Long Run Impact of Macro News on Treasury Bond Yields

Bruno Feunou   Jean-Sébastien Fontaine
Bank of Canada

Guillaume Roussellet
McGill University

October 28, 2019

Abstract

Macro news have large impact on bond yields in high-frequency data. We aggregate the impact of macro news within each month, which we use in a no-arbitrage term structure models. We find that macro news explain 50 percent in the term premium of the 10-year bond at the monthly frequency and 40 percent at longer horizons. By contrast, macro news explain less than 10 percent of variations in the expectation component of the 10-year bond at monthly or longer horizons. The impact on the expectation component of yields is surprisingly low for all maturities and robust across a range of models, suggesting that investors mostly revise expectations based on information outside of data releases. Overall, macro new explains between 25 and 30 percent of yield variances at monthly or quarterly horizons, consistent with event studies, but this share declines to as low as 20 percent for longer horizons, for all maturities.

Keywords: Term Structure, Macro-Finance, Variance Decomposition
JEL codes: C32, E43, G12

– Bank of Canada, 234 Wellington St, Ottawa, On k1a 0h9, Canada, jsfontaine@bank-banque-canada.ca, BFeunou@bank-banque-canada.ca.
– Desautels School of Management, 1001 Sherbrooke St W., Montreal, Qc h3a 1g5, Canada, guillaume.roussellet@mcgill.ca.

1 Introduction

The linkages between macroeconomic news and long-term interest rates are essential for investors to understand the risk that they bear when holding securities. Based on the fact that the time of data releases is known in advance, the impact of macro news is easy to see in high-frequency data, producing a large impact on bond yields, but also on volatility and trading volume (Ederington and Lee, 1993; Fleming and Remolona, 1999; Gürkaynak et al., 2005). Notwithstanding the rich set of results from these event studies, not much is known about the persistence and the propagation of macro news over long horizons.

We contribute to the debate around the strength of linkages between macro news and bond yields. Taking a new look at an old question, we ask how much of yield variances at business-cycle horizons can be attributed to macroeconomic news. Consistent with results from event studies, we find that macro news explains around 30-35 percent of yield at monthly or quarterly frequencies. However, and perhaps counterintuitively, we find that the propagation of macro news explains a smaller share of yield variances at business cycle frequencies, as low as 15 to 25 percent, depending on the model specifications.

Our first contribution is methodological. We embed macro news within a dynamic term structure model (DTSM) using the fact that monthly changes in interest rates sum two components. The first component, a news component, aggregates changes identified from high-frequency data around macroeconomic data releases within each month. The second component aggregates the remaining changes. Based on this identity, we construct a family DTSMs with a distinct role for each component.

This family of DTSMs inherits from benchmark models the idea that a small number of pricing factors can fit observed interest rates. In a way, these factors embed and summarize how past and current macroeconomic information in current yields with different maturities (Bikbov and Chernov, 2010). The latent factors can be rotated in different ways with the main purpose of correctly pricing bond yields across a range of maturities (Dai and Singleton, 2000). For instance, the pricing factors can be three principal components of yields. Our approach naturally nests the benchmark yield-only VAR(1) DTSM when all news components are simply summed together and only the aggregated monthly change matters. One novelty of our model is that macro news and non-macro news have different dynamic implications for the pricing factors.

One defining outcome is that we can ask what share of the factor variances can be attributed to macro news.

Our second contribution is empirical. We estimate the benchmark case as well as more general cases that provide different roles for each component. Given the model estimates, we turn to the variance decomposition of yields. At monthly and quarterly horizons, the macro news component explains around 25-35 percent of the variance of yields, depending on the maturity. These are comparable and consistent with results in Altavilla, Giannone, and Modugno (2017) based on a regression framework. They show that the effect of macro news is more persistent than other shocks to yields and that the macro share of variance can increase between daily and quarterly horizons. One open question is whether this intuition generalizes at longer horizons. We find that it does not. In every other model that we consider, we find that the macro share stabilizes below 25 percent. Our preferred model shows a decline to around 15 percent.

We perform a variance decomposition of the expectation and term premium components of yields, which is not easily available in event studies. In our preferred model, we find that the share of the 10-year term premium variance due to macro news range approaches 50 percent at short horizons and declines gradually at long horizons. By contrast, the share the expectation component of yields is below 15 percent for all the models that we consider. Therefore the declining pattern that we observe in the share of yield variance to macro news is due to the term premium. However, the low overall share due to macro news is due to the expectation component.

Unsurprisingly, the largest revisions in expectations occurred when the Fed funds rate was rising (1999-2000 and 2004-2006) or when it was decreasing (2001-2002 and 2007-2008). Yet, the impact of macro news on the forecast errors was relatively low precisely in these periods. How can expectations of monetary policy respond so little to macro news? One reason could be that other sources of information, such as speeches, play a determinant role. Alternatively, it could be that the information about the economy in bond yields arises due to trading by expert or well-informed investors revealing information outside of the data release windows. There are several reasons how this could happen. For instance, Cieslak et al. (2019) suggest systematic informal communication of Fed officials with the media and financial sector as a channel through which news about monetary policy has reached the market. Alternatively, the Federal Reserve was also perceived to react to equity market valuations because of the information it contains about future economic data (see e.g., Cecchetti 2003 for a review). We leave this for future research.

Related Literature

Our contribution substantially broadens an important point made by Duffee (2018) in the case of inflation news. Using survey data, he finds that news about expected inflation accounts for between 10 to 20 percent of yield variances at the quarterly frequency, which is very low relative to the prediction of existing equilibrium asset pricing models. We extend the results by using all data releases and we find estimates that are barely higher. In addition, we can interpret the results based on the impact that macro news have on expectations and the term premium. Surprisingly, the impact on the expectation component is small. We find that the impact of macro news on long-term yields acts mostly via the term premium.[1]

An important branch of the literature studies linkages between yields and the economy using macro-finance DTSMs.[2] We establish the mapping between our approach and a generic macro-finance term structure model à la Ang and Piazzesi (2003). Intuitively, our framework corresponds to filtering the long-run impact of macro news based observing the evolution of yields and the macro news. We show that the two approaches are equivalent in the framework of Ang and Piazzesi (2003). This means that we do no have to take a stand on the nature and the number of macro variables to model.

Existing results based on macro-finance models often attribute a large share of yield variances to real activity or inflation. In this context, the identification of macro shocks requires additional assumptions, typically imposing that there is no contemporaneous feedback from yield shocks to macro variables. For instance, with monthly data, this assumes that there is no within-month response of unemployment or real activity to yield shocks. In addition, this relies on model-implied forecasts to

[1]Duffee (2018) concluded that the magnitudes of shocks to (short) real rates and expected excess returns cannot be determined reliably based on his results. Nonetheless, he suggested that point estimates of the persistence of real rate shocks suggest a larger role for the risk premium in the variances of long-term yields.

[2]Leading examples of macro-finance term structure models include Ang, Piazzesi, and Wei (2006), Diebold, Rudebusch, and Aruoba (2006), Rudebusch and Wu (2008) and Bikbov and Chernov (2010).

filter macro shocks, which is arguably less robust.[3]

More recently, Moench (2012) uses sign restrictions and introduces a factor model to extract information from a broad range of macro variables. There is one similarity in that we consider all available macroeconomic data releases. The main difference is that we use high-frequency data and we do not need to model the macro factors. Doshi, Jacobs, and Liu (2018) allow for short-term and long-term component in the macro dynamics but do not consider disentangling yield shocks from macro shocks.

The overall message of our paper is robust to the specification of the historical dynamics. Yet, two features of our preferred model also stand out. First, the restriction that the news components across several maturities can be summarized with one linear combination is strongly supported in the data. This is consistent with Bauer (2015) who finds that the effect of macroeconomic news on yields of different maturities can be summarized with one factor. Gurkaynak, Kisacikoglu, and Wright (2018) also find a strong common response of interest rates to data releases. Second, the data support a moving-average component as way to have long lags of the yield factors entering the dynamics, consistent with previous results in Feunou and Fontaine (2018).

In an important contribution, Rudebusch (1998) discusses the differences between monetary policy shocks identified in VAR models and shocks identified using high-frequency changes in the shortest futures rates (see also Kuttner 2001). More recently, the response of interest rates at longer horizons has been used to reveal monetary policy shocks, growth shocks and risk premia shocks (e.g., Gurkaynak et al. 2007; Cieslak and Schrimpf 2018). We build on this approach and extract the yield responses to data releases from high-frequency data. One important difference is that we embed this information in a dynamic term structure model.

The rest of the paper is organized as follows. Section 2 details the computation of the monthly news component of yields and establishes simple stylized facts. Section 3 introduces the news components in a general family of dynamic term structure models and discusses the connections with existing macro-finance models. Section 4 explains the estimation methodology and reports all the results. Section 5 concludes.

[3] Table 8 in Ang and Piazzesi (2003) reports that for the 5-year yield, the share of variance explained by macro variables range around 38-63 percent across different horizons in a VAR model with one lag of macro, but around 2-11 percent in a model with several lags.

2 Preliminary Evidence

2.1 Macro News and Bond Returns

We start with some notation. Time is monthly. The excess returns (in log) to holding a nominal zero-coupon bond with m months to maturity is given by:

$$ex_{t+1}^{(m)} = my_t^{(m)} - (m-1)y_{t+1}^{(m-1)} - i_t, \tag{1}$$

where $y_t^{(m)}$ is the yield to maturity (per month, continuously compounded) and $i_t = y_t^{(1)}$ is the short rate. Iterating over this definition, we get the following identity, in the spirit of Campbell and Amner (1993) and Duffee (2018),

$$y_t^{(m)} = \sum_{i=1}^{m} i_{t+i-1} + \sum_{i=1}^{m} ex_{t+i}^{(m-i+1)}. \tag{2}$$

This identify follows from the definition of excess returns. It states that the returns to holding a bond until its maturity is the sum of short rates and excess returns over the holding period. Keeping the yield to maturity constant, raising the path of future short rates necessarily implies lower future excess return. Given the information available at time-t, taking expectations on both sides of (2),

$$y_t^{(m)} = eh_t^{(m)} + rp_t^{(m)} \tag{3}$$

$$eh_t^{(m)} = E_t\left[\frac{1}{m}\sum_{i=1}^{m} i_{t+i-1}\right] \qquad rp_t^{(m)} = E_t\left[\frac{1}{m}\sum_{i=1}^{m} E_t[ex_{t+i}^{(m-i+1)}\right],$$

where the two terms on the right (3) are commonly described as the expectation and term premium component of yields, respectively. Based on this accounting identity, the forecast error $\tilde{y}_t^{(m)} = y_t^{(m)} - E_{t-1}y_t^{(m)}$ is given by

$$\tilde{y}_t^{(m)} = \Delta eh_t^{(m)} + \Delta rp_t^{(m)} \tag{4}$$

$$\Delta eh_t^{(m)} = \big(E_t - E_{t-1}\big)\left[\frac{1}{m}\sum_{i=1}^{m} i_{t+i-1}\right] \qquad \Delta rp_t^{(m)} = \big(E_t - E_{t-1}\big)\left[\frac{1}{m}\sum_{i=1}^{m} ex_{t+i}^{(m-i+1)}\right],$$

using the $\big(E_t - E_{t-1}\big)$ operator. The news between t and $t+1$ enter the forecast error

$\tilde{y}_t^{(m)}$ either because of a revision in the expectation component $\Delta eh_t^{(m)}$ or a revision of the risk premium component $\Delta rp_t^{(m)}$.

Our paper starts from the premise that the revisions of the expectation component and the term premium can be due to macro news. The goal is to take the common decomposition 4 one step further, separating the impact of macro news released with economic data from the impact of other news that come to the market over the course of a month. Therefore,

$$\tilde{y}_t^{(m)} = \Delta eh_{n,t}^{(m)} + \Delta eh_{y,t}^{(m)} + \Delta rp_{n,t}^{(m)} + \Delta rp_{y,t}^{(m)}, \tag{5}$$

where the subscript n indicates macro news and y indicates the impact of other news that affect bond yields.

Implementing the decomposition 5 presents two challenges. First, we have to identify separately the impact of macro news on monthly yields from the impact of other news. In the following section, we address this challenge using high-frequency data around the release of economic data. The second challenge is to decompose the impact on yields between the impact on the expectation and the risk premium components. This will be the focus of Sections 3-4, where we embed the information from high-frequency data in DTSMs.

