranda-Agrippino and Ricco (2015) exploit the fact that there is a mapping between the impulse-responses at horizon h obtained using a LP and those obtained from a VAR iterated forward h periods. This is done by using the implied VAR coefficients at horizon h as the mean of their prior. The tightness of the prior is set in an optimal manner by maximising model fit and penalising overfitting (see Giannone, Lenza, and Primiceri, 2015). Miranda-Agrippino and Ricco (2015) derive the posterior distributions of the parameters of the BLP and provide a Gibbs sampling algorithm to approximate the posterior distributions.[23]

As discussed by Miranda-Agrippino and Ricco (2015), the *contemporaneous* impulse-response in LPs and VARs is equivalent and the A_0 matrix can be estimated using standard structural VAR (SVAR) methods.[24] We use an external instruments or 'Proxy' SVAR to calculate the A_0 matrix. Introduced by Mertens and Ravn (2013), the approach differs from standard SVARs, where the A_0 matrix is estimated by placing restrictions on the impulse-responses at various horizons. Instead, a Proxy SVAR exploits variation in an instrumental variable that is external to the VAR to isolate the portion of the VAR residuals that are associated with the shock of interest. More formally, we assume that the instrument m_t is

[23]We use a modified version of the Matlab codes provided by Silvia Miranda-Agrippino on her web page (http://silviamirandaagrippino.com/code-data).

[24]The estimation of the A_0 matrix is not subject to the problems with VARs noted above.

relevant and exogenous:

$$\text{cov}\left(m_t, U_{1t}\right) = \alpha \quad \text{and} \quad \text{cov}\left(m_t, U_{-}\right) = 0, \tag{13}$$

where U_{1t} denotes the structural shock of interest, $\alpha \neq 0$ and U_{-} are the remaining structural shocks. Let E_t denote the reduced-form residuals of a VAR estimated using Z_t and note that $E_t = A_0 U_t = A_{0,1} U_{1t} + A_{0,(-)} U_{-}$, where U_t is the matrix of structural shocks, $A_{0,1}$ denotes the column of interest of the A_0 matrix, with the remaining columns denoted by $A_{0,(-)}$. The choice of the first column of the matrix A_0 is without loss of generality. Using this relationship and the assumptions in equation (13), the covariance between the VAR residuals and the instrument can be written as

$$\text{cov}\left(m_t, E_t\right) = \text{cov}\left[m_t, \left(A_{0,1} U_{1t} + A_{0,(-)} U_{-}\right)\right] = A_{0,1}\alpha. \tag{14}$$

With an estimate of the covariance $\text{cov}\left(m_t, E_t\right)$ in hand, equation (14) can be re-arranged to calculate a normalised version of $A_{0,1}$. In other words, for the jth residual: $\text{cov}(m_t, E_{j,t}) = A_{0,1}^{(j)}\alpha$, where $A_{0,1}^{(j)}$ is the jth element of $A_{0,1}$. Taking the ratio of the covariance of m_t and $E_{j,t}$ and the covariance of m_t and $E_{1,t}$ eliminates α, an unknown, from the right-hand side of equation (14) and provides an estimate of the normalised elements of $A_{0,1}$

$$\frac{A_{0,1}^{(j)}}{A_{0,1}^{(1)}} = \frac{\text{cov}(m_t, E_{j,t})}{\text{cov}(m_t, E_{1,t})}, \quad j = 2, 3, ..M. \tag{15}$$

5.3.1 Variables and instruments

In the BLP model, the control variables Z include 12 lags of the following variables: the log of industrial production, the log of CPI, the excess bond premium (Gilchrist and Zakrajšek, 2012), and the first two PCs of yields that were used as risk factors in the estimation of the ATSMs.[25] The first three variables are standard in the empirical macro literature. The two

[25]The excess bond premium is the component of the spread between an index of rates of return on corporate securities and a similar maturity government bond rate that is left after the component due to default risk

PCs are included in order to account for information in the yield curve (they account for 99% of the variation in yields across maturities). The first PC loads positively on yields at all maturities and thus shifts the level of the yield curve (level factor). The second PC has positive loadings on yields at maturities of 2 years and less but loads negatively at the longer end of the curve (the negative of the slope factor). In addition to these benchmark series, we consider responses of housing market variables: the 30-year mortgage rate, the log of new single-family home sales, and the log of real house prices. We include housing variables in the analysis as housing is heavily mortgage dependent and thus should be particularly sensitive to changes in the yield curve.[26] The BLP models are estimated on the monthly sample 1990-2007, which is typical for models that do not consider unconventional monetary policies. As in Miranda-Agrippino and Ricco (2015), the prior distributions are set using a training sample, which spans the period 1982-1989.

We consider three instruments to identify monetary policy shocks. The first two instruments are the first two PCs of the expectations component of the HF yield changes around FOMC meetings. As shown in Figure 11, these PCs represent very different combinations of expected yields across maturities. The first PC (m_{1t}) affects expectations at all maturities with the same sign and resembles a level factor of expectations. In contrast, the second PC (m_{2t}) loads positively on the 3-month maturity and negatively on other maturities. It thus resembles the negative of a slope factor of expectations. The third instrument (m_{3t}) that we consider is the first PC of the HF change in term premia. The findings reported below are based on the instruments extracted from model $\mathcal{M}_1$.[27]

The instruments m_{1t} and m_{2t} are uncorrelated by construction. The correlations between the instruments based on the expectations and term premia components are fairly low, suggesting that they proxy distinct shocks. The correlation between m_{1t} and m_{3t} is estimated

is removed. It is typically interpreted as a measure of credit market conditions in the non-farm business sector.

[26]Except the excess bond premium and the PCs of yields, the data come from either FRED or Haver.

[27]Given the similar patterns across the models in Figure 11, the impulse-responses are not particularly sensitive to which model is used to obtain the instruments from, as confirmed by sensitivity analysis. Only the responses of home sales show some quantitative differences across the models, however without affecting the overall message.

to be only 0.14. As shown in Figure 11, the loadings of m_{3t} rise sharply at short maturities and flatten out at longer maturities, a feature not shared by m_{1t}. It would appear from Figure 11 that at least some of the information provided by m_{2t} may overlap with m_{3t}. The pattern of the loadings on m_{3t} matches those associated with the second PC of expectations, m_{2t}, albeit with the opposite sign. However, the magnitude of the loadings on m_{3t} is smaller at short maturities and larger towards the long end. This implies a correlation between m_{2t} and m_{3t} that is fairly small (-0.08).

Given the low correlations between the instruments based on the expectations and term premia components, we follow the approach of Stock and Watson (2012), and employ these instruments one at a time. We then examine the resulting impulse responses to determine the differences in the three identified shocks, ex-post. We show in the technical online appendix that the impulse-responses are robust to a more complex identification scheme where two shocks are identified simultaneously, based on the instruments m_{1t}, m_{3t} and m_{2t}, m_{3t}, respectively.[28]

Our instruments generalise the proxies for monetary policy shocks employed in previous studies. For instance, Gertler and Karadi (2015) identify monetary policy shocks by using the change in a single interest rate around FOMC meetings. In contrast, we exploit the entire term structure of interest rates. Our approach is closer to the seminal work of Gürkaynak et al. (2005a) who show that FOMC announcements are adequately described by two factors obtained from the cross-section of high frequency yields. These factors have been used in a number of papers to identify conventional and unconventional monetary policy shocks (see Ferreira, 2020; Lakdawala, 2019). However, we go beyond this approach and use the ATSMs to decompose the HF changes in the yield curve into various components of term premia and interest rate expectations. Such decomposition provides sharper implications for economic models than instruments that bundle the different components together.

[28] As the instruments are allowed to be correlated, extra restrictions are required to separate the two shocks. We use the procedure developed by Piffer and Podstawski (2018) where inequality restrictions are used to ensure that each instrument is more correlated with the shock it is meant to identify.

5.3.2 Findings

Figures 12-14 report the responses from the local projection model to monetary policy shocks identified by the three instruments, m_{1t}, m_{2t} and m_{3t}, respectively. The figures plot the median response and the 90 percent error band.[29] In each case, the shock is scaled to increase the negative of the slope factor by 1 percent on impact. This implies a rise in short-term interest rates on impact. We choose this normalisation as it fits the common notion of a policy shock and is a natural normalisation for the interpretation of the responses identified by instruments m_{2t} and m_{3t}. The horizon for the responses in the figures is 36 months.

While it is beyond the scope of the paper to provide an economic model to justify the findings from a theoretical perspective, the impulse-responses appear to broadly fit the narrative of existing theories. First, consider the responses to the shock identified by m_{1t}. Recall from Figure 11 that m_{1t} is characterized by an increase, around FOMCs, in interest rate expectations across maturities, accounting for up to 90 percent of their movement. In Figure 12, the shock identified by this instrument leads to an increase in the level factor of yields, implying that interest rates rise across maturities. CPI increases and stays positive over the horizon considered. The upward trajectory of the median CPI response suggests that annual inflation increases persistently in response to this shock. The shock is thus at least partially characterised by a Fisher-type effect. Industrial production persistently declines. The increase in the level factor is reflected by a similar increase in the 30-year mortgage rate.[30] The black dotted line in the chart for the mortgage rate shows that at least half of the response of this long-term interest rate is due to an increase in interest rate expectations (implied by the ATSM $\mathcal{M}_1$). The response of new home sales and real house prices essentially tracks the

[29] Mertens and Ravn (2013) propose a reliability statistic to check the strength of instruments in a proxy SVAR setting. This statistic can be interpreted as the squared correlation between the instrument and the structural shock of interest and is constrained to lie between 0 and 1. The estimated reliability for m_{1t} is 0.15 with a 90 percent error band of $(0.1, 0.19)$. The median reliability estimate for m_{3t} is 0.12 with an error band of $(0.1, 0.14)$. The reliability of m_{2t} is slightly lower with a point estimate of 0.1 and an error band of $(0.06, 0.11)$. These values are comparable to those for the instruments used by Caldara and Herbst (2019), who also employ a Bayesian approach.

[30] The response of the 10-year yield is almost identical to that of the mortgage rate in all impulse-responses considered.

response of the mortgage rate, but with the opposite sign. That is, housing market activity declines.

In relation to the pattern of the loadings for the instrument m_1 in Figure 11 we have noted that the pattern is suggestive of a policy surprise along the lines of both, a shock inducing a change in inflation expectations (eg, Ireland, 2007) or a positive signal about the future state of the economy, which gets reflected in a persistent increase in the real rate (eg, Nakamura and Steinsson, 2018). While the shock identified by m_1 can be interpreted as partially reflecting a persistent increase in the real rate, the persistent decline of industrial production is inconsistent with a revelation of a positive news about future economic growth, as in the model of Nakamura and Steinsson (2018).[31] It appears that this shock depresses the housing market by increasing the mortgage rate, with the reduced activity in the housing market possibly spilling over into the rest of the economy. The response of the mortgage rate and the housing market variables is consistent with the price effect in the model of Garriga, Kydland, and Šustek (2017), regardless of whether the shock identified by m_1 is a pure policy shock associated with a Fisher-type effect or reflects a persistent change in the real rate. The findings regarding the responses of the mortgage rate, and the accompanying decline in housing market activity, are also supportive of the conclusions reached by Hamilton (2008), who stresses the important role of interest rate expectations and the housing market in the transmission of US monetary policy.

Turning to the instruments m_{2t} and m_{3t}, recall from Figure 11 that both instruments are associated with a HF change in the slope of the yield curve, though for different reasons. The instrument m_{2t} is related to a HF change in expectations (accounting for up to 17 percent of their movement), whereas m_{3t} is related to a HF change in term premia (accounting for up to 99 percent of their movement). The policy shocks identified by these instruments generate similar responses of some variables, as is apparent from Figures 13 and 14. Specifically, of the level and slope factors, inflation, and industrial production. In response to both shocks, the

[31] Interestingly, industrial production weakly increases when, instead of m_1, we use as an instrument the HF change in the first PC of *yields*. The first PC of yields, however, bundles together both expectations and term premia.

level of the yield curve declines, the slope changes whereby the short rate increases while the long rate declines, and both CPI and industrial production decline (the latter temporarily). However, the two shocks markedly differ in their effects on the excess bond premium and the source of the decline in the mortgage rate (the shocks also have a different effect on house prices). In particular, the shock identified by m_{2t} is associated with a small decline in the excess bond premium, whereas the shock identified by m_{3t} is associated with a sharp increase in this variable. Further, in the former case the mortgage rate declines predominantly due to a decline in interest rate expectations (the black dotted line), whereas in the latter case the decline is due to term premia (the blue dashed line). Thus, while the outcome for output and inflation is similar, the two shocks appear to propagate through different mechanisms.

The responses to the shock identified by m_2 seem to fit the narrative of a New-Keynesian mechanism, albeit with not well anchored inflation expectations (Gürkaynak et al., 2005a). According to this mechanism, a temporary surprise increase in the short rate convinces the market that the Fed is more serious about inflation than previously thought, thus reducing long-term interest rates due to a Fisher effect.[32] The temporary decline in industrial production and prices fits the standard New-Keynesian channel (eg, Galí, 2015). This channel, however, appears to be somewhat muted by the increase in housing market activity, resulting from the decline in the mortgage rate.

The responses to the shock identified by m_3, in contrast, are suggestive of a financial accelerator mechanism (Bernanke, Gertler, and Gilchrist, 1999). The worsening of conditions in the market for business credit, captured by the sharp rise in the excess bond premium, leads to a decline in output and prices. An interesting additional feature of the impulse-responses, however, is a decline in the mortgage rate, most of which is due to a decline in term premia. The shock identified by m_3 thus appears to reduce the cost of credit to households, while making accessing credit harder for firms. The latter seems to have a dominant effect on output in the short run.

[32]Some other work (eg, Coibion, Gorodnichenko, Kumar, and Pedemonte, 2020), however, suggests that inflation expectations are not particularly sensitive to monetary policy, especially in a low inflation environment.