2.2 Measuring Macro News in High-Frequency Data

We use high-frequency CME data for U.S. Treasury futures contracts to measure the impact of macro news.[4] Our sample covers the period between 1995 and 2016. We select the most liquid contracts, i.e. those for delivery of Treasury securities with 2 years, 5 years and 10 years to maturity. We denote their yields at the end of month t by $\mathcal{F}_t$. We extract the date and time of each release from the MMS data, supplemented with information in Kilian and Vega (2011) and from Bloomberg. For monetary policy releases, we obtain data from the FOMC and Piazzesi (2005). For each contract and for each data release, we compute the change in futures-implied

[4]There are several reasons for the choice of CME data. It is available at high-frequency over a longer sample, which allows us to measure the change of futures rate during tight window around the release of new macro data. Second, using high-frequency Treasury coupon bond data introduces introduce an unnecessary degree of complexity because of the variations in the coupon rate and maturity of the underlying bonds over time. Futures data can be easily converted to a bond-equivalent yields with fixed coupon and maturity.

bond yields.

In the following, we aggregate these macro news at the monthly frequency. To illustrate the principle underlying our approach, we present the evolution of the 5-year futures yields in a given month. Panel (a) of Figure 1 reports the evolution for the month of July 2004. Blue parts represent variations on days where news were released. July 2004 is a typical month where macro news plays a large role. On July 2nd, the release of the Non-Farm Payroll data surprised the market: payrolls had increased by 112 thousands against the median survey forecast of 250 thousands (month-over-month). All yields fell dramatically. The 5-year yield declined by 18 bps around the release. In the rest of the month, other release added 5 bps to the decrease. Some of this decline was negated by an increase of close to 3 bps between data releases.

Panel (b) of Figure 1 repeats this exercise for September 1998, which is an example where macro news played only a small role in the aftermath of LTCM's financial difficulties. Yields fell dramatically through the month, but on dates without data releases. The largest fall occurred on September 11, 1998, with four other large falls during the month. Overall, the 5-year yield fell by around 45 basis points, partly anticipating the response of the FOMC to financial conditions at their scheduled meeting on 29 September (the target rate was cut by 25 basis point), but also partly reflecting a flight to the safety of Treasury bonds (Longstaff 2004; Fontaine and Garcia 2012).

We formalize the aggregation of news. Consider month t, where $\mathcal{J}_t$ announcements are scheduled on days and times $t + \tau_j$, with $\tau_j \in (0, 1)$ and $j \in \{1, ..., \mathcal{J}_t\}$.[5] We measure the impact of macro news on futures-implied yields by sampling on a small window of time around the announcement:

$$\Delta\mathcal{F}_{n,t+\tau_j} = \mathcal{F}_{t+\tau_j+\mathrm{d}t} - \mathcal{F}_{t+\tau_j-\mathrm{d}t}\,. \tag{6}$$

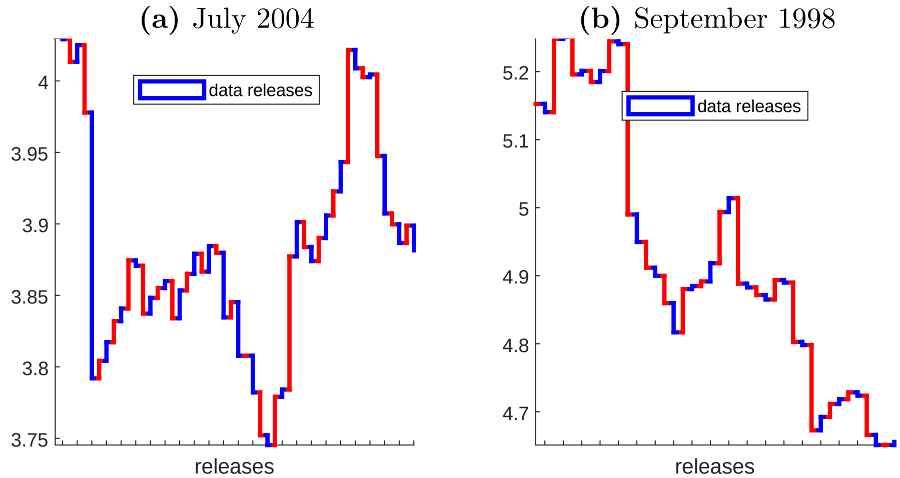
We use a $\mathrm{d}t = 45$min window around each release for robustness, because trading was slower for some contracts early in our sample.

Table 1 reports summary statistics for $\Delta\mathcal{F}_{n,t+\tau_j}$. These implied yield changes are centered around zero with a standard deviation of about 2 and 3 basis points (bps),

[5]The number of data releases in a month changes over time. For instance, there were 7 data releases with distinct date-time in January 1995 but 19 distinct releases in January 2016.

Figure 1: 5-Year Futures around and between Data Releases

Decomposition of changes in futures rates between changes around data releases $\mathcal{F}_{n,t+\tau_j}$ and changes between data releases. Panel a shows the decomposition during the month of July 2004 and Panel b for the month of September 1998.



but the changes have fat tails. The range between the 5th and 95th quantile is around 50 percent wider than what is predicted based on a normal distribution. It is well known that a few specific news releases are more timely and more informative about fundamental and lead to larger market responses (see, e.g., Gilbert, Scotti, Strasser, and Vega 2017).

	Mean	Median	Stdev	Q5	Q95	Min	Max
2-year	-0.02	0.00	2.59	-3.61	3.54	-34.57	27.47
5-year	-0.00	-0.01	2.65	-3.58	3.74	-30.54	28.24
10-year	-0.00	-0.02	2.14	-2.95	3.01	-32.27	20.19

Table 1: Summary statistics of futures yield changes around data releases (bps).

2.3 Monthly Yield Decomposition

Up to this point, we have defined a method that separates the effect of macro news and non-macro news in futures data. We present our methodology to map the futures changes into yield changes below.

First, we aggregate the responses $\Delta\mathcal{F}_{n,t+\tau_j}$ within a month. For month t, the total impact is given by

$$\Delta\mathcal{F}_{n,t} \equiv \sum_{j=1}^{\mathcal{J}_t} \Delta\mathcal{F}_{t-1+\tau_j} \,, \tag{7}$$

where, in general, $E_{t-1}[\Delta\mathcal{F}_{n,t}] \neq 0$, because the compensation for the risk that bondholders bear around macro releases accumulates in the sum (7). In the following, we assume that the premium is constant for every data release, such that $E_{t-1}[\Delta\mathcal{F}_{n,t}] = \mathcal{J}_t \cdot a$.

Next, we map the impact of macro news on Treasury yields. We extract the first three principal components of the Treasury yields $\mathcal{P}_t$, which are normalized so that they have mean zero and unit variance. We identify the news component of $\mathcal{P}_t$ using a linear mapping with $\Delta\mathcal{F}_{n,t}$:

$$\Delta\mathcal{P}_t = \gamma\mathcal{J}_t + \beta\Delta\mathcal{F}_{n,t} + \Delta\mathcal{P}_{y,t}, \tag{8}$$

where $\gamma = -\beta \cdot a$ captures the cumulative impact of the risk premium across all news events within each month and where $\Delta\mathcal{P}_{y,t}$ is the residual from the regression. We define the impact of macro news on the PCA:

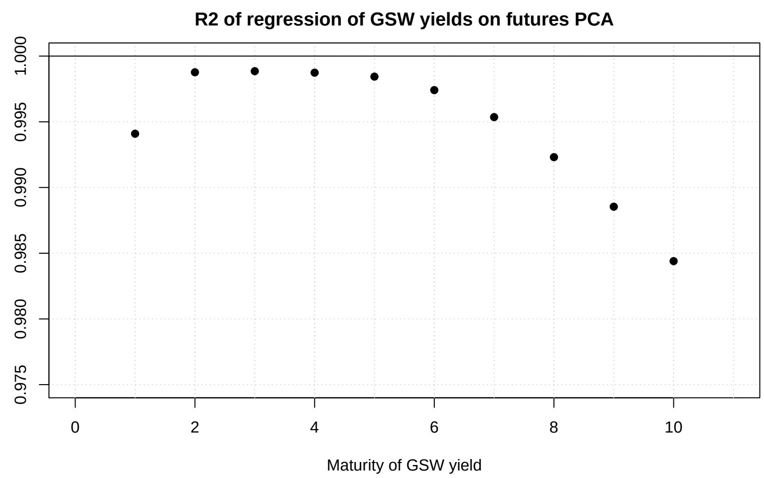
$$\Delta\mathcal{P}_{n,t} \equiv \hat{\gamma}\mathcal{J}_t + \hat{\beta}\Delta\mathcal{F}_{n,t}. \tag{9}$$

This approach means that the macro news component $\Delta\mathcal{P}_{n,t}$ is uncorrelated with the non-macro news component $\Delta\mathcal{P}_{y,t}$ by construction. In addition, it has mean zero, since the first term to the right of (8) captures any drift due to compensation for risk. Figure A.2 in the appendix reports the fitted values $\Delta\mathcal{P}_{n,t}$ as well as the scatter plot of $\Delta\mathcal{P}_t$ against $\Delta\mathcal{P}_{n,t}$. The R^2s for the level and slope component are close to 30 percent, which suggests that the share of variance in monthly yield changes due to macro news is close to 30 percent as well.

Our exercise relies on the ability of movements in futures yields to span the variations observed in bond yields. To confirm this intuition, we run regressions of zero-coupon yields $y_t^{(m)}$ on the three principal components of futures yields, using monthly data. We extract zero-coupon yields from Gurkaynak et al. (2006), GSW hereafter. Figure 2 shows that the R^2 in regressions of bond yields on futures yields range between 0.985 and 0.999. Second, Figure A.1 in the Appendix shows that the first two

Figure 2: R^2s in Regressions of Yields on Futures

R^2 in regressions of GSW zero-coupon yield $y_t^{(m)}$ on three futures-implied yields for bonds with 2, 5 and 10 years.



principals components from Treasury yields and futures (i.e., level and slope) are essentially identical in the time series, while the third principal component (curvature) exhibits only small differences.

2.4 Summary Statistics and Stylized Facts

We report a few key stylized facts about $\Delta\mathcal{P}_{n,t}$ and $\Delta\mathcal{P}_{y,t}$, since this decomposition of monthly yield changes has not been used before.

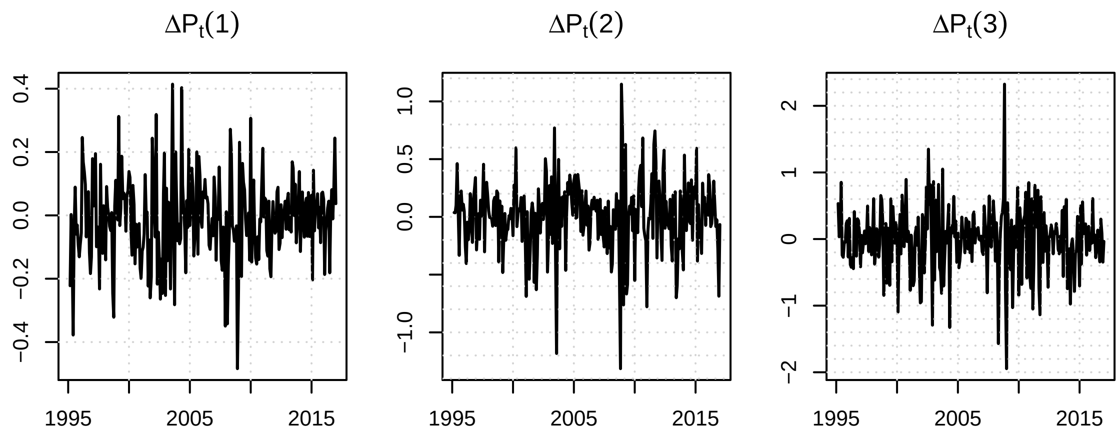
Autocorrelation

Figure 3 shows the time-series of monthly changes in futures $\Delta\mathcal{P}_t$ and $\Delta\mathcal{P}_{n,t}$. The patterns suggest that $\Delta\mathcal{P}_t$ has a slowly-moving component but that $\Delta\mathcal{P}_{n,t}$ is a white noise. Inspection of the autocorrelation and partial autocorrelation functions confirms that each element of $\Delta\mathcal{P}_t$ has auto-regressive and moving-average dynamics However, standard tests based on the autocorrelation and partial autocorrelation functions do not reject that each of the news components $\Delta\mathcal{P}_{n,t}$ is a white noise (unreported).

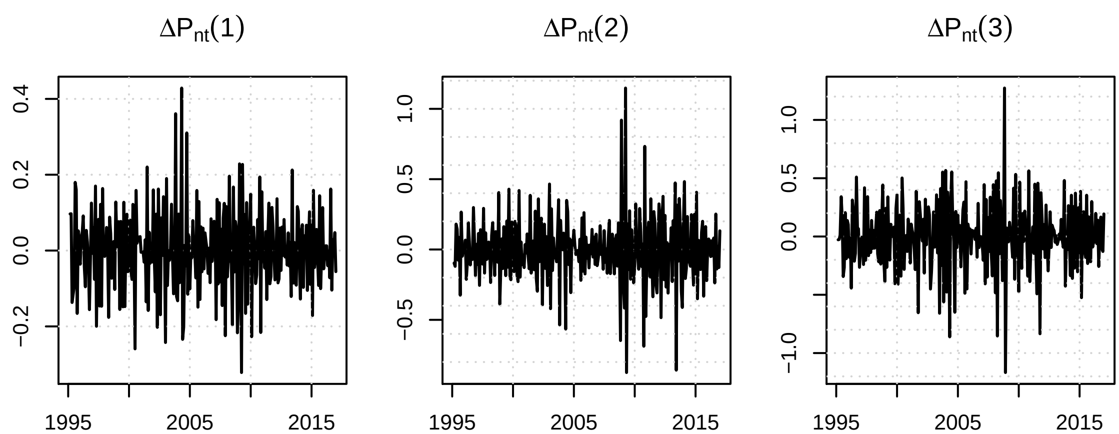
Figure 3: Monthly Changes $\Delta\mathcal{P}_t$ and $\Delta\mathcal{P}_{n,t}$

Monthly time-series of monthly changes and news components in CME futures referencing Treasury notes with 2, 5 and 10 years to maturity.

(a) Monthly Changes $\Delta\mathcal{P}_t$



(b) News Components $\Delta\mathcal{P}_{n,t}$



Returns Spanning Regressions

Consider the excess returns $ex_{t,h}^{(m)}$ to holding the bond with maturity m for $h \leq m$ periods and aggregate the impact of macro news over the holding period,

$$\Delta\mathcal{P}_{n,t,h} \equiv \sum_{i=1}^{h} \Delta\mathcal{P}_{n,t+i}. \tag{10}$$

We measure the contribution of macro news to the realized returns by comparing the R^2 from this regression

$$ex_{t,h}^{(m)} = \omega_{m,h} + \gamma_{m,h}^{\top}\mathcal{P}_t + \beta_{m,h}^{\top}\Delta\mathcal{P}_{n,t,h} + \epsilon_{t,h}^{(m)}, \tag{11}$$

with the R^2 from this restricted regression:

$$ex_{t,h}^{(m)} = \omega_{m,h} + \gamma_{m,h}^{\top}\mathcal{P}_t + \epsilon_{t,h}^{(m)}, \tag{12}$$

where we include $\mathcal{P}_t$ on the right of (11) to capture the predictable bond risk premium based on the level, slope and curvature. Because the component $\Delta\mathcal{P}_{n,t,h}$ is not correlated with $\mathcal{P}_t$, the difference between the R^2s from these regressions measures the share of realized bond returns that is due to macro news.

Figure 4 reports the R^2s for bond with annual maturity between one and then years and for h ranging from one month to two years.[6] At the horizon of one month, the contribution of macro news to excess returns range from 25 to 35 percent from the 1-year bond to the 10-year bond. As we increase the horizon, the impact of macro news exhibits a declining pattern, ranging from around 10 to 20 percent at a 2-year horizon. Perhaps surprisingly, as we increase the horizon, the accumulated impact of macro news explains a decreasing share of realized returns. This is a pattern that potential dynamic term structure must match. It also means that new information that arrives to the bond market outside of the release of economic data explains the majority of yield variances.

It is tempting to compare these results with the central results in Altavilla et al. (2017), showing that the impact of macro news increases between the daily and quarterly horizons, which correspond to the first three observations to the left in Figure 4. At the first glance, our results agree: the share ranges between one quarter to one third at these short horizons. But there are some differences. They find that the share of macro news increase from around 15-25 percent at the monthly horizon to around 30-35 percent at the quarterly frequency. We show a flat pattern across the monthly and quarterly horizons. They find that the share tends to decline with a bond maturity, we find that the share tends to increase with maturity.

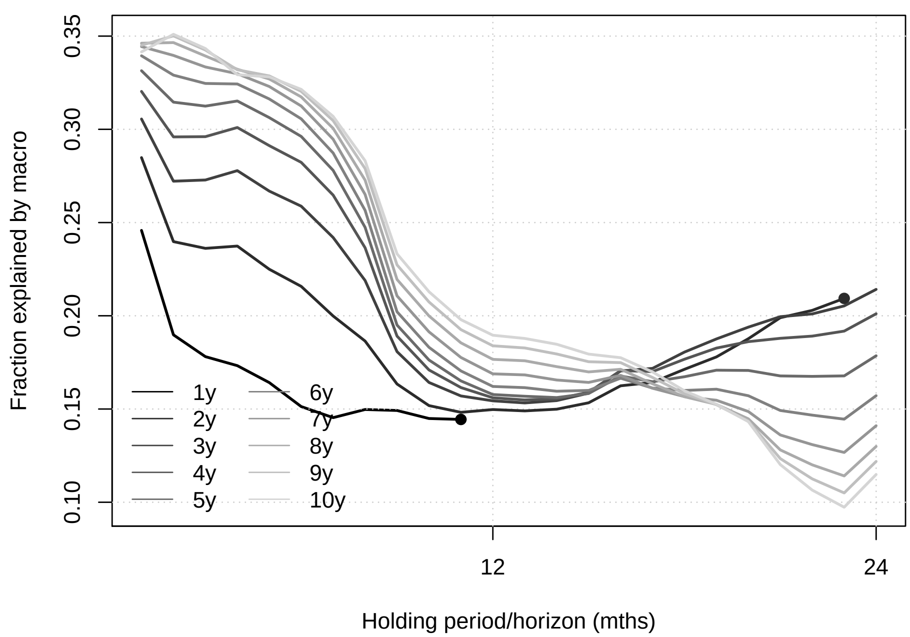
There are a few key differences in our approaches. First, we want to isolate

[6] We compute all yields using the Nelson-Siegel-Svensson parameters provided by GSW.

the impact of macro news on the unpredictable component of yield changes. For this reason, we examine excess returns and we control for predictable risk premium variations. This corresponds to the impact on the yield forecast errors $\tilde{y}_t^{(m)}$ in (5). Second we measure the news using yield changes, and not the headlines, to capture the depth of information released beyond the headlines (Gurkaynak et al., 2018). This corresponds to the full change in the information set around the news releases, as captured by the $(E_t - E_{t-1})$ operator in (4).

Figure 4: Returns Spanning Regressions

Incremental R^2s from contemporaneous regressions of excess returns $er_{t,h}^{(m)}$ on aggregated news $\Delta\mathcal{P}_{n,t,h}$ including $\mathcal{P}_t$ relative to the constrained regression that exclude $\mathcal{P}_t$ for bond maturity m across different horizons h.



3 Model

This Section introduces a family of DTSMs that are flexible enough to reproduce the stylized facts presented in Section 2. These models are closely related to standard affine Gaussian DTSMs but they give $\Delta\mathcal{P}_{n,t}$ and $\Delta\mathcal{P}_{y,t}$ distinct roles in the historical dynamics. This more general specification nests many existing DTSMs.

3.1 Risk-Neutral Dynamics

Assumption 1 postulates a small number of $\mathcal{K}_1$ portfolios $\mathcal{P}_t$ driving the cross-section of bond yields.

Assumption 1 *The yield $y_t^{(m)}$ of the zero-coupon bond of maturity m can be expressed as a linear function of $\mathcal{P}_t \in \mathbb{R}^{\mathcal{K}_1}$,*

$$y_t^{(m)} = A_m + B_m^\top \mathcal{P}_t. \tag{13}$$

Assumption 1 is consistent with the common observation that a small number of portfolios summarizes the cross-section of yields. In practice, we set $\mathcal{K}_1 = 3$. Note that the coefficients A_m and B_m are unrestricted at this stage.

Define the investors' information set at date t: $\mathcal{I}_{\mathcal{P},t} = \{\mathcal{P}_{n,t}, \mathcal{P}_{y,t}, \ldots, \mathcal{P}_{n,0}, \mathcal{P}_{y,0}\}$. Proposition 1 recalls a result from Feunou and Fontaine (2018). In the absence of arbitrage, Assumption 1 implies that the portfolios $\mathcal{P}_t$ must be Markovian under the risk-neutral measure $\mathbb{Q}$.

Proposition 1 *Risk-neutral dynamics*

If bond prices offer no arbitrage opportunity and if the conditional distribution of $\mathcal{P}_{t+1}$ is homoskedastic Gaussian given $\mathcal{I}_{\mathcal{P},t}$ under the risk-neutral measure $\mathbb{Q}$, then Assumption 1 is equivalent to the Markov property for $\mathcal{P}_{t+1}$ under $\mathbb{Q}$ and we have that:

$$\Delta \mathcal{P}_{t+1} = K_0^{\mathbb{Q}} + K_{\mathcal{P}}^{\mathbb{Q}} \mathcal{P}_t + \Sigma^{1/2} \varepsilon_{t+1}^{\mathbb{Q}}, \tag{14}$$

where $\varepsilon_{t+1}^{\mathbb{Q}}$ is a standard Gaussian innovation with respect to $\mathcal{I}_{\mathcal{P},t}$. In addition, A_m and B_m follow standard recursions.

Once we fix the number of linear portfolios required to explain the cross-section of yields, then the factors must have linear Markovian dynamics under any risk-neutral measure. Therefore, we use a standard specification of the risk-neutral dynamics in the following. The result relies on the absence of arbitrage opportunities among bond prices. This is also consistent with the model-free evidence in Joslin et al. (2013) that pricing portfolios are Markovian under the $\mathbb{Q}$ measure. In practice, nearly all GDTSMs use dynamics as in Equation (14) to derive Equation (13) for yields.

3.2 Historical Dynamics

One key implication of Proposition 1 is that separately identifying the impact of macro news is only relevant for the historical dynamics. Based on the stylized facts in Section 2.4, we make an important assumption.

Assumption 2 *The impact of macro news on yield factors is unpredictable* $E_t[\Delta\mathcal{P}_{n,t}] = 0$ *given the investors' information set* $\mathcal{I}_{\mathcal{P},t}$.

Assumption 2 is verified because of the way we constructed $\Delta\mathcal{P}_{n,t}$ via the regression in (8). In particular, it implies that $\Delta\mathcal{P}^{(y)}_{t+1}$ is the only predictable component in $\Delta\mathcal{P}_{t+1}$:

$$E_t\left[\Delta\mathcal{P}_{t+1}\right] = E_t\left[\Delta\mathcal{P}_{n,t+1} + \Delta\mathcal{P}_{y,t+1}\right] = E_t\left[\Delta\mathcal{P}^{(y)}_{t+1}\right] \equiv \mathcal{E}_t\,. \tag{15}$$

Next, we provide a general specification for the dynamics of $\Delta\mathcal{P}_{y,t+1}$, which we will then motivate based on the macro-finance framework in Ang and Piazzesi (2003).

Assumption 3 *The conditional dynamics of* $\Delta\mathcal{P}_{t+1}$ *are given by*

$$\begin{aligned}\Delta\mathcal{P}_{t+1} &= \mathcal{E}_t + u_{t+1} + \Delta\mathcal{P}_{n,t+1} \\ &= K_0 + \Theta\mathcal{E}_{t-1} + K_{\mathcal{P}}\mathcal{P}_t + K_y\Delta\mathcal{P}_{y,t} + K_n\Delta\mathcal{P}_{n,t} + u_{t+1} + \Delta\mathcal{P}_{n,t+1}\,,\end{aligned} \tag{16}$$

where $\mathcal{E}_t$ *is the conditional mean* $E_t[\Delta\mathcal{P}_{y,t+1}]$ *and where the innovations* u_{t+1} *and the impact of macro news* $\Delta\mathcal{P}_{n,t}$ *are uncorrelated.*

Assumption 3 can be motivated based on a standard macro-finance framework. Consider $\mathcal{K}_2 \leq \mathcal{K}_1$ latent factor $\mathcal{M}_t$ that captures the macroeconomic information relevant to explain the dynamics of bond yields. For instance, Ang and Piazzesi (2003) use one principal component from inflation-related variables and one principal component from real activity variables. Then Assumption 3 is equivalent joint dynamics for $\mathcal{P}_t$ and $\mathcal{M}_t$

$$\begin{aligned}\begin{pmatrix} \mathcal{P}_t \\ \mathcal{M}_t \end{pmatrix} &= \begin{pmatrix} I + K_{\mathcal{P}} + K_y & I \\ \Theta K_{\mathcal{P}} + (\Theta - I)\,K_y & \Theta \end{pmatrix}\begin{pmatrix} \mathcal{P}_{t-1} \\ \mathcal{M}_{t-1} \end{pmatrix} \\ &\quad + \begin{pmatrix} I & I \\ 0 & K_n - K_y \end{pmatrix}\begin{pmatrix} u_t \\ \Delta\mathcal{P}^{(n)}_t \end{pmatrix},\end{aligned} \tag{17}$$

where all variables are demeaned for simplicity. The intuition behind this result is that by observing the yield factors $\mathcal{P}_t$ and the impact of macro news $\Delta\mathcal{P}_{n,t}$, the econometrician can filter the forecast of future yields as effectively as if he had observed the macro factors $\mathcal{M}_t$.