6 Conclusions

This paper investigates the relative importance of expected future interest rates and term premia in the transmission of monetary policy. To this end, we adopt a two-stage procedure. First, we decompose high-frequency movements in the yield curve around FOMC meetings into the two components. Unlike existing work on the topic, we carry out this decomposition using term structure models that correct for a small sample bias in the estimates of the two components inherent in time series data. Earlier studies that are subject to this problem have found that the transmission of monetary policy occurs almost exclusively through term premia. We find that once the small sample bias is accounted for, the reaction of expected interest rates to policy surprises is as large as that of term premia.

Second, we decompose the high-frequency reaction of expected interest rates and term premia across maturities into their respective principal components. We find that almost all of the reaction of expected interest rates to FOMC announcements is summarised by two principal components, whereas only one principal component is sufficient to capture the bulk of the reaction of term premia. The first principal component of expected interest rates rises all interest rates across maturities, whereas the second principal component of expected interest rates, as well as the first principal component of term premia, changes the slope of the yield curve. The principal components are then used as instruments to identify three types of monetary policy disturbances in a local projection model of selected macroeconomic variables. These instruments generalise the proxies for monetary policy shocks employed in previous studies, which are typically based on a single maturity.

The first principal component of the change in expected interest rates identifies a monetary policy shock that fits the predictions of models with a role for housing finance. This shock is associated with a persistent increase in the price level, expected interest rates and the 30-year mortgage rate and a decline in housing market activity and output. The shock identified using the first principal component of high-frequency movements in term premia leads to a fall in output and the price level and a sharp rise in credit spreads, suggesting a

financial accelerator type effect. The second principal component of expected interest rates identifies a shock that has similar effects on output and the price level as the shock identified by the term premium component, but appears to be transmitted differently, fitting the standard narrative of New-Keynesian models.

Our analysis has been carried out on the sample preceding the 2008 global financial crisis and the subsequent zero lower bound and unconventional monetary policies. The findings thus characterise the transmission mechanism in a conventional setting. The analysis could be extended to the subsequent period, when the policy interest rates approached the zero lower bound and the Federal Reserve launched the policy of quantitative easing (QE). However, to adequately account for the zero lower bound, the term structure model would need to depart from the convenient affine representation, as, for example, in Wu and Xia (2016). Extending the analysis to cover the QE period would enable us to separate QE-related term premia surprises from the implied policy rate ("forward guidance") surprises. Swanson (forthcoming) is a recent contribution in this direction. This should improve our understanding of the unconventional monetary policy effects and their interplay with conventional monetary policy channels. We see such extensions as a promising avenue for future research.

References

Abrahams, M., Adrian, T., Crump, R. K., Moench, E., Yu, R., 2016. Decomposing real and nominal yield curves. Journal of Monetary Economics 84, 182–200.

Atkeson, A., Kehoe, P. J., 2009. On the need for a new approach to analyzing monetary policy. In: NBER Macroeconomics Annual, Volume 23. National Bureau of Economic Research, Inc.

Bauer, M. D., 2015. Inflation expectations and the news. International Journal of Central Banking 11, 1–40.

Bauer, M. D., 2018. Restrictions on risk prices in dynamic term structure models. Journal of Business and Economic Statistics 36:2, 196–211.

Bauer, M. D., Rudebusch, G. D., Wu, J. C., 2012. Correcting estimation bias in dynamic term structure models. Journal of Business and Economic Statistics 30, 454–467.

Beechey, M., 2007. A closer look at the sensitivity puzzle: The sensitivity of expected future short rates and term premia to macroeconomic news. Finance and Economics Discussion Series 2007-06, Federal Reserve Board.

Beechey, M. J., Wright, J. H., 2009. The high-frequency impact of news on long-term yields and forward rates: Is it real? Journal of Monetary Economics 56, 535–44.

Bernanke, B., Gertler, M., Gilchrist, S., 1999. The financial accelerator in a quantitative business cycle framework. In: Taylor, J., Woodford, M. (Eds.), Handbook of Macroeconomics. Amsterdam: North Holland.

Caldara, D., Herbst, E., January 2019. Monetary policy, real activity, and credit spreads: Evidence from Bayesian proxy SVARs. American Economic Journal: Macroeconomics 11 (1), 157–92.

Campbell, J. Y., Cochrane, J. H., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–51.

Christiano, L. J., Eichenbaum, M., Evans, C. L., 1999. Monetary policy shocks: What have we learned and to what end? In: Taylor, J., Woodford, M. (Eds.), Handbook of Macroeconomics. Amsterdam: North Holland.

Cieslak, A., Schrimpf, A., 2019. Non-monetary news in central bank communication. Journal of International Economics 118, 293–315.

Cochrane, J. H., Piazzesi, M., 2002. The Fed and interest rates—A high-frequency identification. American Economic Review—Papers and Proceedings 92, 90–95.

Cochrane, J. H., Piazzesi, M., 2008. Decomposing the yield curve. Mimeo.

Coibion, O., Gorodnichenko, Y., Kumar, S., Pedemonte, M., 2020. Inflation expectations as a policy tool? Journal of International Economics 124, 1–27.

Diebold, F. X., Piazzesi, M., Rudebusch, G. D., 2005. Modelling bond yields in finance and macroeconomics. American Economic Review: Papers and Proceedings 95, 415–420.

Duffee, G. R., 2011. Information in (and not in) the term structure. Review of Financial Studies 29, 2895–934.

Duffee, G. R., 2012. Forecasting interest rates. In: Timmermann, A., Elliot, G. (Eds.), Handbook of Economic Forecasting. Elsevier, Amsterdam, the Netherlands.

Evans, C. L., Marshall, D. A., 1998. Monetary policy and the term structure of nominal interest rates: Evidence and theory. Carnegie-Rochester Conference Series on Public Policy 49, 53–111.

Fama, E. F., Bliss, R. R., 1987. The information in long-maturity forward rates. American Economic Review 77, 680–99.

Ferreira, L. N., 2020. Forward Guidance Matters: Disentangling monetary policy shocks. Working Papers Series 530, Central Bank of Brazil, Research Department.

Galí, J., 2015. Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications, 3rd Edition. Princeton University Press, Princeton.

Garriga, C., Kydland, F. E., Šustek, R., 2017. Mortgages and monetary policy. Review of Financial Studies 30, 3337–75.

George, E. I., McCulloch, R. E., 1993. Variable selection via Gibbs sampling. Journal of the American Statistical Association 88 (423), 881–889.

Gertler, M., Karadi, P., 2015. Monetary policy surprises, credit costs, and economic activity. American Economic Journal: Macroeconomics 7, 44–76.

Giannone, D., Lenza, M., Primiceri, G. E., May 2015. Prior Selection for Vector Autoregressions. The Review of Economics and Statistics 97 (2), 436–451.

Gilchrist, S., Zakrajšek, E., June 2012. Credit spreads and business cycle fluctuations. American Economic Review 102 (4), 1692–1720.

Gürkaynak, R., Sack, B., Swanson, E., 2005a. Do actions speak louder that words? The response of asset prices to monetary policy actions and statements. International Jousnal of Central Banking 1, 55–93.

Gürkaynak, R., Sack, B., Swanson, E., 2005b. The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. American Economic Review 95, 425–36.

Gürkaynak, R., Wright, J. H., 2012. Macroeconomics and the term structure. Journal of Economic Literature 50, 331–67.

Hamilton, J., Wu, J. C., 2012. Identification and estimation of gaussian affine term structure models. Journal of Econometrics 168, 315–31.

Hamilton, J. D., 2008. Daily monetary policy shocks and new home sales. Journal of Monetary Economics 55, 1171–1190.

Hanson, S. G., Stein, J. C., 2015. Monetary policy and long-term real rates. Journal of Monetary Economics 115, 429–48.

Hördahl, P., Remolona, E. M., Valente, G., 2015. Expectations and risk premia at 8:30AM: Macroeconomic announcements and the yield curve. Working Paper 527, Bank for International Settlements.

Ireland, P. N., 2007. Changes in the Federal Reserve's inflation target: Causes and concequences. Journal of Money, Credit, and Banking 39, 1851–82.

Jarocinski, M., Karadi, P., 2020. Deconstructing monetary policy surprises: The role of information shocks. American Economic Journal: Macroeconomics 12, 1–43.

Jordà, O., March 2005. Estimation and inference of impulse responses by local projections. American Economic Review 95 (1), 161–182.

Joslin, S., Priebsch, M., Singleton, K. J., 2014. Risk premiums in dynamic term structure models with unspanned macro factors. Journal of Finance LXIX, 1197–1233.

Joslin, S., Singleton, K. J., Zhu, H., 2011. A new perspective on gaussian dynamic term structure models. Review of Financial Studies 24, 926–970.

Kendall, M. G., 1954. A note on bias in estimation of autocorrelation. Biometrica 41, 403–404.

Kim, D. H., Orphanides, A., 2012. Term structure estimation with survey data on interest rate forecasts. Journal of Financial and Quantitative Analysis 47, 241–272.

Kuttner, K. N., 2001. Monetary policy surprises and interest rates: Evidence from the fed funds futures market. Journal of Monetary Economics 47, 523–44.

Lakdawala, A., 2019. Decomposing the effects of monetary policy using an external instruments SVAR. Journal of Applied Econometrics 34, 934–950.

Mertens, K., Ravn, M. O., 2013. The dynamic effects of personal and corporate income tax changes in the United States. American Economic Review 103, 1212–47.

Miranda-Agrippino, S., 2016. Unsurprising Shocks: Information, Premia, and the Monetary Transmission. Working Paper 626, Bank of England.

Miranda-Agrippino, S., Ricco, G., Sep. 2015. The Transmission of Monetary Policy Shocks. Discussion Papers 1711, Centre for Macroeconomics.

Nakamura, E., Steinsson, J., 2018. High-frequency identification of monetary non-neutrality: The information effect. Quarterly Journal of Economics 133, 1283–1330.

Nicholls, D. F., Pope, A. L., 1988. Bias in the estimation of multivariate autoregressions. Australian and New Zealand Journal of Statistics 30, 296–309.

Piazzesi, M., 2006. Affine term structure models. In: Ait-Sahalia, Y., Hansen, L. P. (Eds.), Handbook of Financial Econometrics. Elsevier, Amsterdam.

Pierse, R. G., Snell, A. J., 1995. Temporal aggregation and the power of tests for a unit root. Journal of Econometrics 65, 333–45.

Piffer, M., Podstawski, M., 2018. Identifying Uncertainty Shocks Using the Price of Gold. Economic Journal 128, 3266–3284.

Ramey, V. A., 2016. Macroeconomic shocks and their propagation. In: Taylor, J. B., Uhlig, H. (Eds.), Handbook of Macroeconomics. Elsevier, Amsterdam.

Ravn, M. O., Uhlig, H., 2002. On adjusting the Hodrick-Prescott filter for the frequency of observations. Review of Economics and Statistics 84, 371–76.

Romer, C. D., Romer, D. H., 2004. A new measure of monetary shocks: Derivation and implications. American Economic Review 94, 1055–1084.

Shaman, P., Stine, R. A., 1988. The bias of autoregressive coefficient estimators. Journal of the American Statistical Association 83, 842–48.

Stock, J. H., Watson, M. W., 2012. Disentangling the Channels of the 2007-09 Recession. Brookings Papers on Economic Activity 43, 81–156.

Swanson, E., forthcoming. Measuring the effects of Federal Reserve forward guidance and asset purchases on financial markets. Journal of Monetary Economics.

Wachter, J., 2006. A consumption based model of the term structure of interest rates. Journal of Financial Economics 79, 365–99.

Wu, J. C., Xia, F. D., 2016. Measuring the macroeconomic impact of monetary policy at the zero lower bound. Journal of Money, Credit and Banking 48, 253–91.

Table 1: Effect of FOMC announcements on yields across maturities

	3-month	2-year	3-year	5-year	10-year
Average response, bps	-1.1	-1	-1	-0.5	-0.3
Minimum, bps	-48	-22	-23	-16	-16
Maximum, bps	13	19	21	19	13
St. Deviation	6.8	5.9	6	5.2	4.3
Correlations					
3-month	1	0.56	0.49	0.42	0.32
2-year		1	0.92	0.93	0.86
3-year			1	0.92	0.88
5-year				1	0.92

Note: The sample is from January 1996 to October 2007.

Table 2: Business cycle properties, 10-year bond

Correlation with industrial production

Term premia			Expectations			Yields		
$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$	3M	10YR	spread
0.04	-0.33	-0.17	0.46	0.50	0.41	0.65	0.31	-0.60

Notes: The table is for the monthly model. The correlations are based on all series being filtered with the HP-filter, adjusted for monthly frequency as proposed by Ravn and Uhlig (2002). This is to remove apparent trends in the sample period. The comovement with industrial production is reported for the average correlation across the contemporaneous correlation and leads and lags up to three months, in order to take into account any potential cyclical phase shifts in the data.

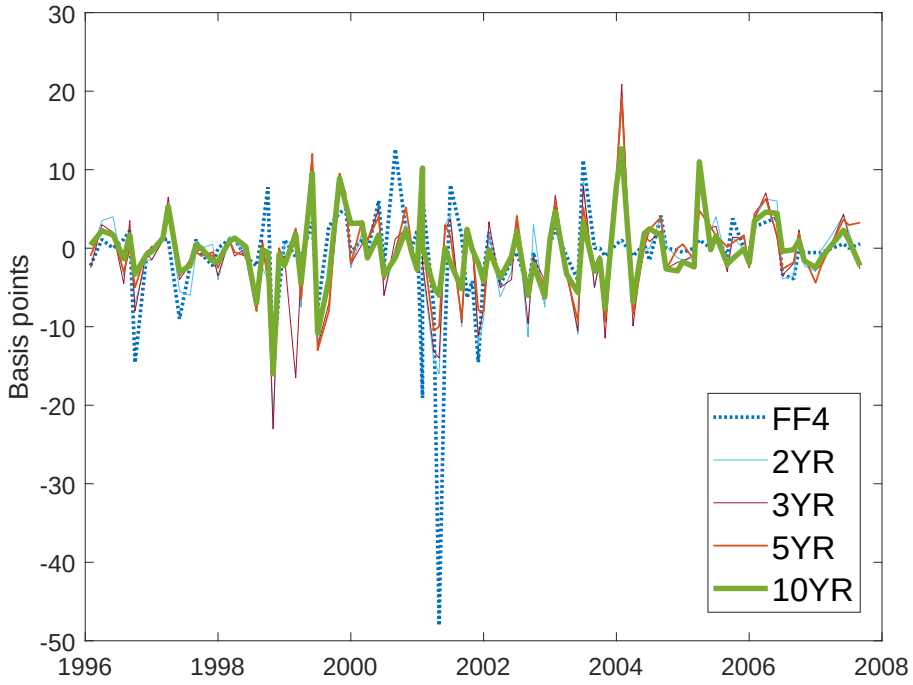

Figure 1: Yield changes around FOMC announcements across maturities.