Note that the dynamics in (17) embed two identification assumptions. First, the macro-news shocks $\Delta\mathcal{P}_t^{(n)}$ enter directly in the yields factors without scaling or rotation. This is merely a consequence of the accounting relationship that innovation in $\mathcal{P}_t$ is given by $u_t + \Delta\mathcal{P}_t^{(n)}$. Second, the top-right of the autoregressive matrix is rotated to the identity matrix. This is merely for convenience and corresponds to a rotation of the macroeconomic factors.

3.3 Nested Yield-Only Models

It is useful to re-write the joint dynamics for $\mathcal{P}_t$ and $\Delta\mathcal{P}_{y,t+1}$ using the following minimal representation,

$$\begin{pmatrix} \mathcal{P}_{t+1} \\ \Delta\mathcal{P}_{y,t+1} \end{pmatrix} = \begin{pmatrix} K_0 \\ K_0 \end{pmatrix} + \begin{bmatrix} I + K_{\mathcal{P}} & K_y + \Theta \\ K_{\mathcal{P}} & K_y + \Theta \end{bmatrix} \begin{pmatrix} \mathcal{P}_t \\ \Delta\mathcal{P}_{y,t} \end{pmatrix} + \begin{pmatrix} u_{t+1} \\ \Delta\mathcal{P}_{n,t} \end{pmatrix} - \begin{bmatrix} \Theta & -K_n \\ \Theta & -K_n \end{bmatrix} \begin{pmatrix} u_t \\ \Delta\mathcal{P}_{n,t} \end{pmatrix}. \tag{18}$$

One important reason why this representation is useful is that we can easily check the mappings with existing DTSMs. First, one can check that the standard DTSM model with VAR(1) dynamics (e.g., Joslin et al. 2011) obtains if $K_y = K_n = 0$ and if $\Theta = 0$. Second, relaxing $K_y = K_n \neq 0$ leads to models with VAR(2) dynamics. Again, the dynamic implications from identifying the monthly news component are limited. Finally, one can also check that relaxing $\Theta \neq 0$ but with $K_y = K_n = 0$ leads to family of CM-DTSM model in Feunou and Fontaine (2018).

These yield-only special cases have in common the restriction $K_y = K_n$. Then, there is no distinct roles in the yield dynamics for the news and non-news components. Every month, the shocks u_{t+1} and $\Delta\mathcal{P}_{n,t+1}$ combine to produce $\mathcal{P}_{t+1}$, but there is no further dynamic implications. However, it is still possible to use high-frequency data to measure $\Delta\mathcal{P}_{n,t+1}$ and identify the effect of macro news on yields. In the following, we are interested in models with $K_y \neq K_n$, which are closely related to macro-finance models where macroeconomic variables play an explicit role.

There is another reason why this representation is useful: we can easily check conditions for the stationarity and invertibility of the process $\left(\mathcal{P}_t,\ \Delta\mathcal{P}^{(y)}_{t+1}\right)$ using standard results for multivariate linear processes (see Lütkephol 2006). For completeness, we provide these conditions in Proposition 2.

Proposition 2 *Stationarity and Invertibility*
The process for $(\mathcal{P}_t,\ \Delta\mathcal{P}_{y,t+1})$ is (i) stationary and (ii) invertible if and only if:

$$(i) \quad \max \left| \mathrm{sp} \begin{pmatrix} I + K_{\mathcal{P}} & K_y + \Theta \\ K_{\mathcal{P}} & K_y + \Theta \end{pmatrix} \right| < 1 \tag{19}$$

$$(ii) \quad \max \left| \mathrm{sp} \begin{pmatrix} \Theta & -K_n \\ \Theta & -K_n \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & -I \end{pmatrix} \right| < 1. \tag{20}$$

3.4 Reducing the Dimension of News

Relative to a standard VAR(1), the model in Section 3 introduces $2 \times \mathcal{K}_1^2$ new parameters in the matrix K_y and K_n. Fortunately, the evidence in Section 2.4 points toward parsimonious restrictions that we can implement: yield changes around data releases can be summarized with one linear combination. From the lenses of a macro-finance model in (17), this corresponds to the case with $\mathcal{K}_1 = 1$ macro factor.

In the following, we implement unrestricted versions of the models as well as restricted $\mathcal{K}_1 = 1$ models. The idea that the news and non-news component can each be summarized with one linear component corresponds to re-writing the matrix K_n and K_y parsimoniously:

$$K_n = \alpha k_n^{\top} \tag{21}$$

$$K_y = \alpha k_y^{\top} \tag{22}$$

where α, k_n and k_y are $\mathcal{K}_1$ vectors. The moving-average matrix Θ also adds $\mathcal{K}_1^2$ parameters relative to the VAR(1) . We impose the following parsimonious restriction:

$$\Theta = \theta \cdot I, \tag{23}$$

where θ is a scalar, which is consistent with Feunou and Fontaine (2018) who show that parsimonious restrictions for Θ are supported in the data.

3.5 What to expect from variance decomposition

We can calibrate a toy version of (17) to help understand the roles of K_n and K_y in the variance decomposition. Suppose that there is only one macro factor and one yield factor $y_t^{(1)} = \mathcal{P}_t$. In this case, we can write:

$$\begin{pmatrix} \mathcal{P}_{t+1} \\ \mathcal{M}_{t+1} \end{pmatrix} = \begin{pmatrix} \rho & 1 \\ 0 & \theta \end{pmatrix} \begin{pmatrix} \mathcal{P}_t \\ \mathcal{M}_t \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & k_n - k_y \end{pmatrix} \begin{pmatrix} u_{t+1} \\ \Delta\mathcal{P}_{n,t} \end{pmatrix}, \tag{24}$$

where, following Ang and Piazzesi (2003), there is no feedback from the macro to the yields $\rho = 1 - k_y/\theta$ and, for simplicity, we take $\alpha = 1$. Therefore, the variance of forecast errors of any $\beta_m \mathcal{P}_{t+h}$ is given by:

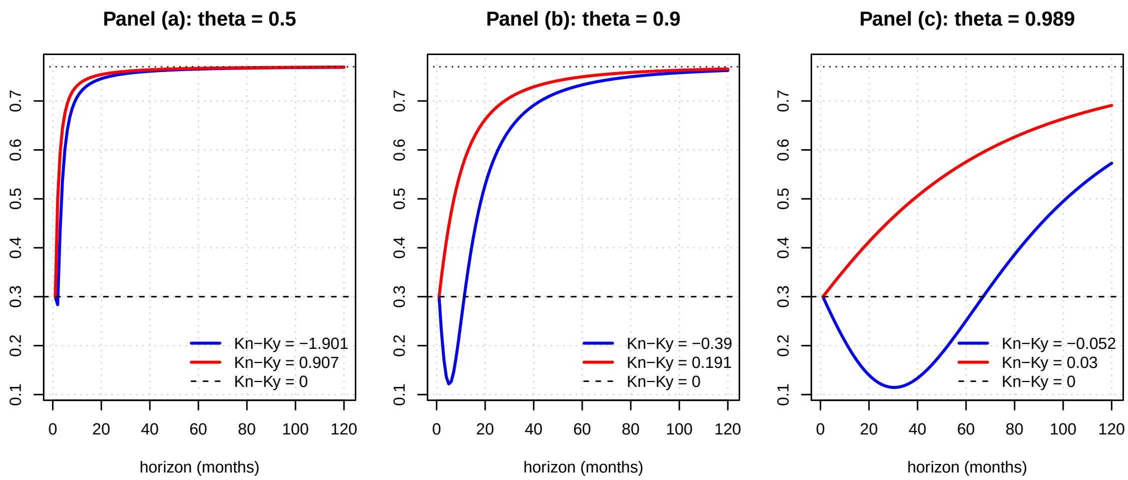
$$\begin{aligned} &= \beta_m^2 \left(\Sigma_u + \Sigma_n\right) \tfrac{1-\rho^{2h}}{1-\rho^2} \\ &\quad + \beta_m^2 \Sigma_n \tfrac{k_n - k_y}{\theta - \rho} \times \left\{ 2\left(\frac{1-\theta^h\rho^h}{1-\theta\rho} - \frac{1-\rho^{2h}}{1-\rho^2}\right) \right. \\ &\qquad \left. + \frac{k_n - k_y}{\theta - \rho}\left(\frac{1-\theta^{2h}}{1-\theta^2} + \frac{1-\rho^{2h}}{1-\rho^2} - 2\frac{1-\rho^h\theta^h}{1-\rho\theta}\right) \right\}. \end{aligned} \tag{25}$$

In this simple model, the variance ratios do not depend on the maturity, because we can factor out β_m^2. In addition, the second term disappears if $k_n = k_y$ and the variance ratios is constant $\Sigma_n/(\Sigma_n + \Sigma_u)$.

To help understand the impact of $k_n - k_y$, we compute and report the variance ratios for different model calibrations. We fix the first eigenvalue to $\rho = 0.99$, which is consistent with the persistence of observed yields, and vary the persistence of the macro variable $\theta \in \{0.5, 0.9, 0.989\}$. Next, we fix the variance of macro news shocks to their in-sample value $\Sigma_n = 0.013$ and, for each value of θ, we calibrate Σ_u such that the variance ratio for macro is 30 percent at the horizon $h = 0$, consistent with existing evidence. Last, for each value of θ, we then set $k_n - k_y$ to match the sample standard deviation of $\mathcal{P}_t$, equal to 1.819%. This condition is quadratic in $k_n - k_y$ and we report results for the positive and negative solutions.

Figure 5 reports the share of variance attributed to macro news across horizons for each calibration. By design, this share is flat at 30 percent when $k_n = k_y$. Otherwise,

Figure 5: Variance Decomposition in a Stylized Model



this share rises and converges to about 75 percent in all cases, by design, since we match the sample standard deviation of $\mathcal{P}_t$ and because of the lack of feedback from yields to macro. Across values of θ, increasing the persistence of macro slows down the transition between the initial 30 percent ratio to the long-run 75 percent ratio.

We are interested in $k_n - k_y$. This parameter controls the shape of the variance ratio between $h = 0$ and $h \to \infty$. When $k_n > k_y$, the macro share increases monotonically from 30 to 75 percent. However, when $k_n < k_y$ the macro share initially declines. For $\theta = 0.989$, the macro share of variance declines to as low as 10 percent and stays below 30 percent at horizons beyond five years. This simple exercise tells us that allowing for $K_n \neq K_y \neq 0$ can produce rising or declining pattern in the macro share of variance. It also suggests that a restrictions to macro-finance models can generate long-run variance ratios that are too high, at least in the light of the stylized facts above. The following section provides a more thorough analysis with the estimation of multivariate models that relax several key restrictions.

4 Results

4.1 Mapping Futures to Bond Yields

Our analysis focuses on the comparison between different historical dynamics, especially in relation with the standard VAR(1) yield-only dynamics. We want to know how the dynamic impact of macro news $\Delta\mathcal{P}_{n,t}$ increases with the horizon and what is its role in the dynamics of bond returns. We estimate and evaluate the specifications detailed in Table 2.

Table 2: Model Specifications
Taxonomy of the main specifications of interest and the number of parameters #, but excluding parameters of the covariance matrix, which is identical for each of these specifications.

Model	K_y	K_n	Θ	# parameters
VAR(1)	-	-	-	12
VAR(2)	-	-	-	21
VAR-RR	$\alpha k_y^\top$	$\alpha k_n^\top$	-	20
CM-RR	$\alpha k_y^\top$	$\alpha k_n^\top$	$\theta \cdot I$	21
VAR-UR	K_y	K_n	-	30
CM-UR	K_y	K_n	$\theta \cdot I$	31

The labels UR and RR refers to unrestricted and restricted rank specifications, respectively. The simple VAR(1) dynamics imply 12 parameters. Allowing for unrestricted coefficients K_n and K_y introduces 18 new parameters. Of course, allowing for macro information in the dynamics for $\mathcal{P}_t$ increases the number of parameters. But note that the number of parameters also increases if we directly include macro variables the model, as in Equation 17. In fact, the number of parameters is the same if the number of macro variables is $\mathcal{K}_1 = \mathcal{K}_2$ the number of risk factors. The restrictions that $\Delta\mathcal{P}_{n,t}$ and $\Delta\mathcal{P}_{y,t}$ can be summarized with one linear combination reduces the number of parameters to 20 (K_n and K_y have rank one).

4.2 Estimation

Estimation proceeds in two steps. In the first step, we estimate parameters of the historical dynamics. In the second step, we estimate parameters of the risk-neutral dynamics. The key reason for this choice is to emphasize how the variance decomposition differs across different specifications of the historical dynamics.

Estimation of the historical dynamics in the first step relies on the dynamics of $\Delta\mathcal{P}_{y,t}$. From Equation 16:

$$\begin{aligned}\Delta\mathcal{P}^{(y)}_{t+1} &= \mathcal{E}_t + u_{t+1} \\ &= K_0 + \Theta\mathcal{E}_{t-1} + K_{\mathcal{P}}\mathcal{P}_t + K_y\Delta\mathcal{P}_{y,t} + K_n\Delta\mathcal{P}_{n,t} + u_{t+1}\,,\end{aligned}$$

The macro news component is measured from high-frequency data and the impact of other news is captured by the residual $\Delta\mathcal{P}_{y,t} = \Delta\mathcal{P}_t - \Delta\mathcal{P}_{n,t}$. Then, the conditional mean $\mathcal{E}_t$ is measurable given the information up to time t and can be computed recursively. Hence, the parameters of the historical dynamics can be estimated using maximum likelihood based on u_{t+1}. This procedure simplifies to the method of ordinary least squares whenever $\Theta = 0$.

We need an estimate of the variance-covariance matrix Σ in order to compute the factor loadings. This is simply

$$\Sigma \equiv \mathbb{V}_t\left(\mathcal{P}_{t+1}\right) = \Sigma_u + \Sigma_n, \tag{26}$$

using the plug-in estimators:

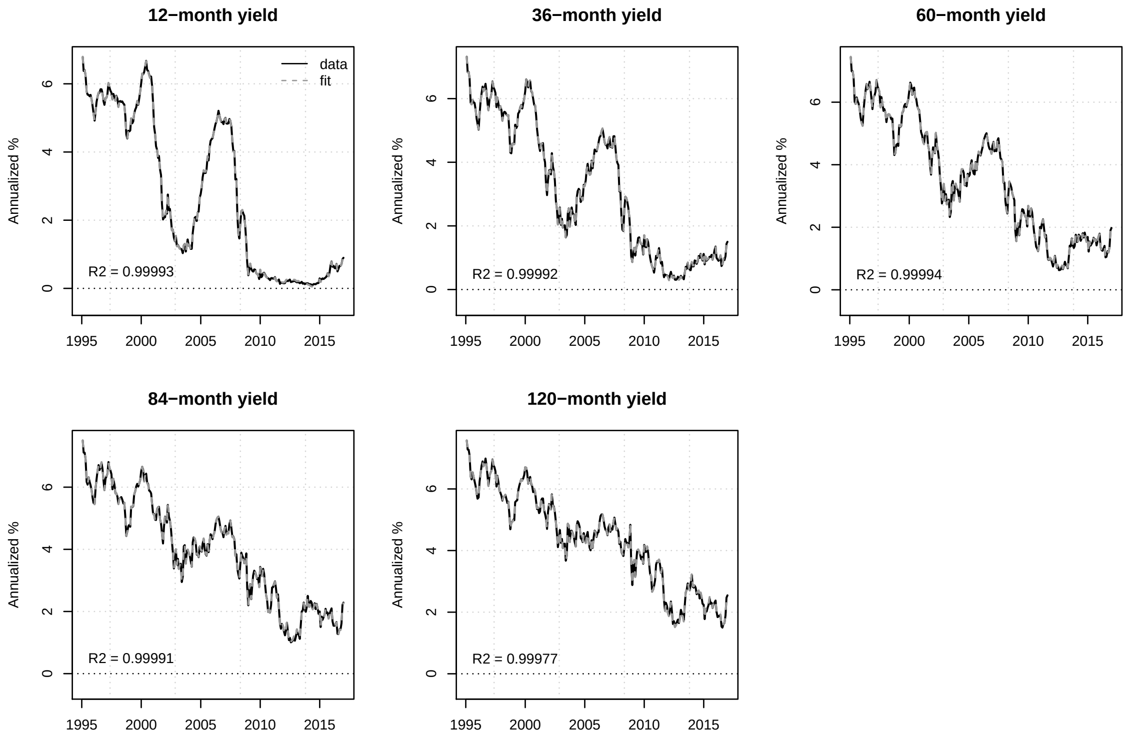
$$\Sigma_n = \frac{1}{T-1}\sum_{i=1}^{T}\left(\Delta\mathcal{P}_{n,t}\right)^2 \quad \text{and} \quad \Sigma_u = \frac{1}{T-1}\sum_{i=1}^{T}\Delta u_t^2, \tag{27}$$

and yields are given by:

$$y_t^{(m)} = A_m(\Sigma) + B_m^{\top}(\Sigma)\,\mathcal{P}_t + \eta_t, \tag{28}$$

where η_t is a measurement errors with diagonal covariance matrix Σ_η. We denote $A_0 = \rho_0$ and $B_0 = \rho_1$ (i.e., $y_t^{(1)} = \rho_0 + \rho_1^{\top}\mathcal{P}_t$). Then, estimation of the risk-neutral parameters ρ_0, ρ_1, $K_0^{\mathbb{Q}}$ and $K_1^{\mathbb{Q}}$ can proceed via non-linear least squares. In practice, we minimize the pricing errors for GSW zero-coupon yields with maturities 1, 3, 5, 7 and 10 years.

Figure 6: Pricing Errors in VAR(1) Model
Observed (solid) and fitted (dashed) yields based on estimates of the VAR(1) model. Data from GSW (1995–2016).



4.3 Pricing Errors

Figure 6 compares the observed and the fitted yields for the VAR(1) . These are nearly indistinguishable and the root mean square errors range from 1.97 to 2.32 basis points across bond maturities. Results are essentially the same for other models (unreported). Unsurprisingly, every model that we estimate fits the yield curve nearly perfectly, since every model uses three portfolio of yields as pricing factors. For our purpose, we can trust the good fit in the cross-section of yields as a building block of reliable variance decompositions.

4.4 Historical Dynamics

This section compares the fit of the historical dynamics across models. In short, the evidence overwhelmingly supports a separate channels for macro news releases in the dynamics for yields.

Table A.1 in the appendix reports all the parameter estimates. Table 3 reports the R^2s across models in forecasting the predictable component $\Delta\mathcal{P}_{y,t+1}$. Higher predictability indicates a better fit of the historical dynamics. The R^2s in the benchmark VAR(1) modelrange around 2-3 percent for the level and slope components $\mathcal{P}_t^{(1)}$ and $\mathcal{P}_t^{(2)}$. The R^2s clearly improve for the VAR-UR model, around 8-10 percent. The same conclusion holds if we compare the CM(1) model with the corresponding CM-UR model. The results clearly show that distinguishing between the information in $\Delta\mathcal{P}_{n,t+1}$ and $\Delta\mathcal{P}_{y,t+1}$ improves predictability. In other words, macro news play a distinct role in the dynamics of yields.

Table 3: Model Predictive R^2s

In-sample time-series $R^2 = 1 - \frac{V[u_t(i)]}{V[\Delta\mathcal{P}_{y,t}(i)]}$ for each component of $\Delta\mathcal{P}_{y,t}$ across models.

	$\mathcal{P}_{y,t}(1)$	$\mathcal{P}_{y,t}(2)$	$\mathcal{P}_{y,t}(3)$
VAR(1)	2.269	2.652	9.347
VAR(2)	5.749	7.021	9.953
VAR-UR	8.295	9.74	10.159
VAR-RR	5.898	4.081	9.365
CM(1)	2.473	2.384	9.494
CM-UR	6.47	11.756	12.033
CM-RR	4.241	4.895	10.181

So far, we compared models using the in-sample predictive R^2s as a measure of fit. Table 4 also reports the sum of squared residuals (SSR), the Gaussian log-likelihood as well as model section criteria. The model rankings are essentially unchanged. Since we use the maximum likelihood as estimation method, the least constrained model CM-UR straightforwardly exhibits the highest likelihood. But the same holds true for the SSR and AIC. The CM-RR model with rank restriction comes very close however, and even finishes first based on the AICc criteria.

Overall, these statistical criteria confirm that the more general CM-UR model should be favored over any of the models where macro news do not play a distinct role (i.e., $K_y = K_n$) and over any of the models without a moving-average component (i.e., $\Theta = 0$). There is some evidence to favor the more parsimonious implementation CM-RR .

We check that the results are not affected by issues related to the stationarity or invertibility of the models. The last two columns of Table 4 show that every model satisfies the stationarity and invertibility conditions.

Table 4: Model Selection Criteria
SSR is the sum of squared residuals, Loglik is the Gaussian log-likelihood, AIC and AICc are the Akaike rough and corrected information criterion, respectively, and —AR— and —MA— provide the maximal modulus of eigenvalues of the autoregressive and moving average matrix (See Equation 18). The symbol ★ indicates the best model and the symbol ▼ indicates the worst model.

	SSR	Loglik	Par.	AIC	AICc	|AR|	|MA|
VAR(1)	67.095▼	806.84▼	18	-1565.68	-1560.659	0.98	0
VAR(2)	65.896	827.091	27	-1588.182	-1578.425	0.982	0
VAR-UR	65.226	836.891	36	-1589.781	-1573.437	0.98	0
VAR-RR	66.72	826.676	26	-1589.351	-1580.209	0.98	0
CM(1)	67.061	807.425	19	−1564.85▼	−1559.388▼	0.979	0.187
CM-UR	63.955★	851.367★	37	−1616.734★	-1599.534	0.987	0.665
CM-RR	66.204	838.254	27	-1610.508	−1600.751★	0.986	0.685

4.5 Variance Decomposition of Pricing Factors

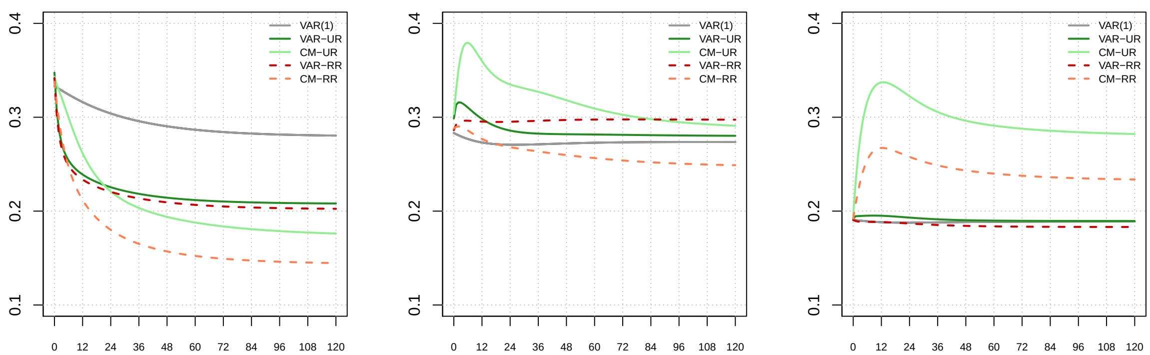
Figure 7 reports the variance decomposition for the element of $\mathcal{P}_t$ (The variance decomposition of yields is reported in the next section. The initial point $h = 0$ in each panel can be computed directly from Equation 26 and it is comparable to results of event studies. The interpretation of the decomposition at this horizon is that we isolate the share of the monthly changes $\Delta\mathcal{P}_t$ that is due to macro news within the month. We find shares close to 35 percent for $\mathcal{P}_t^{(1)}$, around 30 percent for $\mathcal{P}_t^{(2)}$ and close to 20 percent for $\mathcal{P}_t^{(3)}$. These results are consistent with the regression results reported in Figure A.2. Hence, we expect the role of macro news in the variance of yields to also decrease with the horizon.

One important question is whether the share of variance explained by macro news continues to grow at longer horizon. In a regression framework, the results beyond the quarterly horizons are imprecise due to the limited sample size. One benefit of our approach is that we can compute the variance decomposition at any horizon based on the estimated dynamics. For all models, we find that the share of macro news in the variations of level factor $\mathcal{P}_t^{(1)}$ declines with the horizon. In the VAR(1) model,

Figure 7: Variance Decomposition

Variance decomposition of the risk portfolios $\mathcal{P}_t$ for the models VAR(1) (grey solid line), VAR-UR (darkgreen solid line), VAR-RR (red dashed line), CM-UR (lightgreen solid line), and CM-RR (dashed orange line) at horizons ranging from 0 to 120 months (10 years). Panels (a) to (c) presents portfolio one, two and three, respectively.