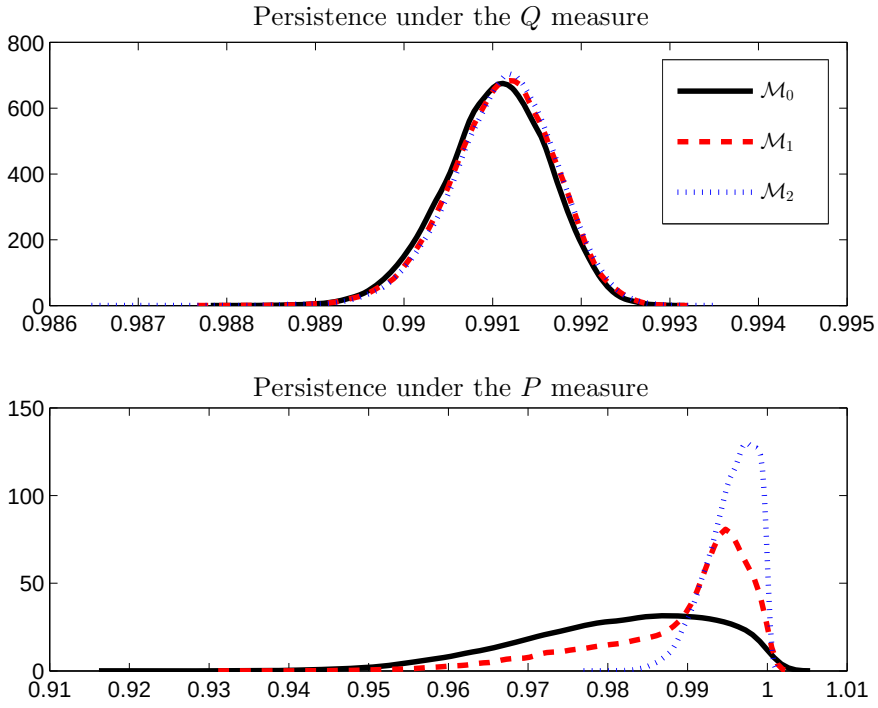

Figure 2: Monthly model: posterior distribution of persistence (largest eigenvalue).

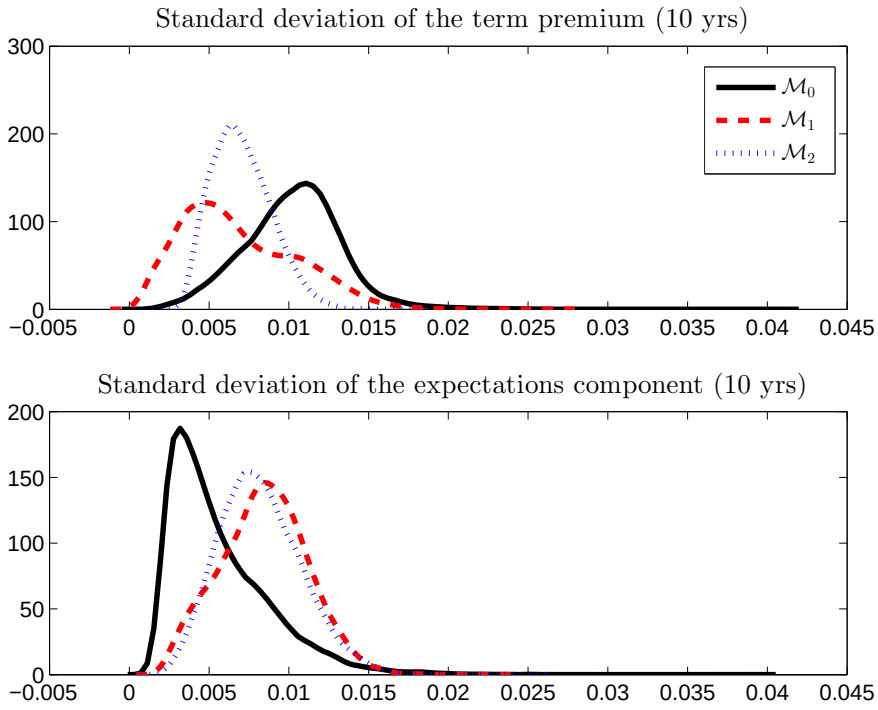

Figure 3: Monthly model: posterior distribution of unconditional volatility of the term premium and the expectations component of the 10-year yield.

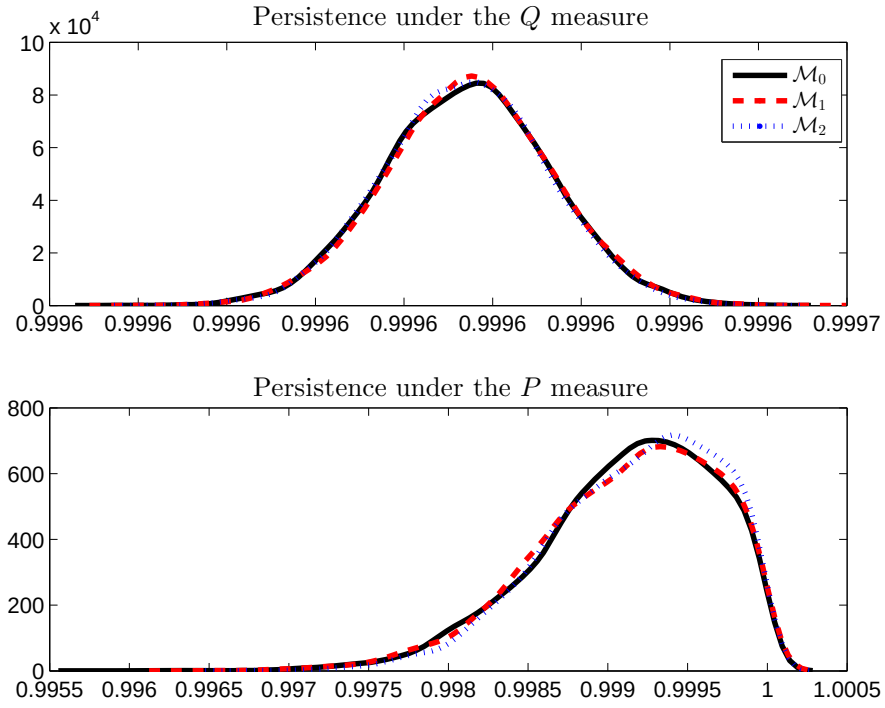

Figure 4: Daily model: posterior distribution of persistence (largest eigenvalue).

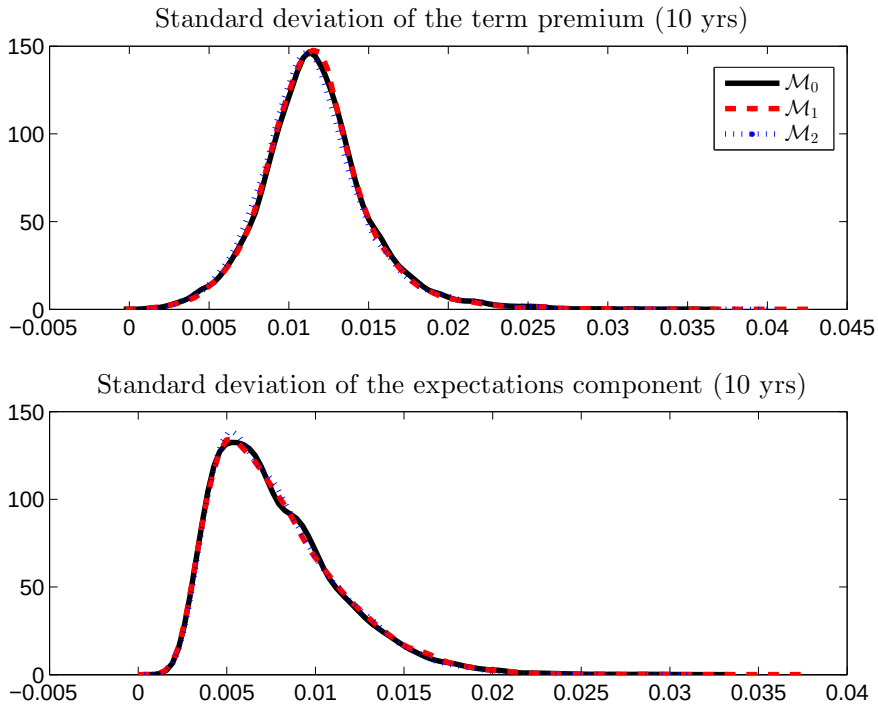

Figure 5: Daily model: posterior distribution of unconditional volatility of the term premium and the expectations component of the 10-year yield.

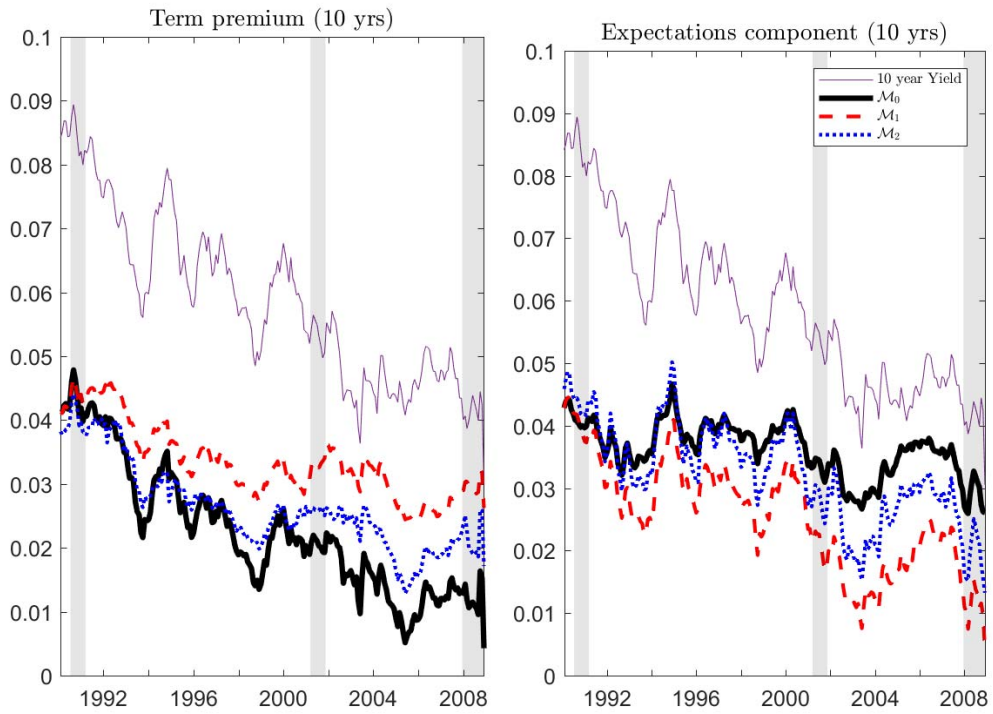

Figure 6: Monthly model: historical decomposition of the 10-year yield.

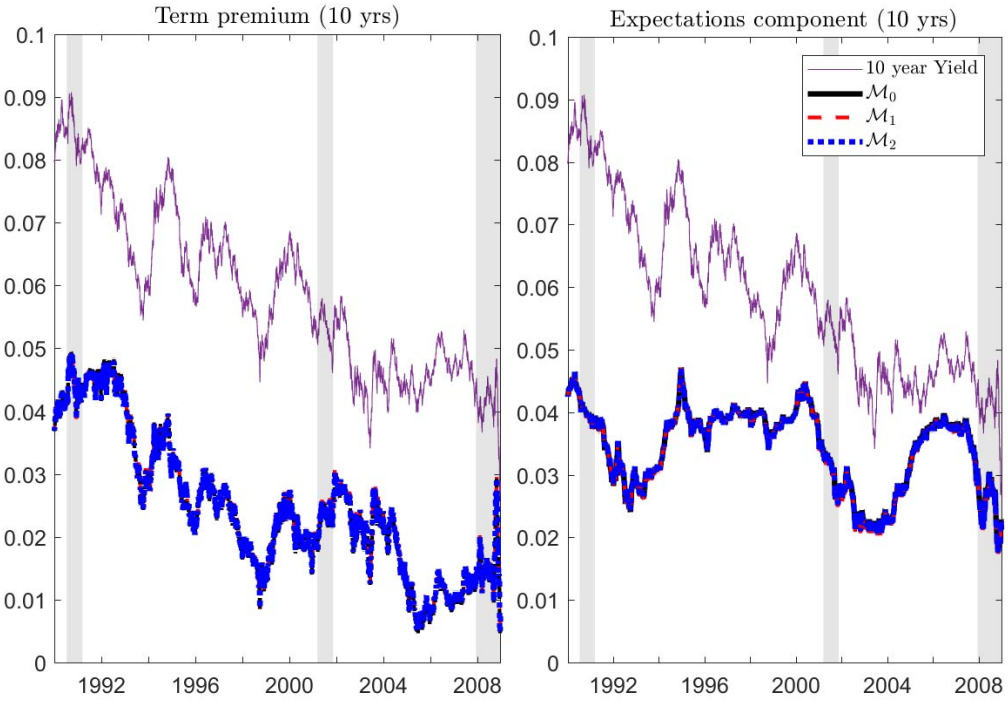

Figure 7: Daily model: historical decomposition of the 10-year yield.