(a) $\mathcal{P}_{1,t}$ **(b)** $\mathcal{P}_{2,t}$ **(c)** $\mathcal{P}_{3,t}$



the decline is modest, reaching a plateau around 28-30 percent at horizon close to 5 years. In every other specification that we consider, we observe a steep decline to a lower level, around 15-20 percent depending on the model. The results show that the share the variance of the slope portfolio $\mathcal{P}_t^{(2)}$ is roughly constant across horizons, and that the share in the variance of the curvature portfolio $\mathcal{P}_t^{(3)}$ increases with the horizons for the CM-UR and CM-RR models.

4.6 Variance Decomposition of Yields

At the monthly horizon, the variance decomposition of a yield with maturity m is straightforward

$$\mathbb{V}_t\left(y_{t+1}^{(m)}\right) = \underbrace{B_m^\top \Sigma_u B_m}_{\text{non-macro}} + \underbrace{B_m^\top \Sigma_n B_m}_{\text{macro}}, \tag{29}$$

which can be computed using the first-stage estimates of the variance and covariance terms as well as the second-stage estimates of the risk-neutral dynamics. At this horizon, the results are comparable with results in Altavilla et al. (2017) who find that 15-25 percent of yield variations can be identified with macro news. The key difference between their results and ours is that they use the headline surprises while we use yield changes around the release time. At the quarterly horizon, they find that 25-35 percent of yield variations can be identified with macro news, similar to our results. One interpretation is that the cumulative content of headline news over one quarter has similar information to the information beyond the headlines over one month.

For longer horizons, we rely on the estimated dynamics to compute the variance decomposition. Figure 8 reports the results for bond yields, for the expectation component and for the term premium separately, using estimates for CM-UR model. Panel (a) shows that the share due to macro news in the variances of yields declines with the horizon for all bond maturities. As in Figure 7 for the level factor $\mathcal{P}_t^{(1)}$, the decline is sharper for our models with a distinct role for macro news. Figure A.3 in the appendix reports very similar pattern using any of the other estimated models, except for the VAR(1) .

Another takeaway from Panel(a) in Figure 8 is that the share of yield variances imputed to by macro news increases with maturity. For the 1-year yield, the CM-RR model imputes 25 percent of the variance to macro news on impact and declining to 15 percent at longer horizons. For the 10-year yield, the shares are 35 percent and close to 20 percent, respectively. In a regression framework, Gurkaynak et al. (2018) find a hump-shape pattern in the responses of interest rates around data releases. The hump shape is apparent in the responses to headline news or to information beyond the headlines (see their Tables 1-4). Hence, our results agree in that short-term yields are less responsive to macro news relative to long-term yields (and with a lower R^2), but we find a monotonous pattern at longer maturities.

Panel (b) of Figure 8 reports the share of the expectation component $eh_t^{(m)}$ due to macro news while Panel (c) reports the share of the risk premium component $rp_t^{(m)}$ due to macro news. Two key results emerge. First, the impact of macro news on the term premium is close to 50 percent at a monthly horizon for the 10-year bond and gradually declines to around 40 percent at longer horizons. Second, the impact of macro news is low, below 10 percent, across all horizons and maturities. The net effect on yields combine the low and flat share of macro news in the expectation components with the declining share of macro news in the term premium.

The variance decomposition is a statement about the unconditional distribution. We can also decompose the time-series of yields forecast errors $\tilde{y}_t^{(m)}$ between the revisions to the expectation component $\Delta eh_t^{(m)}$ and revisions to the term premium component $\Delta rp_t^{(m)}$. Figure 9 reports the results. The first row of figures focuses on $\Delta eh_t^{(m)}$ for maturities of 1, 5 and 10 years.

Unsurprisingly, the largest revisions in expectations occurred when the Fed funds rate was rising (1999-2000 and 2004-2006) or when it was decreasing (2001-2002 and 2007-2008). However, the impact of macro news on the revision in expectations was relatively low. At other times, when the Fed funds rate was stable, the revisions to the expectation component were in line with the impact of macro news. How can expectations respond so little to macro news when the Fed is tightening or easing its policy rate? One reason could be that other sources of information, such as speeches, play a determinant role. Alternatively, it could be that the information about the economy in bond yields arises because of trading by expert or well-informed investors.

The second row in Figure 9 shows the revisions to the risk premium component $\Delta rp_t^{(m)}$ for the same bond maturities. In this case, there is no clear pattern between the contribution macro news and the periods with large changes in the Fed funds rate.

4.7 Term Premium

Figure 10 shows the term premium across different models relative to the term premium from the VAR(1) (as a difference). The term premium is essentially the difference between observed yields and the average expected short-term interest rates.[7] We see large differences across models. In particular, the VAR(1) seems to underestimate the term premium relatively to the other models when rates are high, and underestimate when rates are low. The differences are often larger as 40 bps (either way) for both the CM-UR and the parsimonious CM-RR . One take-away is that neglecting the differential impact of macro news produces term premium that are too smooth and underestimate the impact of macro news on the term premium.

[7]The term structure of risk premia depends not only on $\mathcal{P}_t$ but also on $\Delta\mathcal{P}_{n,t}$ and $\Delta\mathcal{P}_{y,t}$ in a different manner.

Figure 8: Term Premium and Expectation Components Variance Decomposition

Panel (a) shows the share of yields variances to due macro news $\mathcal{P}_{n,t}$. Panel (b) shows the share of yields variances due to the expectation component and Panel (c) shows the covariance between the expectation and risk premium components $eh_t^{(m)}$ and $rp_t^{(m)}$. Panel (d) and (e) shows the share of $eh_t^{(m)}$ variance and of $rp_t^{(m)}$ variances due to macro news, respectively.

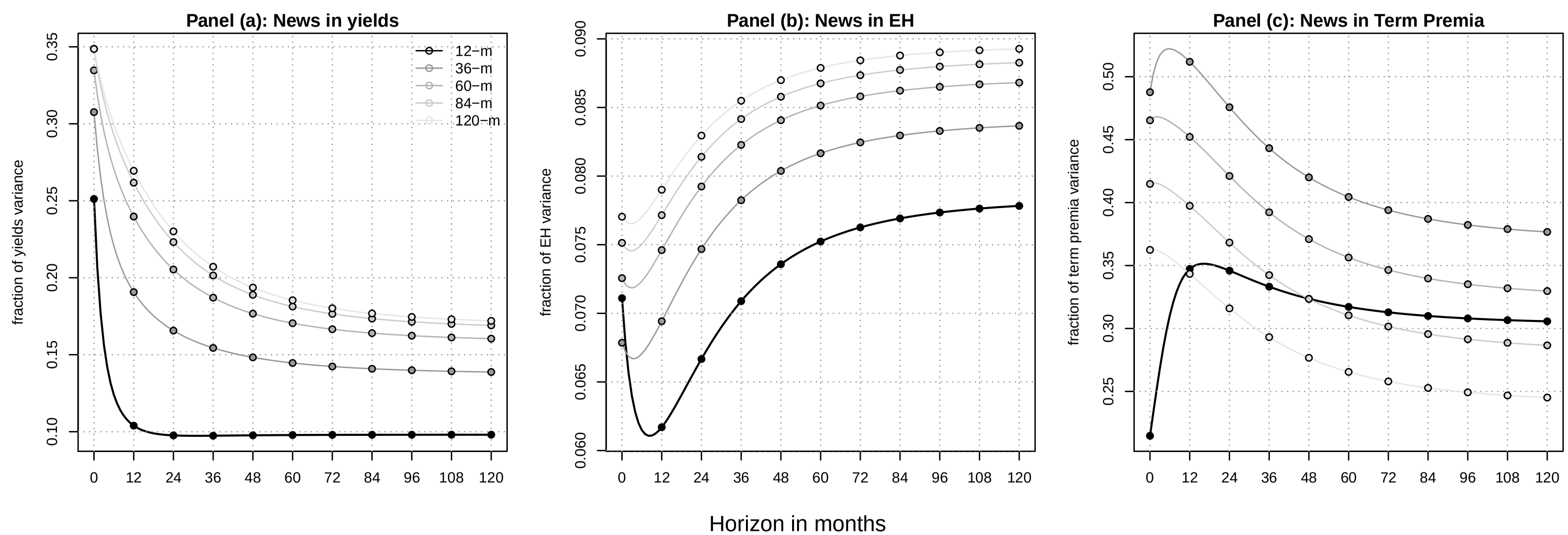


Figure 9: Time Series of Forecast Errors

Time series decomposition of yields forecasts errors $\tilde{y}_t^{(m)}$ in terms of total revisions to the expectation component $\Delta eh_t^{(m)}$ and the revisions due to macro news $\Delta eh_{n,t}^{(m)}$ (Row 1) and in terms of the total revisions to the risk premium component $\Delta rp_t^{(m)}$ and the revisions due to macro news $\Delta rp_{n,t}^{(m)}$ (Row 2).

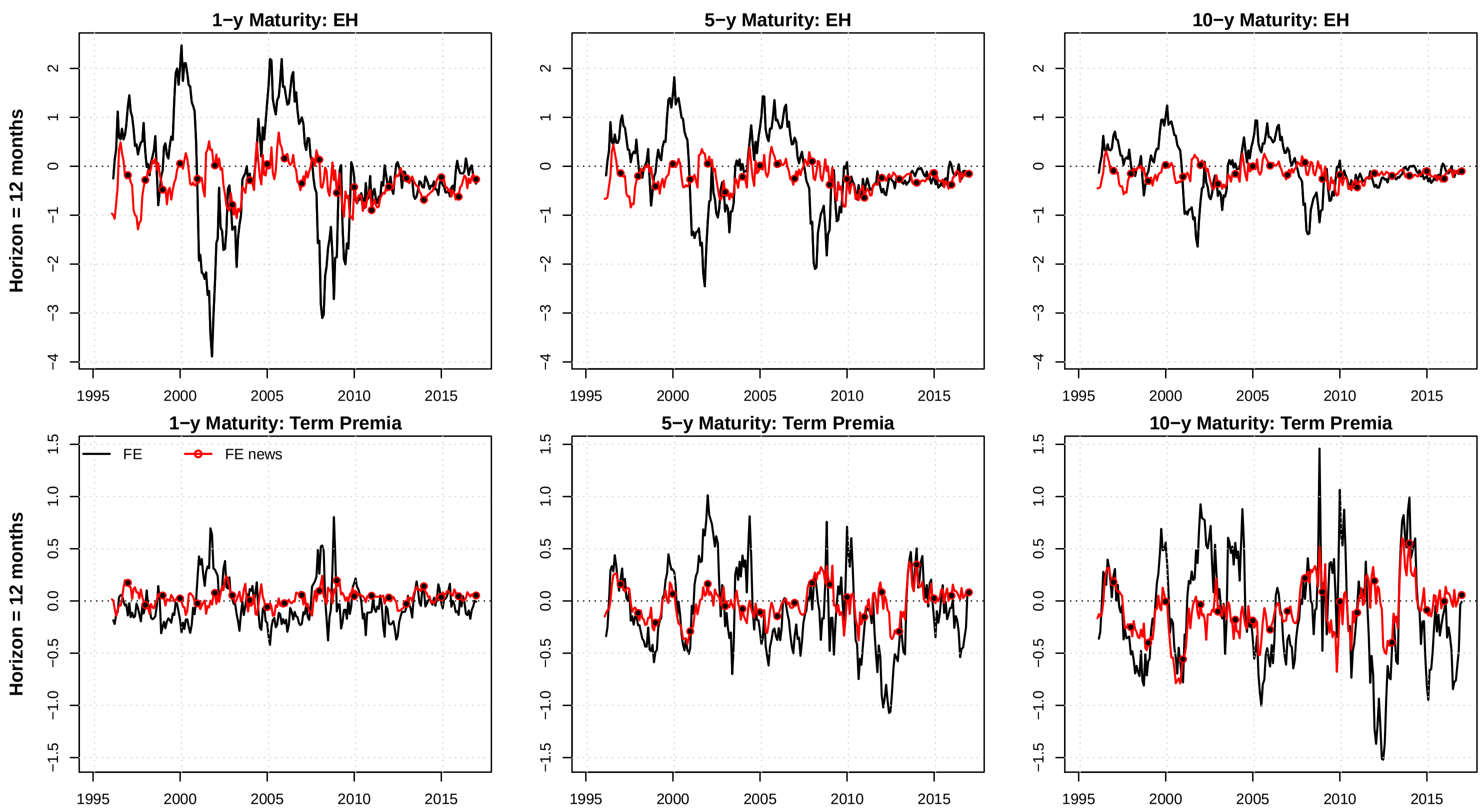
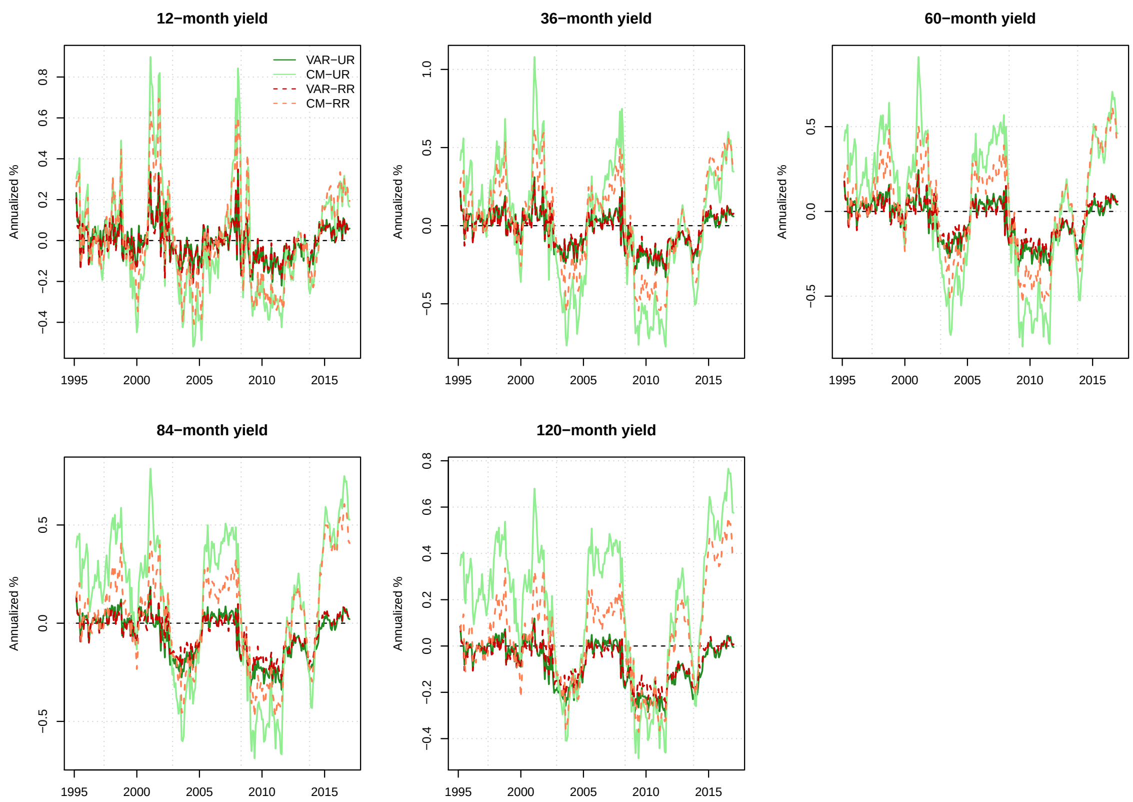