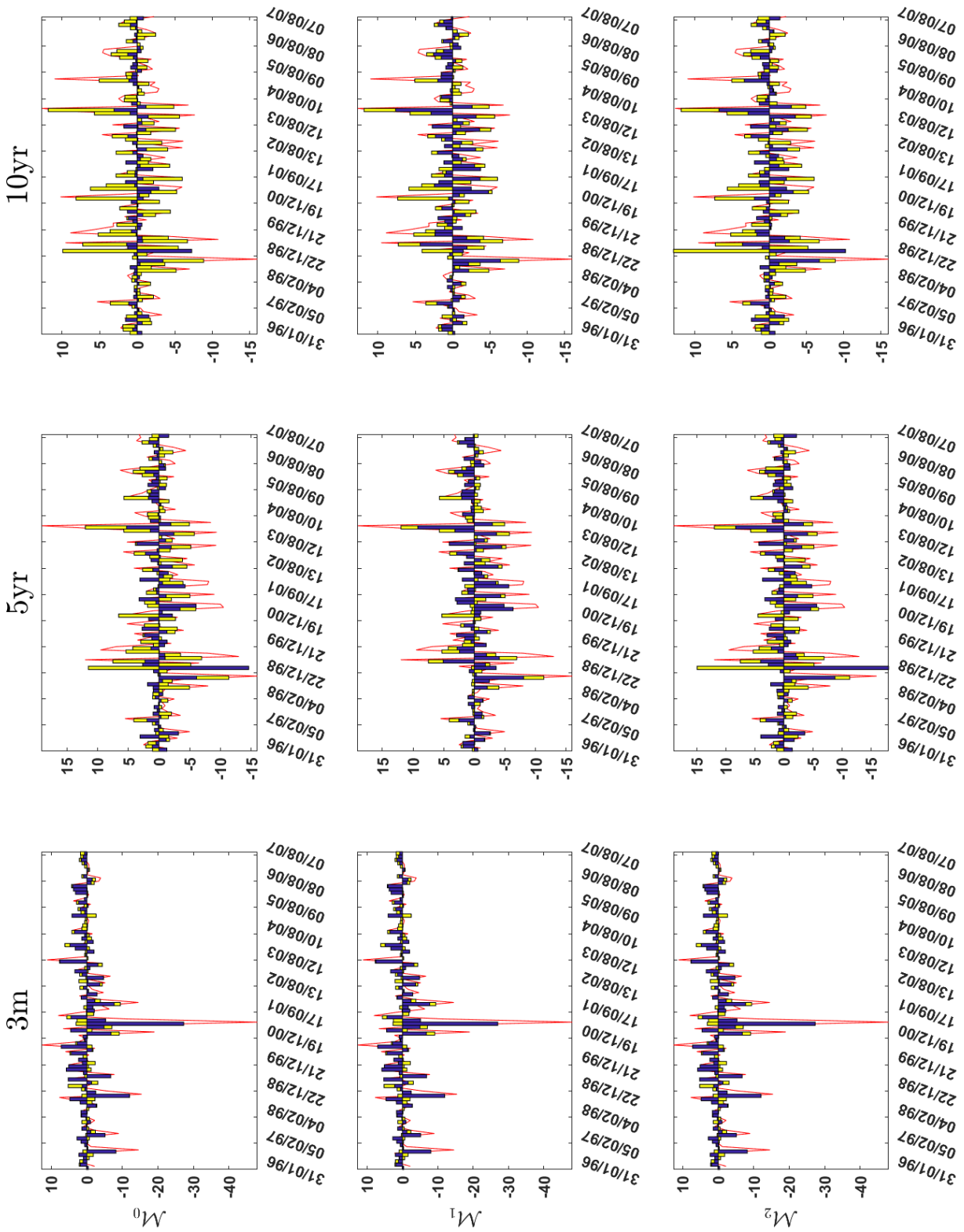

Figure 8: Decomposition of the change in intra-day yields observed before and after FOMC meetings. Based on estimated monthly ATSMs. Yellow bars are for term premia, blue bars for the expectations components. The red line depicts the observed changes in yields. The units are basis points.

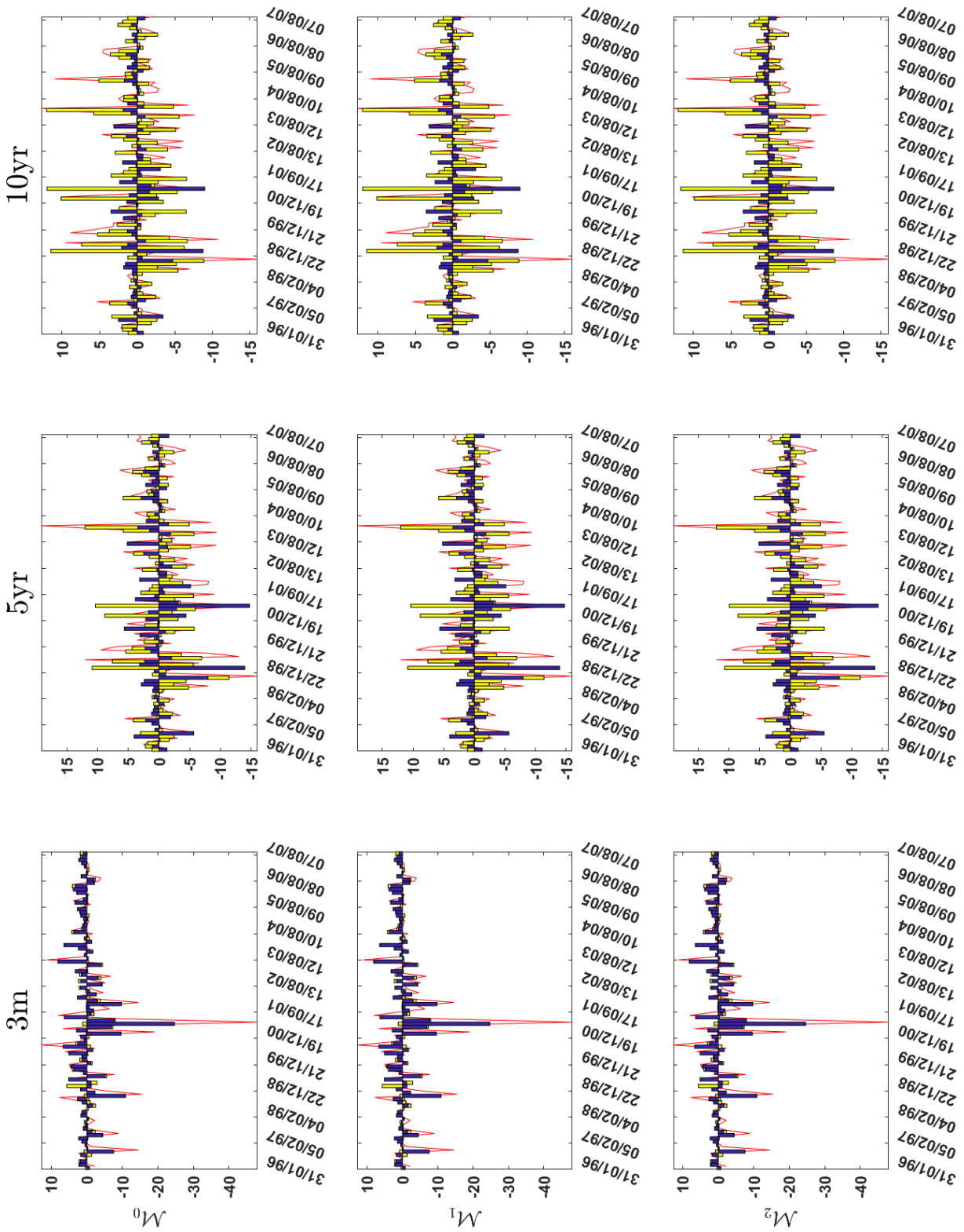

Figure 9: Decomposition of the change in intra-day yields observed before and after FOMC meetings. Based on estimated daily ATSMs. Yellow bars are for term premia, blue bars for the expectations components. The red line depicts the observed changes in yields. The units are basis points.

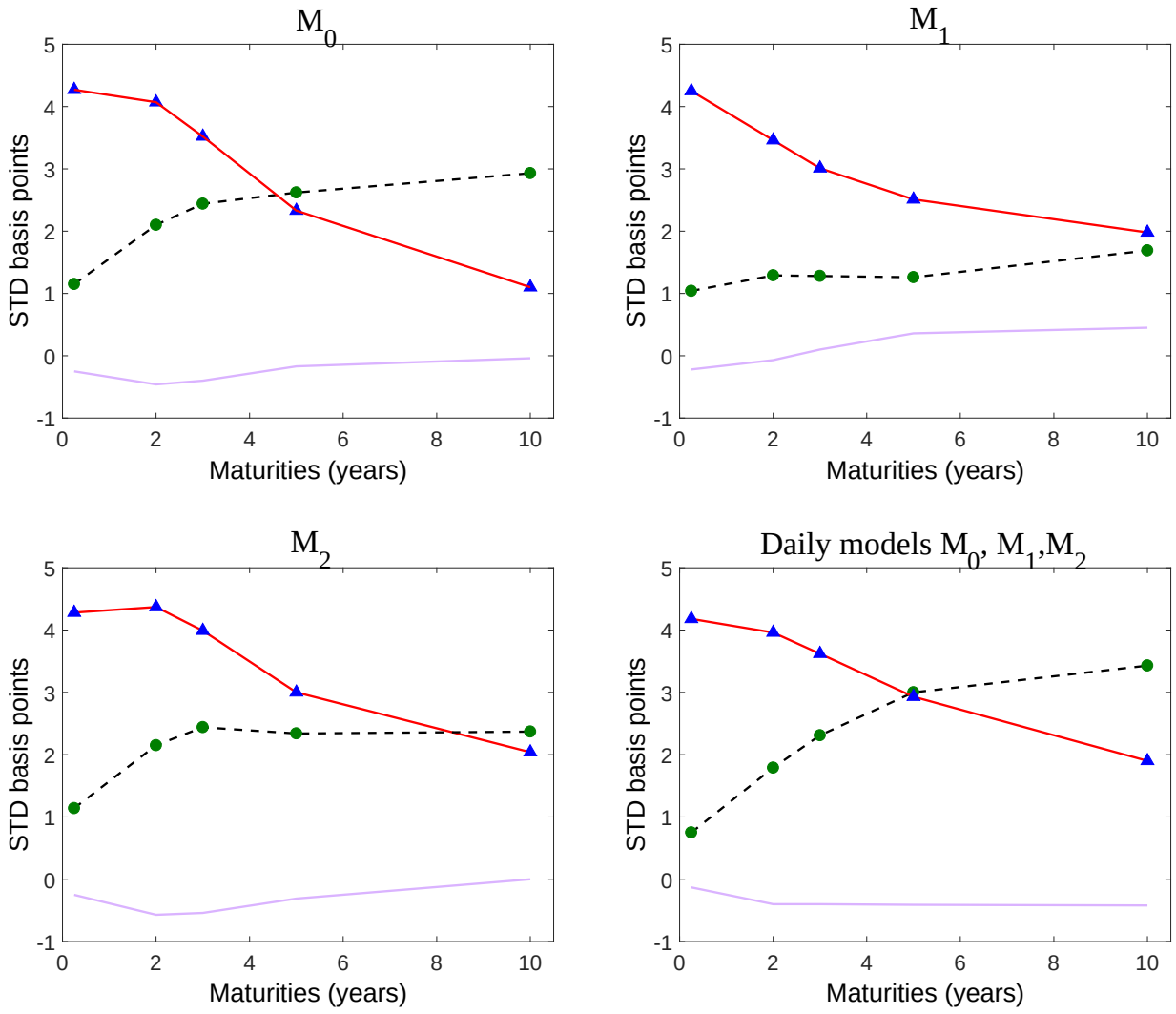

Figure 10: Volatility curve of the change in yield components around FOMC announcements. Solid line with markers: expectations. Dashed line: term premia. Solid line without markers: correlation between the two components. Markers denote the available maturities at the high frequency. The shortest maturity is three months. The results for the three daily models are almost identical. Therefore only one chart is provided.