Figure 10: Term Premium Estimates relative to the VAR(1)

Term premium estimates for the models VAR-UR (darkgreen solid line), VAR-RR (red dashed line), CM-UR (lightgreen solid line), and CM-RR (dashed orange line), all relative to the term premium implied by the VAR(1) .



5 Conclusion

An important branch of the literature emphasize the impact of macro news in the evolution of bond yields. Our results are disappointing in that most of yield variations at business-cycle frequencies cannot be related to macro news. This raises several important questions. What is the role of financial shocks in the variance decomposition of long term interest rates. What is the role of information about the economy revealed outside of data releases. Speeches, geopolitical events, corporate news all may indicate to bond investor new information about the direction of inflation, unemployment and monetary policy. Another question for future research is whether the impact of macro news changes over time.

References

Altavilla, C., D. Giannone, and M. Modugno (2017). Low frequency effects of macroeconomic news on government bond yields. *Journal of Monetary Economics 92*, 31 – 46.

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

Ang, A., M. Piazzesi, and M. Wei (2006). What does the yield curve tell us about GDP growth? *Journal of Econometrics 131*(12), 359–403.

Bauer, M. D. (2015). Nominal interest rates and the news. *Journal of Money, Credit and Banking 47*(2-3), 295–332.

Bikbov, R. and M. Chernov (2010). No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics 159*(1), 166 – 182.

Campbell, J. Y. and J. Amner (1993). What moves the stock and bond markets? a variance decomposition for long-term asset returns. *The Journal of Finance 48*(1), 3–37.

Cecchetti, S. G. (2003). What the fomc says and does when the stock market booms. *Asset prices and monetary policy 77*, 96.

Cieslak, A., A. Morse, and A. Vissing-Jorgensen (2019). Stock returns over the fomc cycle. *The Journal of Finance 74*(5), 2201–2248.

Cieslak, A. and A. Schrimpf (2018). Non-monetary news in central bank communication. Technical report, National Bureau of Economic Research.

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

Diebold, F., G. Rudebusch, and S. Aruoba (2006). The macroeconomy and the yield curve : A dynamic latent factor approach. *Journal of Econometrics 131*, 309–338.

Doshi, H., K. Jacobs, and R. Liu (2018). Macroeconomic determinants of the term structure: Long-run and short-run dynamics. *Journal of Empirical Finance*.

Duffee, G. (2018). Expected inflation and other determinants of treasury yields. *The Journal of Finance 73*(5), 2139–2180.

Ederington, L. H. and J. H. Lee (1993). How markets process information: News releases and volatility. *The Journal of Finance 48*(4), 1161–1191.

Feunou, B. and J.-S. Fontaine (2018). Bond risk premia and gaussian term structure models. *Management Science 64*(3), 1413–1439.

Fleming, M. J. and E. M. Remolona (1999). Price formation and liquidity in the us treasury market: The response to public information. *The journal of Finance 54*(5), 1901–1915.

Fontaine, J.-S. and R. Garcia (2012). Bond liquidity premia. *Review of Financial Studies 25*(4), 1207–1254.

Gilbert, T., C. Scotti, G. Strasser, and C. Vega (2017). Is the intrinsic value of a macroeconomic news announcement related to its asset price impact? *Journal of Monetary Economics 92*, 78 – 95.

Gurkaynak, R., B. Kisacikoglu, and J. Wright (2018). Missing events in event studies: Identifying the effects of partially-measured news surprises. Working Paper w25016, NBER.

Gurkaynak, R., B. Sack, and E. Swanson (2007). Market based measures of monetary policy expectations. *Journal of Business and Economics Statistics 25*, 201–212.

Gurkaynak, R., B. Sack, and J. Wright (2006). The U.S. Treasury curve: 1961 to present. Technical Report 2006-28, Federal Reserve Board.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005, March). The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *American Economic Review 95*(1), 425–436.

Joslin, S., A. Le, and K. J. Singleton (2013). Gaussian macro-finance term structure models with lags. *Journal of Financial Econometrics 11*(4), 581–609.

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

Kilian, L. and C. Vega (2011). Do energy prices respond to us macroeconomic news? A test of the hypothesis of predetermined energy prices. *Review of Economics and Statistics 93*(2), 660–671.

Kuttner, K. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics 47*, 523–544.

Longstaff, F. (2004). The flight-to-liquidity premium in U.S. treasury bond prices. *The Journal of Business 77*, 511–526.

Lütkephol, H. (2006). *New Introduction to Multiple Time Series Analysis.* Springer.

Moench, E. (2012). Term structure surprises: The predictive content of curvature, level, and slope. *Journal of Applied Econometrics 27*(4), 574–602.

Piazzesi, M. (2005). Bond yields and the Federal Reserve. *The Journal of Political Economy 113*, 311–344.

Rudebusch, G. D. (1998). Do measures of monetary policy in a VAR make sense? *International Economic Review 39*(4), 907–931.

Rudebusch, G. D. and T. Wu (2008). A macro-finance model of the term structure, monetary policy and the economy. *The Economic Journal 118*(530), 906–926.

A Appendix

A.1 CME Futures Data

We use high-frequency CME data to measure the immediate response of futures rates to the release of new macroeconomic data. Specifically, we use transaction data for US Treasury futures contracts between 1995 and 2016. The most liquid contracts are for delivery of Treasury securities with either 2 years, 5 years or 10 years to maturity. For each maturity, a futures contract is available with quarterly delivery at the end of March, June, September or December. If an observation date falls in the first two months of a given quarter, we use the contract for the same quarter. For instance, for data releases observed in January and February we use the contract maturing in March. This is the most liquid contract at this date. However, for releases observed in the last month of any quarter, we use the next quarterly contract, because trading activity migrates to this contract. For instance, for releases observed in March we use the contract maturing in June. This yields a panel of futures contracts referencing three Treasury securities with different but constant maturities. These are the most liquid Treasury futures contracts.

Next, we convert the change of futures prices around data releases in terms of bond yields. The procedure is straightforward. Recall that the payoff of futures contracts is contingent on the price of the Treasury bond that will be delivered multiplied by a conversion factor. The conversion factor is public and easy to compute. Essentially, the procedure means that the quoted futures prices closely approximates the price we should observe in the Treasury market for a bond with a six percent coupon (eight percent before September 29, 1999). This means we can use the standard bond pricing formula to translate the futures prices into a bond yield. Any small approximation error in the computation of the implied yield washes out when we compute changes around small time intervals.

Figure A.1: Yield and Futures PCA–Time Series

Times-series of bond yield PCAs and futures yield PCA.

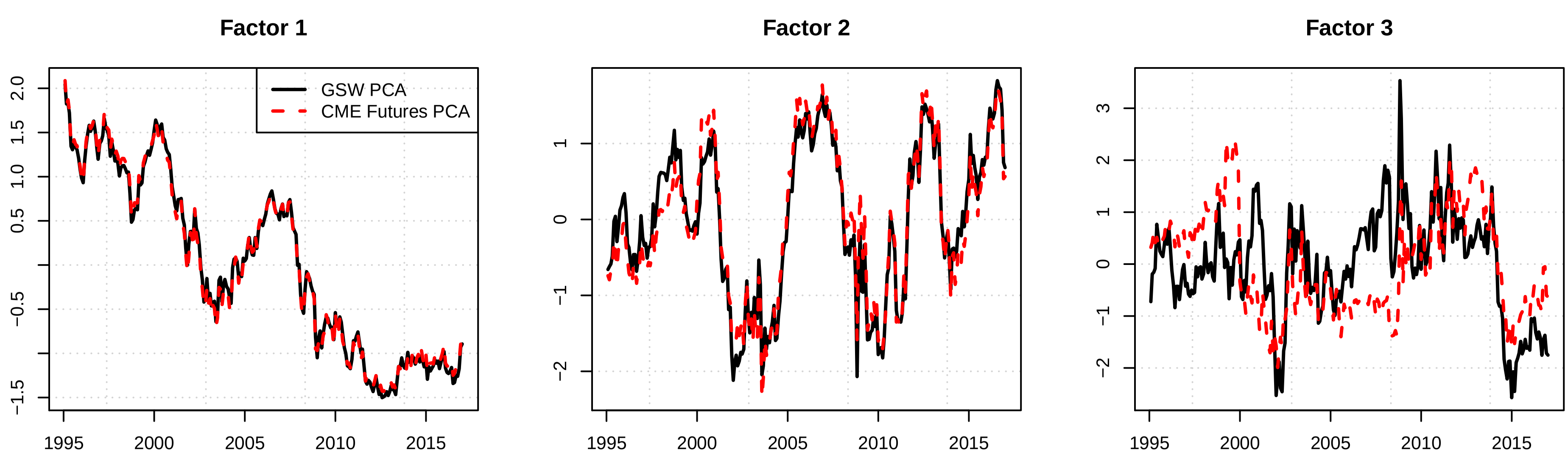


Figure A.2: Yield and Futures PCA–Time Series

Regressions of changes in yield PCAs $\Delta\mathcal{P}_t$ on monthly futures components $\Delta\mathcal{F}_{n,t}$.