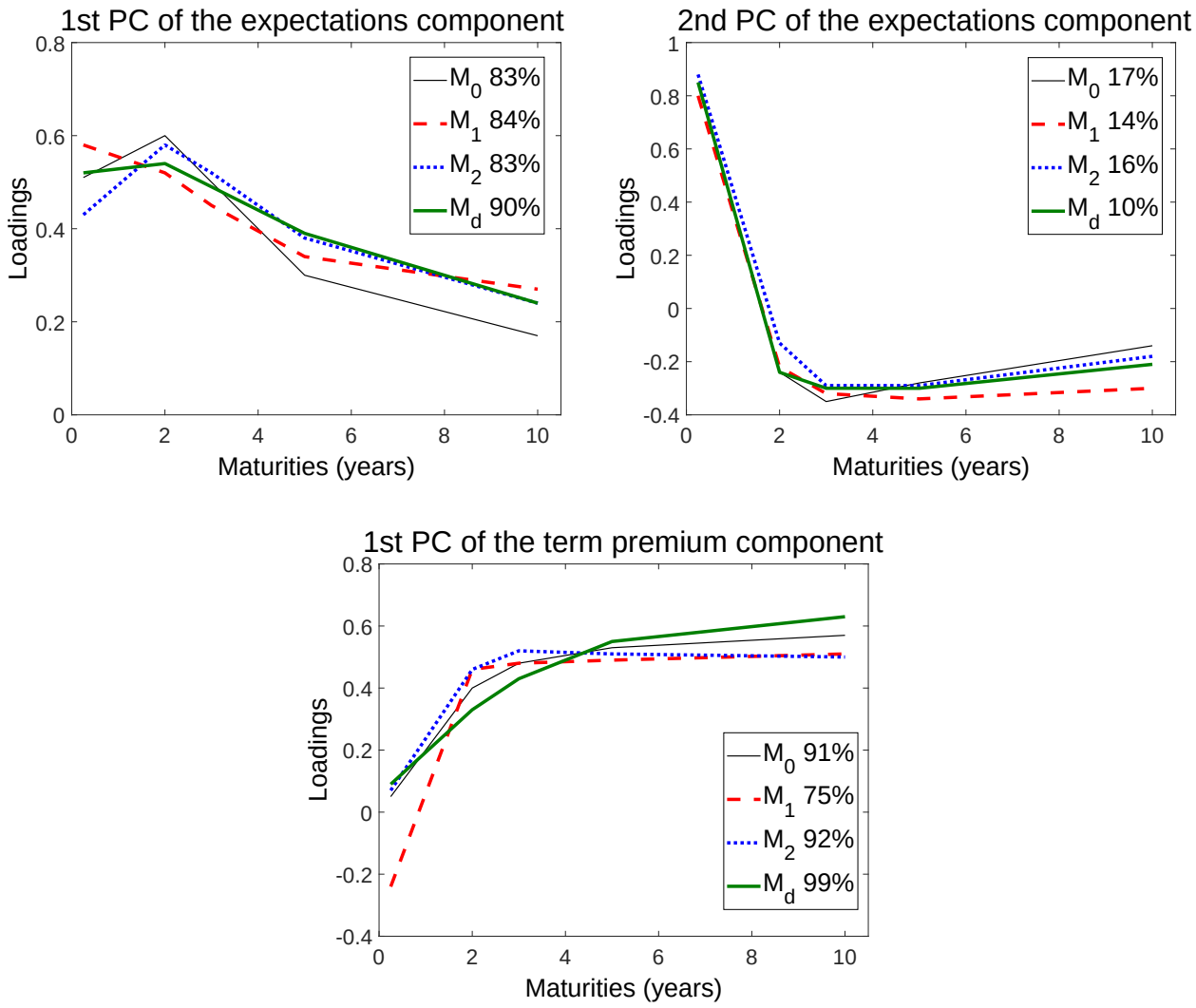

Figure 11: PC decomposition of the change in the expectations component and the term premium component around FOMC. The percentages denote the share of the variance accounted for by a given PC in each model. M_d stands for the daily model. The patterns are almost identical across the three specifications of the daily model.

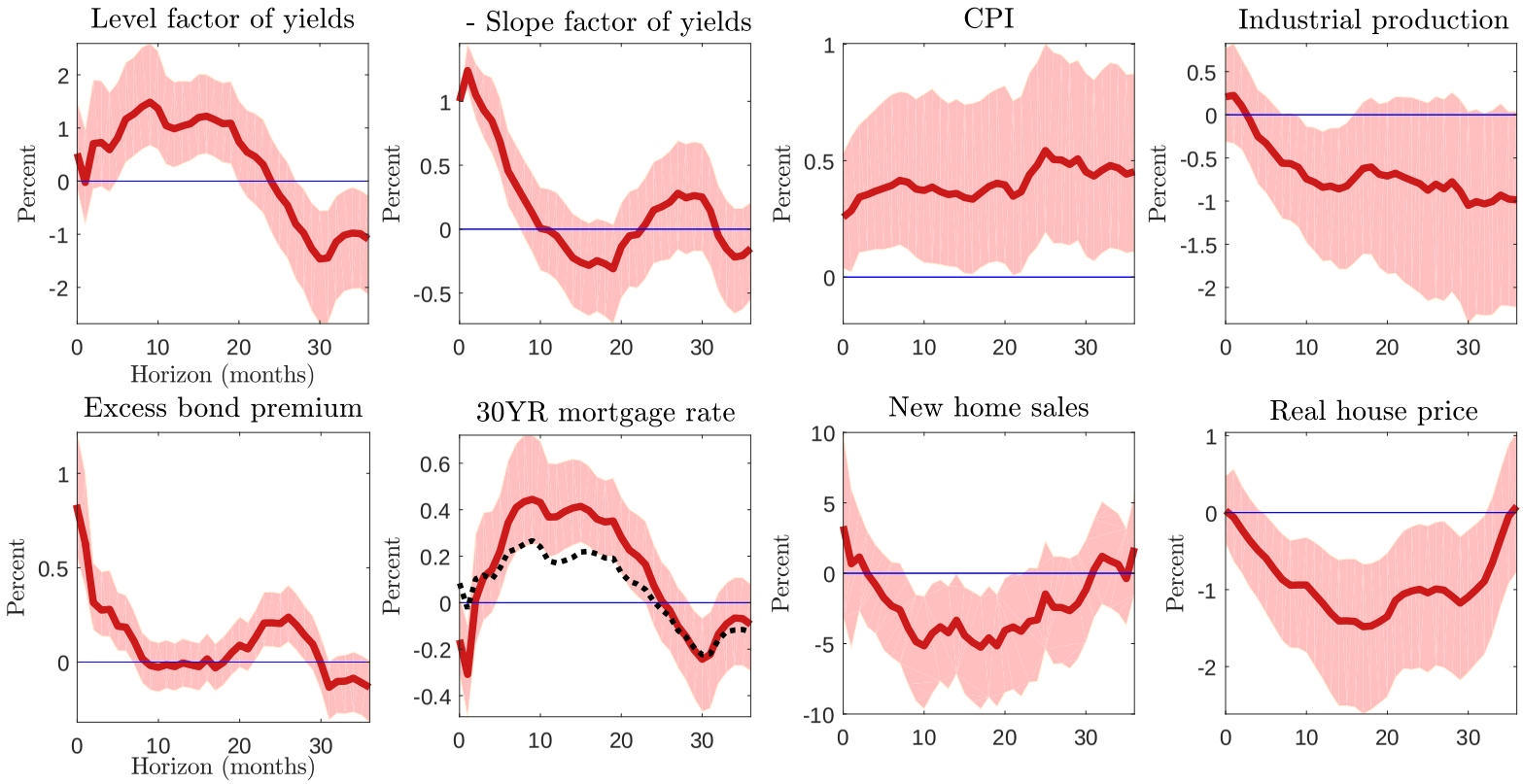

Figure 12: Local projection: 1st PC of expectations used as the instrument (m_1). The instrument is from Model $\mathcal{M}_1$. The level and slope factors are from the principal component decomposition of monthly yields data used as factors in the ATSMs and summarize 99% of the movement in the yield curve. Note that the plot for the slope factor is for the *negative* of the slope factor: an increase in the second chart of the first row means that the short rate increases and the long rate declines. The black dotted line plots the expectations component of the mortgage rate.

Figure 13: Local projection: 2nd PC of expectations used as the instrument (m_2). The instrument is from Model $\mathcal{M}_1$. The level and slope factors are from the principal component decomposition of monthly yields data used as factors in the ATSMs and summarize 99% of the movement in the yield curve. Note that the plot for the slope factor is for the *negative* of the slope factor: an increase in the second chart of the first row means that the short rate increases and the long rate declines. The black dotted line plots the expectations component of the mortgage rate.

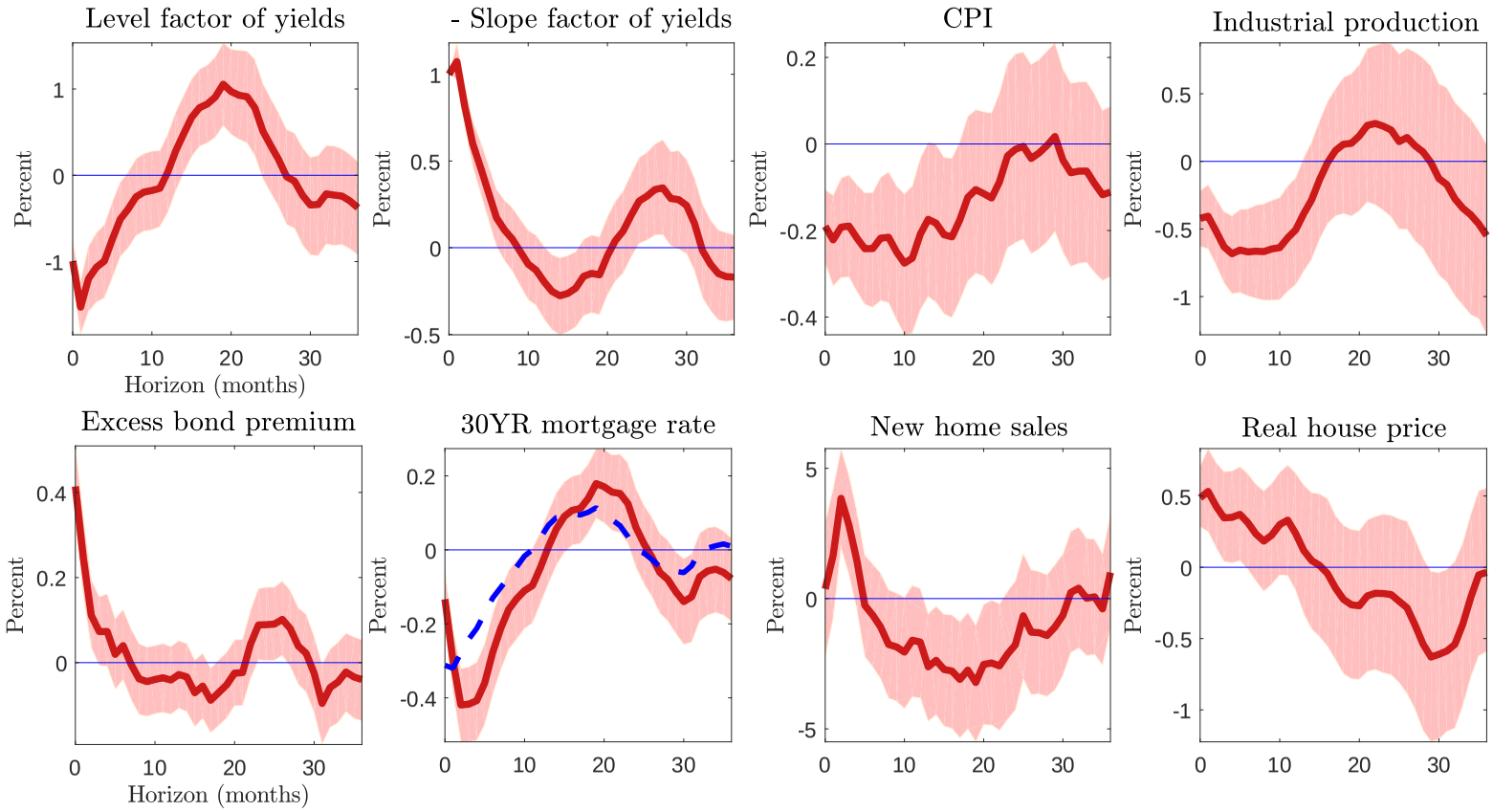

Figure 14: Local projection: 1st PC of term premia used as the instrument (m_3). The instrument is from Model $\mathcal{M}_1$. The level and slope factors are from the principal component decomposition of monthly yields data used as factors in the ATSMs and summarize 99% of the movement in the yield curve. Note that the plot for the slope factor is for the *negative* of the slope factor: an increase in the second chart of the first row means that the short rate increases and the long rate declines. The blue dashed line plots the term premium component of the mortgage rate.

Monetary policy surprises and their transmission through term premia and expected interest rates (Technical Appendix)

Iryna Kaminska* Haroon Mumtaz† Roman Sustek‡

November 11, 2020

Contents

1 MCMC algorithm

1.1 Model $\mathcal{M}_0$

The parameter vector is given by $\Theta = \left(\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma, \sigma_e^2, \mu, \Phi\right)$.

1.1.1 Priors

We specify the following priors. No prior information is assumed for any remaining parameters.

1. $P(\beta)$ is normal $N(\beta_0, V_0)$ where $\beta = vec\begin{pmatrix} \mu \\ \Phi' \end{pmatrix}$, $\beta_0 = 0$ and V_0 is a diagonal matrix with diagonal elements set to 10,000.
2. $P\left(\sigma_e^2\right) \tilde{} IG(T_0, D_0)$ where $T_0 = 0$ and $D_0 = 0$.
3. $P(b) \tilde{} N(b_0, v_0)$ where $b = \begin{pmatrix} \phi^{\mathbb{Q}} \\ k^{\mathbb{Q}} \end{pmatrix}$ with b_0 set equal to the maximum likelihood estimate and v_0 a diagonal matrix with 100 on the main diagonal. We use the Matlab code of JSZ and Bauer *et al.* (2012) to obtain the maximum likelihood estimates and their variance.

*Bank of England
†Queen Mary University of London.
‡Queen Mary University of London.

1.1.2 MCMC algorithm

The algorithm samples from the following conditional posterior distributions:

1. $g(\beta|\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma, \sigma_e^2)$. Given the factorisation of the likelihood, this conditional posterior is standard. In other words, conditional on $\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma$ and σ_e^2, the model collapses to a standard Bayesian VAR:

$$X_t = \mu + \Phi X_{t-1} + \Sigma\varepsilon_t \tag{1}$$

With a normal prior, the conditional posterior is also noromal: $N(M, V)$. The mean and variance of the conditional posterior distribution are:

$$V = \left(V_0^{-1} + (\Sigma\Sigma')^{-1} \otimes x'x\right)^{-1}$$
$$M = V\left(V_0^{-1}\beta_0 + \left((\Sigma\Sigma')^{-1} \otimes x'x\right)\beta_{ols}\right)$$

where x denotes the regressors in equation 1.