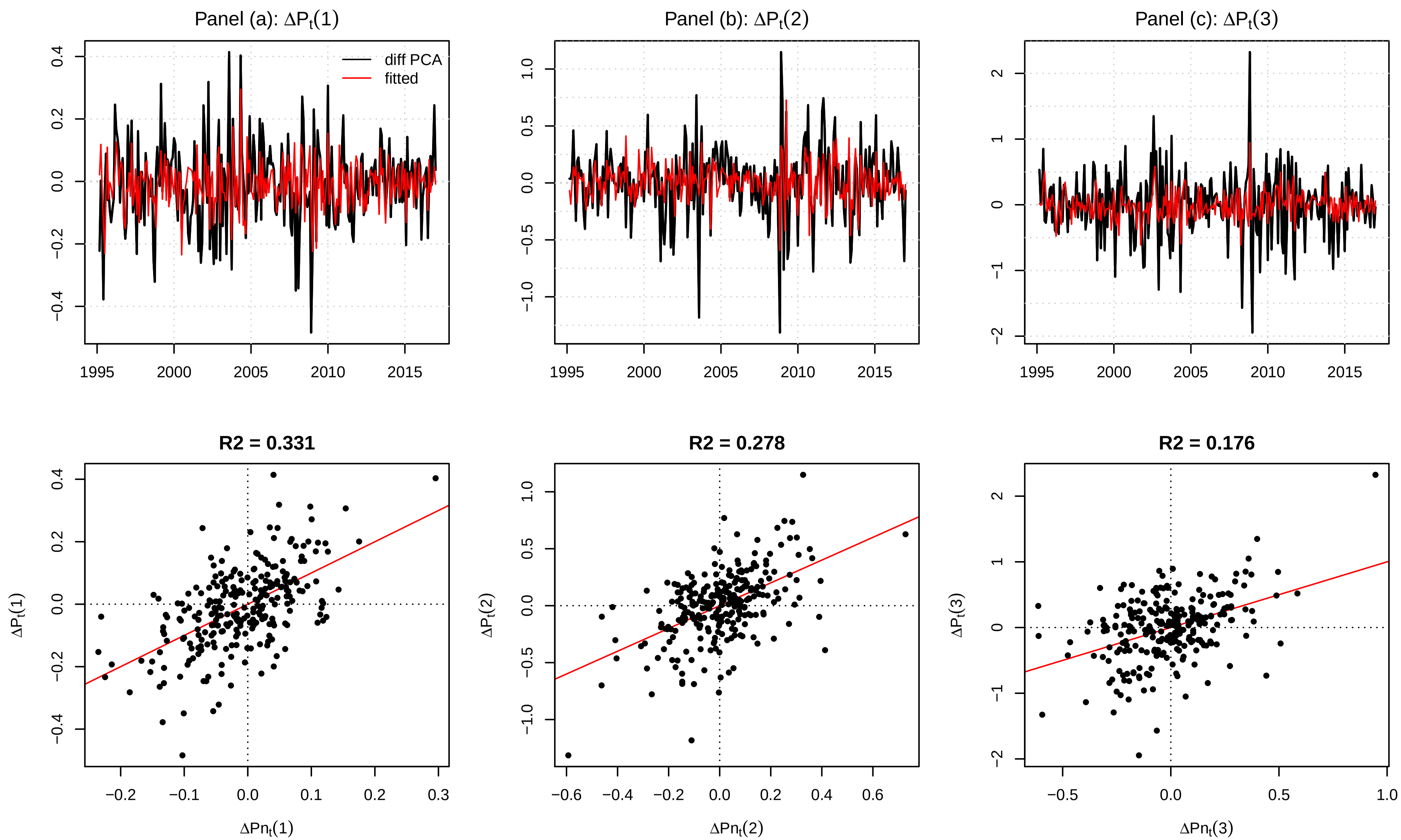


Figure A.3: Variance decomposition of expectation and term premia across models and maturities

Variance ratio due to macro news for the yields $y_t^{(m)}$, the expectation component $eh_t^{(m)}$ and risk premium $rp_t^{(m)}$ across models.

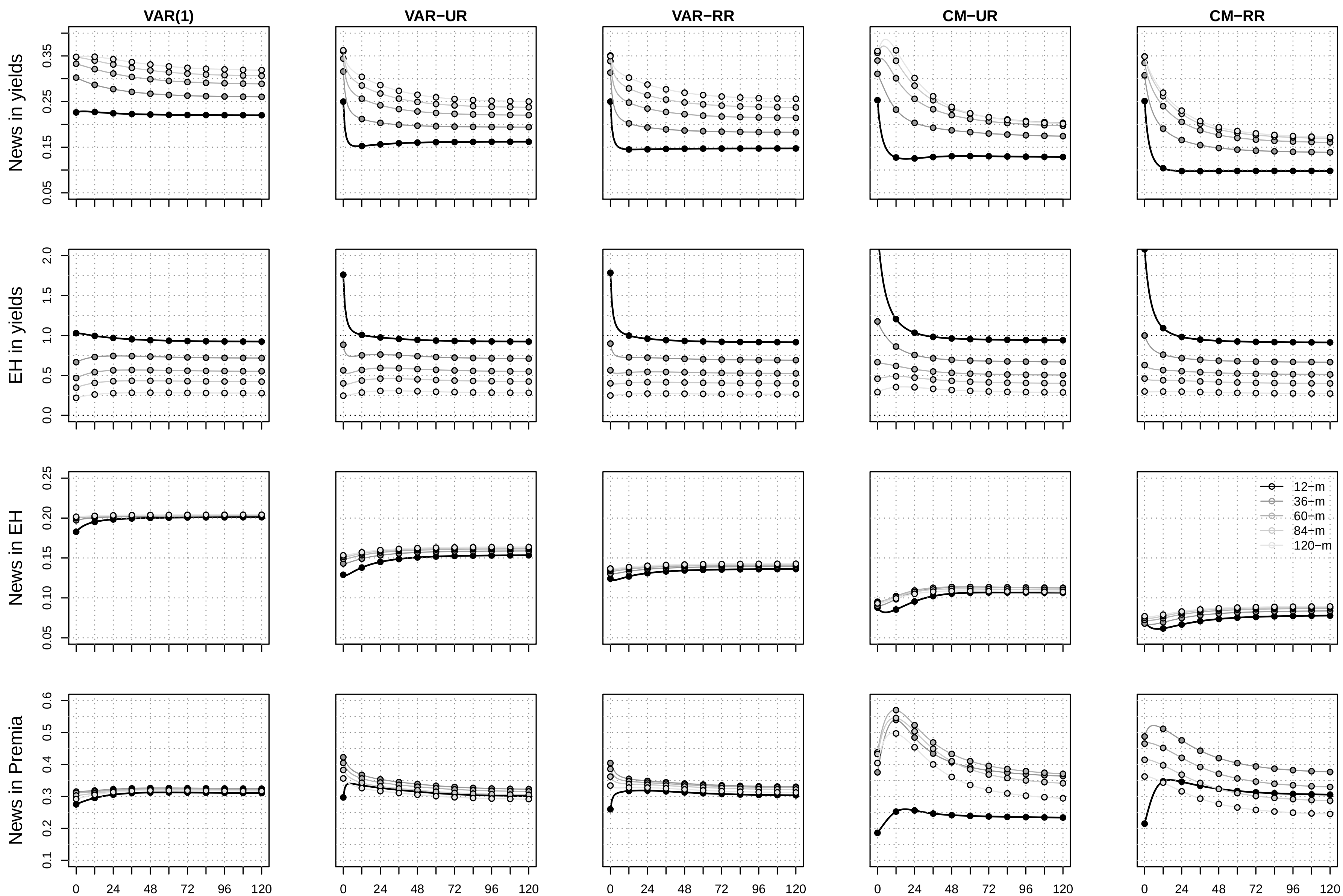


Table A.1: Parametre Estimates—Historical Dynamics

	VAR(1)	VAR(2)	VAR-UR	VAR-RR	VARMA(1)	VARMA(2,1)	CM-UR	CM-RR
$K_0(1)$	-0.0011	-0.0015	2e-04	2e-04	-0.0036	-0.0025	-5e-04	-2e-04
$K_0(2)$	-8e-04	-9e-04	0.0043	0.0035	0.0117	0.009	1e-04	-8e-04
$K_0(3)$	0.0051	0.0062	0.0036	0.0018	-4e-04	-0.0021	0.005	6e-04
$K_z(1,1)$	-0.0118	-0.014	-0.0112	-0.0022	-0.0129	-0.0046	-0.0116	-0.0029
$K_z(2,1)$	3e-04	3e-04	-0.0011	-0.0022	0.0045	0.0017	7e-04	-7e-04
$K_z(3,1)$	0.0603	0.0723	0.0607	0.017	0.0612	0.0254	0.0602	0.0204
$K_z(1,2)$	0.0078	0.0095	0.0066	8e-04	0.0067	0.0016	0.0067	0.0015
$K_z(2,2)$	-0.0373	-0.0427	-0.0416	-0.0143	-0.041	-0.017	-0.039	-0.0137
$K_z(3,2)$	0.019	0.021	0.0232	0.0028	0.0246	0.0073	0.0193	0.003
$K_z(1,3)$	-0.008	-0.0106	1e-04	0.001	0.0044	0.0036	-5e-04	0.0012
$K_z(2,3)$	-0.0204	-0.02	-0.0176	-0.0062	-0.0275	-0.0135	-0.009	-0.0081
$K_z(3,3)$	-0.1201	-0.14	-0.1177	-0.0187	-0.1145	-0.0287	-0.1222	-0.0222
$K_u(1,1)$	0	0	0.1554	0.0446	0.251	0.0917	0.2523	0.0965
$K_u(2,1)$	0	0	0.4751	0.3282	0.3885	0.3442	0.3855	0.1376
$K_u(3,1)$	0	0	-0.1535	0.1642	-0.0694	0.2554	-0.0707	0.4836
$K_u(1,2)$	0	0	0.062	0.034	0.0657	0.0475	0.063	0.0252
$K_u(2,2)$	0	0	0.1344	0.0178	0.0998	-0.0078	0.0962	0.036
$K_u(3,2)$	0	0	-0.0875	0.0967	-0.0529	0.0903	-0.0176	0.1265
$K_u(1,3)$	0	0	-0.0142	-0.0188	-9e-04	-0.0086	0.0123	-0.0116
$K_u(2,3)$	0	0	0.1242	0.0814	0.0837	0.0604	0.0188	-0.0166
$K_u(3,3)$	0	0	-0.0725	-0.1526	-0.0677	-0.1797	-0.0034	-0.0583
$K_n(1,1)$	0	0	0.1554	0.0446	-0.1793	-0.1847	0.0588	-0.0196
$K_n(2,1)$	0	0	0.4751	0.3282	1.0692	0.8035	0.0898	-0.028
$K_n(3,1)$	0	0	-0.1535	0.1642	-0.6344	-0.4387	-0.0165	-0.0983
$K_n(1,2)$	0	0	0.062	0.034	0.008	-0.0259	0.0508	0.0089
$K_n(2,2)$	0	0	0.1344	0.0178	0.3135	0.2164	0.0776	0.0127
$K_n(3,2)$	0	0	-0.0875	0.0967	-0.2939	-0.0558	-0.0142	0.0445
$K_n(1,3)$	0	0	-0.0142	-0.0188	-0.0992	-0.0888	0.0058	-0.0107
$K_n(2,3)$	0	0	0.1242	0.0814	0.3874	0.2848	0.0088	-0.0152
$K_n(3,3)$	0	0	-0.0725	-0.1526	-0.1238	-0.0863	-0.0016	-0.0536
θ	0	-0.1869	0	0.7473	0	0.665	0	0.6852
$\sqrt{\Omega_u}(1,1)$	0.1061	0.106	0.1042	0.105	0.1028	0.1038	0.1041	0.1051
$\sqrt{\Omega_u}(2,1)$	-0.1359	-0.1359	-0.1391	-0.1399	-0.1366	-0.1366	-0.1446	-0.1431
$\sqrt{\Omega_u}(3,1)$	-0.2055	-0.2069	-0.2103	-0.2186	-0.2146	-0.2207	-0.2083	-0.2182
$\sqrt{\Omega_u}(2,2)$	0.2184	0.2188	0.2094	0.2043	0.2066	0.2033	0.2105	0.2101
$\sqrt{\Omega_u}(3,2)$	-0.1169	-0.1164	-0.1205	-0.135	-0.1211	-0.1379	-0.1309	-0.1507
$\sqrt{\Omega_u}(3,3)$	0.3489	0.3479	0.3431	0.3296	0.3397	0.3235	0.3422	0.3251
$\sqrt{\Sigma}(1,1)$	0.13	0.1299	0.1284	0.1291	0.1273	0.1281	0.1284	0.1291
$\sqrt{\Sigma}(2,1)$	-0.1856	-0.1857	-0.1884	-0.189	-0.1866	-0.1865	-0.1929	-0.1916
$\sqrt{\Sigma}(3,1)$	-0.2352	-0.2363	-0.2389	-0.2457	-0.2421	-0.2473	-0.2372	-0.2454
$\sqrt{\Sigma}(2,2)$	0.2405	0.2409	0.232	0.2274	0.2295	0.2266	0.2326	0.2325
$\sqrt{\Sigma}(3,2)$	-0.1364	-0.136	-0.1402	-0.1538	-0.1412	-0.1567	-0.1493	-0.1678
$\sqrt{\Sigma}(3,3)$	0.3815	0.3806	0.3763	0.3639	0.3733	0.3585	0.3754	0.3599