2. $g(b|\mu, \Phi, \Sigma, \sigma_e^2)$. The conditional posterior of the $\mathbb{Q}$ parameters is non-standard. Therefore, a random walk Metropolis Hastings step is used to sample these parameters. The candidate density is defined as:

$$b_{new} \tilde{} N(b_{old}, c\Psi_b) \tag{2}$$

where $b_{new}\,(b_{old})$ denotes the new (old) draw. The variance Ψ_b is obtained from an initial maximum likelihood estimation of the model with the scaling factor c chosen to achieve an acceptance rate between 20 and 40 percent. Draws are accepted with the probability:

$$\alpha = \min\left(\frac{f\left(Y_t|X_t, b_{new}, \Sigma, \sigma_e^2\right) \times P\left(b_{new}\right)}{f\left(Y_t|X_t, b_{old}, \Sigma, \sigma_e^2\right) \times P\left(b_{old}\right)}, 1\right) \tag{3}$$

where the definition of the $\mathbb{Q}$ likelihood can be found in JSZ.

3. $g(\Sigma|\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \mu, \Phi, \sigma_e^2)$. Σ appears in both terms of the likelihood factorisation. As a consequence the conditional posterior is not known in closed form. We therefore sample from this conditional posterior using a random walk Metropolis step. Let $\tilde{\Sigma}$ denote the vectorised free elements of Σ. The candidate density is:

$$\tilde{\Sigma}_{new} \tilde{} N(\tilde{\Sigma}_{new}, d\Psi_{\tilde{\Sigma}})$$

where d is a scaling factor and $\Psi_{\tilde{\Sigma}}$ is the maximum likelihood estimate of the variance of $\tilde{\Sigma}$. The acceptance probability is defined as:

$$\alpha = \min\left(\frac{f\left(Y_t|Y_{t-1}, \Theta\right)}{f\left(Y_t|Y_{t-1}, \Theta\right)}, 1\right) \tag{4}$$

where the formula for the likelihood can be found in JSZ.

4. $g(\sigma_e^2|\Sigma, \phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \mu, \Phi)$. Conditional on the remaining parameters the model collapses to the multi-variate regression $Y_t = A + BX_t + e_t$ where $var(e_t) = \sigma_e^2$. Given an Inverse Gamma (IG) prior for σ_e^2, the conditional posterior is also IG with scale parameter $\tilde{e}_t'\tilde{e}_t + \sigma_0^2$ and degrees of freedom $T\left(J - N\right) + T_0$ where $\tilde{e}_t = vec(e_t)$ and σ_0^2 and T_0 are the prior scale parameter and degrees of freedom, respectively.

1.2 Model $\mathcal{M}_1$

Following Bauer (2018), the model is re-parameterised in terms of the risk prices $\underbrace{\lambda}_{N(N+1)\times 1} = \begin{pmatrix} \lambda_0 \\ vec\left(\lambda_1\right) \end{pmatrix}$.

The parameter vector is $\Theta = \left(\phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma, \sigma_e^2, \lambda, \gamma\right)$.

1.2.1 Priors

1. $P(\gamma)$The prior for each element of γ is specified as:

$$\begin{aligned} P(\gamma_i = 1) &= \gamma_{0i} \\ P(\gamma_i = 0) &= 1 - \gamma_{0i} \end{aligned} \tag{5}$$

where $\gamma_{0,i} = 0.5$, implying that apriori there is an equal chance of ith risk price being zero or non-zero.

2. $P(\lambda)$ The prior for λ is assumed to be normal: $(1-\gamma) N\left(0, \tau_0^2\right) + \gamma N\left(0, \tau_1^2\right)$ where γ is a $N(N+1) \times 1$ vector of ones and zeros, τ_0^2 is a small number while τ_1^2 is chosen to be large number.[1] If the ith element of γ equals zero, the prior for ith risk price is tightly centered around zero. Let V_λ denote a diagonal matrix with the prior variances on the main diagonal.

The remaining priors are as described for the model $\mathcal{M}_0$

1.2.2 MCMC algorithm

1. $g\left(\gamma | \phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma, \sigma_e^2, \lambda\right)$. As shown in George and McCulloch (1993), the conditional posterior for γ is Bernoulli:

$$\begin{aligned} g(\gamma_i = 1) &= \bar{\gamma}_i \\ g(\gamma_i = 0) &= 1 - \bar{\gamma}_i \end{aligned} \tag{6}$$

where:

$$\bar{\gamma}_i = \frac{\frac{1}{\tau_{1i}^2} \exp\left(\frac{-\lambda_i^2}{2\tau_{1i}^2}\right) \gamma_{0i}}{\frac{1}{\tau_{1i}^2} \exp\left(\frac{-\lambda_i^2}{2\tau_{1i}^2}\right) \gamma_{0,i} + \frac{1}{\tau_{0i}^2} \exp\left(\frac{-\lambda_i^2}{2\tau_{0i}^2}\right) (1 - \gamma_{0i})}$$

2. $g\left(\lambda, \gamma, \phi^{\mathbb{Q}}, k^{\mathbb{Q}}, \Sigma, \sigma_e^2\right)$. The VAR under the real world measure can be written as:

$$X_t = BZ + U \tag{7}$$

where $X_t = (X_1, ..., X_t), Z = (Z_1, ..., Z_t), Z_t = (1, X_{t-1}), U = (u_1, .., u_T), \beta = vec(B)$ and $u_1 = \Sigma \varepsilon_t$. The restrictions implied by the definition of risk prices can be written as

$$\beta = \lambda + r \tag{8}$$

where $r = vec\left(\mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}}\right)$. As shown in Bauer (2018), the restricted VAR in vectorised form can be written as:

$$x - (Z' \otimes I_N) r = (Z' \otimes I_N) \lambda + u \tag{9}$$

where $x = vec(X)$. Bauer (2018) derives the conditional posterior distribution of λ and shows that the it is normal with mean and variance given by m and v where

$$\begin{aligned} v &= \left(V_\lambda^{-1} + \left(ZZ' \otimes (\Sigma\Sigma')^{-1}\right)\right)^{-1} \\ m &= v\left(V_\lambda^{-1} \lambda_0 + \hat{V}_\lambda \hat{\lambda}\right) \end{aligned} \tag{10}$$

where $\lambda_0 = 0$ and:

$$\begin{aligned} \hat{\lambda} &= \left(ZZ' \otimes (\Sigma\Sigma')^{-1}\right)^{-1} \left(\left(Z \otimes (\Sigma\Sigma')^{-1}\right) z\right) \\ \hat{V}_\lambda &= \left(ZZ' \otimes (\Sigma\Sigma')^{-1}\right)^{-1} \end{aligned} \tag{11}$$

[1] τ_0 is set equal to $\frac{1}{c}\hat{\sigma}_\lambda$ and $\tau_1 = c\hat{\sigma}_\lambda$ where $\hat{\sigma}_\lambda^2$ is an estimate of the variance of λ based on maximum likelihood estimates of the model parameters. c is set to 100.

with $z = x - (Z' \otimes I_N)\, r$.

The remaining steps of the MCMC algorithm are exactly the same as for the unrestricted model.

1.3 Model $\mathcal{M}_2$

This model is identical to model $\mathcal{M}_0$ apart from the prior on the VAR coefficients. We assume the following prior for β:

$$N\left(\beta_b, V_b\right) \tag{12}$$

where β_b denotes the bootstrap bias corrected estimate of the VAR coefficients used in Bauer *et al.* (2012). The covariance V_b is a diagonal matrix with elements on the main diagonal set to $1e-04$. Therefore, this prior reflects a very strong belief that the VAR coefficients are close to the bias corrected estimates. The steps of the MCMC algorithm are identical to the ones used for model $\mathcal{M}_0$.

				Model $\mathcal{M}_0$							
	$\mathbb{P}$						$\mathbb{Q}$				
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	0.99*	.002	-0.57*	0.01	0.007*	PC_1	1.0*	0.09*	-0.32*	0.67*	-0.0009*
PC_2	-0.004	0.97*	0.32*	-0.35*	-0.001	PC_2	-0.002*	0.96*	0.35*	-0.8*	0.002*
PC_3	0.01*	-0.01*	0.91*	0.11*	0.002	PC_3	-0.003*	0.02*	0.78*	0.98*	-0.002*
PC_4	-0.0006	0.003	-0.002	0.7*	0.0008*	PC_4	0.004*	-0.01*	-0.014*	0.41*	0.003*
				Model $\mathcal{M}_1$							
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	1.0*	0.04	-0.33*	0.66*	-0.008	PC_1	1.0*	0.09*	-0.33*	0.66*	-0.0009*
PC_2	-0.002*	0.96*	0.36*	-0.16	-0.001*	PC_2	-0.002*	0.96*	0.35*	-0.8*	0.002
PC_3	-0.004*	-0.014*	0.87*	0.02	0.0013*	PC_3	-0.003*	0.02*	0.78*	0.98*	-0.002*
PC_4	-0.008	0.003	-0.012*	0.7*	0.001*	PC_4	0.004*	-0.01*	-0.013	0.41*	0.003*
				Model $\mathcal{M}_2$							
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	1.0*	0.016*	-0.57*	0.012*	0.005*	PC_1	1.0*	0.09*	-0.33*	0.66*	-0.0009*
PC_2	-0.0013	0.98*	0.32*	-0.36*	-0.002*	PC_2	-0.002*	0.96*	0.36*	-0.8*	0.002*
PC_3	0.004*	-0.09*	0.92*	0.11*	0.0004*	PC_3	-0.003*	0.02*	0.78*	0.98*	-0.002*
PC_4	-0.0003	0.003*	-0.001	0.72*	0.0007*	PC_4	0.004*	-0.01*	-0.012*	0.41*	0.0025*

Table 1: Parameter estimates of monthly ATSMs. Posterior medians reported. * denotes that the 68-percent error band does not include 0

Model $\mathcal{M}_0$											
	$\mathbb{P}$						$\mathbb{Q}$				
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	1.0*	0.009	-0.06*	0.1*	0.0003*	PC_1	1.0*	0.02*	-0.13*	0.6*	-0.002*
PC_2	-0.0002	1.0*	0.06*	-0.17*	0.0000	PC_2	0.0034*	0.98*	0.14*	-0.62*	0.002*
PC_3	0.0001	0.003*	0.93*	0.27*	-0.0012	PC_3	-0.004*	0.02*	0.85*	0.68*	-0.002*
PC_4	0.0014	-0.002*	0.052*	0.76*	0.0002	PC_4	0.002*	-0.009*	0.06*	0.69*	0.0009*
Model $\mathcal{M}_1$											
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	1.0*	0.009	-0.06*	0.11*	0.0003*	PC_1	1.0*	0.02*	-0.13*	0.6*	-0.0016*
PC_2	-0.0002	1.0*	0.063*	-0.18*	0.0000	PC_2	0.0035*	0.98*	0.14*	-0.63*	0.0018*
PC_3	0.0001	0.0025*	0.92*	0.29*	-0.0002	PC_3	-0.004*	0.02*	0.85*	0.69*	-0.0019*
PC_4	0.00014	-0.002*	0.06*	0.75*	0.0002*	PC_4	0.002*	-0.009*	0.061*	0.69*	0.0009*
Model $\mathcal{M}_2$											
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	1.0*	0.001	-0.06*	0.1*	0.0003*	PC_1	1.0*	0.02*	-0.13*	0.6*	-0.0016*
PC_2	-0.0001	1.0*	0.06*	-0.17*	0.0000	PC_2	0.0035*	0.98*	0.14*	-0.64*	0.0018*
PC_3	0.0001	0.0025*	0.93*	0.27*	-0.0001*	PC_3	-0.004*	0.02*	0.85*	0.7*	-0.002*
PC_4	0.00014	-0.002*	0.05*	0.77*	0.0002*	PC_4	0.002*	-0.009*	0.062*	0.68*	0.0009*

Table 2: Parameter estimates of daily ATSMs. Posterior medians reported. * denotes that the 68-percent error band does not include 0

				Model $\mathcal{M}_0$							
	Monthly								Daily		
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	-0.0023	-0.086*	-0.24*	-0.66*	0.0076*	PC_1	0.0029*	-0.019*	0.071*	-0.46*	0.0019*
PC_2	-0.002	0.014	-0.03	0.45*	-0.0025*	PC_2	-0.0036*	0.015*	-0.082*	0.45*	-0.0017*
PC_3	0.0074*	-0.028*	0.13*	-0.87*	0.0025*	PC_3	0.0041*	-0.017*	0.072*	-0.41*	0.0018*
PC_4	-0.005*	0.013*	0.012	0.29*	-0.0018*	PC_4	-0.0018*	0.0069*	-0.0084	0.076*	-0.00075*
				Model $\mathcal{M}_1$							
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	2.40E-06	-0.053*	-0.00012	-0.00065	2.90E-05	PC_1	0.003*	-0.019*	0.069*	-0.45*	0.0019*
PC_2	-2.40E-06	2.90E-05	1.40E-05	0.64*	-0.0033*	PC_2	-0.0036*	0.016*	-0.077*	0.42*	-0.0017*
PC_3	0.0069*	-0.033*	0.093*	-0.96*	0.0036*	PC_3	0.0041*	-0.017*	0.063*	-0.35*	0.0017*
PC_4	-0.0052*	0.013*	3.20E-06	0.29*	-0.0016*	PC_4	-0.0018*	0.0069*	-1.10E-05	0.00013	-0.00066*
				Model $\mathcal{M}_2$							
	$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c		$PC_{1,t-1}$	$PC_{2,t-1}$	$PC_{3,t-1}$	$PC_{4,t-1}$	c
PC_1	0.0054*	-0.073*	-0.25*	-0.65*	0.0057*	PC_1	0.003*	-0.019*	0.075*	-0.47*	0.0019*
PC_2	0.00071	0.025*	-0.032*	0.44*	-0.0034*	PC_2	-0.0037*	0.016*	-0.085*	0.47*	-0.0018*
PC_3	0.0062*	-0.028*	0.14*	-0.87*	0.0027*	PC_3	0.0042*	-0.017*	0.076*	-0.43*	0.0019*
PC_4	-0.0047*	0.013*	0.012*	0.3*	-0.0018*	PC_4	-0.0018*	0.0071*	-0.01	0.085*	-0.00077*

Table 3: Estimated risk prices. Posterior medians reported. * denotes that the 68-percent error band does not include 0

2 Estimation results

Tables 1 and 2 report the posterior median estimates of the $\mathbb{P}$ and the $\mathbb{Q}$ VAR coefficients of the monthly and daily ATSMs, respectively. Table 3 presents the estimated risk prices for each model. Figure 1 shows the root mean squared estimation errors for the ATSMs. These are below 5 basis points indicating that the models fit the yield data well.

3 Robustness

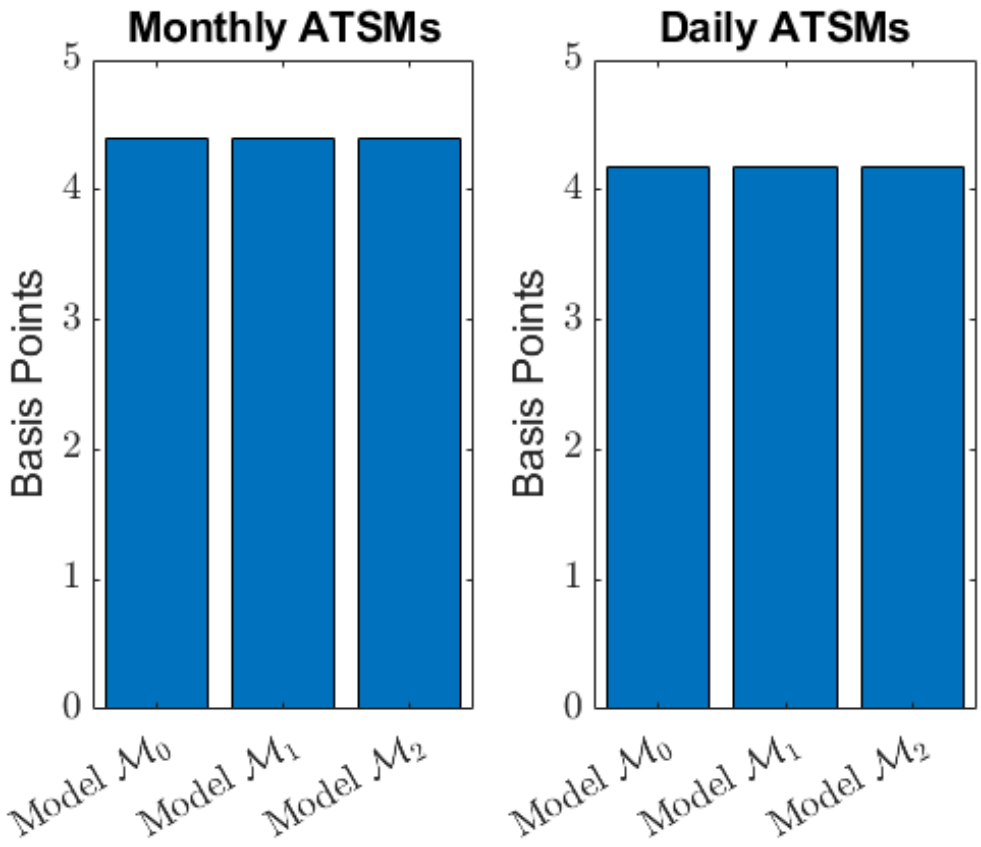

Figure 1: Root mean squared errors

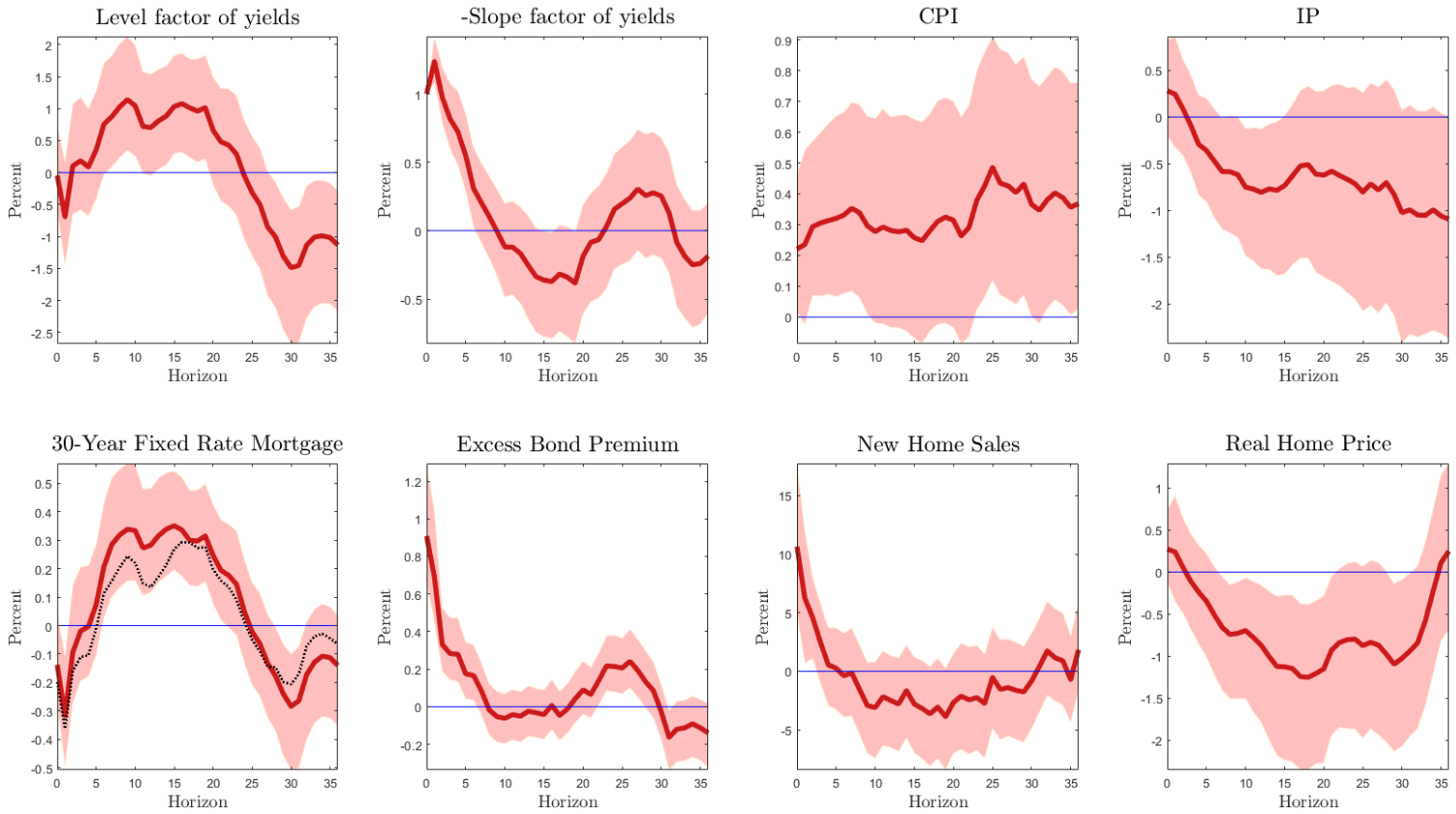

Figure 2: Impulse response using the first PC of expectations

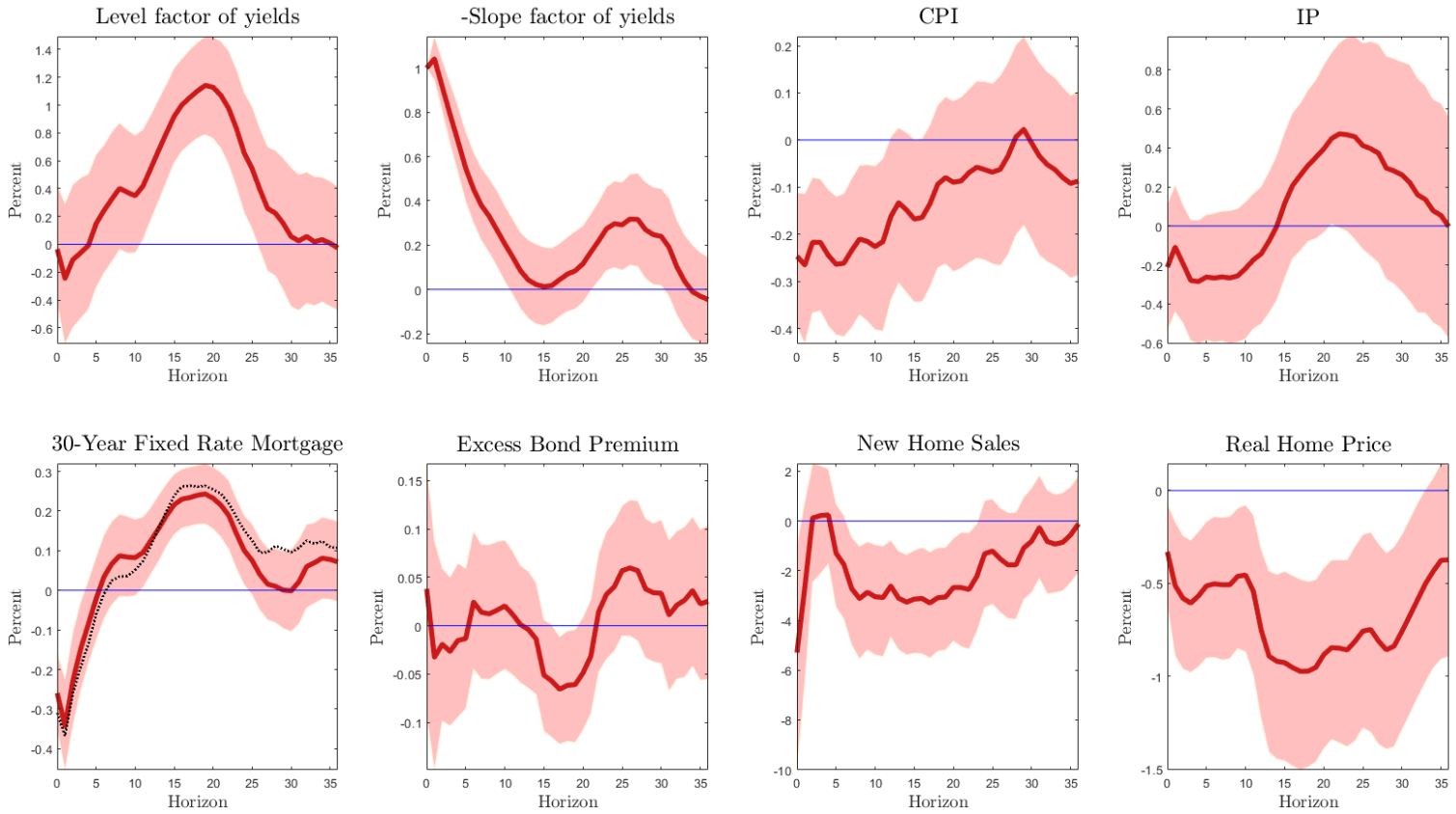

Figure 3: Impulse response using the second PC of expectations

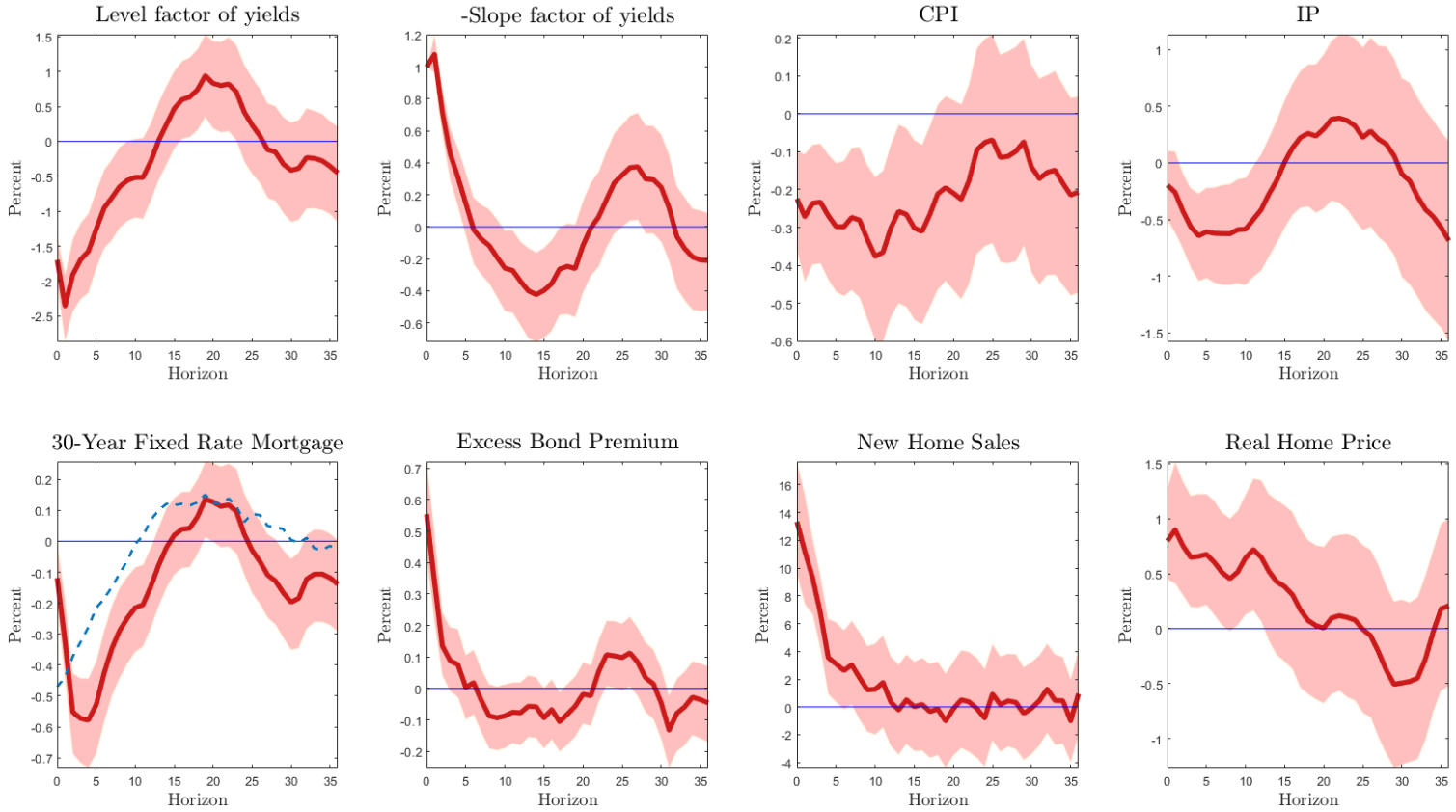

Figure 4: Impulse response using the first PC of term premia

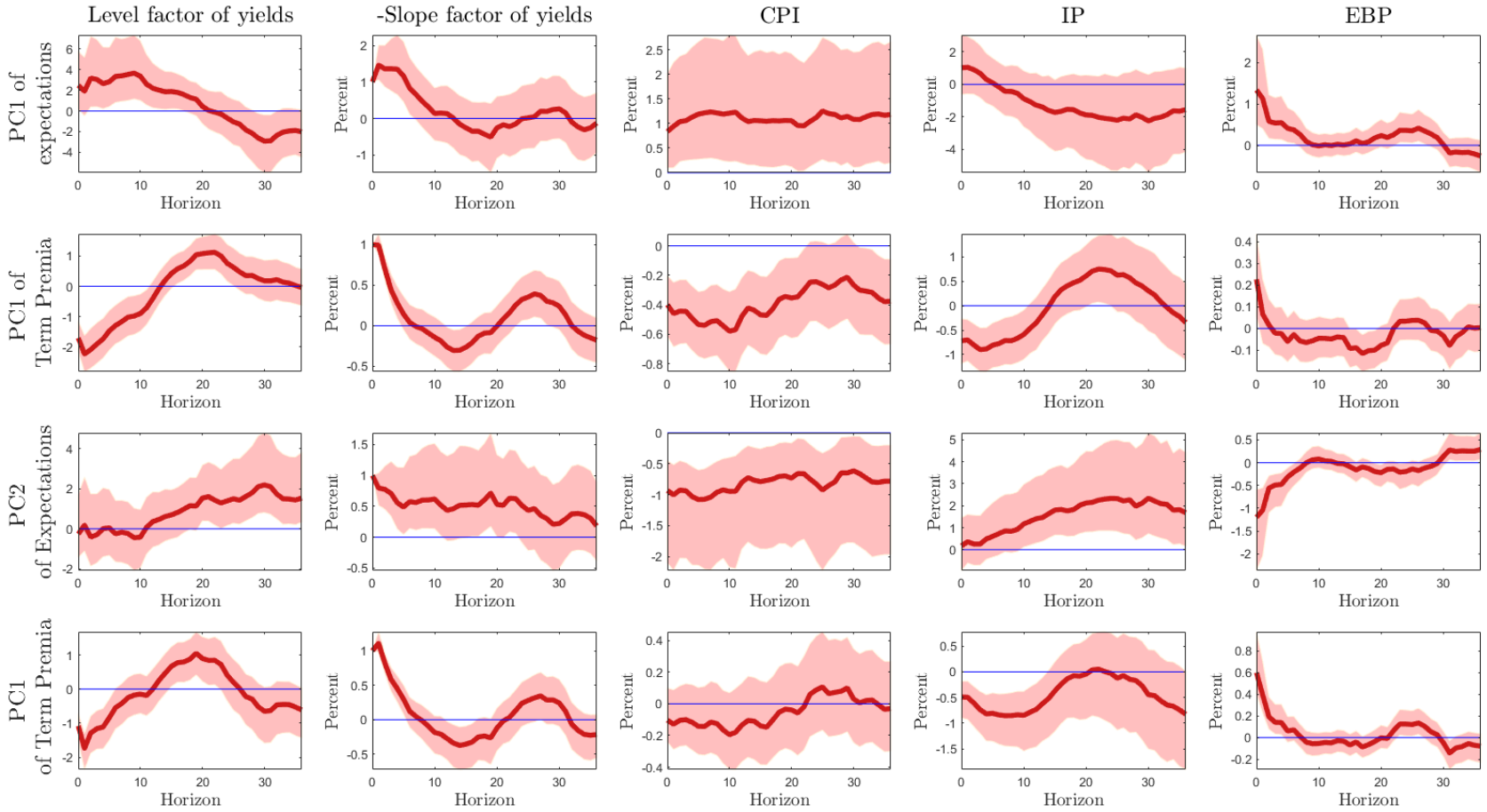

Figure 5: Identifying two shocks simultaneously in the local projections

3.1 Additional Local Projections

Figures 2, 3 and 4 presents the impulse responses estimated using the expectations and term premium components derived from model $\mathcal{M}_2$ as instruments. From a qualitative perspective, the impulse responses are similar to those obtained from the benchmark model. In particular, the shock identified using PC1 of expectations raises the level of the yield curve, CPI and the excess bond premium. There is a decline in industrial production, house prices and, in the medium run, home sales fall. In contrast, the shock identified using PC2 of expectations reduces output and CPI and the level of the yield curve does not rise immediately. Finally, when the shock is identified using the first PC of the term premium component output and CPI declines, but the excess bond premium rises, suggesting a financial accelerator type mechanism. As in the benchmark case, the mortgage rate declines.

3.2 Identifications of two shocks using two instruments

The benchmark analysis uses the instruments m_{1t}, m_{2t} and m_{3t} one at a time to identify the structural shocks. By construction, the instruments m_{1t}, m_{2t} are uncorrelated. However, as mentioned in the text, there is a correlation between m_{1t} and m_{3t} (estimated to be 0.14) and m_{2t} and m_{3t} (estimated to be -0.09). In this section, we consider identifying two shocks jointly using the instrument sets $Z_{1t} = \{m_{1t}, m_{3t}\}$ and $Z_{2t} = \{m_{2t}, m_{3t}\}$, respectively. As discussed in Mertens and Ravn (2013) if one is unwilling to assume that each of the instruments is only correlated with the shock it aims to identify, then the identifying assumptions on the two instruments are insufficient to estimate the two shocks and further restrictions are required. We follow Piffer and Podstawski (2018) and place inequality restrictions on the correlation between the instruments and the identified shocks. In particular,let z_{1t} and z_{2t} denote the two instruments in each set, and ε_{1t} and ε_{2t} the two structural shocks to be identified. We impose the following restrictions:

$$\begin{aligned} |\rho_{z_{1t},\varepsilon_{1t}}| &> 0 \\ |\rho_{z_{2t},\varepsilon_{2t}}| &> 0 \\ |\rho_{z_{1t},\varepsilon_{1t}}| - |\rho_{z_{1t},\varepsilon_{2t}}| &> \varpi \\ |\rho_{z_{2t},\varepsilon_{2t}}| - |\rho_{z_{2t},\varepsilon_{1t}}| &> \varpi \end{aligned}$$

where $\rho_{x,y}$ denotes the correlation coefficient between x and y. The first two restrictions impose the conditions that the instruments are correlated with the shocks they identify. The third and the fourth condition imposes the restriction that, in absolute terms, each instrument is more correlated with the shock it identifies than the other shock. Following Piffer and Podstawski (2018), the threshold value ϖ is set to 0.1.

The top two rows of Figure 5 plots the response to the two shocks identified by the instrument set $Z_{1t} = \{m_{1t}, m_{3t}\}$. The estimates in the top row are very similar to benchmark results for m_{1t}. The shock increase the level factor, CPI and EBP and decreases output in the medium term. The second shock, which is now identified simultaneously, results in impulse responses close to the benchmark responses obtained using m_{3t}. This shock reduces the level of yields, CPI and output while EBP rises. The third and fourth rows of the figure present results from the scheme where the shocks are identified using Z_{2t}. The fourth row shows that the responses to the shock identified using the first PC of term-premia is very similar to the benchmark. As in the benchmark case, the response of EBP to the shock identified via m_{2t} is negative (see third row of the figure) and CPI declines. However, the initial negative effect on output and the level of yields is less pronounced than benchmark.

Overall, the results based on this set identification scheme support the benchmark results from a qualitative perspective.

4 Convergence

Figures 6 and 7 present inefficiency factors for the monthly and daily ATSMs. These are fairly low for most parameters indicating convergence of the MCMC algorithms.

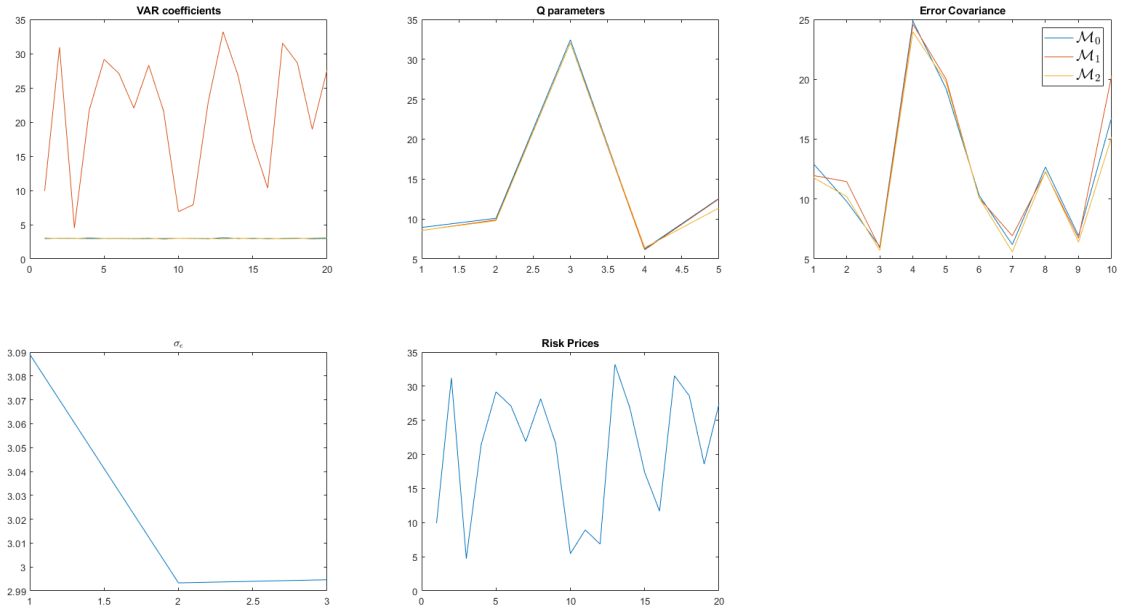

Figure 6: Inefficiency factors monthly model

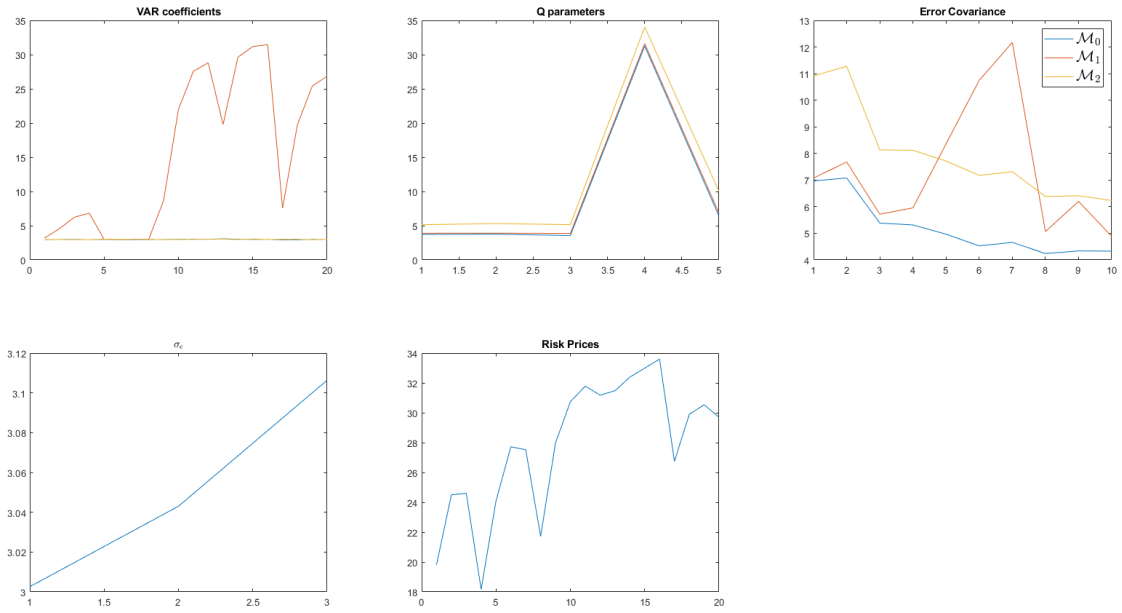

Figure 7: Inefficiency factors for the daily model

References

Bauer, Michael D., 2018, Restrictions on Risk Prices in Dynamic Term Structure Models, *Journal of Business & Economic Statistics* **36**(2), 196–211.

Bauer, Michael D., Glenn D. Rudebusch and Jing Cynthia Wu, 2012, Correcting Estimation Bias in Dynamic Term Structure Models, *Journal of Business & Economic Statistics* **30**(3), 454–467.

George, Edward I. and Robert E. McCulloch, 1993, Variable Selection Via Gibbs Sampling, *Journal of the American Statistical Association* **88**(423), 881–889.

Mertens, Karel and Morten O. Ravn, 2013, The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States, *American Economic Review* **103**(4), 1212–47.

Piffer, Michele and Maximilian Podstawski, 2018, Identifying Uncertainty Shocks Using the Price of Gold, *Economic Journal* **128**(616), 3266–3284.

School of Economics and Finance

Queen Mary University of London

School of Economics and Finance Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/